Modelling of Amyotrophic Lateral Sclerosis (ALS) using induced Pluripotent Stem Cells (iPSC)

Nidaa Anwar Mohammad Ababneh

St Cross College

University of Oxford

A thesis submitted for the degree of
DOCTOR OF PHILOSOPHY IN CLINICAL NEUROSCIENCES

Nuffield Department of Clinical Neurosciences
Medical Sciences Division
University of Oxford
Hilary term, 2017

ABSTRACT

Modelling of Amyotrophic Lateral Sclerosis (ALS) using induced Pluripotent Stem Cells (iPSC)

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford,
Nidaa Ababneh, St Cross College, Hilary Term 2017.

The hexanucleotide repeat expansion (HRE) mutation within *C9orf72* gene is the most common cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Several hypotheses have been proposed for how the mutation contributes to pathogenicity, including the loss of *C9orf72* gene function, RNA-mediate toxicity and the formation of toxic dipeptides by repeat-associated non-ATG (RAN) translation. Patient-specific iPSCs provide a promising tool for the study of the cellular and molecular mechanisms of human diseases in relevant cell types and discovering potential therapies. The CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9-mediated homology directed repair (HDR) system represents an attractive approach for disease modelling and development of therapeutic strategies. In this thesis, iPSCs derived from ALS/FTD patient carrying the HRE mutation were generated and subsequently gene edited to remove a massive repeat expansion from the patient cells and replace it with the wild-type size of the repeats using HDR and a plasmid donor template. The successful genotypic correction of the mutation resulted in the normalization of the *C9orf72* gene promoter methylation level and the gene variants RNA expression level. Removal of the mutation also resulted in abolition of sense and antisense RNA foci formation and reduction of DPRs accumulation. Furthermore, the repeat size correction also rescued the susceptibility of cells to Glutamate excitotoxicity, decreased the apoptotic cell death and stress granules formation under the baseline and stress conditions. This work provides a proof-of-principle that removal of the HRE can rescue ALS disease phenotypes and provides an evidence that HRE mutation is an attractive target for therapeutic strategies and drug screening, to block the underlying disease mechanisms.

The time has come, it may be said,

To dream of many things;

Of genes—and life—and human cells—

Of Medicine—and kings—

Edward L Tatum, Perspectives in biology and medicine

1966.

Declaration Statement

I, Nidaa Ababneh, declare that the results presented here are from original experimental work which has been carried out by me at the Sir William Dunn School of Pathology and the Nuffield Department of Clinical Neurosciences, University of Oxford, between Michaelmas 2013 and Michaelmas 2016. This work was supervised by Professor Kevin Talbot and co-supervised by Dr Sally Cowley. All the work described in this thesis is my own, unless otherwise specifically stated, where I clearly stated the contribution of others to my thesis in the acknowledgment. No part of this work has been submitted for any other degrees or qualification at this or any other university.

Acknowledgments

"This thesis was one of the most important and informative experiences in my life."

I would like to say a great thanks to everyone who has helped and offered support along the way to make the past three years an incredible journey!

I couldn't have done it without you all!

In the beginning, I would like to give a special thanks to Kevin and Sally for being my supervisors. Your integral view on research has made a deep impression on me. I owe you both lots of gratitude for showing me this way of research and giving me the flexibility to work in my way.

Nuffield Department of Clinical Neurosciences (NDCN)

Professor Kevin Talbot – for great supervision, mentorship and support.

Dr Ruxandra Dafinca – for help and advice in motorneuron differentiation and phenotyping.

Dr Jakub Scaber – for general help and advice especially with RNA foci staining experiment.

My very great thanks to all Talbot laboratory members, for their help, guidance and friendship.

Sir William Dunn School of Pathology

Dr Sally Cowley – for great co-supervision, mentorship and support.

Dr Rowan Flynn – for considerable help in CRISPR/Cas9 gene editing experiment.

Jane Vowles and Cathy Browne – for general help and advice with iPSC reprogramming.

Dr Mirvat Surakhy– for help in proofreading of my thesis.

My very great thanks to all James laboratory members, for their help, guidance and friendship.

University of Southampton, Human Development and Health

Dr Andrew Douglas – for help and advice with performing *C9orf72*-related methylation assays.

Molecular Genetics Laboratory, Oxford University Hospitals NHS Trust

I am very grateful to all staff members at the Molecular Genetics laboratory/Churchill hospital for their help and advice in performing the RP-PCR and Southern blotting assays.

The University of Jordan

I am very grateful to the University of Jordan who have funded my DPhil scholarship and made it possible for this research to happen.

I extend many thanks to my friends: Cecilia, Yichen, Maggie, Walther, Julian, Federica, David, Ester, Mirvat, Maha, Reem, Helen, Fatima, Sayra, Maysa, Israa, Mohammad, Amy, and Abu. To all my friends, thank you for your understanding and encouragement. Your friendship makes my life a wonderful experience.

Dedication

"The price of success is hard work, dedication to the job at hand, and the determination that whether we win or lose, we have applied the best of ourselves to the task at hand." Vince Lombardi

❖❖❖❖

First and foremost, this thesis is dedicated to the loving memory of my beloved mother, Alya, who passed away on September 2014. There are no words to express how much I miss you.

To my loving father, Anwar, for his endless love, who always supported me all the way since the beginning of my life. I hope you know that this work is as much yours as it is mine. I am so proud to be your daughter.

❖❖❖❖

To my lovely sisters and brothers who have been great sources of motivation and inspiration, you have always supported me in a way nobody else could. I wish I can offer you the love, the care and the support you need during the rest of your lives.

❖❖❖❖

To My dear friends who were always near, bringing out the best in me.

❖❖❖❖

This thesis must also be dedicated to Prof Abdalla Awidi from Jordan for his magnificent help in getting the PhD scholarship. Thank you for your support through the years!

❖❖❖❖

Finally, I would like to dedicate this thesis to everyone who has contributed on this work.

Nidaa Anwar Ababneh
(December 2016)

Publications and presentations arising from this work

Publications generated from this work:

1. Ruxandra Dafinca, Jakub Scaber, Nida'a Ababneh, Tatjana Lalic, Gregory Weir, Helen Christian, Jane Vowles, Andrew G.L. Douglas, Alexandra Fletcher-Jones, Cathy Browne, Mahito Nakanishi, Martin R. Turner, Richard Wade-Martins, Sally A. Cowley, and Kevin Talbot. "**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. 2016 Aug; 34(8): 2063–2078.

2. Nidaa Ababneh, Rowan Flynn, Ruxandra Dafinca, Jakub Scaber, Andrew Douglas, Kenny Moore, Martin Turner, Sally Cowley, Kevin Talbot. "**Correction of ALS-related phenotypes in iPSC-derived motorneurons carrying a hexanucleotide expansion mutation in *C9orf72* by CRISPR/Cas9 genome editing and homology-directed repair (HDR) system."** (Manuscript in Preparation).

Platform and Poster Presentations arising from this work

- **Platform presentation**. ENCALS meeting Dublin 21st-2nd May, 2015. **"Genetic correction of C9orf72 Repeat Expansion Mutation in ALS/FTD Patient iPSC."**

- **Shortlisted for the best 'Science Poster' prize competition.** The International ALS/MND Symposium, Orlando, USA, 11th-13th December, 2015. **Genetic Correction of C9orf72 Repeat Expansion Mutation In ALS/FTD Patient iPSCS.**

- **The best poster Prize winner.** 3rd BIRAX Conference, April, 2016. "**CRISPR/Cas9 genome editing results in correction of ALS phenotypes in *C9orf72*mutant iPSC-derived motorneurons."**

- **Platform presentation.** ENCALS 2016 meeting in Milan from 19th to 21th of May. "**CRISPR/Cas9 genome editing results in correction of ALS phenotypes in *C9orf72* mutant iPSC-derived motorneurons."**

- **Poster prize certificate winner.** OxDARE/ARUK Oxford Network Memory Research Day 2016 on 13th July. "**CRISPR/Cas9 results in genotypic and phenotypic correction of ALS/FTD patient iPSCS repeat expansions mutation in *C9orf72* gene."**

- **Poster presentation.** The International ALS/MND Symposium, Dublin, Ireland, 7th-9th December, 2016. "**CRISPR/Cas9 genome editing results in genotypic and phenotypic correction of ALS phenotypes in *C9orf72* mutant iPSC-derived motorneurons."**

List of Abbreviations

°C – Degrees Celsius

< – is less than

> – is more than

~ – approximately

3’ – 3 prime

5’ – 5 prime

5mC _ 5-methylcytosine

aa – Amino acids

AA – Ascorbic Acid

AD –Alzheimer’s Disease

aDMEM – Advanced Dulbecco's Modified Eagle's Medium

ALS – Amyotrophic Lateral Sclerosis

ALS/FTD– Amyotrophic Lateral Sclerosis / Frontotemporal dementia

Amp _ Ampicillin

AMPA _ α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid

ANOVA – Analysis of variance

ASOs – Antisense oligonucleotides

ATP _ Adenosine triphosphate

ATXN8 _ Ataxin-8

BDNF – Brain-derived neurotrophic factor

BER _ Base excision repair

bFGF _ Basic fibroblast growth factor

BMP-4 _ Bone morphogenetic protein 4

bp _ base pair

BSA _ bovine serum albumin

C _ Cytosine

c-Myc _ Avian Myelocytomatosis Viral Oncogene Homolog

C. elegans _ Caenorhabditis *elegans*

C9orf72 _ Chromosome 9 open reading frame 72

CaCl2 _ Calcium chloride

cAMP _ Cyclic AMP

cDNA _ Complementary DNA

CGIs _ CpG islands

ChAT _ Choline O- Acetyltransferase

CHMP2B _ Charged multivesicular body protein 2B

Clvd-Casp-3 _ Cleaved Caspase-3

CM _ Conditioned media

cMNs _ Cranial motorneurons

CNS _ Central nervous system

CO2 _ Carbon dioxide

CRISPR _ Clustered regularly interspaced short palindromic repeats

CSF _ Cerebrospinal fluid

DAPI _ 4’ 6-diamidino-2-phenylindole

DAPT _ N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester

ddH2O _ double distilled water

ddPCR _ Digital droplet polymerase chain reaction

DENN _ Differentially Expressed in Normal and Neoplasia

DIG _ Digoxigenin

DM _ Myotonic dystrophy

DM1 _ Myotonic Dystrophy Type 1

DMEM _ Dulbecco’s modified eagle medium

DMEM/F12 _ Dulbecco’s Modified Eagle Medium/F12 Supplement

DMPK _ Dystrophia Myotonica Protein Kinase

DMSO _ Dimethyl sulfoxide

DNA _ Deoxyribonucleic acid

DNase _ Deoxyribonuclease

DNMTs _ DNA methyltransferases

dNTP _ Deoxynucleotide triphosphate

DPR _ Dipeptide repeat

DR4 _ Drug-resistant (MEFs)

DSB _ Double-strand break

dsDNA _ Double-strand DNA

dUTP _ Deoxyuridine triphosphate

EAAT2 _ Excitatory amino acid transporter 2

EB _ Embryoid Body

EDTA _ Ethylenediaminetetraacetic acid

EF1α _ Elongation Factor-1 alpha

ER _ Endoplasmic Reticulum

ERAD _ ER-associated degradation

ESCs _ Embryonic stem cells

FACS _ Fluorescence Activated Cell Sorting

fALS _ Familial ALS

FAM _ Carboxyfluorescein

FASL _ Fas ligand

FBS _ Fetal bovine serum

FBS-ES _ Fetal bovine serum-Embryonic stem

FCS _ Fetal calf serum

FGF _ Fibroblast growth factor

FISH _ Flourescent in situ hybridization

FM _ Freezing medium

FMR1 _ Fragile X mental retardation 1

FMRP _ Fragile X mental retardation protein

FTD _ Frontotemporal dementia

FTLD _ Frontotemporal lobar dementia

FUS _ Fused in sarcoma

FXS _ Fragile X syndrome

G _ Guanine

G3BP1 _ GTPase-activating protein-binding protein 1

GA _ Glycine-Alanine

GAPDH _ Glyceraldehyde 3-Phosphate Dehydrogenase

GDNF _ Glial-derived neurotrophic factor

GDP _ Guanosine diphosphate

GEFs _ GDP/GTP exchange factors

GEO _ Gene Expression Omnibus

GFP _ Green Fluorescent Protein

GGGGCC _ (G4C2)

GLT-1 _ Astroglial glutamate transporter 1

GluR2 _ Glutamate receptor 2

GP _ Glycine-Proline

GR _ Glycine–Arginine

GTP _ Guanosine triphosphate

HB9 _ Homeobox 9

HBSS _ Hank's balanced salt solution

HD _ Huntington's disease

HDR _ Homology-directed repair

HEK _ Human embryonic kidney

hESC _ human Embryonic Stem Cell

hnRNPA1 _ heterogeneous nuclear ribonucleoprotein A1

hnRNPA2B1 _ heterogeneous nuclear ribonucleoprotein A2B1

HR _ homologous recombination

HRE _ Hexanucleotide repeat expansion

HRP _ Horseradish peroxidase

ICC _ Immunocytochemistry

ICM _ Inner cell mass

ICV _ Intracerebroventicular

IDT _ Integrated DNA Technologies

IGF _ Insulin-like growth factor

iPSC _ induced Pluripotent Stem Cells

iPSC-derived MNs _ induced pluripotent stem cell-derived motorneurons

IPTG _ Isopropyl β-D-1-thiogalactopyranoside

Islet1/2 _ Insulin Gene Enhancer Protein

IVF _ In vitro fertilization

kb _ kilobase

KCl _ Potassium chloride

Klf4 _ Kruppel-Like Factor 4

LN – Liquid Nitrogen

mC _ methylated C

MEFs _ Mouse embryonic fibroblasts

mESCs _ Mouse embryonic stem cells

MgCl2 – Magnesium chloride

ml – milliliter

MN – Motorneuron

MND _ Motorneuron disease

MOI _ Multiplicity of infection

mRNA _ messenger ribonucleic acid

mSOD1 _ mutant SOD1

NEAA _ Non-essential amino acids

NFW _ Nuclease-free water

NHEJ _ Non- homologous end joining

NI _ Neural Induction

NPs _ Neural progenitors

Nt _ nucleotide

Oct4 _ Octamer-Binding Transcription Factor 4

Olig2 _ Oligodendrocyte transcription factor

OPTN _ Optenin

P/S _ Penicillin/Streptomycin

PA _ Proline-alanine

PABP-1 _ Poly(A)-binding protein-1

PAM _ Protospacer adjacent motif

PAX6 _ Paired box protein

PBS _ Phosphate buffered saline

PCR _ Polymerase chain reaction

PD _ Parkinson's disease

PFA _ Paraformaldehyde

PI _ Propidium Iodide

poly-Q _ Polyglutamine

PR _ Proline-Arginine

Puro _Puromycin

Puro/TK _ Puromycin/thymidine kinase

qRT-PCR _ Quantitative reverse transcriptase polymerase chain reaction

RA _ Retinoic acid

Rab-GEF _ Rab-GTPase GDP/GTP exchange factor

RAN _ Repeat-Associated Non-ATG initiated translation

RBP _ RNA-binding protein

REC _ Research Ethics Committee

RFLP _ Restriction fragment length polymorphism

RI _ Random integration

RNA _ Ribonucleic acid

RNP _ Ribonucleoprotein

ROS _ Reactive oxygen species

RP-PCR _ Repeat-Primed PCR

RT _ Reverse Transcription or Reverse Transcriptase

RT _ Room temperature

RVDs _ Repeat variable di-residues

SAG _ Sonic Agonist

sALS _ Sporadic ALS

SCA1 _ Spinocerebellar ataxia 1

SD _ Standard deviation

SDM _ Site-directed mutagenesis

SDS _ Sodium dodecyl sulfate

Sec _ Seconds

SG _ Stress granule

sgRNA _ single gRNA

SHH _ Sonic hedgehog

SMA _ Spinal muscular atrophy

SMI-32 _ Neurofilament H Non-Phosphorylated

sMN _ Spinal motorneurons

SNP _ Single nucleotide polymorphism

SOC _ Super optimal broth

SOD1 _ Superoxide/Dismutase 1

Sox2 _ SRY (Sex Determining Region Y)-box 2

SQSTM1 _ Sequestosome-1, or p62

SSC _ Saline-Sodium Citrate

SSEA4 _ Specific embryonic antigen-4

ssODNs _ single-stranded oligonucleotide

T7E1 – T7 endonuclease 1

TAE – Tris base, acetic acid, ethylenediaminetetraacetic acid

TALEN _ Transcription activator-like effector nucleases

TARDBP – Transactive response DNA binding protein

TBE – Tris-borate EDTA

TDP-43 – Transactive response DNA-binding protein 43

TE _ Tris-EDTA

TGF-β _ Transforming growth factor beta

TNF- α _ Tumor necrosis factor-α

TRA-1-60 _ Tumor-related antigen 1-60

tracrRNA _ *trans*-activating crRNA

TX-100 _ Triton X-100

UBQLN2 _Ubiquilin-2

USP10 _ Uniquitin-specific peptidase 10

UV _ Ultraviolet

VAPB _ Vesicle -associated membrane protein-associated protein B

VCP _ Valosin containing protein

WT _ Wild-Type

Y-27632 _ Rock inhibitor

ZFNs _ Zinc Finger Nucleases

Table of Contents

List of Figures

List of Tables

1. CHAPTER ONE

INTRODUCTION

1.1 An introduction to Amyotrophic Lateral Sclerosis (ALS)

Amyotrophic lateral sclerosis (ALS), also known in the USA as Lou Gehrig's disease, is a neurodegenerative disease characterized by the progressive degeneration of upper and lower motorneurons (MNs), in the spinal cord, brain stem and cerebral cortex, leading to fatal paralysis and death within 1-5 years of onset (Kiernan et al. 2011). The name "ALS" is derived from the pathological observation that a distinct myelin pallor in the lateral part of the spinal cord reflects the degeneration and loss of the axons of upper MNs in the spinal cord (Boillée et al. 2006; Bruijn & Cudkowicz 2006). Common clinical features are progressive limb muscular atrophy and weakness, difficulty in speech, and swallowing, and respiratory failure. The proportion of affected individuals is 2 to 5 per 100,000/population per year worldwide and the average age of onset is 64 years old (Bennion Callister & Pickering-Brown 2014; Marin et al. 2016). Familial ALS (fALS) cases account for approximately 10%, while about 90% of ALS cases are sporadic (sALS) with no family history of ALS.

ALS is a multifactorial and multisystem disease involving MNs and other cell types and tissues, such as glial cells, skeletal muscles and immune cells. Up to 15% of all ALS cases may also show symptoms of frontal temporal lobar degeneration (FTLD) (Dormann & Haass 2013).

1.2. Frontotemporal Dementia

Frontotemporal dementia (FTD) is the second most common form of dementia, affecting approximately 20/100,000 patients less than 65 years-old (Onyike & Diehl-Schmid 2013; Luukkainen et al. 2015). FTD is characterized by neurodegeneration of the frontal and temporal lobes, described by neuropathologists as frontotemporal lobar degeneration (FTLD). The most common clinical features are behavioural and personality changes, with relative sparing of memory (Seelaar et al. 2011).

FTD is divided into three clinical syndromes; behavioural variant, semantic dementia, and progressive nonfluent aphasia (Neary et al. 1998). Some FTD patients show motorneuron disease symptoms typical of ALS (Lomen-Hoerth et al. 2002; Hodges et al. 2004), suggesting that the two clinical phenotypes are part of one spectrum disorder, termed as ALS/FTD (Ling et al. 2013).

Approximately 50% of ALS cases develop FTD-like cognitive and behavioural changes (Lomen-Hoerth et al. 2003; Ringholz et al. 2005), while up to 50% of FTD patients develop motor dysfunction (Lomen-Hoerth et al. 2003). FTD shares similar pathological features with ALS, mainly the presence of cellular mislocalisation of the trans-activating responsive (TAR) sequence DNA binding protein (TDP-43) into cytoplasmic inclusions throughout the central nervous system (CNS) (Neumann et al. 2006; Ringholz et al. 2005) and the most common genetic mutation of C9*orf*72 hexanucleotide repeat expansion (HRE), which leads to ALS and FTD (DeJesus-Hernandez et al. 2011; Renton et al. 2011; Gijselinck et al. 2015).

1.3. Clinical presentation of ALS/FTD

The disease can be classified clinically according to the site of onset. About two thirds of patients can develop symptoms in one limb at onset including muscle wasting and weakness, stiffness and cramps, and bulbar symptoms occurring later in disease (Brooks et al. 2000). ALS with bulbar onset due to combined degeneration of MNs in the corticobulbar tract and brainstem, affects the muscles of speech and swallowing, but in most cases also leads to generalised weakness involving the limbs and respiratory muscles (Polkey et al. 1999). The majority of ALS patients do not suffer from clinically significant cognitive or behavioural symptoms. However, FTD cognitive impairment occurs in some ALS cases (Ringholz et al. 2005).

1.4. Pathogenic Cellular Mechanisms in ALS

ALS mainly affects upper MNs of the corticospinal tract that project from the motor cortex onto synapses in the brainstem and spinal cord, and the lower MNs that project from the brainstem and spinal cord to skeletal muscles (Taylor et al. 2016).

ALS pathogenesis is complicated and multifactorial involving a number of biological pathways. The majority of ALS cases are sporadic and approximately 5-10% of ALS cases are familial with autosomal dominant inheritance. Therefore, genetic mutations represent key pathogenic mechanisms of fALS, but can also be found in a minority of sALS cases. A group of disease pathways contributing to the pathology of ALS have been identified, including excitotoxicity, oxidative stress, mitochondrial dysfunction, protein aggregation, neuroinflammation, impairments in RNA processing and axonal transport defect.

1.4.1. Genetic factors

Several genetic mutations have been identified in familial and sporadic ALS cases. *SOD1* mutations were the first gene mutations described over 20 years ago, *TDP-43, FUS, UBQLN, OPTN, VCP, C9orf72* and other rare gene mutations have been identified more recently (Taylor et al. 2016). All of these mutations and some other less frequent mutations are summarised in Table 1.1 (adapted fom Taylor et al., (2016).

1.4.1.1.SOD1

The *SOD1* gene encodes for the antioxidant enzyme Cu/Zn superoxide dismutase 1 (SOD1). SOD1 protein represents an ubiquitously expressed and widely distributed protein, consisting of 153 amino acids, mainly exists in the cytosol of cells, including motoneurons. SOD1 protein is responsible for protecting cells from oxidative stress through neutralising cytoplasmic oxygen free radicals by converting them to water and hydrogen peroxide (Reddi & Culotta 2013).

SOD1 mutations account for 10-20% of patients with fALS and 1-3% of sALS cases, in total about 2% of all patients (Deng et al. 1993; Andersen 2006). To date, more than 170 ALS-causing mutations have been identified in ALS patients. The ability of mutant SOD1 (mSOD1) protein to form insoluble ubiquitinated cytoplasmic aggregates and the contribution of other pathogenic cellular mechanisms, such as mitochondrial dysfunction and oxidative stress, protein misfolding, axonal transport defects and other pathways in indicate a gain of function mechanism (Saccon et al. 2013; Andersen et al. 2003; Urushitani et al. 2006).

To further confirm the involvement of *SOD1* as a disease causing mutation, more than 20 *SOD1* transgenic models have been developed to recapitulate the ALS disease phenotypes, including muscle atrophy, loss of spinal MNs, damage to astrocytes, mitochondrial dysfunction and axonal defects (Gurney et al. 1994; Wong et al. 1995;

Bruijn et al. 1997). Most studies on SOD1 show that a toxic gain of function is the main underlying disease mechanism, as knockout mice do not develop ALS pathology (Saccon et al. 2013). However, the exact mechanism of *SOD1* mutations pathogenicity in ALS is still unclear.

1.4.1.2.TARDBP

The transactive response DNA-binding protein (*TARDBP*) gene encodes for TDP-43, a protein of 414 amino acids and a molecular weight of 43 kDa. TDP-43 is an ubiquitously expressed DNA- and RNA-binding protein, predominantly in the nucleus, with lower levels in the cytoplasm (Ayala et al. 2008; Winton et al. 2008). TDP-43 plays a significant role in RNA metabolism including RNA transcription regulation, alternative splicing and transport and RNA stabilization (Buratti & Baralle 2001; Ignatius et al. 1995). TDP-43 can also be localised to stress granules where it interferes with several proteins associated with their regulation and formation (McDonald et al. 2011).

b mutations in TARDBP have been identified in fALS and sALS cases, and the majority are located within the C-terminal region of the protein (Andersen & Al-Chalabi 2011; Lagier-Tourenne et al. 2010). Several transgenic mouse models have been developed for *TARDBP* mutations. These models can recapitulate ALS pathology to varying degrees, with protein fragmentation and aggregation, neuronal loss, and reduced survival. Pathogenic mutations in the *TARDBP* gene can lead to relocalisation of the protein to the cytosol and to the loss of its nuclear localisation in MNs in the majority of sporadic cases of ALS (Neumann et al. 2006).

1.4.1.3.FUS

The fused in sarcoma (*FUS*, also known as *TLS*, for 'translocated in sarcoma') protein plays an important role in different cellular processes including DNA/RNA binding, transcriptional regulation (Law et al. 2006), DNA damage repair (Mastrocola et al. 2013), and RNA splicing and transport (Ranjith & Sama 2014). FUS protein is normally localize to neuronal nuclei, but in ALS/FTD with FUS mutations large insoluble protein aggregates staining for FUS form in the cytoplasm (Vance et al. 2009; Kwiatkowski et al. 2009). Mutations in the *FUS* gene account for about 4-6% of fALS cases and between 0.7-1.8% of sALS cases (Vance et al. 2009; Kwiatkowski et al. 2009). Most mutations in the *FUS* gene are located within the coding sequence of the nuclear localization domain, resulting in mislocalization of the protein to the cytoplasm (Vance et al. 2009; Kwiatkowski et al. 2009), and several mutations have been identified within the 3' untranslated region of the gene (Sabatelli et al. 2013). Mutations cause impairment in nuclear localization of the protein and result in more rapid disease progression (Dormann & Haass 2013).

1.4.1.4 C9orf72

In 2011, two independent groups have discovered that a hexanucleotide repeat expansion (HRE) mutation in a previously uncharacterised gene, at the C9*orf*72 locus of chromosome 9, is the most common genetic cause of FTD and ALS (DeJesus-Hernandez et al. 2011; Renton et al. 2011). The C9*orf*72 gene consists of 12 exons that can be transcribed into three main variants by alternative splicing (DeJesus-Hernandez et al. 2011; Renton et al. 2011) (Figure 1.1). Variants 2 (NM 018325.4) and 3 (NM 001256054.2) encode for the same 481-amino acid protein isoform A consists of exons

2-11. Variant 1 (NM 145005.6) encodes a shorter protein isoform B, containing 222 amino acids and encompassing exons 2-5 (Figure 1.2).

It has been shown that these protein isoforms are expressed in different regions of the brain and the peripheral tissues (DeJesus-Hernandez et al. 2011; Renton et al. 2011; Gijselinck et al. 2012; Waite et al. 2014; Rizzu et al. 2016). The C9*orf*72 protein is also expressed in neuronal cells that are affected in ALS and FTD such as cells in hippocampus, striatum, thalamus, brain stem, cerebellum, the cortex and the spinal cord (Suzuki et al. 2013). The HRE mutation consisting of G4C2 sequence repetition is located between the two non-coding exons, exon 1a and 1b (Figure 1.1). Repeats containing pre-mRNA transcripts 1 and 3 represent a small proportion of the total amount of C9*orf*72 gene levels, and variant 2 is the most abundant variant.

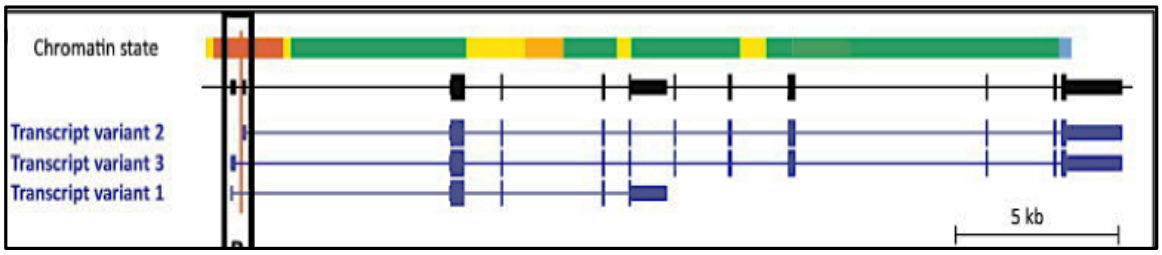

Figure 1. 1 Genomic organization of C9orf72 and the HRE at chromosome 9p21. The C9orf72 gene and its three major transcript variants. Black and blue boxes represent transcribed regions; wide boxes are coding and narrow boxes are untranslated exonic regions. Horizontal connecting lines represent intronic regions. The red vertical line indicates the location of the noncoding G4C2 hexanucleotide repeat. This figure was taken from (Cruts et al. 2013)

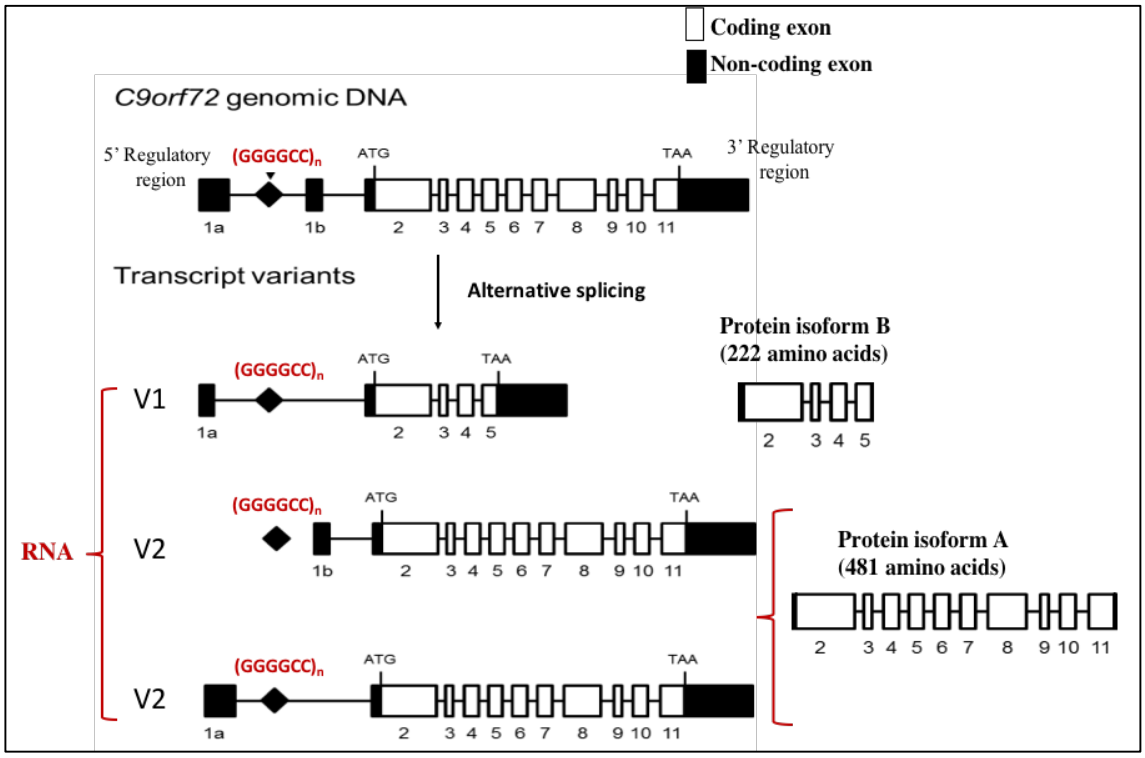

Figure 1. 2 *C9orf72* gene structure and transcription variants. The genomic DNA structure of C9orf72 gene and its three different transcripts. The black boxes represent the non-coding region and white boxes represent the coding region. Transcript 2 and 3 encode the long protein isoform A and transcript 1 encodes protein isoform B. The figure was adapted from (DeJesus- Hernandez et al., 2011).

1.4.1.4.5 Other mutations

The RNA-binding proteins hnRNPA1 and hnRNPA2B1 have a prion-like domain that enhances mRNA metabolism and transport. Mutations of these proteins have been reported in one patient with dominantly familial ALS and in animal models as stress granules and cytoplasmic inclusions (Liu et al., 2016). Mutations have also been identified in valosin-containing protein (VCP), which has a role endosomal trafficking. ALS-related mutations have been also detected in *VAPB*, which functions in the unfolded protein response in the delivery of ER-ejected substrates to the proteasome (Maruyama et al. 2010; Wild et al. 2011; Johnson et al. 2010; Parkinson 2006).

UBQLN2 (Ubiquilin-2) and *SQSTM1* (Sequestosome-1, or p62) have been also involved in familial cases of ALS, suggesting that defects in autophagy may also be implicated in the pathogenesis of ALS and FTD (Fecto et al. 2011; Rubino et al. 2012; Deng et al. 2011). Sequestosome-1 (p62), is a major component of cytoplasmic neuronal inclusions in both familial and sporadic FTD and ALS. Several other ALS genes are also involved in autophagy, including OPTN, *CHMP2B*, which encodes for proteins that have been implicated in maturation of the autophagosome (Oakes et al. 2017).

Table 1. 1 Genetics of ALS (adapted from Taylor et al., 2016).

Locus	Gene	Protein	Protein function	Mutations	Proportion of ALS		Reference
					Familial	Sporadic	
21q22.1	*SOD1*	Cu–Zn superoxide dismutase	Superoxide dismutase	>150	20%	2%	(Rosen et al. 1993)
14q11	*ANG*	Angiogenin	Ribonuclease	>10	<1%	<1%	(Greenway et al. 2006)
q36	*TARDBP*	TDP-43	RNA-binding protein	>40	5%	<1%	(Pokrishevsky et al. 2016)
16p11.2	*FUS*	FUS	RNA-binding protein	>40	5%	<1%	(Vance et al. 2009)
9p13.3	*VCP*	Transitional endoplasmic reticulum ATPase	Ubiquitin segregase	5	1–2%	<1%	(Johnson et al. 2010)
10p15-p14	*OPTN*	Optineurin	Autophagy adaptor	1	4%	<1%	(Maruyama et al. 2010)
9p21-22	*C9orf72*	C9orf72	Possible guanine nucleotide exchange factor	Intronic GGGGCC repeat	25%	10%	(Renton et al. 2011; DeJesus-Hernandez et al. 2011)
Xp11.23-Xp13.1	*UBQLN2*	Ubiquilin 2	Autophagy adaptor	5	<1%	<1%	(Deng et al. 2011)
5q35	*SQSTM1*	Sequestosome 1	Autophagy adaptor	10	<1%	?	(Fecto et al. 2011)
17p13.2	*PFN1*	Profilin-1	Actin-binding protein	5	<1%	<1%	(Wu et al. 2012)
12q13.1	*HNRNPA1*	hnRNP A1	RNA-binding protein	3	<1%	<1%	(Liu et al. 2016)
22q11.23	*CHCHD10*	Coiled-coil-helix-coiled-coil-helix domain-containing protein 10	Mitochondrial protein of unknown function	2	<1%	<1%	(Bannwarth et al. 2014)
TBK1	*TBK1*	Serine/threonine-protein kinase TBK1	Regulates autophagy and inflammation	10	?	?	(Freischmidt et al. 2015)

1.4.2. Excitotoxicity

Glutamate is the primary excitatory amino acid neurotransmitter in the CNS, produced in the astrocytes, neurons and oligodendrocytes and it is important in maintaining normal brain function (Fontana 2015; Peng et al. 2011). Release of glutamate into the synapse produces excitatory signal, which is terminated by the removal of glutamate from the

synaptic cleft through glutamate transporters, predominantly excitatory amino acid transporter 2 (EAAT2, also known as GLT1) in the CNS.

Glutamate-induced excitotoxicity is a specific form of neuronal injury resulting from increased levels of glutamate in the synaptic cleft, which leads to excessive activation of glutamate receptors or increase sensitivity of postsynaptic neurons to glutamate, resulting in disruption of intracellular calcium homeostasis and mitochondrial function and production of ATP leading to motorneuron injury (Soni et al. 2014; Ferraiuolo et al. 2011; Van Damme et al. 2005).

It has been shown that the knockout of GLT1 transporter, which is responsible for the clearance of synaptic glutamate, in transgenic mice results in neuronal death suggesting a potential role of glutamate-induced toxicity in ALS (Rothstein et al. 1996). The expression of GLT1 was downregulated in SOD1G93A mice with disease progression (Guo et al. 2010) and in human post-mortem ALS brain and spinal cord tissues compared with controls (J D Rothstein et al. 1995). It has also been shown that glutamate levels are elevated in CSF of some ALS patients, and the expression and activity of GLT1 receptors are reduced in the affected areas of the CNS (Shaw, Forrest, et al. 1995; Jeffrey D. Rothstein et al. 1995).

1.4.3. Oxidative Stress

Oxidative stress results from an imbalance between the generation and removal of reactive oxygen species (ROS) (Turner et al. 2013). Oxidative stress is considered as an important disease mechanism in ALS, particularly in *SOD1* mutation cases (Rosen 1993). TDP-43-related ALS cellular models have also shown oxidative stress in motor neuronal cell lines (Duan 2010).

In ALS, oxidative stress interacts with other pathophysiological processes such as excitotoxicity, mitochondrial dysfunction, protein aggregation, endoplasmic reticulum stress, and changes in cellular signaling pathways leading to motorneuron injury (Turner et al. 2013). Analysis of human cerebrospinal fluid (CSF) and post-mortem brain tissue from ALS patients have shown biochemical changes as a result of free radical damage or abnormal free radical metabolism compared to control cases (Shaw, Ince, et al. 1995; Ferrante et al. 1997). Additionally, elevated levels of oxidative damage to proteins, lipids, and DNA have been observed in post-mortem tissue from sALS and SOD1-linked fALS cases (Ferrante et al. 1997; Shibata 2001).

1.4.4. Mitochondrial Dysfunction

Mitochondria have a major role in producing the intracellular energy, and controlling several cellular pathways, such as normal apoptosis, calcium homeostasis and the generation of intracellular free radicals. Impaired mitochondrial function is considered a pathogenic mechanism involved in ALS. High levels of mitochondrial DNA mutations in brain tissue from sALS cases and diminished mitochondrial DNA in muscle and spinal cord in sALS represent an evidence of the involvement of mitochondrial dysfunction in ALS (Wiedemann et al. 2002; Dhaliwal & Grewal 2000).

Expression of mSOD1 proteins in motor-neuronal cell lines and aggregation of mSOD1 protein in mitochondria demonstrate a shift in the redox state of the mitochondria and impairment of respiratory complexes (Ferri et al. 2006). This suggests a strong involvement of defective mitochondria in mSOD1 and the pathogenesis of ALS (Pandya et al. 2013). In SOD1 models, the mSOD1 protein accumulates in mitochondrial intermembrane space, and leads to organelle dysfunction by impeding protein import (Liu

2004; Wong et al. 1995; Pasinelli 2004). It has also been shown that calcium buffering is impaired in mitochondria purified from the CNS of mSOD1 mice, which leads to altered calcium homeostasis in motorneurons (Damiano 2006).

Mitochondrial dysfunction also activates caspases and apoptosis which promotes neuron cell death in ALS disease (Sathasivam et al. 2005). It has also been reported that impaired mitochondria play a major role in ALS cases with *TARDBP* mutations (Wang et al. 2013). Axonal transport of mitochondria is impaired in animal models of ALS. The resulting impaired mitochondrial may lead to the axonal defect similar to what has been observed in many animal models of ALS (De Vos 2007; De Vos et al. 2008) (see 1.4.7).

1.4.5. Neuroinflammation

ALS is characterized by motorneuron death in the brain and spinal cord and the involvement of neuroinflammatory response through microglial activation and T-cell infiltrates in affected regions in the CNS (Kawamata et al. 1992). Additionally, patients with sALS show increased levels of circulating inflammatory CD16+ monocytes in peripheral blood with increased levels of lipopolysaccharide (Zhang et al., 2009). In ALS mouse models and human patient samples elevation in tumor necrosis factor (TNF)-α and Fas ligand (FASL) immunogenicity indicates proinflammatory responses and the involvement of inflammatory reactions in ALS disease pathogenecity (Glass et al. 2010).

1.4.6. Protein aggregation

One of the hallmarks of ALS is the dysregulation of protein localization and homeostasis. Several gene mutations including *SOD1*, *TARDBP*, and fused in sarcoma *(FUS)*, have been reported to form abnormal protein aggregates (Mackenzie et al. 2010; Keller et al.

2012). TDP-43 and FUS proteins mislocalization and aggregation has been observed in familial and sporadic ALS cases. Ubiquilin2 (UBQLN2), valosin-containing protein (VCP), and optineurin (OPTN) have also been linked to protein mislocalization and aggregation in cases with ALS disease (Johnson et al. 2010; Deng et al. 2011; Maruyama et al. 2010).

1.4.7. Impaired axonal transport

Axonal transport is required for the delivery of the essential materials such as RNA, proteins and organelles to the synaptic structures at the neuromuscular junctions (NMJ). In mSOD1 mice, impaired axonal transport has been observed early in the disease process, providing evidence of the involvement of the axonal transport as an early disease mechanism (De Vos et al. 2008; Kieran 2005; Williamson & Cleveland 1999).

1.4.8. Non-cell-autonomous toxicity

Investigating mouse ALS models based on SOD1 mutations indicates that the disease pathology is related to non-cell-autonomous mechanisms, through the involvement of glial cells and muscle rather than motorneurons only.

1.4.8.1.Role of glia cells in ALS

Microglial cells are activated in all types of ALS. The importance of glial cells in ALS emerged from a study observed that mSOD1 was silenced in microglia, astrocytes or oligodendrocyte precursor cells (Boillee 2006). Impairment in microglial function has also been implicated as a potential mechanism in *C9orf72* associated ALS, resulting from the reduction of *C9orf72* gene expression induced by methylation caused by the mutation and loss of C9orf72 function. C9orf72 protein inactivation in mice results in abnormal microglia and age-related neuroinflammation, providing an evidence that non-cell-

autonomous, microglial-mediated inflammation might contribute to ALS (ORourke et al. 2016; Burberry et al. 2016).

Astrocytes, another type of glial cells, provide MNs with nutrients, ion buffering and recycling of the neurotransmitter glutamate. Reduction of mSOD1 synthesis by astrocytes in mice resulted in slow disease onset or progression and a delay in the activation of microglia (Wang et al. 2011; Yamanaka 2008). In glutamate-mediated excitotoxicity, astrocytes limit the firing of motorneurons through clearing of glutamate from synaptic cleft through the GLT1 transporter. Failure of astrocytes to clear s glutamate triggers the leads to firing of MN action potentials and a subsequent increase in calcium influx, and ER stress (Bruijn et al. 1997; Howland et al. 2002; J D Rothstein et al. 1995).

1.4.9. Neurodegeneration-associated repeat expansion mutations

Neurodegenerative diseases can result from pathological expansion of a repetitive sequence in exons, intronic sequences, 5' or 3'-untranslated regions, resulting in coding or noncoding mutations. A number of hereditary diseases result from this mechanism, including fragile X syndrome (FXS), myotonic dystrophy (DM), Huntington's disease (HD), and spinocerebellar ataxia 1 (SCA1). Three mechanisms can be involved in disease pathogenesis:

1) Loss-of-function, as seen in FXS with CGG repeat expansion in *FMR1* leads to hypermethylation of the promoter region, transcriptional silencing and loss of FMRP protein (Bagni et al. 2012).

2) RNA mediated gain-of-function, as reported in DM1, due to CTG repeat expansion in the *DMPK* gene which results in the formation of nuclear RNA foci and sequestration of RBPs and consequent RNA splicing defects (Rahimov & Kunkel 2013). RNA foci have

been identified in other noncoding repeat disorders when transcription of the expanded repeat occurs and forms RNA/protein aggregates in the cell nucleus due to sequestration of RNA-binding proteins (RBPs) and formation of RNA secondary structures and hairpin loops by repeat sequences. It has been shown that some repeat expansion mutations can be transcribed bidirectionally in sense and antisense directions leading to the formation of toxic repeat sequences.

3) Protein mediated gain-of-function mechanism in CAG repeat expansion disorders such as Huntington disease (HD). In this disease the repeat in the *huntingtin* gene coding region is translated into a polyglutamine (poly-Q) region and lead to protein aggregation and disruption of the normal function of huntingtin protein (Nelson et al. 2013). Gain-of-function mechanism can also happen when repeat expansions in non-coding regions through production of toxic proteins by Repeat-Associated Non-ATG initiated (RAN) translation. This mechanism leads to translation in all reading frames in the absence of the AUG start codon next (Zu et al. 2011).

1.5. GGGGCC Hexanucleotide repeat expansion mutation (HRE)

An HRE mutation, GGGGCC (G4C2), in the C9orf72 gene is the commonest genetic cause of both ALS and FTD. The overall frequency of the mutation in familial FTD cases is about 25% and approximately 6-10% are sporadic cases, whereas it accounts for approximately 40% of familial ALS cases in European populations (DeJesus-Hernandez et al. 2011; Renton et al. 2011). Therefore, understanding the *C9orf72* gene and the underlying downstream pathways leading to ALS/FTD is crucial for the development of new therapeutic strategies.

The HRE is located either in the promoter region of transcript variant 2 or in the first intron of transcript variant 1 and 3 of the C9orf72 gene (Figure 1.1) (DeJesus-Hernandez et al. 2011; Renton et al. 2011). This suggests that the repeat expansion could influence the transcription of the *C9orf72* gene or affect its pre-mRNA processing and result in the reduction of C9orf72 expression levels (Almeida et al. 2013; Belzil et al. 2014; Ciura et al. 2013; Christopher J Donnelly et al. 2013; DeJesus-Hernandez et al. 2011).

Due to its high GC- content, the exact repeat expansion size can only be estimated by southern blotting. A recent study based on semi-quantification of repeats using SB showed that most control subjects carry fewer than 10 repeats, while ALS/FTD mutation carriers ranges between several hundreds to several thousand repeats (Suh et al. 2015). However, the exact lower number of the repeats causing pathogenicity is still unclear.

It has been also demonstrated that HRE sizes vary in different tissue and cell types (Beck et al. 2013; Suh et al. 2015), suggesting that repeat expansions might be caused by some cell division-independent mechanisms and somatic instability (Wen et al. 2016). For instance, it has been reported that a difference in repeat size between peripheral blood and CNS tissues was several thousands and that repeat size smaller than 16 repeats was somatically stable, suggesting that intrinsic tissue properties might affect the stability of the repeats (Wen et al. 2016; Nordin et al. 2015).

1.5.1. Clinical phenotypes and pathology of *C9orf72*

The age of onset for ALS/FTD patients carrying the expansion mutation is highly variable, from 27 to 83 years, with a disease duration ranging from 1-5 years (Majounie et al. 2012; Sabatelli et al. 2012). Some studies have suggested that ALS patients with a

C9orf72 mutation have a shorter survival than sporadic ALS (Cooper-Knock et al. 2012; Byrne et al. 2012). Neuropathological analysis shows TDP-43-positive inclusions in almost all *C9orf72* patients. In addition to TDP-43, p62-positive neuronal cytoplasmic inclusions was reported to be correlated with overall disease distribution in *C9orf72* patients (Al-Sarraj et al. 2011).

1.5.2. The underlying mechanisms of HRE-related neurodegeneration

The mechanisms by which G4C2 repeat expansions can cause neurodegeneration are still unclear, but two hypothesis have been proposed (Figure 1.3). Decreased C9*orf*72 expression has been observed in lymphoblasts from ALS/FTD patients with the HRE mutation (DeJesus-Hernandez et al. 2011), suggesting that repeat expansions can lead to downregulation of the C9orf72 protein and neurodegeneration through a loss-of-function mechanism (Ciura et al. 2013). The other suggested mechanisms are the gain-of-function mechanisms, which involves two main routes:

1. RNA-mediated toxicity directly driven by the HRE through the accumulation of nuclear sense or antisense foci of G4C2 RNA (DeJesus-Hernandez et al. 2011) which sequester RNA-binding proteins, leading to disturbances of cellular function (Gendron et al. 2013).

2. The sense and antisense RNA sequences can undergo unconventional non-ATG translation to generate dipeptide repeat (DPR) proteins, which leads to neurodegeneration by accumulation and producing cell toxicity.

1.5.2.1. Loss of function mechanisms

The exact function of the C9orf72 protein is still unclear. However, bioinformatic studies have demonstrated a homology of a region near the N-terminus (amino acids 108-229) in C9orf72 protein isoform A and B, except for the last eight amino acids absent in isoform B short protein, to the Differentially Expressed in Normal and Neoplasia (DENN)-like proteins domain, a class of GDP/GTP exchange factors (GEFs) that regulate activation and inactivation of the Rab-GTPase family by adding guanosine triphosphate (GTP) or guanosine diphosphate (GDP) to these proteins to mediate autophagy (Levine et al., 2013 and Zhang et al., 2012). Rab-GTPases have a role in membrane trafficking inside cells (Levine et al. 2013; Zhang et al. 2012), suggesting a function for C9orf72 in Rab GTPase-dependent endosomal membrane trafficking. Farg et al., 2014 identified several Rab proteins (Rabs 1, 5, 7 and 11) that co-localize with C9 and involved in endocytosis and autophagy in neuronal cell lines and ALS spinal cord tissue (Farg et al. 2014). Moreover, knockdown of the C9orf72 gene in neuronal cell lines results in impaired endocytosis and autophagy-mediated trafficking, which could also suggest that *C9orf72* downregulation is a potential pathogenic mechanism in ALS/FTD (Farg et al. 2014). Almeida et al. (2013) also observed that iPSC-derived MNs carrying the HRE were more sensitive to autophagy inhibiting drugs (Almeida et al. 2013).

The repeat expansion is located in the 5′ regulatory region upstream of the transcription start site in V2 and within intron 1 of the other two transcripts (Figure 1.1 & 1.2), suggesting that the presence of the mutation could affect the transcription of C9orf72 or the processing of precursor mRNA variants. It has been shown in several studies that the HRE mutation is associated with reduced expression level of all C9orf72 transcripts (Almeida et al. 2013; Belzil, Gendron, et al. 2013; Ciura et al. 2013; DeJesus-Hernandez

et al. 2011; Donnelly et al. 2013; Waite et al. 2014). The reduction of expression of all transcripts and not just repeat-containing sequences could be related to epigenetic changes related to the presence of repeat expansions.

Downregulation of *C9orf72* expression might lead to haploinsufficiency if the transcription of the wild-type allele is not sufficient to produce enough C9orf72 protein. The downregulation is either related to DNA hypermethylation or histone methylation in the promoter region of *C9orf72* gene. Several groups have observed that the reduction in *C9orf72* mRNA transcripts in human post-mortem brain tissue, blood, cortex, cerebellum and spinal cord, and lymphoblasts of mutation carriers compared to controls (Renton et al. 2011; Donnelly et al. 2013; Belzil, Bauer, et al. 2013; Mori, Weng, et al. 2013).

Knockdown of *C9orf72* in neuronal cell lines showed defects in endocytosis and autophagy and gene knockout in *C. elegans* and zebrafish lead to downregulation of *C9orf72* and motor deficits with neurodegeneration (Ciura et al. 2013; Farg et al. 2014; Therrien et al. 2013). It has also been shown in previous studies that the *C9orf72* expression level is decreased in iPSC-derived MNs from ALS/FTD patients carrying the HRE mutation (Almeida *et al.*, 2013; Donnelly *et al.*, 2013; Waite *et al.*, 2014; Tran *et al.*, 2015). This would suggest that the loss-of-function mechanism might play a crucial role in disease pathogenesis.

1.5.2.1.1. C9orf72 DNA promoter hypermethylation

Epigenetic alterations including DNA methylation, histone modification, chromatin remodeling, and microRNA activity, result in changes in gene expression and function

(Mehler, 2008). DNA methylation regulates mammalian gene expression and has an important role in cell development and differentiation. DNA hypermethylation has been identified in many human neuromuscular and neurodegenerative diseases such as Fragile X syndrome, Huntington's disease, Friedreich's ataxia, spinocerebellar ataxias and ALS/FTD (Dion & Wilson 2009; Belzil et al. 2014; Martin & Wong 2013; Qureshi and Mehler, 2013a and Yang et al., 2010).

The methylation of the CpG sites can be varied between individuals but stable over time (Feinberg et al. 2010). DNA methylation is an important mechanism for gene expression in eukaryotes, mostly affects Cytosine (C) when it is followed by a guanine (G). DNA methylation involves the addition of methyl groups (–CH3) to the 5′ position of cytosines catalyzed by specific enzymes called novo DNA methyltransferases (DNMTs) in the presence of ATP and S-adenosylmethionine as methyl donor, resulting in 5-methylcytosine (5mC). DNMT1 is an enzyme responsible for the maintenance of DNA methylation and sustaining the methylation after DNA replication (Long et al. 2014). DNMT3a and DNMT3b enzymes are responsible for *de novo* DNA methylation (Wienholz et al. 2010).

The most common consequence of DNA methylation is genes silencing especially at the promoter regions, where CpG sites can be clustered as CpG islands (CGIs) of >200 bp with >50% CG content. These islands are normally not methylated in active genes and when methylated leads to silencing of gene expression.

The G4C2 repeat sequence is flanked by two CpG islands (CGIs) within a ~1-kb region that spans from the promoter sequence into intron 1 of the *C9orf72* gene. Xi et al. (2013) studied two CpG islands flanking the HRE region in a cohort study of 37 ALS

HRE carriers, 64 ALS patients noncarriers and 76 control subjects. The CpG island was unmethylated in 97% of noncarriers cases while it was methylated in 73% of HRE carrier cases. No methylation was observed in control subjects.

A high level of CpG hypermethylation has been identified in blood, cerebellum, frontal cortex and spinal cord of ALS patient samples carrying the HRE mutation (Belzil et al. 2014; Liu et al. 2014; Xi et al. 2016). Methylation of the promoter region in CpG islands of the *C9orf72* gene results in the reduction of all transcript levels due to silencing of gene expression. This has been associated with reduced RNA foci and DPR accumulation (Liu et al. 2014). These findings would suggest that DNA promoter hypermethylation might have a protective effect through inhibiting the production of C9orf72-related pathogenic products. Neuroimaging studies have also shown that patients with C9 hypermethylation have less brain atrophy and neuronal loss (McMillan et al. 2015). Further evidence comes from the observation that promoter hypermethylation is associated with later age at death and longer disease duration (Russ et al. 2015; Xi et al. 2013).

Furthermore, methylation has been observed in the HRE region itself in blood and brain samples from repeat expansion carriers, since the G4C2 sequence provides an additional CpG island where methylation can occur (Xi et al. 2015). The methylation in the repeat itself is more common in larger repeat sizes compared to smaller ones. In a previous study, patients with repeat expansion sizes over than 90 repeats were found to be methylated in different tissues, whereas no methylation was observed in cases with G4C2 size of less than 90 repeats (Xi et al. 2015).

Attachment of methyl groups to histone tails can also be involved in gene silencing through reduction of *C9orf72* gene levels. Histone H3 and H4 trimethylation at the

C9orf72 locus was detected in blood samples of HRE mutation carriers (Belzil et al. 2013) which also suggests another potential mechanism of downregulation of *C9orf72* gene levels (Belzil et al. 2013). The reduction of C9 gene expression was rescued by treatment of patient fibroblasts with 5-AZA demethylation agent, which confirmed that the repression of the C9orf72 gene in HRE carriers is caused by the binding of mutant C9orf72 to trimethylated lysine residues within H3 and H4 histones (Belzil, Bauer, et al. 2013).

Another study demonstrated that increased CpG methylation in patient-derived lymphoblasts is correlated with a reduction in RNA foci and DPR inclusion levels, suggesting that methylation may be protective. Demethylation of these lymphoblast cells increased their sensitivity to oxidative and autophagic stress which also confirms a protective role of CpG promoter methylation and supports a model where expression of *C9orf72* is toxic due to the generation of RNA aggregates and DPRs (Russ et al. 2015).

One case-study found that the clinical and pathological features of disease were less severe in a homozygous expansion FTD patient than heterozygous expansion-carriers (Fratta et al. 2013). This might suggest that *C9orf72* haploinsufficiency produces less severe disease phenotype in homozygous carriers compared to heterozygotes and the toxic gain-of-function mechanism is the major cause of neurodegeneration in C9orf72 patients.

The HRE is able to form DNA and RNA G-quadruplex secondary structures and RNA:DNA hybrid stable intramolecular structures *in vitro* and *in vivo* (Fratta et al. 2012). In these structures, four guanines interact through Hoogsteen bonding to form G-quarters, and guanine residues can interact with other guanines to form unimolecular and multimolecular G-quadruplexes. G-quadruplexes are known to be normal regulators of

some biological processes such as genetic instability, gene regulation, RNA translation and telomere regulation (Rhodes & Lipps 2015).

RNA-DNA hybrids (R-loops) and G-quadruplexes formed by HRE play an important roles in initiating nucleolar stress. They induce RNA transcription impairment by impeding RNA polymerase transcription, which leads to decrease the formation of full-length transcripts and accumulation of repeat-containing transcripts (Haeusler et al. 2014). The accumulation of repeat-containing RNA further forms stable RNA G-quadruplexes (Haeusler et al., 2014; Reddy et al., 2013). G-quadruplexes interact with a variety of essential RNA-binding proteins (RBPs) and sequester them into RNA foci in neurons of HRE carriers, which makes them unable to bind and regulate RNA processing, which may lead to RNA dysfunction (Gendron et al. 2014; Lee et al. 2013).

1.5.2.1.2. Gene inactivation studies

A single *C9orf72* orthologue in zebrafish (z*C9orf72*) has been used to study *C9orf72* function (Ciura et al. 2013). Zebrafish embryos were injected with three different morpholino antisense oligonucelotides (ASOs), to lower z*C9orf72* levels by blocking either the translation or the splicing of the z*C9orf72* gene. Oligonucelotide treatment resulted in motor deficits and shortening of MN axons, which indicates that loss-of-function of *C9orf72* gene could impair motorneuron function. Another loss-of function model is a null mutation in the *C. elegans C9orf72* orthologue which resulted in motorneuron degeneration and motility deficits (Therrien et al. 2013). However, in one study two ALS patients with HRE homozygous mutations harboring the repeats on both alleles showed clinical and pathological findings similar to heterozygous cases. These findings do not support the principle of the complete loss of function as a major disease mechanism.

ASOs targeting mouse *C9orf72* were injected by intracerebroventicular (ICV) injection and resulted in 30-40% reduction in *C9orf72* mRNA levels in the spinal cord and the brain (Lagier-Tourenne et al. 2013). The *C9orf72* levels remained lower several months after initial injection but it did not result in any behavioural phenotype or motor impairment. Conditional ablation of *C9orf72* in neurons and glia in mice resulted in RNA foci formation but it did not show neurodegeneration or defects in motor function (Koppers et al. 2015). Two recent studies have generated homozygous knockout mice (Atanasio et al. 2016; ORourke et al. 2016). These mice did not develop motorneuron disease and neurodegeneration but they developed splenomegaly and deficits in the immune response and microglial function, which could suggest that altered immune regulation might contribute to disease development.

1.5.2.2.Toxic gain of function mechanisms

This hypothesis involves the following mechanisms:

1.5.2.2.1. Toxicity related to RNA Foci accumulation

RNA foci represent the first described phenotype in C9orf72 disease, and can be visualized using RNA fluorescence in-situ hybridization (FISH). Repeat expansions can be transcribed bidirectionally in the pre-mRNA of transcript 1 and 3 and the presence of the expansions prevents the normal splicing into mature mRNA, so that transcribed repeat-containing RNA sequences could accumulate in the nucleus and produce toxicity. These aggregates are called RNA foci and they have been identified in brain post-mortem specimens of patients carrying the HRE mutation in cerebellum, cortex, and hippocampus, as well as in the spinal cord post-mortem tissue samples and in iPSC-derived MNs (DeJesus-Hernandez et al. 2011; Christopher J. Donnelly et al. 2013; Almeida et al. 2013; Sareen et al. 2013; Mizielinska et al. 2013; Lagier-Tourenne et al.

2013; Zu et al. 2013), which makes them a potential target developing therapeutic strategies. Antisense RNA foci were discovered later and also detected in frontal cortex, spinal cord, and cerebellum of *C9orf72* patients (Mizielinska et al. 2013; Zu et al. 2013; Cooper-Knock, Higginbottom, et al. 2015).

Accumulation of sense and antisense RNA sequences could sequester RNA binding proteins in foci inside the nucleus. Many RNA-binding proteins have been suggested to be sequestered by the repeats, including hnRNP H1/F, hnRNPA3, hnRNPA1, hnRNP-H, nucleolin, Pur-α, FUS, ADARB2, and RanGAP1 (Donnelly et al. 2013; Haeusler 2014; Lee et al. 2013; Sareen et al. 2013; Cooper-Knock, Shaw, et al. 2014; Li et al. 2015). In a Drosophila model, the repeat expansions were bound to the RNA binding protein Pur α, and resulted in neurodegeneration (Xu et al. 2013). Expression of G4C2 repeats of 38 and 72 length resulted in RNA foci formation and cell death in neuronal cell lines and in zebrafish (Lee et al. 2013).

1.5.2.2.2. RAN translation and dipeptide repeats

Repeat associated non-ATG (RAN) translation is an unconventional form of translation, forming stable stem-loops or hairpins (Cleary & Ranum 2013; Zu et al. 2013). It occurs using all reading frames and results in long peptides of five DPR proteins, as observed in human post-mortem brain and cultured patient iPSC-derived MNs (Ash et al. 2013; Christopher J Donnelly et al. 2013; Mori, Arzberger, et al. 2013). Translation of sense G4C2 repeat gives rise to poly-glycine-alanine (GA), poly-glycine-proline (GP), and poly-glycine-arginine (GR), whereas the translation of antisense C4G2 reverse transcript results in poly-proline-arginine (PR), poly-glycine-proline (GP), and poly-proline-alanine (PA), which have been proved to be toxic in vitro leading to cell death (Ash et al. 2013; Christopher J Donnelly et al. 2013; Mori, Arzberger, et al. 2013; Gendron et al.

2013; Zu et al. 2013).

Previous studies provide evidence that DPR proteins have an important role in the neurodegeneration process. Poly-(Gly-Arg) and poly-(Pro-Arg) are observed to bind prominently to nucleoli of cultured cells, impeding RNA biogenesis, and SG formation (Tao et al. 2015; Kwon et al. 2014). In the nucleolus, DPRs induce translocation of a key nucleolar component nucleophosmin, which results in nucleolar stress and cell death (Tao et al. 2015). It has also been reported that the expression of poly-(Gly-Ala) induces neurotoxicity via inhibiting proteasome activity and activating ER stress (Zhang et al., 2014) and inducing apoptosis in primary cortical and hippocampal neurons (May et al. 2014).

Two independent studies have observed the accumulation of DPRs and RNA foci in transgenic mice with *C9orf72* repeat expansion mutation, but without developing neurodegeneration or motor abnormalities (O'Rourke et al. 2015; Peters et al. 2015), suggesting that *C9orf72*-associated pathology can occur presymptomatically and additional pathological mechanisms are required to induce the neurodegenerative process.

Four independent studies have shown that *C9orf72*-associated HRE and DPR proteins disrupt nucleocytoplasmic transport of RNA export and nuclear protein import (Boeynaems et al., 2016; Freibaum et al., 2015; Jovičić, Paul, & Gitler, 2016; Zhang et al., 2015). All four studies have suggested that nucleocytoplasmic transport is a potent mediator of *C9orf72* hexanucleotide expansion-associated cellular toxicity. In these studies, alterations of nuclear import or export lead to neurodegeneration as a result of genetic modulation of a variety of key regulators within the nucleocytoplamic transport pathway.

Although each of these mechanisms have been identified and supported with some evidence, the exact disease pathogenesis of either loss or gain-of-function or combination of both mechanisms still unclear and more studied need to be conducted.

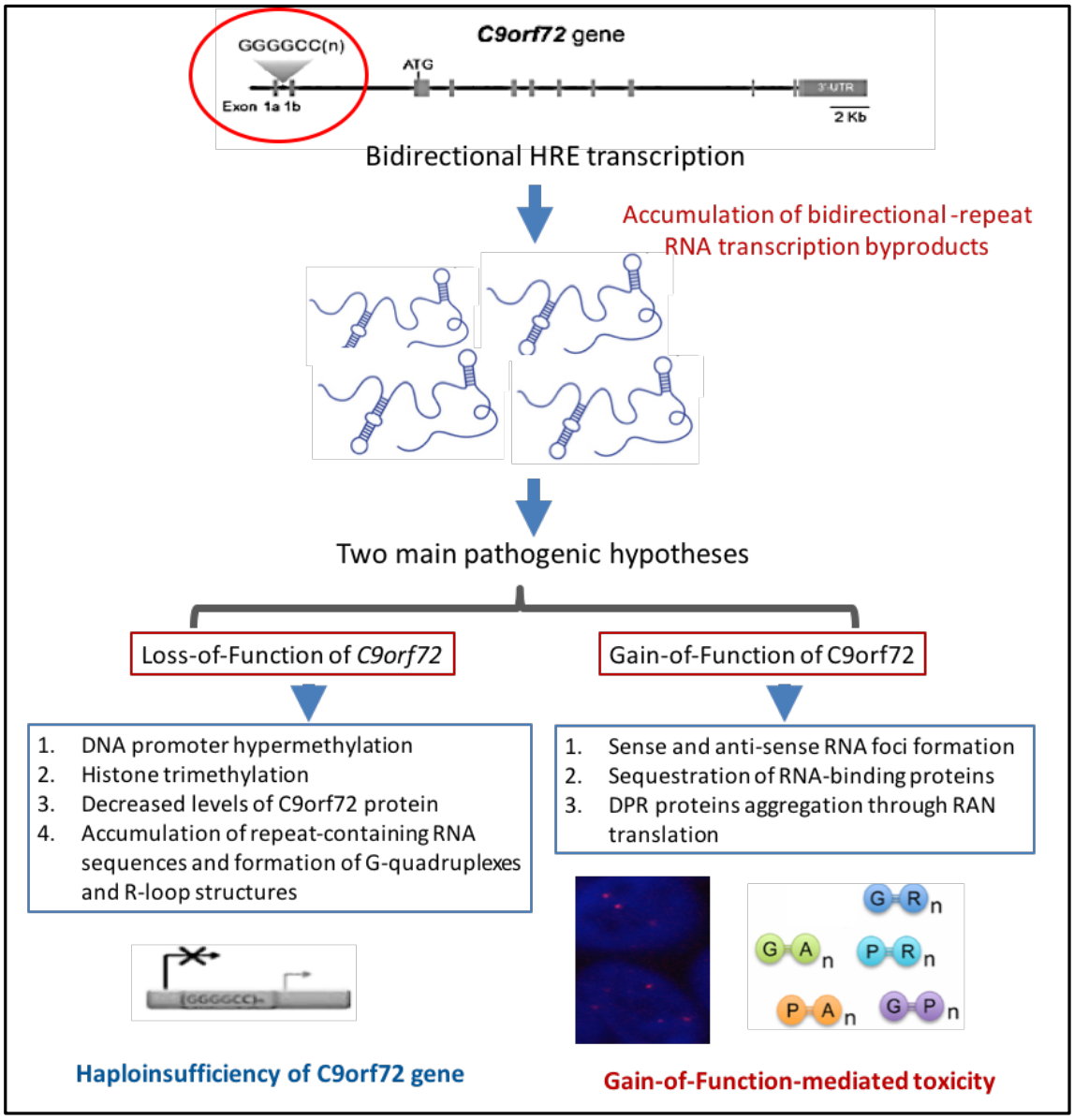

Figure 1. 3 Potential disease mechanisms of HRE mutation. (1) The repeat expansion might lead to a reduction in *C9orf72* gene expression and loss of protein function. (2) The transcribed sense and antisense repeat RNA sequences accumulate and produce sense and antisense RNA foci and may cause toxicity and sequestration of RNA-binding proteins, which results in a toxic gain of function mechanism or (3) The repeat can undergo repeat-associated Non-ATG translation into five potential DPR proteins, poly-GA, poly-GR, poly-GP, poly-PA and poly-PR leading to a toxic protein gain of function. Adapted from (Taylor et al. 2016; Rohrer et al. 2015).

1.5.2.3.Susceptibility to Glutamate toxicity

It has been shown that ALS patients have reduced levels of astroglial glutamate transporter 1 (Lin et al. 1998; Jeffrey D. Rothstein et al. 1995). iPSC-derived neurons from HRE patients showed increased vulnerability to glutamate excitotoxicity in a dose-dependent manner (Christopher J Donnelly et al. 2013) and treatment of these cells with ASOs targeting the repeat sequence rescued the sensitivity of cells to glutamate (Christopher J Donnelly et al. 2013). It has also been shown that expression of HRE resulted in increased vulnerability to cellular stressors such as autophagy inhibitors in iPSC-derived MNs and reduction in cell viability of *C9orf72* iPSC-derived MNs when exposed to chloroquine, an autophagy pathway inhibitor (Almeida et al. 2013).

1.5.2.4.Ribonucleoprotein Stress granules

Ribonucleoprotein (RNP) granules are formed through RNA processing and have distinct roles in mRNA regulation (Anderson and Kedersha, 2008; Kedersha et al., 2005). The main types of RNP-containing granules are transport RNPs, processing bodies (p-bodies) and stress granules (SGs). The formation of SGs is part of the protective response to various cell stresses, such as oxidative, mitochondrial, proteosomal and viral stress. SGs may sequester unwanted mRNAs such as viral RNAs during infection (Beckham and Parker, 2008), or stress conditions (Li et al., 2013; Unsworth et al., 2010; Wolozin, 2012). GA proteins have shown neuronal inclusions with positive staining for the autophagy-related protein p62 (Schludi et al. 2015). It has also been observed that SGs can sequester proteins such as PABP-1 to localize to TDP-43 positive inclusions in patients carrying the *C9orf72* mutation (McGurk et al., 2014).

1.5.3. Therapeutic strategies for C9orf72 neurodegeneration

Several therapeutic strategies have been developed for targeting *C9orf72* mediated neurodegeneration by downregulation of *C9orf72* RNA transcripts, to prevent the toxic effect of RNA foci and DPRs.

1.5.3.1. Antisense oligonucleotides (ASOs)

Have been developed to bind to the target RNA and initiate RNase H-dependent degradation of *C9orf72* RNA, or prevent RNA-binding protein sequestration by interfering with G-quadruplex structure and reduction of the pathogenic effect of the repeat expansion in fibroblasts and iPSC-derived MNs from *C9orf72* ALS patients (Christopher J Donnelly et al. 2013). However, the majority of ASOs showing reduction of sense RNA foci and rescue of *C9orf72* toxicity have also reduced *C9orf72* transcription level (Christopher J Donnelly et al. 2013; Sareen et al. 2013; Lagier-Tourenne et al. 2013).

1.5.3.2. Small molecules

Small molecules have also been developed to target RNA secondary structures formed by the C9orf72 repeat expansion and led to inhibition of RNA foci formation and RAN translation in neurons derived from patients with *C9orf72* mutation. A cationic porphyrin that target G-quadruplexes, named TMPyP4, is able to bind to *C9orf72* RNA repeat (Zamiri et al. 2014). Another group have also identified small molecule ligands that bind to the HRE and decrease the formation of RNA foci and RNA translation in cultured cells and neurons derived from patient carrying the HRE mutation (Su et al. 2014).

However, most of the current therapeutic strategies using ASOs and small-molecule based approaches are focused on gain of function mediated toxicity mechanism. The ideal therapy may need to focus on maintaining the wild type allele expression level and decreasing the mutant allele expression.

1.6. Modelling ALS using induced pluripotent stem cells (iPSCs)

Understanding ALS disease mechanisms has been progressing slowly due to the difficulty in obtaining cell-types directly involved in disease pathology. The only available source is the post-mortem tissues that represent the end-stage of the disease in which MNs have gradually degenerated, which makes it difficult to study the early-stage of the disease.

The development of iPSC technology has provided the opportunity for disease modelling to overcome the limitations of using animal models and to enable modelling of familial and sporadic disease cases using patient-specific cells. The use of iPSC technology has allowed us to increase our understanding of the key pathological processes in neurodegenerative diseases and to accelerate the search for effective therapeutics. In this project, CRISPR/Cas9 genome editing was applied to excise the abnormal repeat expansion on the mutant allele and replace it with the wild-type repeat size by homologous recombination using a donor template carrying the normal repeat size. This strategy provides an attractive opportunity to correct disease phenotypes, restore normal regulation of gene expression and ameliorate the features of neurodegeneration *in vitro*.

1.7.1. Pluripotent Stem Cells

1.7.1.1. Human embryonic stem cells (ESCs)

Human embryonic stem cells (ESCs) are pluripotent cells derived from the inner cell mass of a developing blastocyst of a 5-day-old pre-implantation embryo. They have two unique characteristics; they can expand indefinitely in culture without losing their potential and they can differentiate into any cell type of the three germ layers, which makes them a popular cell source for regenerative medicine and disease modelling.

Mouse ESCs (mESCs) were the first isolated pluripotent stem cells from early embryos.

In 1981, two independent groups have established the generation of mESC cultures from mouse blastocysts (Evans & Kaufman 1981; Martin 1981). In 1998, Thomson and colleagues were the first to report the generation of embryonic stem cells (ESCs) from the inner cell mass (ICM) of human blastocysts produced through *in vitro* fertilization (IVF) (Thomson 1998). ESCs are a source of non-autologous cells and their therapeutic use in transplantation may require immunosupression. In addition, the ethical issues associated with obtaining oocytes restrict their use in practical applications.

iPSCs possess the same properties as ESCs, in terms of indefinite proliferation and differentiation into any adult cell type. These properties make iPSCs a promising source for research, including drug screening and other therapeutic purposes such as transplantation.

1.7.1.2. Induced pluripotent stem cells (iPSCs)

In 2006, Yamanaka and Takahashi (Takahashi & Yamanaka 2006) developed a revolutionary technology by inducing the pluripotent state in mouse somatic cells through ectopic expression of four genes using retroviral vectors: Octamer-Binding Transcription Factor 4 (*Oct4*), SRY (Sex Determining Region Y)-box 2 (*Sox2*), Kruppel-Like Factor 4 (*Klf4*) and Avian Myelocytomatosis Viral Oncogene Homolog (*c-Myc*). These four genes are now known as the "Yamanaka Factors" or OSKM factors. A year later, human iPSCs were generated by the same group and using the same reprogramming protocol (Takahashi et al. 2007). Some of the transcription factors used in reprogramming process Oct3/4, Sox2, and Nanog are normally involved in maintaining pluripotency in early embryos (Nichols et al. 1998; Mitsui et al. 2003; Avilion et al. 2003; Chambers et al. 2003). Furthermore, some genes involved in pluripotency such as C-myc and Klf4 are also up-regulated in tumors, (Cartwright et al. 2005).

Retroviral and lentiviral methods are commonly used to deliver reprogramming factors to somatic cells, which requires the integration into the host genome carrying the risk of genetic instability and reactivation of the reprogramming factors. Reactivation of c-myc can induce tumour formation in mice. To overcome this issue, efficient and integration-free reprogramming protocols have been established by transient expression of reprogramming factors. These methods involved adenoviruses, plasmids, transposons, Sendai viruses, synthetic mRNAs and recombinant proteins. Sendai viruses and synthetic mRNAs are currently the most common protocols for generating integration-free iPSCs (Ban et al. 2011; Tucker et al. 2013; Fusaki et al. 2009). Transduction with Sendai virus is transient and viral particles can be removed from the infected cells following passaging.

In addition, several non-viral reprogramming protocols have also been developed. Transient transfection with episomal plasmid DNA (Fontes & Lakshmipathy 2013; Lorenzo et al. 2013) and synthetic self- replicating RNA (Yoshioka et al. 2013) represent promising approaches to cell reprogramming without affecting the host genome. Another way to avoid integration is mRNA transfection, which is a safe and effective method of generating iPSCs with a reduced oncogenic potential (Warren et al. 2010). Reprogramming using small molecules represents an attractive way to improve reprogramming efficiency (Maherali et al. 2008). Table 1.2 summarizes the most common methods of inducing pluripotency in somatic cells (adapted from Seah, et al, 2015).

In addition to fibroblasts, iPSCs have also been generated from other cell sources including blood cells, hair follicles and epithelial cells (Staerk et al., 2010; Zhou et al., 2012). The reprogrammed iPSCs have demonstrated similar characteristics to hESCs, in their capacity for self-renewal, the ability to differentiate into any cell type of the three germ layers (endoderm, mesoderm, and ectoderm), and the capacity for teratoma

formation *in vivo*.

Table 1. 2 Methods of inducing pluripotency in somatic cells (adapted from Seah, et al, 2015).

Method Type	Method	Factors/Other Agents	Cell Type	References	Efficiency
Transgene-based	Retroviral	OSKM	Mouse fibroblasts	(Takahashi & Yamanaka 2006)	0.02%
	Retroviral + small molecules	OSKM + SB431542 + PD0325901	Human fibroblasts	(Lin et al. 2009)	~1%
	Lentiviral	OSNL	Mouse fibroblasts, human fibroblasts	(Yu et al. 2007)	0.0095%
	Small molecules	C6FZ or VC6TFZ	Mouse fibroblasts	(Maherali et al. 2008)	0.2%
Transgene-free	Episomal plasmids	OSKM	Mouse fibroblasts	(Okita et al. 2007)	~0.1%
	Sendai viruses	OSKM	Human fibroblasts	(Fusaki et al. 2009)	~1%
	Non-integrating DNA adenoviral	OSKM	Tail tip fibroblasts, hepatocytes, fatal liver cells	(Stadtfeld et al. 2008)	0.0001%–0.001%
	PiggyBAC transposon	OSKM or OSKML	Mouse fibroblasts	(Yusa et al. 2009)	~1%
	Synthetic mRNA	OSKM	Human fibroblasts	(Warren et al. 2010)	2%
	MicroRNAs (lentiviral)	miR302/367	Mouse fibroblasts, human fibroblasts	(Anokye-Danso et al. 2011)	Up to 10%

1.7.1.2.1. Advantages of iPSCs

iPSCs represent an exciting cell source for disease modeling due to their ability to self-renew and to differentiate into any source of cells. iPSC generation does not require the destruction of human blastocysts, avoiding the major ethical issues encountered with using human ESCs. They also have the capacity to generate a long term source of inaccessible cells. Furthermore, iPSCs have been utilized in detecting and analysing molecular changes related to early disease phenotypes, including alterations in gene expression and cellular signalling pathways and other underlying disease mechanisms.

iPSCs share the same genetic background as the original cells, which makes them a promising source for the development of future transplantation therapies. Moreover, iPSCs can be genetically manipulated to remove or insert mutations to generate isogenic lines and directly study the disease .

1.7.1.2.2. Applications of iPSCs

Recent progresses in the iPSC field have opened up many avenues for the medical research. iPSCs showed the ability to recapitulate disease phenotypes *in vitro*, which is important in disease modelling and studying the underlying mechanisms. iPSCs have been used in several applications for disease modelling, drug screening and genetic engineering (Figure 1.3).

i. *Disease modelling:* the ability of iPSCs to self-renew and to differentiate into any somatic cell types make them an attractive source for disease modelling. Many studies have already used iPSCs in disease modelling and studying the disease underlying mechanisms.

ii. *Gene editing studies:* incorporation of genome editing and creation of isogenic lines and normalization of disease phenotypes can provide evidence that the observed

phenotypes are caused by the presence of the mutation under study. Additionally, disease-causing mutations can be introduced into iPSCs from healthy donors for the study of disease-associated phenotypes. In this way, iPSCs are providing the opportunity to study the molecular mechanisms of genetically complex and rare disorders.

iii. *Drug discovery using iPSC-disease modelling:* iPSCs are a promising model for drug discovery and identification of new therapeutic small molecule compounds by large-scale drug screening. This application requires differentiation towards a specific cell type affected in the disease under invistigation.

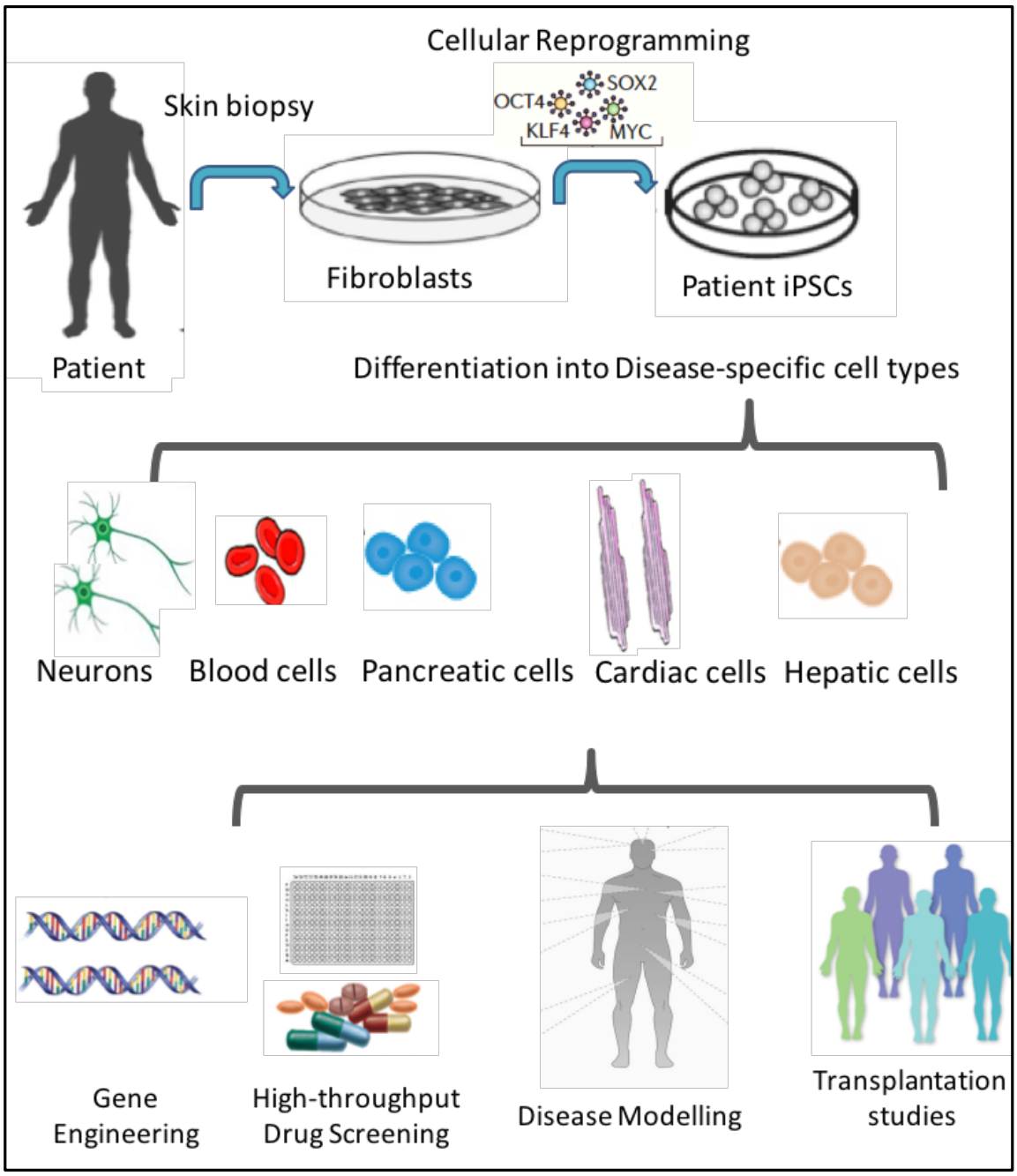

Figure 1. 4 Generation of patient-specific induced pluripotent stem cells (iPSCs) and their clinical applications. Generation of patient-specific iPSCs from somatic cells can be achieved by cell transduction with Yamanaka factors. iPSCs can be used in gene editing of genetic defects using several gene editing techniques such as zinc finger nucleases (ZFNs), activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats (CRISPR) system. iPSCs can also be used in disease modeling, drug screening, and stem cell therapy in the future.

1.7.1.2.3. Challenges in iPSC-disease modelling

Heterogeneity across different cell lines is the main challenge in iPSC disease modelling. These heterogeneity changes can be related to reprogramming methods, variations between different cell lines, culture conditions or cell passaging. A large-scale karyotype analysis was carried out on more than 1,700 human ESC and iPSC cultures from different laboratories (Moralli et al. 2011). Results showed similar karyotype abnormalities in both sources, suggesting that the presence of abnormal genetic aberrations is independent of the reprogramming procedure and indicating that the reprogramming process itself is mutagenic. Forced expression of reprogramming factors leads to genetic and epigenetic changes where chromatin re-organises to ESC-like state (Apostolou & Hochedlinger 2013). It has been reported that iPSCs retain some epigenetic "memory" from the original somatic cells (Kim et al. 2010).

The selection of study controls is critical in iPSC modelling due to the variation between different iPSC clones. The optimal control standards is to use both healthy control lines and gene-edited isogenic controls derived from the original disease lines (Soldner et al., 2011).

1.7.2. Studying ALS using iPSC-derived motorneurons

The majority of the established ALS animal models were based on the transgenic overexpression of genes causing familial ALS. These models have recapitulated several ALS disease phenotypes but they represent a subset of the disease-related mutation not the whole disease (McGoldrick et al. 2013). Recapitulating sporadic ALS disease using animal models has been limited due to the disease complexity and cell heterogeneity involved in ALS disease. Moreover, studying a slowly progressive diseases as ALS requires months of study, which is highly expensive and tedious. Additionally, several

therapeutic strategies have been developed and successfully applied on ALS animal models but they have failed when applied in human clinical trials.

iPSCs technology has opened the avenue to develop disease-specific cell types, offering the opportunity to understand and explore the underlying disease mechanisms that contribute to pathogenicity of the disease. iPSCs are able to differentiate into any cell types of the three germ layers, including neurons. In particular, the ability to produce neurons from iPSCs provides unprecedented advantages to generate patient-specific iPSC-MNs that carry the same genetic mutations and from the same genetic background. Multiple differentiation protocols into motorneuron lineage have been developed for modeling ALS and FTD to investigate the disease mechanisms *in vitro*.

Numerous cell types play a role in ALS disease pathogenesis and progression, including astrocytes, oligodendrocytes, and microglia. The contribution of these cell types has been demonstrated primarily in mice models overexpressing the mutant SOD1 gene (Ilieva et al. 2009; Papadeas et al. 2011). As mentioned earlier in this chapter, multiple genetic mutations and pathophysiologic processes have been implicated in ALS. Of the multiple gene abnormalities identified in ALS, iPSCs have been used for modelling the most important mutations such as *SOD1*, TDP-43 and *C9orf72*.

1.7.2.1.Modeling SOD1 mutations using iPSCs

Patient-specific iPSC-derived MNs from cases with SOD1 mutations provide an attractive opportunity to study the neurodegenerative features of ALS cases with SOD1 mutations. iPSC from ALS patients with SOD1 mutation were used to identify neurofilament aggregation and neurite degeneration in spinal MNs (Chen et al. 2014).

Mutant SOD1 iPSC-derived MNs showed a sequestration of neurofilament subunit mRNA, resulting in less mRNA available for translation and less protein produced and indicating a novel mechanism for axonal degeneration in SOD1 MNs. iPSC-derived MNs from two patients with SOD1A4V mutation demonstrated that MNs have a higher tendency to undergo apoptosis, similar to those observed in the SOD1G93A mouse model (Kiskinis et al. 2014).

1.7.2.2. Modeling TDP-43 mutations

iPSC-derived MNs from ALS patients carrying TARDBP mutations have been generated and studied to understand the subcellular localization of TDP-43 proteins and to determine if these mutations directly contribute to the disease pathogenesis. iPSC derived MNs from patients with TARDBP mutations showed an increased levels of soluble and detergent-resistant TDP-43 and decreased cell survival (Bilican et al. 2012). Neuronal cultures from ALS patient iPSCs showed TDP-43 aggregates in the cytoplasm and the same morphological characteristics observed in animal ALS models (Egawa et al. 2012).

TDP-43 M337V mutated iPSC-derived MNs from ALS patient have demonstrated an increase in cell death, which is in line with the previously reported cellular and transgenic TDP-43 models when elevated levels of cytoplasmic TDP-43 correlates with cellular toxicity (Barmada et al. 2010). Knockdown of TDP-43M337V motorneurons with siRNA results in a significant reduction of the accumulated cytoplasmic TDP-43 proteins (Nishimura et al. 2014).

Egawa and colleagues have generated iPSCs from three ALS patients carrying TDP-43 mutations (Q343R, M337V, and G298S). iPSC-derived MNs showed cytosolic aggregates containing TDP-43, similar to what has been found in pathological brain

samples from ALS patients. iPSC-derived MNs were also used to study cell stress after treating cells with arsenite and they showed increased sensitivity to arsenite when compared to healthy controls (Egawa et al. 2012).

MNs in iPSC derived cultures from both TARDBP and C9ORF72 ALS patients showed hyperexcitability followed by loss of action potential at the synapse. This was associated with a decrease in voltage dependent Na^+ and K^+ movement, suggesting reduced or dysfunctional ion channels (Devlin et al. 1AD).

1.7.2.3. Modeling C9orf72 mutations using iPSC-derived neurons

Almeida *et al.*, 2013, were the first group to generate iPSC-derived MNs from *C9orf72* HRE patients. iPSC derived neurons showed the ability to recapitulate disease-related phenotypes previously observed in brain samples, including the presence of RNA foci and the formation of DPR proteins (Almeida et al. 2013). The presence of RNA foci in iPSC-derived MNs from HRE patients may affect RNA metabolism through sequestration to some RNA binding proteins, such as ADARB2, hnRNPA1, hnRNPA1B2, Pur-α, FUS, and TDP-43, (Almeida et al. 2013; Christopher J. Donnelly et al. 2013; Sareen et al. 2013).

An increased sensitivity to autophagic inhibitors was also detected and increased levels of p62 was observed, suggesting the impaired autophagy pathway and sensitivity to stress in patient derived cultures. *C9orf72* iPSC-derived MNs showed elevated sensitivity to glutamate-mediated cytotoxicity in a dose-dependent manner (Christopher J Donnelly et al. 2013) and reduced viability as indicated by increase in caspase-3 activation when exposed to autophagy pathway inhibitor chloroquine (Almeida et al. 2013). Additionally,

previous studies showed that ASOs against the G4C2 repeat-containing transcripts can reduce RNA foci formation, normalize disease-specific transcriptional changes and rescue the increased vulnerability to glutamate toxicity (Donnelly, Zhang, Pham, Heusler, et al., 2013; Sareen et al., 2013). Recent evidence has suggested a role of nucleocytoplasmic transport dysfunction in C9ALS/FTD pathology. Patient iPSC-derived MNs have contributed to the identification of nuclear RNA export defects (Freibaum et al. 2015) and nuclear import impairment caused by nucleocytoplasmic transport protein mislocalization (Zhang et al. 2015).

One of the limitations in studying the HRE mutation is the lack of patient isogenic lines from HRE carriers. iPSCs provide a renewable source of patient-specific cells that retain identical genetic information and give rise to all neural cell types. The rapid advancement of CRISPR/Cas9 system combined with iPSC technology have provided unprecedented opportunities to model ALS/FTD in vitro.

1.7.3. Correction of HRE mutation and creation of the isogenic controls for disease modelling

Recent advances in genetic engineering is the use of CRISPR gene editing technology in combination with iPSCs to generate iPSC isogenic lines by inserting mutations in a control iPSC line or correcting a mutation in disease cells by replacing it with the wild-type sequence in patient iPSCs to study the reversal of disease phenotypes. The generation of isogenic lines can eliminate the genetic heterogeneity and genomic variations between parental cells and the edited clones.

1.8. Gene Editing Techniques

Current gene editing techniques can be classified into two groups: protein- and nucleotide- directed targeting mechanisms (Esvelt & Wang, 2013). The first efficiently developed genome editing were protein-based, Zinc Finger Nucleases (ZFNs) and Transcription Activator-Like Effector Nucleases (TALENs). In contrast, CRISPR/Cas9 editing is a nucleotide-directed DNA editing technique that produce a double-strand break (DSB) in the genomic DNA sequence (Xu et al. 2014). DSBs can be repaired either by non- homologous end joining (NHEJ), which results in insertion or deletion of sequence around the DSB. This method can be used to create gene knockouts effectively. For precise correction of large deletions or insertions, a donor repair template carrying the intended modification can be used and transfected with sgRNAs and Cas9 constructs (Cong et al. 2013; Byrne et al. 2014). The repair occurs by homologous recombination to restore the wild-type sequence and protein expression and function.

1.8.1. Zinc-Finger Nucleases (ZFNs)

ZFNs are consisting of a fusion between the DNA-binding domain of a zinc-finger protein and the nuclease domain of the FokI restriction endonuclease. Two ZFN monomers can combine and form a heterodimer to cleave DNA and create a double-stranded break. ZFNs have been used for gene modifications on several cell types. ZFNs are less popular than the other site-specific nucleases (TALENs and CRISPR/Cas9), due to its high cost and complexity of applying the technique (Table 1.3). However, the efficiency of the system is higher than the other nucleases because each ZFN monomer is much smaller in size than other nucleases, which facilitates their transfection and packaging into viruses.

1.8.2. Transcription Activator-Like Effector Nucleases (TALENs)

TALENs are derived from the plant pathogen Xanthomonas bacteria. The term stands for Transcription Activator-Like Effector which describes a family of proteins that can bind to specific genes in the plant genome and regulate their expression. TALENs contain 33 to 35 repetitive amino acid residues in a central domain, with an almost identical amino acid sequence on each repeat. However, the amino acids at the 12^{th} and 13^{th} residues in the repetitive region have a function in DNA-binding specificity, known as the repeat-variable di-residues (RVDs), and determine the specificity of the TALE proteins. The structure of TALENs is very similar to ZFNs in that they are composed of heterodimers of a DNA binding domain and the nuclease of the FokI restriction endonuclease. Genome editing with TALENs has become more popular than ZFNs (Table 1.3). TALENS can bind to any DNA sequence in the genome, which provides an advantage for designing gene engineering strategies.

Although ZFNs and TALENS have been used in genome editing successfully, they have some technical drawbacks:

1. The high cost and difficulty of preparing custom proteins for the recognition of specific DNA sequences.
2. Optimization procedures of these techniques are time consuming compared to CRISPR system, where DNA cleavage is guided by a sgRNA sequence rather than an engineered protein.

1.8.3. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9

The CRISPR/Cas9 technology was originally identified in bacteria as an adaptive immune response to defend against bacteriophage infection. In 2012, Jennifer Doudna and Emmanuelle Charpentier published a study demonstrating the potential of using

CRISPR/Cas9 to cut genes efficiently (Jinek et al. 2012). CRISPR system has become the most popular gene editing technique due to its potential to precisely edit prokaryotic and eukaryotic cells (Lin et al. 2014; Upadhyay et al. 2013; Mali et al. 2013).

The CRISPR/Cas9 genome-editing technology consists of the Cas9 nuclease, a single gRNA (sgRNA), and a *trans*-activating crRNA (tracrRNA). Each gRNA contains a 20-nucleotide guide sequence that recruits the Cas9 to the 20-base-pair DNA target by Watson-Crick base pairing (Jinek et al. 2012) and the formation of DSB at the target site, which must be preceded by 5'-NGG PAM sequence (Cong et al. 2013; Mali et al. 2013; Jinek et al. 2012).

CRISPR/Cas9- mediated targeting in iPSCs is more efficient than other genome editing techniques (Ding et al. 2013), but the potential for off-target effects remains a major issue in using CRISPR system. To address this issue of off-target effects, researchers developed a modified Cas9 nuclease (Ran et al. 2013). The modified Cas9 has a "nickase" activity in which Cas9 cleaves only the strand non-complementery to the guide RNA. These nickases carry a catalytic amino acid substitution (D10A) in the conserved RuvC nuclease domain making them able to produce single-stranded breaks that can be repaired by the base excision repair pathway (BER) (Dianov & Hübscher 2013). Cas9-nickase creates DSBs by using a pair of sgRNAs targeting opposite strands of the target region to minimise off-target results (Ran et al., 2013). Another approach has been developed to decrease the off-target effects is shortening of the sgRNAs sequence to 17 or 19 nucleotides in length to reduce off-target activity (Fu et al. 2014). The main advantages of the CRISPR/Cas9 system over the other nucleases are listed in Table 1.3.

Table 1. 3 Comparison of the CRISPR-Cas9 system with ZFNs and TALENs (adapted from Heidenreich & Zhang, 2015).

	ZFNs	**TALENs**	**CRISPR-Cas9 system**
Target DNA recognition	Zinc fingers (ZFs)	Repeat variable diresidues (RVDs)	crRNA or sgRNA
DNA cleavage endonuclease	Fok I nuclease domain	FoK I nuclease domain	Cas9
Construction of endonuclease	3–4 ZFs domains	8–31 RVD repeats	sgRNA
Minimum number of DNA base be recognized	Triple	Single	Single
Size of recognized DNA sequences	9–18 bp	(8–31 bp) × 2	20 bp + NGG
Restriction of target site	/	Binding site should start with a T base	PAM sequences (NGG or NAG)
Cytotoxicity	High	Low	Low
Advantages (comparison of ZFNs and TALENs with CRISPR-Cas9)	High efficiency Lower rate of off-target cleavage than CRISPR-Cas9		Very high efficiency and rapid construction and easy delivery Multiplexing possible *in vitro* and *in vivo*
Disadvantages (comparison of ZFNs and TALENs with CRISPR-Cas9)	More difficult to assemble than CRISPR-Cas9 and remaining possibility of off-target cleavage		Off-target cleavage possibly more frequent than ZFNs and TALENs

Abbreviations: Cas, CRISPR-associated protein; CRISPR, clustered regularly interspaced short palindromic repeat; TALEN, transcription activator-like effector nuclease; ZF, zinc-finger; ZFN, ZF nuclease.

1.9. Gene editing in iPSCs

The combination of iPSCs and genome editing technologies represents a powerful approach for studying the pathogenesis of inherited diseases and for finding therapeutic strategies. CRISPR has become the most popular and powerful system for gene editing

of iPSCs, due to its ease of use and lower cost and complexity compared to other nucleases. CRISPR has been used in several studies for creating knockouts, deletions and functional mutations. Homologous recombination was successfully applied to correct the CFTR locus in cultured intestinal stem cells of patients with cystic fibrosis (Schwank et al. 2013), an inherited disorder primarily affecting the lungs and digestive system. The corrected gene showed normal expression and functionality.

The variability between individual iPSC lines in their differentiation potential and the differences in genetic background and phenotypic differences between iPSCs differentiated from patient or control may result in a failure to recapitulate the disease-related phenotypes. Therefore, the ability to generate the isogenic lines from disease-iPSC lines has decreased the possibility of these variations (Soldner et al. 2012). Differentiation of isogenic lines with their original disease-specific iPSC lines can be used to directly compare cellular pathogenic phenotypes related to the gene mutation (Salewski et al. 2013; Yusa et al. 2011)

Aims and Objectives

- The overarching goal of this thesis was to create isogenic cell lines using CRISPR/Cas9 from ALS/FTD patient iPSCs carrying the *C9orf72* HRE mutation.
- To determine if ALS disease-related phenotypes are dependent on the presence of the abnormal repeat expansions and to provide a clear evidence for the pathogenicity of these expansions.

Chapter 3 Aims:

- The main aim of this chapter was to generate and characterize iPSC lines from ALS patients carrying the G4C2 repeat expansion mutation. This has been achieved by transducing fibroblast cells with Sendai viruses carrying the four Yamanaka reprogramming factors (OSKM). Three cell lines were expanded and characterized for their pluripotency and karyotyping profile.
- To choose a suitable line for the gene editing experiment, two ALS/FTD iPSC lines were differentiated into MNs using an EBs-based protocol, to demonstrate that the isolated iPSCs can be differentiated efficiently into MNs before using them in gene editing experiment.

Chapter 4 Aims:

- This chapter aimed to correct the HRE mutation in ALS/FTD patient iPSC line (C9-02-02), using CRISPR/Cas9-mediated HDR system and a plasmid donor repair template carrying the wild-type size of the repeats. This chapter has involved the following steps:
  - Design and construction of CRISPR gRNAs to target upstream of the repeats region of the *C9orf72* gene and evaluation of their cleavage efficiency.

- Constructing the donor template from the same patient DNA by amplification of the WT-allele sequence to replace the repeat expansion in the mutant allele by HDR.
- iPSC nucleofection with gRNAs and donor template and puromycin selection of the drug resistant clones.
- Molecular screening of the isolated clones and expansion of the successfully targeted clones.

Chapter 5 and 6 Aims:

- To elucidate the pathogenicity of *C9orf72* gene mutation and to investigate whether the disease pathogenicity is directly correlated with the presence of the repeat expansions using iPSC-derived MNs.
- To study the potential of reversing the disease-related phenotypes after gene editing of the abnormal repeat expansions.
- To understand how HRE mutation contributes to disease progression. This could help in finding treatments that can ameliorate disease-related phenotypes.

Methods covered on this thesis:

- ✓ Induced pluripotent stem cell generation and characterization. This involved, primary isolation of iPSC from skin fibroblasts, isolation of the individual clones, expansion on feeder-free layer and characterization.
- ✓ Gene editing using CRISPR/Cas9-mediated homologous recombination, and characterization and expansion of the successfully generated isogenic lines.
- ✓ Motorneuron differentiation and characterization using different protocols.
- ✓ Phenotyping and and confocal imaging of the neuronal cultures.
- ✓ Basic statistical analysis using Prism software

2. CHAPTER TWO

MATERIALS AND METHODS

2.1. Reagents and recipes

Tris-Acetic Acid-EDTA (TAE) buffer (50X)

To make up 50X TAE buffer, 242 g Tris-base (Sigma Aldrich), 57.1 ml of 100% acetic acid (Merck) and 100 ml of 0.5M sodium EDTA (pH 8.0) were mixed and made up to 1L with ddH_2O. For preparation of 1L of 1X TAE, 20 ml of the 50X TAE solution was mixed with 980 ml of ddH_2O and mixed thoroughly.

Phosphate buffered saline PBS (1X)

One PBS tablet (Sigma Aldrich) was dissolved in 100 ml ddH_2O. The solution was autoclaved and stored at room temperature (RT) for routine work use.

Human Embryonic stem cell (hESC) medium

The media was made up of the following ingredients: Knock-out Dulbecco's modified Eagle's medium (DMEM) (Gibco), 10% knock-out-serum replacement (Gibco), 2mM Glutamax-I (Gibco) 100 U/ mL, Penicillin/Streptomycin (P/S) (Gibco), 1% Nonessential amino acids (NEAA) (Gibco), 0.5 mM 2-Mercaptoethanol (Gibco), and 10 ng/mL Basic fibroblast growth factor (bFGF) (R&D).

Advanced DMEM complete medium

Advanced Dulbecco's Modified Medium (Gibco), supplemented with 10% FCS (Gibco), and 1% of 100U/ml, P/S.

2X iPSC Freezing medium (2X FM)

The following ingredients were mixed together to prepare 2X FM: 6.6ml Dimethyl sulfoxide (DMSO) (Sigma Aldrich), 6.6ml hES medium, 20ml ES-Fetal Bovine Serum (FBS) (Hyclone) in a 50ml tube. The mixture was stored in the dark at 4°C.

Matrigel coating

Matrigel (BD) was thawed at 4°C and diluted 1:50 in DMEM medium. Each well of 6-well plate was coated with 1mL of Matrigel solution and incubated for 1h at 37°C or overnight at 4°C. Matrigel was aspirated immediately before seeding the cells.

Flow cytometry buffer

Consisting of 1X PBS, 10ug/ml human IgG (Sigma Aldrich), 1% fetal calf serum (FCS) (Hyclone), and 0.01% sodium azide (Sigma Aldrich).

2.2. Cell culture methods

All the antibodies used in this thesis are listed in Appendix I.

2.2.1. Subculturing of dermal fibroblast cells

Fibroblast cells were maintained at 37°C with 5% CO2 in humidified conditions and maintained in Advanced DMEM complete medium. Fibroblast cells were passaged and expanded to get enough number of cells for freezing and reprogramming. Cell passaging was performed at approximately 80% confluency. Cells were washed with 1X PBS (Gibco), and then detached by incubation with 1-2ml of TrypLE solution (Invitrogen) at 37°C for ~5 min. After that, cells were detached and re-suspended in 8ml of 1XPBS then

transferred to a 15ml falcon tube and centrifuged at 450rpm for 5 min at RT. Then, the supernatant was discarded and the cell pellet was re-suspended in 10ml fresh complete Advanced DMEM growth medium and seeded at a ratio of 1:3 into a new three T75 culture flasks (Corning) in 15ml growth medium each. Cells were examined the next day under the microscope for their morphology, and maintained at 37°C until they reached 80-90% confluency. After that, cells were trypsinized and stored in Liquid Nitrogen (LN) or used in iPSC reprogramming experiment. The supernatant of cell culture medium for the last passage before cell freezing was collected and used for Mycoplasma testing and cell pellets were made for DNA extraction and karyotype analysis, to compare them with their derivative iPSC lines karyotypes.

2.2.2. Preparation of MEFs and Conditioned Medium

Primary mouse embryonic fibroblasts (MEFs) were expanded from MEFs derived from CF-1 mice cell line (Merck) in Advanced DMEM complete medium and passaged for three times to achieve enough quantity for final splitting into twelve T175 cell culture flasks. Cultured cells were mitotically inactivated with 10 µg/mL Mitomycin C for 3h at 37°C. Then inactivated MEF cells were washed with PBS and detached with TrypLE, counted and frozen in 2X FM in a total of $2x10^6$ cells per vial.

Conditioned media (CM) was prepared from one thawed vial of the inactivated MEFs in T75 flask and maintained in 20ml Advanced DMEM complete medium for 24h. Then, medium was aspirated and rinsed in 15ml PBS and then hES media was added. The same procedure of adding and collecting hES medium was repeated for 7 days. After final medium collection, CM was filtered, aliquoted into 50ml tubes and stored at -20°C until further use. Before using for feeding iPSCs, CM was thawed and bFGF was added in a

final concentration of 15ng/ml and stored at 4°C for 1-2 weeks.

2.2.3. Freezing and Thawing of hiPSCs

For freezing of iPSC on MEF layer, colonies were detached using TrypLE and centrifuged at 450g for 5 min. Then, every $1x10^6$ cells were resuspended in 1ml of 1X freezing media and 1ml aliquots were prepared in cryotubes and stored at -80°C for overnight in a 2-propanol filled freezing container. And then, banked on the next day in LN for long-term storage. For freezing of iPSC on Matrigel, cells were washed with PBS and incubated with 0.5mM EDTA for 10 min at 37°C. After that, the EDTA solution was aspirated and iPSCs were detached and resuspended in 1X FM.

For hiPSCs thawing, MEF–feeder layer or matrigel coated plates were prepared in advance. Frozen aliquots were thawed at 37°C in a waterbath, then cells transferred into 10ml 1X PBS in a 15ml falcon tube and centrifuged at 450g for 5 min. After that, cells were re-suspended in the appropriate media either hES or mTeSR1 and 10μM Y-27632, Rock inhibitor, was added directly onto the cell culture media before seeding cells.

2.2.4. Immunofluorescence

For immunostaining experiments, cells were seeded on 12mm autoclaved glass coverslips coated with either Matrigel for iPSC culture or Geltrex for iPSC-derived MN cultures. At the appropriate time-point, cells were washed with PBS and fixed with 4% paraformaldehyde in PBS at RT for 20 min. Then, cells were washed 3 times in PBS and blocked with 10% Normal Goat serum or Donkey serum as appropriate in 0.1% TritonX-100 in 1X PBS (0.1%TX-100 PBS) for 1h at RT. Primary and secondary antibodies (Appendix I) were also diluted in 0.1%TX-100 PBS. All primary antibody incubations were performed overnight at 4°C. Then cells were washed 3X in 1X PBS for 10 min each.

After that, the appropriate secondary antibody (Alexa-Fluor 488 or Alexa-Fluor 568 conjugated antibodies, Life Technologies) was added in 1:500 dilution and incubated for 1hr at RT and then samples were washed for 3X in 1X PBS for 10 min each. Then 4',6-Diamidino-2-Phenylindole (DAPI) nuclear stain (Life Technologies) was added to PBS (1:1000) and incubated for 10 min at RT with shaking. Finally, coverslips were placed face-down and mounted on glass slides using few drops of DACO mounting medium (life technologies), dried for overnight and protected from light. Images were captured at 60X oil lens magnification using Ziess confocal microscope (LSM). For each assay, images were taken for at least ten representative fields and percent positive was analysed using imageJ analysis software.

2.3. Molecular biology methods

All primers used in this project are listed in Appendix III

2.3.1. Polymerase Chain Reaction (PCR)

A proof-reading Phusion High Fidelity Polymerase (Thermo Scientific) was used in all PCR assays performed in this study. Annealing temperatures for Phusion polymerase were calculated for all primers used in this project using the following website: https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html. Total PCR volume used in most assays was 25µl per reaction. RCR products were visualised by gel electrophoresis using 1-2% agarose gel (w/v) (depending on the product size).

2.3.2. Agarose gel electrophoresis

DNA fragments from PCR products or plasmid extraction were separated by 1-2% agarose gel electrophoresis containing 0.5µg/ml ethidium bromide in 1X TAE buffer. Before loading on agarose gel, PCR products were mixed with 5X DNA loading buffer (Bioline) and run at ~120V for ~1h until the required separation was achieved. Bioline hyperladders 1kb or 100bp were used as appropriate to identify band sizes. Bands were visualised using UV light and images were captured accordingly. Bands were excised if required and DNA purification was performed using QIAquick Gel Extraction kit (Qiagen).

2.3.3. Cloning and transformation

All cloning experiments were performed using 2X Rapid Ligase enzyme kit (Promega). Plasmid vectors were digested with the appropriate restriction enzymes and digested fragments were purified by gel electrophoresis. DNA concentration was measured using NanoDrop spectrophotometer (ThermoFisher). Insert to vector ratio was made as ~3:1 in up to 10µl ligation mix. Then, 1µl of ligase was added per 10µl reaction volume, followed by addition of 2X Rapid Ligation Buffer (NEB). Reaction components were mixed and centrifuged briefly. Ligations were performed for 1h at RT or overnight at 17°C, followed by transformation into either Alpha-Select Gold Efficiency (Bioline) or home-made Stabl-3 competent E. coli. Twenty microliters of competent cells were thawed on ice before the addition of 3µl of ligation product and then incubated for ~30 min before heat-shock at 42°C for 30s. After that, transformed cells were incubated on ice for 2 min and then transferred to 600µl super optimal broth (SOC) medium and incubated at 37°C for 1h with agitation. Then, 100µl of cells were seeded on LB agar plates supplemented with

100mg/ml Ampicillin and kept to grow for overnight at 37°C.

2.3.4. Minipreparation

On the next day, some single colonies were picked up and suspended in 3-5 ml LB broth supplemented with 100 mg/ml Ampicillin (LB/Amp) for overnight at 37°C with shaking. Then, 2ml of the cultured broth were centrifuged for 5 min at 13000 rpm. The supernatant was discarded and the pellet was used for Plasmid DNA extraction using Qiagen DNA minikit according to the manufacturer instructions. Briefly, the pellet was re-suspended in 250µl of P1 buffer containing RNase and then lysed with 250µl of P2 buffer, after which they were neutralized with 350µl of N3 buffer to terminate lysis. The mixture was centrifuged at 13,000rpm for 10 min at RT. The supernatant was transferred to QIAprep spin column and centrifuged at 13,000rpm for 1 min, then 750µl of PE buffer was added to the column and centrifuged. DNA was then eluted from the column using 50µl of NFW and centrifuged at 13,000rpm for 3 min. Finally, DNA concentration was measured using Nanodrop spectrophotometer and stored at 4°C for further use.

2.3.5. Midipreparation

One ml of miniprep broth was inoculated into 50 ml of LB/Amp and grew overnight at 37°C with shaking. Plasmid DNA was extracted using PureLink® HiPure Plasmid Filter Purification Kit (Invitrogen) as per manufacturer's instructions. DNA concentration was quantified using a Nanodrop spectrophotometer and diluted to 1 µg/µl stocks.

2.3.6. Direct DNA-Sequencing

DNA fragments were purified by PCR purification or gel purification kits if needed (Qiagen) per manufacturer's instructions. Approximately 100ng of purified plasmid DNA samples (minipreps or midipreps) with the associated 10µM primer were sent to the

Source Bioscience company for Sanger DNA sequencing, to assess the correct cloning product for no deletions, insertions or point mutations resulted from the polymerase activity. Samples were sequenced in forward and/or reverse directions as listed in each specific protocol.

2.3.7. QIAamp DNA Mini Kit

Cells grown in a 6-well plate were washed with 1X PBS then harvested using cell scraper, transferred to a 1.5ml microcentrifuge and centrifuged at 450g for 5 min. The supernatant was then discarded, and the pellet was resuspended in 200µl of 1x PBS. After that, 20µl of Proteinase K was added and followed by the addition of 200µl of AL buffer, mixed gently by vortexing for 15 sec and then incubated at 56°C for 10 min. Then 200µl of 100% ethanol was added and mixed by vortexing for 15 sec. The microcentrifuge tube was centrifuged briefly and the mixture was applied to a QIAamp Spin Column and centrifuged again at 8000g for 1 min, then the column was placed in a clean 2ml collection tube and 500µl of AW1 buffer was added and centrifuged. The column was washed with 500µl of AW2 Buffer and centrifuged at high speed for 3 min. The column was then transferred into a clean 1.5ml microcentrifuge tube and the DNA was eluted by adding 150µl of NFW and incubated for 5 min at RT. Finally, the DNA was collected by centrifugation at 8000g for 3 min and stored at -20°C until use.

2.3.8. RNA isolation and cDNA synthesis

RNA was extracted directly from cells using RNeasy Minikit QIAGEN), according to the provided manual instructions. RNA was eluted in 50µl NFW and stored at -80°C until further use. RNA quantity and quality was assessed using a Nanodrop 2000

spectrophotometer by analyzing 260/230 and 260/280 values. RNA was reverse-transcribed into complementary DNA (cDNA) using Superscript III First-strand synthesis kit (Invitrogen). cDNA samples were diluted 1:10 in NFW and 2µl of each sample were used in 96-well plate. Each cDNA sample was run in triplicate.

2.3.9. PCR purification and cleanup

PCR products were purified using PCR Cleanup KIT (Qiagen), according to manufacturer's instructions. Briefly, five volumes of PB buffer were added to 1 volume of PCR reaction and mixed by pipetting. A QIAquick column was placed in a 2ml collection tube and sample was applied to the column then centrifuged for 30sec at 13,000rpm and the flow-through was discarded. Then, 0.75 ml of PE buffer was added to the column to wash the DNA and then centrifuged. DNA was eluted from the column by adding 50µl NFW and incubating at RT for 5 min then centrifuged at 8000g and stored at -20°C.

2.4. Isolation of skin fibroblasts and generation of iPSCs

Skin biopsies were taken from two ALS patients (C9-01 and C9-04) and one ALS/FTD patient (C9-02) carrying the HRE mutation in the C9orf72 gene, under ethical approval described in (Dafinca et al. 2016), which involved a signed informed consent, with approval from the South East Wales Research Ethics Committee (REC ref no. 12/WA/0186) and the South Central Berkshire Research Ethics Committee, U.K. (REC 10/H0505/71).

For fibroblast expansion, each biopsy sample was cut into three equal pieces of approximately 3mm in size. Each piece was placed on one well of 6-well culture plates and dried for approximately 5 min to facilitate their adherence. Fibroblast culture medium (Advanced DMEM supplemented with 10% (v/v) FCS and 1% (v/v)

Penicillin/Streptomycin) was then added and changed every other day without disturbance of the adherent biopsies. When cells became densely growing around the tissue pieces, cells were washed with PBS and passaged using TrypLE into a new well of a 6-well plate. Fibroblast cultures were subsequently passaged at a ratio of 1:3 when cells reached 80-90% confluency as described earlier in this chapter in Section 2.3.1.

2.5. Generation of clonal iPSC Lines

Reprogramming was started using low passage number (p3-5) fibroblasts, derived from C9-01, C9-02 and C9-04 motorneuron disease patients. Briefly, fibroblasts were seeded on 12-well plate ($1x10^5$ cells per well), and transduced with commercially available Cytotune1 Sendai viruses-based reprogramming system. This transduction system contains the Yamanaka reprogramming genes *Klf4, Oct4, Sox2*, and *c-Myc* packaged into individual viruses (Life Technologies). Cell Transduction was performed as per manufacturer's protocol. Briefly, CytoTune1 Sendai virus (MOI=3∞) was diluted in advanced DMEM with 10%FBS (without P/S) and added to cells when they were 70-80% confluent.

On the next day, the medium was changed and replaced with fresh medium. Seven days after transduction, cells were trypsinised and seeded on 6-well plates treated with 0.1% (w/v) Gelatin and coated with mitotically inactivated MEFs in a seeding density of 800 or 1600 of transduced cells per well. After 24h, medium was changed and replaced with hES medium. For ten days, medium was changed every other day, then CM was used and supplemented with 10 ng/ml bFGF and 1% P/S. For generation of the clonal iPSCs lines, at least six single iPSC colonies showing a typical ES-like morphology were dissected after three weeks of reprogramming, using a sterile needle on syringe, and picked up individually into new wells of MEF layer on 6-well plate. Clonal iPSCs were cultured

under feeder conditions, maintained in hES medium at 37°C in 5% CO2 incubator and manually dissected every 5-7 days. Each day, half of the culture media was removed and a new fresh hES media was added. Cells were passaged with TrypLE at passage 6 after being confluent, mycoplasma tested, and banked. Three iPSC clones were kept growing on MEFs for several passages (p10-12) to allow them to expand and to have enough number for freezing and expansion on feeder-free conditions. The generated iPSC lines were identified as the following: C9-01-01, C9-01-06 and C9-01-07 derived from C9-01 fibroblasts; C9-02-02, C9-02-03 and C9-02-10 derived from C9-02 fibroblasts, and finally C9-04-01, C9-04-09 and C9-04-12 derived from C9-02 fibroblasts.

2.5.1. Culture of hiPSCs under feeder-free Conditions

After p10-12 on MEFs and manual dissection of iPSC colonies, iPSC lines were adapted to feeder-free conditions onto Matrigel-coated 6-well plates in mTeSR (StemCell Technologies) supplemented with 10µM Y-27632 on the day of passaging. For coating, 1ml per well (6 well plates) of diluted Matrigel was incubated for at least 1h at RT. Medium was changed completely every 24h and lines were passaged before reaching 90% confluency using 0.5mM EDTA and allowed to expand and then stored in LN and banked. Representative samples of the frozen stocks were taken for further characterisation of the isolated iPSCs and mycoplasma testing. iPSCs from healthy controls used on this study (SFC-840-03-03 (WT-01) and SFC-856-03-04 (WT-02)) were kindly provided by Dr Sally Cowley, James Martin stem cell facility. After each single cell dissociation, 10µM of Y-27632 was supplemented onto culture medium on the day of passaging.

2.6. iPSCs Characterisation

2.6.1. Immunostaining of Pluripotency markers

iPSCs were cultured on 12mm coverslips in 24-well plates for few days until they produced compact colonies, then fixed with 4% paraformaldehyde for 15min at RT, and permeabilised with 0.2% TX-100/PBS, then blocked with 10% (v/v) donkey serum for 1h to prevent non-specific binding. After that, cells were incubated with the primary antibodies against NANOG, SSEA-4, TRA-1-60 and Oct3/4 for 1h (Appendix I). After three washes with PBS, cells were incubated with the relevant Alexa Fluor-conjugated secondary antibodies for 1h at RT, counterstained with DAPI and examined by fluorescence microscopy. The corresponding isotype antibodies were used as a negative control staining to check for the non-specific binding.

2.6.2. Flow cytometry for TRA-1-60 and NANOG

iPSCs from one well of a confluent six-well plate of each iPSC lines cultured on Matrigel was treated with TrypLE for 5 min. Then cells were collected in PBS and centrifuged at 400g for 5 min. After dissociation, cells were fixed in 2%PFA, then permeabilised in methanol for 30 min, centrifuged and washed with FACS buffer. Then, 0.5–1 x 10^6 cells were suspended in FACS buffer and distributed in 96 well plates (50µl per well). Cell suspensions were then treated with antibodies against both TRA-1-60 FITC (1:100) and NANOG PE (1:150) and relevant isotype controls were used at the same concentration and incubated for 60 min at RT in the dark. Finally, cells were washed three times and 10,000 ‘live’–gated events were analysed using a Becton Dickinson FACSCalibur instrument and CellQuest Pro software (BD Biosciences, San Diego). Analysis and

plotting were performed using FlowJo software, gated on forward scatter-side scatter (FSC-SSC) dot plots.

2.6.3. Karyotyping

Genome integrity was assessed using genomic DNA extracted from all the iPSC lines and their corresponding original fibroblasts using DNAeasy Blood minikit (Qiagen), followed by an Illumina Human OmniExpress 24 array (~700,000 markers) (Wellcome Trust Centre for Human Genomics) and analysed using KaryoStudio software (Illumina). SNP profiles for the iPSC lines were compared to the original dermal fibroblasts. SNP datasets have been deposited in Gene Expression Omnibus (GEO) under the accession number GSE64582.

2.6.4. Sendai virus clearance RT-PCR

RT–PCR was performed to check for the clearance of Sendai virus genome in the isolated iPSC clones using specific primers listed in Appendix III. RNA was extracted using RNeasy minikit (Qiagen) and cDNA was synthesised from the purified RNA using SuperscriptIII Reverse-Transcription kit with oligo(dT) primers (Invitrogen). RT-PCR was performed using 10μM of each forward and reverse primers for Sev genome backbone, and forward primers for OSKM factors were used with a common reverse primer in the Sendai virus genome. Then, 100ng cDNA was added to the PCR reaction in a total volume of 25μl. RNA from fibroblasts and from fibroblasts infected with Sendai reprogramming viruses were used as negative and positive controls, respectively. The following PCR programme was used: 95 °C for 2 min followed by 30 cycles of 94 °C for 15 sec, 56 °C for 15 sec and 72 °C for 40 sec. β-actin gene was used as an internal control (Eurogentec) and products were visualised on 1.5% agarose gel stained with ethidium bromide, and 100bp ladder as a size marker (Bioline).

2.6.5. Testing for Mycoplasma contamination

The frozen cells were passaged at least once after thawing and then left without media change for 2 days. One ml of the culture medium was aspirated from the well and centrifuged to remove cell debris. Supernatant was isolated and Mycoplasma Test Kit (Mycoalert, Lonza) was used following the manufacturer´s protocol.

2.7. Motorneuron differentiation and characterization

Embryoid bodies (EBs) were generated and differentiated into motorneurons (MNs) from two ALS/FTD patient iPSC lines named as C9-02-02 and C9-02-03 before further use in gene editing experiment (Hu & Zhang 2009). WT-01 was used as a healthy control.

2.7.1. EBs differentiation protocol

To produce EBs, two undifferentiated confluent iPSC wells of 6-well plate were used from each line. Cells were dissociated with TrypLE, centrifuged and transferred into 100mm low attachment plates and maintained in mTeSR media with 10μM Y-26732 for 24h. On the following day, EBs were collected and maintained in neural induction (NI) serum-free medium (DMEM/F12 with L-glutamine, NEAA, 1% P/S, heparin (2 μg/ml), N2 supplement (Invitrogen)) for 4 days, and then plated on 6-well plates covered with Geltrex and maintained on NI media for additional 3 days, after which 0.1μM retinoic acid (RA) was added (Sigma). The differentiated cells were evaluated under the microscope for the formation of neural rosettes. After that, rosettes were lifted and allowed to grow in suspension in the presence of B27, 10μM Y-26732, 0.1μM RA, 100 ng/mL Sonic Agonist (SAG) and Puromorphamine for ventralisation to produce neurospheres. After 5 days, larger neurospheres were collected and allowed to grow for 10 days in suspension. The spheres were then plated on Geltrex coated coverslips in the

above neural differentiation medium supplemented with brain derived neurotrophic factor (BDNF 10 ng/mL), glial-derived neurotrophic factor (GDNF, 10 ng/mL), insulin-like growth factor (IGF- 1, 10 ng/mL), cyclic AMP (cAMP), and 0.5μM ascorbic acid (Sigma). After 14 days in neuronal differentiation medium, the cells were assessed for MN markers expression.

2.7.2. Immunofluorescence staining of MN markers

Immunostaining was performed as described in section 2.2.4. Cells were incubated with primary antibodies for either Tuj-1 and Islet1 or HB9 for overnight at 4°C. And on the next day, the appropriate secondary antibodies were used (Alexa Fluor 488 and Alexa Fluor 568). Nuclei were counterstained with DAPI. The structures of MNs were visualised using 60x oil objective Confocal Microscope Zeiss System. The quantification of the immunofluorescence staining was based on quantifying the cells positively stained with Tuj-1 and HB9 or Islet-1, using well-established antibodies that have been used successfully for immunostaining of iPSC-derived MNs in several MN differentiation protocols and primary MN cultures.

2.8. Gene editing of G4C2 mutation using CRISPR/Cas9 system.

Plasmids

pX335 was kindly gifted from Zhang lab (Addgene, Plasmid #42335), pENTR-EF1α-Puro/Tk was kindly provided by Dr Kenny Moore (Stem Cell Institute, Sir William Dunn School of Pathology), pGEMT easy vector (Promega), GFP^{max} plasmid (Lonza) and Cre-mRNA (TriLink Biotechnologies).

2.8.1. gRNAs Construction and CRISPR design

The CRISPR-Cas9 vectors used in this study were based on the use of Cas9 pX335 nickase (D10A), which is a dual Cas9n and sgRNA-expressing plasmid, designed as described previously by (Ran et al. 2013). Two gRNA sequences were selected (gRNA1 and gRNA2) using the CRISPR design tool (http://crispr.mit.edu) with limited predicted off-target effects to target upstream the G4C2 repeats region. PCR primers and oligos for gRNA1 and gRNA2 construction are listed in Appendix II. For construction of gRNAs, px335 vector was digested with BbsI and then the linearized vector was gel-purified and oligos for each sgRNA were annealed, phosphorylated, and ligated to the linearized vector. Ligated DNA was transformed into Stabl3 E.Coli competent cells. DNA miniprep was prepared for each gRNA, and sequenced using RF27 primer (5'-CTCGACCATGGTAATAGCG-3') to check for the presence of a precise gRNA sequence. Finally, midiprep was prepared for each gRNA and plasmid DNA was extracted as described previously.

2.8.2. Cell Culture and Transfection of HEK293T cells for cleavage activity testing

HEK293T cells were transfected with either gRNA1 or gRNA2 plasmid vectors or with both together simultaneously using Lipofectamine 2000 to evaluate the cleavage efficiency of the designed constructs. Briefly, 24h prior transfection, HEK293T were seeded at $5x10^5$ cells per well and incubated at 37°C with 5% CO2. Cells (80%–90% confluent) were transfected using Lipofectamine 2000 (Life Technologies) per the manufacturer's protocol. A total of 1.5µg plasmid DNA was transfected either with both gRNAs or with individual gRNAs and the untransfected cells were used as a negative

control. Then 72h post-transfection, cells were collected and DNA was extracted. The GFP-expressing plasmid (pmaxGFP, Lonza) was used as a positive control to monitor the efficiency of the transfection, by transfecting one-well of HEK293T cells with GFP only. GFP expression was examined after 48h using inverted fluorescence microscope.

2.8.3. T7EI and Cel1 (SURVEYOR) cleavage assays

Genomic DNA samples from untreated and treated iPSCs were extracted using DNeasy Blood & Tissue kit. The genomic region flanking the CRISPR cleavage site was amplified by PCR using specific primers to amplify 900 bp surrounding the genomic cleavage site (Appendix III) and the resulting products were purified using PCR Cleanup kit (Qiagen), following the manufacturer's protocol. Cleavage assays were performed using 400ng of PCR products that were denatured by heating at 95°C then reannealed to form heteroduplex DNA using a thermocycler. Then, the reannealed PCR products were digested with either T7 endonuclease 1 (T7E1, New England Biolabs) or SURVEYOR (Cel1) nuclease. For T7E1 assay, digestion was performed using 10 unites of T7E1 enzyme incubated for 15 min at 37°C. The T7EI reaction was stopped by adding 2µl of 0.25M EDTA solution. For CelI assay, the same protocol was applied with an incubation of the reannealed PCR products with 1µl SURVEYOR enzyme (Transgenomic, SURVEYOR® Mutation Detection Kit), 1µl Enhancer and 2µl $MgCl_2$ and then samples were incubated for 1hr at 42°C. The digestion was stopped by adding 2µl stop solution. The digested products from both assays were analysed on 3% agarose gel electrophoresis and stained with Ethidium Bromide to check for heteroduplexes formation.

2.8.4. Design of repair donor template and Site-directed mutagenesis (SDM)

The donor template was designed to span the target region with two long homology arms of approximately 2kb on each side. The left and right homology arms of the donor template were amplified from fibroblast-derived genomic DNA of the same patient (C9-02) wild-type allele, which has shown two G4C2 repeats after sequencing. The amplification was performed using 2 units of Phusion polymerase with 10X buffer, 10µM forward primer, 10µM reverse primer, 10mM dNTP mix and 50ng DNA. PCR cycles were as follows: initial denaturation at 98^{o}C for 2 min, 35 cycles of 98°C for 15 sec, 67°C for 20 sec, 72°C for 3 min, and finally 5 min of final extension at 72°C. The PCR product was inserted into the pGEMT easy vector and ligated using 2X Rapid T4 DNA Ligase as described earlier. Then, 3µl of the ligated PCR products were transformed into Alpha-Select Gold Efficiency E.Coli (Bioline). These cells contain a lacZ marker that provides α-complementation of the ß-galactosidase gene for blue/white color screening. After incubation for 20 min at RT, cells were heat-shocked at 42°C for 30 sec and then processed using the same way described in section 2.3.3., except that they were plated on LB/AMP agar provided with 20 mg/ml X-Gall and 10 µl IPTG for white/blue colonies selection, then miniprep was prepared and DNA was sequenced to check for the presence of the correct sequence.

After confirming the precise cloning and ligation of the donor template, back to back site-directed mutagenesis (SDM) technique was applied to introduce two NsiI sites before and after the repeats, to serve as diagnostic sites when analysing the targeted clones. One NsiI site was inserted in the PAM region of gRNA2 and the other one inserted after the repeat region, using the same PCR reaction conditions and ingredient concentrations listed

above in this section (primers in Appendix III). After SDM, 1µl of DpnI restriction enzyme was added to the reaction and incubated for 30 min at 37^{o}C to destroy the original plasmid template. A loxP-flanked EF1α-Puro-Tk cassette from pENTR plasmid was inserted between the 2 homology arms, approximately 140bp upstream of the repeats region. The cassette was amplified by PCR then ligated to donor template by Gibson assembly mastermix (NEB) using specific primers listed in Appendix III, to assemble the two plasmids together per the manufacturer's instructions. The reaction was incubated at 50°C for 15 min, followed by bacterial transformation into Alpha-Select Gold Efficiency E.Coli as described above. Then miniprep and midiprep samples were prepared and donor template was sequenced to check for any mutations occurred during the amplification process.

2.8.5. CRISPR-Mediated genome targeting and positive selection of iPSCs

On day 0 before nucleofection, iPSCs were pre-treated with 10µM Y-27632 for at least 1h prior to nucelofection experiment. This step is important to prevent apoptosis of hiPSCs in single cell suspension. Cells were dissociated into single cells by incubating them with TrypLE for 5 min at 37°C, then cells were mixed with 1X PBS followed by gentle pipetting. Two million iPSCs were nucleofected with 2.5µg of each gRNA (gRNA1 and gRNA2) and 4µg of the donor plasmid by Neon Transfection system (Life Technologies) and plated onto 10cm plates coated with drug-resistant (DR4 MEFs) feeder layer and maintained in hES media supplemented with 10 µM Y-27632 with no P/S added. On the next day, media was changed and the regular hES with P/S was used. Three days after the nucleofection, puromycin selection was started using 0.35 µg/mL of puromycin until the individual colonies were observed. GFP puromycin-free plasmid was

used as a control of cell death after puromycin treatment, to ensure efficient killing of the negative clones. One hundred puromycin resistant colonies were picked individually under the dissecting microscope into 24-well plates coated with DR4 MEFs. Puromycin selection was continued after cell passaging for at least three passages. Confluent cells were harvested using TrypLE into two wells of 24-well plates, one for cell expansion and freezing and the other one for genomic DNA extraction and molecular screening.

2.9. Molecular screening and genotyping of gene edited clones

2.9.1. Repeat-Primed PCR (RP-PCR)

Two RP-PCRs were performed, RP-PCR1 to assess the presence of the repeats on the 5' direction, and RP-PCR2 to assess the presence of the repeats on the 3' direction. Primer sequences used in both RP-PCRs are listed in Appendix III. RP-PCRs were performed as previously described with a brief modification in the primer sequences (Renton et al. 2011; DeJesus-Hernandez et al. 2011). PCR mix was prepared in a reaction volume of 15µl, containing 100ng of genomic DNA, 8µl Extensor mastermix (Thermo Scientific), 6µl of 5M Betaine, 20µM of both forward and repeat specific primer and 2µM of reverse primer. The reactions were subjected to touchdown PCR programme consisting of 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec, then 1µl of each PCR product was mixed with 8µl of Formamide, and 0.5µl GeneScan size standard (Applied Biosystem).

2.9.2. In-Out PCRs and direct sequencing of the intended modification

Total DNA was extracted using DNeasy Blood & Tissue kit (Qiagen) per the manufacturer's instructions. The clones were screened first by PCR amplification inside

EF1α-Puro-TK cassette and outside either 5' or 3' homology arms (5'-In-out, 3'In-out PCRs). Sequences and specifications of the primers are shown in Appendix III. PCR cycling conditions were performed as follows: 2 min at 98°C, 35 cycles of 15 sec at 98°C, 30 sec at 67°C and 2 mins at 72°C and one final step for 5 min at 72°C. PCR products were digested with NsiI (3h with 10 units of NsiI enzyme (NEB)). Then, PCR fragments were purified using QIAquick PCR Purification kit per the manufacturer's instructions and sent to DNA sequencing to confirm site-specific integration and the successful insertion of Nsi1 sites.

2.9.3. Digital droplet PCR for detection of random integration (RI) of Puro/Tk selection cassette and Cas9 cassette

Digital droplet PCR (ddPCR) was used to measure the copy number of puromycin cassette in the edited DNA samples and compare it to endogenous Lrrk2 housekeeping gene using EvaGreen PCR mastermix (BioRad) and per the manufacturer's instructions. Primers were designed with PRIMERX software (Premier Biosoft international, Palo Alto, CA) to amplify a segment of puromycin gene or Cas9 gene. PCR conditions were 30 sec at 95°C for initial denaturation, followed by 40 cycles of 10 sec at 95°C denaturation, 15 sec at 60°C annealing, and 30 sec at 72°C extension. The ddPCR was performed in 20µl reaction mixture and 70µl of ddPCR droplet generation oil (Bio-Rad) per well. Reaction droplets were generated using a QX100 Droplet generator (Bio-Rad) and a droplet generator DG8 cartridge (Bio-Rad). Forty microliters of the generated droplet reaction were transferred to 96-well PCR plates and sealed with heat-sealing foil sheets. The fluorescence of each droplet was measured using the QX100 droplet reader (Bio-Rad). Data were analysed using QuantaSoft software (Bio-Rad).

2.9.4. Removal of Puro/Tk cassette using Cre-Lox P system

The Puromycin cassette inserted in corrected lines was flanked by two loxP sites to facilitate its removal by Cre-mediated excision system to generate a cassette-free iPSC lines. To excise the selection cassette, iPSCs from one of the heterozygous targeted clones (Ed-02) with HDR only in the mutant allele was maintained on Matrigel-coated wells of a 6-well- plate in mTeSR1 medium. Similar nucleofection experiment was applied (as described in Section 2.3.8) using *Cre* mRNA. One hour before passaging, iPSCs were incubated with 10μM Y-26732 containing mTeSR1 medium. Cells were then harvested and dissociated into single cell suspension. Then, $2x10^6$ iPSCs were nucleofected with 0.5μg *Cre* mRNA and seeded on 10cm plates pre-coated with MEFs and maintained in hES supplemented with 10μM Y-26732 and no antibiotics were added. After few days of nucleofection, cells were treated with 3 μg /ml Ganciclovir added directly to the feeding media. After 7-10 days, some of the growing colonies were picked up and transferred into 96-well plate coated with MEFs then split into two new plates, one used for freezing and one for genotyping and PCR analysis of Puro/Tk cassette-free clones.

2.9.5. Evaluation of cassette-free clones by PCR and sequencing

Multiple edited isogenic lines were generated from Ed-02 heterozygous clone after excision of the selection marker. Two edited clones were used in phenotyping experiments after excision of the cassette Ed-02-01 and Ed-02-02. DNA was extracted from iPSCs of the Cre-treated and untreated edited clones and from the parental cells (C9-02-02). PCR amplification was performed using Puro-F and Puro-R primers to amplify inside the Puro/Tk cassette and outside the 5'-homology arm. Another PCR was carried out using Exc-F and Exc-R primers to amplify upstream and downstream the repeats,

generating a fragment of 75bp and 109bp for the wild-type and the targeted allele, respectively. The difference in size is related to the presence of a *loxP* site remained after excision of the cassette using *Cre-loxP* system.

2.9.6. Southern Blot

Southern blotting was applied on genomic DNA derived from edited cells, healthy controls and diseased cells. Briefly, DNA was extracted using DNeasy Blood & Tissue kit and 5-10µg of genomic DNA was digested overnight at 37°C with *Alu*I- and *Dde*I. Then the digested samples were separated by electrophoresis on 0.8% agarose gel and transferred overnight to a positively charged nylon membrane by capillary blotting (Roche Applied Science). In the next day, membrane was cross-linked by UV irradiation, and hybridized for overnight at 55°C with 1ng/ml digoxigenin-labeled $(GGGGCC)^5$ oligonucleotide probe (Integrated DNA Technologies) in 25ml hybridization reaction (EasyHyb Granules, Roche). Following denaturation of the probe at 95°C for 5 min, probe was snap cooled on ice for 30 sec and immediately added to the prehybridised membrane. The following day, the membrane was washed twice using stringency washes (2X Saline-Sodium Citrate buffer (SSC) and 0.1% sodium dodecyl sulfate (SDS), each for 15 min at 65°C in the hybridisation oven. The membrane was then washed twice at the same temperature and conditions with 1X SSC and 0.1% SDS and finally washed twice with 0.5X SSC and 0.1% SDS for further 15 min at room temperature. The membrane was washed with 1X wash buffer (Roche Applied Science) for 2 min, then incubated with blocking solution (Roche Applied Science) for 1–2 hr. After that, the membrane was incubated with Anti-DIG-AP Fab fragments (Roche) for 25-30 min and washed three times with 1X washing buffer. Ready-to-use CDPD (Roche Applied

Science) chemiluminescent substrate was added to the membrane and signals were visualised on Detection Film. All samples were compared to DIG-labeled DNA molecular-weight marker II (Roche Applied Science) and repeat number was estimated compared to the DNA marker sizes.

2.10. Characterisation and motorneuron differentiation of corrected clones

2.10.1. Differentiation of corrected iPSC lines into MNs.

The iPSCs from two edited lines (Ed-02-01 and Ed-02-02), two different healthy controls (WT-01 and WT-02) and two diseased lines from the same patient (C9-02-02 the original parental clone and C9-02-03 iPSC line) were differentiated into mature MNs as described in a previously published protocol with some modifications (Maury et al. 2014). Briefly, the iPSCs were grown on Matrigel until confluency, media was changed and neural induction (NI) was started using DMEM/F12: Neurobasal 1:1 media containing 1X N2, 1X B27, 0.5mM ascorbic acid, 100µM 2-mercaptoethanol, 3mM Compound C and 3mM Chir99021. From day 2 onward, 1mM RA and 500nM SAG (Calbiochem) were added to the medium. On day 3, Chir99021 and Compound C were removed completely from the medium, and the cells were cultured for another 4–5 days in NI medium. After that, cells were split 1:2 on Geltrex coated plates and media was changed every other day for another 10 days. On day 20, cells were split harshly for complete maturation of MNs into new 6-well plates and 24-well plates on 12mm coverslips as required for MN differentiation assessment and phenotyping assays. The same neural induction media with RA/SAG was used and supplemented with growth factors BDNF (10 ng/ mL) (R&D systems), GDNF (10 ng/mL) (R&D systems), DAPT (10 mM), and laminin (0.5 mg/mL), medium was

changed every other day. On day 7, DAPT was removed from the medium, and the neurons were maintained on neural maturation medium for another 4-5 days before they have been used on phenotypes analysis at day 30. All cultures were grown at 37°C in 5% CO2 conditions. Neural induction medium was supplemented with 10 μM Y27632 after each passage.

2.10.2. Confocal microscopy and Immunoflourescence

Immunostaining was performed at day 30 as described in section 2.2.4. Cells were stained with primary goat anti-ChAT, mouse anti-SMI-32, rabbit anti-Islet1/2 and rabbit anti-HB9, and goat anti-Olig2, rabbit anti-PAX6, and mouse anti-Nestin (Appendix I). Adequate Alexa Fluor conjugated Secondary antibodies (AlexaFluor 488, 568, and 647) were used as required (1:500, Life Technologies), Images were obtained with Ziess confocal microscope. The percent of positive cells were quantified using ImageJ software and normalised to the number of cells positive staining for $Olig2^{+}$ or $SMI\text{-}32^{+}$ or $ChAT^{+}$ on each image. The quantification of cells after immunofluorescence staining was based on identifying the iPSC-derived MNs positively stained for either SMI-32 and/or ChAT at day 30 using well-established antibodies that have been used successfully by our group in immunostaining of iPSC-derived MNs from several MN differentiation protocols (Dafinca et al. 2016). For each assay, images were taken for at least ten representative fields and percent positive was analysed using imageJ analysis software. The immunostaining of each MN marker was repeated at least three times from three independent differentiation experiments.

2.11. Studying ALS-disease related phenotypes

2.11.1. C9orf72 Promoter methylation assay

Quantitative PCR with methylation-sensitive restriction enzymes was applied to measure the DNA methylation level as described previously (Russ et al. 2015; Liu et al. 2014). Briefly, 100ng of genomic DNA was digested with 10 units of *Hha*I (NEB) and 2 units of *Hae*III (NEB) or 10 units of *Hpa*II (NEB) and 2 units of *Hae*III (NEB), to assess the methylation status in two CpG sites within the promoter region. DNA samples digested with only 2 units of *Hae*III were used as a mock reaction. Digestion mixtures were incubated overnight at 37°C followed by heat inactivation at 80°C for 20 min. qPCR was carried out using 30ng of digested DNA per reaction and 2X FastStart SYBR Green Master mix (Applied Biosystem). The specific primers for amplifying the methylated promoter region of C9orf72 are listed in Appendix III. The difference in the number of cycles to threshold amplification between double (*Hae*III and *Hha*I) versus single digested DNA (*Hae*III) was used as measure of CpG methylation. *Hae*III was added to digest the hexanucleotide repeat to increase the efficiency of PCR amplification. To determine the linearity of this assay, a methylated DNA standard was generated by *in vitro* methylating DNA from a non-expanded healthy control DNA sample using M.SssI (NEB) for overnight at 37 °C. Methylated DNA samples were purified using PCR Cleanup kit (Qiagen), and different ratios of methylated to unmethylated DNA were mixed and used in *C9orf72* promoter methylation assay. Both *Hha*I and *Hpa*II restriction sites are within the unmethylated *C9orf72* promoter and the amplification of the unmethylated digested products after enzyme digestion represents the methylation level.

2.11.2. Bisulfite sequencing

Direct bisulfite sequencing was achieved using an assay adapted from Xi et al., (2013). Briefly, 400ng of genomic DNA was bisulfite-converted using the EpiTect bisulfite kit (Qiagen). PCR was performed on bisulfite-treated DNA to amplify the 5' CpG island as described previously (Liu et al. 2014), using the following primers *C9orf72*: forward 5'-GGAGTTGTTTTTTATTAGGGTTTGTAGT-3', reverse 5'-TAAACCCACACCTACTCTTACTAAA-3'. GoTaq Hot Start polymerase (Promega) was used as per manufacturer's instructions using the following thermal cycling conditions: 95°C for 5 min, followed by 10 cycles of touchdown PCR with 30 sec denaturation at 95°C, 30 sec annealing reducing the annealing temperature from 70°C to 60°C by -1°C per cycle and with 1 min extension at 72°C, followed by 30 cycles with the annealing temperature at 60°C and with a final extension step at 72°C for 10 min. PCR products of 466bp were purified by agarose gel extraction and DNA-sequencing by Source BioScience using the forward primer. Resulting chromatograms were analysed using Chromas v2.6 software and the percentage of methylcytosine (mC) determined for each CpG using relative peak heights whereby %mC = C/(C + T). In all samples over 95% of unmethylated C nucleotides were converted to T.

2.11.3. Quantitative RT-PCR for C9orf72 all and the three transcripts.

Quantitative real time PCRs (qRT-PCRs) were performed using 2µl of each cDNA samples and SYBR Fast mastermix (Applied Biosystems) and specific primers for overall *C9orf72* mRNA levels, V1 transcript, V2+V3 transcripts and V1+V3 transcripts, in a total volume of 25µl per each reaction. Master mixes and cDNA samples were loaded in a 96-well plates and samples were run on Applied Biosystems StepOne machine and analysed

using relative quantitation methods. All samples were amplified in triplicates from at least three different differentiation experiments. All *C9orf72* gene variants expression analyses were performed using ΔΔCt method with normalisation to GAPDH reference gene. Melting curve analysis was used to verify the specificity of the primers. PCR cycling conditions were as the following: 95°C 20 sec initial denaturation, followed by 40 cycles of [95°C 3 sec, 60°C 30 sec]. Primer sequences are listed in Appendix III.

2.11.4. Hexanucleotide repeat methylation assay using RP-PCR

One microgram of each DNA samples was digested individually with the following restriction enzymes: 10 units of *Hpa*II, 10 units of *Msp*I, or 4 units of *MspJ*I (NEB) overnight at 37°C followed by enzyme heat inactivation. Approximately, 100ng of the digested DNA was used for RP-PCR as described in Section 2.8.6.1. *In vitro* methylated DNA control sample with *M.Sss*I (NEB) restriction enzyme was used as a control to check for the accuracy of *MspJ*I in digesting the methylated DNA. One microgram of *M.Sss*I digested DNA was digested again with 2 units of *MspJ*I at 37 °C for 4h and purified using heat inactivation prior to RP-PCR.

2.11.5. Fluorescence in Situ Hybridization (FISH) for detection of sense and antisense RNA Foci

On the last splitting at day 20, MN cells were grown on coverslips in 24-well plates until maturation, and then fixed using 4% PFA for 20 minutes, washed in PBS and permeabilised using 0.2% Triton X-100 (Sigma) in PBS for 30 minutes. Cells were then dehydrated in a graded series of alcohols (50%, 70%, 90% and 100%), air dried, and rehydrated in PBS, and washed in 2X standard sodium citrate (SSC) and then incubated in prehybridization solution (50% formamide, 2X SSC) for 1hr and then incubated in

hybridization solution for 1h and 30 min (50% formamide, 10% dextransulfate, 2X SSC, and RNAse inhibitor) in dark in hybridization oven with agitation. Probe sequences were (C2G4x3)-Cy3 or (G4C2x3)-alexa-488. Following hybridisation washes, staining with SMI-32 (1:1000, DSHB) was performed to assess RNA foci in SMI-32+ cells.

2.11.6. Dot-blot analysis

For dot-blot analysis, nitrocellulose membrane was used, and 10µg of protein were loaded as a dot in 10µl total volume of RIPA-1 buffer. Four dots were prepared for each sample to stain with four different antibodies. The blot was allowed to dry for at least 1h and 30 min and blocked in 5% skim milk in 0.1% Tween/PBS for 1h. After that, the blocking buffer was removed, and the membrane was cut into four segments, each one involved the same samples but incubated with four different primary antibodies as the following: rabbit anti-GA (ProteinTech, 1:100), rabbit anti-GP (ProteinTech, 1:100), rabbit anti-GR (ProteinTech, 1:100), and β-actin (1:500), added specifically to each blot and incubated overnight at 4°C. The membrane was washed three times for 10 min each. Then the membrane was incubated with HRP-conjugated anti-rabbit secondary antibodies for 1h at RT and then washed three times 10 min each in 0.1% Tween/PBS. Signal was detected using ECL reagents (Amersham Pharmacia Biotech).

2.11.7. Cleaved caspase-3 assay for apoptotic cell death assessment

For cleaved Caspase-3 (Clvd-Casp-3) assay, cells were fixed in 4%PFA, washed, and then used for immunofluorescence staining. Cells were co-stained with anti-clvd-casp-3 and SMI-32 to identify mature MNs and DAPI for nuclear staining. At least five representative images were taken for each sample. Staining was performed in triplicate from three independent differentiation experiments. Cell death was estimated by counting

clvd-Casp3^{+} cells over total SMI-32^{+}/ DAPI^{+} cells.

2.11.8. Glutamate toxicity and Propidium Iodide (PI) Staining.

Cell cultures on coverslips were incubated with 10μM, 50μM and 100μM of L-Gltamate for 5h, and then cell culture medium was aspirated and PI staining was performed by incubating the neurons with 20 μg/ml PI and 20μl RNase A for 20 min at 37°C. The solution was then washed with PBS, and cells were fixed in 4% PFA. Immunostaining was then performed for SMI-32 to stain MN+ cells.

2.11.9. PABP Stress granule (SG) Immunostaining

Mature MNs were fixed with 4% PFA and incubated overnight at 4°C with rabbit anti-PABP (1:500) and mouse anti-SMI-32 (Covance, 1:1,000). Adequate Fluor 488 or Alexa Fluor 568 conjugated goat anti-rabbit or goat anti-mouse secondary antibodies were used and analysed by confocal microscopy.

2.11.10. Assessment of cell viability and G3BP1 stress granule formation after exposure to stressing agents

For SG formation under stress condition, cells were treated with 3mM Sodium Arsenite for 1h 30 min, then cells were stained with anti-G3BP1 antibody directly after stress recovery step for 1h and counterstained with SMI-32. Images were taken using a 60X objective under confocal microscope and from different areas of the coverslips. Data are represented as total number of SGs per SMI-32 positive cells.

2.12. Data analysis and statistics

All statistical analyses were performed using GraphPad Prism. A minimum of three replicates were performed for differentiation and phenotyping assays. Statistical analysis for comparison between groups was performed by One-Way ANOVA test and followed by a Bonferroni post-hoc comparison analysis. Results from cell culture experiments are shown as % (mean±SEMs). P-values ≤ 0.05 were considered significant (*≤ 0.05, ** ≤ 0.001, *** ≤0.0001)

3. CHAPTER THREE

Generation and characterization of C9orf72 patient induced Pluripotent Stem cells (iPSC)

3.1. Introduction

In 2006/2007, Yamanaka et al., reported that mouse and human skin fibroblasts could be reprogrammed to a pluripotent state with retroviruses encoding for *KLF4, SOX2, OCT4,* and *c-MYC* (Takahashi & Yamanaka 2006; Takahashi et al. 2007). The main advantage for using human iPSCs (hiPSCs) is their capacity of indefinite self-renewal, and the ability to differentiate into any cell lineage.

iPSC technology represents a useful tool for disease modelling and studying the pathophysiological disease mechanisms. Recently, the combination between gene editing techniques and iPSCs modelling represents a promising approach for understanding the underlying disease mechanisms. Since the original disease-causing mutation is present in patient iPSCs, gene correction studies can be performed to generate isogenic lines carrying the same genetic background of the original disease cells. Generating the isogenic lines would facilitate the generation of disease models *in vitro* and allows for drug screening and therapeutic applications.

The main aim of this chapter was to generate and characterize iPSC lines from ALS patients carrying the HRE mutation. This has been achieved by transducing fibroblast cells with Sendai viruses carrying the four Yamanaka reprogramming factors. The generated iPSC lines were characterized and evaluated for their pluripotency and karyotyping profile and then differentiated into MNs using EBs-based protocol.

3.2. Reprogramming of G4C2 patient fibroblast cells to pluripotency

The generation of iPSCs from fibroblast somatic cells has been achieved using different reprogramming protocols to express pluripotency factors, mainly *OCT4, SOX2, KLF4, c-MYC*, and *NANOG*. In this project, Sendai viruses were used to reprogram iPSCs from ALS and ALS\FTD patients' dermal fibroblasts into a pluripotent state. The schematic protocol for generating iPSCs from skin biopsies is summarised in Figure 3.1.A. First, skin biopsies from two ALS patients and one ALS/FTD patient carrying the HRE mutation in the C9*orf*72 gene were obtained and designated as the following; C9-01, C9-02 and C9-04.

The presence of the HRE was confirmed by southern blotting and RP-PCR in those samples before taking the biopsies (South Wales Research Ethics Committee approval number: 12/WA/0186). The HRE size in patients C9-01 and C9-02 fibroblasts was identified by Southern blotting before iPSC reprogramming. Both samples showed a massive HRE between 4-6kb in size. Skin fibroblasts were isolated and multiple iPSC lines were generated by transduction of the cells with Sendai viruses (Cytotune 1 kit) carrying the four Yamanaka transcription factors (OCT3/4, SOX2, KLF4, and c-MYC).

After three weeks of culture on MEFs, approximately 10-12 hESC-like clones emerged (efficiency was about 0.001 %), and up to ten single colonies were picked up into feeder cells individually (Figure 3.1.B). After that, individual clones were cultured and passaged by manual dissection to increase the number of colonies and to reach enough quantity for clonal expansion (Figure 3.2.A).

The isolated iPSC clones were different in terms of morphological characteristics and proliferation capacity. Some clones reached the confluency in less than five days with typical iPSC morphology and no random differentiation was observed. At passage 6, three clones for each patient of the best observed morphology and proliferation capacity were expanded and characterized. After that, the successfully characterized cell lines were identified as: C9-01-01, C9-01-06 and C9-01-07 derived from C9-01 ALS patient dermal fibroblasts; C9-02-02, C9-02-03 and C9-02-10 derived from C9-02 ALS/FTD patient dermal fibroblasts; and C9-04-01, C9-04-09 and C9-04-12 derived from C9-04 ALS patient dermal fibroblasts (Table 3.1).

The established iPSC lines displayed a morphology and cell characteristics similar to ESCs. The iPSCs formed tightly packed and flat colonies with large cellular nuclei and scant cytoplasm when cultured on MEF feeder layer (passage 6) as shown in Figure 3.2.B. All iPSC lines were transferred to feeder-free conditions on Matrigel after 10-12 passages and maintained in mTeSR1 medium (Figure 3.2.C). Control iPSC lines were chosen by matching age and gender with our patient lines. The wild-type control iPSCs SFC-840-03-03 (WT-01) and SFC-856-03-04 (WT-02) used in this project were prepared and characterized in the James Martin Stem Cell Facility and they have also been used and

characterized in previous studies (Fernandes et al. 2016; Dafinca et al. 2016). The G4C2 size for wild-type controls were evaluated by RP-PCR; WT-01 harbors six repeats and WT-02 harbors two repeats.

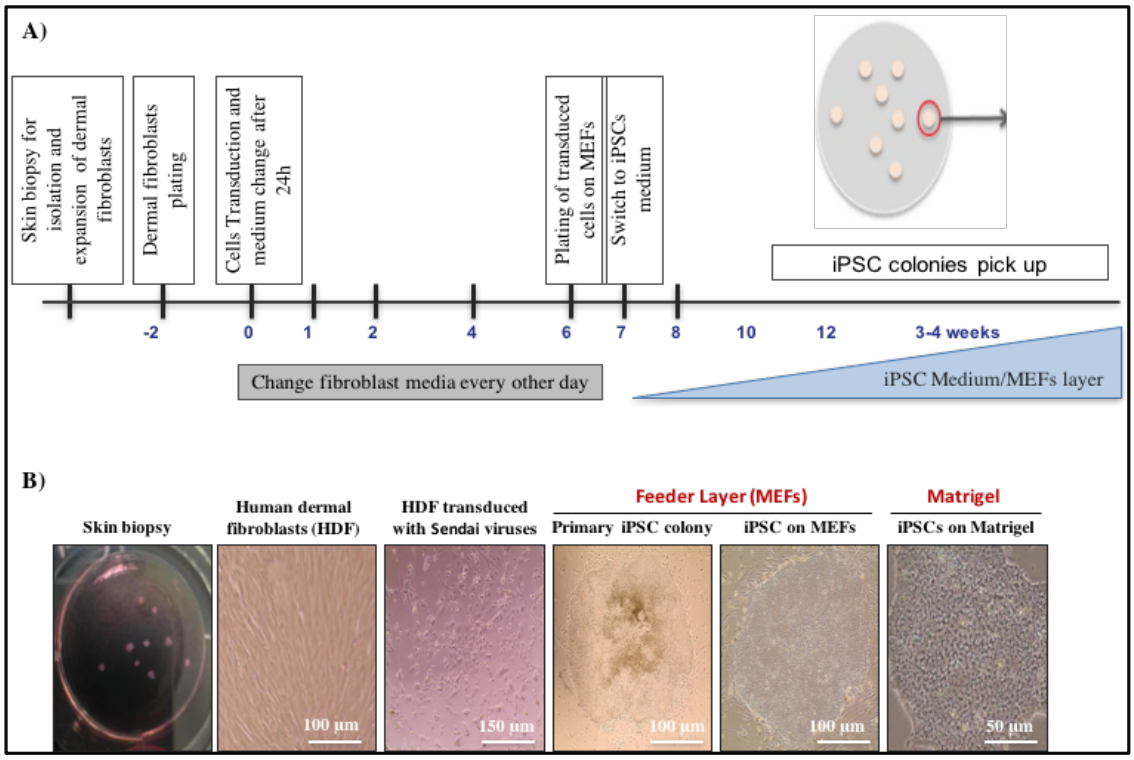

Figure 3. 1 Generation of iPSC lines from patient's dermal fibroblasts carrying the HRE mutation. A) Timeline overview of iPSCs reprogramming from skin fibroblasts using the CytoTune1 reprogramming protocol. **B)** Morphological representation and phase-contrast microscopy images of the reprogramming and expansion processes. Dermal fibroblasts were isolated from skin biopsies, and then transduced with Sendai viruses carrying Yamanaka factors. iPSC colonies were observed 3–4 weeks following viral transduction, manually picked up and cultured on feeder cells then moved to feeder-free conditions on Matrigel and maintained on mTeSR1 medium. All scale bars are shown on the figure.

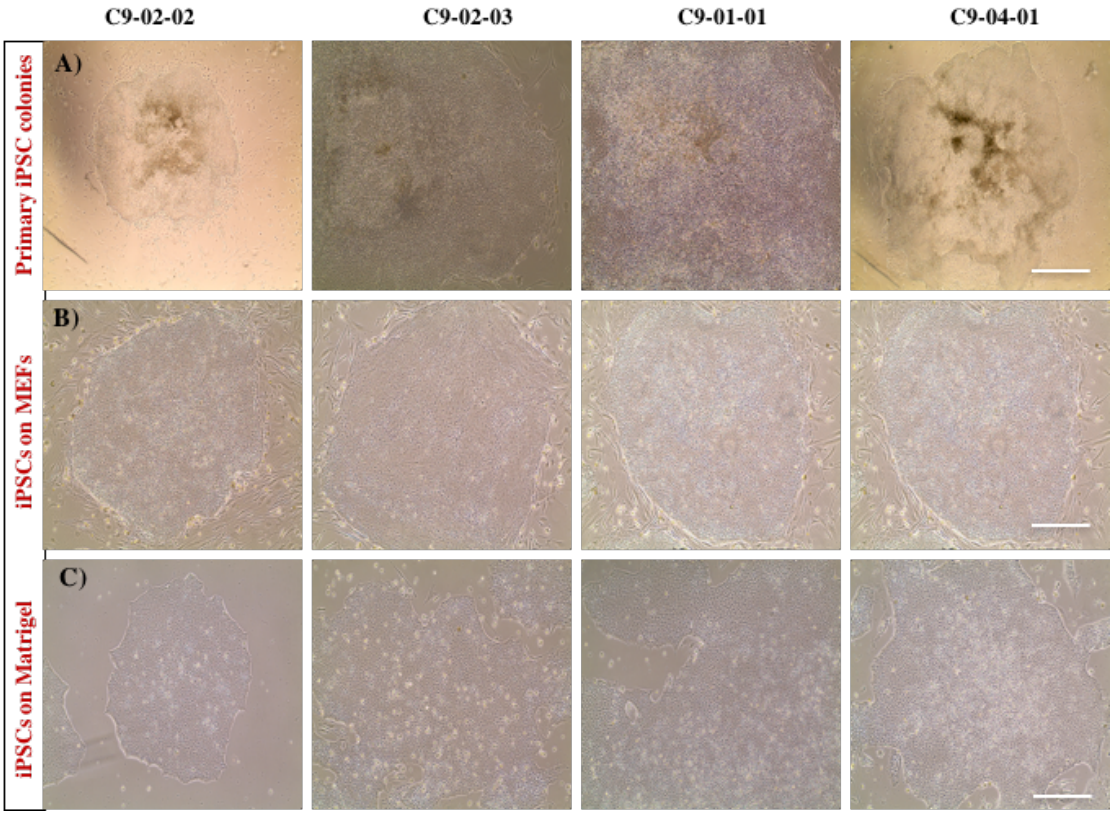

Figure 3. 2 Phase-contrast representative images for the morphological characteristics of the generated iPSC lines. A) Primary isolated ESC-like colonies emerged 3–4 weeks following viral transduction. **B)** iPSC colonies after passaging on feeder layer (at passage 6), showing compact colonies in morphology. **C)** iPSCs cultured on Matrigel and maintained on mTeSR1 (at passage 15) were grown as flat and dense colonies with smooth and round boarder and high nucleus:cytoplasm ratio. Scale bars represent 100μm.

Three iPSC lines were generated for each patient, expanded and banked in LN for long-term storage to give rise to stable cell lines. The isolated iPSC lines were extensively characterised as described in Table 3.1, based on their morphology and ability to express pluripotency markers, karyotyping results, Sendai virus clearance and Mycoplasma testing result.

Table 3. 1 Characterisation of established human iPSC lines from patients and controls.

SAMPLE ID	AGE / GENDER	iPSC CLONES	KARYOTYPE ANALYSIS	TRA-1-60 AND NANOG	SENDAI CLEARANCE	MYCOPLASMA TESTING
C9-01 (ALS)	72/M	C9-01-01	✓	✓	Cleared	Negative
		C9-01-06	✓	✓	Cleared	Negative
		C9-01-07	✓	✓	Cleared	Negative
C9-02 (ALS/FTD)	62/F	C9-02-02	✓	✓	Cleared	Negative
		C9-02-03	✓	✓	Cleared	Negative
		C9-02-10	✓	✓	Cleared	Negative
C9-04 (ALS)	39/M	C9-04-01	✓	✓	Cleared	Negative
		C9-04-09	✓	✓	Cleared	Negative
		C9-04-12	✓	✓	Cleared	Negative
SFC840	67/F	SFC840-03-03 (WT-01)	✓	✓	Cleared	Negative
SFC856	62/F	SFC856-03-04 (WT-02)	✓	✓	Cleared	Negative

3.3. Patient iPSC lines express pluripotency markers and show normal karyotypes

All iPSC lines were assessed by immunocytochemistry (ICC) for the expression of undifferentiated hESC-specific pluripotency markers: specific embryonic antigen-4 (SSEA4), tumor-related antigen (TRA-1-60), OCT3/4 (an ESC transcription factor) and NANOG. All iPSC lines were also assessed by flow cytometry for TR1-160 and NANOG. iPSC colonies showed cell surface staining for both SSEA4 and TRA-1-60 and nuclear staining for NANOG and OCT4 by ICC (Figure 3.3). They also stained positive for TRA-1-60 and NANOG by flow cytometry (Figure 3.4).

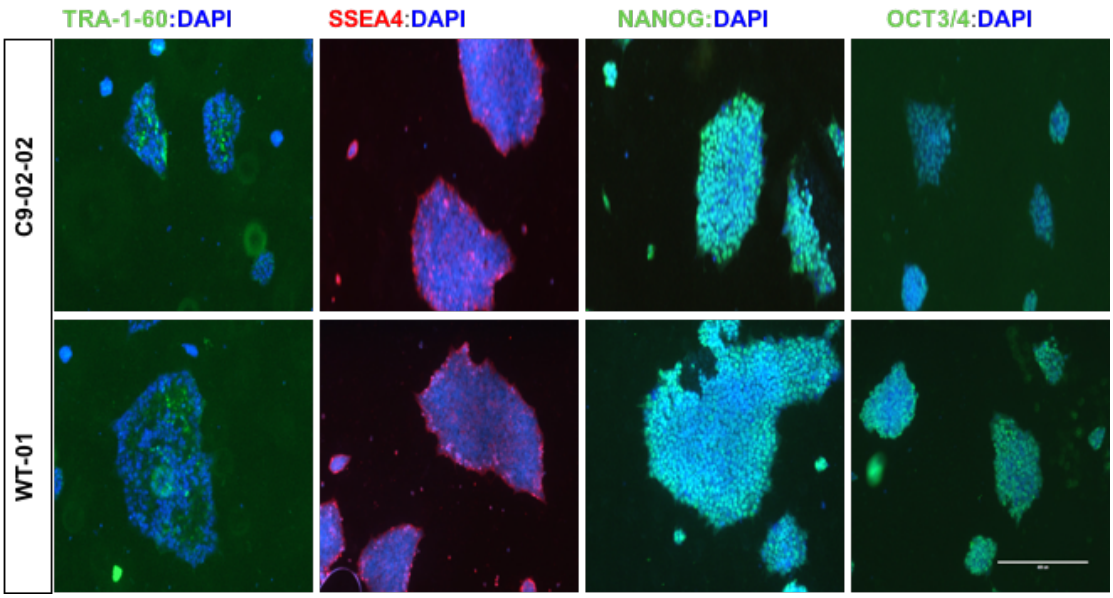

Figure 3.3 Characterisation of iPSC lines by immunocytochemistry staining of pluripotency markers. iPSCs showed positive expression of nuclear factors NANOG (green) and OCT3/4 (green) and cell surface markers SSEA4 (red) and TRA-1-60 (green). Nuclei (blue) counterstained with DAPI. Scale bar 100µm.

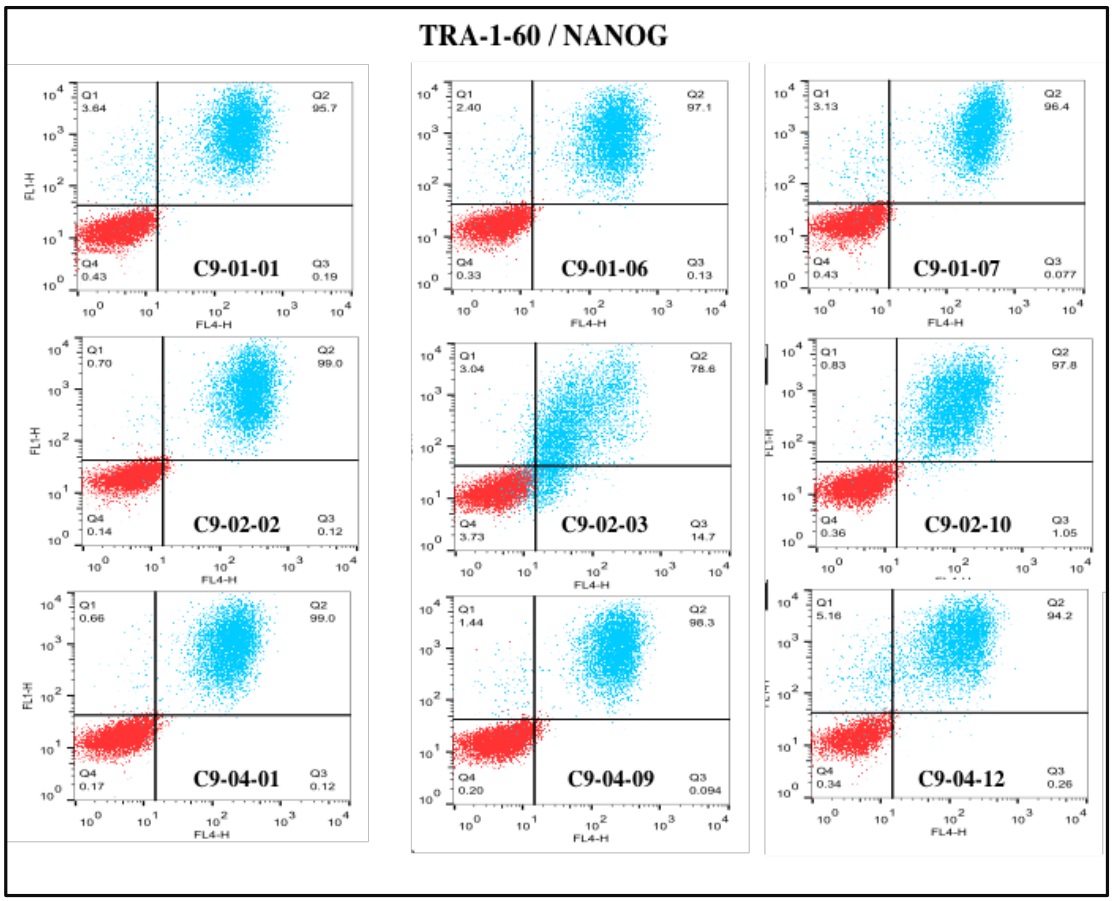

Figure 3. 4 Characterisation of the isolated iPSCs by flow cytometry. Blue dot plots demonstrate the positivity of the analysed iPSCs for TRA-1-60 and NANOG, while the red dot plots show the cells negativity for isotype controls.

All lines were also karyotyped by SNP profiling array (Illumina CytoSNP-12-v2.0 array) to rule out any genetic abnormalities or chromosomal aberrations. This assay contains nearly 300,000 genetic markers and provides detailed resolution of genome integrity. When compared to the karyotyping results of the original fibroblast DNA samples, karyotype analysis profiles of the iPSC lines showed no aneuploidies or large duplications/deletions (Figure 3.5.A-C). This karyotyping assay is considered faster and more representative than the traditional methods of iPSC karyotyping. Autosomal detected regions deviating from reference data are annotated with green (amplification) or with orange (deletion) bands. Finally, Sendai virus clearance was also confirmed by RT-PCR amplification of the Sendai virus genome and of the Yamnaka transcription

factors, confirming that the isolated lines were free from the virus genome (Figure 3.6).

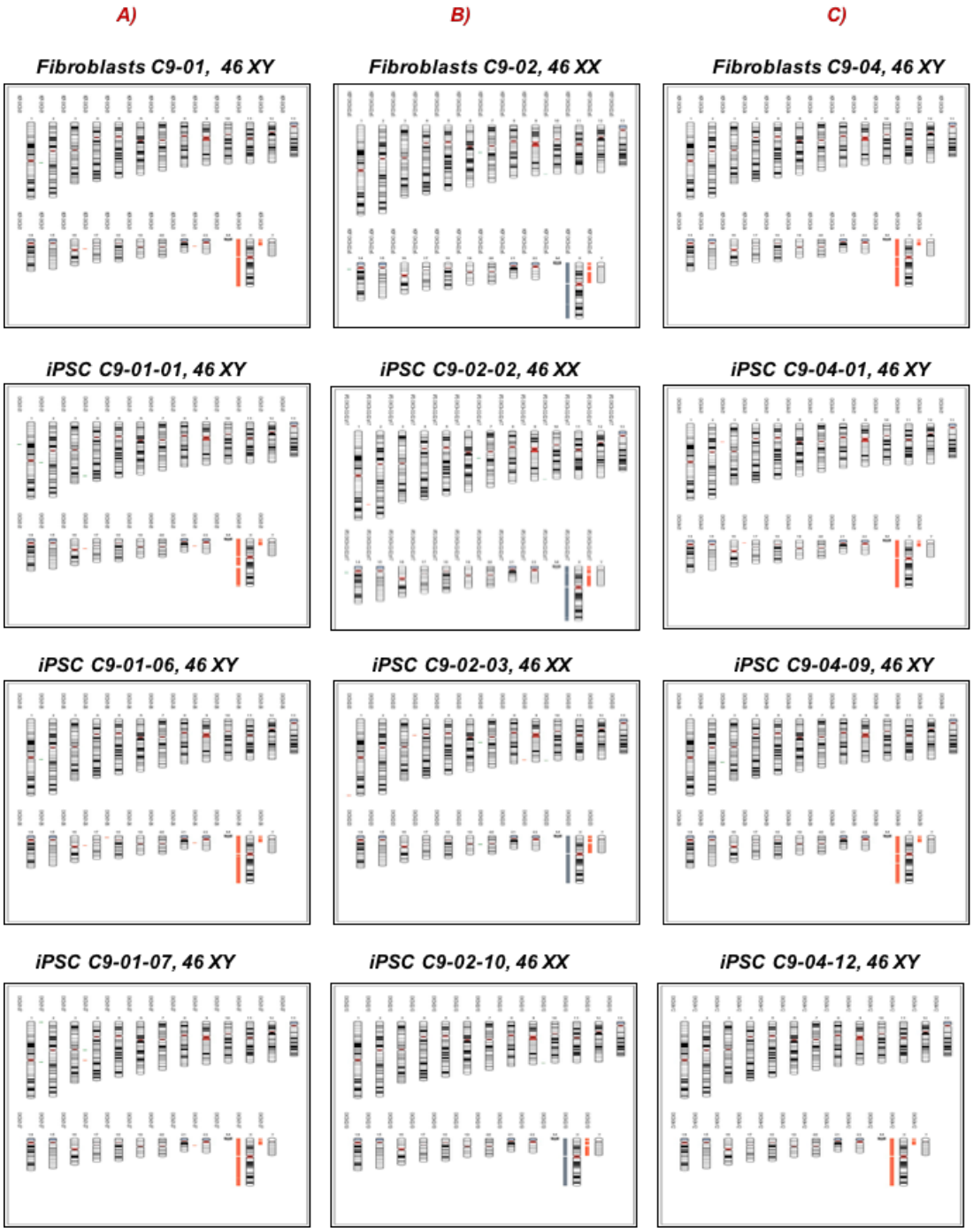

Figure 3. 5 Karyotype analysis of the established iPSC lines. All iPSCs showed normal karyotype and their chromosomal profiles showed no abnormalities or chromosomal aberrations when compared to the original dermal fibroblasts. **A)** Karyograms generated from SNP analysis of C9-01-01, C9-01-06 and C9-01-07 clones when compared with the original human dermal fibroblasts for the same patient C9-01. **B)** Karyograms generated from C9-02-02, C9-02-03 and C9-02-10 clones when compared to dermal fibroblasts for the same patient C9-02. **C)** Karyograms of C9-04-01, C9-04-09 and C9-04-12 clones when compared to C9-04 dermal fibroblasts.

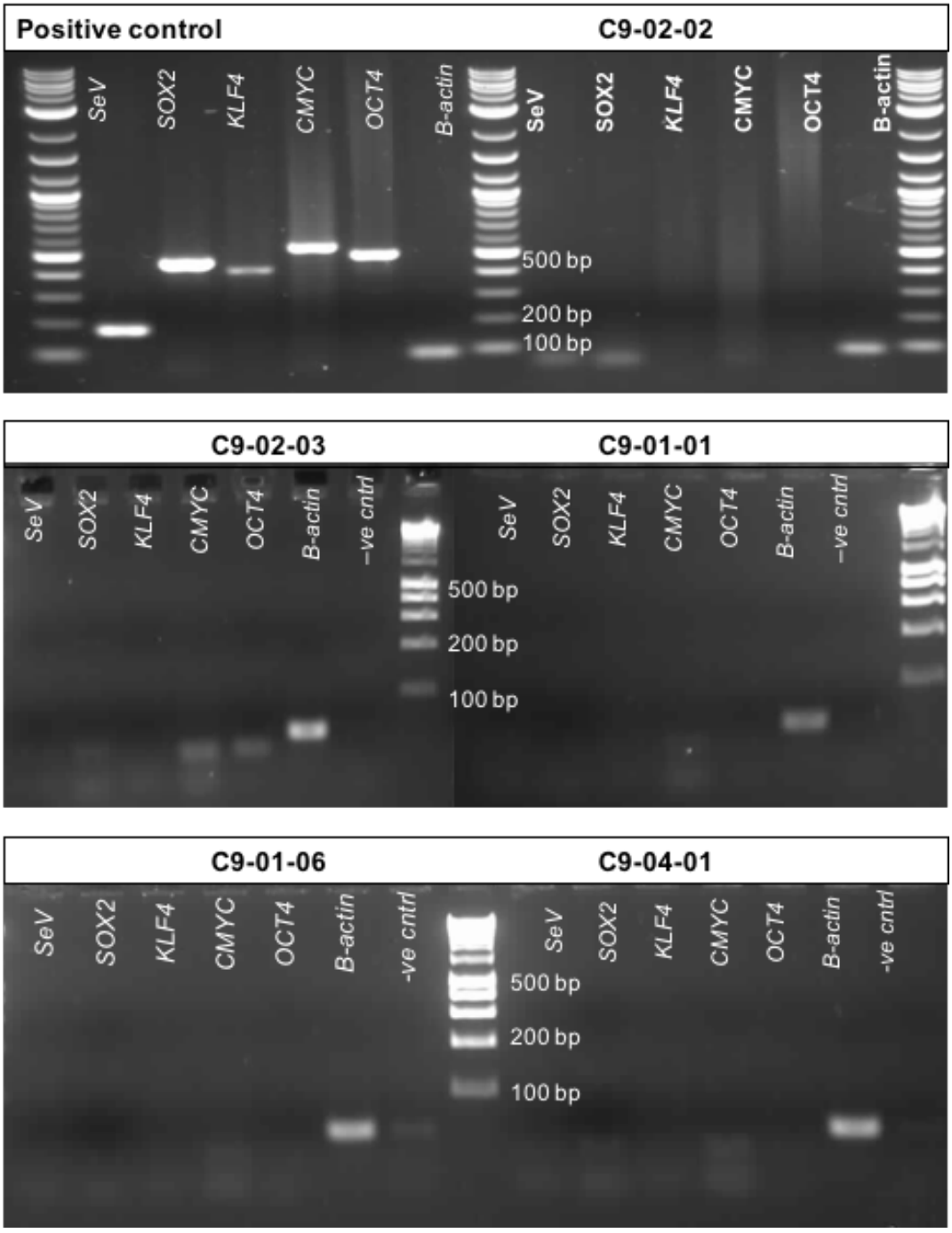

Figure 3. 6 RT-PCR confirming Sendai virus clearance from the isolated iPSC lines. PCR amplification of Sendai virus genome and SOX2, KLF4, OCT3/4 and c-MYC transcription factors were performed. All genes were amplified separately with a common reverse primer that binds to the Sendai virus genome. The amplification of cDNA from a fibroblast sample after 5 days of transfection with Sendai viruses was used as a positive control. A **β**-actin internal control was used to establish the integrity of the used cDNA. Band sizes are: Sev: 181 bp, SOX2: 481 bp, KLF4: 410 bp, c-MYC: 532 bp, OCT4: 483 bp, and **β**-actin: 93 bp.

A previously established EBs differentiation protocol was applied (Hu & Zhang 2009) to differentiate two ALS/FTD C9-02 patient-derived iPSC lines (C9-02-02 and C9-02-03) into MNs. This was done to determine the capacity of iPSC lines to produce EBs representing the three germ layers, and to determine their differentiation potential into MNs before using them in gene editing experiments. The WT-01 wild type iPSC line was also used in the differentiation experiment as a healthy control.

The differentiation process of iPSCs into spinal MNs involves three steps: neuralisation, caudalisation, and ventralisation. This differentiation protocol uses serum-free media to generate rosettes in the presence of retinoic acid (RA) and sonic hedgehog (SHH) pathway activators, smoothened agonist (SAG) and Puromorphamine, which caudalise and ventralise MN progenitors (Figure 3.7.A).

iPSCs showed the capacity to re-aggregate and generate EB-like structures. EBs-based neural induction with RA and SHH pathway activators formed rosettes after two weeks of induction, and then neurospheres on the third week. After that, neurospheres were isolated and plated on Geltrex coated plates to produce mature MNs. At day 28, brain-derived neurotrophic factor (BDNF), glial-derived neurotrophic factor (GDNF), insulin-like growth factor-1 (IGF1), and cAMP were added to the medium for complete maturation and cell survival enhancement of the generated MNs. Maturation steps of MN cultures were primarily identified according to morphological changes (Figure 3.7.B). Two days after plating the neurospheres, MNs started to differentiate from the neurospheres by forming bipolar appearance with one long neurite, which then started to increase in length after few days of seeding the neurospheres.

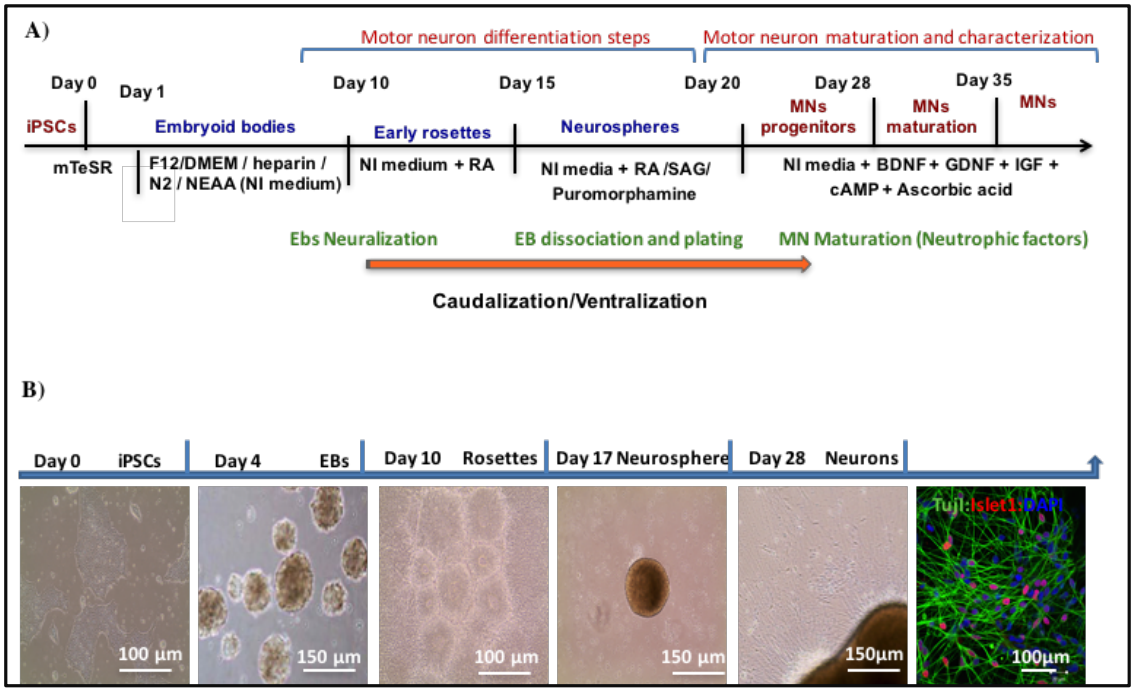

Figure 3.7 Overview of EB-based MN differentiation protocol. A) Schematic representation of MN generation timeline from ALS/FTD patient iPSCs. **B)** Phase-contrast imaging showing morphological changes during neural induction and differentiation process from iPSCs into MNs. From left to right: iPSCs on Matrigel, EBs neuralised to form neural tube-like rosettes, neurospheres, maturation into MNs and immunostaining of mature MNs with Islet-1 (MN marker) in red and DAPI (nucleus) in blue. All scale bars are shown on the figure.

At week 5, more neurites were visible and they were more branched, forming a network-like structure. Differentiated cells were characterised by immunostaining of Islet1-1 transcription factor, which is previously described as the earliest marker of developing MNs (Francius & Clotman 2010); Tuj1, a general neuronal marker, and HB9 (also known as Mnx1), which is also a MN-specific transcription factor. Quantification of positive cells after immunostaining showed that 15-20% of cells were positive for HB9 and Islet-1 and over 90% of differentiated cells were Tuj1-positive. These results suggest the successful differentiation of ALS/FTD iPSC lines (C9-02-02 and C9-02-03) into MNs with an efficiency similar to the control line (Figure 3.8 A&B). This differentiation process requires more than 5 weeks for the generation of mature MNs.

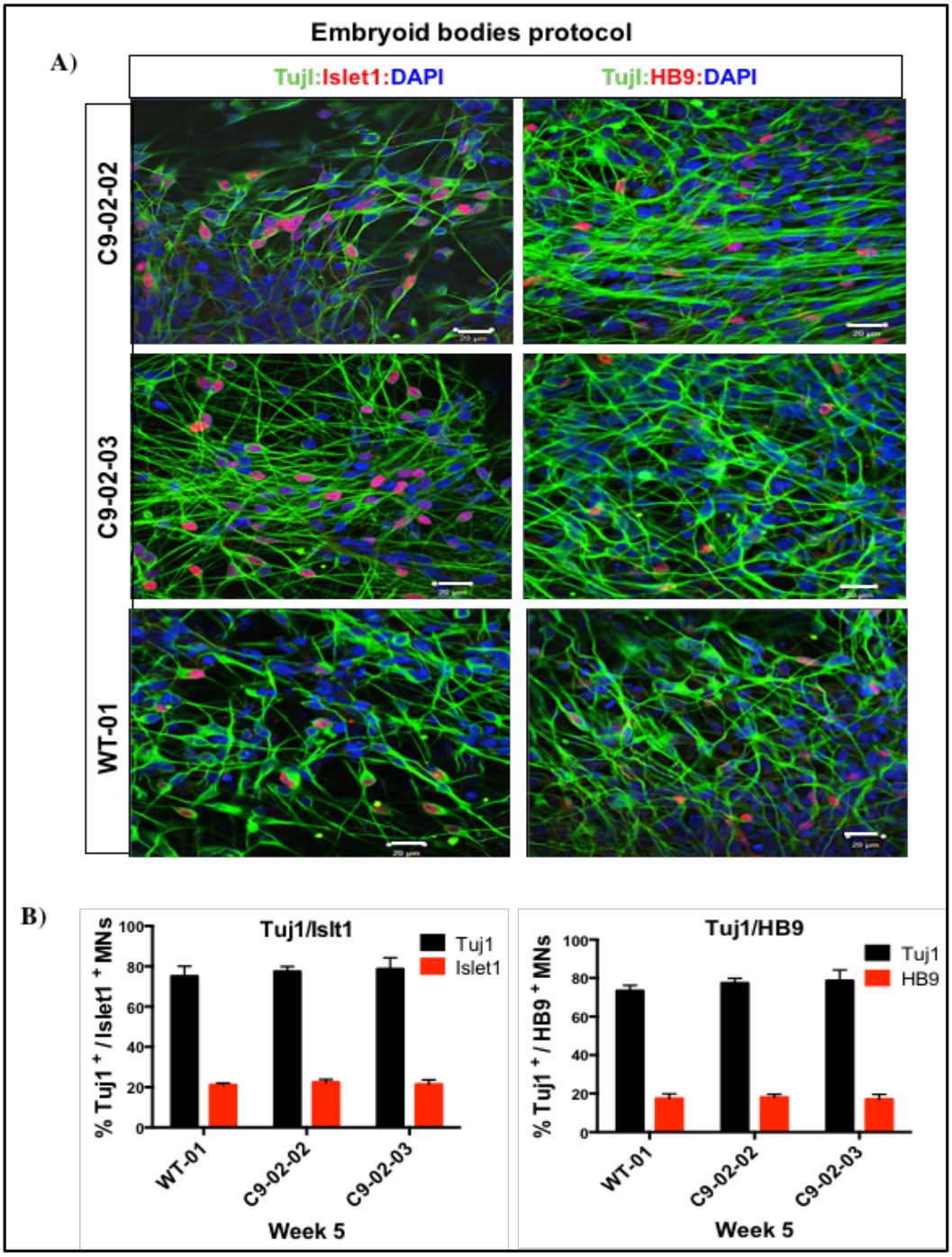

Figure 3.8 EB-based differentiation of iPSCs into MNs. A) Positive immunostaining of Tuj1 and Islet-1 or HB9 was observed without obvious differences between the analysed clones. **B)** Quantification of Tuj1/Islet-1 positive cells and Tuj1/HB9 positive cells after 5 weeks of differentiation. Approximately 15-20% of Islet-1 and HB9-positive cells were also co-stained with Tuj1. No significant differences were observed between control lines and patient lines ($p>0.05$ One-Way ANOVA). Scale bars 20µm**.**

3.4. Studying the stability of repeat expansions after cell reprogramming

The stability of the repeat expansions after the reprogramming process was confirmed for one of the isolated clones (C9-02-02) using different repeat-specific evaluation assays and before its further use in gene editing experiments. First, the repeat expansion was identified on C9-02 fibroblast DNA sample by southern blotting using a $GGGGCC^5$ DIG-labeled repeat specific probe. The southern blot was done by Dr. Andrew Douglas (Oxford, DPAG) and showed a band of approximately 6kb, which corresponds to ~1000 G4C2 repeat units (Figure 3.10.A). The stability of the repeats after the reprogramming process was confirmed using genomic DNA extracted from C9-02-02 iPSCs and showed similar repeat size (approximately 1000 repeats) to the original DNA from fibroblast cells (Figure 3.10.B). Repeat-primed PCR (RP-PCR1) was also performed and showed a typical repeats saw-tooth like electropherogram (Figure 3.10.C). Finally, FISH analysis was carried out to evaluate the presence of the sense RNA foci in the generated iPSCs as shown in Figure 3.10.D. Multiple sense RNA foci were detected in C9-02-02 iPSCs with no foci detected in WT-01 iPSCs.

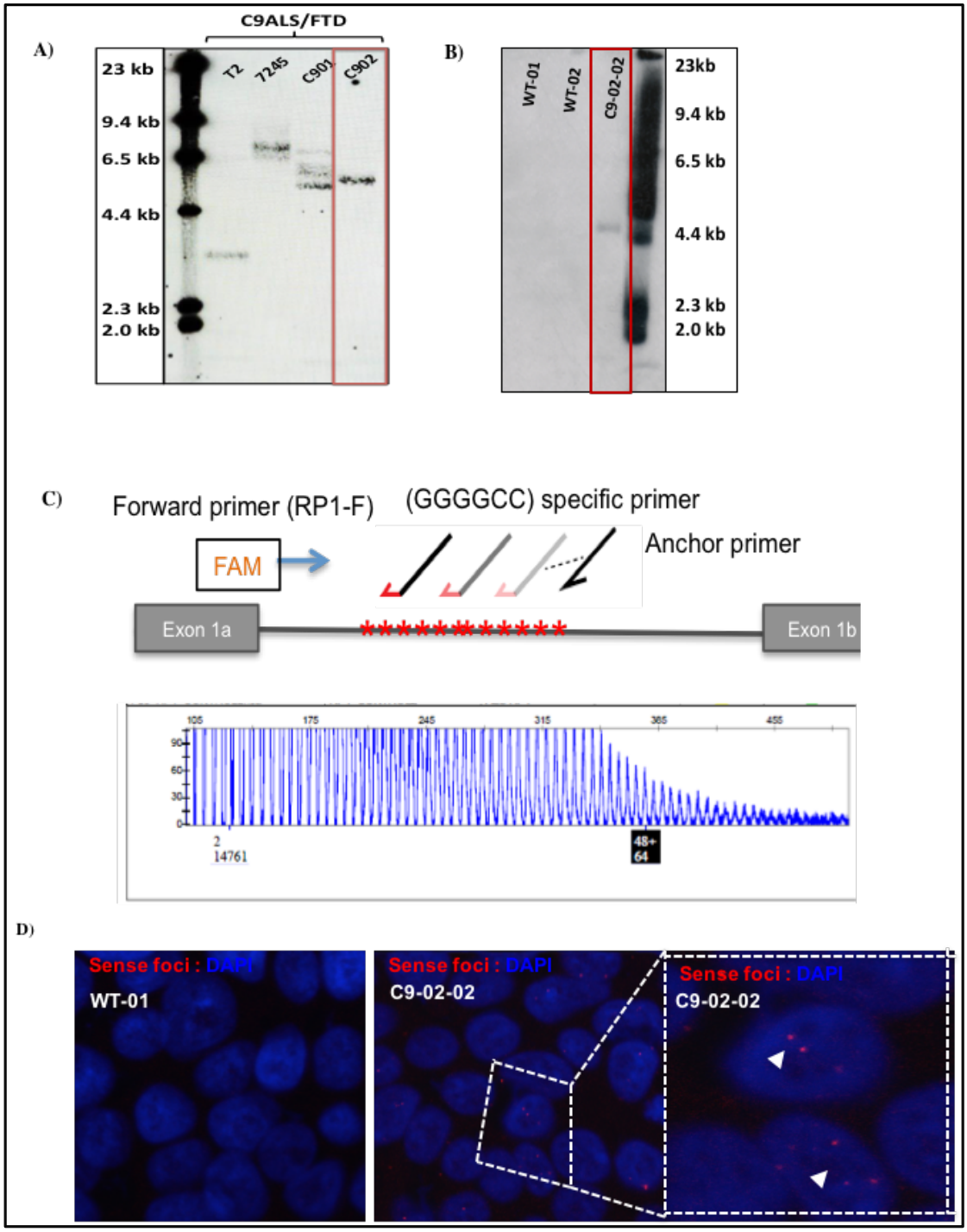

Figure 3.9 Stability of G4C2 repeat expansion after reprogramming. A) Southern blotting analysis for C9-02 fibroblast DNA sample showed a massive G4C2 repeat expansion of approximately 6kb, which is corresponding to repeat size of ~1000 repeat units. **B)** The stability of the repeats was confirmed on genomic DNA extracted from C9-02-02 iPSCs using the same Southern blotting procedure. **C)** Saw-tooth pattern RP-PCR result of C9-02-02 iPSC line. **D)** FISH analysis showed the presence of sense RNA foci in the generated iPSCs, with no foci detected in WT-01 iPSCs.

Conclusion

Challenges with iPSC technology are related to several technical problems, such as reprogramming efficiency which is often less than 0.02%. The cost is usually high and the reprogramming procedures are time consuming and laborious. Generation of iPSC lines free from gene integration that involved in the reprogramming process is also another issue to consider for successful reprogramming. Sendai virus protocol has been used in this thesis for the generation of footprint-free cell lines because it does not integrate into the host genome. However, transfection efficiency was less than 1% and finding alternative protocols with higher efficiency are required for more efficient iPSC lines production.

In conclusion, iPSCs were identified according to their ability to demonstrate typical stem cells features, such as self-renewability and pluripotency by expression of specific pluripotency markers and their differentiation potential. All of the generated iPSC lines from three different ALS patients carrying the HRE mutation were evaluated and showed positive expression of specific pluripotency surface markers, (SSEA4 and TRA-1-60), and the expression of nuclear transcription factors NANOG and OCT4. They also showed normal karyotyping profiles and Sendai virus clearance after reprogramming and expansion. C9-02-02 ALS/FTD patient iPSC line harboring a massive G4C2 repeat expansion was chosen for gene editing experiment using CRISPR/Cas9-mediated HDR and a plasmid donor repair template carrying the wild-type size of the repeats. The selection of C9-02-02 line was based on its typical iPSC morphology and growth characteristics as well as the ability of the clone to produce MNs *in vitro*.

4. CHAPTER FOUR

Gene editing of HRE mutation using CRISPR/Cas9 system

Introduction

iPSC technology offers the opportunity to repair disease-causing mutations *in vitro* and to produce corrected iPSCs isogenic controls. The generation of isogenic lines could help in studying the reversal of disease phenotypes in human cells, by differentiating the corrected patient-derived iPSCs into specific-cell types and studying the disease related phenotypes.

Several approaches of gene editing in iPSCs have been established using artificial nucleases, Zinc finger nucleases (ZFNs), Transcription activator-like effector nucleases (TALEN) and CRISPR/Cas9 system. Recently, CRISPR/Cas9 has rapidly emerged as a powerful genome editing approach to introduce site-specific double-strand breaks (DSBs). Cas9 protein is guided by a 20 nt single-guide RNA (sgRNA) complementary to the target region. This sgRNA binds to the target site upstream of the 5'-NGG protospacer adjacent motif (PAM) and initiates a DSB. The DSB can then be repaired either by non-homologous end-joining (NHEJ) or by homologous recombination (HR) or homology-directed repair (HDR). NHEJ directly connects the broken ends, which results in insertion or deletion (indel) mutations. HDR involves the precise integration of an exogenous DNA sequence into the target site using either a single-stranded oligonucleotide (ssODNs) sequence with short homology arms or a double-stranded DNA donor template. Several

studies have demonstrated the effectiveness of the CRISPR system in targeting hiPSCs for gene knockout and gene correction of disease mutations (Huang et al. 2015; Shen et al. 2013; Yang et al. 2013).

In this chapter, experiments were performed to correct a massive HRE in patient iPSCs using CRISPR/Cas9n and homologous recombination system. To correct the expanded repeats in the *C9orf72* gene, C9-02-02 ALS/FTD patient-derived iPSCs carrying a heterozygous mutation were nucleofected with CRISPR/Cas9n sgRNA plasmids and a dsDNA donor template plasmid carrying the wild-type repeat size. The production of the isogenic lines could facilitate studying of disease phenotypes which are directly correlated to the presence of HRE mutation.

4.1. CRISPR/Cas9 strategy and design for targeting HRE mutation

The *C9orf72* gene consists of 12 exons, and the repeats region is located between two non-coding exon 1a and 1b. The *C9orf72* gene can be transcribed into three variants: V1, V2 and V3 transcripts. The HRE mutation is located in the promoter region of V2 transcript, or in the first intron of V1 and V3 variants. To correct the HRE mutation and generate the isogenic lines, double-nicking CRISPR/Cas9 excision and homology-directed repair (HDR) was applied. A double-stranded DNA (dsDNA) plasmid donor template harbouring two G4C2 repeats was used to replace the massive repeat expansion (~1000 repeats) via homologous recombination. The template was composed of two long homology arms flanking each side of the target site, and a Puro/TK selection/counterselection cassette was also inserted into the donor template to facilitate the isolation of the corrected clones.

4.1.1. Design and construction of sgRNA expression plasmids

Two sgRNA sequences of 20nt each, gRNA1 and gRNA2, were designed using the online CRISPR tool website (http://crispr.mit.edu). The sgRNA target sites were designed to target approximately 100bp upstream of the G4C2 locus. In order to edit the mutation by HDR with a lower chance of off-target effects, gRNA1 and gRNA2 were chosen in very close proximity to each other. Oligonucleotides for each gRNA were annealed, phosphorylated and ligated into a pX335 plasmid vector, which can express both gRNA and Cas9n protein at the same time. Figure 4.1 summarizes the steps applied for constructing and evaluating the gRNAs generation.

Sanger sequencing was performed to confirm the accurate ligation of the gRNA1 and gRNA2 sequences into the plasmid vector (Figure 4.2). The specificity and cleavage efficiency of these gRNAs were assessed by transfecting them into HEK293T cells and performing cleavage assays for detecting heteroduplexes formation**.**

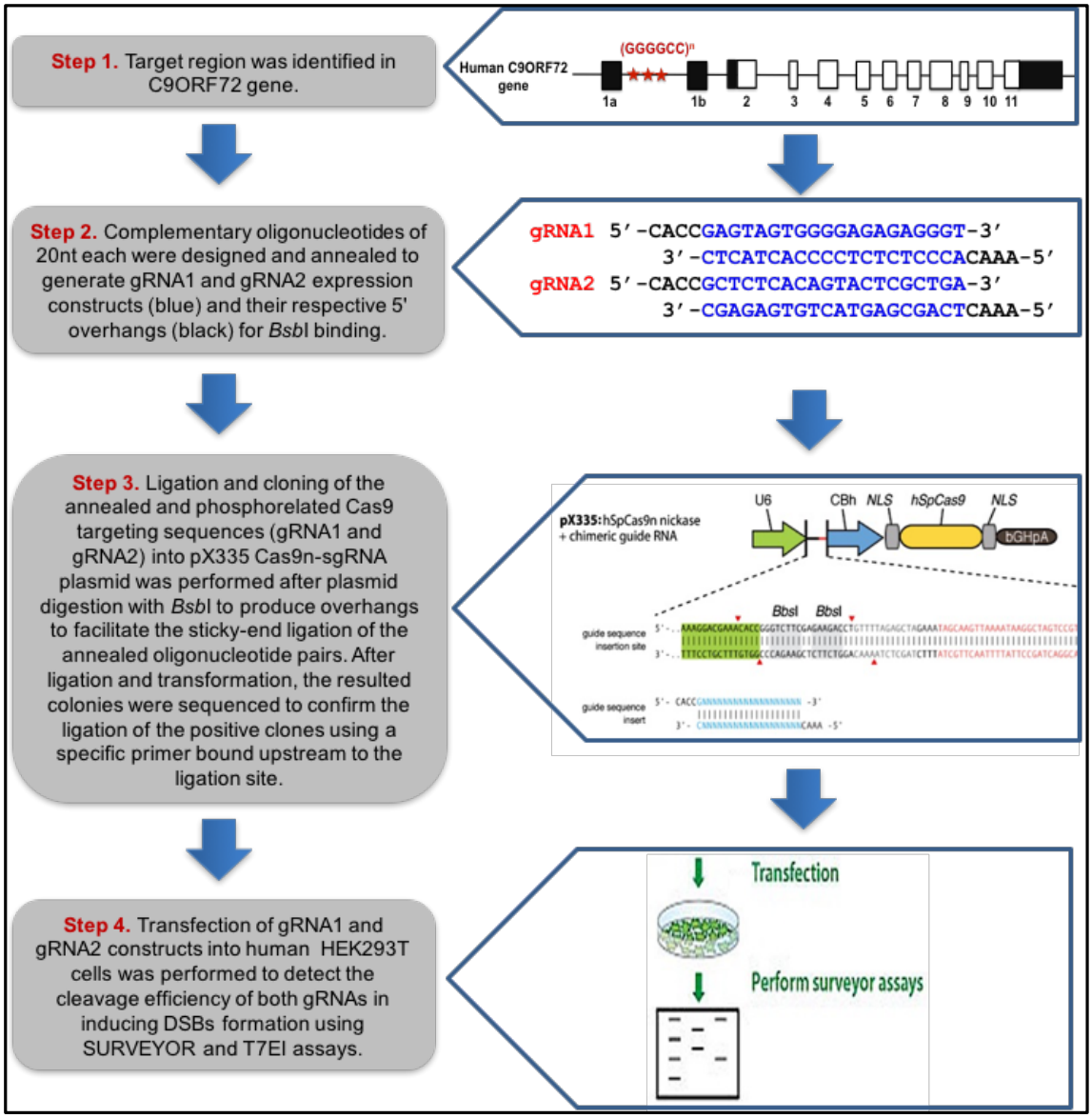

Figure 4.1 A schematic diagram describing the steps required to design and construct Cas9n-sgRNA plasmids. First, the target region was identified and the primers for gRNAs construction were designed, annealed and phosphorylated and then ligated into a pX335 Cas9n expressing vector. After that, gRNAs cleavage efficiency was assessed by using T7E1 and SURVEYOR cleavage enzymes.

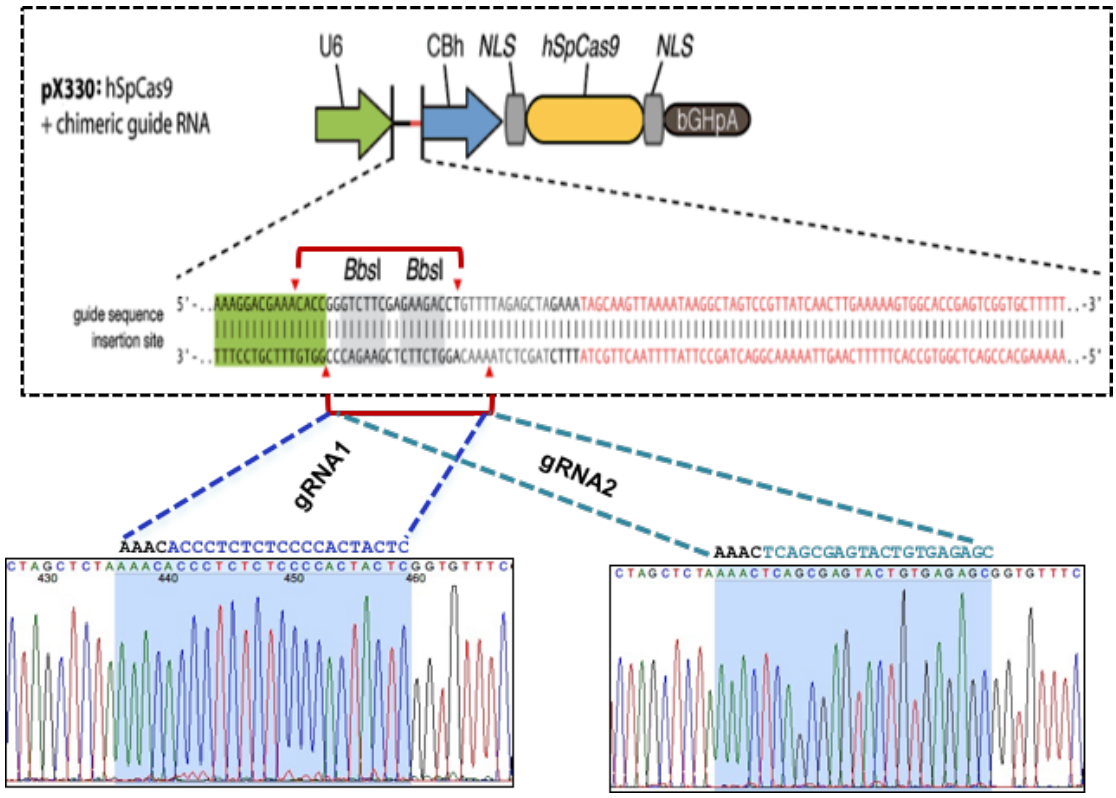

Figure 4.2 Precise ligation and cloning of gRNA1 and gRNA2 into pX335 plasmid as shown by direct sequencing. The guide sequence was inserted between two BbsI sites, after phosphorylation and annealing of the designed oligonucleotides. The expression of the chimeric gRNA is driven by the U6 promoter while the Cas9 is expressed using the CBh promoter. Adapted from (Cong et al. 2013)

4.1.2. Double nicking CRISPR/Cas9 system generates detectable levels of DSBs

To reduce off-target effects, double-nicking CRISPR with Cas9-nickase was applied using Cas9D10A nickase to make SSB in the complementary strand of the sgRNA sequence. By using a pair of sgRNAs, DSBs can be formed by nicking the opposite strand simultaneously. In this case, the chance of off-target effects is reduced due to the lower possibility of finding two gRNA sequences in close proximity to each other in the human genome. In Figure 4.3.A, gRNAs location on the sequence (blue) and their PAM region sites (red) are identified, Cas9 cleavage sites are indicated by the small red arrows, three nucleotides upstream the PAM regions**.** The DSB can be either repaired by the error-prone endogenous NHEJ repair mechanism by rejoining of the cleaved strands at the

cleavage site, which results in random insertion/deletion mutations at the cleavage site, or it can be repaired by using HDR and a donor template carrying the desired modification (Figure 4.3.B).

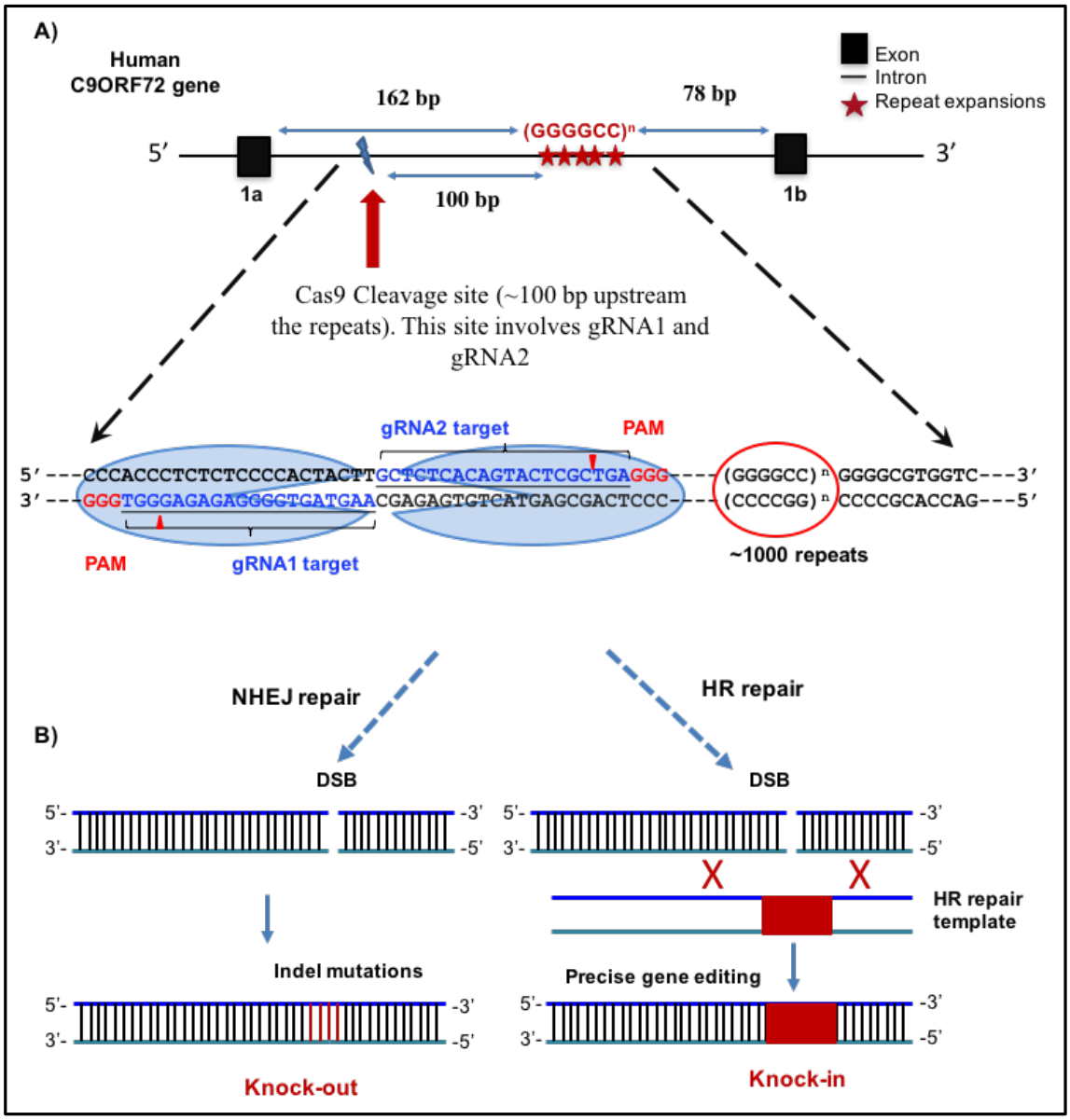

Figure 4. 3. Schematic illustration of the double nicking CRISPR/Cas9n strategy for targeting the HRE region. A) The double nicking CRISPR/Cas9 system using gRNA1 and gRNA2 was applied to remove a massive repeat expansion sequence (>1000 repeats) and replacing it with the wild-type sequence (two repeats), using a plasmid DNA donor template to drive HDR. By using a pair of sgRNAs, DSBs can be formed by nicking the opposite strand simultaneously. gRNAs location on the sequence (blue) and their PAM region sites (red) are identified. Cas9 cleavage sites are indicated by the small red arrows, three nucleotides upstream the PAM regions. **B)** DSB can be repaired by NHEJ endogenous repair mechanism or it can be repaired by using HDR and a donor template carrying the desired modification.

After transfection of HEK293T cells with Lipofectamine 2000 and either individual gRNA1 or gRNA2 or both of gRNAs simultaneously, the cleavage efficiency of the gRNAs was evaluated using T7 endonuclease I (T7EI) or SURVEYOR (CelI) nuclease (Transgenomic) cleavage assays to detect the heteroduplex DNA formation resulted from DSB. Genomic DNA of the transfected cells was extracted 72h after transfection and a 900 bp region around the cleavage site was amplified by PCR then denatured and reannealed to form DNA heteroxuplexes. The reannealed PCR products were then cleaved by CelI or T7E1 enzymes to detect the heteroduplex mismatch DNA formation and the cleavage efficiency was determined by agarose gel electrophoresis. GFP was transfected separately to assess transfection efficiency, which showed high percentage of successfully transfected cells. A schematic outline of the mismatch cleavage assays used for validation of gRNAs cleavage efficiency is summarised in Figure 4.4. The primer set used to amplify the target region was identified on the figure as Clvg-F and Clvg-R as indicated by the blue arrows (Figure 4.4).

Visualisation of the PCR products after performing T7E1 and Cel1 assays was carried out on agarose gel and showed positive cleavage results by producing a detectable level of DSBs when using gRNA1 and gRNA2 (Figure 4.5). The upper band of each lane represents the original PCR products from uncleaved DNA, while the lower digested bands were resulted from DSB cleavage and the formation of heteroduplexes (as indicated by the red stars) (Figure 4.5). In both images, cells transfected with gRNA1 and gRNA2 simultaneously produced a detectable level of heteroduplex DNA. These results demonstrate that both gRNAs were able to introduce DSBs when applied simultaneously, ensuring that they can be used in the subsequent experiments of nucleofecting iPSCs for editing the HRE mutation. No heteroduplexes were observed in cells treated with any of

the gRNAs individually. Untransfected cells were used as a negative control (Figure 4.5).

Another set of gRNAs (gRNA3 and gRNA4) was also designed and evaluated using the same cleavage assays. However, gRNA1 and gRNA2 showed more obvious cleavage results after digestion with T7E1 and Cel1enzymes than gRNA3/gRNA4 combination in Hek293T cells.

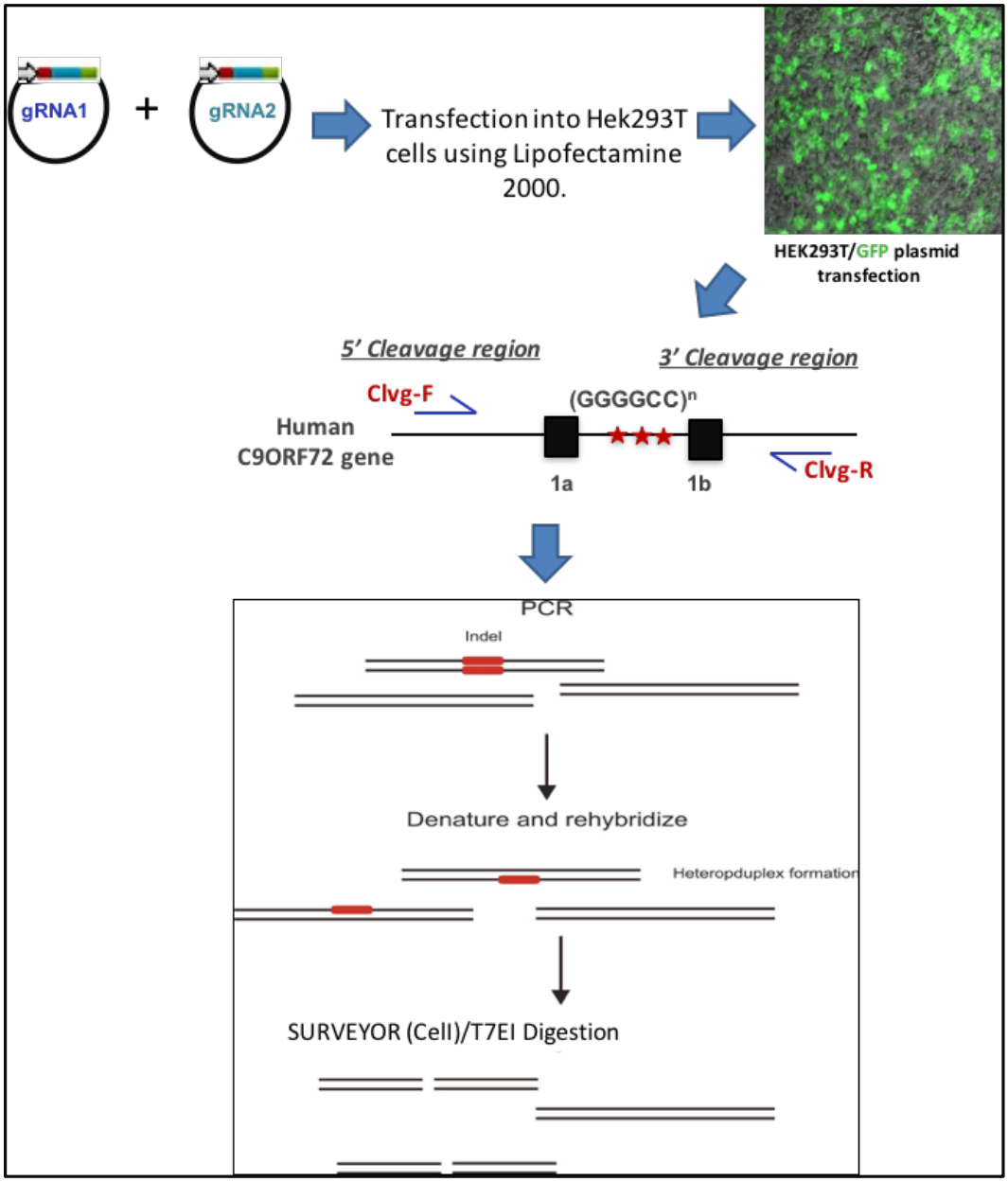

Figure 4. 4. Schematic outline of the mismatch cleavage assays used for validation of gRNAs. First, HEK293T cells were transfected with gRNA1 and gRNA2 and genomic DNA of the transfected cells was extracted then the region around the cleavage site was amplified by PCR and analysed by CelI (SURVEYOR) or T7E1 enzymes.

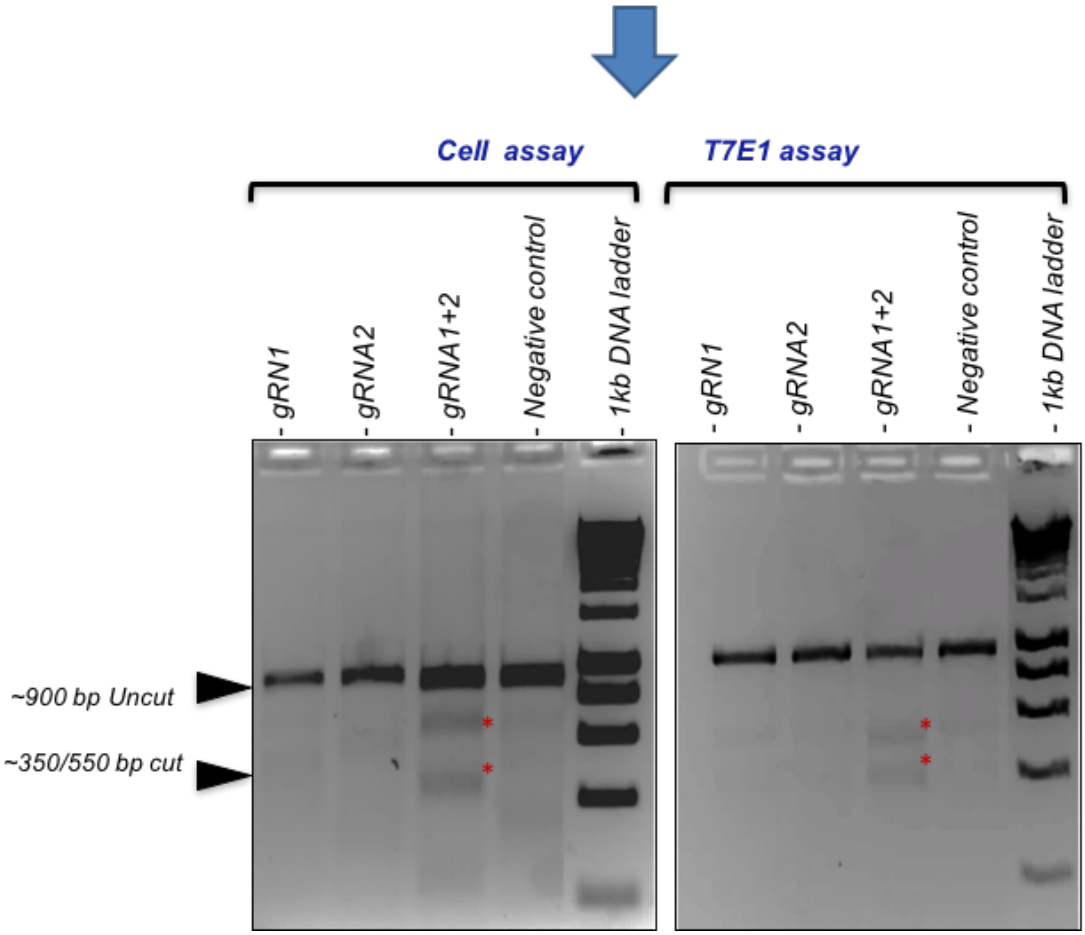

Figure 4. 5 Evaluation of the gRNAs cleavage efficiency by agarose gel electrophoresis. T7E1 and Cel1 assays showed positive cleavage results on PCR products, indicating that both gRNAs led to detectable level of DSBs (red stars). No heteroduplexes were observed in cells treated with any of the gRNAs individually. Untransfected cells were used as a negative control.

4.1.3. dsDNA donor template assembly

In this project, a dsDNA donor template was designed to replace the mutated G4C2 sequence with two repeats of the wild-type allele. Two long homology arms of ~1.8kb each, flanking each side of the target site, were amplified from the wild-type allele of the same patient C9-02 genomic DNA and ligated into pGEMT-easy plasmid vector. The NsiI site was then introduced by site-directed mutagenesis (SDM) of ~100bp upstream the HRE locus at gRNA2 PAM region site to ensure that the Cas9 will not cleave the repair template. Another Nsi1 site was also introduced ~105bp downstream of the repeats. Both restriction sites were used as diagnostic markers in the screening of successfully targeted clones for HDR. The schematic process of constructing the donor template and the sequencing results are described in Figure 4.6.

In order to reduce the screening of a large cell population, an EF1a-PGK-Puro-Tk selection cassette was inserted ~53bp upstream of the repeat region. The cassette was driven by EF1a promoter and flanked by two LoxP sites. This would allow the subsequent removal of the cassette by Cre-loxP system after successful selection of the targeted clones. Puromycin selection was started 48h after transfection to kill iPSCs that do not express the puro gene, enabling the selection of cells which transiently express the vector. The full cassette was amplified from pENTR-Ef1a-Puro/Tk plasmid and then the PCR product was ligated to the linearized pGEMT-easy donor plasmid using Gibson assembly master mix. Gibson assembly is considered an effective and efficient way of assembling large dsDNA donor constructs in one step, without the need of performing multiple cloning and ligation steps. The assembled construct was then transformed into competent cells and sequence integrity was confirmed by Sanger sequencing (Figure 4.6). The completed DNA donor template sequence is listed in Appendix IV.

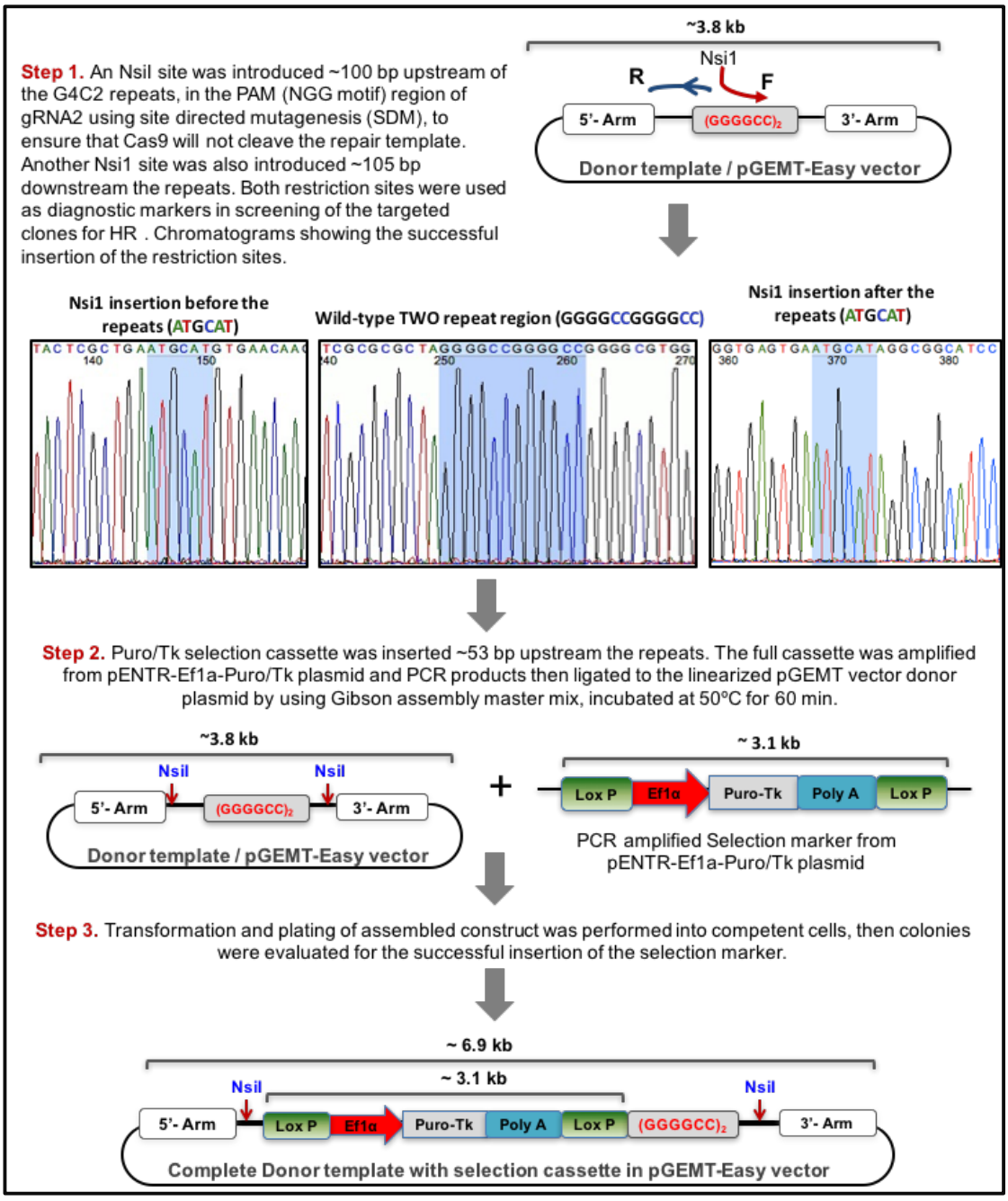

Figure 4.6 Schematic outline of the donor template construction. Long homology arms flanking each side of the target site were amplified from the wild-type allele of the same patient, then a selection/counterselection cassette was inserted upstream the repeats. The donor plasmid comprises two long homology arms (~2kb 5'homology arm and ~1.8kb 3'-homology arm carrying the wild-type G4C2 repeat sequence (two G4C2 repeats)). The Puro/Tk selection cassette was inserted directly upstream the repeats location and expressed under an EF1a promoter and flanked by two loxP sites to facilitate its removal after successful selection process.

4.1.4. Targeted Correction of the G4C2 mutation in iPSCs.

Following successful gRNAs validation and donor template construction, a nucleofection-based protocol was applied on C9-02-02 iPSC culture, using the same double-nicking gene editing strategy and HDR system (Figure 4.7). The Neon transfection system (Life Technologies) was applied on patient iPSCs using 100µl single Nucleocuvettes to increase the number of nucleofected cells and increase the chance of getting the successfully targeted clones. iPSCs were nucleofected with both gRNAs and the donor template and then plated on a drug-resistant feeder layer (DR4 cells). The nucleofection efficiency was also improved by incubating iPSCs in culture media containing 10µM Y-27632 at least 1h before dissociating the cells, as well as after the nucleofection procedure.

Three days after seeding nucleofected single cells on a DR4 layer, puromycin treatment was started and continued until the iPSC colonies became large enough to be picked up individually. Individual colonies were manually picked and transferred to individual 24-well plates for expansion and genomic DNA analysis and screening using different molecular assays (Figure 4.8). A control GFP expressing plasmid was used as a negative control for verification of successful puromycin selection. No colonies were observed after puromycin treatment in the GFP control plate.

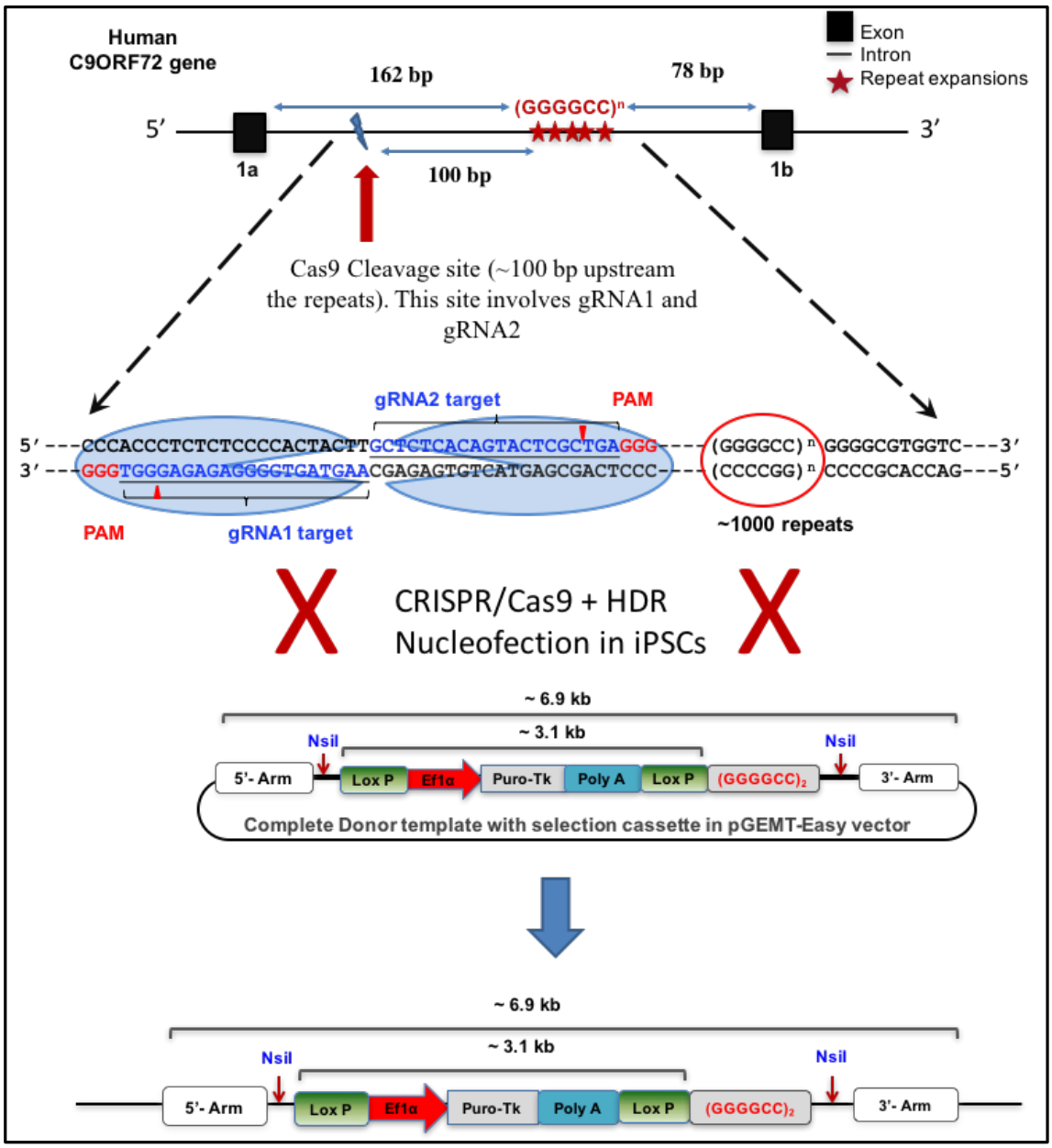

Figure 4.7 CRISPR/Cas9-HDR strategy for correcting the HRE mutation in patient iPSCs. A homologous donor template containing a Puro/Tk cassette was used along with gRNA1 and gRNA2 in iPSCs transfection using Neon Neucleofection system.

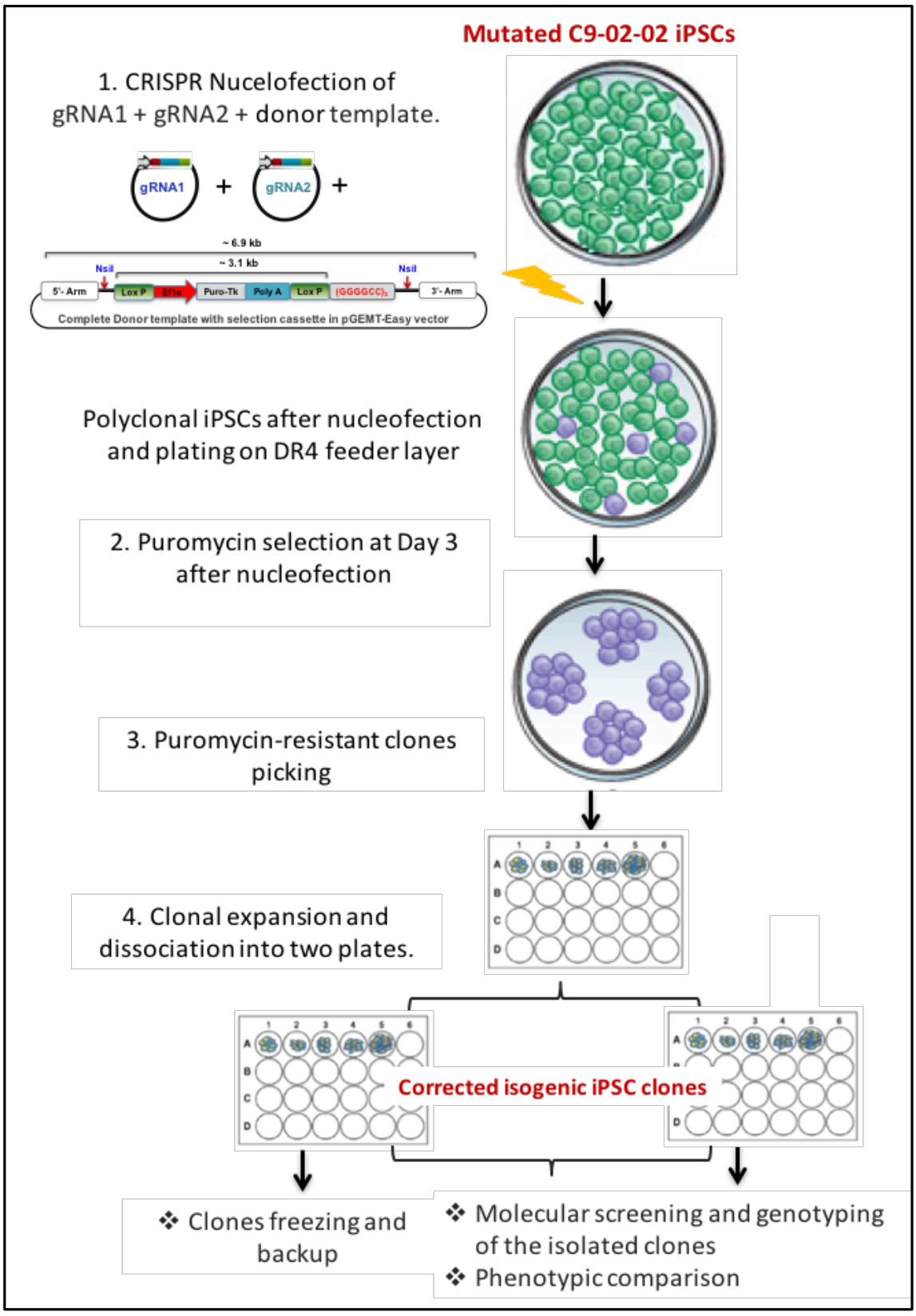

Figure 4.8 Overview of the selection and expansion procedure of puromycin resistant clones. iPSCs were nucleofected with the gRNAs and donor template, then plated on a drug-resistant (DR4) feeder layer for puromycin selection and isolation of the resistant clones. The drug-resistant clones were assessed by several molecular screening assays, to check for the absence of HRE mutation and the successful replacement of the wild-type repeat region by HDR.

4.1.5. Molecular screening of selected clones

4.1.5.1. Repeat-primed PCR-based screening assay (RP-PCR)

Effective gene targeting in 100 drug-resistant clones was first identified by RP-PCR. Two RP-PCRs (RP-PCR1 and RP-PCR2) were performed to amplify the G4C2 repeat region in 5' and 3' directions, to detect the presence or absence of the repeats. Of all the 100 clones analyzed, 24 clones showed no repeat expansion on their electropherogram in both PCRs, which confirmed the elimination of the expanded hexanucleotide region. Figure 4.9.A. shows a schematic representation of RP-PCR1 strategy and Figure 4.10.A shows RP-PCR2. The red stars indicate the repeat expansion, Repeat-primed, anchor and FAM-fluorescently labelled reverse primer positions are identified on the figure. The RP-PCR generates amplicons of varying size, based on the hybridisation of the repeat primer to different parts of the expansion region. Lower panel B of both figures 4.9 and 4.10 shows capillary-based sequence traces of the RP-PCR. WT-01 healthy control had a normal electropherogram with six repeats detected. Ed-02 edited clone showed the absence of the abnormal repeat expansion and the replacement with the normal repeat size which is two repeats, whereas parental uncorrected C9-02-02 iPSCs showed a typical saw-tooth pattern electropherogram with more than 50 G4C2 repeats detected.

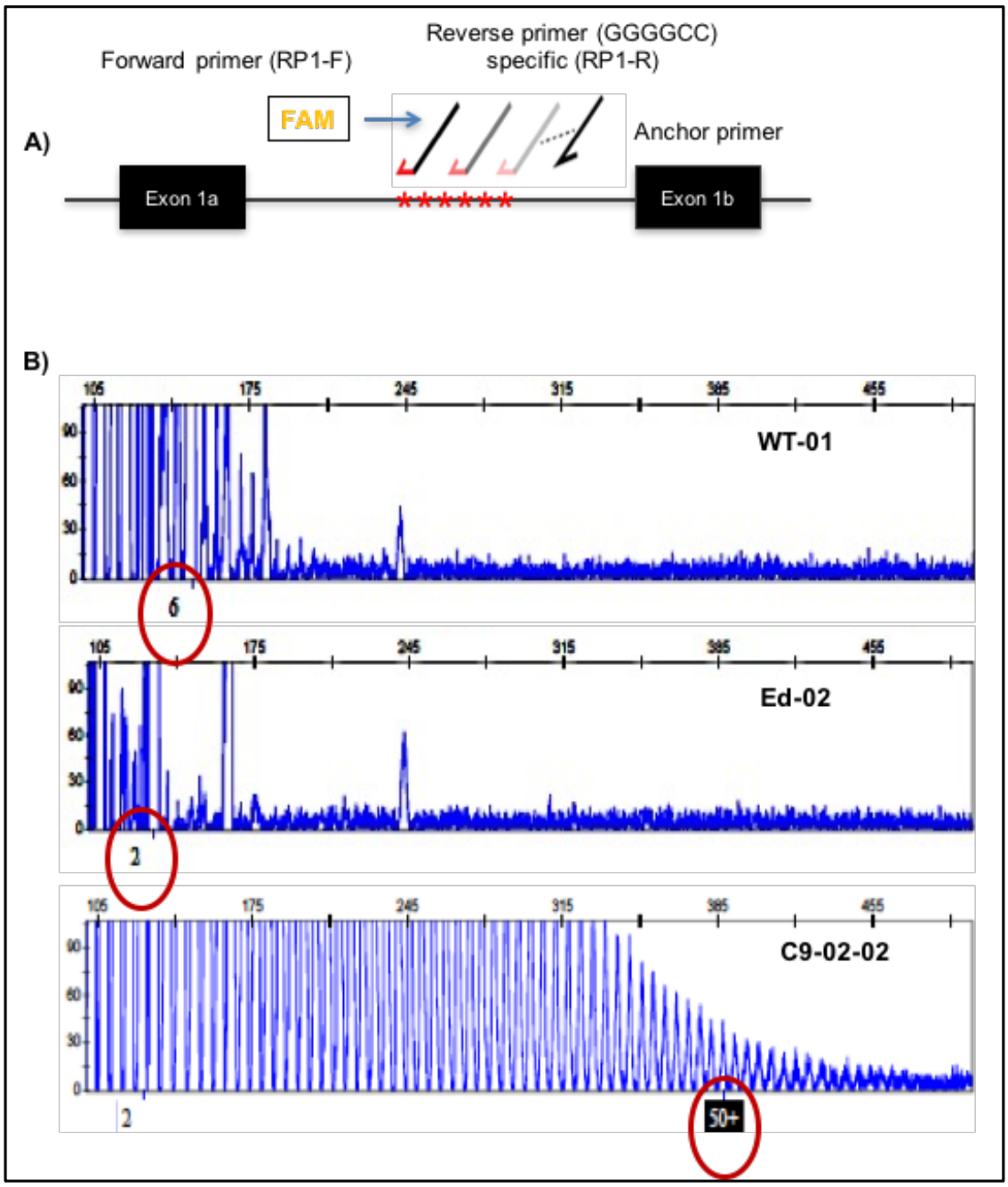

Figure 4.9 Genotyping of the resistant clones by RP-PCR1. A) Schematic representation showing RP-PCR amplification strategy and the primers position. The red stars indicate the repeat expansion. The position of repeat-primed, anchor and FAM-fluorescently labelled reverse primers are shown in the diagram. **B)** Electropherogram for WT-01 healthy control showed normal electropherogram and six repeats were detected, Ed-02 edited clone also showed normal repeat size (two repeats), and the parental C9-02-02 iPSCs showed typical saw-tooth pattern electropherogram with more than 50 G4C2 repeats detected.

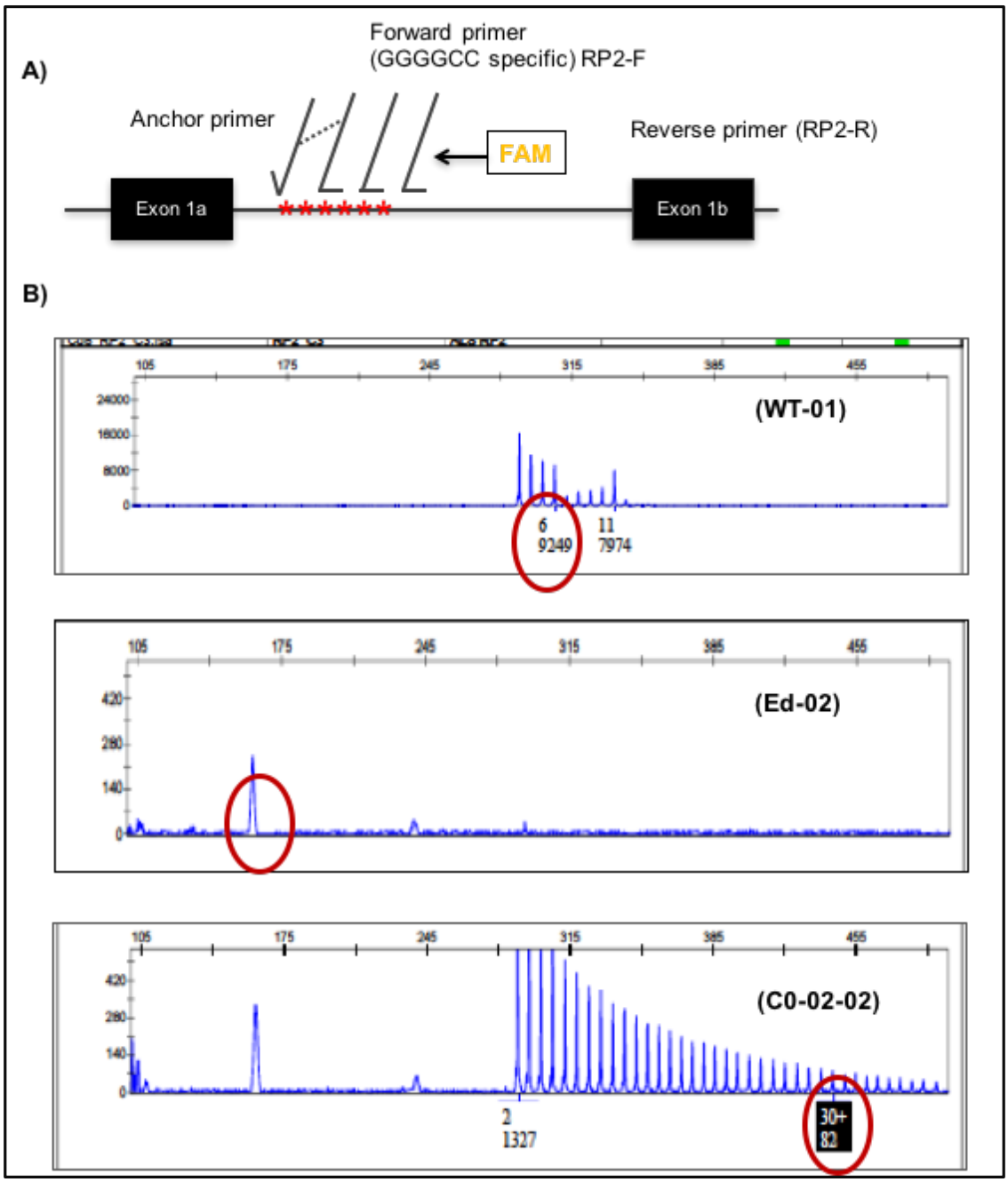

Figure 4.10 Genotyping of the resistant clones by RP-PCR2. A) A similar RP-PCR principle was applied with different primer sequences for amplification of the end of the repeat region. **B)** Electropherogram for WT-01 healthy control, Ed-02 edited clone, and the parental uncorrected C9-02-02 iPSCs showed similar results to RP-PCR1 assay.

4.1.5.2. Sequence confirmation of the intended modification for the verification of successful HDR.

After RP-PCR screening, fourteen repeat-free clones were further analysed by direct sequencing of the the repeat region to confirm the successful insertion of the wild-type repeat size by HDR.

The first image in Figure 4.11.A, shows the wild-type allele repeat size in C9-02-02 patient parental iPSCs (two repeats). The second image of the same panel A shows the mono-allelic targeted edited clone with HDR exclusively in the mutant allele (Ed-02). The third image shows a bi-allelic targeted clone with HDR in both alleles (Ed-03). To assess the cleavage of the CRISPR/Cas9 system and detect any indels at the target site, sequencing of the region around the cleavage site was performed and demonstrated successful insertion of Nsi1 site upstream the repeat (Figure 4.11.B). These sequencing results confirmed the presence of the normal repeat size derived from HDR and demonstrated the potential of the CRISPR/Cas9 system for genome editing of HRE mutation on iPSCs.

In summary, two of the edited lines showed monoallelic HDR targeting (Heterozygous, Ed-01 and Ed-02), and three lines showed biallelic HDR targeting (Homozygous, Ed-03, Ed-04 and Ed-05). No indel mutations were observed in any of the five isogenic lines at the cleavage site. While the remaining clones of the 24 puromycin-resistant clones showed off-target mutations at the cleavage site but]with no HDR. Molecular screening results for all clones are summarized in Table 4.1.

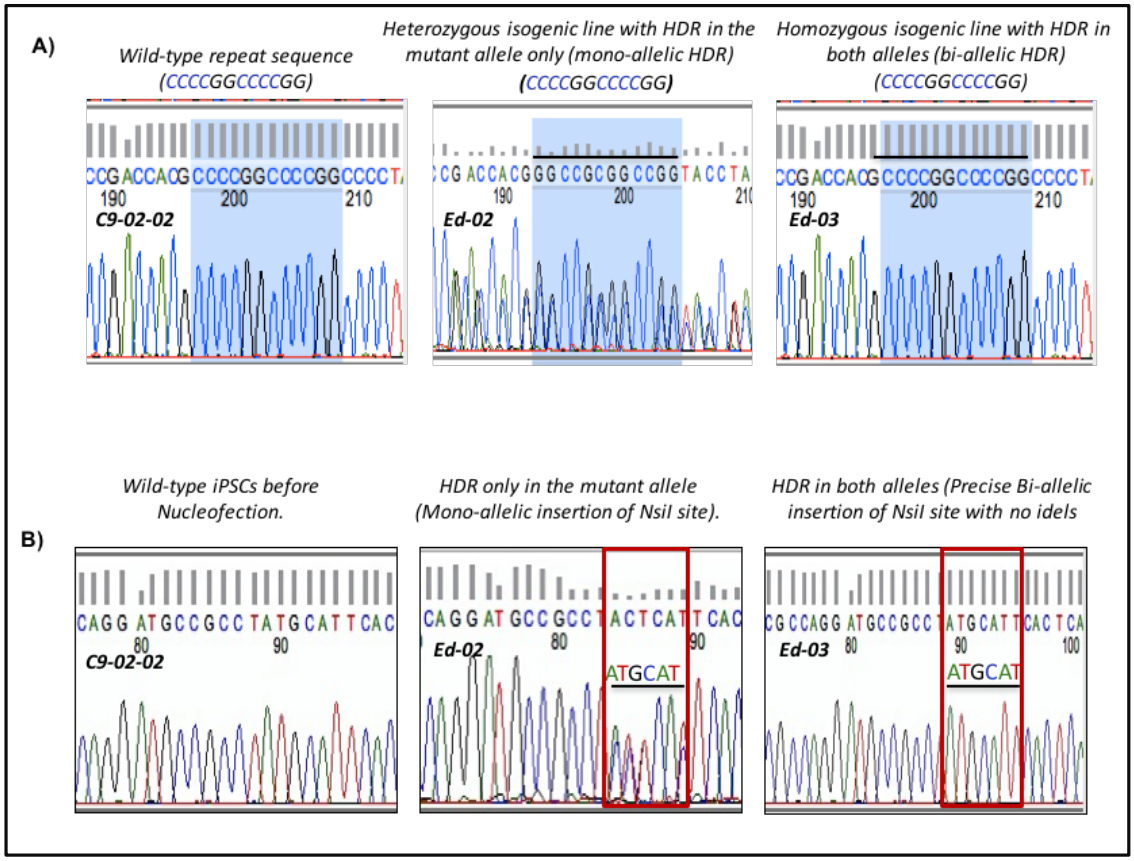

Figure 4.11 Sequence confirmation of the intended modification. A) After RP-PCR, repeats-free clones were sequenced around the repeat region to confirm the HDR and the presence of the wild-type repeat size. The first image shows the wild-type repeats in C9-02-02 patient parental iPSCs (two repeats), the second image shows a mono-allelic targeted clone with HDR in the mutant allele (Ed-02). The third image shows a bi-allelic targeted clone with HDR in both alleles (Ed-03). **B)** Sequencing of the region upstream the inserted Nsi1 site was performed to confirm the HDR results observed in part A of this figure. Mono-allelic insertion of Nsi1 site was observed in Ed-02 clone and bi-allelic Nsi1 insertion observed in Ed-03 clone.

4.1.5.3. Site-specific integration of the donor template to 5' and 3' junctions

The successful targeting results obtained for the isolated edited lines were further confirmed by PCR-RFLP amplification of upstream and downstream Nsi1 sites and with outside the homology arms. This was performed using one primer within the selection cassette and one outside of either 5' or 3' homology arms (Figure 4.12.A-C). Several primers were designed and the PCR reactions were identified as 5'In-out or 3' In-out PCRs (Figure 4.12.A). The 5'Out-F/EF1a-R primer combination was used to amplify the 5' region, and Puro-F/3'Out-R was used to amplify the 3' region. The PCR products from

both PCRs of Ed-02 and Ed-03 mono-allelic targeted clones were digested with *Nsi1* enzyme and visualised using agarose gel electrophoresis. Positive digestion was observed in both edited lines by digesting the original 3423bp band into 2012bp and 813bp bands. These results indicate the successful integration of the donor template into the intended site (Figure 4.12.B&C). Direct sequencing of the same PCR product was also performed to confirm the specific integration of the donor template into the 5' or 3' junctions (Figure 4.12.B&C). The sequencing results showed the precise integration of the donor template in both edited clones.

Finally and before Puro/Tk cassette excision, digital droplet PCR was applied to measure the copy number of puromycin gene after gene correction. This was done to assess if there is any random integration (RI) resulted during gene targeting process using Cas9 cassette or puro/Tk cassettes transfection. No RI events were detected in any of the mono-allelic targeted clones, which indicate successful HDR without random integration.

Table 4. 1 Molecular screening results of 24 puromycin resistant iPSC clones after gene correction using CRISPR/Cas9-mediated HDR.

Clone #	5'In-Out PCR	RFLP	3'In-Out PCR	RFLP Result	Sequencing	RP1	RP2
1	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
2	+ve	+ve	+ve	+ve	HDR+	NRE	NRE
3	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
4	+ve	+ve	+ve	+ve	HDR+	NRE	NRE
5	+ve	+ve	+ve	+ve	HDR+	NRE	NRE
6	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
7	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
8	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
9	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
10	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
19	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
19	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
13	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
14	+ve	+ve	+ve	+ve	HDR+	NRE	NRE
15	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
16	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
17	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
18	+ve	+ve	+ve	+ve	HDR+	NRE	NRE
19	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
20	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
21	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
23	-ve	-ve	-ve	-ve	No HDR	NRE	NRE
24	-ve	-ve	-ve	-ve	No HDR	NRE	NRE

Abbreviations: -ve: Negative result, +ve: positive result, HDR: Positive homology-directed repair, HDR-: Negative homology-directed repair, NRE: No Repeat expansion.

Note: Blue highlighting reveals positive clones with NRE and HDR+.

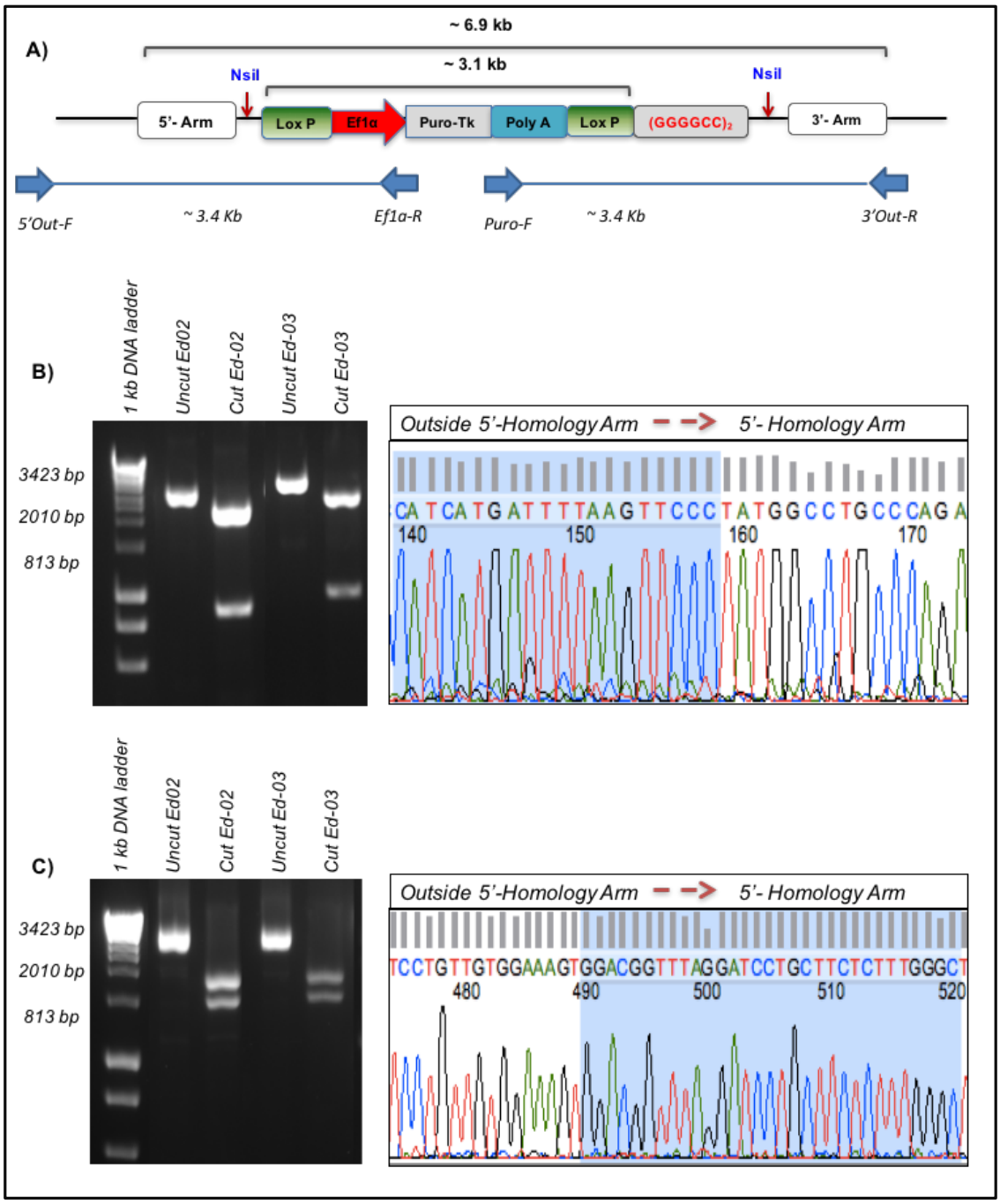

Figure 4. 12 Site-specific integration of the donor template at the 5' and 3' junctions. A) Primers were designed to amplify inside the cassette and outside the 5' or 3' homology arms in PCR reactions identified as 5' In-out or 3' In-out PCRs. **B&C)** PCR products of Ed-02 and Ed-03 clones were digested with *Nsi1* enzyme and visualised using agarose gel electrophoresis. Positive digestion was observed in both clones, which indicates successful integration of the donor template. Sequencing of the same PCR products also confirmed the precise integration of donor template with outside of 5' or 3' homology arms.

4.1.6. Excision of the selection cassette using the Cre-LoxP system.

In order to achieve genomic integrity in the corrected clones, the LoxP-flanked Puro/Tk gene cassette was excised by the Cre-LoxP system. This system leaves a 34bp loxP sequence at the site of excision, which has been used in screening of the cassette free clones (Figure 4.13.A). Ed-01 clone was excluded from this project because it showed a 5.6Mb duplication on chromosome 6 after karyotype analysis. Cre/LoxP mediated excision was applied on Ed-02 and Ed-03 to isolate multiple isogenic lines free of the selection marker. This has been achieved by nucleufecting iPSCs with Cre mRNA, then negative selection of the clones lacking *TK* gene by Ganciclovir treatment. Accordingly, several clones were picked up, expanded and maintained in mTeSR1 medium.

Clones were then evaluated for the absence of Puro/Tk marker by PCR amplification inside puromycin gene and outside the 5' junction. Most of the clones showed successful removal of the cassette as demonstrated by agrose gel electrophoresis of Ed-02-01 and Ed-02-02 mono-allalic clones and Ed-03-02 bi-allelic clone (Figure 4.13.B). The original ED-02 clone carrying the selection cassette was used as a positive control and showed a ~4kb PCR band.

Successful removal of the Puro/Tk cassette was also confirmed by performing a PCR amplification and direct sequencing of the region around the LoxP site involving the repeat region. PCR amplification produced a 75 bp band for the wild-type allele and 109 bp band for the targeted allele (including the 34 bp loxP site remained after Cre treatment) (Figure 4.13.C).

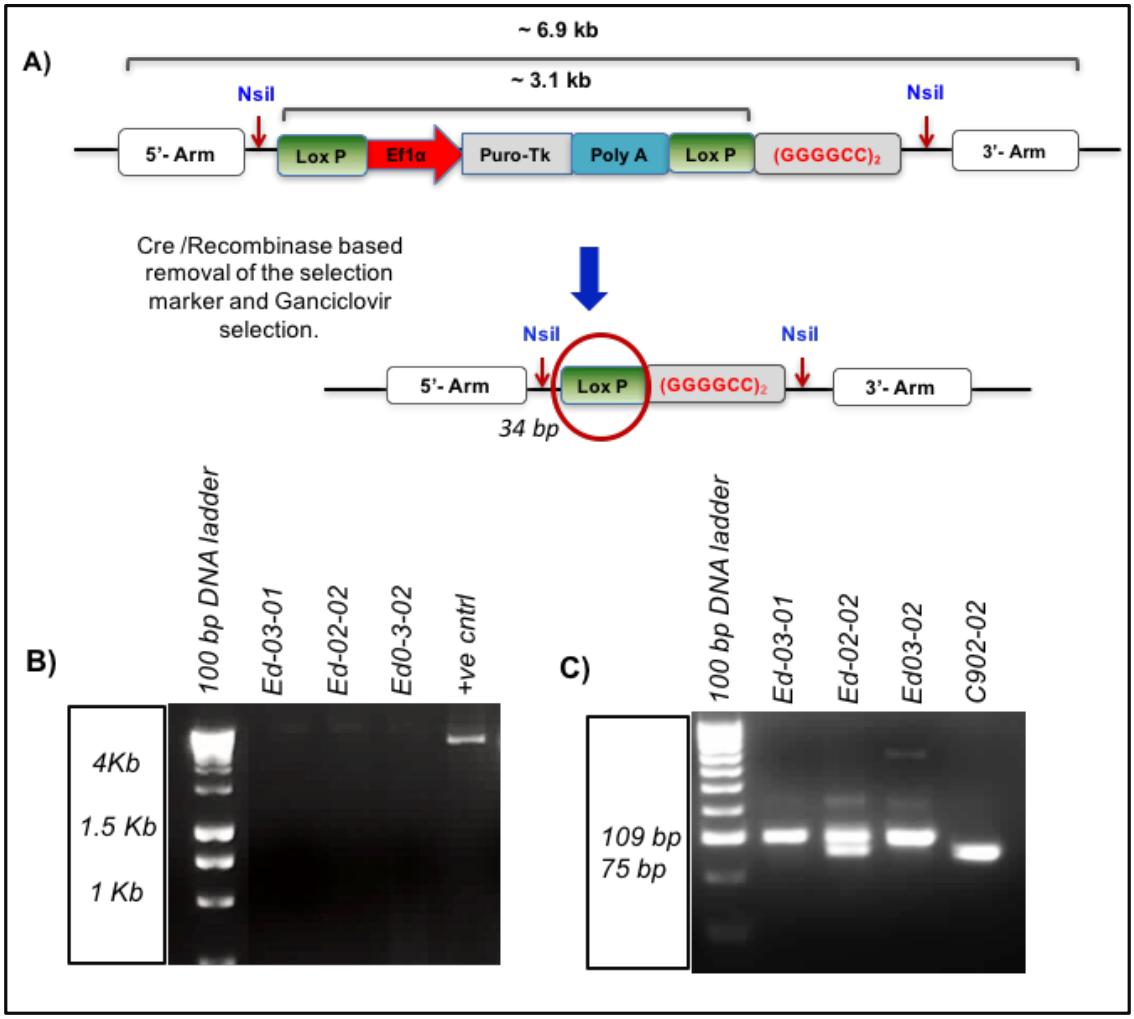

Figure 4.13 Confirmation of successful excision of the Puro/Tk cassette. A) Cre-LoxP-system was applied to remove the Puro/Tk cassette **B)** PCR product shows the deletion of the selection marker by the lack of amplification inside the puro/Tk and outside the 5' junction. The positive control showed a ~4kb band. **C)** Successful PCR results showing a 75bp band for wild-type allele and a 109bp band for the targeted allele with 34bp loxP site remained after Cre treatment of edited clones.

Sequence confirmation was performed using the same PCR products and also confirmed the successful removal of the cassette as indicated by the presence of the *Nsi*I and LoxP sites in Ed-02-01 and Ed-02-02 genomic DNA samples (Figure 4.14.A).

Finally, Southern blotting analysis using a $GGGGCC^5$ DIG-labelled repeat specific probe was carried out and revealed the absence of the repeat expansion in the edited isonegic cell lines, while a band of ~6kb was observed in the original parental C9-02-02 cells (Figure 4.14.B).

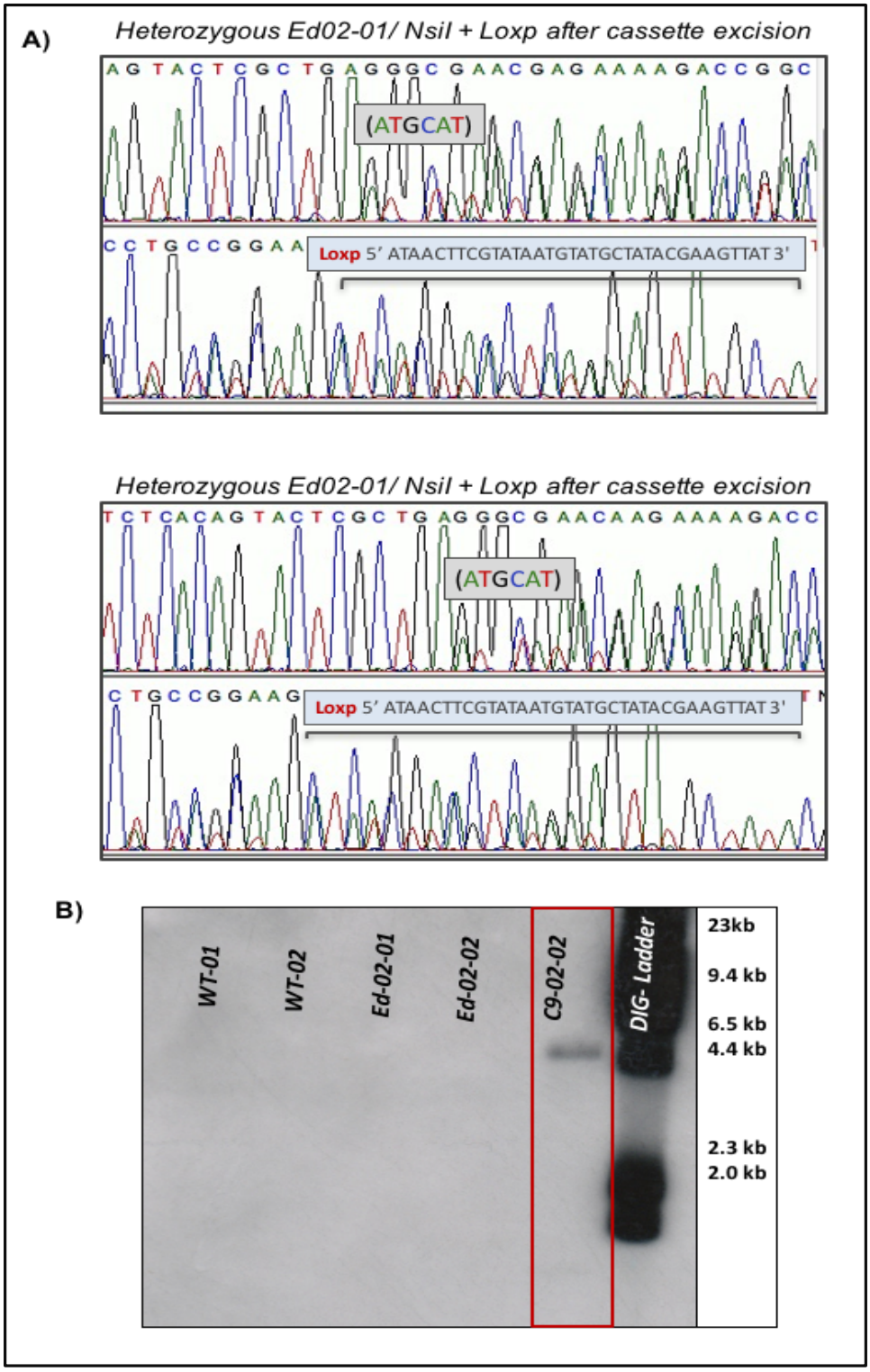

Figure 4.14 Confirmation of successful excision of the Puro/Tk cassette by direct sequencing and southern blotting analysis for final confirmation of the excision of HRE. A) Sequence analysis showed successful removal of Puro/Tk cassette as indicated by the presence of NsiI and LoxP sites in Ed-02-01 and Ed-02-02 clones. **B)** Southern blotting showed repeat expansion in C9-02-02 cells but no bands were detected in the edited isonegic cell lines.

4.1.7. Corrected iPSCs retain pluripotentcy characteristics and normal karyotyping profiles

After successful excision of the Puro/Tk cassette, two of the mono-allelic targeted clones (Ed-02-01 and Ed-02-02) were expanded and assessed for their ability to express the pluripotency markers; NANOG, OCT4 and TRA-1-60. Both edited lines retained their pluripotency as observed by their normal morphology and their ability to produce uniform immunostaining of pluripotency markers (Figure 4.15). This could suggest that CRISPR-mediated genome editing did not alter their pluripotent state. Karyotyping analysis of Ed-02-01 and Ed-02-02 clones also revealed that both lines have retained their normal chromosomal profile when compared to the parental fibroblast (C9-02) and iPSCs (C9-02-02) karyotyping profiles (Figure 4.16).

Taken together, these results showed that corrected iPSCs maintained their pluripotentcy characteristics and normal karyotypes after genetic engineering and they can be used in phenotyping experiments to assess the reversal of ALS disease-related phenotypes.

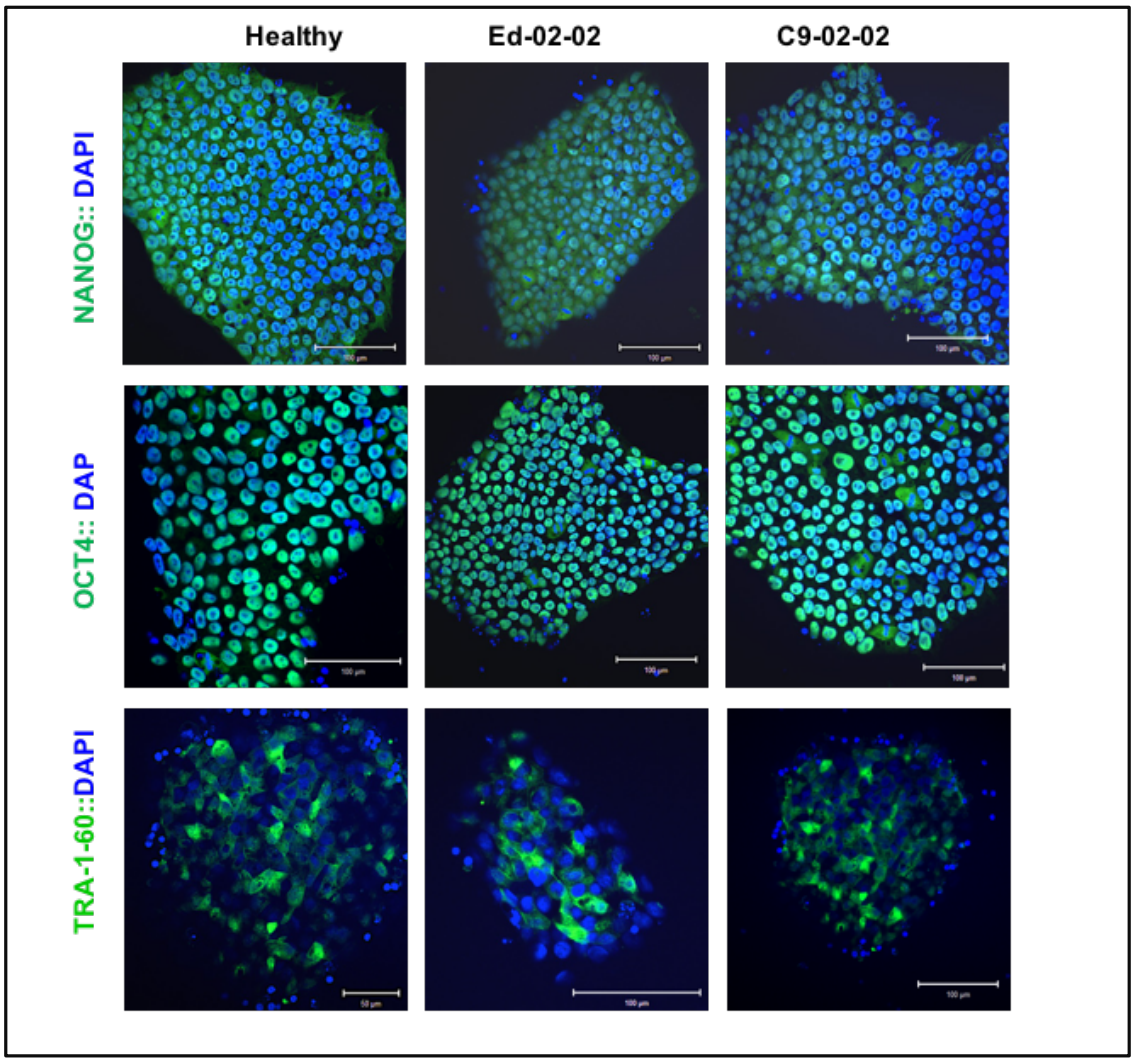

Figure 4.15 Edited iPSCs lines retain their pluripotency. Representative confocal images show that CRISPR/Cas9-edited Ed02-02 iPSCs retain normal pluripotency after gene editing experiments.

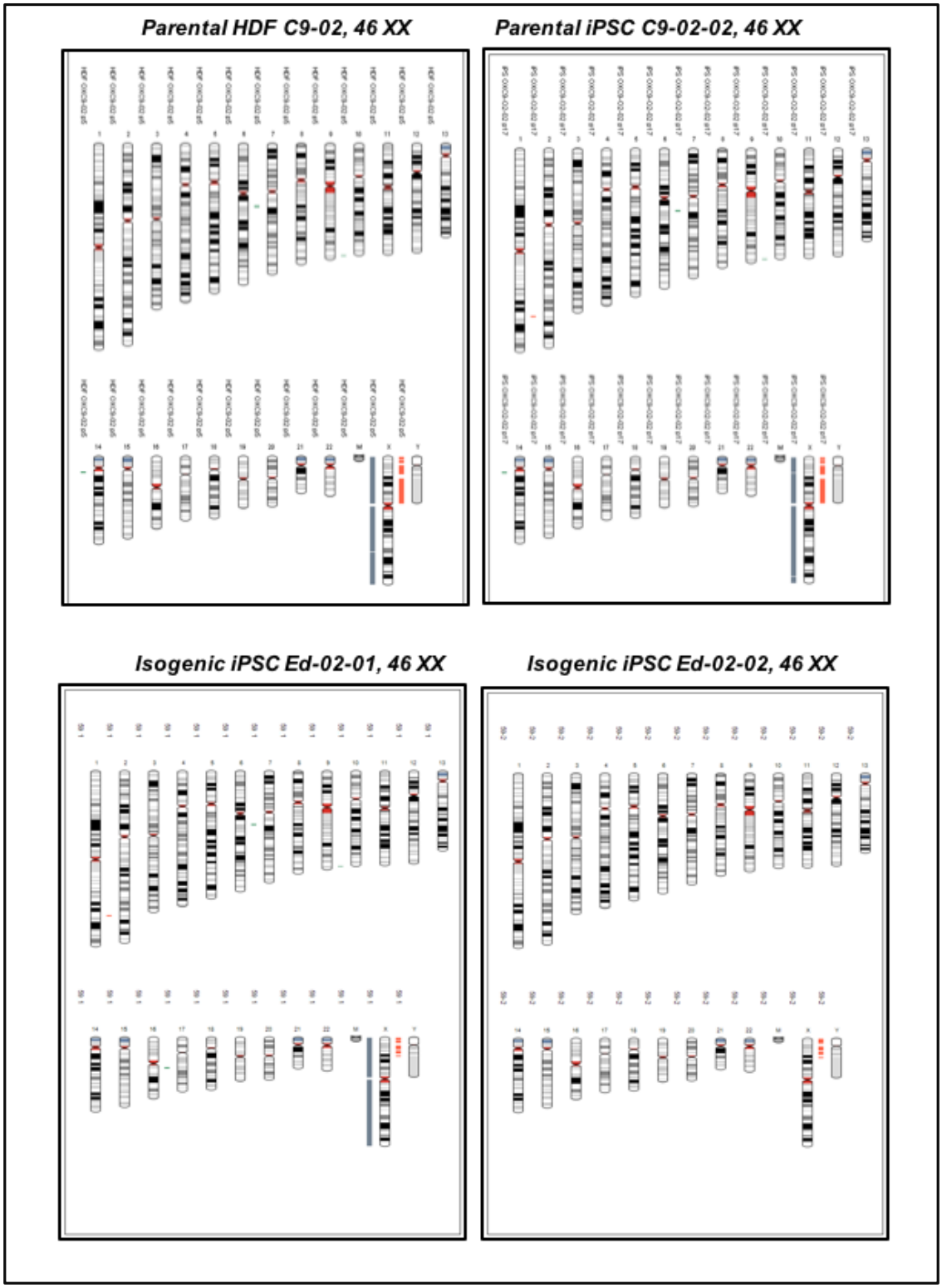

Figure 4. 16 Karyotyping results of the edited iPSC lines. Karyotyping analysis of Ed-02-01 and Ed-02-02 edited clones revealed that both lines retain their normal chromosomal profile when compared to the parental fibroblast (C9-02) and iPSC (C9-02-02) karyotyping results.

Summary

Combining CRISPR/Cas9 technology and iPSCs derived from G4C2 ALS/FTD patient has resulted in the efficient generation of genetically corrected iPSCs. This work demonstrates the successful generation of isogenic lines from ALS/FTD patient iPSC carrying a massive repeat region with precise correction. These isogenic lines can be used to study the reversal of HRE-associated disease phenotypes by differentiating them into MNs.

5. CHAPTER FIVE

Characterization of MNs differentiated from gene-edited iPSC lines

5.1. Introduction

iPSCs share similar characteristics to ESCs in that they can proliferate indefinitely in culture and differentiate into any cell types of the three germinal layers; ectoderm, endoderm, and mesoderm (Takahashi et al. 2007). Patient-derived iPSCs have been generated from somatic cells using several protocols to serve as invaluable resource of cells for translational medicine in applications such as disease modelling, drug screening and studying disease-related mechanisms.

Currently, iPSCs represent an ideal target for gene correction studies in genetic disorders by creating isogenic controls from patient cells. Isogenic controls would allow the direct comparison with the disease causing genetic variants by reducing genetic background variations that might result when using healthy control subjects. Normalization of cellular phenotypes after gene correction also provides a proof that these phenotypes are directly caused by the mutation under investigation. They also provide a potential source of autologous corrected cells for transplantation and future regenerative medicine studies.

Another advantage of iPSCs is the capacity to model human diseases of unknown etiology by differentiating them into cells of inaccessible origin to study molecular and functional disease pathways, such as neuronal cells in neurodegenerative diseases.

Motorneuron diseases (MNDs), including amyotrophic lateral sclerosis (ALS) and spinal muscular atrophy (SMA), are characterized by the progressive loss of MNs. The possibility of generating MNs from iPSCs provides an avenue for modelling ALS *in vitro*. Additionally, it can increase our understanding of the pathogenic mechanisms leading to MN degeneration, especially early disease pathways and to overcome the limitations of using human post-mortem CNS tissues, which reflects the end-stage disease.

Lower MNs in the spinal cord and brainstem that have the ability to project their axons to muscles to control the voluntary movement. Spinal motorneurons (sMN)s are located in the ventral horn of the spinal cord, and cranial motorneurons (cMNs) are located in the brainstem. These neurons are specified and patterned under the control of Wnt, fibroblast growth factor (FGF), retinoic acid (RA) and sonic hedgehog (SHH) (Maury et al. 2014). Several protocols have been developed to generate functional MNs from human iPSC lines (Sances et al. 2016). In most MN differentiation protocols, neural induction (NI) and neural progenitors (NPs) are produced through the use of N2 and B27 supplements, retinoic acid (RA), and Sonic SHH or SHH pathway activators (Sonic Agonist (SAG) and puromorphamine). RA plays an important role in the development and regeneration of nervous system cells (Maden 2007), by promoting the caudalisation and induction of differentiation into mature MNs. The presence of SHH activators promotes the patterning and ventralization of MN progenitors. Finally, neurotrophic factors are usually added to the medium at the last stage for further maturation of MNs and to promote cell survival.

In this chapter, after successful genotypic correction of iPSCs and characterization of the isogenic lines, edited cells were differentiated into MNs and characterised. The embryoid body (EB)-based differentiation protocol was successfully applied earlier in this thesis through neural induction of EBs followed by neural rosette formation and patterning then the production of neurospheres. Neurospheres were further induced to produce mature neurons using neurotrophic factors (Hu & Zhang 2009). Characterization of the generated MNs was performed by immunostaining of MN specific biomarkers. However, this EB-based differentiation protocol involves multiple steps and more than 5 weeks are required to produce mature MNs with limited efficiency (~15–20% of cells are MNs).

Recently, several groups reported rapid and efficient protocols for neural induction of iPSCs in adherent monolayer cultures (Maury et al. 2014; Du et al. 2015; Qu et al. 2014) to obtain a higher proportion of MNs from pluripotent stem cells in a shorter period of time. A monolayer differentiation protocol was adapted as previously described with minor modifications (Maury et al. 2014). The protocol was first applied by Dr Ruxandra Dafinca and Dr Jakub Scaber on ALS patient derived-iPSCs and healthy controls as described by Dafinca et al., 2016.

MNs can be assessed according to their neuronal morphology and their potential to express specific MN markers. MN cell exhibits unipolar morphology with a single axon extending from the soma, which then branches and forms network of dendrite trees. The efficiency of cell differentiation into MNs has been identified in several studies by immunocytochemical analysis of early and mature MN markers including PAX6, Nestin, Olig2, Islet-1, HB9, SMI-32, and ChAT. Immunostaining of dendrites can be achieved using antibodies specific to neuronal cyctoskeletal proteins, such as β-III tubulin (Tuj1; Tubb3) or using SMI antibodies to evaluate the phosphorylation status of neurofilaments and to allow the identification of the neuronal morphology typically present in MNs.

MNs can be distinguished from other neurons based on expression of MN specific transcription factors, including Islet1 and Hb9, as well as markers of more mature and cholinergic MN phenotype, such as ChAT (Wichterle et al. 2002; Son et al. 2011; Mackenzie W Amoroso et al. 2013; Maury et al. 2014; Boulting et al. 2011; Karumbayaram et al. 2009). HB9 is the most specific marker and is often used as the primary method of identification of specific MN identity (Vult von Steyern et al., 1999; Amoroso et al., 2013). The use of coimmunostaining of more than one marker is crucial for the successful identification of MNs because most of these markers are also expressed in other neuronal cell types (Sun et al., 2008).

5.2. Genome-edited iPSC-derived MNs show normal differentiation capacity.

In the previous chapter, edited iPSCs (Ed-02-01 and Ed-02-02) were thoroughly characterized, to confirm their pluripotency and their normal karyotype after gene editing experiment. Most of the experimental procedures in this chapter and Chapter 6 were carried out using iPSC clones derived from two unrelated wild-type healthy controls (WT-01 & WT-02), two edited clones (Ed-02-01; Ed-02-02) and two G4C2 mutation disease lines from the same patient (C9-02-02, the parental cells and C9-02-03).

Motorneuron differentiation of gene-edited iPSCs was performed using a multistep differentiation protocol and a combination of small molecules in chemically defined neural induction medium to induce early neural differentiation and production of Olig2-expressing neural progenitors (NPs) from confluent iPSC culture, based on the dual inhibition of BMP and Activin signaling with Compound C (Dorsomorphim), and the activation of Wnt pathway by Chir99021 molecule. Briefly, iPSCs were seeded on matrigel and maintained on mTESR until they reached the confluency. Then iPSC cultures were switched onto neural induction media supplemented with N2 and B27 for growth and long-term viability. The medium was also supplemented with Compound C (Dorsomorphin) and Chir99021. From day 2, patterning of the induced cells was initiated by exposing them to retinoic acid (RA), a caudalizing factor, and SAG, a ventralizing morphogen, to specify the ventral spinal fate and produce neural rosettes. Compound C and CHIR99021 were withdrawn at day 4 of the differentiation process. On day 9, cells were re-plated on Matrigel and the same induction media with RA and SAG was used. Then cells were split harshly at day 20 and the same induction media with RA/SAG was used and supplemented with DAPT, brain-derived neurotrophic factor (BDNF) and glial cell-derived neurotrophic factor (GDNF) for further maturation of MN culture and the

production of highly enriched population of functional MNs. After 30 days of differentiation and induction of cultured cells, MNs were characterized and assessed for the expression of mature-MN markers (Figure 5.1).

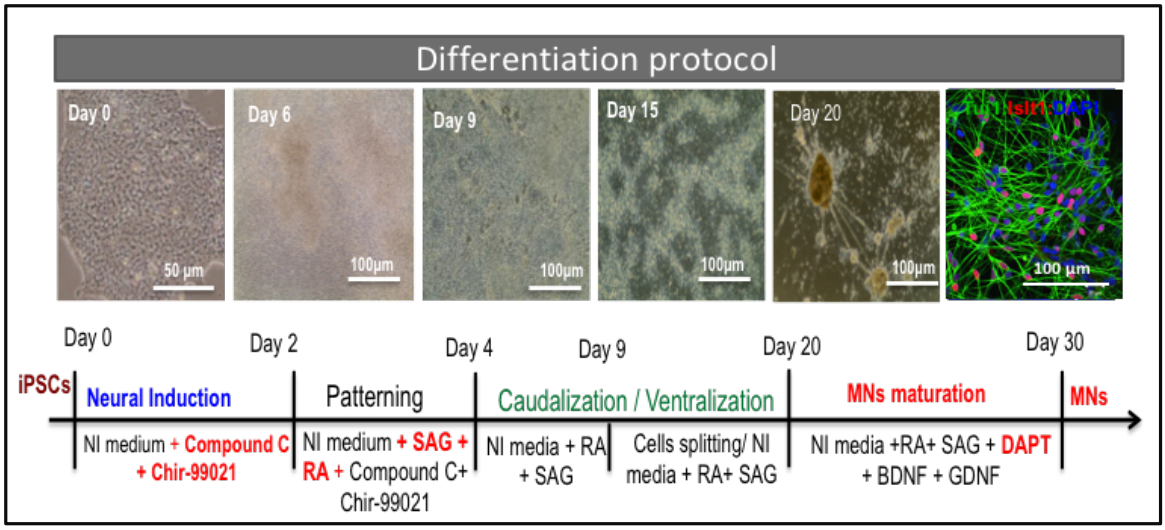

Figure 5.1 Schematic overview and cell morphology changes at different stages of MN differentiation process. iPSCs were induced using Compound C and CHR19901 in neural induction (NI) media to produce neural progenitors. RA/SAG treatment was started at day 2 of the differentiation process to produce neural rosettes. Cells were further induced for maturation in the presence of DAPT and neurotrophic factors BDNF and GDNF. MNs were quantified by immunostaining of HB9, Islet-1, SMI-32 and ChAT after 30 days of MN culture.

The efficiency of the early differentiation process was evaluated after 10 days by immunostaining of early neural progenitor (NP) markers PAX6, Olig2 and Nestin. PAX6 and Nestin showed an efficiency of up to 55% of cells positive for both markers, indicating a successful differentiation towards a neural stem cell fate. Approximately 45% of cultured cells were positive for Olig2 (Figure 5.2.A-C). The obtained NP cells have shown characteristics of early neuroepithelial cells of the CNS and the capacity to produce neural rosettes-like structures *in vitro*. NPs were further induced to become mature neurons by continuous treatment with RA and SAG for neural patterning.

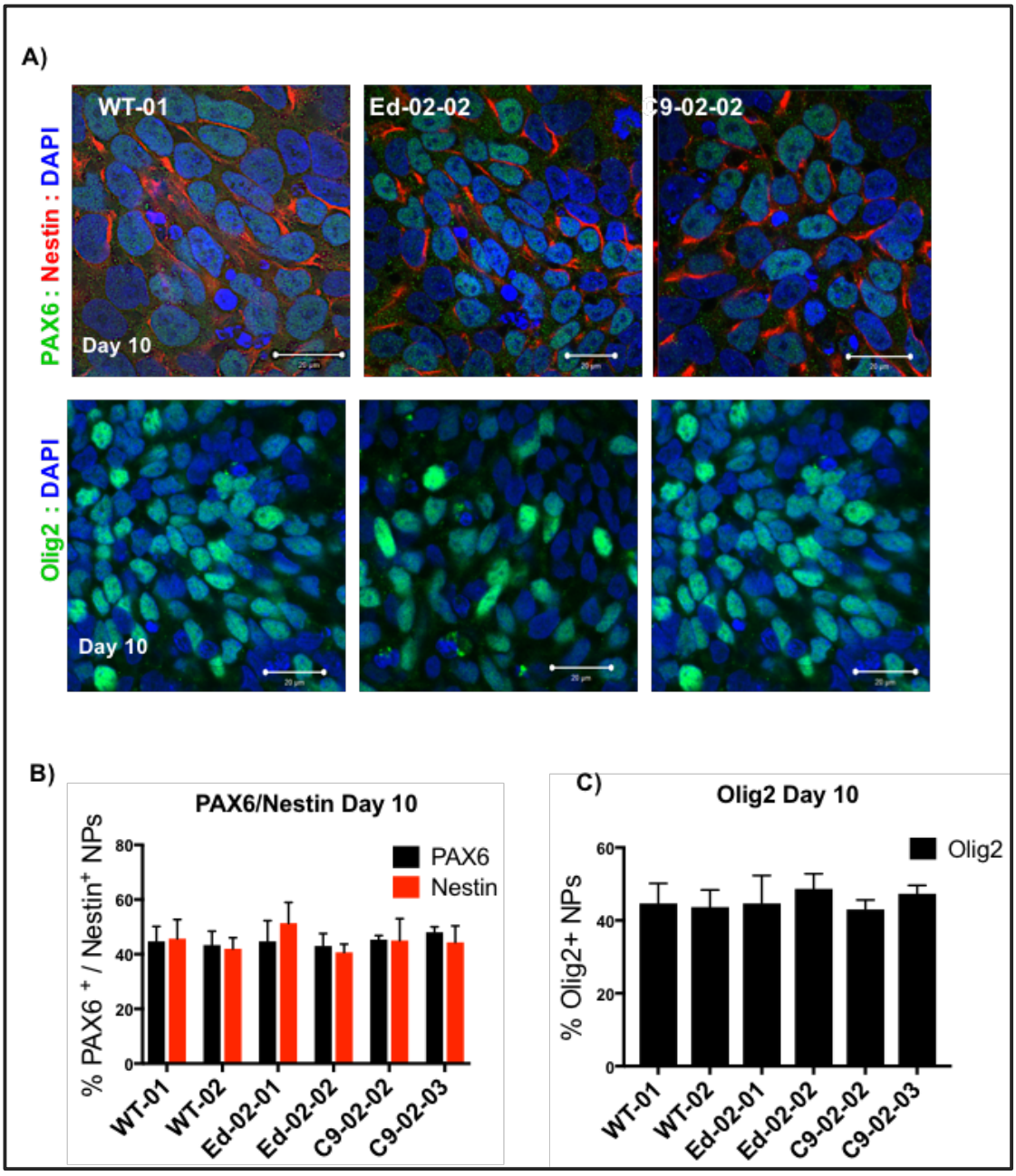

Figure 5.2 Representative immunostaining images and proportion of early NP markers. A) Confocal images of PAX6 (green) and Nestin (red) immunostaining on the first row, and Olig2 staining (green) on the second row, at day 10 of the differentiation process. Cell nuclei were stained with DAPI (blue). Scale bar, 20 µm. **B)** 45-55% of cells showed positive co-staining for PAX6 and Nestin. **C)** ~45% of the NPs showed Olig2 staining.

After approximately 16-20 days of the differentiation process, all the differentiated cell lines showed rosette-like structures as shown in Figure 5.3.

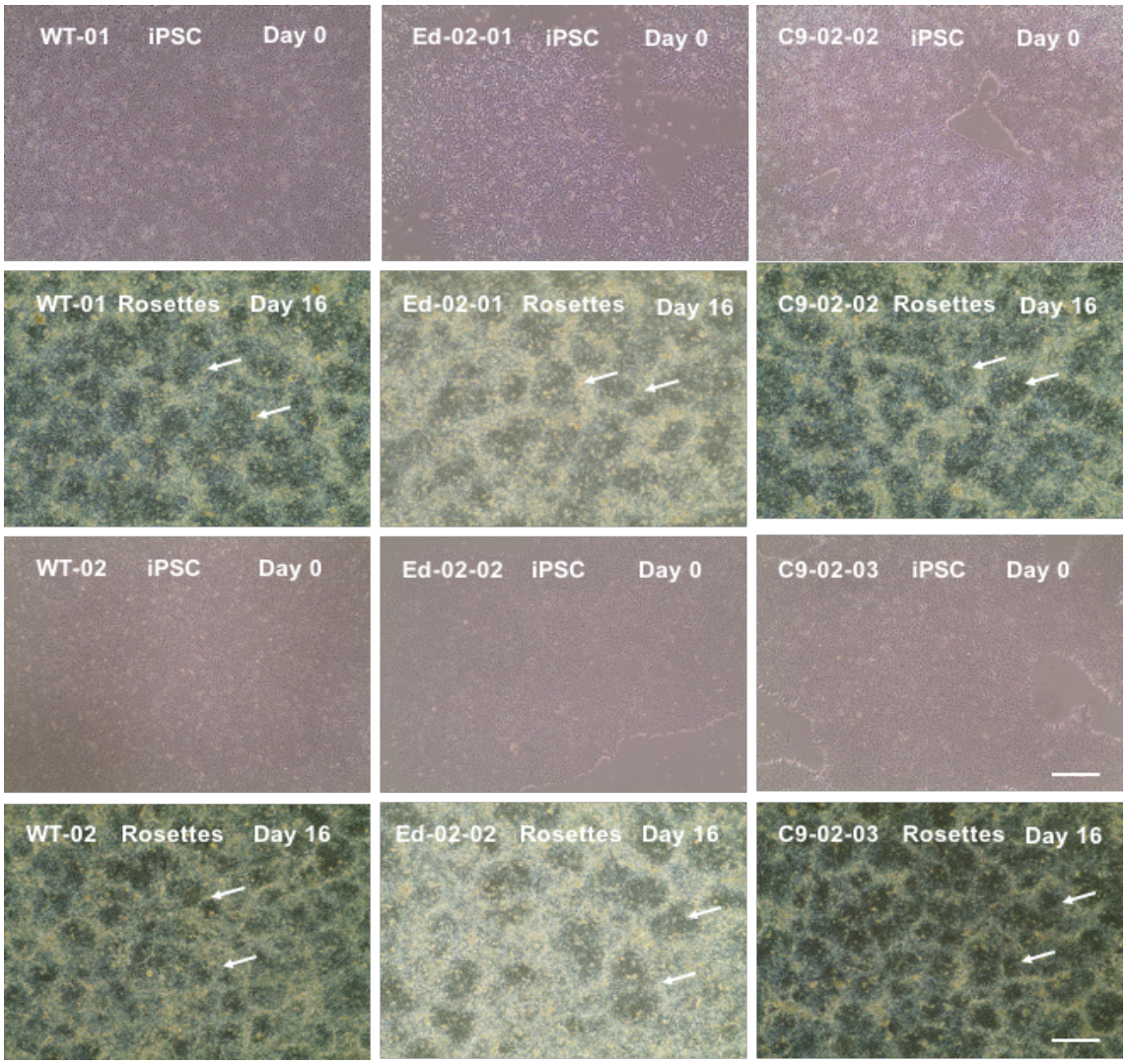

Figure 5.3 Representative phase contrast images from iPSC and neural tube-like rosette cultures. High-density iPSC culture was used for efficient differentiation from all the experimental cell lines. At day 16, all the differentiated cell lines showed typical neural tube-like rosettes. Scale bar, 100 µm.

Cells were also assessed on day 20 for their ability to express Islet-1, an early-stage mature MN marker and the key regulator of human MN development, which is produced after the neural patterning stage (Qu et al. 2014). SMI-32, a non-phosphorylated neurofilament marker that stains MNs in the spinal cord was also examined.

Approximately 30-35 % of cells showed positive staining for Islet-1 and 35-40 % of cells for SMI-32 at day 20. A higher proportion of differentiated cells showed positive Olig2 staining at day 20 (55-60%) compared to day 10 (40-45%), which indicates successful generation and enrichment of motorneuron progenitors (MNPs). No significant differences were observed between the analysed lines ($P > 0.05$, One-Way ANOVA) (Figure 5.4.A-C).

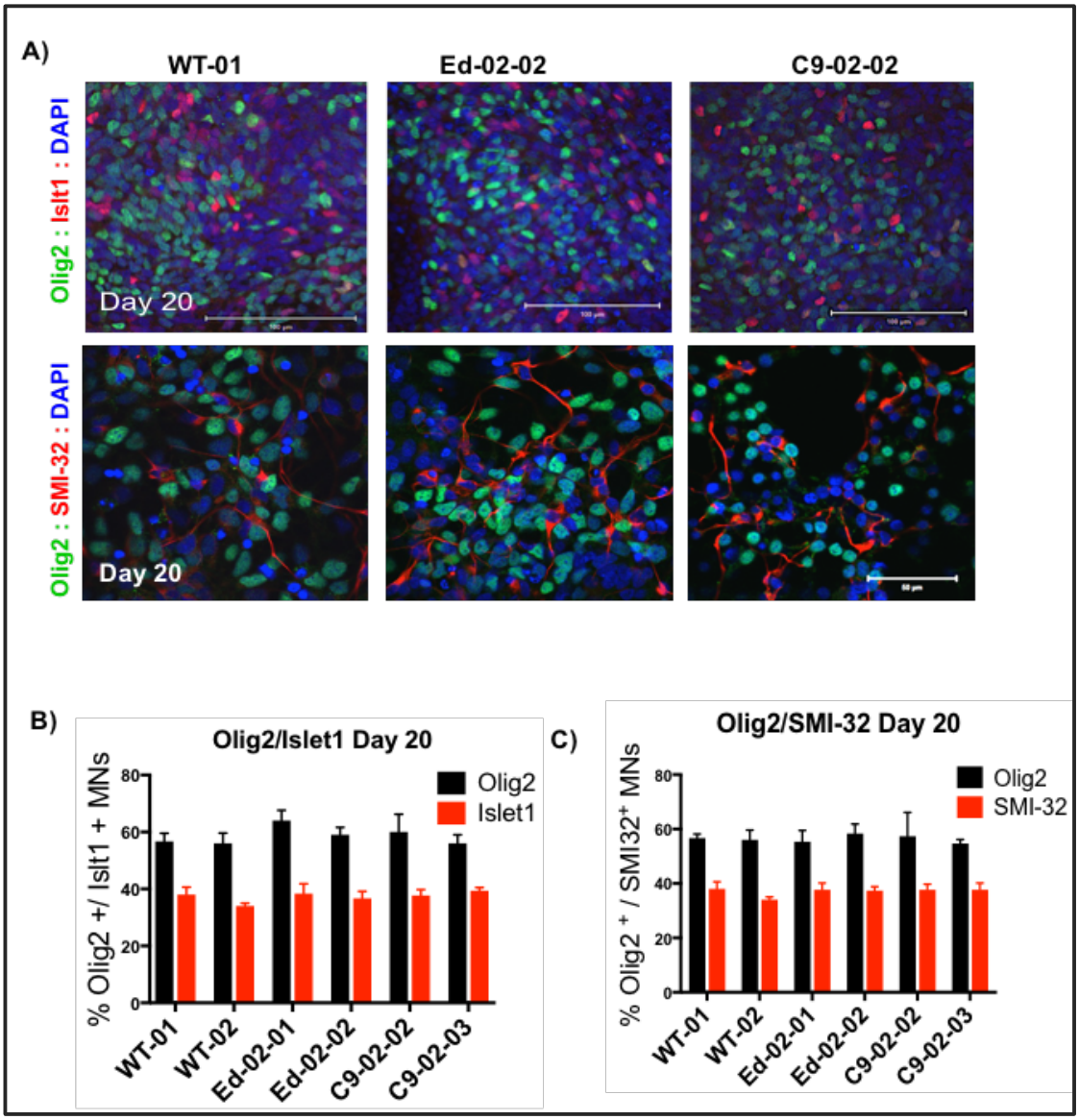

Figure 5.4 Immunostaining of MNP markers on day 20. A) Representative images for co-immunostaining of Olig2/Islet1 and Olig2/SMI-32 at day 20 of the differentiation process. **B)** On day 20, 55-60% of the differentiated MNs were stained positive for Olig2 and 30-35% of cells

stained positive for Islet1, with no significant difference between the analysed lines. **C)** 35-40% of cells showed positive staining for SMI-32. ($P > 0.05$, One-Way ANOVA).

After the last splitting at day 20, cultured cells formed long neurites or filopodia-like structures after the addition of BDNF and GDNF which promote axon elongation. MNs formed long axonal projections with a bipolar morphology, branching and extending their axons to form dendritic trees in a network-like structure within less than four weeks of the differentiation process. Phase-contrast images were taken at week 4 of the differentiation process and all cell lines showed typical mature MN morphology (Figure 5.5).

Motorneurons were further assessed by immunostaining at day 30 of the differentiation process for their ability to co-express mature MN markers ChAT and SMI-32, and to evaluate axonal dendritic MN morphology (Figure 5.5). Choline acetyltransferease (ChAT) is an enzyme catalyzing acetylcholine synthesis for signal conduction across the neuromuscular junction, which is found in MNs of the ventral spinal cord. A high proportion of positive ChAT and SMI-32 immunostaining (60-70%) was observed, suggesting a high proportion of mature MN generation (Figure 5.5).

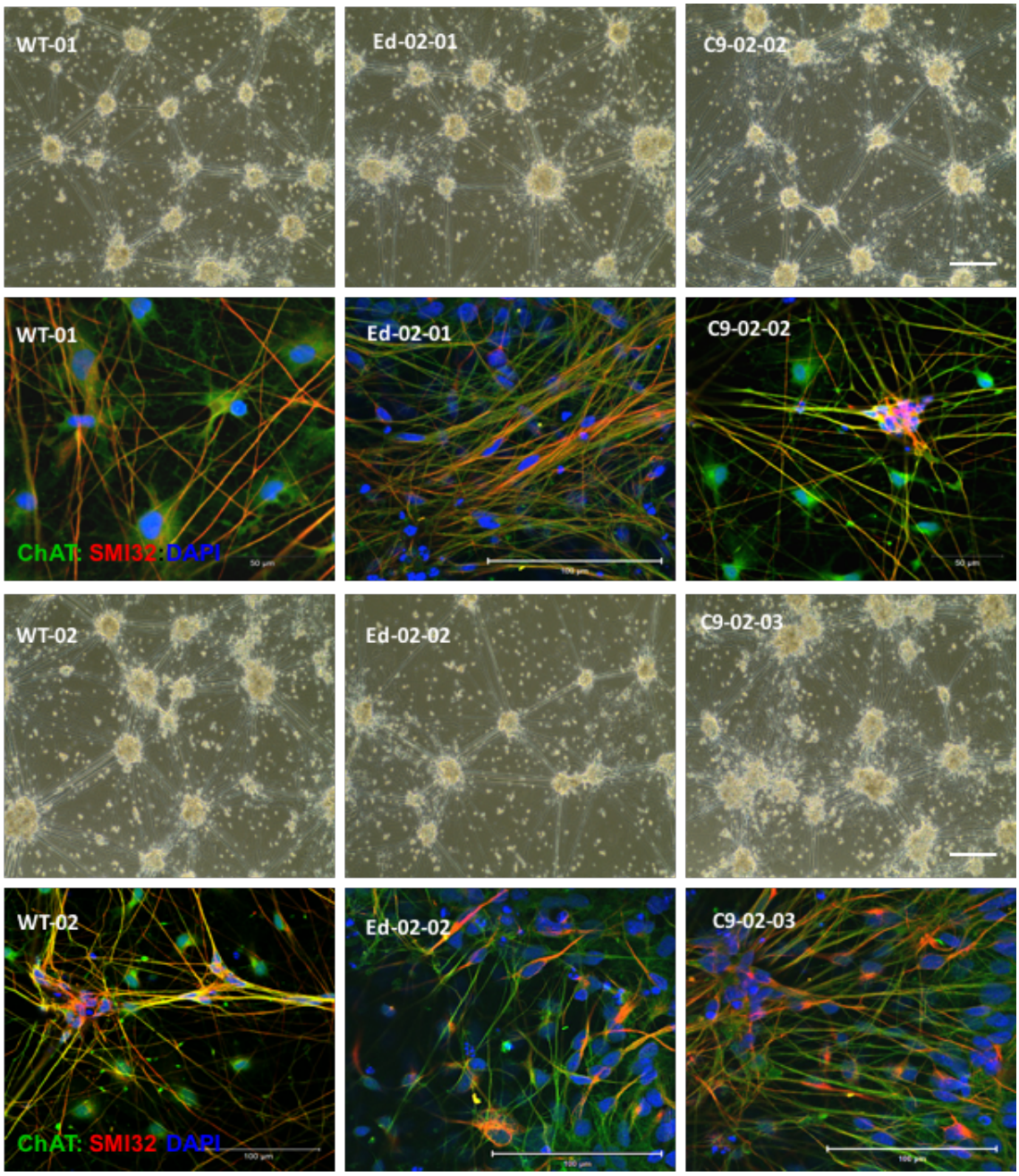

Figure 5.5 Phase-contrast images and confocal images for co-immunostaining of SMI-32 and ChAT at day 30. Typical MN morphology of bipolar and neurite-like cell structure was observed. A high proportion of MNs showed positive staining for SMI-32 and ChAT after 30 days of differentiation, which indicates the formation of axonal and dendrites structure in the cultured MNs. All scale bars are 100 µm and nuclei were stained with DAPI.

All differentiated iPSC-derived MNs showed similar proportions of MN markers expression, as indicated by the graphs from the two healthy lines (WT-01 and WT-02), the two edited lines (ED-02-01 and ED-0-02) and the two patient lines (C9-02-02 and C9-02-03) (Figure 5.6.B-E). To further demonstrate the successful differentiation of iPSCs into spinal cord neurons, differentiated cells were stained with HB9, which encodes for a transcription factor specifically expressed by mature MNs. Cells showed HB9 positive staining in the nucleus (~35-40%) and ChAT positive staining in the cytoplasm with neurites staining (Figure 5.6.A&). No significant differences between the analysed samples were observed ($P > 0.05$, One-Way ANOVA). Furthermore, Olig2 was completely absent on Day 30 of differentiation (Figure 5.6.A&E). While SMI-32 positive staining has increased at day 30 (70-80%) compared to day 20, which indicates further maturation of the differentiated cells (Figure 5.6 A&B). Islet-1 staining showed an increase in proportion of positive cells to 35-40% compared to day 20 results (Figure 5.6.A&D).

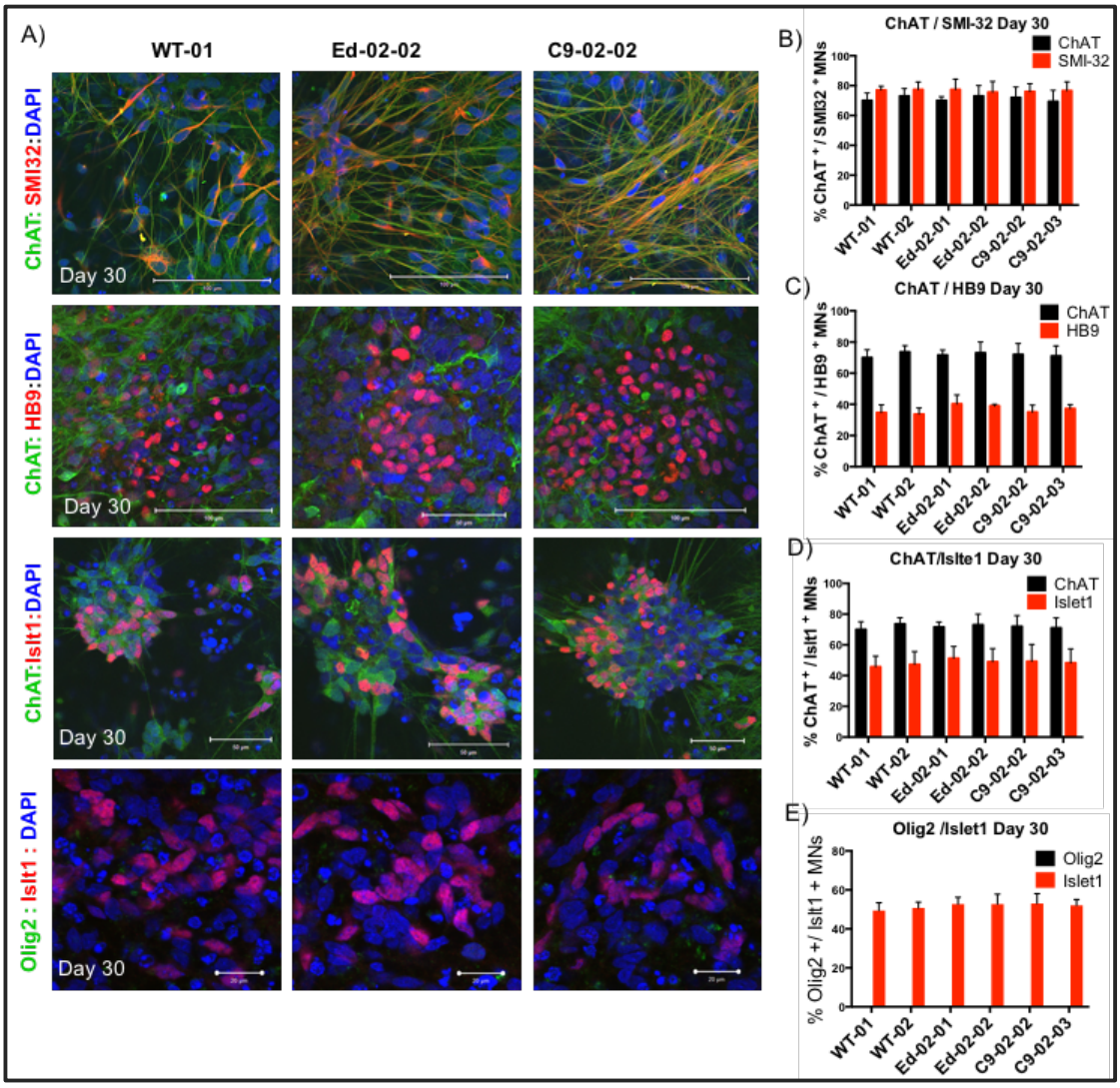

Figure 5.6 Characterization of iPSC-derived MNs by immunostaining of mature MN differentiation markers at day 30. A) Representative confocal immunostaining images for mature MN differentiation markers (ChAT, SMI-32, HB9, Islet1 and Olig2) for cells derived from WT-01 healthy control, Ed-02-02 edited clone and the C9-02-02 parental clone. **B-E)** Proportion of positive MN specific markers in iPSC-derived MN culture from two healthy lines WT-01 and WT-02, two edited lines (Ed-02-01 and Ed-0-02) and two patient lines (C9-02-02 and C9-02-03). **B)** MN population showed 60-70 % positive staining for ChAT, 75-80% of cells stained for SMI-32. **C)** Approximately, 35-40% of cells were stained positive for HB9. **D-E)** Olig2 was completely absent on Day 30 of differentiation, while Islet1 showed 35-40% of nuclear staining. No significant differences were observed between the analyzed lines ($P > 0.05$, One-Way ANOVA). Scale bars are shown on the figure.

5.3. Summary

Taken together, the results obtained here reveal that neural patterning had been successfully obtained by simultaneous inhibition of bone morphogenetic protein (BMP) and activin signaling pathways in cultured iPSCs. High density plating of iPSCs on Matrigel, early induction of cells with compound C and CHIR99021, and the early treatment with RA and SAG have resulted in increasing the efficiency of neural induction and neuroepithelial cell proliferation and production of mature MNs.

In this Chapter, the monolayer differentiation protocol has produced a higher proportion of Islet1 and HB9 positive MNs compared to the previously applied EBs protocol in Chapter 2 and showed higher percentage of positive immunostaining for SMI-32 and ChAT. No significant differences in the MN differentiation efficiency were observed in any of the analyzed patient iPSC lines when compared to healthy controls and edited clones.

These results suggest that corrected iPSC lines are able to undergo successful differentiation into MNs, which then can be used for studying disease-related phenotypes and directly comparing them to the uncorrected original cells.

6. CHAPTER SIX

Studying ALS disease-related phenotypes in C9orf72 edited iPSC-derived MNs

6.1. Introduction

The G4C2 hexanucleotide repeat expansion (HRE) mutation in *C9orf72* gene noncoding region is the most common genetic cause of ALS/FTD disease. The normal function of C9orf72 protein and the mechanism by which the HRE mutation causes the neurodegeneration is still unknown. However, it has been reported that the *C9orf72* gene product is related to DENN-like Rab-GTPases that involved in vesicular trafficking during autophagy pathway, a mechanism has been correlated to neurodegenerative diseases (Levine et al. 2013). It has also been suggested that *C9orf72* plays a role in membrane trafficking (Farg et al. 2014).

Three potential pathogenic mechanisms have been reported in HRE-related ALS/FTD disease neurodegeneration. The first two proposed mechanisms are related to toxic gain-of-function. First, the RNA-repeat sequences can be transcribed bidirectionally, which results in the formation of GGGGCC sense and CCCCGG antisense RNA transcripts. Accumulation of these transcripts inside the nucleus leads to the formation of toxic intranuclear sense and antisense RNA foci (Mizielinska et al. 2013) and sequestration of RNA binding proteins (Cooper-Knock, Walsh, et al. 2014), suggesting a toxicity-

mediated gain of function mechanism. Second, the abnormal RNA sequences can undergo repeat-associated Non-ATG translation (RAN) into dipeptide repeat proteins, which aggregate intracellularly and form toxic cytoplasmic inclusion bodies (Ash et al. 2013; Mori, Weng, et al. 2013). Third, the expansion may result in decreased tissue levels of *C9orf72* gene variants, which may contribute to disease development via haploinsufficiency (DeJesus-Hernandez et al. 2011; Renton et al. 2011; Belzil, Bauer, et al. 2013; Xi et al. 2013).

Toxic gain-of-function due to the accumulation of repeat-associated RNA sequences or DPR proteins is considered as the main mechanism of pathogenicity. However, haploinsufficiency remains a possible important mechanism in the disease pathogenicity as motor deficit has been observed in a zebrafish model of *C9orf72* loss-of-function (Ciura et al. 2013). Previous studies have also shown that cells with HRE mutation and C9orf72 haploinsufficiency are more vulnerable to glutamate toxicity (Christopher J Donnelly et al. 2013) and autophagic stress (Almeida et al. 2013).

The discovery of G4C2 mutation as the most common pathogenic mutation in familial ALS and FTD cases holds a promise for finding new therapeutic strategies based on effective targeting of the repeat expansion region (DeJesus-Hernandez et al. 2011; Renton et al. 2011).

In this thesis, isogenic lines were generated by CRISPR/Cas9 gene editing from an ALS/FTD patient iPSC line (C9-02-02) as previously described in Chapter Four, and then successfully differentiated into MNs as described in Chapter Five. To ascertain whether the genotypic correction of the G4C2 mutation can rescue G4C2 ALS disease-related

phenotypes, iPSC-derived MNs from the same experimental groups described in Chapter Five; edited clones (ED-02-01 and ED-02-02), the parental diseased cells (C9-02-02 and C9-02-03) and healthy controls (WT-01 and WT-02) were used to study the disease-related phenotypes and to try to understand the pathogenicity of repeat expansions mutation after excision of the abnormal repeat region.

6.2. *C9orf72* methylation levels

6.1.1 Validation of *C9orf72* promoter methylation by *Hha*I and *Hpa*II-qPCR

It has been hypothesised that the HRE mutation can cause disease by loss of function of the *C9orf72* gene. This mechanism results from DNA hypermethylation of the CpG islands in the GC-rich promoter region and histone trimethylation of the *C9orf72* gene, leading to downregulation and silencing of *C9orf72* gene transcription (Liu et al. 2014; Belzil, Bauer, et al. 2013). The loss-of-function mechanism is also associated with a decrease in accumulation of toxic RNA byproducts, RNA foci and DPR protein aggregates formation, which suggests that *C9orf72* hypermethylation could reduce the disease pathogenicity by protecting against gain-of-function toxicity (McMillan et al. 2015).

To evaluate *C9orf72* DNA promoter methylation levels in edited lines and to compare them with the original parental cells, a methylation-sensitive restriction enzyme qPCR was applied as previously described (Liu et al. 2014; Russ et al. 2015). A schematic representation adapted from previously published study (Liu et al. 2014) for the 5' promoter region of the *C9orf72* gene upstream the HRE locus was shown in Figure 6.1A. CpG dinucleotides are indicated by vertical bars, *Hha*I and *Hpa*II recognition sites within the qPCR amplicon as well as *Hae*III restriction sites are all identified on the schematic.

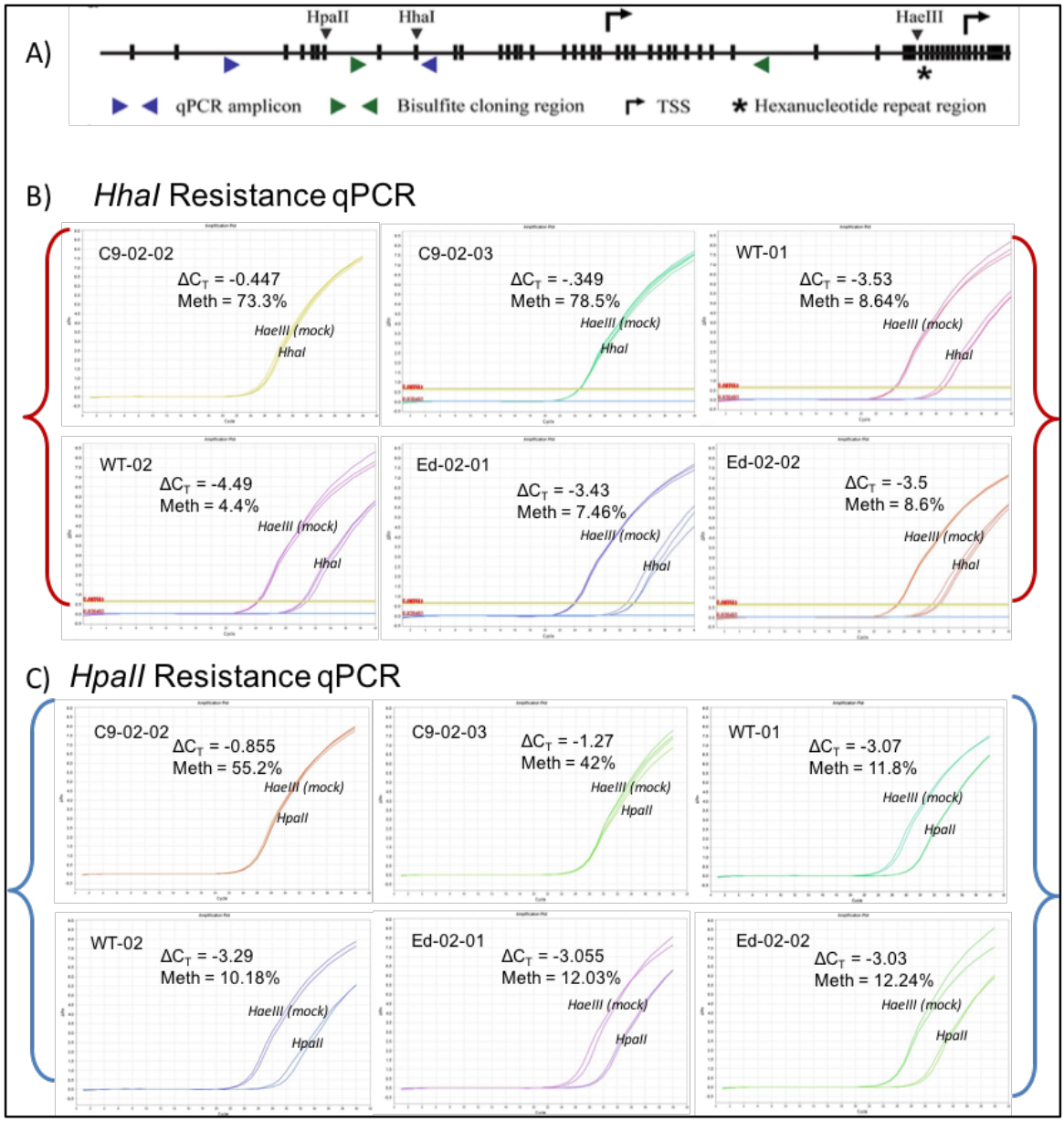

Figure 6.1 C9orf72 promoter hypermethylation in iPSC-derived MNs. A) Schematic representation of the 5' promoter region of *C9orf72* gene upstream the G4C2 repeat locus (Liu et al. 2014). CpG dinucleotides are indicated by vertical bars. The *Hha*I and *Hpa*II sites within the qPCR amplicon as well as *Hae*III recognition sites are identified on the schematic. **B&C)** Representative qPCR amplification curves for mock (*Hae*III) versus *Hha*I-digested DNA in **B** or *Hpa*II digested in **C**, showing shift in the amplification curves used to calculate the percentage of methylation level. The resulting ΔCT values were used to calculate the percentage of DNA resistance to *Hha*I digestion. Disease lines C9-02-02 and C9-02-03 show hypermethylation after *Hha*I and *Hpa*II digestion, whereas the edited clones and the healthy controls show normal methylation levels as indicated by the curve shifts.

Briefly, DNA samples derived from iPSC-MNs were digested with *HhaI* enzyme, which cuts the amplified promoter region in the absence of CpG DNA methylation. The

percentage of amplified methylated DNA sequences resistant to digestion due to the presence of CpG methylation was quantified by qPCR and analysed by comparing with a mock- *Hae*III digested DNA control for each line that does not cut within the PCR region. An *Hpa*II digestion site, which also cuts the unmethylated DNA sequence of the promoter region upstream of the *Hha*I site, was examined and used as another evidence to confirm the methylation results in the CpG promoter region obtained by *Hha*I digestion. Shifts in qPCR amplification curves for mock (*Hae*III) versus *Hha*I-digested or *HpaII*-digested DNA were compared and used to calculate ΔCT values. The percentage of DNA resistance to *Hha*I or *Hpa*II digestion ($2^{\Delta CT} \times 100$) was also calculated as a measure of DNA methylation (Figure 6.1.B&C). All samples were run in triplicates from at least three different differentiation experiments.

Quantification of the methylated DNA products showed that CpG methylation was significantly higher in C9-02-02 and C9-02-03 patient lines compared to WT-01 and WT-02 healthy controls (****P < 0.0001, One-Way ANOVA) and Ed-02-01 and Ed-02-02 lines (****P < 0.0001, One-Way ANOVA) after *Hha*I and *Hpa*II digestion (Figure 6.2.A&B), suggesting that CpG methylation at these two sites has been normalized in edited clones to levels similar to the healthy controls. The average methylation level of healthy or edited clones was between 8-10%, compared to >70% in uncorrected disease cells after *Hha*I digestion. The methylation level after *Hpa*II digestion was between 10-12% and > 40% in healthy or edited clones respectively. No significant differences in the methylation status were observed between healthy controls and edited lines.

Similar methylation results were also observed in DNA samples derived from iPSCs before differentiation, using the same experimental cell lines and the same methylation assays (Figure 6.2.C&D). A significant decrease in methylation status of the promoter

was observed in edited clones to levels similar to healthy controls when compared to parental disease clones after digestion with either *Hha*I or *Hpa*II (****P < 0.0001, One-Way ANOVA).

To assess the sensitivity of this assay, unmethylated DNA from WT-01 control was *in vitro* methylated by *Mss.I* methyltransferase enzyme and different methylation concentrations were prepared by mixing the *Mss.I* methylated and unmethylated DNA of the same control. After that, *HhaI* digestion of the prepared mixes was performed and samples were amplified, analyzed and plotted to show the increasing amount and the linearity of input methylated DNA (Figure .2.E&F).

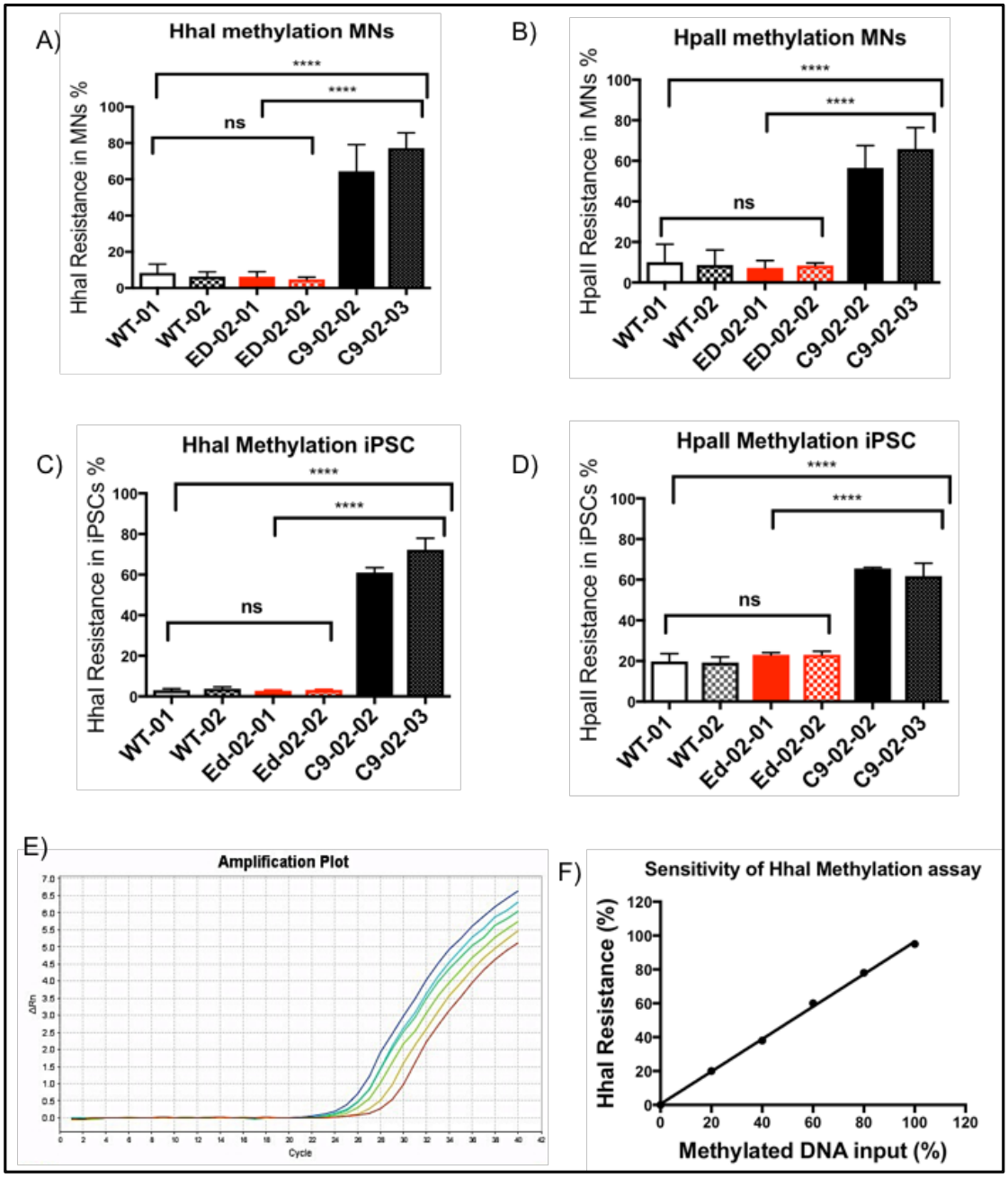

Figure 6.2. Evaluation of *C9orf72* promoter DNA methylation. A&B) Reduction back to normal levels of C9*orf*72 CpG promoter methylation levels in iPSC-derived MNs from Ed-02-01 and Ed-02-02, compared to CpG hypermethylated C9-02-02 and C9-02-03 iPSC-derived MNs, assayed by digestion of the DNA samples with *Hha*I and *HpaII* (****P < 0.0001, One-Way ANOVA). **C&D)** Methylation levels of iPSCs after digestion of DNA samples with either *HhaI* or *HpaII***.** Significant differences in methylation levels were observed between the edited clones and the original disease clones (****P < 0.0001, One-Way ANOVA). **E-F)** DNA from WT-01 healthy control was methylated *in vitro* with *Mss.I* and various ratios of methylated and unmethylated DNA were mixed and digested with *HhaI* a to determine the sensitivity and linearity of the qPCR methylation assay.

6.1.2 Validation of *C9orf72* promoter methylation by bisulfite sequencing

All data related to this assay were obtained by Dr Andrew Douglas from the University of Southampton. In order to ascertain the methylation status of multiple CpG sites within the 5' CpG island region of *C9orf72*, genomic DNA samples isolated from iPSC-derived MNs were treated with sodium bisulfite and subjected to PCR using primers specific to the 5' CpG region. Direct Sanger sequencing of PCR products allowed relative quantification of C versus T nucleotides at each CpG site as a measure of methylation, since unmethylated C is converted to T by bisulfite, while methylated C (mC) remains unchanged. A total of 26 CpG sites could be assessed by this method, including one site (CpG 5) formed by an A>G SNP positioned 151 nucleotides upstream of the exon 1a start site (rs1373537, hg38 9:27574017T>C), which was present in all DNA samples analysed in this study. Although the C allele of this SNP is not present in the reference sequence, it is in fact the predominant genotype in all populations studied, with the T allele having a minor allele frequency of only 0.01 as determined by the 1000 Genomes Project.

On bisulfite sequencing, all 26 CpG residues from iPSC-MNs derived from C9-02-02 and C9-02-03 parental G4C2 mutation positive clones were found to be methylated (>50% mC). In contrast, the same CpG residues were found to be unmethylated (<30% mC) in control iPS-derived MNs (WT-01 and wt-02) and in edited clones (Ed-02-01 and Ed-02-02). Notably, the majority (86%) of CpG residues in the analysed samples were either 100% or 0% methylated (Appendix V). Representative chromatograms of a sequence from the promoter region containing some CpG sites from the wild-type controls, edited and diseased lines are shown in Figure 6.3. C9-02-02 and C9-02-03 sequences show CpG methylation in 5 CpG regions as indicated by the presence of C nucleotides, whereas the edited clones show the presence of a T nucleotide which indicates the demethylation of

the CpG sites in the promoter region. CpG 1 was the site most likely to have a mixed methylation pattern, being mixed in every sample tested. Note that in this assay, CpG 2 corresponds to the site assayed by the methylation-sensitive restriction digestion assay. These results were compatible with restriction enzyme-based methylation results.

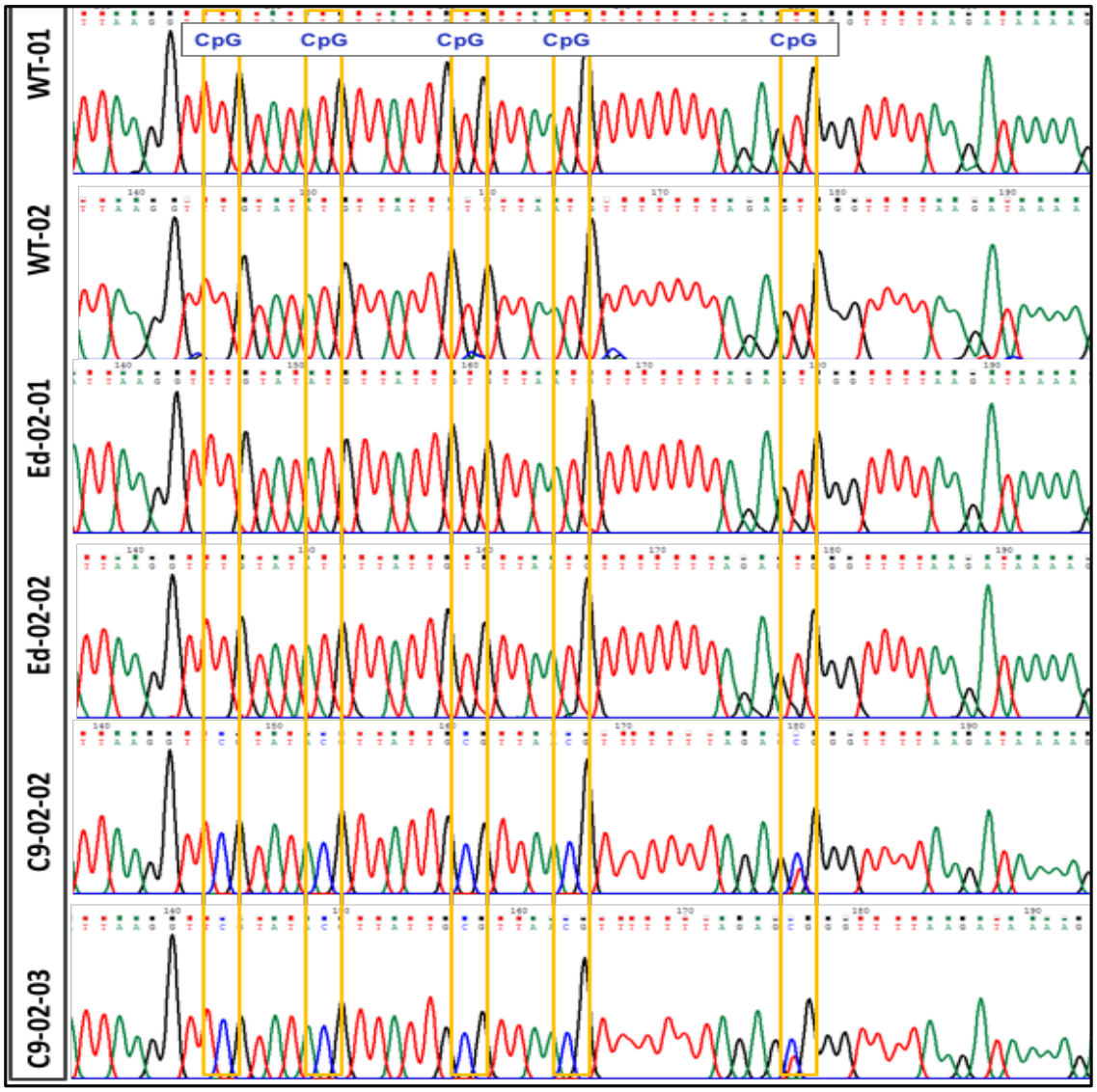

Figure 6.3 Bisulfite sequencing of the *C9orf72* gene promoter region. Representative chromatograms of a sequence from the promoter region containing CpG sites from the wild-type controls, edited and diseased lines. C9-02-02 and C9-02-03 sequences show CpG methylation in 5 CpG regions as indicated by the presence of C nucleotides, whereas the edited clones show the presence of a T nucleotide which indicates the demethylation of the CpG sites in the promoter region.

6.1.3 Validation of G4C2 repeat methylation using methylation-sensitive restriction enzymes and an RP-PCR assay

It has been reported that the G4C2 repeat expansion itself can be methylated (Xi, Zhang, et al. 2015). To determine whether G4C2 repeat region is methylated, genomic DNA from the same WT controls, edited clones and the two parental G4C2 lines were digested using the following three restriction enzymes: *Msp*I that cuts the hexanucleotide repeat region regardless of the methylation status, *Hpa*II that cuts the repeat region only if it is not methylated, and *MspJ*I that cuts only if the repeat region contains methylated CpGs. After digestion, the digested products were amplified by RP-PCR using RP-PCR2 primer combination used previously in Chapter 4. RP-PCR was used on this assay, because the conventional PCR is unable to amplify the GC-rich repeat expansion region. After analyzing the RP-PCR results, the parental cells of C9-02-02 showed amplification in the repeat region after mock and *MspJ*I digestion which indicates that the repeats region is unmethylated. The repeat region was not amplified after *Hpa*II and *Msp*I digestion, which also confirms that the repeat region is unmethylated. The same results were observed as expected for controls and edited lines with no methylation observed in any of the amplified samples (Figure 6.4).

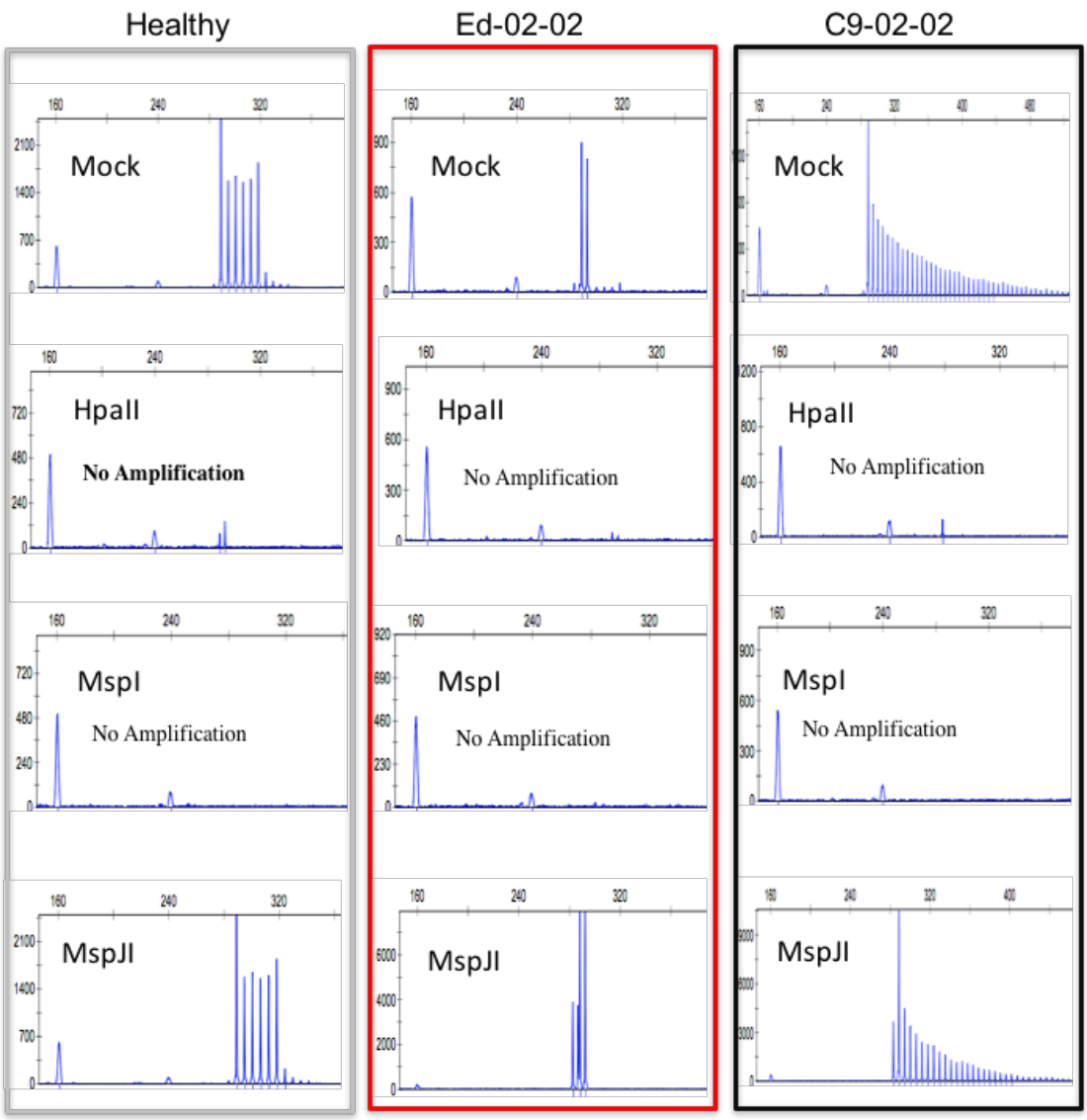

Figure 6.4. Hexanucleotide repeat methylation assay. Mock no enzyme, *Hpa*II, *Msp*I, and *MspJ*I digested samples were amplified by RP-PCR. The G4C2 repeat region showed no methylation in our patient lines C9-02-02 and C9-02-03 as indicated by *MspJ*I amplification of unmethylated DNA and the lack of amplification after digestion with *Hpa*II which reveals that the repeat expansion region in the disease lines was not methylated.

In conclusion, the gene correction approach effectively and significantly decreased and rescued the methylation levels as confirmed by qPCR and bisulfite sequencing assays. These results are correlated with a significant rescue of the hypermethylation disease phenotypes.

6.2 *C9orf72* gene mRNA transcripts

DNA methylation in CpG islands leads to reduction in *C9orf72* mRNA levels in ALS patients carrying the repeat expansion mutation. Three *C9orf72* mRNA transcripts have been identified, variant 1, 2, and 3 (NM_145005.5, NM_018325.3, and NM_001256054.1, respectively), with variant 1 and 3 containing open reading frames (ORFs) starting upstream of the expanded G4C2 repeat. Transcript variants 2 and 3 code for a 481-amino acid long protein by translating exon 2-11 of the *C9orf72* gene, whereas variant 1 is predicted to produce a 222-amino acid protein by translating exons 2–5 **(**Figure 6.5.A). The G4C2 repeat region is located in the noncoding region between two alternatively spliced non-coding exons, 1a and 1b. The location of the repeats can be either in the promoter region (V2) or in intron 1 (V1 and V3) (Figure 6.5.A).

Since G4C2 repeat expansion mutation in *C9orf72* induces downregulation of the *C9orf72* transcript due to the presence of DNA hypermethylation, the *C9orf72* gene expression levels were subsequently examined to confirm the normalization of *C9orf72* gene levels after gene editing. The analysis was performed using RNA samples from the same edited, diseased and control iPSC-derived MN groups. RNA samples were analyzed by SYBRGreen Quantitative RT-PCR assay with primers for amplification of overall *C9orf72* mRNA isoforms, the individual V1 short variant, V2 and V3 long variants by the amplification of exon 8 and exon 9 and finally V1 and V3 variants through the amplification of exon 1a region (Figure 6.4.A). Relative quantification was applied using ΔΔCt method after normalization of the analyzed variants to endogenous GAPDH control and healthy control samples. After analyzing the Q-PCR results, total *C9orf72* RNA level was significantly higher in edited clones when compared to C9-02-02 and C9-02-03

disease clones (*P****<0.0001, One-Way ANOVA) (Figure 6.5.B). A significant difference was also observed between healthy (WT-01 and WT-02) controls and the downregulated disease lines (*P****<0.0001, One-Way ANOVA). The genetic correction and the normalization of DNA methylation levels in edited clones also resulted in a significant increase in the RNA levels of V1 variant (*p***<0.001, One-Way ANOVA), V2+V3 long variants (*p****<0.0001, One-Way ANOVA) and V1+V3 variants (*P** < 0.001, One-Way ANOVA) in edited clones compared to disease lines. Significant differences were also observed between parental iPSC-derived MNs and WT controls (*p****<0.0001, One-Way ANOVA) (Figure 6.5.C-E). These results demonstrate that *C9orf72* promoter methylation due to the presence of the repeats is directly linked to transcriptional silencing and downregulation of mutant *C9orf72* gene. No significant differences were observed between edited and control groups (*P*>0.05, One-Way ANOVA). To sum up, these results confirm that CRISPR gene editing of the repeat expansion has significantly decreased the level of *C9orf72* promoter DNA methylation in iPSC-derived MNs and subsequently increased *C9orf72* gene variant expression levels through the restoration of normal DNA methylation levels.

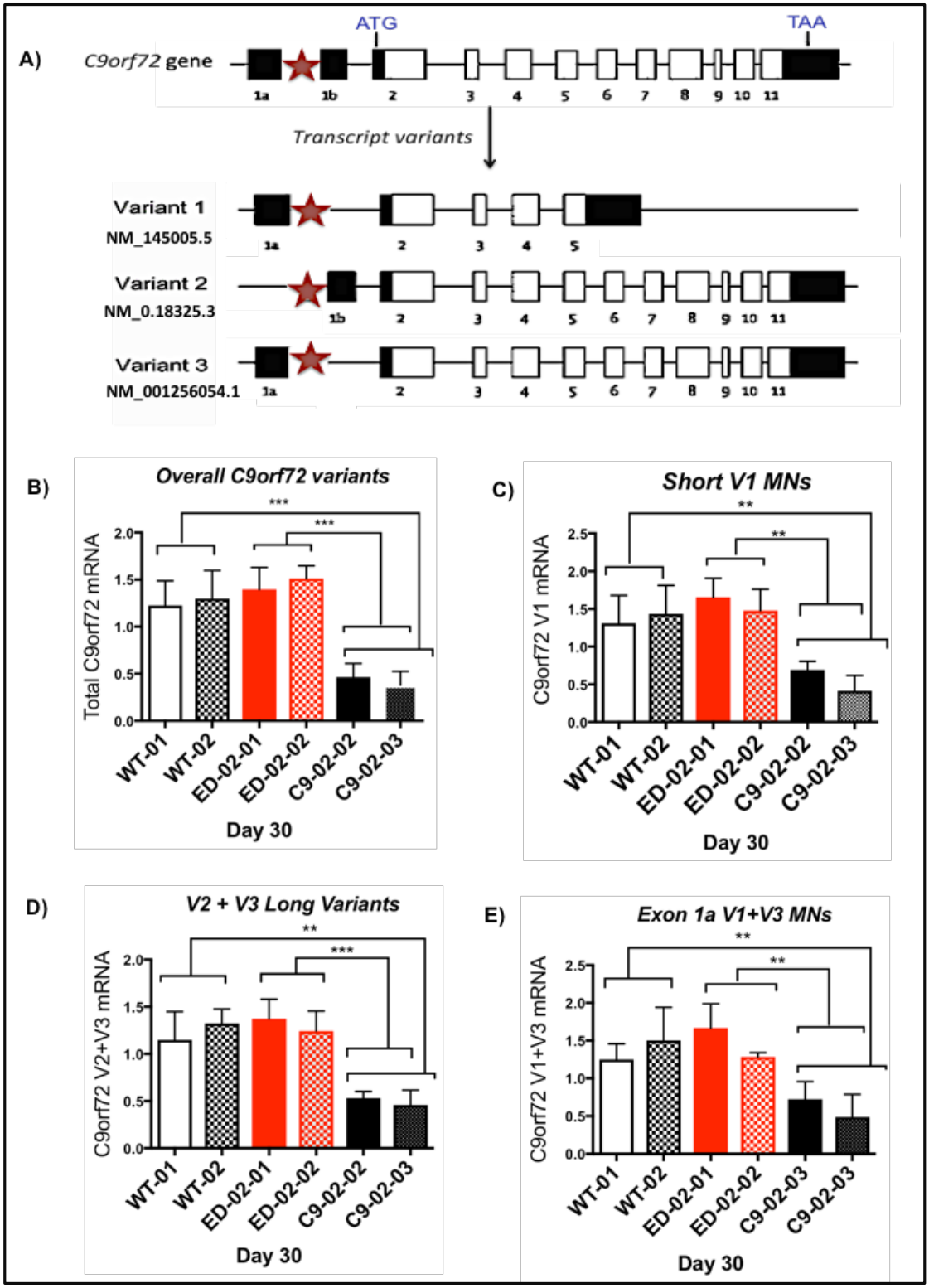

Figure 6.5 Evaluation of *C9orf72* gene variant expression levels. A) Schematic of the three *C9orf72* transcript variants produced by alternative pre-mRNA splicing and the location of the G4C2 repeats on each transcript. **B)** A significant increase in total C9*orf*72 RNA levels was observed in edited clones Ed-02-01 and Ed-02-02 compared to disease C9-02-02 and C9-02-03 clones. **C-E)** A significant increase in expression of V1, V2 + V3, and repeat-containing variants V1+V2. Statistical differences were calculated by one-way ANOVA with Bonferoni correction. **p*< 0.05, ***p*< 0.01, ****p* <0.001 * = p < 0.05 One-way- ANOVA.

6.3 Detection of the presence of RNA foci and DPRs in CRISPR/Cas9 corrected iPSC-derived MNs

Sense (GGGGCC) and antisense (GGCCCC) nuclear RNA foci have been identified in ALS disease patient cells derived from different tissue sources, including white blood cells, fibroblasts, and several nervous system cell types (spinal motor, cortical, hippocampal, and cerebellar neurons) and also in iPSC-derived MNS (Gendron et al. 2013; Lagier-Tourenne et al. 2013; Mizielinska et al. 2013; Zu et al. 2013).

The presence of RNA foci and dipeptide repeat proteins (DPRs) are the hallmarks of *C9orf72* neuropathology-associated gain-of-function mechanism. It has been previously reported that *C9orf72* promoter hypermethylation and *C9orf72* downregulation are associated with the reduction in RNA foci formation and DPR aggregates accumulation (Liu et al. 2014). This would suggest that the number of RNA foci and the frequency of DPR in our ALS/FTD C9-02 patient cells are expected to be low because of the high methylation level observed in patient iPSC and iPSC-derived MNs.

The presence of sense and antisense RNA foci was evaluated by fluorescence in situ hybridization (FISH) using a probe targeting the sense GGGGCC repeat (G2C4) and the CCCCGG antisense repeats in iPSC-derived MNs from all the experimental cell lines (WT-01, WT-02, Ed-02-01, Ed-02-02, C9-02-02 and C9-02-03). Co-immunostaining of SMI-32 with FISH was performed to detect sense and antisense RNA foci in SMI-32 positive MNs. Representative images of FISH staining demonstrate that multiple sense RNA foci were detected in MNs derived from unedited C9-02-02 and C9-02-03 patient lines (Figure 6.6.A). No foci were detected in any of the healthy controls WT-01 and WT-02 nor the edited clones Ed-02-01 and Ed-02-02 in at least 100 nuclei. Several antisense

RNA foci were also detected in patient lines but not in the healthy or edited clones (Figure 6.6.B). Most nuclei showed single or double RNA foci but multiple nuclear foci and cytoplasmic foci have also been observed in iPSC-derived MNs from C9-02-02 and C9-02-03 patient lines. The proportion of cells containing sense or antisense RNA foci per field was 10-12 % of the MN culture (Figure 6.6.C). At least 10 different fields were examined (approximately 10-12 cells per field) of three independent differentiation experiments.

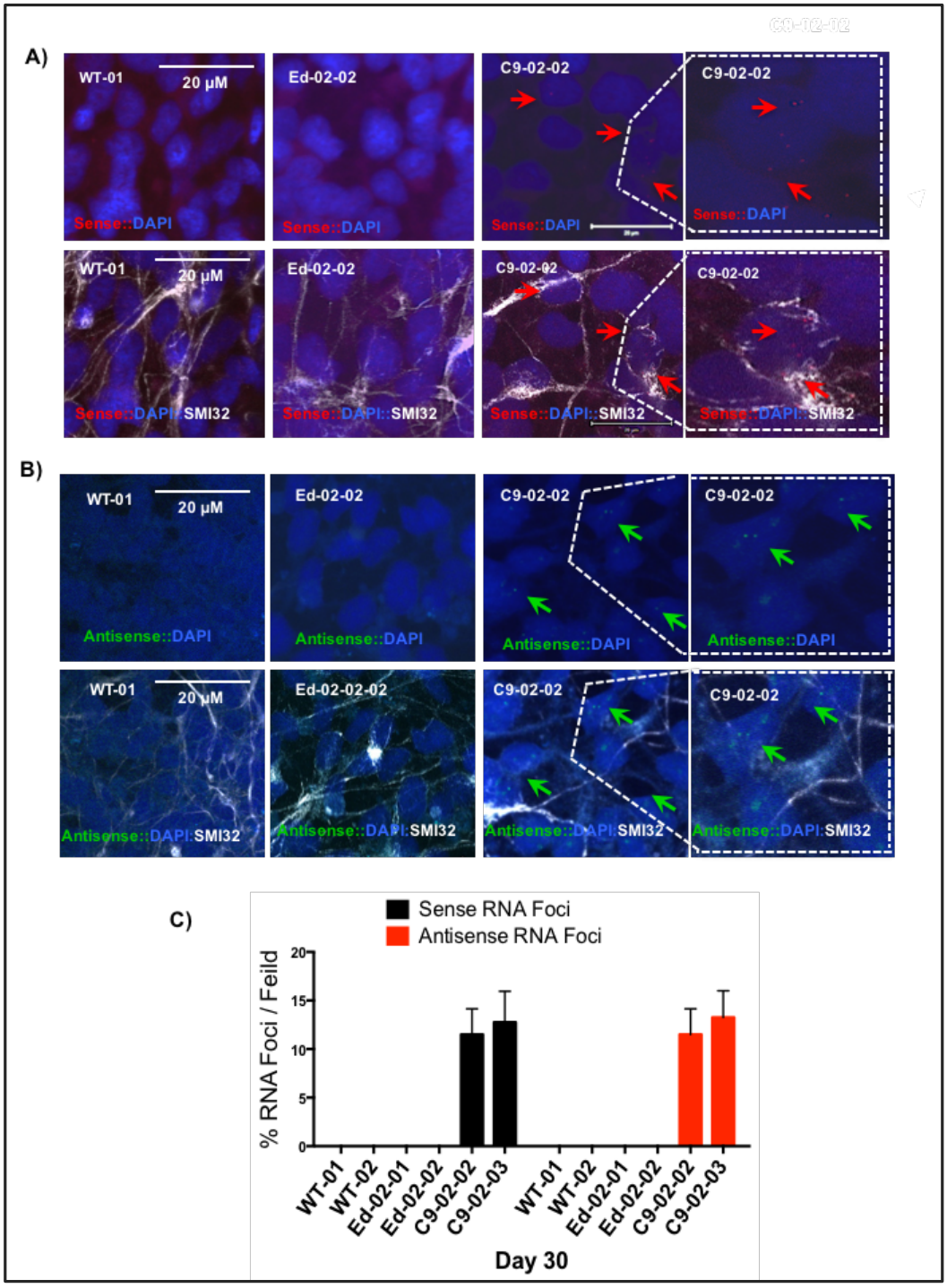

Figure 6.6 Abolition of sense and antisense RNA foci in edited iPSC-derived MNs. Quantification of cells containing sense and antisense foci was performed by FISH staining using repeat-specific probes. Nuclear foci are indicated by arrows and SMI-32 was used as a specific MN staining marker. **A)** Representative images of sense RNA foci using G2C4-Cy3 probe (Red). Multiple foci were detected in C9-02-02 clone. **B)** Antisense RNA foci were also detected in C9-02-02 iPSC-MNs using G_4C_2-alexa488 (green). No foci were detected in healthy or edited clones as shown in WT-02 healthy and Ed-02-02 edited clone images. **C.** The percentage of positive nuclei for sense and antisense RNA foci was between 10-12%, with no foci were detected in edited and healthy clones (****$P < 0.0001$, One-Way ANOVA).

It has been known that the expanded sense and antisense RNA sequences may undergo repeat-associated non-ATG translation (RAN) (Ash et al. 2013; Mori, Arzberger, et al. 2013) and accumulation of cytoplasmic repeat-associated dipeptide protein products, identified as poly-Gly-Ala (GA), poly-Gly- Pro (GP), and poly-Gly-Arg (GR). A dot-blot analysis was performed to detect GA, GP and GR proteins using protein samples from all the experimental cell lines, to examine whether the editing of HRE region prevents the formation of toxic DPRs in iPSC-derived MNs. Higher concentration of DPR accumulation was detected in iPSC-derived MNs from C9-02-02 and C9-02-03 disease lines compared to healthy controls and edited lines (Figure 6.7). However, DPR aggregates were relatively low in our patient protein samples, probably due to the presence of DNA promoter hypermethylation and consequently lower levels of transcripts. The antibodies showed some staining in control and edited cell lines, which likely reflects poor specificity of the antibodies. The gene editing of G4C2 region has provided a proof-of-principle that excision of G4C2 hexanucleotide repeats can correct the pathology of nuclear RNA foci and decrease the cytoplasmic dipeptide proteins accumulation in iPSC-derived MNs.

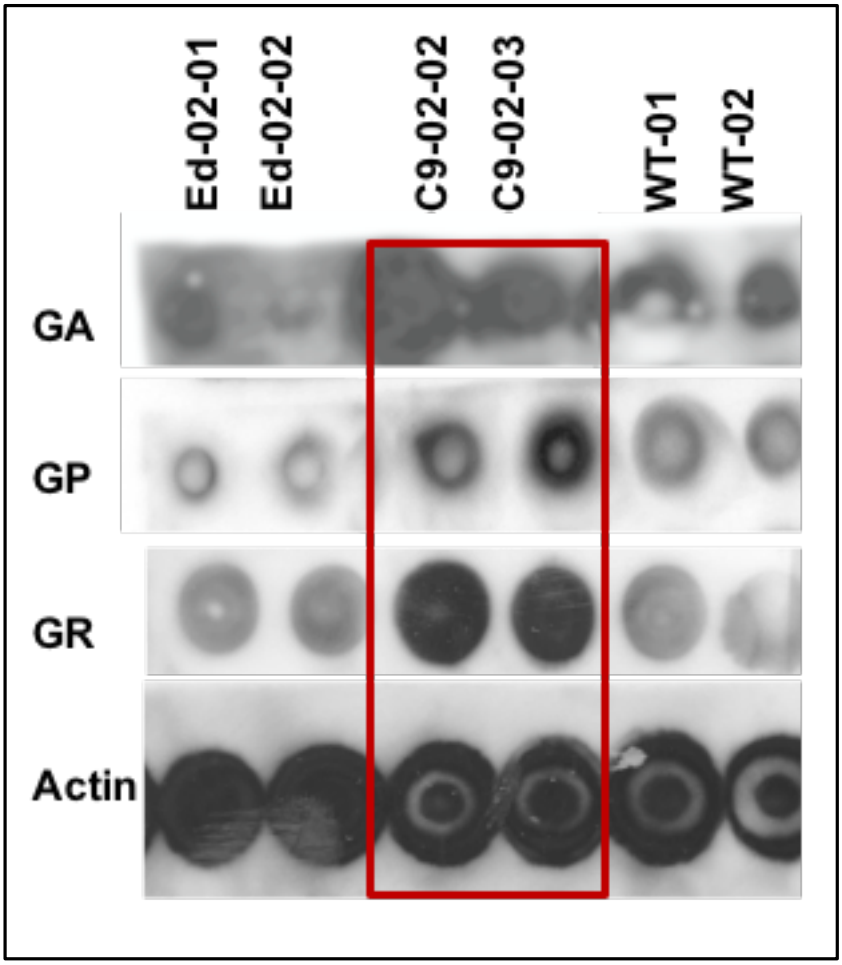

Figure 6.7 Dot-blot analysis of protein samples extracted from edited iPSC-derived MNs. Samples were blotted on nitrocellulose membrane and poly(GA), poly(GP), and poly(GR) dipeptide repeat proteins were detected and showed higher levels of accumulation in C9-02-02 and C9-02-03 disease-related protein samples compared to healthy and edited lines.

6.4 Susceptibility to glutamate toxicity has improved after gene editing

Recently, it has been shown that iPSC-derived MNs from HRE carrying patients are highly susceptible to glutamate-mediated excitotoxicity (Donnelly et al. 2013). Glutamate toxicity has been shown to play a major role in ALS and C9 mutation carrying patients exhibit a loss of astroglial glutamate transporter 1 (GLT-1/EAAT2), which buffers synaptic glutamate and subsequently prevent excitotoxicity (Lin et al. 1998; Renton et al. 2011; Jeffrey D. Rothstein et al. 1995; Donnelly et al. 2013).

To determine whether CRISPR-editing of disease cells lead to altered susceptibility to glutamate excitotoxicity, sensitivity to glutamate was evaluated by treating iPSC-MN cells with different concentrations of glutamate (10μM, 50μM, 100μM and no glutamate treatment) for five hours at 37C° and 5% CO_2. Representative images in Figure 6.8A demonstrate the iPSC-derived MNs from WT-01, Ed-02-02 and C9-02-02 after treatment with 50 μM glutamate. Propidium iodide (PI) staining uptake was performed to stain dying cells. A significant increase in cell death of C9-02-02 and C9-02-03 iPSC-derived MNs was observed and compared to healthy and edited clones in a dose-dependent manner as the following: after treatment with 10 μM glutamate (**$P < 0.001$, One-Way ANOVA), 50 μM (**$P < 0.001$, One-Way ANOVA) and 100 μM (***$P < 0.0001$, One-Way ANOVA) (Figure 6.8.B), using iPSC-derived MNs from three independent differentiation experiments. No significant differences were observed between the edited clones and the healthy controls at any glutamate concentration used ($P > 0.05$, One-Way ANOVA). No significant difference in cell death was observed between any of the experimental groups without glutamate treatment ($P > 0.05$, One-Way ANOVA). This would suggest that CRISPR/Cas9 gene editing has resulted in the normalization of cells susceptibility to glutamate excitotoxicity induced by G4C2 expansion mutation.

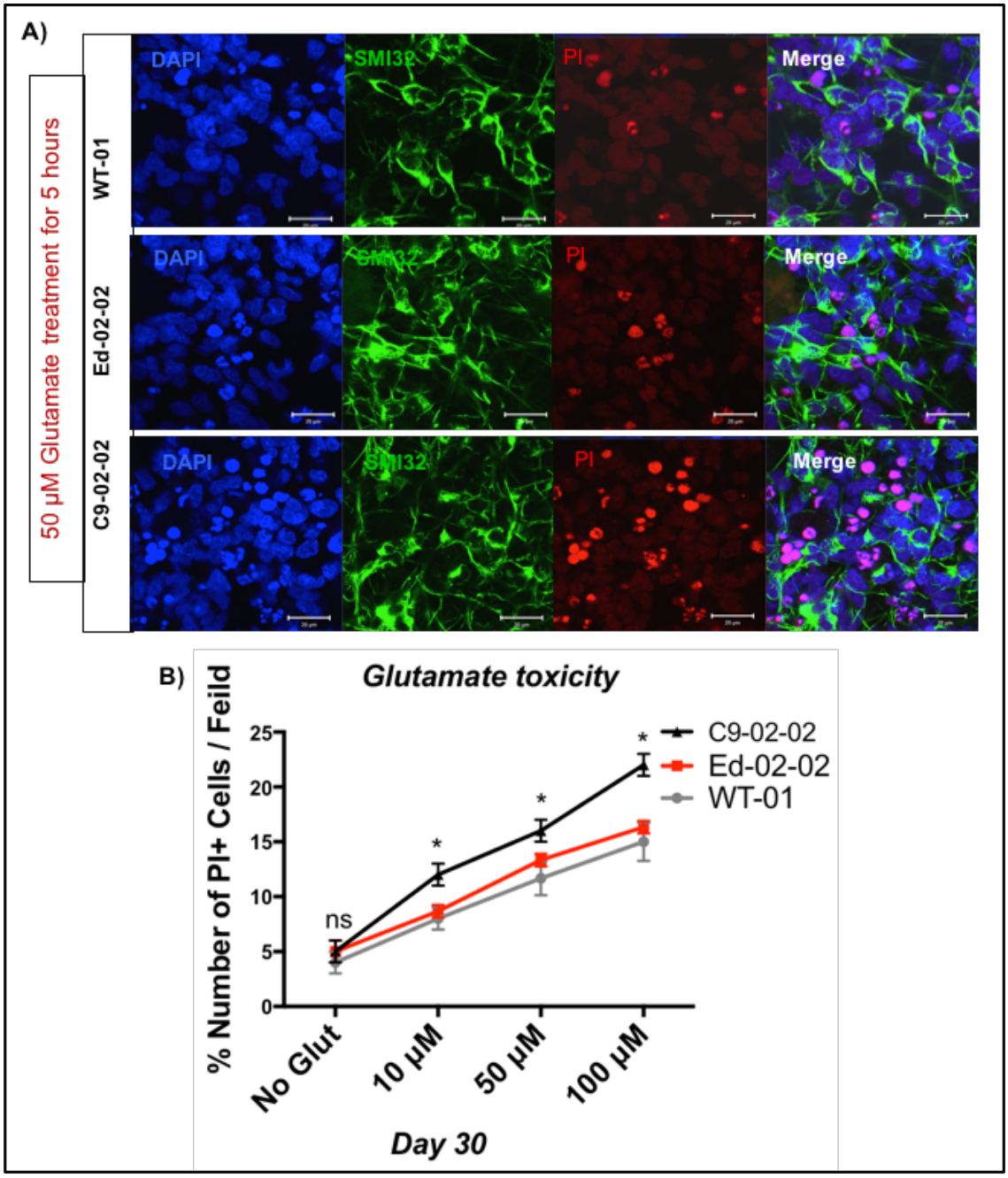

Figure 6.8 Susceptibility of Edited iPSC-derived MNs to glutamate excitotoxicity and evaluation of apoptotic cell death. A) Representative images of glutamate-induced exitotoxicity assay after treating WT-01, Ed-02-02 and C9-02-02 iPSC-derived MNs with 50 µM glutamate. Images reveal that gene editing of Ed-02-02 has rescued the susceptibility of C9-02-02 cells to glutamate toxicity as indicated by the reduction of propidium iodide (PI) uptake. **B)** Statistically significant difference in cell death has been observed on MN culture derived from C9-02-02 cells after treatment with 10 µM glutamate (**$P < 0.001$, One-Way ANOVA), 50 µM (**$P < 0.001$, One-Way ANOVA) and 100 µM (***$P < 0.0001$, One-Way ANOVA) when compared to the healthy (WT-01) and (Ed-02-02) lines in a dose-dependent manner.

6.5 Reduction in markers of cellular stress in gene-edited isogenic MNs

Stress granules (SGs) are cytoplasmic aggregates of nontranslating messenger ribonucleoprotein complexes. They are formed when RNA binding proteins aggregate through their glycine rich domains and they dissociate when conditions return to normal or when they cleared by autophagy (Buchan et al. 2013). SGs function to sequester, silence, or degrade specific RNA transcripts to facilitate the response to cellular stress (Buchan & Parker 2009). PABP is a major component of the stress granules and serves as marker for the activation of stress response. SGs formation has been linked to the pathogenesis of ALS diseases. Oxidative stress is one of the conditions that induce SG formation. In our previous study, both *C9orf72* MNs and CNs showed the formation of poly-A-binding protein (PABP) positive stress granules (Dafinca et al. 2016).

Formation of pathogenic RNA foci and DPR proteins may lead to cellular stress and formation of stress granules (SGs). A recent study has shown that RAN translation may be associated with SG formation after transfecting neuronal cells with poly-GA, poly-GP, and poly-PA (Tao et al. 2015). In this section, the number of polyA-binding protein (PABP) stress granules were counted under the baseline conditions without using stressing agents in iPSC-derived MNs from all of the experimental groups (Figure 6.9.A). A significant decrease in PABP SG number was observed in Ed-02-01 and Ed-02-02 iPSC-derived MNs compared to unedited C9-02-02 and C9-02-03 lines (**$P < 0.001$, one-way ANOVA) (Figure 6.9.B). The accumulation of $PABP^{+}$ SGs suggests the impairment in autophagy and proteasomal degradation pathways, and the convergence of these two dysfunctional pathways may partially explain the vulnerability of *C9orf72* MNs to degeneration (Dafinca et al. 2016).

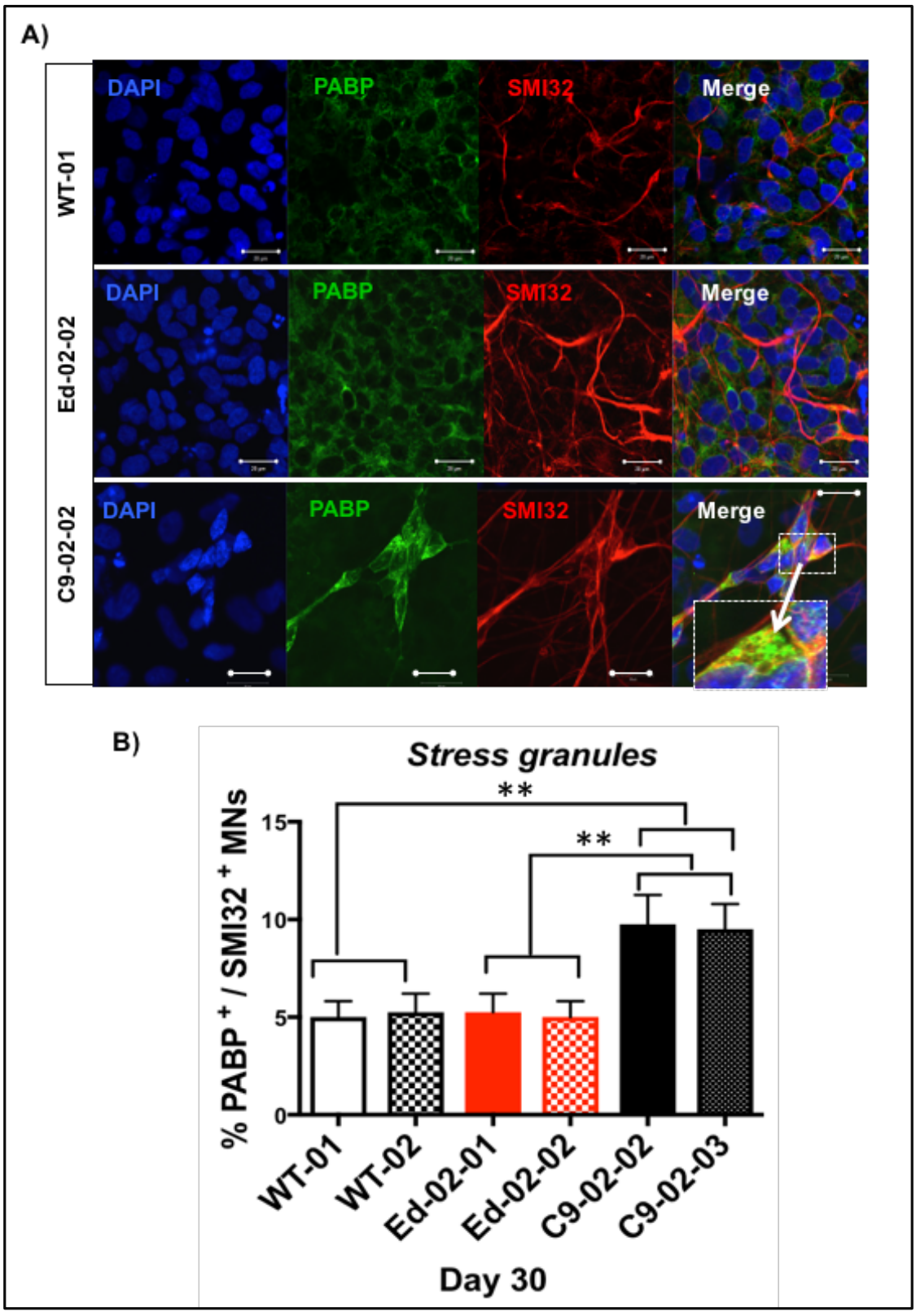

Figure 6.9 Edited iPSC-derived MNs are less susceptible to cellular stress. A) Representative images for PABP SG staining (red) in SMI-32 positive (green) iPSC-derived MNs. **B)** PABP accumulation in SGs was detected at higher frequency in C9-02-02 and C9-02-03 patient iPSC-MNs compared to Ed-02-01 and Ed-02-02 clones in three independent experiments (**$P < 0.001$, One-Way ANOVA).

6.6 Reduction in apoptotic cell death as indicated by immunostaining of cleaved-caspase-3

Caspases are present in the cell as inactive proenzymes which can be activated by proteolytic cleavage. The cleavage of caspase-3 is an essential process to initiate the apoptotic disassembly of the cell. Elevated cleaved caspase-3 levels and cell death upon stress has been observed in previous ALS studies (Hart & Gitler 2012; Dafinca et al. 2016). We have previously shown that iPSC-derived MNs from ALS patients had increased cleaved caspase-3 activity compared to wild-type controls (Dafinca et al. 2016). The frequency of cleaved-caspase-3 expression was examined by co-immunostaining with SMI-32 in MN cultures derived from the same diseased, corrected and wild-type controls cell lines. A significant difference in cleaved caspase-3 frequency was observed in disease clones C9-02-02 and C9-02-03 compared to edited Ed-02-02 and Ed-02-03 MN cultures (**$P < 0.001$, One-Way ANOVA), and WT-01 and WT-02 controls (***$P < 0.0001$, One-Way ANOVA). No significant differences have been observed between edited lines and WT controls, which indicates that the apoptotic cell death activity due to the presence of *C9orf72* mutation has been decreased (**$P < 0.0012$, One-Way ANOVA), suggesting that edited iPSC-derived MNs are less vulnerable to apoptotic cell death triggered by G4C2 toxicity (Figure 6.10.A&B). Taken all together, the presence of the repeat expansion has enhanced the activation of caspase-3 and significantly increased apoptotic cell death. No significant differences were observed between edited and healthy controls.

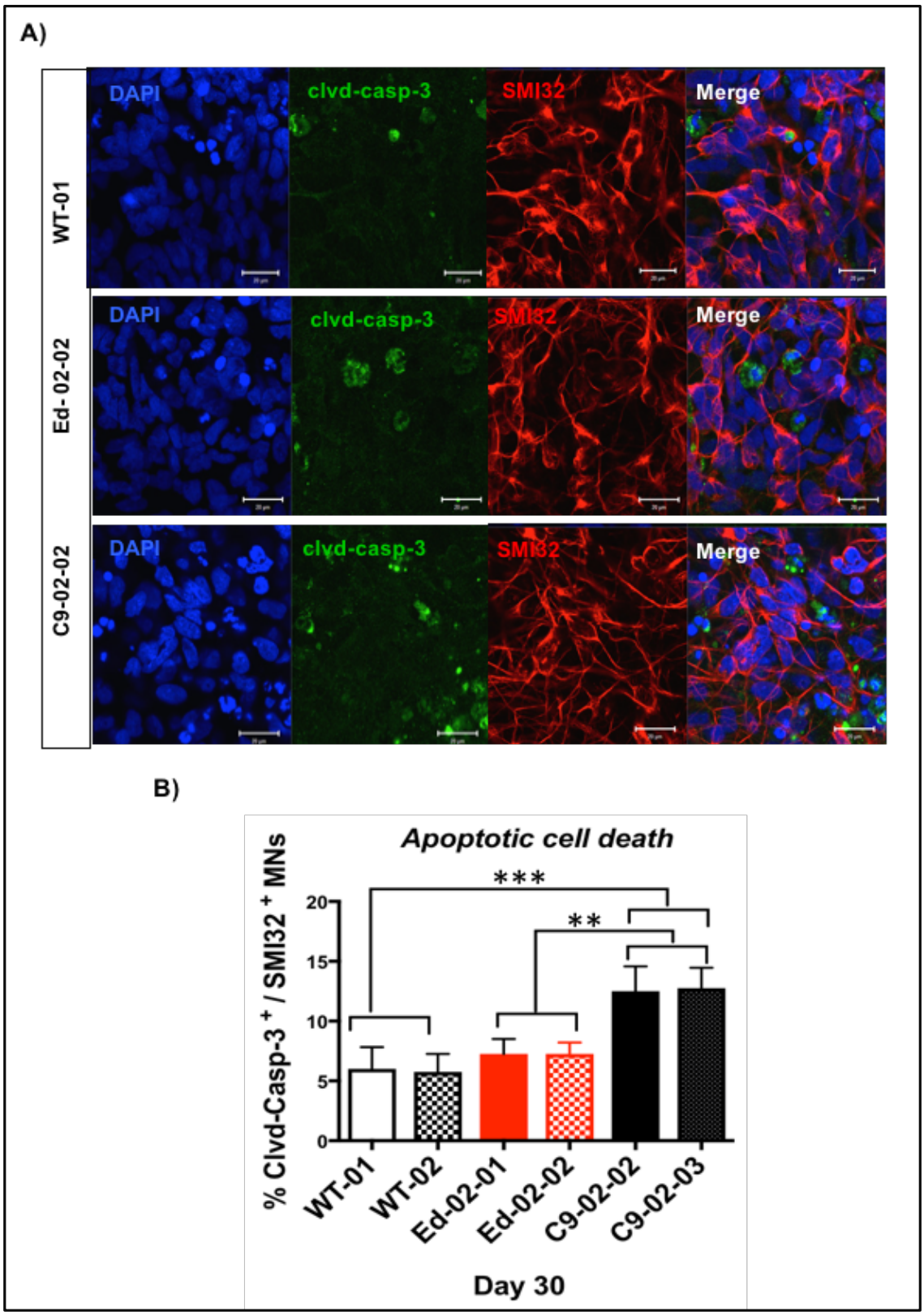

Figure 6.10 Edited iPSC-derived MNs are less susceptible to cell death. A) Representative images of cleaved-caspase-3 apoptotic detection in MN cultures. Cleaved caspase-3 frequency (green), SMI-32 (red), DAPI (blue). **B)** Significant decrease in cleaved caspase-3 frequency in Ed-02-01 and Ed-02-02 cells compared to C9-02-02 and C9-02-03 iPS-MN cultures (**$P < 0.001$, One-Way ANOVA).

6.7 Reduction in G3BP1 SG formation in edited cells under cellular stress conditions

SGs have antioxidant activity that is controlled by two SG components, USP10 antioxidant enzyme and its cofactor G3BP1. Under oxidative stress such as sodium arsenite treatment, SGs formation will be induced and lead to inactivation of the G3BP1 and activation of USP10, which results in a decrease in ROS production and apoptosis. Emerging evidence indicates that SG formation plays an active regulatory role in the response of cells to oxidative stress (Takahashi et al. 2013; Wheeler et al. 2016; Lian & Gallouzi 2009).

iPSC-derived MNs were assessed for SG formation under stress conditions by treating them with sodium arsenite for 1hr 30 min and then immunostaining with G3BP1, a well-known marker for SGs, which aggregates in the cytoplasm and stains positive granules. ChAT was co-stained with G3BP1 to confirm the MN identity in the analyzed cells (Figure 6.11.A). Immunostaining of G3BP1 was also performed on untreated cells but it remained diffuse and did not show any clear SG structure suitable for counting (Figure 6.11.B). SGs were quantified and analyzed in all the experimental cell line groups. Significant differences in the number of SGs were detected between the uncorrected disease lines carrying the mutation and the WT healthy controls (**$p<0.001$) and edited cells (*$p<0.05$) under the stress conditions (Figure 6.11.C). These results might suggest that the CRISPR editing has made cells less sensitive to stress triggered by the presence of the repeat expansion mutation, suggesting that the stress sensitivity might be either related to the presence of RNA foci and the accumulation of DPRs or to *C9orf72* hypermethylation and gene haploinsufficiency.

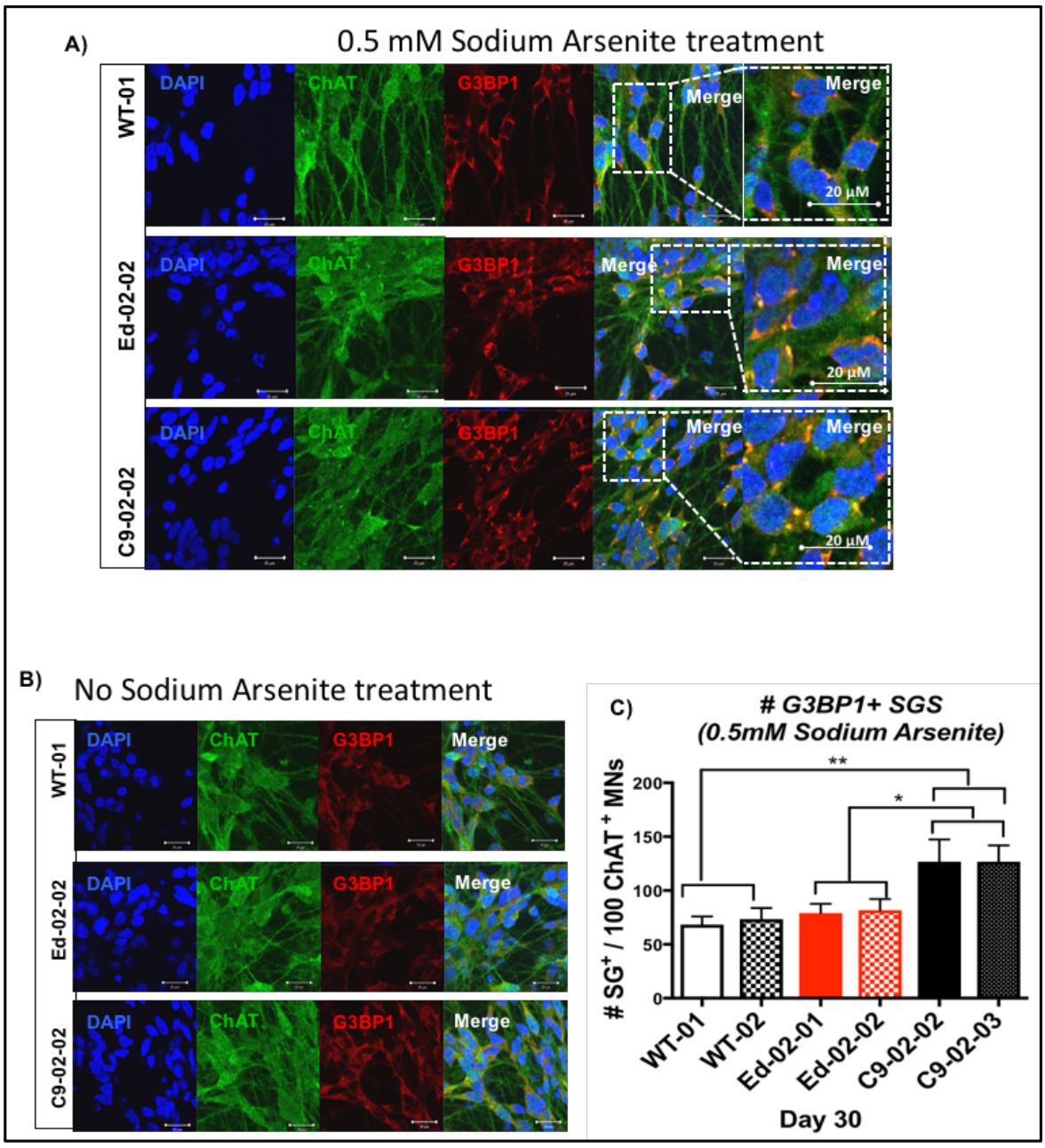

Figure 6.11 G3BP1 cytoplasmic SGs immunostaining after sodium arsenite treatment. A) Treatment of iPSC-derived MNs with 0.5 mM sodium arsenite for 1hr 30 min resulted in stress granule (SGs) formation as indicated by positive immunostaining for ChAT (*green*) and anti-G3BP (*red*). **B)** In unstressed conditions G3BP1 remained diffuse in the cytoplasm and no clear SGs were formed in any of the analysed lines. **C)** The number of G3BP-positive SGs per 100 ChAT+ cells was quantified and analysed.

Summary

In order to connect the genetic mutation with the cause of disease, reversal of previously studied ALS phenotypes was assessed in CRISPR/Cas9-corrected iPSC-derived MNs. Correction of the HRE mutation results in the normalization of loss-of-function phenotypes by significant reversal of DNA promoter hypermethylation levels and correcting the altered C9 gene variant expression levels in iPSC-derived MNs. CRISPR efficiently results in abolishing of RNA foci and decreasing RAN-translation products in iPSC-derived MNs. CRISPR gene editing of the *C9orf72* mutation has also resulted in the reversal of cellular phenotypes, specifically glutamate excitotoxicity, apoptotic cell death and stress granule formation. These phenotypes have previously been associated with the presence of HRE mutation. These results demonstrate that I have created a valuable isogenic model to study ALS/FTD repeat expansion associated disease phenotypes for better understanding of disease mechanisms.

To summarize, the findings of this chapter clearly show that the presence of HRE mutation is directly relevant to the presence of the main disease phenotypes and the correction of HRE mutation can alter C9orf72 variants expression and DNA methylation status to normal levels. In addition, this work highlights on the fact that the methylation is directly related to the presence of RNA foci and DPR proteins, which are the main ALS disease hallmarks and CRISPR can ameliorate these disease pathogenic phenotypes.

7. Chapter Seven

Discussion, conclusion and future perspective

7.1. Discussion

A G4C2 repeat expansion within the *C9orf72* gene has been identified as the most common genetic cause of ALS and FTD (DeJesus-Hernandez et al. 2011; Renton et al. 2011). There are several suggested mechanisms of neurodegeneration related to the presence of the repeat expansions. Loss-of-function and haploinsufficiency mechanism has been suggested to be related to *C9orf72* promoter hypermethylation and downregulation of *C9orf72* gene variants expression. Decreased levels of *C9orf72* transcripts have been reported in ALS/FTD patient tissues carrying the HRE mutation (Belzil, Bauer, et al. 2013; DeJesus-Hernandez et al. 2011), suggesting the haploinsufficiency of *C9orf72* as a potential disease mechanism.

On the other hand, toxic gain-of-function mechanism is related to the formation of toxic RNA foci and the accumulation of DPR protein aggregates in neurons. Bidirectional transcription of the mutated repeat expansion region of *C9orf72* gene generates sense and antisense RNA transcripts that form sense and antisense intranuclear RNA foci (Mizielinska et al. 2013) and potentially sequester RNA binding proteins inside the nucleus (Cooper-Knock, Walsh, et al. 2014). These RNA transcripts can also undergo repeat-associated non-ATG (RAN) translation, producing dipeptide repeat (DPR) proteins that aggregate and form cytoplasmic inclusion bodies (Ash et al., 2013; Mori et

al., 2013; Zu et al., 2011).

The C9orf72 protein function is unknown but it has been reported to have a role in regulating endosomal trafficking. In a previous study, knockdown of *C9orf72* gene in zebrafish results in motorneuron degeneration and motor deficits (Ciura et al. 2013). In another study, toxicity through increased vulnerability to glutamate excitotoxicity and cellular stressors in iPSC-derived MNs have been also linked to the presence of HRE mutation.

The key question regarding the pathogenicity of the HRE mutation is whether the repeat expansion causes the disease due to the loss or gain of function of *C9orf72* gene, or both. The mechanisms of MN degeneration in *C9orf72*-related disease is complicated and it most likely involves both RNA gain-of-function toxicity and loss of function mechanisms.

One of the most important issues for the lack of effective treatment in ALS is the difficulty of producing disease-specific cells *in vitro* that can be used for investigating the underlying pathogenesis and for drug screening strategies. Studies with animal models and neuronal cultures have demonstrated promising results using ASOs to target the repeat expansion. However, more studies need to be conducted to enhance the targeting of the repeat mutation without affecting the *C9orf72* gene expression levels by maintaining normal expression level of the wild-type allele.

iPSC technology has opened an avenue to generate disease-specific neurons from ALS patients to provide an *in vitro* model of cells reflecting the disease phenotypes. Protocols to differentiate iPSCs into MNs are well-established. The differentiation process involves inducing neural patterning and ventralization using small molecules (Wichterle et al.

2002; M W Amoroso et al. 2013; Maury et al. 2014). The ability to generate MNs from iPSCs has allowed the researchers to better understand MN degeneration in ALS at an early-stage of the disease development.

The combination of iPSC technology and genome engineering represents a promising tool to study the phenotypic changes and the pathological mechanisms by generation the isogenic lines from patient-specific iPSCs and studying the underlying disease phenotypes with the same genetic background of the parental disease clones. Repeat expansions in *C9orf72* gene represents an ideal target for gene editing. However, the excision of the repeats is challenging due to the presence of massive repeat expansions in ALS disease cases.

Recently, correction of genetic mutations using programmed nucleases has been successfully achieved on several genetic diseases with abnormal repeat expansion (Yusa et al., 2011, Sebastiano et al., 2011, Yin et al., 2014, Long et al., 2014, Maetzel et al., 2014 and Li et al., 2015). CAG repeat expansion in Huntington's disease patients has been corrected using the traditional homologous recombination (HR) or Cas9-mediated HR methods (An et al. 2012; An et al. 2014). In another study, CRISPR/Cas9 was applied for targeted deletion of CGG repeat expansion in the fragile X mental retardation 1 (FMR1) gene using both embryonic stem cells and iPSCs derived from fragile X (FXS) patients. This resulted in demethylation and restoration of FMR1 mRNA and FMRP protein expression (Park et al. 2015).

In this project, I attempted to correct the abnormal G4C2 repeat expansion in ALS/FTD patient-derived iPSCs using CRISPR/Cas9 and an exogenous donor sequence carrying the normal repeat size. I applied CRISPR/Cas9-HDR mediated gene editing system on

ALS/FTD patient iPSC line (C9-02-02) carrying a massive repeat expansions, to produce the isogenic lines bearing the wild-type repeat size inserted by HDR. The generated isogenic lines have been used in this thesis to investigate if the gene correction can abrogate the HRE related-phenotypes and reverse some of the pathogenic mechanisms involved in ALS disease.

To study the HRE pathology, skin biopsies were isolated from ALS patients carrying the HRE mutation that has been confirmed by repeat-primed PCR and Southern blotting (Renton et al., 2011). I isolated the dermal fibroblasts from the biopsy sections and I reprogrammed them into iPSCs using Sendai viruses carrying the Yamanaka factors (Sox2, Oct4, Klf4, and c-Myc). After that, I successfully assessed them for their pluripotency and karyotyping profiles using a panel of quality control tests. Subsequently, I differentiated two ALS/FTD iPSC lines (C9-02-02 and C9-02-03) derived from C9-02 patient into MNs, to demonstrate their ability of producing MNs before using them in gene editing experiment. The proportion of cells positive for HB9 or Islet-1 MN markers was approximately 15%–20% of the whole heterogeneous cell population. I have also confirmed that iPSCs derived from C9-02 patient retain the stability of the repeat expansion after reprogramming from fibroblast to iPSCs as indicated by southern blotting and RP-PCR, and they have also showed the presence of sense RNA foci by FISH analysis. And finally, I have demonstrated that the generated iPSC clones were free of any Sendai virus's genome after the reprogramming process.

The iPSC reprogramming and differentiation into specific cell types requires multiple steps including proper biological samples collection from patients, isolation and expansion of patient cells, and iPSC reprogramming using standardized protocols. Quality control of the established iPSC lines is mandatory and should include karyotype

analysis of chromosomes, isolation of several clones from the same subject and selection of iPSC lines based on cell morphology, expression of pluripotency markers and the ability to produce three germ layers *in vitro* (Corti et al. 2015).

Recently, it has been reported that using double-nicking CRISPR with a pair of gRNAs to cut the opposite strands of DNA may increase the homologous recombination frequency and reduce off-target mutations (Ran et al. 2013). To correct the HRE mutation, a double-nicking CRISPR system was applied to remove the repeat expansion and replace it with the normal repeat size by HDR using plasmid donor template to initiate a precise gene editing by inserting the wild-type repeat size. Two gRNAs and a donor template carrying the wild-type repeat sequence and a puroΔTK selection cassette were constructed and used for targeting C9-02-02 patient-derived iPSCs. Then drug-resistant clones were picked and screened for successful homologous recombination by RP-PCR and sequencing assays.

The edited clones with successful targeting were further confirmed by Southern blot analysis to confirm the absence of the massive repeat expansion. The generated edited lines showed positive expression of the pluripotency markers and they retained their normal karyotype (46, XX). After the successful characterization, I expanded several biallelic and monoallelic targeted clones and I confirmed the precise gene correction with no off-target effects detected. I then differentiated two iPSC lines of each of the diseased, edited and control lines into motorneurons using previously published protocol (Maury et al. 2014) to study the C9orf72 disease-related phenotypes.

Gene-edited cells showed normal differentiation potential into MNs

Development of efficient differentiation protocols from iPSC to produce homogeneous cell populations is crucial for effective disease modelling studies. Additionally, the presence of disease-related phenotypes in the differentiated cells must be evaluated to see if they are disease specific and reflect the disease mechanisms. For instance, epigenetic instability and DNA hypermethylation in iPSCs is important as many iPSCs aberrantly acquire hypermethylation as a consequence of reprogramming (Gore et al. 2011; Blasco et al. 2011).

Several protocols have been developed for differentiation of MNs from iPSCs. Some protocols are based on EB induction followed by RA/SHH-induced neural patterning and neurosphere development, as described in Chapter 3. But this protocol is tedious and requires up to 2 months to produce functional MNs with low efficiency. More recently, protocols have been developed for rapid and efficient neural induction of MN progenitors from monolayer iPSCs in adherent culture for the efficient generation of more homogeneous population of MNs in a shorter time.

The MN differentiation protocol applied in Chapter 5 involved the earlier addition of RA at day 2 for neural induction and patterning which results in the generation of higher MN population than the first applied EBs-based protocol. This monolayer protocol has also involved the simultaneous inhibition of BMP and Activin signalling for induction of the neural lineage, which results in the production of a higher percentage of PAX6+/Nestin+ and Olig2+ NPs in less than 10 days. The efficiency of cell differentiation into MNs was verified by successful expression of MN markers including Olig2, Islet-1, HB9, SMI-32,

and ChAT. This approach has enabled us to study the disease phenotypes by producing enough MN cells derived from highly condensed NPs rosettes-like structures (Maury et al. 2014; Dafinca et al. 2016). Immunocytochemical analysis of MN markers did not reveal any significant differences in the differentiation efficiency between any of the analysed cell lines. No differences were observed in the ability of isogenic lines to differentiate into MNs when compared to their parental cell lines.

Although the basic steps of MN differentiation protocols are well identified, variations in efficiency can be still available. These variations may arise due to different reasons such as the use of different growth factors at different timing points, cellular heterogeneity or the use of different media composition. The presence of standardised protocol for in vitro disease modelling of ALS disease is crucial. MNs can be assessed according their morphology and ability to express MN marker, according to their electrophysiological activity and their ability to produce functional neuromuscular junctions.

Modeling of C9orf72-disease related phenotypes

It has been reported that repeats number is directly correlated with the severity of the disease (Cooper-Knock, Bury, et al. 2015). We have shown in our recent publication that iPSC-derived MNs from ALS or ALS/FTD patients carrying the HRE mutation can reveal the relevant disease phenotypes (Dafinca et al. 2016). After the successful differentiation of iPSC lines into MNs, I investigated some of the C9/ALS disease related-loss and gain-of-function phenotypes in edited clones and compared them to the original disease lines using iPSC-derived MNs.

Understanding the role of epigenetic events in the disease development may help in the identification of biomarkers and the development of future therapeutics. DNA

hpermethylation has been previously reported in the promoter region of *C9orf72* gene in patients with repeat expansions, which results in the transcriptional silencing of *C9orf72* gene. DNA methylation of *C9orf72* leads to reduced expression of *C9orf72* mRNA (Belzil, Bauer, et al. 2013; Belzil et al. 2014; Liu et al. 2014; Xi et al. 2013). Knockdown of the *C9orf72* homologue in zebrafish results in motorneuron degeneration (Ciura et al. 2013). It has been also reported that *C9orf72* hypermethylation is associated with a protective phenotype in lymphoblast cells with HRE mutation, and this means that *C9orf72* hypermethylation is associated with reduced RNA foci and DPR aggregates formation (Liu et al. 2014). However, it is still unclear whether silencing of *C9orf72* has a detrimental or protective effect.

In this project, I studied the DNA promoter methylation status in *C9orf72* gene on iPSC-derived MNs from the isogenic lines and compared them to the original disease cells. Our C9FTD/ALS patient sample (C9-02-02) was found to be hypermethylated and a rescue of the DNA promoter hypermethylation in edited clones was observed, as indicated by methylation-specific Q-PCR, which has been recently used for the detection of DNA hypermethylation in expansion carriers.

To confirm the methylation results, we performed bisulfite sequencing on DNA samples from the same research groups. DNA samples were treated with bisulfite, which converts cytosine residues to uracils without converting the 5-methylcytosine residues. After PCR amplification, amplified products were sequenced and analysed. All uracil residues were read as thymines in edited and healthy controls while all 5-methylcytosine residues were read as cytosines in the original disease lines (Xi et al. 2013). These findings confirm the results obtained by QPCR and restriction enzymes digestion.

Consequently, the promoter demethylation by CRISPR gene editing has resulted in the normalization of RNA expression levels of *C9orf72* gene variants in iPSC-derived MNs of the isogenic lines to levels similar to the wild-type controls. This suggests that the DNA methylation status of the *C9orf72* promoter is directly dependent on the presence of the repeat expansion mutation, which confirms that the presence of HRE mutation induces the downregulation of *C9orf72* gene variants.

ALS patient cells with HRE mutation are characterized by the nuclear accumulation of sense and antisense RNA foci and the sequestration of RNA-binding proteins, leading to transcriptome defects, and generation of sense and antisense RAN proteins (Gendron et al. 2014). The normalization of methylation levels was also associated with the abolition of nuclear and cytoplasmic RNA foci. Dot blot analysis using high quantities of the protein samples revealed a higher frequency of GA, GP, and PR toxic RAN DPR aggregates in the patient lines than the other research group cell lines. The abundance of RAN translation products was relatively low in diseases lines, probably due to C9orf72 gene silencing resulted from DNA hypermethylaion and consequently the production of less repeat-containing mRNA variants. However, there is a clear difference between the diseased lines and the edited clones.

The presence of RNA foci and RAN translation products in iPSCs-derived human neurons with HRE suggests that these cells are recapitulating two major neuropathological features of patients with HRE mutation and both phenotypes seem to show a correlation with the methylation level on this patient lines and can be rescued using gene editing system. However, more patient lines need to be included to confirm the specificity of these phenotypes in disease pathogenicity.

It has been demonstrated that poly-GA caused toxicity when expressed in cultured cells and primary neurons as measured by increased release of LDH, caspase-3 activation and endoplasmic reticulum stress (Zhang et al. 2014). Cleaved caspase-3 frequency under baseline conditions was also significantly lower in the edited cells which may indicate that the abolition of the pathogenic RNA foci and DPRs has an influence in decreasing apoptotic cell death in which the presence of the repeats would alter the apoptotic cell death activity and leads to an increase in cell death. CRISPR correction of HRE has also resulted in a rescue of cells vulnerability to glutamate excitotoxicity in a dose-dependent manner.

In a previous study they have observed that iPSC-derived MNs from ALS patients with HRE mutation are more sensitive to cellular stress induced by autophagy inhibitors (Almeida et al. 2013). Decreased cell viability of the neurons might indicate the sensitivity of these cells to stress. Recent studies showed that arginine-rich DPR proteins were cytotoxic when non-neuronal cells such as astrocytes were incubated with these synthetic peptides by impairing the biogenesis of ribosomal RNA (Kwon et al. 2014), and also they showed toxicity in vivo in *Drosophila* models (Mizielinska et al. 2014). However, more studies are required to investigate the individual RAN proteins toxicity on neuronal cultures derived from disease and isogenic lines to confirm if the presence of the DPR proteins are sufficient to cause disease pathogenicity and neurodegeneration and if the presence of DPR is the main disease hallmark or if it also depends on the contribution of other pathogenic disease mechanisms.

It has also been reported in neuronal cells transfected with poly-GA, poly-GP, and poly-PA that RAN translation may be associated with stress granule formation and the induction of neurodegeneration (Tao et al. 2015). In this thesis, I have observed that correction of the repeat size using CRISPR resulted in a reduction of the number of stress granules in iPSC-derived MNs from edited clones compared to the parental disease cells under baseline and cellular stress conditions as indicated by immunostaining of PABP and G3BP1 antibodies. SGs are specifically induced upon cellular stress, which triggers translational silencing of several pathways.

Two independent studies have observed the accumulation of DPRs and RNA foci in transgenic mice with C9orf72 repeat expansion mutation, but without developing neurodegeneration or motor abnormalities (O'Rourke et al. 2015; Peters et al. 2015). These results could suggest that C9orf72-associated pathology can occur presymptomatically and additional pathological mechanisms are required to induce the neurodegenerative process. In this project, the presence of RNA foci and accumulation of RAN-proteins suggest an RNA-gain of function-mediated toxicity. The presence of DNA hypermethylation with reduced *C9orf72* levels would also suggest the importance of the loss-of-function mechanism as one of the main ALS hallmarks of ALS disease development.

According to the results described in this thesis, iPSC-derived MNs can recapitulate the disease-specific cellular phenotypes that are typically seen in human post-mortem tissues of C9orf72-related brain samples. Altogether, I have provided a clear evidence that the DNA hypermethylation is directly correlated with the presence of the abnormal repeat expansions, and the reversal of the methylation levels results in restoration of C9 mRNA variants expression. Results of this thesis have also demonstrated the feasibility of

CRISPR/HDR for the correction of HRE mutation and normalisation of the disease abnormal phenotypes.

The correction method I have applied here can be also applied on other C9orf72 patient-specific iPSCs who carry repeat expansions with different sizes and with variable methylation and C9 expression levels, to try to correlate the methylation levels with ALS disease severity and progression. Further work on edited clones should be done to understand the other underlying disease mechanisms involved in ALS disease pathogenicity.

7.2. Conclusion

Understanding HRE mutation associated disease mechanisms requires cellular or animal models carrying repeat expansions. However, The generation of long G4C2 repeat sequence for disease modelling in animals is technically challenging. iPSCs technology allows us to study diseases using patient-derived cells carrying the same disease-specific mutations.

The combination between iPSCs and CRISPR/Cas9 gene editing technology represents an excellent approach for generating disease models for studying monogenic and complex genetic disorders in familial and sporadic diseases. This discovery can also facilitate the production of site-specific genetic changes in iPSC lines, including gene modifications in normal and disease cell lines. The combination of both approaches has also enabled the generation of isogenic iPSC lines that can be distinguished from the original lines only by alterations within the modified target locus.

In this project, I found that CRISPR/Cas9-mediated HDR can precisely edit the HRE mutation in *C9orf72* gene of ALS/FTD patient iPSCs. C9-02-02 iPSC-derived MNs that carry the mutation showed their ability to recapitulate certain phenotypes related to the presence of the mutation, including DNA promoter hypermethylation and RNA toxicity-related phenotypes. These disease phenotypes were restored in iPSC-derived MNs from the edited lines when compared to the original disease lines. These results could provide an evidence that the presence of HRE mutation is directly responsible for the loss- or gain-of-function disease mechanisms and the correction of the repeat size could ameliorate the disease-associated pathways. This could help in understanding the relative contribution of the underlying loss and gain of function mechanisms of *C9orf72* gene and

in finding possible therapeutic targets by performing high-throughput drug screening on disease-specific iPSC-derived MNs and directly comparing them to the edited lines. The results described in this thesis suggest that reversal of disease phenotype can help in eliminating the disease pathogenicity and highlighting the importance of recapitulating and studying disease phenotypes using iPSC in combination with CRISPR/Cas9 system for efficient gene modification results.

7.3. Limitations and future perspective

With increasing life expectancy, the number of patients with neurodegenerative diseases will grow which will increase the demand on finding a cure and effective therapy for neurodegenerative diseases. Efforts have been made to develop animal models for HRE mutation. However, the problem of species differences is a major issue especially when developing therapeutics for ALS diseases.

Strategies to block the expression of HRE toxic RNA and RAN products need to be developed to restore the normal function of the *C9orf72* gene as well as to mitigate the RNA-mediated toxicity resulted from the accumulation of repeat-RNA sequences. Targeting strategies to increase *C9orf72* expression from the normal allele and the suppression of mutant *C9orf72* allele are required to decrease the pathogenicity of the mutant allele only and to protect the normal expression level of *C9orf72* gene.

iPSCs hold a great promise for their capacity to differentiate into any cell types, which makes them a good candidate for *in vitro* disease modeling. However, iPSCs have several limitations. One of the major limitations is the phenotypic inconsistency and clonal variations between iPSC lines differentiated from the same individual. Optimization of the reprogramming and differentiation protocols is required to limit these clonal variability.

Most of the optimized MN differentiation protocols result in the production of heterogeneous population of neuronal and non-neuronal cells. Improvement of theses protocols is required to produce homogeneous functional MNs from iPSCs, which is crucial for their use in research studies. However, the use of iPSCs is limited to specific cell types for which efficient differentiation protocols are established. Another important

issue is that *in vitro* differentiation of iPSC only reflects certain aspects of the disease due to the production of specific type of cells which might do not represent the heterogeneous diseases where more than one cell type is involved in the disease pathogenicity. To address this problem, co-culture and heterogeneous cellular models need to be developed for more accurate and extensive disease modeling studies. Further improvements are also required to enhance the specificity and efficiency of CRISPR and to develop safe and effective methods of delivery into cells to eliminate the risk of off-target mutations.

Building on the results of this thesis, further investigation of the transcriptomic profiles of the edited clones compared to the original cells will facilitate our understanding of the earliest ALS disease pathways by providing an insight into transcriptional and post-transcriptional mechanisms regulating the disease. These include molecular characterization of iPSC-derived MNs from all the research group cell lines. More isogenic cell lines from different ALS patients need to be included in this study, to confirm the specificity of the phenotypes observed here and to determine their contribution in ALS disease development. Additionally, studying the neurotoxicity of DRPs independently in absence of the abnormal HRE is crucial, to determine if RAN products are the hallmark of C9/ALS related disease pathogenicity.

Investigating other ALS pathogenic disease pathways is also required, to elucidate disease-specific cellular mechanisms and to allow a full understanding of the key pathogenic mechanisms directly related to the presence of HRE mutation. This could help in the development of effective treatments for ALS. For example, it would be interesting to investigate mitochondrial function using the isogenic lines to improve our

understanding of the role of this organelle in disease development and to allow the investigation of the mechanisms that regulate mitochondrial mechanisms in ALS-patient specific iPSCs and the role of the HRE in developing the disease phenotypes when compared to the isogenic lines phenotype.

It has been reported that G4C2 leads to the accumulation of p62 and the neurons are more sensitive to autophagy inhibitors, which might be related to the impaired autophagy pathway. Further work need to be conducted to study the autophagic pathway with autophagy inhibitors to confirm if the HRE mutation interferes with autophagy function and to see if the neuronal cells are sensitive to autophagy inhibitors.

Most of the potential therapeutics have failed in clinical trials for ALS with no effective treatment is available to date. Drug screening for the compounds that reverse the disease phenotypes in ALS disease models is important to find potential therapies that interfere with specific pathogenic disease mechanisms involved in ALS disease. iPSCs from diseased and isogenic lines can serve as a source for screening of potential drug candidates to target the HRE mutation and to predict the compounds toxicity and cellular response *in vitro* (Giri & Bader 2015).

Appendices

Appendix I: Sources and dilutions of antibodies used in this study.

Antibody	Source	Secondary Antibody	Dilution	Application
Islet-1	DSHB	Alexaflour 568 goat Anti-Mouse	1:200	ICC
TUJ1	Covance	Alexaflour 488 goat Anti-rabbit	1:1000	ICC
HB9	DSHB	Alexaflour 568goat Anti-Mouse	1:50	ICC
Olig2	R&D	Alexaflour 568 DNKY Anti-Mouse Alexaflour 488 DNKY Anti-rabbit	1:1000	ICC
PAX6	R&D	Alexaflour 488 goat Anti-rabbit	1:1000	ICC
Nestin	R&D	Alexaflour 568goat Anti-Mouse	1:1000	ICC
SMI-32	R&D	Alexaflour 568goat Anti-Mouse	1:1000	ICC
Tuj-1	Dako	Alexaflour 568 DNKY Anti-Mouse	1:1000	WB
ChAT	Millipore	Alexaflour 568 DNKY Anti-Mouse	1:500	WB
SSEA-4	R&D	Alexaflour 633 Anti-Mouse	1:50	ICC
IgG3 (SSEA-4 isotype)	R&D	Alexaflour 633 Anti-Mouse	1:50	ICC
Oct4	R&D	Alexaflour 488 Anti-Rat	1:50	ICC
IgG (Oct4 Isotype)	R&D	Alexaflour 488 Anti-Rat	1:50	ICC
Tra-1-60	Biolegend	-	1:100	FC and ICC
IgM isotype	Biolegend	-	1:100	FC and ICC
Nanog	Cell Signaling	-	1:150	FC and ICC
IgG isotype	Cell Signaling	-	1:150	FC and ICC

Appendix II: gRNAs oligonucleotide target sequences on C9*orf*72 gene used to make the gRNA vectors.

gRNA/Target site location	Target site sequence (5'>3') PAM in Underlined	Oligo 1 (5'>3')	Oligo 2 (5'>3')
gRNA1	AAGTAGTGGGGA GAGAGGGTGGG	CACCGAGTAGTGGG GAGAGAGGGT	AAACACCCTCTCT CCCCACTACTC
gRNA2	GCTCTCACAGTAC TCGCTGAGGG	CACCGCTCTCACAG TACTCGCTGA	AAACTCAGCGAGT ACTGTGAGAGC

Appendix III: Sequences of primer sets used in this thesis.

Primer name	Primer sequence (5'>3')	Experiment
SNP-F1 *SNP-R1* *SNP-F2* *SNP-R2*	F-TATGGCCTGCCCAGAAGC R-CGCGCGACTCCTGAGTTC F-GTGGCGAGTGGGTGAGTG R-CTTTCCACAACAGGAGCTGC	SNP detection, direct sequencing and donor template amplification.
HDR-F *HDR-R*	F-TATGGCCTGCCCAGAAGC R-CTTTCCACAACAGGAGCTGC	Donor template amplification
SEQ-F *SEQ-R*	F-GTGGCGAGTGGGTGAGTG R-CTTTCCACAACAGGAGCTGC	Direct sequencing. Of the donor template
Surv-F *Surv-R*	F-ACTGCCTACCAAGCACAAACAA R-GAAGGAGACAGCTCGGGTACT	Cleavage assays (Surveyor and T7E1)
SDM-F *SDM-R*	F_ATGCATGTGAACAAGAAAAGACCTGATAAAG R-TCAGCGAGTACTGTGAGAGC	Site-directed mutagenesis
Loxp/Puro/tk-F *Puro-tk-R*	F_CAACCGCAGCCTGTAGATAACTTCGTATAATGTATGCTATACGAAGTTATAGGAGTGGGAATTGGCTCC R_CATAGAGCCCACCGCATCC	Gibson assembly experiment.
Loxp/Donor-F *Donor-R*	F_TGCGGTGGGCTCTATGATAACTTCGTATAATGTATGCTATACGAAGTTATCAAGCTCTGGAACTCAGGAGTC R_CTACAGGCTGCGGTTGTTTC	Gibson assembly experiment.
Sev	F_GGATCACTAGGTGATATCGAGC R_ACCAGACAAGAGTTTAAGAGATATGTATC	Sendai virus clearance
c-Myc	F_TAACTGACTAGCAGGCTTGTCG R_TCCACATACAGTCCTGGATGATGATG	Sendai virus clearance
Klf4	F_TTCCTGCATGCCAGAGGAGCCC R_AATGTATCGAAGGTGCTCAA	Sendai virus clearance
Sox2	F_ATGCACCGCTACGACGTGAGCGC RE_AATGTATCGAAGGTGCTCAA	Sendai virus clearance
Oct4	F_CCCGAAAGAGAAAGCGAACCAG	Sendai virus clearance

	R_AATGTATCGAAGGTGCTCAA	
RP1-F *RP1-R* *RP1-Anchor*	F_FAM_TGTAAAACGACGGCCAGTAGCTC TGGAACTCAGGAGTCGCG R_CAGGAAACAGCTATGACCGGGCCCGC CCCGACCACGCCCCGGCCCCGGCCCCGG Anchor-CAGGAAACAGCTATGACC	RP-PCR1
RP2-F *RP2-Anchor* *RP2-R*	F_TACGCATCCCAGTTTGAGACGGGGGCC GGGGCCGGGGCCGGGG F_TACGCATCCCAGTTTGAGACG R_FAM-AGTCGCTAGAGGCGAAAGC	RP-PCR2
5'-Out-F *EF1a-R*	F_ AGGTGAGTGGATTATGGGGTGG R_ACTACCCCCGTCCGATTCTC	5' In-out
Puro-F *3'-Out-R*	F-ATCCATGCCCACGCTACTG R_CACCTTCTCCAACCTGGCTC	3'In-out
Meth-F	5′-CAGTGTGAAAATCATGCTTGAGAGA-3′	Methylation
Meth-R	5′-TTTGTGCTTGGTAGGCAGTG-3′	Methylation
exon 2F	CCCACTTCATAGAGTGTGTGTTG	C9ORF72 ALL
exon 3R	TTCCATTCTCTCTGTGCCTTC	C9ORF72 ALL
exon 4F	GAAATCACACAGTGTTCCTGAAGAA	C9ORF72 SHORT
exon 5-UTR-R	ATCTGCTTCATCCAGCTTTTATGA	C9ORF72 SHORT
exon 8F	CATGGCTCAGGATACGATCA	C9ORF72 LONG
exon 9R	GGAAGGCTTTCACTAGAGTGTCTC	C9ORF72 LONG
exon 1AF	GGGTCTAGCAAGAGCAGGTG	C9ORF72 EX1A
exon 2R	CGACATCACTGCATTCCAAC	C9ORF72 EX1A
Exc-F	ACCGCAGCCTGTAGATAACTTC	Cassette excision
Exc-R	AGTTCCAGAGCTTGATAACTTCGT	Cassette excision

Appendix IV: The completed DNA donor template sequence.

pGEM-T Easy plasmid sequence

Right homology arm

Left homology arm

Exons 1a and 1b

Repeats region

gRNA regions

NsiI sites

LoxP sites

Ef1α Promoter

Puro/TK

BGH/P(A)

Primer locations are underlined

GGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGA

Ttatggcctgcccagaagctgatccagccatgcttcttgtacagcctgcagaactgtga
gccattaaacttttctttataaattacccagtttcagttatttctttatagcagtgtaa
gaatggactaacacaattattaacgctagtcctcatgttgtacattaaatctctagatg
tattagacgtaactgcaactttgtaccctaccctacaattttctttccccccaagcccc
ccaaccaagggtctactctgtttctataaattcagttgtttttaattccacgtataag
tgaagtacaactcagtgtagaaacttggtaaatgctagctacttgttataagctgtcag
tcaaaataaaaatacagagatgaatctctaaattaagtgatttatttgggaagaaagaa
ttgcaattagggcatacatgtagatcagatggtcttcggtatatccacacaacaaagaa
aaggggaggttttgttaaaaaagagaaatgttacatagtgctctttgagaaaattcat
tggcactattaaggatctgaggagctggtgagtttcaactggtgagtgatggtggtaga
taaaattagagctgcagcaggtcattttagcaactattagataaaactggtctcaggtc
acaacgggcagttgcagcagctggacttggagagaattacactgtgggagcagtgtcat
ttgtcctaagtgcttttctaccccctaccccactattttagttgggtataaaaagaat
gacccaatttgtatgatcaactttcacaaagcatagaacagtaggaaaagggtctgttt
ctgcagaaggtgtagacgttgagagccattttgtgtatttattcctccctttcttcctc
ggtgaatgattaaaacgttctgtgtgatttttagtgatgaaaaagattaaatgctactc

actgtagtaagtgccatctcacacttgcagatcaaaaggcacacagtttaaaaaaccttt
gtttttttacacatctgagtggtgtaaatgctactcatctgtagtaagtggaatctatac
acctgcagaccaaaagacgcaaggtttcaaaaatctttgtgtttttacacatcaaaca
gaatggtacgttttcaaaagttaaaaaaaaacaactcatccacatattgcaactagc
aaatgacattccccagtgtgaaaatcatgcttgagagaattcttacatgtaaaggcaaa
attgcgatgactttgcaggggaccgtgggattcccgcccgcagtgccggagctgtccc
ctaccagggtttgcagtggagttttgaatgcacttaacagtgtcttacggtaaaaacaa
aatttcatccaccaattatgtgttgagcgcccactgcctaccaagcacaaacaaaacca
ttcaaaaccacgaaatcgtcttcactttctccagatccagcagcctcccctattaaggt
tcgcacacgctattgcgccaacgctcctccagagcgggtcttaagataaaagaacagga
caagttgccccgccccatttcgctagcctcgtgagaaaacgtcatcgcacatagaaaac
agacagACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAG
TTCCGCCCACGTAAAAGATGACGCTTGGTGTGTCAGCCGTCCCTGCTGCCCGGTTGCTT
CTCTTTTGGGGGCGGGGTCTAGCAAGAGCAGGTGTGGGTTTAGGAGgtgtgtgttttg
ttttcccaccctctctccccactacttgctctcacagtactcgctgagATGCATGTGA
ACAAGAAAAGACCTGATAAAGATTAACCAGAAGAAAACAAGGAGGGAAACAACCGCAGC
CTGTAGATAACTTCGTATAATGTATGCTATACGAAGTTATAGGAGTGGGAATTGGCTCC
GGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGG
GGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATG
TCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTA
GTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGT
GTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTA
CTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGG
GAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTG
GCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTG
CTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTC
TGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGG
GGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCGGCGAGGCGGGGCC
TGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTG
GTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTC
GGCACCAGTTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAA
AATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGG
GCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAG
GCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGTTGGGGGGAGGGGT

TTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGG
CACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCT
CAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTCAGGTGTCGTGAGGAATTT
CGACTAGTGGCCACAACCATGGGGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCG
CGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCA
CGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTC
TTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGC
GGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCG
GCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGC
CTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTC
GCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGG
CCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTC
TACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCAC
CTGGTGCATGACCCGCAAGCCCGGTGCCGGATCCATGCCCACGCTACTGCGGGTTTATA
TAGACGGTCCTCACGGGATGGGGAAAACCACCACCACGCAACTGCTGGTGGCCCTGGGT
TCGCGCGACGATATCGTCTACGTACCCGAGCCGATGACTTACTGGCAGGTGCTGGGGGC
TTCCGAGACAATCGCGAACATCTACACCACACAACACCGCCTCGACCAGGGTGAGATAT
CGGCCGGGGACGCGGCGGTGGTAATGACAAGCGCCCAGATAACAATGGGCATGCCTTAT
GCCGTGACCGACGCCGTTCTGGCTCCTCATATCGGGGGGGAGGCTGGGAGCTCACATGC
CCCGCCCCCGGCCCTCACCCTCATCTTCGACCGCCATCCCATCGCCGCCCTCCTGTGCT
ACCCGGCCGCGCGATACCTTATGGGCAGCATGACCCCCCAGGCCGTGCTGGCGTTCGTG
GCCCTCATCCCGCCGACCTTGCCCGGCACAAACATCGTGTTGGGGGCCCTTCCGGAGGA
CAGACACATCGACCGCCTGGCCAAACGCCAGCGCCCCGGCGAGCGGCTTGACCTGGCTA
TGCTGGCCGCGATTCGCCGCGTTTACGGGCTGCTTGCCAATACGGTGCGGTATCTGCAG
GGCGGCGGGTCGTGGCGGGAGGATTGGGGACAGCTTTCGGGGACGGCCGTGCCGCCCCA
GGGTGCCGAGCCCCAGAGCAACGCGGGCCCACGACCCCATATCGGGGACACGTTATTTA
CCCTGTTTCGGGCCCCCGAGTTGCTGGCCCCCAACGGCGACCTGTACAACGTGTTTGCC
TGGGCCTTGGACGTCTTGGCCAAACGCCTCCGTCCCATGCACGTCTTTATCCTGGATTA
CGACCAATCGCCCGCCGGCTGCCGGGACGCCCTGCTGCAACTTACCTCCGGGATGGTCC
AGACCCACGTCACCACCCCCGGCTCCATACCGACGATCTGCGACCTGGCGCGCACGTTT
GCCCGGGAGATGGGGGAGGCTAACTGAGCTCTAGATGAAACGATATGGGCTGAATAACT
AGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA
ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTG

GGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGT
GGGCTCTATGATAACTTCGTATAATGTATGCTATACGAAGTTATcaagctctggaactc
aggagtcgcgcgctaggggccggggccggggccggggcgtggtcggggcgggcccgggg
gcgggcccggggcggggctgcggttgcggtgcctgcgcccgcggcggcggagGCGCAGG
CGGTGGCGAGTGgtgagtgaATGCATaggcggcatcctggcgggtggctgtttggggtt
cggctgccgggarcgggtagaagcgggggctctcctcagagctcgacgcatttttactt
tccctctcatttctctgaccgaagctgggtgtcgggctttcgcctctagcgactggtgg
aattgcctgcatccgggccccgggcttcccggcggcggcggcggcggcggcggcgcagg
gacaagggatggggatctggcctcttccttgctttcccgccctcagtacccgagctgtc
tccttcccggggacccgctgggagcgctgccgctgcgggctcgagaaaagggagcctcg
ggtactgagaggcctcgcctgggggaaggccggagggtgggcggcgcgcggcttctgcg
gaccaagtcggggttcgctaggaacccgagacggtccctgccggcgaggagatcatgcg
ggatgagatgggggtgtggagacgcctgcacaatttcagcccaagcttctagagagtgg
tgatgacttgcatatgagggcagcaatgcaagtcggtgtgctccccattctgtgggaca
tgacctggttgcttcacagctccgagatgacacagacttgcttaaaggaagtgactatt
gtgacttgggcatcacttgactgatggtaatcagttgtctaaagaagtgcacagattac
atgtccgtgtgctcattgggtctatctggccgcgttgaacaccaccaggctttgtattc
agaaacaggagggaggtcctgcactttcccaggaggggtggcccttcagatgcaatcg
agattgttaggctctgggagagtagttgcctggttgtggcagttggtaaatttctattc
aaacagttgccatgcaccagttgttcacaacaagggtacgtaatctgtctggcattact
tctacttttgtacaaaggatcaaaaaaaaaaaagatactgttaagatatgatttttctc
agactttgggaaacttttaacataatctgtgaatatcacagaaacaagactatcatata
ggggatattaataacctggagtcagaatacttgaaatacggtgtcatttgacacgggca
ttgttgtcaccacctctgccaaggcctgccactttaggaaaaccctgaatcagttggaa
actgctacatgctgatagtacatctgaaacaagaacgagagtaattaccacattccaga
ttgttcactaagccagcatttacctgctccaggaaaaaattacaagcaccttatgaagt
tgataaaatattttgtttggctatgttggcactccacaatttgctttcagagaaacaaa
gtaaaccaaggaggacttctgtttttcaagtctgccctcgggttctattctacgttaat
tagatagttcccaggaggactaggttagcctacctattgtctgagaaacttggaactgt
gagaaatggccagatagtgatatgaacttcaccttccagtcttccctgatgttgaagat
tgagaaagtgttgtgaactttctggtactgtaaacagttcactgtccttgaagtggtcc
tgggcagctcctgttgtggaaagATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACC
ATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTA
AATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC

AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAG
TGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTG
TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC
TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCT
CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
AGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCT
CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAAC
CACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC
TCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT
AAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCC
ATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAA
TAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCC
ATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC
AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGT
GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAA
GATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG
CGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCT
TTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAA
GGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATT

GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAA
AATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGATGCGGTGTG
AAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAGCGTTAATA
TTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC
GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGT
TCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAA
AAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTG
GGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGC
TTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGG
GCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCG
CTTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAA
GGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGC
AAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGG
CCAGTGAATTGTAATACGACTCACTATA

Appendix V: Representative chromatograms of a sequence from the promoter region.

>CpG1

ggagctgtcccctaccagggtttgcagtggagttttgaatgcacttaacagtgtcttacggtaaaaacaa
aatttcatccaccaattatgtgttgagcgcccactgcctaccaagcacaaacaaaaccattcaaaaccac
gaaatcgtcttcactttctccagatccagcagcctcccctattaaggttcgcacacgctattgcgccaac
gctcctccagagcgggtcttaagataaaagaacaggacaagttgcccgccccatttcgctagcctcgtg
agaaaacgtcatcgcacatagaaaacagacagACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCG
TCAAACAGCGACAAGTTCCGCCCACGTAAAAGATGACGCTTGgtgtgtcagccgtccctgctgccccggtt
gcttctcttttgggggcgggctctagcaagagcaggtgtgggttta

>CpG1 after bisulfite cloning

ggagttgttttttattagggtttgtagtggagttttgaatgtatttaatagtgttttatggtaaaaataa
aattttatttattaattatgtgttgagtgtttattgtttattaagtataaataaaattatttaaaattat
gaaatgtgttttatttttttttagatttagtagttttttttattaaggttgtatatgttattgtgttaat
gttttttagagtgggttttaagataaaagaataggataagttgtttgttttatttgttagtttgtgg
agaaaatgttattgtatatagaaaatagatagATGTAATTTATGGTGTTTTGTTAGGAAAGAGAGGTGTG
TTAAATAGTGATAAGTTTTGTTTATGTAAAAGATGATGTTTGgtgtgttagttgttttgttgtttggtt
gttttttttttgggggtgggtttagtaagagtaggtgtgggttta

C9-02-02

**********TTNATAGTGTTTTACGGTAAAAATAA
AATTTTATTTATTAATTATTTGTTGAGCGTTTATTGTTTATTAAGTATAAATAAAATTATTTAAAATTAC
GAAATCGTTTTTATTTTTTTCGGATTTAGTAGTTTTTTTTTATTAAGGTTCGTATACGTTATTGCGTTAA
CGTTTTTTTAGAGCGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTCGTTTTATTTCGTTAGTTTCGT
GAGAAANCGTTATCGTATATAGAAAATAGATAGACGTAATTTACGGTGTTTCGTTAGGAAAGAGAGGTGC
GTTAAATAGCGATAAGTTTCGTTTACGTAAAAGATGACGTTTGGTGTGTTAGTCGTTTTTGTTGTTCGGT
TGTTTTTTTTTGGGGGCGGGTTTAGTAGAGTAGGTGTGGGTTTA

C9-02-03

**********TTTNAAAGTGTTTTACGGTAAATAAAT
TTTATTTATTAATTATTTGTTGAGCGTTTATTGTTTATTAAGTATAAATAAAATTATTTAAAATTACGAA
ATCGTTTTTATTTTTTTCGGATTTAGTAGTTTTTTTATTAAGGTTCGTATACGTTATTGCGTTAACGTTT
TTTTAGAGCGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTCGTTTTATTTCGTTAGTTTCGTGAGAA
AACGTTATCGTATATAGAAAATAGATAGACGTAATTTACGGTGTTTCGTTAGGAAAGAGAGGTGCGTTAA
ATAGCGATAAGTTTCGTTTACGTAAAAGATGACGTTTGGTGTGTTAGTCGTTTTTGTTGTTCGGTTGTTT
TTTTTTTGGGGGCGGGTTTAGTAAGAGTAGGTGTGGGTT

Ed-02-01

**********NNNTTANNGTGTTTTATGGTANNNAAAA
TTTTATTTATTAATTATTTGTTGAGTGNNNNNNGTTTATTAAGTATAAATAAAATTATTTAAAATTATGA
AATTGTTTTTATTTTTTTTGGATTTAGTAGTTTTTTTTATTAAGGTTTGTATATGTTATTGTGTTAATGT
TTTTTTAGAGTGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTTGTTTTATTTTGTTAGTTTTGTGAG
AAAATGTTATTGTATATAGAAAATAGATAGATGTAATTTATGGTGTTTTGTTAGGAAAGAGAGGTGTGTT
AAATAGTGATAAGTTTTGTTTATGTAAAAGATGATGTTTGGTGTGTTAGTTGTTTTTGTTGTTTGGTTGT
TTTTTTTTTGGGGGTGGGGTTTC

Ed-02-02

**********NNNTANNGTGTTTTATGGTAANTAAAA
TTTTATTTATTAATTATTTGTTGAGTGTNNNNNGTTTATTAAGTATAAATAAAATTATTTAAAATTATGA
AATTGTTTTTATTTTTTTTGGATTTAGTAGTTTTTTTTATTAAGGTTTGTATATGTTATTGTGTTAATGT

TTTTTTAGAGTGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTTGTTTTATTTTGTTAGTTTTGTGAG
AAAATGTTATTGTATATAGAAAATAGATAGATGTAATTTATGGTGTTTTGTTAGGAAAGAGAGGTGTGTT
AAATAGTGATAAGTTTTGTTTATGTAAAAGATGATGTTTGGTGTGTTAGTTGTTTTTGTTGTTTGGTTGT
TTTTTTNTTGGGGGTGGGGTT

WT-01

**********NNNTTATANTNNTTTATGNNAANTANNN
TTTATTTATTAATTATTTGNTGAGTGNTNANTGNTTATTAAGTATAAATAAAATTATTTAAAATTATGAA
ATTGTTTTTATTTTTTTTGGATTTAGTAGTTTTTTTTATTAAGGTTTGTATATGTTATTGTGTTAATGTT
TTTTTAGAGTGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTTGTTTTATTTTGTTAGTTTTGTGAGA
AAATGTTATTGTATATAGAAAATAGATAGATGTAATTTATGGTGTTTTGTTAGGAAAGAGAGGTGTGTTA
AATAGTGATAAGTTTTGTTTATGTAAAAGATGATGTTTGGTGTGTTAGTTGTTTTTGTTGTTTGGNTGTT
TTTTTTTTGGGGGTGCGGTTTCCCCCACCAGGGGCCCCTCCCCCCCCGCCNCCCCCCCCCCCCCCCCCTC
CCCCCCCCCT

WT-02

**********NNNTANNGTGTTTTATGGTNNNATAAA
ATTTTATTTATTAATTATTTGTTGAGTGTTTATTGTTTATTAAGTATAAATAAAATTATTTAAAATTATG
AAATTGTTTTTATTTTTTTTGGATTTAGTAGTTTTTTTTTATTAAGGTTTGTATATGTTATTGTGTTAATG
TTTTTTTAGAGTGGGTTTTAAGATAAAAGAATAGGATAAGTTGTTTTGTTTTATTTTGTTAGTTTTGTGA
GAAAATGTTATTGTATATAGAAAATAGATAGATGTAATTTATGGTGTTTTGTTAGGAAAGAGAGGTGTGT
TAAATAGTGATAAGTTTTGTTTATGTAAAAGATGATGTTTGGTGTGTTAGTTGTTTTTGTTGTTTGGTTG
TTTTTTTTTTGGGGGTGGGGTTTANTAAA

BIBLIOGRAPHY

Al-Sarraj, S. et al., 2011. P62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. *Acta Neuropathologica*, 122(6), pp.691–702.

Almeida, S. et al., 2013. Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathologica*, 126(3), pp.385–399.

Amoroso, M.W. et al., 2013. Accelerated high-yield generation of limb-innervating motor neurons from human stem cells. *J Neurosci*, 33(2), pp.574–586. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23303937%5Cnhttp://www.jneurosci.org/content/33/2/574.full.pdf.

Amoroso, M.W. et al., 2013. Accelerated High-Yield Generation of Limb-Innervating Motor Neurons from Human Stem Cells TL - 33. *The Journal of Neuroscience*, 33 VN-r(2), pp.574–586. Available at: http://dx.doi.org/10.1523/JNEUROSCI.0906-12.2013.

An, M.C. et al., 2012. Genetic correction of huntington's disease phenotypes in induced pluripotent stem cells. *Cell Stem Cell*, 11(2), pp.253–263.

An, M.C. et al., 2014. Polyglutamine Disease Modeling: Epitope Based Screen for Homologous Recombination using CRISPR/Cas9 System. *PLoS currents*, 6, pp.1–19. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3994193&tool=pmcentrez&rendertype=abstract.

Andersen, P.M., 2006. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Current Neurology and Neuroscience Reports*, 6(1), pp.37–46.

Andersen, P.M. et al., 2003. Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes. *Amyotrophic lateral sclerosis and other motor neuron disorders : official*

publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases, 4(2), pp.62–73.

Andersen, P.M. & Al-Chalabi, A., 2011. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat Rev Neurol*, 7(11), pp.603–615. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21989245%5Cnhttp://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21989245.

Anokye-Danso, F. et al., 2011. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell*, 8(4), pp.376–388.

Apostolou, E. & Hochedlinger, K., 2013. Chromatin dynamics during cellular reprogramming. *Nature*, 502(7472), pp.462–71. Available at: http://www.nature.com.libezproxy.open.ac.uk/nature/journal/v502/n7472/full/nature12749.html.

Arai, T. et al., 2006. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochemical and Biophysical Research Communications*, 351(3), pp.602–611.

Ash, P.E.A. et al., 2013. Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron*, 77(4), pp.639–646. Available at: http://dx.doi.org/10.1016/j.neuron.2013.02.004.

Atanasio, A. et al., 2016. C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Scientific reports*, 6, p.23204. Available at: http://www.nature.com/srep/2016/160316/srep23204/full/srep23204.html.

Avilion, A.A. et al., 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes and Development*, 17(1), pp.126–140.

Ayala, Y.M. et al., 2008. Structural determinants of the cellular localization and shuttling of TDP-43. *Journal of Cell Science*, 121(22), pp.3778–3785.

Bagni, C. et al., 2012. Fragile X syndrome: Causes, diagnosis, mechanisms, and therapeutics. *Journal of Clinical Investigation*, 122(12), pp.4314–4322.

Ban, H. et al., 2011. Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc Natl Acad Sci U S A*, 108(34), pp.14234–14239. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21821793.

Bannwarth, S. et al., 2014. A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain*, 137(8), pp.2329–2345.

Barmada, S.J. et al., 2010. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *The Journal of Neuroscience*, 30(2), pp.639–49. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2821110&tool=pmcentrez&rendertype=abstract.

Beck, J. et al., 2013. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *American Journal of Human Genetics*, 92(3), pp.345–353.

Belzil, V. V. et al., 2014. Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Research*, 1584(October), pp.15–21.

Belzil, V. V., Bauer, P.O., et al., 2013. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathologica*, 126(6), pp.895–905.

Belzil, V. V., Gendron, T.F. & Petrucelli, L., 2013. RNA-mediated toxicity in neurodegenerative disease. *Molecular and Cellular Neuroscience*, 56, pp.406–419. Available at: http://dx.doi.org/10.1016/j.mcn.2012.12.006.

Bennion Callister, J. & Pickering-Brown, S.M., 2014. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Experimental Neurology*, 262(PB), pp.84–90.

Bilican, B. et al., 2012. Mutant induced pluripotent stem cell lines recapitulate aspects of TDP-43 proteinopathies and reveal cell-specific vulnerability. *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), pp.5803–8. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3326463&tool=pmcentrez&rendertype=abstract.

Blasco, M.A., Serrano, M. & Fernandez-Capetillo, O., 2011. Genomic instability in iPS: time for a break. *The EMBO Journal*, 30(3011), pp.991–993. Available at: www.embojournal.org.

van Blitterswijk, M. et al., 2015. Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta Neuropathologica*, 130(6), pp.863–876.

Boeynaems, S. et al., 2016. Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Scientific reports*, 6(August 2015), p.20877. Available at: http://www.nature.com/srep/2016/160212/srep20877/full/srep20877.html.

Boillee, S., 2006. Onset and progression in inherited ALS determined by motor neurons and microglia. *Science*, 312, pp.1389–1392. Available at: http://dx.doi.org/10.1126/science.1123511.

Boillée, S., Vande Velde, C. & Cleveland, D.W., 2006. ALS: A Disease of Motor Neurons and Their Nonneuronal Neighbors. *Neuron*, 52(1), pp.39–59.

Boulting, G.L. et al., 2011. A functionally characterized test set of human induced pluripotent stem cells. *Nat Biotechnol*, 29(3), pp.279–286.

Brooks, B.R. et al., 2000. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis and other motor neuron disorders : official publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases*, 1(5), pp.293–299.

Bruijn, L.I. et al., 1997. ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Neuron*, 18(2), pp.327–338.

Bruijn, L.I. & Cudkowicz, M., 2006. Therapeutic targets for amyotrophic lateral sclerosis: current treatments and prospects for more effective therapies. *Expert review of neurotherapeutics*, 6(3), pp.417–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16533145.

Buchan, J.R. et al., 2013. XEukaryotic stress granules are cleared by autophagy and

Cdc48/VCP function. *Cell*, 153(7).

Buchan, J.R. & Parker, R., 2009. Eukaryotic Stress Granules: The Ins and Outs of Translation. *Molecular Cell*, 36(6), pp.932–941.

Buratti, E. & Baralle, F.E., 2001. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator of CFTR Exon 9. *Journal of Biological Chemistry*, 276(39), pp.36337–36343.

Burberry, A. et al., 2016. Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Science translational medicine*, 8(347), p.347ra93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27412785.

Byrne, S. et al., 2012. Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: A population-based cohort study. *The Lancet Neurology*, 11(3), pp.232–240.

Byrne, S.M., Mali, P. & Church, G.M., 2014. *Genome editing in human stem cells* 1st ed., Elsevier Inc. Available at: http://dx.doi.org/10.1016/B978-0-12-801185-0.00006-4.

Cartwright, P. et al., 2005. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development (Cambridge, England)*, 132(5), pp.885–96. Available at: http://dev.biologists.org/content/132/5/885.abstract.

Chambers, I. et al., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113(5), pp.643–655.

Chen, H. et al., 2014. Modeling ALS with iPSCs reveals that mutant SOD1 misregulates neurofilament balance in motor neurons. *Cell Stem Cell*, 14(6), pp.796–809. Available at: http://dx.doi.org/10.1016/j.stem.2014.02.004.

Ciura, S. et al., 2013. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Annals of Neurology*, 74(2), pp.180–187.

Cleary, J.D. & Ranum, L.P.W., 2013. Repeat-associated non-ATG (RAN) translation in neurological disease. *Human Molecular Genetics*, 22(R1), pp.45–51.

Cong, L. et al., 2013. Multiplex Genome Engineering Using CRISPR/Cas System. *Science*, 339(February), pp.819–824.

Cooper-Knock, J., Higginbottom, A., et al., 2015. Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathologica*, 130(1), pp.63–75.

Cooper-Knock, J., Bury, J.J., et al., 2015. C9ORF72 GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. *PLoS ONE*, 10(5), pp.1–17.

Cooper-Knock, J. et al., 2012. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain*, 135(3), pp.751–764.

Cooper-Knock, J., Walsh, M.J., et al., 2014. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain*, 137(7), pp.2040–2051.

Cooper-Knock, J., Shaw, P.J. & Kirby, J., 2014. The widening spectrum of C9ORF72-related disease; Genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathologica*, 127(3), pp.333–345.

Corti, S. et al., 2015. Human pluripotent stem cells as tools for neurodegenerative and neurodevelopmental disease modeling and drug discovery. *Expert Opinion on Drug Discovery*, 10(6), pp.615–629. Available at: http://informahealthcare.com/doi/abs/10.1517/17460441.2015.1037737.

Cruts, M. et al., 2013. Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. *Trends in Neurosciences*, 36(8), pp.450–459.

Dafinca, R. et al., 2016. C9orf72 Hexanucleotide Expansions are Associated with Altered ER Calcium Homeostasis and Stress Granule Formation in iPSC-Derived Neurons from Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Stem cells (Dayton, Ohio)*, p.2016. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27097283.

Damiano, M., 2006. Neural mitochondrial Ca2+ capacity impairment precedes the onset of motor symptoms in G93A Cu/Zn-superoxide dismutase mutant mice. *J. Neurochem.*, 96, pp.1349–1361. Available at: http://dx.doi.org/10.1111/j.1471-4159.2006.03619.x.

Van Damme, P. et al., 2005. Excitotoxicity and amyotrophic lateral sclerosis. *Neuro-*

degenerative diseases, 2(3–4), pp.147–159.

DeJesus-Hernandez, M. et al., 2011. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2), pp.245–256.

Deng, H.-X. et al., 2011. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*, 477(7363), pp.211–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3169705&tool=pmcentrez&rendertype=abstract.

Deng, H.X. et al., 1993. Amyotrophic lateral sclerosis and structural defects in Cu,Zn superoxide dismutase. *Science (New York, N.Y.)*, 261(5124), pp.1047–1051.

Devlin, A.-C. et al., 1AD. Human iPSC-derived motoneurons harbouring TARDBP or C9ORF72 ALS mutations are dysfunctional despite maintaining viability. *Nature Communications*, 6, pp.1–12. Available at: http://dx.doi.org/10.1038/ncomms6999%5Cnpapers2://publication/doi/10.1038/ncomms6999.

Dhaliwal, G.K. & Grewal, R.P., 2000. Mitochondrial DNA deletion mutation levels are elevated in ALS brains. *Neuroreport*, 11(11), pp.2507–2509.

Dianov, G.L. & Hübscher, U., 2013. Mammalian base excision repair: The forgotten archangel. *Nucleic Acids Research*, 41(6), pp.3483–3490.

Ding, Q. et al., 2013. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*, 12(2), pp.238–251. Available at: http://dx.doi.org/10.1016/j.stem.2012.11.011.

Dion, V. & Wilson, J.H., 2009. Instability and chromatin structure of expanded trinucleotide repeats. *Trends in Genetics*, 25(7), pp.288–297.

Donnelly, C.J. et al., 2013. {RNA} toxicity from the {ALS/FTD} {C9ORF72} expansion is mitigated by antisense intervention. , 80(2), pp.415–428.

Donnelly, C.J. et al., 2013. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron*, 80(2), pp.415–428. Available at: http://dx.doi.org/10.1016/j.neuron.2013.10.015.

Dormann, D. & Haass, C., 2013. Fused in sarcoma (FUS): An oncogene goes awry in

neurodegeneration. *Molecular and Cellular Neuroscience*, 56, pp.475–486.

Du, Z.-W. et al., 2015. Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nature communications*, 6, p.6626. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4375778&tool=pmcentrez&rendertype=abstract.

Duan, W., 2010. Mutant TAR DNA-binding protein-43 induces oxidative injury in motor neuron-like cell. *Neuroscience*, 169, pp.1621–1629. Available at: http://dx.doi.org/10.1016/j.neuroscience.2010.06.018.

Egawa, N. et al., 2012. Drug Screening for ALS Using Patient-Specific Induced Pluripotent Stem Cells. *Science Translational Medicine*, 4(145), p.145ra104-145ra104. Available at: http://stm.sciencemag.org/content/4/145/145ra104%5Cnhttp://stm.sciencemag.org/content/4/145/145ra104.full.pdf%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/22855461.

Esvelt, K.M. & Wang, H.H., 2013. Genome-scale engineering for systems and synthetic biology. *Molecular systems biology*, 9(641), p.641. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3564264&tool=pmcentrez&rendertype=abstract.

Evans, M.J. & Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, pp.154–156.

Farg, M.A. et al., 2014. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Human Molecular Genetics*, 23(13), pp.3579–3595.

Fecto, F. et al., 2011. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Archives of neurology*, 68(11), pp.1440–1446. Available at: http://archneur.jamanetwork.com/article.aspx?doi=10.1001/archneurol.2011.250%5Cnpapers3://publication/doi/10.1001/archneurol.2011.250.

Feinberg, A.P. et al., 2010. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Science translational medicine*, 2(49), p.49ra67. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3137242&tool=pmcentrez&rendertype=abstract.

Fernandes, H.J.R. et al., 2016. ER Stress and Autophagic Perturbations Lead to Elevated Extracellular α-Synuclein in GBA-N370S Parkinson's iPSC-Derived Dopamine Neurons. *Stem Cell Reports*, 6(3), pp.342–356.

Ferraiuolo, L. et al., 2011. Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nature reviews. Neurology*, 7(11), pp.616–630. Available at: http://www.nature.com/nrneurol/journal/v7/n11/pdf/nrneurol.2011.152.pdf%5Cnpapers://2869f907-86c9-45b4-b3a1-9fd72ada1eca/Paper/p53665.

Ferrante, R.J. et al., 1997. Evidence of increased oxidative damage in both sporadic and familial amyotrophic lateral sclerosis. *Journal of neurochemistry*, 69, pp.2064–2074.

Ferri, A. et al., 2006. Familial ALS-superoxide dismutases associate with mitochondria and shift their redox potentials. *Proceedings of the National Academy of Sciences of the United States of America*, 103(37), pp.13860–13865.

Fontana, A.C.K., 2015. Current approaches to enhance glutamate transporter function and expression. *Journal of Neurochemistry*, 134(6), pp.982–1007.

Fontes, A. & Lakshmipathy, U., 2013. Advances in genetic modification of pluripotent stem cells. *Biotechnology Advances*, 31(7), pp.994–1001. Available at: http://dx.doi.org/10.1016/j.biotechadv.2013.07.003.

Francius, C. & Clotman, F., 2010. Dynamic expression of the Onecut transcription factors HNF-6, OC-2 and OC-3 during spinal motor neuron development. *Neuroscience*, 165(1), pp.116–129.

Fratta, P. et al., 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Scientific reports*, 2, p.1016. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3527825&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23264878%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3527825.

Fratta, P. et al., 2013. Homozygosity for the C9orf72 GGGGCC repeat expansion in

frontotemporal dementia. *Acta Neuropathologica*, 126(3), pp.401–409.

Freibaum, B.D. et al., 2015. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature*, 525(7567), pp.129–33.

Freischmidt, A. et al., 2015. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nature neuroscience*, 18(5), pp.631–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25803835.

Fu, Y. et al., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*, 32(3), pp.279–284. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24463574%5Cnhttp://www.nature.com/nbt/journal/v32/n3/pdf/nbt.2808.pdf.

Fusaki, N. et al., 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, 85(8), pp.348–62. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3621571&tool=pmcentrez&rendertype=abstract.

Gendron, T.F. et al., 2014. Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathologica*, 127(3), pp.359–376.

Gendron, T.F., Cosio, D.M. & Petrucelli, L., 2013. c9RAN translation: a potential therapeutic target for the treatment of amyotrophic lateral sclerosis and frontotemporal dementia. *Expert opinion on therapeutic targets*, 17(9), pp.991–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23844663.

Gijselinck, I. et al., 2012. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: A gene identification study. *The Lancet Neurology*, 11(1), pp.54–65.

Gijselinck, I. et al., 2015. The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Molecular Psychiatry*, (August), pp.1–13. Available at: http://www.nature.com/doifinder/10.1038/mp.2015.159.

Di Giorgio, F.P. et al., 2007. Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nature neuroscience*, 10(5), pp.608–14. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3139463&tool=pmcentrez&rendertype=abstract.

Giri, S. & Bader, A., 2015. A low-cost, high-quality new drug discovery process using patient-derived induced pluripotent stem cells. *Drug Discovery Today*, 20(1), pp.37–49. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84921511195&partnerID=40&md5=756c7cf2b2fb0b36d007bd4366de96e0.

Glass, C.K. et al., 2010. Mechanisms Underlying Inflammation in Neurodegeneration. *Cell*, 140(6), pp.918–934.

Gore, A. et al., 2011. Somatic coding mutations in human induced pluripotent stem cells. *Nature*, 470(7336), pp.63–67. Available at: http://dx.doi.org/10.1038/nature09805.

Greenway, M.J. et al., 2006. ANG mutations segregate with familial and "sporadic" amyotrophic lateral sclerosis. *Nature genetics*, 38(4), pp.411–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16501576.

Guo, Y. et al., 2010. Decreased GLT-1 and increased SOD1 and HO-1 expression in astrocytes contribute to lumbar spinal cord vulnerability of SOD1-G93A transgenic mice. *FEBS Letters*, 584(8), pp.1615–1622.

Gurney, M.E. et al., 1994. Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science (New York, N.Y.)*, 264(5166), pp.1772–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8209258.

Haeusler, A.R. et al., 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. TL - 507. *Nature*, 507 VN-(7491), pp.195–200. Available at: /Users/yurikoharigaya/Documents/ReadCube Media/Sid_6_13_14.pdf%5Cnhttp://dx.doi.org/10.1038/nature13124.

Haeusler, A.R., 2014. Nucleotide Structural Polymorphisms Formed by GGGGCC Repeats Cause C9orf72 Abortive Transcription and Nucleolar Stress. *Biophysical Journal*, 106(2), p.488a. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0006349513057366.

Harraz, M.M., 2008. SOD1 mutations disrupt redox-sensitive Rac regulation of NADPH oxidase in a familial ALS model. *J. Clin. Invest.*, 118, pp.659–670.

Hart, M.P. & Gitler, A.D., 2012. ALS-associated ataxin 2 polyQ expansions enhance stress-induced caspase 3 activation and increase TDP-43 pathological modifications. TL - 32. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32 VN-r(27), pp.9133–9142. Available at: /Users/yurikoharigaya/Documents/ReadCube Media/hart2012.pdf%5Cnhttp://dx.doi.org/10.1523/JNEUROSCI.0996-12.2012.

Hodges, J.R. et al., 2004. Clinicopathological correlates in frontotemporal dementia. *Annals of Neurology*, 56(3), pp.399–406.

Howland, D.S. et al., 2002. Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proceedings of the National Academy of Sciences of the United States of America*, 99(3), pp.1604–1609. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=122237&tool=pmcentrez&rendertype=abstract.

Hu, B.-Y. & Zhang, S.-C., 2009. Differentiation of spinal motor neurons from pluripotent human stem cells. *Nature protocols*, 4(9), pp.1295–304. Available at: http://dx.doi.org/10.1038/nprot.2009.127.

Huang, X. et al., 2015. Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient ipscs after genome editing of the sickle point mutation. *Stem Cells*, 33(5), pp.1470–1479.

Ignatius, S.O. et al., 1995. Cloning and Characterization of a Novel Cellular Protein, TDP-43, That Binds to Human Immunodeficiency Virus Type 1 TAR DNA Sequence Motifs. *J. Virol.*, 69(6), pp.3584–3596.

Ilieva, H., Polymenidou, M. & Cleveland, D.W., 2009. Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. *Journal of Cell Biology*, 187(6), pp.761–772.

Ivanov, P. et al., 2014. G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proceedings of the National Academy of*

Sciences of the United States of America, 111(51), pp.18201–6. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4280610&tool=pmcentrez&rendertype=abstract.

Jinek, M. et al., 2012. A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(August), pp.816–822.

Johnson, J.O. et al., 2010. Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron*, 68(5), pp.857–864.

Jovičić, A., Paul, J.W. & Gitler, A.D., 2016. Nuclear transport dysfunction: a common theme in amyotrophic lateral sclerosis and frontotemporal dementia. *Journal of neurochemistry*, 43. Available at: http://doi.wiley.com/10.1111/jnc.13642%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27087014.

Karumbayaram, S. et al., 2009. Directed differentiation of human-induced pluripotent stem cells generates active motor neurons. *Stem Cells*, 27(4), pp.806–811.

Kawamata, T. et al., 1992. Immunologic reactions in amyotrophic lateral sclerosis brain and spinal cord tissue. *The American journal of pathology*, 140(3), pp.691–707. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1886170&tool=pmcentrez&rendertype=abstract.

Keller, B.A. et al., 2012. Co-aggregation of RNA binding proteins in ALS spinal motor neurons: Evidence of a common pathogenic mechanism. *Acta Neuropathologica*, 124(5), pp.733–747.

Kieran, D., 2005. A mutation in dynein rescues axonal transport defects and extends the life span of ALS mice. *J. Cell Biol.*, 169, pp.561–567. Available at: http://dx.doi.org/10.1083/jcb.200501085.

Kiernan, M.C. et al., 2011. Amyotrophic lateral sclerosis. In *The Lancet*. pp. 942–955.

Kim, K. et al., 2010. Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313), pp.285–290.

Kiskinis, E. et al., 2014. Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell*, 14(6), pp.781–795.

Available at: http://dx.doi.org/10.1016/j.stem.2014.03.004.

Koppers, M. et al., 2015. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Annals of Neurology*, 78(3), pp.426–438.

Kwak, S. et al., 2010. AMPA receptor-mediated neuronal death in sporadic ALS. *Neuropathology*, 30, pp.182–188. Available at: http://dx.doi.org/10.1111/j.1440-1789.2009.01090.x.

Kwiatkowski, T.J. et al., 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 323(5918), pp.1205–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19251627.

Kwon, I. et al., 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (New York, N.Y.)*, 345(6201), pp.1139–45. Available at: http://science.sciencemag.org/content/345/6201/1139.abstract.

Lagier-Tourenne, C. et al., 2013. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(47), pp.E4530-9. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1318835110%5Cnpapers3://publication/doi/10.1073/pnas.1318835110.

Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D.W., 2010. TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration. *Human Molecular Genetics*, 19(R1).

Law, W.J., Cann, K.L. & Hicks, G.G., 2006. TLS, EWS and TAF15: A model for transcriptional integration of gene expression. *Briefings in Functional Genomics and Proteomics*, 5(1), pp.8–14.

Lee, Y.B. et al., 2013. Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. *Cell Reports*, 5(5), pp.1178–1186.

Levine, T.P. et al., 2013. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics*,

29(4), pp.499–503.

Li, J. et al., 2015. Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system. *Cell Research*, pp.634–637. Available at: http://www.nature.com/doifinder/10.1038/cr.2015.43.

Lian, X.J. & Gallouzi, I.E., 2009. Oxidative stress increases the number of stress granules in senescent cells and triggers a rapid decrease in p21waf1/cip1 translation. *Journal of Biological Chemistry*, 284(13), pp.8877–8887.

Lin, C.L.G. et al., 1998. Aberrant RNA processing in a neurodegenerative disease: The cause for absent EAAT2, a glutamate transporter, in amyotrophic lateral sclerosis. *Neuron*, 20(3), pp.589–602.

Lin, S. et al., 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*, 3, pp.1–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25497837.

Lin, T. et al., 2009. A chemical platform for improved induction of human iPSCs. *Nature methods*, 6(11), pp.805–8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3724527&tool=pmcentrez&rendertype=abstract.

Ling, S.C., Polymenidou, M. & Cleveland, D.W., 2013. Converging mechanisms in als and FTD: Disrupted RNA and protein homeostasis. *Neuron*, 79(3), pp.416–438. Available at: http://dx.doi.org/10.1016/j.neuron.2013.07.033.

Liu, E.Y. et al., 2014. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathologica*, 128(4), pp.525–541.

Liu, J., 2004. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron*, 43, pp.5–17. Available at: http://dx.doi.org/10.1016/j.neuron.2004.06.016.

Liu, Q. et al., 2016. Whole-exome sequencing identifies a missense mutation in hnRNPA1 in a family with flail arm ALS. *Neurology*, 87(17), pp.1763–1769.

Lomen-Hoerth, C. et al., 2003. Are amyotrophic lateral sclerosis patients cognitively normal? *Neurology*, 60(7), pp.1094–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12682312.

Lomen-Hoerth, C., Anderson, T. & Miller, B., 2002. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology*, 59(7), pp.1077–1079.

Long, X.R. et al., 2014. MicroRNA-148a is silenced by hypermethylation and interacts with DNA methyltransferase 1 in hepatocellular carcinogenesis. *International Journal of Oncology*, 45(6), pp.1915–1922.

Lorenzo, I.M., Fleischer, A. & Bachiller, D., 2013. Generation of mouse and human induced pluripotent stem cells (iPSC) from primary somatic cells. *Stem cell reviews*, 9(4), pp.435–50. Available at: http://link.springer.com/article/10.1007/s12015-012-9412-5/fulltext.html.

Luukkainen, L. et al., 2015. Epidemiology of Frontotemporal Lobar Degeneration in Northern Finland. *Dement Geriatr Cogn Dis Extra*, 5(3), pp.435–441. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26675285.

Mackenzie, I.R.A., Frick, P. & Neumann, M., 2014. The neuropathology associated with repeat expansions in the C9ORF72 gene. *Acta Neuropathologica*, 127(3), pp.347–357.

Mackenzie, I.R.A., Rademakers, R. & Neumann, M., 2010. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *The Lancet Neurology*, 9(10), pp.995–1007.

Maden, M., 2007. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nature reviews. Neuroscience*, 8(10), pp.755–765. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17882253.

Maherali, N. et al., 2008. A High-Efficiency System for the Generation and Study of Human Induced Pluripotent Stem Cells. *Cell Stem Cell*, 3(3), pp.340–345.

Majounie, E. et al., 2012. Large C9orf72 repeat expansions are not a common cause of Parkinson's disease. *Neurobiology of Aging*, 33(10).

Mali, P. et al., 2013. RNA-guided human genome engineering via Cas9. *Science*, 339(6121), pp.823–826. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23287722%5Cnhttp://classic.sciencemag.org/content/339/6121/823.full.pdf.

Marin, B. et al., 2016. Clinical and demographic factors and outcome of amyotrophic

lateral sclerosis in relation to population ancestral origin. *Eur J Epidemiol*, 31(3), pp.229–245. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26458931.

Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A*, 78(12), pp.7634–7638. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6950406.

Martin, L.J. & Wong, M., 2013. Aberrant Regulation of DNA Methylation in Amyotrophic Lateral Sclerosis: A New Target of Disease Mechanisms. *Neurotherapeutics*, 10(4), pp.722–733.

Maruyama, H. et al., 2010. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature*, 465(7295), pp.223–226. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20428114&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1038/nature08971.

Mastrocola, A.S. et al., 2013. The RNA-binding protein fused in sarcoma (FUS) functions downstream of poly(ADP-ribose) polymerase (PARP) in response to DNA damage. *Journal of Biological Chemistry*, 288(34), pp.24731–24741.

Matsumoto, S. et al., 1993. Ubiquitin-positive inclusion in anterior horn cells in subgroups of motor neuron diseases: A comparative study of adult-onset amyotrophic lateral sclerosis, juvenile amyotrophic lateral sclerosis and werdnig-hoffmann disease. *Journal of the Neurological Sciences*, 115(2), pp.208–213.

Maury, Y. et al., 2014. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol*, 33(1), pp.89–96. Available at: http://dx.doi.org/10.1038/nbt.3049.

May, S. et al., 2014. C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathologica*, pp.485–503.

McDonald, K.K. et al., 2011. TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Human Molecular Genetics*, 20(7), pp.1400–1410.

McGoldrick, P. et al., 2013. Rodent models of amyotrophic lateral sclerosis. *Biochimica et biophysica acta*, 1832(9), pp.1421–36. Available at: http://www.sciencedirect.com/science/article/pii/S0925443913000860.

McMillan, C.T. et al., 2015. C9orf72 promoter hypermethylation is neuroprotective: Neuroimaging and neuropathologic evidence. *Neurology*, 84, pp.1622–1630.

Merkle, F.T. & Eggan, K., 2013. Modeling human disease with pluripotent stem cells: From genome association to function. *Cell Stem Cell*, 12(6), pp.656–668. Available at: http://dx.doi.org/10.1016/j.stem.2013.05.016.

Mitsui, K. et al., 2003. The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113(5), pp.631–642.

Mizielinska, S. et al., 2013. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathologica*, 126(6), pp.845–857.

Mizielinska, S. et al., 2014. C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science*, 345(6201), pp.1192–1194.

Moralli, D. et al., 2011. An Improved Technique for Chromosomal Analysis of Human ES and iPS Cells. *Stem Cell Reviews and Reports*, 7(2), pp.471–477.

Mori, K., Arzberger, T., et al., 2013. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathologica*, 126(6), pp.881–893.

Mori, K., Weng, S.-M., et al., 2013. The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science*, 339(6125), pp.1335–1338. Available at: http://www.sciencemag.org/content/339/6125/1335%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23393093%5Cnhttp://www.sciencemag.org/content/339/6125/1335.full.pdf%5Cnhttp://www.sciencemag.org/content/suppl/2013/02/06/science.1232927.DC1/Mori.SM.pdf.

Neary, D. et al., 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, 51(6), pp.1546–1554. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=985550

0&retmode=ref&cmd=prlinks%5Cnpapers3://publication/uuid/5585A7A8-6AAF-4056-A34B-6FB202BC70D2.

Nelson, D.L., Orr, H.T. & Warren, S.T., 2013. The unstable repeats-Three evolving faces of neurological disease. *Neuron*, 77(5), pp.825–843.

Neumann, M. et al., 2006. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 314(5796), pp.130–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17023659.

Nichols, J. et al., 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, 95(3), pp.379–391.

Nishimura, A.L. et al., 2014. Allele-specific knockdown of ALS-associated mutant TDP-43 in neural stem cells derived from induced pluripotent stem cells. *PLoS ONE*, 9(3).

Nishitoh, H. et al., 2008. ALS-linked mutant SOD1 induces ER stress- and ASK1-dependent motor neuron death by targeting Derlin-1. *Genes and Development*, 22(11), pp.1451–1464.

Nordin, a. et al., 2015. Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Human Molecular Genetics*, 24(11), pp.3133–3142. Available at: http://www.hmg.oxfordjournals.org/cgi/doi/10.1093/hmg/ddv064.

O'Rourke, J.G. et al., 2015. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron*, 88(5), pp.892–901. Available at: http://dx.doi.org/10.1016/j.neuron.2015.10.027.

O 'rourke, J.G. et al., 2015. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD Neuron Report C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron*, 88, pp.892–901. Available at: http://dx.doi.org/10.1016/j.neuron.2015.10.027.

Oakes, J.A., Davies, M.C. & Collins, M.O., 2017. TBK1: a new player in ALS linking autophagy and neuroinflammation. *Molecular Brain*, 10(1), p.5. Available at: http://dx.doi.org/10.1186/s13041-017-0287-x.

Okita, K., Ichisaka, T. & Yamanaka, S., 2007. Generation of germline-competent induced pluripotent stem cells. *Nature*, 448(7151), pp.313–317. Available at:

http://www.nature.com/doifinder/10.1038/nature05934.

Onyike, C.U. & Diehl-Schmid, J., 2013. The epidemiology of frontotemporal dementia. *International review of psychiatry (Abingdon, England)*, 25(2), pp.130–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23611343%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3932112.

ORourke, J.G. et al., 2016. C9orf72 is required for proper macrophage and microglial function in mice. *Science*, 351(6279), pp.1324–1329. Available at: http://science.sciencemag.org/content/351/6279/1324.abstract.

Oyanagi, K. et al., 2008. Spinal anterior horn cells in sporadic amyotrophic lateral sclerosis show ribosomal detachment from, and cisternal distention of the rough endoplasmic reticulum. *Neuropathology and Applied Neurobiology*, 34(6), pp.650–658.

Pandya, R.S. et al., 2013. Therapeutic neuroprotective agents for amyotrophic lateral sclerosis. *Cellular and Molecular Life Sciences*, 70(24), pp.4729–4745.

Papadeas, S.T. et al., 2011. Astrocytes carrying the superoxide dismutase 1 (SOD1G93A) mutation induce wild-type motor neuron degeneration in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 108(43), pp.17803–8. Available at: http://www.pnas.org/content/108/43/17803.full?sid=15bf62be-d4f7-49f5-9f69-37803c5583bf.

Papp, B. & Plath, K., 2013. Epigenetics of reprogramming to induced pluripotency. *Cell*, 152(6), pp.1324–1343.

Park, C.Y. et al., 2015. Reversion of FMR1 Methylation and Silencing by Editing the Triplet Repeats in Fragile X iPSC-Derived Neurons. *Cell Reports*, 13(2), pp.234–241.

Parkinson, N., 2006. ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology*, 67, pp.1074–1077. Available at: http://dx.doi.org/10.1212/01.wnl.0000231510.89311.8b.

Pasinelli, P., 2004. Amyotrophic lateral sclerosis-associated SOD1 mutant proteins bind and aggregate with Bcl-2 in spinal cord mitochondria. *Neuron*, 43, pp.19–30.

Available at: http://dx.doi.org/10.1016/j.neuron.2004.06.021.

Peng, S. et al., 2011. Glutamate receptors and signal transduction in learning and memory. *Molecular Biology Reports*, 38(1), pp.453–460.

Peters, O.M. et al., 2015. Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron*, 88(5), pp.902–909.

Pokrishevsky, E., Grad, L.I. & Cashman, N.R., 2016. TDP-43 or FUS-induced misfolded human wild-type SOD1 can propagate intercellularly in a prion-like fashion. *Scientific reports*, 6, p.22155. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4772009&tool=pmcentrez&rendertype=abstract.

Polkey, M.I. et al., 1999. Respiratory aspects of neurological disease. *Journal of neurology, neurosurgery, and psychiatry*, 66(1), pp.5–15.

Qu, Q. et al., 2014. High-efficiency motor neuron differentiation from human pluripotent stem cells and the function of Islet-1. *Nature communications*, 5, p.3449. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24622388.

Rademakers, R., Neumann, M. & Mackenzie, I.R., 2012. Advances in understanding the molecular basis of frontotemporal dementia. *Nature reviews. Neurology*, 8(8), pp.423–34. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3629543&tool=pmcentrez&rendertype=abstract.

Rahimov, F. & Kunkel, L.M., 2013. The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. *The Journal of cell biology*, 201(4), pp.499–510. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3653356&tool=pmcentrez&rendertype=abstract.

Ran, F.A. et al., 2013. Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell*, 154(6), pp.1380–1389. Available at: http://dx.doi.org/10.1016/j.cell.2013.08.021.

Ranjith, R. & Sama, K., 2014. FUS / TLS in Stress Response - Implications for

Amyotrophic Lateral Sclerosis : A Dissertation.

Reddi, A.R. & Culotta, V.C., 2013. SOD1 integrates signals from oxygen and glucose to repress respiration. *Cell*, 152(1–2), pp.224–235.

Reddy, K. et al., 2013. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *Journal of Biological Chemistry*, 288(14), pp.9860–9866.

Renton, A.E. et al., 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72(2), pp.257–268.

Rhodes, D. & Lipps, H.J., 2015. Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Research*, 43(18), pp.8627–8637.

Ringholz, G.M. et al., 2005. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*, 65(4), pp.586–590. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16116120.

Rizzu, P. 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 Neuropathologica Communications*, 4(1), p.37. Available at: http://actaneurocomms.biomedcentral.com/articles/10.1186/s40478-016-0306-7.

Rohrer, J.D. et al., 2015. C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *The Lancet Neurology*, 14(3), pp.291–301.

Rosen, D.R., 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, 364, p.362. Available at: http://dx.doi.org/10.1038/364362c0.

Rosen, D.R. et al., 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis [published erratum appears in Nature 1993 Jul 22;364(6435):362] [see comments]. *Nature*, 362(6415), pp.59–62.

Rothstein, J.D. et al., 1996. Knockout of glutamate transporters reveals a major role for astroglial transport in excitotoxicity and clearance of glutamate. *Neuron*, 16(3), pp.675–686.

Rothstein, J.D. et al., 1995. Selective Loss of Glial Glutamate Transporter Glt-1 in Amyotrophic-Lateral-Sclerosis. *Annals of Neurology*, 38, pp.73–84.

Rothstein, J.D. et al., 1995. Selective loss of glial glutamate transporter GLT???1 in amyotrophic lateral sclerosis. *Annals of Neurology*, 38(1), pp.73–84.

Rubino, E. et al., 2012. SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology*, 79, pp.1556–62. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3655323&tool=pmcentrez&rendertype=abstract.

Russ, J. et al., 2015. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathologica*, 129(1), pp.39–52.

Sabatelli, M. et al., 2012. C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiology of Aging*, 33(8).

Sabatelli, M. et al., 2013. Mutations in the 3′ untranslated region of FUS causing FUS overexpression are associated with amyotrophic lateral sclerosis. *Human Molecular Genetics*, 22(23), pp.4748–4755.

Saccon, R.A. et al., 2013. Is SOD1 loss of function involved in amyotrophic lateral sclerosis? *Brain*, 136(8), pp.2342–2358.

Salewski, R.P. et al., 2013. The generation of definitive neural stem cells from PiggyBac transposon-induced pluripotent stem cells can be enhanced by induction of the NOTCH signaling pathway. *Stem cells and development*, 22(3), pp.383–96. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3549637&tool=pmcentrez&rendertype=abstract.

Sances, S. et al., 2016. Modeling ALS with motor neurons derived from human induced pluripotent stem cells. *Nature neuroscience*, 16(4), pp.542–53. Available at: http://dx.doi.org/10.1038/nn.4273.

Sareen, D. et al., 2013. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Science Translational Medicine*, 5(208), p.208ra149. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=24154603&retmode=ref&cmd=prlinks%5Cnpapers2://publication/doi/10.1126/scitranslmed.3007529.

Sasaki, S., 2010. Endoplasmic reticulum stress in motor neurons of the spinal cord in sporadic amyotrophic lateral sclerosis. *Journal of neuropathology and experimental neurology*, 69(4), pp.346–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20448480.

Sathasivam, S., Grierson, A.J. & Shaw, P.J., 2005. Characterization of the caspase cascade in a cell culture model of SOD1-related familial amyotrophic lateral sclerosis: expression, activation and therapeutic effects of inhibition. *Neuropathol. Appl. Neurobiol.*, 31, pp.467–485. Available at: http://dx.doi.org/10.1111/j.1365-2990.2005.00658.x.

Schludi, M.H. et al., 2015. Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathologica*, 130(4), pp.537–555. Available at: http://link.springer.com/10.1007/s00401-015-1450-z.

Schwank, G. et al., 2013. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell*, 13(6), pp.653–658.

Seah, Y.F.S. et al., 2015. Induced pluripotency and gene editing in disease modelling: Perspectives and challenges. *International Journal of Molecular Sciences*, 16(12), pp.28614–28634.

Seelaar, H. et al., 2011. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *Journal of neurology, neurosurgery, and psychiatry*, 82(5), pp.476–486.

Shaw, P.J., Forrest, V., et al., 1995. CSF and plasma amino acid levels in motor neuron disease: elevation of CSF glutamate in a subset of patients. *Neurodegeneration*, 4, pp.209–216. Available at: http://dx.doi.org/10.1006/neur.1995.0026.

Shaw, P.J., Ince, P.G., et al., 1995. Oxidative damage to protein in sporadic motor neuron disease spinal cord. *Ann. Neurol.*, 38, pp.691–695. Available at: http://dx.doi.org/10.1002/ana.410380424.

Shen, B. et al., 2013. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell research*, 23(5), pp.720–3. Available at: http://dx.doi.org/10.1038/cr.2013.46.

Shibata, N., 2001. Morphological evidence for lipid peroxidation and protein glycoxidation in spinal cords from sporadic amyotrophic lateral sclerosis patients. *Brain Res.*, 917, pp.97–104. Available at: http://dx.doi.org/10.1016/S0006-8993(01)02926-2.

Soldner, F. et al., 2012. Medicine. iPSC disease modeling. *Science (New York, N.Y.)*, 338(6111), pp.1155–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23197518.

Son, E.Y. et al., 2011. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell*, 9(3), pp.205–218. Available at: http://dx.doi.org/10.1016/j.stem.2011.07.014.

Soni, N., Reddy, B.V.K. & Kumar, P., 2014. GLT-1 transporter: An effective pharmacological target for various neurological disorders. *Pharmacology Biochemistry and Behavior*, 127, pp.70–81.

Stadtfeld, M. et al., 2008. Induced pluripotent stem cells generated without viral integration. *Science (New York, N.Y.)*, 322(5903), pp.945–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18818365%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3987909.

Su, Z. et al., 2014. Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron*, 83(5), pp.1043–1050. Available at: http://dx.doi.org/10.1016/j.neuron.2014.07.041.

Suh, E.R. et al., 2015. Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta Neuropathologica*, 130(3), pp.363–372.

Suzuki, N. et al., 2013. The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nature neuroscience*, 16(12), pp.1725–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24185425%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4397902&tool=pmcentrez&rendertype=abstract.

Takahashi, K. et al., 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, 131(5), pp.861–872.

Takahashi, K. & Yamanaka, S., 2006. Induction of Pluripotent Stem Cells from Mouse

Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4), pp.663–676.

Takahashi, M. et al., 2013. Stress granules inhibit apoptosis by reducing reactive oxygen species production. *Molecular and cellular biology*, 33(4), pp.815–29. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23230274%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3571346.

Tao, Z. et al., 2015. Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Human Molecular Genetics*, 24(9), pp.2426–2441.

Taylor, J.P., Brown, R.H. & Cleveland, D.W., 2016. Decoding ALS: from genes to mechanism. *Nature*, 539(7628), pp.197–206. Available at: http://www.nature.com/doifinder/10.1038/nature20413%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27830784.

Therrien, M. et al., 2013. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. *PLoS ONE*, 8(12), pp.1–10.

Thomson, J.A., 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science*, 282(5391), pp.1145–1147. Available at: http://science.sciencemag.org/content/282/5391/1145.abstract.

Tran, H. et al., 2015. Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron*, 87(6), pp.1207–1214. Available at: http://dx.doi.org/10.1016/j.neuron.2015.09.015.

Tucker, B.A. et al., 2013. Patient-specifc iPSC-derived photoreceptor precursor cells as a means to investigate retinitis pigmentosa. *eLife*, 2013(2), pp.1–18.

Turner, B.J. et al., 2008. TDP-43 expression in mouse models of amyotrophic lateral sclerosis and spinal muscular atrophy. *BMC neuroscience*, 9, p.104. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2583980&tool=pmcentrez&rendertype=abstract.

Turner, M.R. et al., 2013. Mechanisms, models and biomarkers in amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis & frontotemporal degeneration*, 14 Suppl 1, pp.19–32. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4284067&tool=pmcentrez&rendertype=abstract.

Upadhyay, S.K. et al., 2013. RNA Guided Genome Editing for Target Gene Mutations in Wheat. *G3: Genes|Genomes|Genetics*, 3(December), pp.2233–2238. Available at: http://g3journal.org/cgi/doi/10.1534/g3.113.008847.

Urushitani, M. et al., 2006. Chromogranin-mediated secretion of mutant superoxide dismutase proteins linked to amyotrophic lateral sclerosis. *Nature neuroscience*, 9(1), pp.108–118.

Vance, C. et al., 2009. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*, 323(5918), pp.1208–1211. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4516382&tool=pmcentrez&rendertype=abstract.

De Vos, K.J., 2007. Familial amyotrophic lateral sclerosis-linked SOD1 mutants perturb fast axonal transport to reduce axonal mitochondria content. *Hum. Mol. Genet.*, 16, pp.2720–2728. Available at: http://dx.doi.org/10.1093/hmg/ddm226.

De Vos, K.J. et al., 2008. Role of axonal transport in neurodegenerative diseases. *Annu. Rev. Neurosci.*, 31, pp.151–173. Available at: http://dx.doi.org/10.1146/annurev.neuro.31.061307.090711.

Waite, A.J. et al., 2014. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiology of Aging*, 35(7).

Wang, L., Gutmann, D.H. & Roos, R.P., 2011. Astrocyte loss of mutant SOD1 delays ALS disease onset and progression in G85R transgenic mice. *Human Molecular Genetics*, 20(2), pp.286–293.

Wang, W. et al., 2013. The ALS disease associated mutant TDP-43 impairs mitochondrial dynamics and function in motor neurons. *Human molecular genetics*, pp.1–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23827948.

Warren, L. et al., 2010. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, 7(5),

pp.618–630.

Wen, X. et al., 2016. Pathogenic determinants and mechanisms of ALS/FTD linked to hexanucleotide repeat expansions in the C9orf72 gene. *Neuroscience letters*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27619540.

Wheeler, J.R. et al., 2016. Distinct stages in stress granule assembly and disassembly. *eLife*, 5(Se).

Wichterle, H. et al., 2002. Directed differentiation of embryonic stem cells into motor neurons. *Cell*, 110(3), pp.385–397.

Wiedemann, F.R. et al., 2002. Mitochondrial DNA and respiratory chain function in spinal cords of ALS patients. *J. Neurochem.*, 80, pp.616–625. Available at: http://dx.doi.org/10.1046/j.0022-3042.2001.00731.x.

Wienholz, B.L. et al., 2010. DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B in vivo. *PLoS Genetics*, 6(9).

Wild, P. et al., 2011. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science*, 333(6039), pp.228–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21617041%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/21617041.

Williams, T.L. et al., 1997. Calcium-permeable [alpha]-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors: a molecular determinant of selective vulnerability in amyotrophic lateral sclerosis. *Ann. Neurol.*, 42, pp.200–207. Available at: http://dx.doi.org/10.1002/ana.410420211.

Williamson, T.L. & Cleveland, D.W., 1999. Slowing of axonal transport is a very early event in the toxicity of ALS-linked SOD1 mutants to motor neurons. *Nat. Neurosci.*, 2, pp.50–56. Available at: http://dx.doi.org/10.1038/4553.

Winton, M.J. et al., 2008. Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *Journal of Biological Chemistry*, 283(19), pp.13302–13309.

Wong, P.C. et al., 1995. An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of

mitochondria. *Neuron*, 14(6), pp.1105–1116.

Wu, C.-H. et al., 2012. Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature*, 488(7412), pp.499–503. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3575525&tool=pmcentrez&rendertype=abstract.

Xi, Z. et al., 2014. Hypermethylation of the CpG-island near the C9orf72 G4C2-repeat expansion in FTLD patients. *Human Molecular Genetics*, 23(21), pp.5630–5637.

Xi, Z. et al., 2016. Hypermethylation of the CpG Island Near the G_4C_2 Repeat in ALS with a C9orf72 Expansion. *The American Journal of Human Genetics*, 92(6), pp.981–989. Available at: http://dx.doi.org/10.1016/j.ajhg.2013.04.017.

Xi, Z. et al., 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *American Journal of Human Genetics*, 92(6), pp.981–989.

Xi, Z., van Blitterswijk, M., et al., 2015. Jump from Pre-mutation to Pathologic Expansion in C9orf72. *American Journal of Human Genetics*, 96(6), pp.962–970.

Xi, Z., Zhang, M., et al., 2015. The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathologica*, 129(5), pp.715–727.

Xu, T. et al., 2014. Cas9-based tools for targeted genome editing and transcriptional control. *Applied and Environmental Microbiology*, 80(5), pp.1544–1552.

Xu, Z. et al., 2013. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), pp.7778–83. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3651485&tool=pmcentrez&rendertype=abstract.

Yamanaka, K., 2008. Astrocytes as determinants of disease progression in inherited amyotrophic lateral sclerosis. *Nat. Neurosci.*, 11, pp.251–253. Available at: http://dx.doi.org/10.1038/nn2047.

Yang, L. et al., 2013. Optimization of scarless human stem cell genome editing. *Nucleic Acids Research*, 41(19), pp.9049–9061.

Yoshioka, N. et al., 2013. Efficient generation of human iPSCs by a synthetic self-replicative RNA. *Cell Stem Cell*, 13(2), pp.246–254. Available at: http://dx.doi.org/10.1016/j.stem.2013.06.001.

Yu, J. et al., 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, NY)*, 318(5858), pp.1917–1920. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1151526%5Cnpapers3://publication/doi/10.1126/science.1151526.

Yusa, K. et al., 2009. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nature methods*, 6(5), pp.363–369.

Yusa, K. et al., 2011. Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. *Nature*, 478(7369), pp.391–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21993621%5Cnhttp://dx.doi.org/10.1038/nature10424.

Zamiri, B. et al., 2014. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNAbinding proteins. *Journal of Biological Chemistry*, 289(8), pp.4653–4659.

Zhang, D. et al., 2012. Discovery of novel DENN proteins: Implications for the evolution of eukaryotic intracellular membrane structures and human disease. *Frontiers in Genetics*, 3(DEC).

Zhang, K. et al., 2015. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature*, 525(7567), pp.56–61. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26308891%5Cnhttp://www.nature.com/nature/journal/v525/n7567/pdf/nature14973.pdf.

Zhang, R. et al., 2009. Circulating endotoxin and systemic immune activation in sporadic amyotrophic lateral sclerosis (sALS). *Journal of Neuroimmunology*, 206(1–2), pp.121–124.

Zhang, Y.J. et al., 2014. Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathologica*, 128(4), pp.505–524.

Zu, T. et al., 2011. Non-ATG-initiated translation directed by microsatellite expansions.

Proceedings of the National Academy of Sciences of the United States of America, 108(1), pp.260–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21173221%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3017129.

Zu, T. et al., 2013. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences*, 110(51), pp.E4968–E4977. Available at: http://www.pnas.org/content/110/51/E4968%5Cnhttp://www.pnas.org/cgi/doi/10.1073/pnas.1315438110.