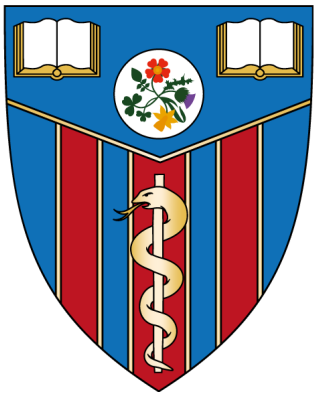
Using VCP mutant hiPSC-derived motor neurons and glia to capture early molecular pathogenic events in MND

Claire Elizabeth Hall

Institute of Neurology

UCL



Doctor of Philosophy

July 2016

Declaration:

I, Claire Hall confirm that the work presented in this thesis is my own. Where information has been derived from other sources or work has been produced collaboratively, it has been indicated in the text and preface. This thesis is not substantially the same as any other that may have been submitted for a qualification at any other university or similar institution. It does not exceed the limit of 100,000 words.

Claire Hall

Abstract

Motor neuron disease (MND) is rapidly progressive and invariably fatal, and with no significantly impactful therapies available to date there is desperate and unmet clinical need. Recent discoveries in the underlying pathology and genetics of MND suggest that altered proteostasis, RNA regulation and glial contribution play key roles in molecular pathogenesis. I have used a combination of human induced pluripotent stem cell (hiPSC) models and functional genomic technologies to characterize early pathogenic events in VCP-related MND. I hypothesize that the early pathogenic events in MND differ during the differentiation of motor neurons and astrocytes; therefore I first generated enriched populations of both motor neurons (MNs) and astrocytes (ACs) from control and VCP mutant patient iPSCs. Next to understand if changes in RNA regulation contribute to pathogenic events in VCP-related MND, I carried out a temporal analysis using RNA sequencing (RNAseq) of control and VCP mutant iPSCs neural derivatives. Additionally, I examined the results of this transcriptional study in light of collaborative live cell imaging experiments, performed across a matched time-course, that have uncovered cell-type specific organellar dysfunction in VCP mutant MNs. This paradigm importantly enabled us to discriminate primary from secondary pathogenic events in VCP-related MND. Further, after the onset of pathogenesis I also used individual nucleotide resolution UV cross linking with immunoprecipitation (iCLIP) to study RNA interactions of the candidate RNA-binding protein TDP43, which forms the pathological hallmark of MND. Together, this provides new insights into the earliest molecular pathogenic events in MND and has highlighted the contribution of both motor neurons and astrocytes. By identifying key disease mechanisms this study aims to guide future therapeutic strategy.

Preface

This thesis describes my doctoral research performed at the Institute of Neurology, UCL between October 2013 and July 2016.

Some of the work presented here has been achieved collaboratively. The VCP-mutant hiPSC lines used throughout my thesis were generated by Dr Selina Wray. In chapter 3, Electrophysiology was carried out by Sarah Crisp and axonal transport by Alexander Fellows, supervised by Professor Dimitri Kullmann and Professor Gipi Schiavo respectively. The public RNAseq data used in chapter 3 was downloaded from the NCBI GEO database (GEO: GSE73721) (Ye Zhang et al. 2016). Analysis of iCLIP data in Chapter 4 and 6 was partially done through the iCOUNT web server, which was established by Dr. Tomaž Curk and is helped maintained by Igor Ruiz. In addition the calculation of binding by RBPs to repeat elements (Alu's elements presented) was carried out by Nejc Haberman and analysis of RNAseq strandedness and distribution of reads binding to exonic, intronic or intergenic reads was carried out by Raphaelle Luiser from Nicholas Luscombe's lab. All Live cell imaging work in Chapter 5 (Cell death assays, Mitochondria dynamics, ER-mitochondria colocalisation, Glutathione levels, ROS production and ER stress), plus the analysis of ATP/KCl/Glutamate response and the mitochondrial membrane dependence upon complex I presenting in Chapter 3 was done in collaboration with Dr. Sonia Gandhi, Dr. Minee Choi and Dr. Zhi Yao. Confocal synaptic analysis presented in chapter 5 was carried out by Dr. Andras Lakatos. Finally in chapter 6, motif analysis of selected sites was performed by Nejc Haberman.

Acknowledgements

First of all I would like to thank my supervisors Professor Jernej Ule, Dr. Rickie Patani and Professor Nicholas Wood for their joint guidance and support throughout my PhD. Your enthusiasm and passion for research has been contagious. You have provided me with critical input while giving me the independence to explore on my own. I also appreciate all the opportunities you have given me to allow me to develop my career as a scientist and as a future medic, including through numerous opportunities to both attend and present at international conferences.

Next I would like to specially thank Dr. Chris Sibley for his continual advice, patience and encouragement throughout my PhD studies. I am extremely lucky to have learnt many of the core molecular and bioinformatics techniques from you - your precision, rigor and dedication is inspirational. Further you gave me the confidence to trust in myself and develop my own ideas, for which I am grateful. Additionally I would like to express my gratitude to Dr. Giulia Tyzack who has been a tremendous support in so many ways over this past year.

Further, I have been lucky enough to work within several fantastic laboratories and teams. I would like to express my gratitude to all of the past and present members of the Ule lab. In particular, to Nejc haberman for your invaluable bioinformatics input and Ina Huppertz, Jan Attig and Aarti Singh who continually supported me and made me laugh at the bench. Also I would like to thank everyone in the Patani lab, especially to Rubika Balendra, Giulia Tyzack, Sarah Weithoff, and to Selina Wray, Charlie Arber, Nuria Seto-Salvia and Elisavet Preza for your support in the iPSC lab. I'm so grateful to have spent all the high and low points in cell culture with you! Subsequently, 2 years in to my studies came the establishment of the iPSC-dream team! It has been thoroughly enjoyable to work so closely with you all and to have shared many good times outside of the lab too.

I wish to thank all of our collaborators over the past 3 years, which have all enabled us push the boundaries of our research in new exciting ways that we could not have done alone. In particular I would like to address our ongoing collaboration with Dr. Sonia Gandhi, Dr. Minee Choi and Dr. Zhi Yao, from

which the results have formed a substantial section within this thesis. Additionally I would like to gratefully acknowledge Deborah Hughes and Alan Pittman, at the Institute of Neurology's NGS core facility, who conducted all High throughput sequencing and Fastq file generation.

Also I would like to acknowledge the generosity of the Grand Challenges Programme who financial supported me, making the PhD possible.

During my time at UCL I have made many long-lasting friendships and I will never forget all the fun times we have had along the way! I also wish to thank all of my friends outside of the lab who have been there to celebrate all the small steps and to support me during the challenges. You have provided me the best distractions from my research.

Last but by no means least I would like to thank my family and George for your unconditional support and giving me the courage to follow my dreams. Mum and Dad, I know it hasn't been easy to understand what I've been doing and why but I am enormously grateful to have had your full backing. Coming home for part of my writing was an immense help and was made even more enjoyable with a constant supply of home-baked treats. George, I am exceptionally thankful for your patience and for believing in me more than I thought anyone ever could. Your selfless encouragement has never wavered and I'm looking forward to embarking on the next adventure together.

Contents

Abbreviations

3'UTR	3' untranslated region
AC	Astrocyte
AP	Anterior to posterior
ATP	Adenosine triphosphate
bHLH	Basic helix-loop-helix
BMP	Bone Morphogenetic Protein
bp	Base pair
CDK	Cyclin-dependent kinase
cDNA	Complementary DNA
CHAT	Choline acetyltransferase
CLIP	UV crosslinking and immunoprecipitation
CNS	Central nervous system
CNTF	Ciliary Neurotrophic Factor
D	Day
DAPI	4′,6-Diamidino-2-phenylindole dihydrochloride
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Triphosphates
DTT	Dithiothreitol
DV	Dorsal to ventral
EAAT-2	Excitatory Amino-Acid Transporter 2
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ESC	Embryonic stem cell
FGF	Fibroblast growth factor
fMND	Familial MND
FTLD	Frontotemporal lobar degeneration
FUS	Fused in sarcoma
GFAP	Glial fibrillary acidic protein
GLAST	Glutamate aspartate transporter
GO	Gene ontology
GOF	Gain of function

GPC	Gliogenic precursor
HD	homeodomain
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
hnRNPC	Heterogeneous nuclear ribonucleoprotein C
HOX	Homeotic
IBMPFD	Inclusion body myopathy with Pagets disease of the bone and frontal temporal dementia
ICC	Immunocytochemistry
iCLIP	Individual-nucleotide resolution CLIP
ICM	Inner cell mass
iPSC	Induced pluripotent stem cell
kDA	Kilodalton
KO	Knock out
LIF	Leukemia inhibitory factor
LOF	Loss of function
MAP2	Microtubule-associated protein 2
MN	Motor neuron
MND	Motor neuron disease
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein
NES	Nuclear export signal
NF1	Neurofibromin
NFIA	Nuclear Factor I A
NLS	Nuclear localization sequence
NPC	Neural precursor
nt	Nucleotides
OCT4	Octamer-binding transcription factor 4
OLIG2	Oligodendrocyte transcription factor
PAR-CLIP	Photoactivatable-ribonucleoside-enhanced CLIP
PCR	Polymerase chain reaction
PK	Proteinase K
PNK	Polynucleotide kinase
QC	Quality check
qPCR	Quantitative PCR
RA	Retinoic acid

RAN	Repeat-associated non-ATG
RBP	RNA binding protein
RIP	RNA immunoprecipitation
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RNase	Ribonuclease
ROS	Reactive oxygen species
RRM	RNA recognition motif
rRNA	Ribosomal RNA
RT	Reverse transcription
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SG	Stress granule
SHH	Sonic hedgehog
SMI32	Neurofilament H Non-Phosphorylated
SOD1	Superoxide dismutase 1
SYT1	Synaptotagmin-1
TARDBP	Gene encoding TAR DNA-Binding Protein 43
TDP43	TAR DNA-binding protein 43
TGF-B	Transforming growth factor beta
UPR	Unfolded protein response
UV	Ultraviolet
VCP	Vasolin containing protein

Chapter 1; Introduction

Neurodegenerative diseases are characterized by the progressive deterioration of region- and/or subtype-specific neurons. They vary in terms of severity, age of onset, prognosis and a number of pathogenic mechanisms have been implicated including oxidative stress, axonal transport deficits, mitochondrial dysfunction, excitotoxicity, calcium dysregulation, neuroinflammation, DNA damage and aberrant RNA processing (Ferraiuolo et al. 2011; R. Johnson et al. 2012). However one common pathological hallmark observed across the spectrum of neurodegenerative diseases is the accumulation and aggregation of misfolded protein (Soto 2013; Yerbury et al. 2016).

Motor Neuron Disease

Motor Neuron Disease (MND) encompasses many different neurodegenerative conditions, including progressive bulbar palsy (PSP), progressive muscular atrophy (PMA), primary lateral sclerosis (PLS) and Kennedy's disease. The most common form of MND is often referred to as amyotrophic lateral sclerosis (ALS) or sometimes as Lou Gehrig's disease and this is what I refer to throughout this thesis as MND.

MND is a rapidly progressive and invariably fatal neurodegenerative disease. The average age of onset is 55 years (S. Chen et al. 2013) and prevalence is 1/400 in those with European ancestry (Hardiman, van den Berg, and Kiernan 2011), with the incidence being 20% higher in males. It affects both the upper and lower motor neurons (MNs) (Hardiman, van den Berg, and Kiernan 2011); Upper MNs are those originating from the motor cortex or brainstem that convey impulses to the lower MNs, which originate in the anterior horn of the spinal cord and synapse at muscles. Patients suffer from premature loss of MNs resulting in progressive paralysis, loss of speech, inability to swallow and usually die within 3-5 years of symptom onset due to respiratory complications. Despite extensive research there are no effective treatments to date; the only available disease modifying therapy is Riluzole, a glutamate antagonist approved over 20years ago, which only prolongs life by 3 months (Hardiman, van den Berg, and Kiernan 2011). Other treatments are currently just symptomatic.

Genetics underlying MND

The cause of MND is largely unknown but several risk factors have been identified; older age, male sex and genetic (Ingre et al. 2015). The majority of MND cases are sporadic with only ~10% being familial (S. Chen et al. 2013), where the disease has been passed down through generations through genetic mutation. Over 20 causative genes have been identified in familial MND (fMND) (listed in Table 1.1.A) and these are usually inherited in an autosomal dominant pattern. Further, several of these genes associated with fMND have also been reported in a small number of sporadic MND cases, including FUS, C9ORF72, SOD1 and TDP43 (Turner et al. 2013). On top of this many other genes have been identified that increase susceptibility in sporadic MND (Listed also in Table 1.1.B)(S. Chen et al. 2013).

A)

Chromosome locus	Gene	Protein	Other Diseases caused by the gene	Key molecular pathways implicated in
2q33-2q35	Alsin	Alsin	PLS, IAHSP	Oxidative stress
14q11.2	ANG	Angiogenin	-	Aberrant RNA Processing
9p21	C9ORF72	Chromosome 9 open reading frame 72	FTD	Aberrant RNA Processing
2p13	DCTN1	Dynactin	-	Axonal transport
6q21	FIG4	Phosphoinositide-5phosphatease	CMT 4 J	Apoptosis
16p11.2	FUS	Fused in Sarcoma	-	Aberrant RNA Processing
10p13	OPTN	Optineurin	Primary open angle Glaucoma	Protein aggregation
9q34	SETX	Senataxin	SCAR1	Aberrant RNA

			and AOA2	Processing
9p13.2-21.3	SIGMAR1	Sigma Non Opiod Intracellular Receptor	FTD	Metabolic disturbance
21q22.1	SOD1	Cu/Zn SOD-1	-	Apoptosis, Oxidative stress, Mitochondrial disruption, microglia activation, protein aggregation, metabolic disturbance, axonal transport
15q15-21	SPG 11	Spatacsin	HSP	Axonal transport
1p36.2	TARDBP	DNA-binding protein	FTD	Aberrant RNA Processing, protein aggregation
Xp11	UBQLN2	Ubiquilin 2	ND	Protein aggregation, protein degradation
9q21-22	Unknown	Unknown	FTD	-
18q21	Unknown	Unknown	-	-
20q13.3	VAPB	Vesicle associated membrane protein associated protein B	SMA	Apoptosis, metabolic distrubance
9p13.3	VCP	Valosin Containing Protein	IBMPFD	Protein aggregation, protein

				degradation, apoptosis

B)

14q11.2	APEX1	Apurinic Endonuclease DNA repair enzyme 1	-	DNA repair, oxidative stress
12q24.12	ATXN2	Ataxin-2	SCA2	
3p11.2	CHMP2B	Chromatin Modifying Protein 2B	FTD	Protein degradation
6p22.2	HFE	Haemochromatosis	HHC	Oxidative stress, metabolic disturbance
22q12.2	NEFH	Neuro filament Heavy	-	Axonal transport
4q32.1-q34.3	NEK1	NIMA Related Kinase 1	-	DNA repair, microtubule stability, neuronal morphology
17q21.31	PGRN	Progranulin	FTD	Inflammation, microglia activation
7q21.3	PON 1,2,3	Paraoxonase	-	Oxidative stress
12q13.12	PRPH	Peripherin	-	Axonal transport
5q12.2-q13.3	SMN1	Survival Motor Neuron 1	SMA	Aberrant RNA Processing, protein aggregation
5q12.2-q13.3	SMN2	Survival Motor Neuron 2	SMA	Aberrant RNA Processing, protein aggregation
6p21	VEGF	Vascular	-	Hypoxia-

		Endothelial Growth Factor		response

Table 1.1; Genes identified in MND (adapted/updated from (S. Chen et al. 2013)**)**
A) Lists genes identified to cause fMND. B) Lists additional susceptibility genes identified in sporadic MND. Abreviations for table 1.1; PLS Primary Lateral Sclerosis, IAHSP Infantile onset ascending hereditary spastic paralysis, SCAR 1 Autosomal Recessive Spino-cerebellar ataxia, AOA2 Ataxia Ocular Apraxia 2, HSP Hereditary spastic paraplegia, SMA Spinal Muscular Atrophy, CMT 4 J Charcot-Marie Tooth disease type 4 J, SCA2 spinocerebellar ataxia type 2, HHC hereditary hemochromatosis, IBMPFD Inclusion body myopathy with Pagets disease and fronto temporal dementia.

From the genetic mutations identified, a number of biological pathways have been highlighted to potentially play a role in disease pathogenesis (Table 1.1, column 5); for example oxidative stress (SOD1 mutations) and intracellular trafficking (Dynactin mutations)(Pasinelli and Brown 2006). However, for reasons discussed below, the two most common hypotheses are that abnormal RNA metabolism and/or altered proteostasis play a key role in the pathogenesis of MND (Ferraiuolo et al. 2011).

The role of abnormal ribostasis and proteostasis in MND

All neurodegenerative disorders, including MND, exhibit the aggregation of misfolded protein, but the role of these aggregates in disease pathogenesis remains unresolved. It is vital for all cells to maintain a fine balance of protein production, concentration, conformation, location and degradation for normal cell metabolism and to be able to respond to changes in their environment. There are numerous stages involved in the production of a functional protein; including transcriptional regulation where RNA undergoes elaborate processing and translational control. Transcriptional regulation is necessary to generate the diversity and specificity of proteins required within different cell types and compartments, and is predominately regulated by a complex network of RNA binding proteins (RBPs). Translational control is particularly important to spatially regulate protein production and is markedly evident under stress (Spriggs, Bushell, and Willis 2010). Additionally, there are hundreds of chaperones, enzymes, and specialized proteins involved in folding, clearance and translocation networks within the cell that work together to ensure proteins are managed appropriately (summarized in Figure 1). However during stress,

in aging and disease these networks become overloaded and dysfunctional leading to the accumulation and aggregation of RNA, misfolded and aggregation-prone proteins. These aggregated proteins in MND are predominately ubiquitinated and are seen in the spinal cord and multiple brain regions (Baloh 2011).

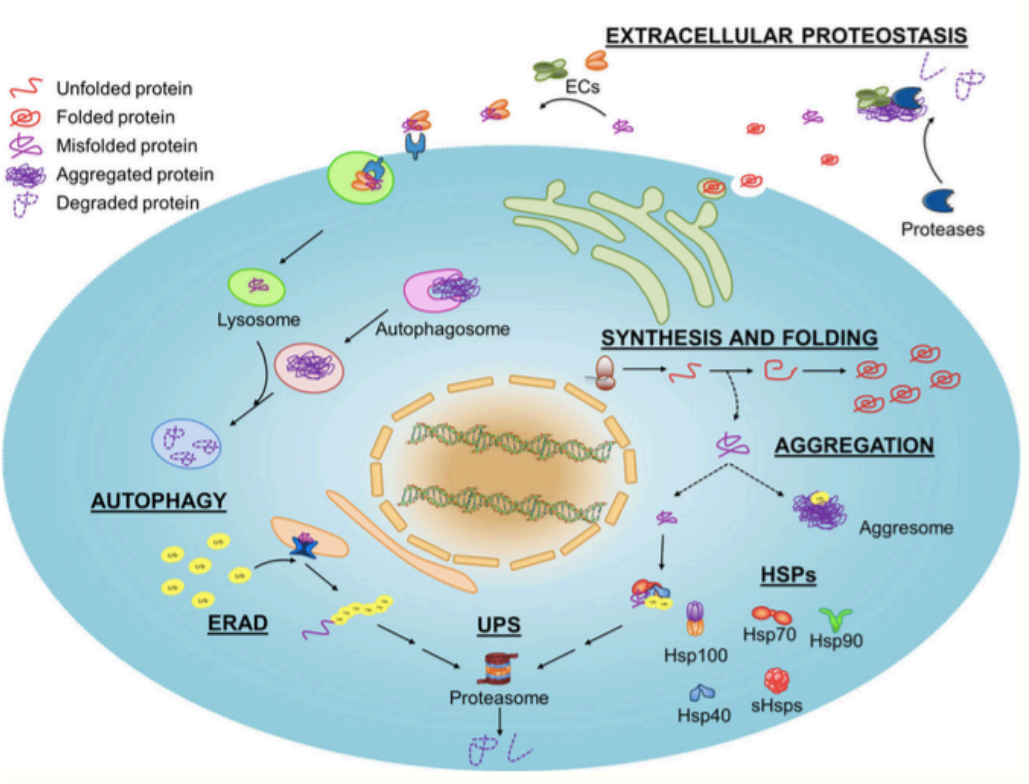


Figure 1; Schematic representation of key pathways controlling protein homeostasis, taken from (Yerbury et al. 2016).

The balance and management of effective protein production and degradation in the cell is crucial. Protein synthesis begins with transcription in the nucleus and continues in the cytoplasm where it is translated and folded. Protein folding takes place with the help of many molecular chaperones and the ubiquitin-proteasome system (UPS). Mis-folded proteins are dealt with either by being bound to by Heat shock protein (HSPs) in the cytoplasm or can be sorted by the endoplasmic reticulum's degradation pathways (ERAD), however both result in the recruitment to the UPS. Alternatively if a mis-folded protein is not dealt with it will aggregate and be removed by autophagy. Extracellular protein homeostasis pathways are also monitored by molecular chaperones, called extracellular chaperones (EC), which will direct mis-folded proteins to specific cell surface receptors where they are engulfed and degraded in the lysosome.

In 2006 a landmark finding was published when the primary component of the cytoplasmic and ubiquitin-positive neuronal inclusions seen in MND was

discovered to be a predominately nuclear RNA binding protein (RBP), TDP43 (Manuela Neumann et al. 2006). Here, both the full length and C-terminal fragment of TDP43 is phosphorylated, cleaved, mislocalised to the cytoplasm and aggregated. This is now recognized as the pathological hallmark of MND and is present in over 97% of all cases.

This was followed by a discovery in 2008 showing that mutations in the gene *TARDBP*, that encodes TDP43, are causative of MND and account for 4% of fMND (Sreedharan et al. 2008). The majority of mutations within *TARDBP* are found within the C-terminal Glycine rich domain, which is a highly unstructured region commonly found in RBPs but the function of this region is largely unknown. A feature of intrinsically unstructured proteins is that its cellular levels are strictly regulated. TDP43 binds to its own 3'UTR causing a negative feedback loop to auto regulate its own expression(Ayala et al. 2011); when bound an intron in 3'UTR is alternatively spliced leading to nonsense-mediated decay (NMD)(Bembich et al. 2014). This prevents the deleterious accumulation of TDP43 (Ayala et al. 2011) however in MND it has been shown that there is a loss of this autoregulation pathway leading to increased protein production (Budini and Buratti 2011). Increased levels of TDP43 have been shown to be toxic in multiple species; yeast, chicken and drosophila (B. S. Johnson et al. 2008; Voigt et al. 2010). In addition, mutant TDP43 has been shown to have increased stability that contributes to the formation of pathological inclusions (Ling, Polymenidou, and Cleveland 2013).

Further, the studies around TDP43 have also raised the hypothesis that dysregulated RNA metabolism plays a key role in the pathogenesis of MND. This notion was further reinforced when additional mutations in another RBP, FUS, were found to cause MND (Vance et al. 2009). Patients with the FUS mutation are among the minority that does not present typical TDP43 mislocalisation, but they do present aggregation of the mutated protein FUS (Kwiatkowski et al. 2009; Vance et al. 2009). FUS has similar structural domains to TDP43; a glycine rich domain, nuclear localization sequence plus 2 RRMs, however unlike TDP43, mutations in FUS are found throughout the gene (Lagier-tourenne, Polymenidou, and Cleveland 2010). Both TDP43 and FUS also have striking functional similarities. They are involved in pre-mRNA splicing, microRNA biogenesis, transcriptional regulation, mRNA stabilization, transport and translation (Lagier-tourenne, Polymenidou, and Cleveland 2010).

However it is not clear whether these mutations cause disease through a loss of function (loss of RBP from nucleus), a gain of toxic function by the formation of cytoplasmic granules and/or nuclear inclusions, or both.

To further strengthen the relevance of RNA metabolism in MND, an intronic hexanucleotide expansion in C9ORF72 was discovered in 2011 that accounts for 40% of fMND making it the most common known cause of MND (Dejesus-hernandez et al. 2011; Renton et al. 2011). Three different pathogenic mechanisms have been pursued for the expansion; haploinsufficiency (loss of function of the protein that contains the repeat), toxic gain of function of the mutant/expansion containing protein, and RNA toxicity due to the RNA produced containing the repeat. Both products from the transcription (RNA repeats) and translation (RAN translation) of expanded gene lead to toxic species found to be aggregated in MND.

Additional causative mutations have been identified that also implicate abnormalities in proteostasis, ribostasis or both, including mutations within genes encoding VCP, VAPB, UBQLN2 and CHMP2B have been discovered to cause adult onset MND (S. Chen et al. 2013). Moreover several of these mutated proteins have been found to be constituent of the aggregated granules in both mutation carrying and sporadic patients, for example OPTN and UBQLN2 (Blokhuis et al. 2013). Together this indicates aggregation and de-regulated proteostasis plays a widespread role in degeneration in MND.

Given the strong evidence to date, it is expected that dysfunction in both RNA and protein metabolism contribute to MND pathology. However an important and unresolved question against this background is: **what is the precise sequence of molecular events that initiates pathogenesis in MND?**

Cell autonomous and non cell autonomous mechanisms of disease

Classically it has been thought that neurodegenerative diseases cause the selective degeneration and death of a subset of neurons, it is becoming increasingly recognized that glia also play key roles in neurodegeneration. Astrocytes (ACs) are the most abundant cell in the nervous system and provide critical support for many neuronal functions. For example they assist by buffering the extracellular environment, particularly ion homeostasis, rapid

removal of neurotransmitter molecules/glutamate from synapse terminal to prevent continuous firing, and to provide metabolic support/supply of nutrients (lactate, antioxidants etc). Given the complexity and importance of AC homeostatic control in the brain it is unsurprising that their involvement has been implicated in several neurodegenerative disorders. In MND the contribution of glia has been shown to act via both cell autonomous and/or non cell autonomous mechanism of injury (Bilican et al. 2012; Di Giorgio et al. 2007; Maragakis and Rothstein 2006; Nagai et al. 2007).

Firstly there have been several lines of evidence to support a non-cell autonomous role for glia in MND. For MND caused by mutations in the gene encoding superoxide dismutase 1 (SOD1), it has been shown that the selective expression of mutant SOD1 protein within ACs is deleterious to healthy MNs (Di Giorgio et al. 2007; Nagai et al. 2007), but that the presence of MNs are important for initiating pathogenesis. It was further showed that damage came from the release of soluble toxic factors. Further for SOD1 related MND early astroglial atrophy has been reported directly surrounding spinal MNs and is observed prior to their degeneration (Rossi et al. 2008), suggesting that the presence of sick ACs has an impact upon surrounding MNs. Glia proliferation and activation has also been associated with disease progression in MND. Since, non cell-autonomous injury by ACs has also been implicated in sporadic MND (Haidet-Phillips et al. 2011; K. Meyer et al. 2014), suggesting the possibility of common pathogenic mechanisms.

Additional evidence for non-cell autonomous role of ACs comes the observation that loss of the astrocytic glutamate transporter GLT-1 leads to MN death by excitotoxicity (J D Rothstein et al. 1995; Jeffrey D. Rothstein et al. 1996). Further evidence to suggest a loss of function mechanism in ACs comes from transgenic and KO animal models which have shown that many AC-specific genes play a role in neuroprotection and degeneration, for example a deficiency in the glutamate transporter EAAT2 (the human equivalent of GLT-1) has been seen in ACs surrounding synapses in human sporadic MND tissue. These converging lines of evidence suggests that a role of ACs in neurodegeneration is due to a loss of their regular supportive functions, but importantly it does not exclude any toxic gain of function mechanism(s).

Next some studies have explored non-cell autonomous pathogenic mechanisms using hiPSC-derived ACs from patients with familial and sporadic MND. ACs differentiated from hiPSCs from either familial (SOD1 and C9ORF72) or sporadic cases, were found to be deleterious to both MN survival and neurite outgrowth in co-culture paradigms. Using this co-culture paradigm, the study was able to determine whether this non-cell autonomous effect was dependent on MND-AC toxicity or lack of support. The addition of wild type AC-conditioned media failed to rescue motor-neuron cell death, suggesting a toxic gain of astrocytic function (K. Meyer et al. 2014).

Separately, there is reason to hypothesize a cell autonomous effect in ACs in neurodegeneration/MND, one being due to many of the known causative genes being ubiquitously expressed. Therefore since the mutation is present in ACs too, it can be reasoned that the mutant protein could also be having aberrant effects in ACs too. Evidence to suggest a cell autonomous role for ACs in MND initially comes from studies of the TARBP (the gene encoding TDP43) mutant M337V. The expression of *TARDBP* mutant M337V in AC resulted in mislocalisation of TDP43 protein, the key pathological feature of MND (Bilican et al. 2012). Additionally, TDP43 pathology has since been seen in ACs of both familial & sporadic MND cases, advocating a possible role for ACs in all TDP43 proteinopathies (Serio et al. 2013). However, when *TARDBP* mutant ACs are co-cultured with either control or *TARDBP* mutant MNs, mutant ACs were not toxic (Serio et al. 2013).

Overall the role and underlying mechanism of both astrocytic cell autonomous and/or non-cell autonomous mechanisms of injury in MND currently remains inconclusive. TDP43 mutations have been shown to affect both MNs (Bilican et al. 2012) and ACs (Serio et al. 2013) in human iPSC cellular models, but there remains some controversy as to whether TDP43 mutations result in non-cell autonomous mechanisms injury. Two studies in 2013 give seemingly contrasting results from cellular and animal models (Serio et al. 2013; Tong et al. 2013). This is also in apparent contrast with the non-cell autonomous toxicity previously reported in the context of sporadic, C9ORF72 and SOD1 mutations (Di Giorgio et al. 2007; K. Meyer et al. 2014; Nagai et al. 2007). However, these results could point towards mutation-specific AC pathology and at least partially divergent underlying mechanisms of disease in ACs in familial MND. The lack of concrete conclusions so far could be due to the lack of

systematic approaches used so far to investigate AC-neuron interactions; a variety of model systems have been used eg. different species, different causative mutations. Hence, in order to precisely elucidate key aspects of cellular autonomy, further systematic strategies are required and could be achieved using hiPSC platform systems.

These data provide a clear rationale for studying both ACs and MNs to investigate disease mechanisms. **I hypothesise that the early pathogenic events in MND differ during the differentiation of MNs and ACs.**

Vasolin containing protein (VCP) and its physiological functions

VCP is a highly abundant, ubiquitously expressed protein and is classified as an AAA+ ATPase (ATPases associated with a variety of cellular activities). It has a substrate and cofactor binding N terminal domain and 2 ATPase domains (with un-equivalent enzymatic activity), that come together to form a hexameric ring structure (H. Meyer and Weihl 2014). It functions as a chaperone for a diverse range of cellular functions, including cell cycle progression/regulation, membrane turnover and plays a primary role in protein homeostasis (Yamanaka, Sasagawa, and Ogura 2012). These various activities of VCP are largely determined by cofactor/substrate availability. One interactor of VCP was discovered to be neurofibromin-1, implicating VCP functions in the regulation of synaptogenesis (Weihl 2011).

VCP is involved several stages of protein control and its plays a major role in protein degradation (Yamanaka, Sasagawa, and Ogura 2012). Intracellular protein degradation is particularly important for non-dividing cells, such as neurons, and functions to remove and prevent the accumulation of damaged or abnormal protein, and it also serves to regulate cellular processes by removing enzymes and regulatory proteins that are no longer needed. It can be accomplished by two pathways; ubiquitin-dependent targeting to the proteasome or via autophagy targeting to the lysosome. VCP plays a role in both systems.

VCP facilitates proteolysis firstly by its interaction with ubiquitin, usually through adapter proteins (often ligases) and secondly via its segregase functions. VCP

can extract proteins for degradation by utilizing energy from the hydrolysis of ATP to aid conformational changes as a segregase or unfoldase (Yamanaka, Sasagawa, and Ogura 2012). It extracts aberrant proteins from organellar membranes (notably both the ER and mitochondrial membranes), chromatin, complex assemblies and also from stalled defective translation products from the ribosome, and assists in their translocation to the cytosol. For example at the mitochondria, VCP's ubiquitin-proteosome functions play an important role in quality control. This is primarily through its interaction with vsm1, which locates to the mitochondria under stress and is required for protection against oxidative stress (Yamanaka, Sasagawa, and Ogura 2012). Additionally at the ER VCP works in a complex with Ufd1 and Npl4 to retrotranslocate ER-associated degradation (ERAD) substrates (Ye et al., 2001). ERAD is the cells mechanism for removing proteins that fail to fold properly in the ER. It occurs by transporting the misfolded proteins out of the ER to the cytosol where they get ubiquitinated and degraded by the proteasome. In addition to ERAD VCP has been associated with other roles at the ER; within neurons, it works in a complex with p47 and ATL1 to regulate tubular ER formation and ER distribution, extension into dendrites and protein synthesis efficiency which in turn influences dendritic spine formation (Shih and Hsueh 2016).

Another core function of VCP is within the cell cycle. Its precise role is not yet fully characterised but many functions have been revealed for VCP throughout several stages in the cell cycle (Vaz, Halder, and Ramadan 2013; Yamanaka, Sasagawa, and Ogura 2012). The transition between phases of the cell cycle is tightly regulated, primarily by CDKs that are co-ordinated by transcription and ubiquitin dependent degradation. VCP is a critical regulator of G1 through degradation of Far1p, an arrest factor (CDK inhibitor) and substrate of VCP-mediated proteolysis (Fu et al. 2003). Additionally VCP regulates DNA replication and S phase progression by interaction with adapter proteins Ufd-1 or Npl-4. Further during mitosis there are profound changes in cellular physiology, eg. spindle disassembly, extensive remodeling of golgi network, ER-golgi morphogenesis, nuclear envelope disassembly and reformation. These events are partially controlled by VCP dependent degradation, for example through its association with aurora-B (spindle and chromatin disassembly), XMAP215/TPX2/Plx1 (spindle assembly) and p47/p37 (regulation of membrane fusion of intracellular organelles, including the ER and Golgi).

More recently there has been an emerging role for VCP in RNA regulation and the translational stress response. VCP has been shown to act in post-transcriptional regulation at the ribosome (Lykke-Andersen and Bennett 2014) and in guiding dendritic pruning through regulating RNA metabolism (Rumpf et al. 2014). Also functions for VCP have been described in the disassembly of mRNA–protein complexes (mRNPs), via binding and extraction of HuR (component of mRNP complexes) leading to proteosomal degradation. mRNPs control transport and stability of mRNAs, but under global inhibition of translational, mRNPs aggregate in stress granules or p-bodies. These are also cleared by a VCP-dependent mechanism, but instead of proteosomal degradation VCP directs aggregates to the lysosome for degradation by autophagy (Buchan et al. 2013).

The role of VCP in MND

1-2% of familial cases of MND are caused by mutations in the gene encoding valosin-containing protein (VCP) (J. O. Johnson et al. 2010; Koppers et al. 2012). Mutations in the VCP gene lead to a broad range of phenotypes including the multisystem degenerative disorder Inclusion body myopathy with Pagets disease of the bone and frontal temporal dementia (IBMPFD) as well as MND (J. O. Johnson et al. 2010; Koppers et al. 2012).

There have been several studies to date exploring the underlying pathogenesis of VCP'opathies (Bartolome et al. 2013; Ju et al. 2009; Ritson et al. 2010; H. F. Wang et al. 2011; Yi et al. 2012; Yin et al. 2012). Interestingly, it has been shown that VCP'opathies are also characterised by the accumulation and mislocalisation of TDP43, the pathological hallmark of MND. Post mortem material from patient cases with either MND or IBMPFD, caused by VCP mutations, have been described to present striking cytoplasmic mislocalization and aggregation of TDP43 in the frontal and temporal lobes (M Neumann et al. 2007). This is replicated in disease models of VCP causing MND; overexpression of disease causing VCP mutations leads to TDP43 mislocalization from the nucleus to the cytoplasm of transfected cells, transgenic mice (Ju et al. 2009) and human post-mortem tissue (J. O. Johnson et al. 2010). This suggests there could be common mechanisms underlying VCP, sporadic MND and TDP proteinopathies.

VCP has been idenitified as a recognition factor; it recognises abnormal proteins and polyglutamine-containing proteins, and its co-localisation has been seen with aggregates in multiple neurodegenerative diseases; nuclear inclusions in Huntington's disease, lewy bodies in lewy body diseases and ubiquitin positive intranuclear inclusions in MND (M. Hirabayashi et al. 2001; Mizuno et al. 2003). On top of this VCP has also been explored as "an effector of neurodegeneration". In the absence of VCP, disrupted aggresome formation and degradation of expanded polyglutamine-containing proteins is observed (Kobayashi et al., 2007). Similarly a number of studies have shown a role for VCP in the disassembly of stress granules (Buchan et al. 2013). Stress granules are aggregates of protein and RNA (stored translational initiation complexes) that appear in the cytoplasm when a cell is under stress. They are usually highly dynamic, continually dissociating and reforming and are present in most neurodegenerative diseases, including MND. In cells carrying a mutation in VCP it has been shown that there is constitutive presence of stress granules (Buchan et al. 2013). A similar phenotype is observed when autophagy was inhibited (Buchan et al. 2013).

Given VCPs functions in protein degradation, it is not surprising that defective protein clearance has been highlighted in VCP LOF models. Firstly loss of VCP activity has been shown to lead to the accumulation of ubiquitinated proteins and impaired ERAD (Dalal et al., 2004; Wójcik et al., 2004). It has also been shown that the loss of VCP causes the accumulation of autophagosomes, implicating that VCP is involved in the maturation and targeting of autophagosomes to the lysosome for degradation (Ju et al. 2009). Defects in protein clearance also seen with other MND causing mutations (TDP43, SOD1) and in sporadic MND suggesting there could be a common theme across MND pathogenesis (Kiskinis et al. 2014; Ling, Polymenidou, and Cleveland 2013; Manuela Neumann et al. 2006). On top of protein clearance further defects in protein homeostasis have been shown in VCP mutant cells; rER morphology is impaired in VCP mutant mice neurons resulting in reduced protein synthesis, which in turn leads to lower expression of synaptic proteins and a reduced synaptic density, a characterised feature in MND (Shih and Hsueh 2016).

Other factors that have been explored in VCP'opathies include mitochondrial dysfunction (Bartolome et al. 2013) and synaptogenesis (H. F. Wang et al. 2011), which could both cause cellular death. A loss of VCP or presence of a

clinically relevant mutation in VCP was shown to lead to reduced mitochondrial membrane potential, increased oxidative phosphorylation and reduced ATP levels (Bartolome et al. 2013). Mitochondrial deficiency has also been more broadly implicated in MND (reviewed in. Cozzolino and Carri 2012) however to date it is not clear whether it is a primary event in disease pathogenesis. In addition VCP mutants have been shown to lose interaction with neurofibromin-1 leading to dendritic spine loss in hippocampal neurons (H. F. Wang et al. 2011). Since synaptic deficiency is observed in MND (Devlin et al. 2015; Pun et al. 2006; Shoichi Sasaki and Iwata 1995; Shoichi Sasaki and Maruyama 1994) this function of VCP in synaptogenesis may be of importance in disease pathogenesis.

For my PhD I have selected to study VCP-related MND. This is because it is pathologically representative of the majority of MND cases (i.e. displays nuclear to cytoplasmic mislocalisation of TPD-43); which excludes mutations in FUS and SOD1. Secondly VCP mutations are rare and we are one of few labs with the unique opportunity to study MND using hiPSCs carrying a VCP mutation. **Here I aim explore RNA regulation and to define the earliest molecular events in VCP-related MND**.

Models of disease

There are many model systems which can be used to study MND and until recently a lot of research into MND has been conducted using mouse models. These approaches have generally had a poor transitional yield, which could be due to a number of reasons. Potential therapies have often showed promise in pre-clinical animal models but did not translate successfully in clinical trials. It is possible that fundamental interspecies differences and poor timing in pre-clinical study design may explain the failure in clinical translation. For example ceftriaxone, a modulator of the glutamate receptor EAAT2, failed in phase III human clinical trails showing no efficacy but in the SOD1-D93A mice is was shown to reduce excitotoxicty and increasing lifespan when administered before symptom onset (Fonslow et al. 2012; Jeffrey D Rothstein et al. 2005).

Furthermore, many models used have not been truly representative of human MND. This can be highlighted as a further potential weakness in previous trial designs, where many pre-clinical animal models have overexpressed mutant

SOD1. In human MND, SOD1 mutations accounts for only a minority of the total cases and these patients do not present with the pathological hallmark that >97% of sporadic and familial MND cases do (a nuclear to cytoplasmic mislocalisation of TDP43). These facts suggest that SOD1 mutants may not be pathomechanistically representative of MND. Another model generated to study MND has been in mice carrying mutations in *TARDBP* gene, although some events in MND are recapitulated (eg. detection of ubiqutinated proteins and MN specific degeneration is seen) there are striking differences in the symptoms observed compared to human MND. Only 25% of MNs are seen to degenerate compared to 90% in SOD1 and human forms of MND, and they develop a severe gastrointestinal pathology, which is the cause of their premature death (Esmaeili et al. 2013). On the other hand animal models are an invaluable tool as they enable us to study the early stages and progression of disease within an vivo environment.

In addition, post-mortem tissue from MND patients has been used to study disease pathology, although it is only representative of the end stages of disease. Therefore the use of human induced pluripotent stem cells offers an unparalleled approach to study early molecular pathogenesis in MND.

Human induced Pluripotent Stem cells

Human induced pluripotent stem cells (hiPSCs) are stem cells derived from reprogramming adult somatic cells, usually fibroblasts, to an embryonic-like state.

They hold several major advantages as a new approach to modelling human development and disease;

i) They have theoretically a limitless self renewal capacity
ii) capable of differentiating into any cell type in the body across the three germ layers, which is a particular advantage for deriving neuronal cell types as they are difficult to obtain from human tissue.
iii) fully humanized model system and
iv) mutations are expressed at representative pathophysiological doses removing the need for artificial overexpression, knock down or knockout studies.

Additionally as this is essentially a developmental system, iPSCs allow us to recapitulate the earliest pathogenic events occurring during disease manifestation in a clinically relevant model system.

As with all models there are limitations too. One of the biggest criticisms of using hiPSCs to model neurodegenerative diseases, where aging is a major risk factor, is that the relative age of the hiPSC-derivatives has been shown to be comparable to a fetal maturational stage (Defects et al. 2015; Miller et al. 2013; Rickie Patani, Lewis, et al. 2012). Another major limitation using current hiPSCs models is that they lack the dynamic and complex in-vivo environment, found within an organism. However, this could be advantageous when trying to model complex diseases such as MND since we can focus on understanding the contribution of the different cell types separately to begin with.

The use of hiPSCs as a model system for MND has previously been demonstrated, for example cells carrying TDP43 mutations have been shown to recapitulate key aspects of MND, such as selective vulnerability, premature death and elevated levels of TDP43 (Bilican et al. 2012). By revealing specific disease-associated phenotypes this thereby validates the use of this model system for investigating the underlying mechanisms of MND and developing new screening platforms for therapeutic intervention. Although many disease-relevant phenotypes have been shown, the precise time-course of pathogenic events has not yet been studied therefore for my PhD **I aimed to temporally analyse neural derivatives from control and VCP mutant hiPSCs, in order to study the earliest events in pathogenesis of MND**. I will differentiate both MNs and ACs in order to analyse cell specific events.

Neuronal development in humans

In order to take full advantage of iPSC technology as a disease model system, as well as for drug discovery and regenerative medicine, it is critical to establish accurate in vitro methods for directed differentiation. It has been a major challenge in neuroscience to understand the developmental 'logic' of human neuronal and glial specification, which in turn is used to guide optimal differentiation strategies of pluripotent stem cells (PSCs).

The development of the human nervous system can be divided into 3 main stages. Beginning from fertilization, cells undergo asymmetric division to produce the early blastocyst which is divided into the inner cell mass (ICM) and trophectoderm. The ICM gives rise to the epiblast, which in turn gives rise to

three germ layers (ectoderm, mesoderm and endoderm) during gastrulation. The nervous system is derived from the ectoderm lineage. Therefore firstly, like other mammals, the development of the human nervous system begins with the induction of neuroectoderm; an area of the ectoderm is specified to form the neural plate, then this folds to form the neural tube. Secondly the neural precursor cells (NPC) within the neural tube are patterned, which refers to the spatial separation of NPCs into distinct progenitor domains. And finally these 'patterned precursors' go on to give rise to defined neuronal cell types; first neurons followed by glia. All these processes progressively restrict cell fate and are tightly regulated by a variety of coordinated mechanisms including gene expression programmes, cell-cell contact and positional relationships, cell signaling and epigenetic modifications.

i) Neural induction

Neural induction is a complex process where ectodermal cells are specified to the neural lineage. In 1924 Spemann and Mangold discovered 'the organiser', which was an area of specialised cells (located in the dorsal blastopore lip) capable of inducing the neural plate in the ectoderm. Since it has been shown that the organiser can induce neural fate when transplanted ectopically within and across species (Stern 2005) by the release of signals to inhibit BMP4 (noggin, chordin, follistatin). The concept of inductive signals orchestrating neural induction was widely accepted until the proposition of the Default model, which hypothesises that ectodermal cells become neurally specified by default whereas the epidermis is induced by BMP4 signals (Grunz and Tacke 1989; Hemmati-brivanlou and Melton 1994). Since it has been accepted that neural induction occurs through inhibitory signalling that suppress epidermal cell fate. A gradient is established between BMP ligands and antagonists controls epidermal versus neural fate. However there are many challenges to this model (Stern 2005), for example BMP4 inhibition isn't sufficient for neural induction (allows expression of sox2 but not sufficient for sox3 an early marker), suggesting the involvement of additional mechanisms.

In addition to BMP antagonism other members of the TGF-b signaling family have been identified to play a role in neural induction. Nodal inhibition in neural induction is well established and has been tested across a variety of species (Camus et al. 2006; Chambers et al. 2009). Further factors have also been discovered to contribute to neural induction and include FGF and WNT

signaling, however their exact role remains controversial and is likely to depend heavily upon the timing (Zirra, Wiethoff, and Patani 2016).

ii) Patterning (Motor neuron precursor specification)

There are many different types of neurons, located at defined positions in the nervous system with specific functions in functional circuits. Patterning is the process that specifies neural precursor positioning in the developing embryo, which in turn influences their differentiation fate. The default positional identity in the developing embryo after neural induction is anterior (i.e. in the forebrain), so upon terminal differentiation, the resulting neural precursors will give rise to cortical neurons. Developmentally, MNs in the spinal cord arise from the pMN domain in the ventral horn, where transcription factors NKX6.1, Pax6 and Olig2 are expressed (Wichterle et al. 2002).

In vivo neural patterning is chiefly controlled by secreted extracellular signaling molecules called morphogens. Morphogens are molecules that spread over variable distances, here across the neural tissue, establishing different signaling gradients, which often function in combination and in opposing directions (eg. BMP4 and SHH across the neural tube). The activity of these molecules is spatiotemporally integrated to determine the specific combination of transcription factors that are activated in distinct compartments of the CNS. The same signals are used reiteratively, at different times and positions during development.

Neural patterning occurs across both the anterior to posterior (AP) and dorsal to ventral (DV) directions in the neuroaxis, giving rise to the diversity of cells forming the nervous system. AP patterning occurs prior to DV patterning. Anterior to posterior fates are progressive re-specified by a number of signaling molecules; including fibroblast growth factors, retinoic acid, transforming growth factor 10 and WNTs (Zirra, Wiethoff, and Patani 2016). An approximate AP axis is set up during neural induction, and then secondary organizers are established. Elongation of the AP axis is regulated by retinoic acid (RA) and FGF, which also coordinate the timing of differentiation; RA (among many other functions) caudalises progenitors and promotes cell to become neuronal/cell cycle exit and FGF inhibits cell differentiation to prevent it occurring too early while neural tube still being formed etc.

The distinct domains along the AP axis are defined by expression of the homeotic (HOX) genes. This is established by graded signaling throughout the hindbrain and spinal cord to establish positional boundaries of HOX gene expression. Most anterior progenitors express HOXa1 and HOXb1 only, then the more posterior regions express higher numbered HOX paralog genes; at 1-5 in hindbrain and Cervical region, 4-8 in brachial region, 8-9 in the thoracic, and 10-13 in lumbar region of the neural tube (Philippidou and Dasen 2013).

AP patterning is followed by DV patterning and again it requires the interplay of graded concentrations of selected cues in a specific window of time to direct lineage restriction. Signals for DV patterning are mostly the same at all levels of AP axis. Opposing gradients of BMP and sonic hedgehog (SHH) establish cell patterning across the neural tube; BMP signals from roofplate/dorsal and SHH in notochord and floor plate/ventral. These morphogenic signals work together to create distinct progenitor domains where a unique combination of transcription factors are expressed leading to differentiation of a select subtype of neurons. Opposing signaling gradients result in the expression of unique combinations of homeodomain and basic helix-loop-helix transcription factors, which results in the establishment of discrete progenitor domains. SHH signaling promotes ventralisation and the expression of class II transcription factors OLIG2, NKX2.2 and NKX6.1, which typically arise in the pMN domain of spinal cord and give rise to motor neuron precursors.

Within the spinal cord MNs can be further subdivided. They reside in distinct motor columns: at cervical and lumbar levels they are split into the lateral (LMC) and medial motor columns (MMC), and at thoracic levels into the preganglionic (PGC), MMC, and hypaxial motor column (HMC) (Philippidou and Dasen 2013). Each motor column is distinguished by the different combinatorial expression of a few genes; LHX3, HB9, ISL1 and FOXP1.

iii) Terminal differentiation of neurons and glia

Although NPCs cell fate has been restricted in their developmental potential they are still a type of multi-potent cell that can self-renew and are capable of differentiating into both neurons and glia. To begin with NPCs rapid proliferate and expand through symmetric divisions but in later stages of development NPCs undergo asymmetric division that gives rise to two distinct daughter cells; one NPC and one differentiated cell. NPCs potency is temporally regulated,

first giving rise to neuronal cell types but later their phenotypic potential becomes restricted and they give rise to glial cells (ACs and oligodendrocytes). However the mechanisms underlying this regulation are not completely characterised yet.

The developmental program in NPCs that controls their terminal fate into either neurons or glia is known as the gliogenic switch. There is limited progress in understanding the regulation and timing of this switch but both cell intrinsic programs (epigenetic programs and transcription factors) and by external cues, including growth factors and cytokines released from neighbouring tissues, have been shown to contribute (Kasai, Satoh, and Akiyama 2005; Majumder et al. 2013; Morrison et al. 2000). Consistent with the sequence of developmental events, pro-neural genes first act whilst gliogenic genes remain supressed, then later pro-neural genes will be inhibited while a gliogenic fate is promoted.

The balance of neuronal and astrocytic signals need to be precisely to controlled, firstly to ensure that a dividing undifferentiated population of NPCs remains available (to later differentiate into glia) and secondly at later stages pro-neural signals must be inhibited to efficiently promote a gliogenic potential but the timing is crucial to ensure all neurons that are required in the adult brain are generated before neurogenesis is terminated. Notch signalling is critical for several stages of control; it promotes self-renewal of NPCs, inhibits neurogenesis and promotes gliogenesis (Zhou et al. 2010). Signals to maintain the self-renewable fate of NPCs come from neighbouring terminally differentiated neuronal cells, through a process known as lateral inhibition. Once a cell is committed to neuronal differentiation it signals to the surrounding NPCs to prevent all progenitors differentiating into neurons. Notch signals via specific bHLH transcription factors, including HES1, HES3, HES5, which signal to maintain the NPC pool and in addition to promote gliogenesis, partially through stat3 signalling (Zhou et al. 2010).

In contrast to Notch signalling, BMP and WNT signals have been shown to promote neuronal differentiation but also act to promote astrocytic differentiation (Wen, Li, and Liu 2009). It is a common theme that the same pathways often operate at different times during development to produce contrasting cellular effects. Here the differing effects of BMP/WNT signalling could be due to the contribution of epigenetic factors. Over time in NPCs the

demethylation of AC-specific sites, eg. the GFAP promotor where stat3 binds to promote gene expression, enables NPCs to become more responsive to signals promoting gliogenesis (Hatada et al. 2008). Among the most prominent signals promoting astroglioenesis is the JAK-STAT pathway, which becomes robustly activated during the transition from neuronal to glial differentiation, triggered by cytokines and/or growth factors such as CNTF, LIF and IL6. DNA methylation tightly regulates JAK-STAT signaling during neurogenesis (Fan et al. 2005).

After the gliogenic switch, precursors migrate to populate all areas of the CNS, expand and terminally differentiate into ACs. AC specification initially generates immature ACs, which are essential for coordinating with neurons the formation of functional synapses. During early post natal stages ACs undergo drastic morphology changes and maturation (Freeman 2010). Mature adult ACs are involved in a variety of functions including insulating synapses and removing excess transmitters preventing excitotoxicity.

RNA regulation in neuronal cells

Gene expression is controlled at many stages before the production of a functional protein; these stages include transcription (initiation and elongation), RNA processing, translation, post-translational regulation & modifications to control degradation/abundance, stability and localization. Such regulation is essential for maintaining cellular homeostasis, and coordinating rapid cellular responses to stimuli. It is also particularly important in highly polarized, complex cells such as neurons as it enables diverse cell type specific functions (Rickie Patani, Sibley, et al. 2012).

Neurons heavily rely on gene regulation at the RNA level, expressing more microRNA, alternatively spliced mRNA and small RNA than any other cell type. Due to improved methods, our knowledge of how RNA is processed and regulated differently within neurons has greatly increased (Kiebler, Scheiffele, and Ule 2013). Post-transcriptional mechanisms play a key role in neuronal RNA regulation, heavily influencing mRNA transport and local translation (Kapeli and Yeo 2012). For example the localization of mRNA to dendrites is critical for the integration of multiple stimuli from thousands of cells in a timely manner and the stimulation by synaptic input can trigger local mRNA

translation at the synapse terminal (Bramham and Wells 2007). Additionally many neuron-specific proteins have been identified, including neuron-specific RBPs such as NOVA, ELAV and NeuN (Darnell 2013). Their specialized RNA regulatory networks allow for increased cellular complexity in neurons, localization of messenger RNA (mRNA) to distinct subcellular environments (eg. axons, dendrites) and regulation of their gene expression profile in response to synaptic stimuli.

The importance of RNA regulation and selected RBPs in neurons is further highlighted through the association with several neurological diseases. A key example being MND, where causative mutations have been discovered within RBPs and intronic regions, abnormal splicing and other post transcriptional defects have also been shown to contribute to the underlying pathology (Almeida et al. 2013; Kwiatkowski et al. 2009; Manuela Neumann et al. 2006; Rogelj et al. 2012; Tollervey et al. 2011; Vance et al. 2009). Thus emphasizing the importance of studying RNA regulation within neuronal cells, which could possibly uncover new cell-specific and/or disease specific regulatory mechanisms.

Regulation by RBPs

All stages of RNA regulation are tightly controlled by RNA binding proteins (RBPs). These RBPs interact extensively with different RNAs in a spatiotemporally regulated manner to carefully orchestrate gene expression. They regulate gene expression at many different stages to influence cellular functions and their roles generally differ within nuclear and cytoplasmic compartments of the cells (Glisovic et al. 2008). In the nucleus, RBPs regulate RNA helicase activity, RNA polymerase elongation, mRNA splicing, maturation, and nuclear export. In the cytoplasm, they control RNA transport, silencing, translation, and degradation.

RBPs interact with many mRNAs and proteins via conserved domains, glycine rich regions and RNA recognition motifs (RRMs). In the cytoplasm they form dynamic complexes called RNA granules, which help regulate RNA activity and distribution. Depending on their composition they are subdivided into ribonucleoprotein complexes (RNPs), stress granules (SG) and P-bodies (Vanderweyde et al. 2013). Neuronal function depends on RNPs for storage of mRNAs (for rapid response to metabolic or environmental changes) and for the

transport of mRNAs, RBPs recognize localization elements in UTRs of mRNA enabling them to direct the localization while also suppressing translation until required. This has been well studied in relation to dendritic transport of RNAs and synaptic plasticity as mentioned above (Doyle and Kiebler 2011). P-bodies are involved in control of mRNA degradation and often interact with miRNA machinery. SG control the stability and translational activity of selective mRNAs. They play a critical role in fine-tuning protein expression under conditions of stress and more recently have been implicated in neurological disease conditions too, again emphasizing selective neuronal functions (Vanderweyde et al. 2013). A number of disease associated RBPs have been demonstrated to colocalise with and/or influence dynamics of SG (eg. Ataxin2, TDP43) (Colombrita et al. 2009; Dewey et al. 2010; Nonhoff et al. 2007; Vanderweyde et al. 2013). Additionally unfolded protein aggregates are a common feature in the pathology of MND, over 95% of patients present with cytoplasmic aggregates of the RBP TDP43 (Manuela Neumann et al. 2006). Protein aggregates share many similarities with SG; interactions within complexes are mediated by low complexity regions such as glycine-rich domains, they contain ubiquitinated proteins and are dissociated by molecular chaperones. Nevertheless there are some fundamental differences too, under healthy conditions SG serve a characterized biological function and their formation is reversible, whereas in disease the function of the aggregates seen in MND/neurodegeneration is not yet fully understood but insights have shown that their dynamics are altered and formation becomes irreversible (Baron et al. 2013; Molliex et al. 2015).

Methods to study RNA regulation

There are many different techniques available to study RNA regulation, including RNA sequencing (RNAseq), ribosome profiling and RT-PCR. Methods to study RNA regulation have advanced significantly in the past decade or so, from traditional biochemical approaches where a single RNA/RBP is studied at once to techniques that allow us to view entire cellular events on a global scale. Using high-throughout methodologies enables us to better grasp the complexity of RNA regulation in a given biological sample. These genomic techniques produce large amounts of data that require considerable computational analysis, which must be handled appropriately to ensure correct interpretation of biological relevance. Here I have chosen to use

RNA sequencing and iCLIP methodologies to begin to study RNA regulation in MND.

i) RNAseq

Initially microarrays opened up the opportunity to get a general view of the overall RNA activity, however there were many challenges associated with the method including a high signal-to-noise ratio and varying probe specificity and sensitivity. Subsequently the development of RNA sequencing has revolutionized genomics technologies by allowing us to capture a 'snap-shot' of the entire transcriptome in a high throughput, unbiased way. This in turn permits analysis of cellular dynamics by looking at gene expression and splicing changes in a highly quantitative and accurate manner. Advantages of using RNAseq over other techniques, such as microarrays, include:

- The ability to detect transcripts that are not known (therefore there is no bias into what you select for as it doesn't rely on any known genomic sequences), novel splice variants etc.
- Allows some analysis of post transcriptional changes
- Can detect Mutations/SNPs

ii) Individual crosslinking and immunoprecipitation (CLIP)

Given the key role of RBPs in the regulation of gene expression it is important to understand their precise functions, which we can begin to characterize by identifying a selected RBPs RNA targets and binding sites to nucleotide resolution. Single RBPs bind many RNAs and most RBPs bind to several sites on an individual RNA, so it is important to understand the complete landscape of protein-RNA interactions. Individual crosslinking and immunoprecipitation (CLIP) allows this. Further we can integrate this positional information with a functional read out such as RNAseq data to determine its effect on the transcriptome.

There have been many techniques developed to study RNA-protein interactions in vivo beginning with RNA immunoprecipitation protocol (RIP). RIP detects RNA molecules bound to a protein of interests by immunoprecipitation followed by rt-PCR, microarray (RIP-chip) or high-throughput sequencing (RIP-seq). However there are several major drawbacks to this methodology; firstly is the reliance on the specificity of protein-RNA interaction being maintained throughout cell lysis which can lead to the

detection of non-specific, false-positive interactions, and secondly RIP is unable to pinpoint the site of RNA interaction which is an important consideration in the mechanistic implications of RBP-RNA interactions.

In 2003 Cross-linked immuno-precipitation (CLIP) was developed, which involves in vivo UV-cross-linking of RNA-binding proteins (RBPs) to interacting RNA targets (within 1A distance of each other). UV crosslinking only crosslinks directly bound nucleic acid-protein interactions - a major advantage compared to the use of formaldehyde, which had previously been used but was found to induce protein-protein crosslinking too and other artifacts leading to poor distinction between direct and indirect interactions (Ule et al. 2005). UV crosslinking allowed for more stringent purification conditions throughout and massively improved signal-to-noise ratio and the detection rate of non-specific interactions. Since, cross-linking of RNA-protein complexes has emerged as the standard for identifying functional interactions. Following crosslinking in the CLIP method is cell lysis, RNase digestion to generate optimal size fragments for sequencing, immuno-purification of the RBP-of-interest, treatment with proteinase K to digest the bound protein and linker ligation to allow RT-PCR amplification of the bound RNA to generate a library for sequencing (sanger was originally used) (Ule et al. 2003). An important quality control step included is the SDS-page purification of RBP-RNA complexes, after linker ligation. This visualization allows for selection of RBP of interest and removed free RNA.

Subsequently there have been several variations of the CLIP methodology. Firstly high-throughput sequencing of RNA isolated by CLIP (HITS-CLIP) combined CLIP with high-throughput sequencing giving a greater sequencing depth which enabled a more comprehensive map of RBP-RNA interactions in vivo (Licatalosi et al. 2008). Still a major drawback of the approach is the resolution of binding which is correlated to the length of fragmented RNA (typically 30-60nts). New strategies were developed shortly afterwards to overcome this, in 2010 Hafner et al present photoactivatable-ribonucleoside-enhanced CLIP (PAR-CLIP) (Hafner et al. 2010) Briefly, PAR-CLIP uses photoactivatable nucleotide analogues such as 4-thiouridine (4-SU) or 6-thioguanosine (6-SG), which are cross-linked efficiently with UVA at 365nm (oppose to the standard with UVC at 254nm) and cause a base transition at the crosslink site during reverse transcription. This enables us to detect the crosslink site by analyzing the site of mutation in the resulting cDNA

sequences. The approach is of course limited to cell cultures that can effectively incorporate photoactivable nucleotide analogues.

In addition, in 2011 Konig et al demonstrate individual nucleotide-resolution CLIP (iCLIP) (König et al. 2011). Until 2011 it was established that reverse transcriptase read through sites of residual peptide-RNA cross-links. However Zhang and Darnell discovered that the reverse transcriptase used in CLIP skipped a nucleotide at the crosslink site due to a couple of peptides remaining attached to the RNA after proteinase K digestion (Chaolin Zhang and Robert B. Darnell 2011). Thus errors at these sites can be leveraged to map protein-RNA interactions with single nucleotide resolution. Moreover iCLIP was developed where they reasoned that reverse transcription would be truncated at the site of binding due to a couple of peptides remaining attached to the RNA after proteinase K digestion. Therefore, they were able to show reverse transcription arrest at the binding site could be used to define cross-link sites to nucleotide resolution; the mapped truncation site is located one nucleotide downstream of the cross-link site. In iCLIP the adapter ligation is altered to accommodate the transcriptional arrest. A unique adapter containing P3 and P5 sequences is used, and after reverse transcription the RNA and adapter are circularized, then cut between P3/P5 sequences before PCR and sequencing. All approaches discussed are multi-step and multi-day protocols, which require tightly controlled conditions to ensure accurate profiling of the RBP-of-interest (Huppertz et al. 2014), but more comprehensive review and comparison of the approaches is discussed elsewhere (König et al. 2012; Modic, Ule, and Sibley 2013).

To date CLIP has been successfully used to study multiple RBPs revealing their diverse role in post transcriptional regulation, for example regulation of alternative splicing by Nova (Licatalosi et al. 2008) or competitive binding between RBPs to control cryptic element exonization such as hnRNPC and U2AF65 (Zarnack et al. 2013). It has also revealed important insight into disease linked RBPs such as TDP43 and FUS (Lagier-Tourenne et al. 2012; Rogelj et al. 2012; Tollervey et al. 2011). Lastly, CLIP has recently been adapted for the simultaneous capture and monitoring of all RBPs following oligo-dT capture of mRNA (Baltz et al. 2012). Fittingly, CLIP is now considered the method of choice to study the control of post-transcriptional networks by RBPs.

Aims of this thesis

The overarching aim of this thesis is to capture the earliest pathogenic events that occur in MND using control and VCP-mutant carrying patient derived hiPSCs. I use a combination of patient specific hiPSC-derived neural derivatives and functional genomic technologies, including RNA sequencing (RNAseq) and individual nucleotide resolution UV cross linking with immunoprecipitation (iCLIP), to understand the role of dysfunction RNA regulation in VCP-related MND. Further, I address collaboratively the presence and onset of other phenotypes, that have been previously associated with MND in our VCP hiPSC derived neural cells.

More specifically my aims for the PhD are;

1) To develop and refine directed differentiation strategies for hiPSC to spinal cord MNs and ACs.

2) Optimise the functional genomic techniques, RNAseq and iCLIP, to explore molecular events in hiPSCs.

3) To conduct a temporal and unbiased transcriptome-wide analysis of early molecular events in MND caused by VCP mutations using RNA sequencing.

4) In Parallel to 2), conduct a temporal analysis through the differentiation and maturation of MNs and ACs to examine a range of other phenotypes (rationalized by pathways previously implicated in MND) that could contribute to VCP-related MND.

5) Characterise the role of TDP43 in VCP-related MND

Chapter 2; Methods

iPSC culture

Derivation of Human Fibroblasts and iPSC Generation

Dr. Selina Wray and team acquired dermal fibroblasts and cultured them in OptiMEM +10% FCS medium. They transfected the following episomal plasmids into the fibroblast samples for iPSC generation: pCXLE hOct4 shp53, pCXLE hSK and pCXLE hUL (Addgene), as previously reported (Okita et al. 2011). Informed consent was obtained from all patients prior to skin biopsy.

iPSC maintenance

hiPSC were propagated using an adherent monolayer system on plates pre-coated with Gel-trex basement membrane matrix (150µg/ml) with Essential 8 medium (containing DMEM/F-12, L-ascorbic acid, selenium, transferrin, NaHCO3, insulin, fibroblast growth factor-2 and transforming growth factor β1) (A1517001, Thermo Scientific).

Passaging: hiPSC were passaged when cells reached approximately 70% confluency using 0.5mM Ethylenediaminetetraacetic Acid (EDTA) in DPBS.

Freezing procedure: EDTA is used to dissociate the cells, taken up in 90% medium and 10% dimethylsulfoxide (DMSO), a cyroprotectant and placed in a freezing vial. Vials were placed at -80°C overnight in a cyropreservant container, that controls the rate of cooling, then transferred the following day for long term storage in liquid nitrogen.

Thawing procedure: The cryovial of cells was held in a 37°C water bath until only an ice crystal remains, 1ml of media was then added dropwise onto the cells, which were next added to a larger tube containing approx. 10mls of medium, centrifuged at 180x G for 5 minutes, resuspended in 1ml of medium and plated onto pre-coated plates.

Motor Neuron differentiation

The differentiation of hiPSCs to motor neurons was divided into 3 stages;

1) Neural conversion

For neural conversion cells were plated to 100% confluency and fed for 10days with neural maintenance media (consisting of DMEM/F12 Glutamax, Neurobasal media, Pen Strep, L-Glutamine, N2 supplement, Non essential amino acids, B27 supplement, B-mercaptoethanol and Insulin) supplemented with 10μM SB431542 and 1μM Dorsomophin – (dual SMAD inhibition). Media was changed daily during these 10days. After 10days the neural epithelial sheet is lifted, washed 3x in neural maintenance media and spun for 1minute at 1000g, before plating down onto laminin coated plates (laminin used at 10-40μg/ml in DPBS). The neural epithelial sheet was lifted by spiking dispase (1:5) into the media and incubating at 37°C for approximately 15minutes, once the edges started to lift detachment of the sheet was aided by gently tapping the culture dish and/or gently pipetting up and down around the edges using a p1000.

2) Patterning

The next phase in differentiation is patterning. For generation of spinal cord motor neurons we fed the neurally converted cells daily with neural maintenance media supplemented with 1μM Pur (ventralise) and 0.5μM RA (caudalise) for 8days. During these 8days cells are cultured in clumps and are were fed with extra care in order to not disrupt cellular clumps/cell-to-cell signaling, sometimes this involved only partially removing the media and replacing with fresh.

3) Terminal differentiation

After patterning, NPCs were split using accustase to promote a single cell suspension (media was removed and NPCs were incubated with accutase at 37°C for 2-3minutes). Cells were then terminally differentiated on laminin coated plates (laminin used at 20-40μg/ml in DPBS) fed with neural maintenance media for a minimum of 35 days.

Accelerated motorneurogenesis protocol (A more detailed version of the protocol can be found in appendix 8.1)

hiPSCs were first differentiated to neuroepithelia by plating hiPSC to 100% confluency and feeding with neural maintenance media supplemented with 3 compounds; 2μM SB431542, 1μM dorsomorphin and CHIR99021 (3 μM) for 7 days. At day 8, the neuroepithelial layer was lifted using dispase, plated onto

laminin coated plates and patterned for 7days with 0.5μM RA and 1μM Purmorphamine, to respectively caudalise and ventralise precursors to the pMN domain. At day 14 progenitors were expanded in neural maintenance media supplemented with 0.1μM Purmorphamine for an additional 4 days patterning before being plated on fresh laminin for terminal differentiation. To generate synchronized postmitotic neurons, the cultures were treated from day 18 with 0.1μM compound E to promote cell cycle exit. During terminal differentiation media was changed 2-3 times weekly.

Neural Propagation

After Neural Patterning (Day14), neural precursors can be expanded/propagated before terminal differentiation to increase the yield/material generated. For this purpose neural precursors are plated on gel-trex and fed with neural maintenance media supplemented with 10ng/μl FGF for a maximum of 30days. Precursors are then terminally differentiated as above in neural maintenance media plus 0.1μM compound E on laminin-coated plates.

Astrocyte differentiation

For astrocyte differentiation, hiPSC under went neuronal conversion and patterning as described above. Afterwards, they were propagated in gel-trex coated flasks in neural maintenance media supplemented with 20ng/μl FGF for 50-90days. They were passaged approximately every 7days using EDTA and as they aged, progressively harsher trituration with a p1000 was used. Media was changed 2-3 times per week. For terminal differentiation propagated precursors were plated onto freshly coated gel-trex plates and fed 2-3 times weekly with neural maintenance media supplemented with 10ng/μl BMP4 and 10ng/μl LIF for a minimum of 2weeks.

Astrogliogenesis after an accelerated conversion to spinal neural precursors

For astrocyte differentiation, hiPSCs also undergo an accelerated neural conversion and patterning using the method described above. As before an additional propagation phase is required before terminal differentiation to allow NPCs to undergo the gliogenic switch, this was again done by plating precursors on gel-trex coated plates and feeding 2-3 times weekly with neural maintenance media supplemented with 10ng/μl FGF. Terminal differentiation is

as before (treated with 10ng/μl BMP4 and 10ng/μl LIF for a minimum of 2weeks).

Electrophysiology

MNs were plated at approximately 50,000 cells/cm2 on glass coverslips and allowed to mature for 2-3 weeks. Whole-cell patch-clamp recordings and analysis were carried out by Sarah Crisp, in Dimitri Kullmann's lab. Coverslips were continuously perfused in a recording chamber with extracellular solution containing 125 mM NaCl, 2 mM CaCl2, 1 mM MgCl2, 2.5 mM KCl, 10 mM HEPES, 30 mM glucose, pH 7.3 with NaOH, 290 mOsm, at room temperature. Neurons were visualised using an Evolve Delta EMCCD camera (Photometrics) and MultiManager software, connected to an Olympus IX73 microscope with difference interference contrast optics. Recordings were made using 5-6 MΩ pipettes filled with K-gluconate intracellular solution (126 mM K-gluconate, 4 mM NaCl, 1 mM MgSO4, 0.02 mM CaCl2, 0.1 mM BAPTA, 15 mM glucose, 5 mM HEPES, 3 mM MgATP, 0.1 mM NaGTP, pH 7.3 with KOH, 280 mOsm with sucrose), to achieve an access resistance of less than 20 MΩ. Cells showing >20% change in series resistance during the course of recording were rejected. Current clamp recordings were obtained using a MultiClamp 700B amplifier (Molecular Devices), acquired at 50 kHz and filtered at 10 kHz. Data acquisition and offline analysis were performed using LabView 2014 (National Instruments). The calculated liquid junction potential was 15.1 mV. All values shown are corrected for liquid junction potential. After setting the bridge balance and adjusting the holding current to keep the cell at −70 mV, 1-second-long current injections, in 1-10 pA incremental steps, were delivered. A hyperpolarising step was used to determine passive properties. Depolarising steps were used to elicit action potentials. Parameters of action potentials were measured for the first action potential (defined as peak > 0 mV) elicited with each current injection. Values are reported as mean ± s.e.m.

Imaging techniques

Immunocytochemistry (ICC)

Cells were plated onto pre-coated Ibidi 8 well chamber slides and left to attach/grow for a minimum of 4hours. Medium was removed and cells were washed twice in phosphate buffered saline (PBS) before being fixed in 4% paraformaldehyde in PBS for 30 minutes. Cells were washed twice each for ten minutes in PBS supplemented with 0.3% Triton-X (PBST). Non-specific

antibody binding was blocked with 5% normal goat serum in PBST for 60 minutes and then incubated overnight at 4°C with primary antibody in 5% normal goat serum in PBST. Cells were washed three times for ten minutes each in PBST then incubated with complementary Alexa Fluor secondary antibody at 1:200 dilution in 5% normal goat serum in PBST for one hour at room temperature in the dark. Cells were washed once in PBST containing 4',6-diamidino-2-phenylindole, dihydrochloride (DAPI) nuclear stain (two drops in 1 ml PBST) for ten minutes then washed a further two times for ten minutes in PBST and after removal of PBST one drop of ProLong Gold AntiFade was applied to each well and chamber slides were allowed to dry out in the dark overnight. Slides were inverted and cells were visualized under immunofluorescence.

Table 2.1; Antibody table

Target	Company antibody bought from	Catalog number	Species	Dilution
Islet1	DSHB	40.2D6	Mouse IgG1	1:50
HoxB4	DSHB	I12	Mouse IgG2a	1:50
Olig2	Millipore	AB9610	Rabbit	1:200
FOXP1	Abcam	AB16645	Rabbit	1:200
SMI32	Cambridge Bioscience	SMI-32R-500	Mouse IgG1	1:1000
CHAT	EMD Millipore	AB144P	Goat	1:100
BIII tubulin	Sigma	T8578	Mouse	1:500
3CB2 (VIM)	DSHB	3CB2	Mouse IgM	1:5
NFIA	Abcam	ab41851	Rabbit	1:1000
GFAP	Dako	M 0761	Mouse IgG1	1:200
GLAST	Antibodies-online	ABIN350309	Rabbit	1:500
TDP43	Proteintech	10782-2-AP	Rabbit	1:500
ATP5B	Abcam	ab14730	Mouse	1:500
PDI	New England Biolabs	2446	Rabbit	1:50
Syt1	Synaptic	clone 41.1	Mouse	1:250

	Systems			
MAP2	Abcam	ab5392	Chicken	1:500

Microscopy

Fluorescence

All ICC was analyzed using fluorescence microscopy unless stated otherwise. Coverslips were imaged using the Hamamatsu camera controller C10600 and Leica DM5500B microscope. For each coverslip a minimum of 3 images were acquired at a magnification of 64x. Images were counted manually and statistics performed using Prism Graphpad version 6.

Confocal analysis of TDP43

Coverslips were imaged using a Zeiss 710 Confocal, at 63x magnification, 1.4 N.A. oil objective. A minimum of 3 Z-stack images were acquired per coverslip. Settings for acquisition and thresholding were kept standard for each experimental set. Image analysis was performed in Fiji software version 2.0.0-rc-41/1.50d. Area of TDP43 staining was calculated for both nucleus (outlined by DAPI) and cytoplasmic compartments using analyze particle function, and the relative percentages of total were calculated.

Analysis of ER-mitochondria colocalization.

ER-mitochondria colocalisation was analysed by Zhi Yao, under the supervision of Sonia Ghandi. Cultures (n=3) were immunolabelled for ATP5B and PDI. Cells were imaged using a Zeiss 710 confocal system with a 63x, 1.4 N.A. oil objective. Z series of images were acquired using a pinhole diameter of 1 Airy units (AU). Settings for acquisition and thresholding were kept standard for each experimental set. Colocalization of red and green signals was quantified with Zen software.

Analysis of synaptotagmin-1 staining

Assessment of SYT1 positive puncta was performed by Dr. Andras Lakatos following published methods (Tyzack et al. 2014). Briefly, I immunolabelled cultures (3-4 technical repeats) for pre-synaptic (synaptotagmin-1), MN (Chat) and dendritic (MAP-2) markers. Following, Andras used confocal laser microscopy in 5mm depth to scan cultures and images were collected by 1mm apart. Scanning parameters were defined for the control groups and the same

settings were used for the mutant hiPSCs. SYT1 positive puncta were analysed using ImageJ (v1.0, NIH, plugin written by B. Wark and was provided to Dr A. Lakatos by Dr C. Eroglu). Threshold values were kept the same across the optical sections. The number of SYT-1 positive puncta colocalising with either the MAP-2/Chat immunoreactive soma or main dendrites was counted (N=26-32 cell). SYT1 densities were defined by the number of puncta per cell area/dendrite length for each line.

Live Cell imaging assays

All live cell imaging experiments were done in collaboration with Minee Choi, Zhi Yao and Sonia Gandhi, except the axonal transport which was carried out by Alexander Fellows under the supervision of Gipi Schiavo.

Cell Survival Assays

Cells were plated on Geltrex substrate and differentiated in 96 well plate. To determine the percentage of cell death, cells were loaded with 5ug/ml Hoechst 33342 (cell permeable nuclear dye) and 1ug/ml membrane impermeant propidium iodide (PI) for 15 mins at room temperature in normal culture media. Image acquisition was performed using the ImageXpress system (Molecular Divices). Hoechst-stained nuclei were imaged using the 377nm excitation/530 emission filter set and images were used to determine objects for total cell counting. PI staining were imaged using the 531nm excitation/590nm emission filter set. The number of PI positive cells were determined using the multiwavelength cell scoring module of the MetaXpress software. The threshold for counting PI+ cells was kept standard for each experimental set.

Measurement of mitochondrial membrane potential

Mitochondrial membrane potential was measured by loading the cells with 25nM tetramethylrhodamine methyl ester (TMRM, Life Technologies) for 40 min in physiological buffer (138 mM NaCl, 5.6 mM KCl, 4.2 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, 1.2 mM $MgCl_2$, 2.6 mM $CaCl_2$, 10 mM D-glucose and 10 mM HEPES). TMRM fluorescence signal was acquired using a Zeiss 710 confocal system with a 63x, 1.4 N.A. oil objective. Z series of images were acquired using a pinhole diameter of 1 Airy units (AU). TMRM fluorescence intensity was quantified by measuring the average TMRM fluorescence above threshold area on the maximal projection of the Z series using Metamorph software. Settings for acquisition and thresholding were kept standard for each experimental set.

Imaging of intracellular ROS production

ROS generation was measured with dihydroethidine (DHE, 10-40μM, Invitrogen). All imaging was performed in physiological buffer (138 mM NaCl, 5.6 mM KCl, 4.2 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, 1.2 mM $MgCl_2$, 2.6 mM $CaCl_2$, 10 mM D-glucose and 10 mM HEPES). To avoid accumulation of oxidized products, DHE was added immediately before measuring and was present in solution throughout the experiments. Imaging was performed either with a CCD camera on an epifluorescence inverted microscope or with the FLIPR system (Molecular devices).

Determination of GSH level

Reduced glutathione (GSH) level was measured by incubation live cells with 40μM monochlorobimane (mCB) for 1 hour in physiological buffer (138 mM NaCl, 5.6 mM KCl, 4.2 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, 1.2 mM $MgCl_2$, 2.6 mM $CaCl_2$, 10 mM D-glucose and 10 mM HEPES). The fluorescence signal of GS-mCB adduct was imaged using a Zeiss 710 confocal system with a 63x, 1.4 N.A. oil objective and excitation 405nm / emission at 490 nm. GS-mCB fluorescence intensity was quantified by measuring the average fluorescence intensity above threshold area on the maximal projection of the Z series using Metamorph software. Settings for acquisition and thresholding were kept standard for each experimental set.

Axonal transport

Motor neurons derived from human iPSCs were incubated with 30 nM HcT labeled with AlexaFluor 555 for 30 min at either day 6 or 19 days post terminal differentiation. Cultures were washed and fresh media was applied. Cells were then imaged using a Zeiss LSM 780 microscope equipped with a Zeiss X63, 1.40 NA DIC Plan-Apochromat oil-immersion objective whilst kept at 37 °C. Retrograde transport was assessed using Motion Analysis software (Kinetic Imaging).

Mouse primary motor neurons were cultured and kept until days 6-7 in vitro when method above was applied.

Statistics using prism

All experiments have a minimum of technical n=3 and biological n=2, unless otherwise stated. GraphPad Prism 6 was used to perform all statistical analysis. Error bars represent mean ± SEM. Two-way analysis of variance was used to analyse independent variables in different groups. For SYT1 staining

analysis, Kruskal-Wallis (Dunn) post hoc tests was applied to analyse data pairs. Statistical significance was accepted at P-values of <0.05. *, **, *** indicate significance of P<0.05, P<0.01 and P<0.001 respectively.

Biochemistry

Western blot analysis

Protein levels of a number of markers were quantified by Dr. Zhi Yao. Briefly, protein samples were extracted using 2x laemmli buffer (4% (w/v) SDS, 20% Glycerol, 120 mM Tris-Cl (pH 6.8)). Total protein concentration was quantified using BCA assay (Pierce). Equal amount of protein samples were then loaded and separated by SDS PAGE and transferred onto a nitrocellulose membrane. Samples were then incubated with primary antibodies overnight at 4°C followed by horseradish peroxidase conjugated secondary antibodies (Dako). The blots were imaged and the density was analysed using Bio-red imaging software.

Functional Genomics/preparation of libraries for high throughput sequencing

RNA isolation

1. Column-based RNA isolation

Cell pellets were resuspended in 1ml qiazol (for up to 1x10^6cells) and left at room temp for 5minutes. (If extracting from brain tissue use Qiagen TissueRupter (9001273) to homogenize). An equal volume of 100% EtOH was added and mixed well by vortexing. The sample was applied directly to a Direct-zol RNA column (R2052, Cambridge Bioscience), placed in 2ml collection tube and spun through for 1min at 10,000g at 4°C. (A maximum of 700μl can be applied to the column at one time so if necessary spin through part of sample, remove waste and repeat with remaining sample.) Waste and collection tube were discarded and the sample was in column DNAse treated. The column was then washed twice with 400μl RNA pre-wash buffer, once with 700μl RNA wash buffer and spun through for 1min at 10,000g and 4°C to dry column. The column was placed inside a new 1.5ml collection tube and the RNA sample was eluted in 50μl of nuclease free H20.

2. An automated RNA isolation approach

RNA was extracted using an automated approach, using the promega Maxwell RSC instrument and Maxwell® RSC simplyRNA Cells Kit (AS1390, Promega). Before the automated process, cell pellets were homogenized with a chilled 1-Thioglycerol/homogenisation solution, triturated and vortexed to break up the pellet then immediately before loading cells were lysed and vortexed again. A DNase step is included in the automated process. Up to 16 samples were processed at once.

Post RNA extraction QC

Samples are measured with nanodrop to check RNA concentration, pure RNA samples will have a 260/280 ratio of 2.0-2.1 and on the agilent bioanalyser to assess quality, RNA integrity (RIN) scores are >8 for downstream applications (few exceptions were processed with RIN scores >7.5).

RNA sequencing (including steps for ribosomal RNA depletion)

Approach 1; NEXTflex Directional RNAseq

Firstly ribosomal RNA (rRNA) was depleted using the Qiagen generead rRNA depletion kit. After rRNA depletion (or PolyA selection of mRNA), 100ng of each sample was taken forward for library preparation using the NEXTflex Directional RNA-Seq Kit (dUTP-Based) v2. Briefly, samples were fragmented (enzymatically) to obtain optimal size for sequencing, first strand synthesis was followed by second strand synthesis with dUTP to retain directionality, ends of fragments were adenylated before the ligation of adapters (which include barcodes). Finally samples underwent UDG digestion and 15cycles PCR reaction. Quality of each library was checked individually using a DNA1000 bioanalyser prior to multiplexing and sequencing on a Hiseq.

Approach 2; Truseq stranded mRNA library preparation

Firstly, using the Truseq stranded mRNA RNAseq kit (RS-301-2001/2, Illumina), the poly-A containing mRNA molecules were purified from 0.1-4ug total RNA input using poly-T oligo attached to magnetic beads. Following purification, the mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Strand specificity was achieved by replacing dTTP with dUTP in the Second Strand Marking Mix (SMM), followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. The incorporation of dUTP in second strand

synthesis quenched the second strand during amplification, because the polymerase used is not incorporated past this nucleotide. The addition of Actinomycin D to First Stand Synthesis Act D mix (FSA) prevented spurious DNA-dependent synthesis, while allowing RNA-dependent synthesis, improving strand specificity. These cDNA fragments next underwent 'end repair' (adenylation) where a single 'A' base was added to each strand, followed by ligation of an adapter (containing 6nucleotide barcode – allows multiplexing for sequencing). The products were then purified and enriched with PCR amplification to create the final cDNA library.

Approach 3; Truseq Access RNAseq

The Truseq Access kit (RS-122-2101/2, Illumina) uses total RNA sequencing to create a template library, from which the coding regions are captured and sequenced. So firstly the total RNA (100ng input) was fragmented into small pieces, cDNA was generated using random priming during first and second strand synthesis and sequencing adapters were ligated to the resulting double-stranded cDNA fragments. The coding regions of the transcriptome were then captured from this library using sequence-specific probes, then a final round of PCR amplification and second strand digestion occurred to create the final library.

iCLIP

Reagents:

RT primers used:

Rt22clip	NNAGAGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt23clip	NNAGTCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt24clip	NNATACNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt25clip	NNATCANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt26clip	NNCAAGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt27clip	NNCAGANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt28clip	NNCCACNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt29clip	NNCCCTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt35clip	NNGAGTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt36clip	NNGATGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt37clip	NNGCAGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt38clip	NNGCTCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt39clip	NNGGACNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt40clip	NNGGCGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt41clip	NNGTATNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt42clip	NNGTTANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt43clip	NNTAACNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt44clip	NNTACANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt46clip	NNTCTANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Pre-adenylated 3' adaptor DNA (ordered from IDT):
L3-App: rAppAGATCGGAAGAGCGGTTCAG/ddC/

Cut_oligo:
GTTCAGGATCCACGACGCTCTTCaaaa

PCR primers:
P3 Solexa:
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCG
CTCTTCCGATCT

P5 Solexa:
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCT
TCCGATCT

Buffers:

Lysis Buffer	50 mM Tris-HCl, pH 7.4 100 mM NaCl 1% Igepal CA-630 (Sigma I8896) 0.1% SDS 0.5% sodium deoxycholate
High salt wash	50 mM Tris-HCl, pH 7.4 1M NaCl

	1M EDTA 1% Igepal CA-630 (Sigma I8896) 0.1% SDS 0.5% sodium deoxycholate
PNK Buffer	50 mM Tris-HCl, pH 7.4 100 mM $NaCl_2$ 20 mM Tris-HCl, pH 7.4 10 mM $MgCl_2$ 0.2% Tween-20
5x PNK pH 6.5 Buffer: (freeze aliquots of the buffer, do not thaw and freeze again)	350 mM Tris-HCl, pH 6.5 50 mM $MgCl_2$ 5 mM dithiothreitol
4x Ligation Buffer (freeze aliquots of the buffer, do not thaw and freeze again)	200 mM Tris-HCl, pH 7.8 40 mM $MgCl_2$ 4 mM dithiothreitol
PK Buffer	100 mM Tris-HCl, pH 7.4 50 mM NaCl 10 mM EDTA
PK Buffer + 7 M Urea	100 mM Tris-HCl, pH 7.4 50 mM NaCl 10 mM EDTA 7 M urea

Method (as displayed in Figure 2.1):

Cell samples were UV crosslinked with 0.15mJ/cm^2 254nm UV light, to preserve RNA-RBP interactions by covalently binding protein-RNA in vivo. Cells were lysed in lysis buffer plus 1/100 protease inhibitors and 1/1000 anti-RNase, then sonicated (30second bursts followed by 30secounds rest, x

10cycles) and partially fragmented by incubation with RNase I at 37°C for 3minutes, to produce optimal size fragments for sequencing (RNase dilution/conditions adjusted per experiment, I used 1/1000 dilution low RNase and 1/10 dilution for high RNase). Protein-RNA complexes were immunoprecipitated under stringent conditions using antibodies specific to target protein. Next the RNA in the protein-RNA complexes is dephosphorylyated, an adapter is ligated to the 3' end of each RNA molecule and the 5'end is labeled with p32. Gel electrophoresis and nitrocellulose transfer, under denaturing conditions, were carried out to remove RNA not bound by the protein of interest and to visualize the protein-RNA complexes. The p32 signal guided the excision of protein-RNA complexes off the membrane. Proteins were digested away using proteinase K (This usually leaves a single peptide/amino acid at the crosslink site) and remaining RNA underwent reverse transcription, which truncated at the crosslink site preserving the position of binding. The oligonucleotides for reverse transcription contain two inversely oriented adaptor regions (P3/5) separated by a BamHI restriction site as well as a barcode region at their 5′ end containing a 2-nt barcode to mark the experiment and a 3-nt random barcode to mark individual cDNA molecules enabling multiplexing and removal of PCR duplicates. Gel purification was used to remove primers and size-select cDNA samples. Single stranded cDNA is then circularized, and an oligo complementary to the BamHI site was annealed to allow BamHI to cut at the restriction site, linearizing the cDNA. Finally libraries were PCR amplified and multiplexed (usually mixing the medium and high fragments of each library in a ratio of 1:2 respectively).

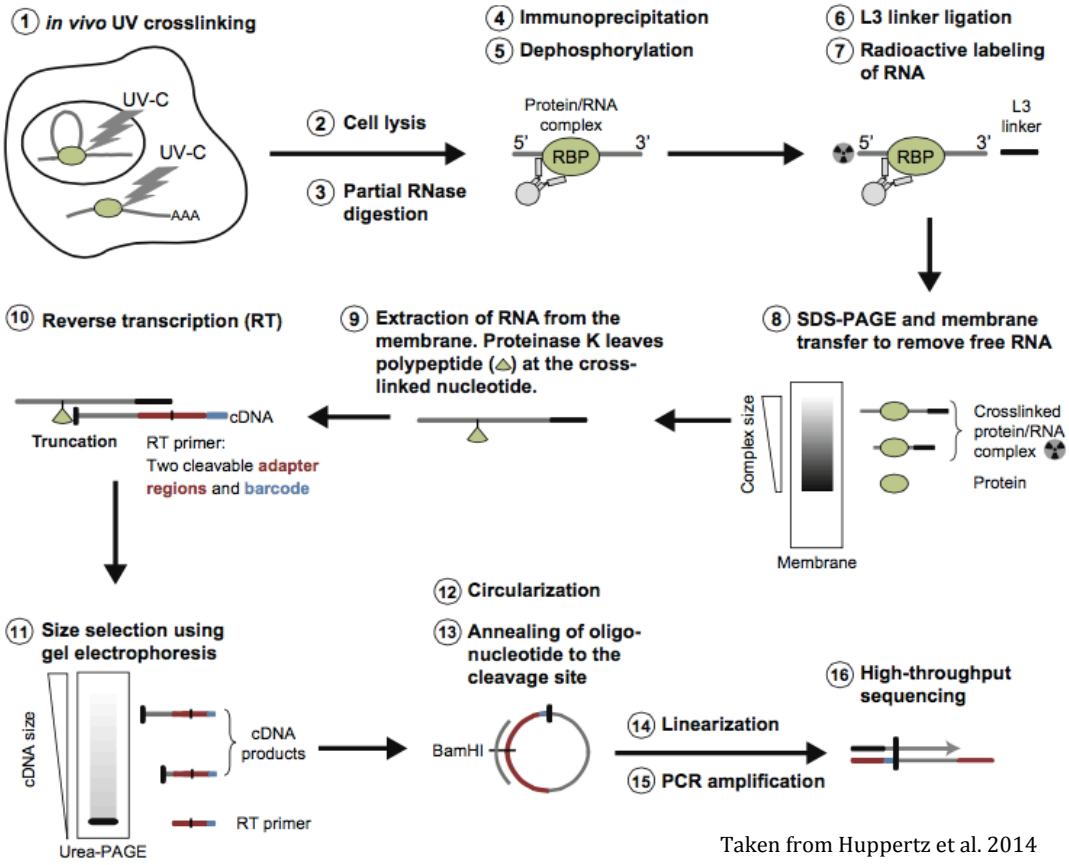


Figure 2.1; Schematic representation of the iCLIP protocol identifying RNA–protein interactions in vivo.

Library quantification and sequencing

It is important to accurately quantify libraries before sequencing to ensure optimal sequencing capacity. Overestimation of library concentration results in lower cluster density after bridge PCR, however underestimation of library concentration results in too many clusters on the flow cell, which can lead to poor cluster resolution. I have used several methods to check quality of libraries before sequencing.

Bioanalyser

Agilent bioanalyser 2100 was run with the agilent RNA nano 6000 kit (5067-1511) to check total RNA quality before samples are used for downstream experiments. The standard protocol from agilent was followed. Two peaks representing ribosomal RNA usually dominate the trace and these are used to calculate RIN score. The agilent High sensitivity DNA 1000 kit (5067-4627) was then used check library quality, which is usually normally distributed with an average fragment size of 275-300nt (again the standard protocol from agilent was followed).

Qubit

The Qubit dsDNA HS (High Sensitivity) Assay Kit (Q32854, Thermo Scientific) was used with the Qubit® Fluorometer to measure library concentration. The assay is highly selective for double-stranded DNA (dsDNA) over RNA and is designed to be accurate for initial sample concentrations from 10 pg/μL to 100 ng/μL.

qPCR Kapa

qPCR is widely regarded as the gold standard for accurate quantification of DNA libraries as it is the only technique capable of measuring the number of amplifiable molecules. KAPA SYBR FAST qPCR kit (KK4824, Insight Biotech) was used for final library quantification before sequencing.

High-throughput Sequencing

Libraries were sent for high throughput sequencing at the Institute of Neurology's NGS core facility using the Hiseq2500. They were typically sequenced on a single-end read rapid flow cell for 50 cycles plus separate indexing cycles to enable de-multiplexing (an additional cycle was ran after the final cycle and the final index cycle to ensure accurate basecalling for the final cycles).

Computational methods

Mapping and annotation of RNAseq data

For the processing and analysis of RNAseq samples I have set up a bioinformatics pipeline, with the guidance of Dr. Chris Sibley (the same workflow is used to analyze all three RNAseq approaches) (Figure 2.2).

Primary analysis was carried out during sequencing on the Illumina Hiseq2500; RTA realtime analysis software converts TIFF images to intensity scores and then to basecalls. Alan Pittman next generated FastQ files from the basecalls, using casava version 1.8.2. FastQ files are de-multiplexed raw data files that contain information on the bases and their quality. Samples were taken forward if 90% of the data had a quality score >Q30 and had passed filter (an automated Illumina QC step which analyses each cluster quality individually and removes poor quality data).

Next I removed the 3' Solexa adapter sequence and sequences shorter than 24 nucleotides using fastx_clipper (part of the Fastx toolkit, version 0.0.13),

generating ‘clipped’ data. Afterwards the quality of the sequencing data was checked using FastQC, version 0.11.2. (Fastq raw data files were used as input). FastQC displayed multiple summary graphs to easily visualize the data quality; information includes sequence quality per base & per tile & across sequences, sequence content per base, GC % per sequence, distribution of sequence lengths across all sequences, sequence duplication levels, adapter content and kmer content.

Secondary analysis included aligning, annotating, and counting the data. I aligned the clipped data using tophat2 a splice aware aligner; this extracted transcript sequences and used bowtie indexes to align reads to a reference genome (bowtie2-2.2.3 was used to compile hg19 or mm10). I allowed 1 mismatch and have used annotation taken from ensembl (version 60). At this stage the % of each library that maps to the genome and number of duplicated reads was assessed (all samples I have used map >80-90% to hg19/mm10). Samtools flagstat was used to detect any duplicated reads. Additionally the % of ribosomal reads, strandedness, % mapping to exons/introns/intergenic regions in each library was evaluated before continuing with any downstream analysis. I used Seqmonk to get an overview of quality across all experimental samples but a more accurate calculation was performed by Dr. Raphealle Luisier in Prof. Nicholas Luscombe’s lab. Raphaelle used Bowtie to map ribosomal reads and then tophat2 to map those reads that did not correspond to rRNA, strandedness was calculated by counting the reads which align on (+) versus (-) strand of each transcript and then counting the fraction which map to the colinear strand (antisense strand given RNAseq approached used) and the percentage of reads mapping to intronic, exonic and intergenic regions was counted using HTSeq count. A threshold of >70% mapping to exonic regions, <1% rRNA and >90% strandedness is used throughout this thesis.

Following this, I counted the aligned reads and genomic features (eg. genes) using htseq-count (version HTseq-0.6.1 & python version 2.7.9). I used *S* resolution mode, which allows one count per feature, but if the read contains more than one feature it is counted as ambiguous (and not counted for any features) and if the read doesn’t overlap any feature, it is counted as no_feature. The output BAM/SAM files were also used to visualize the data in the IGV browser (indexing BAI files were generated using samtools).

Finally the data is normalized and interpreted using a variety of downstream analysis tools. Initially I normalized the data by calculating rpkm values (reads per kilobase per million mapped reads) using the edgeR package in R, which takes into account the gene length. These RPKM values were firstly used to carry out a final quality check of the data; I confirmed the expression of cell-specific markers in each sample before continuing with any further analysis, for example CHAT, Islet1, BetaIII and Olig2 expression in the MN samples. RPKM values were also used to analyze the expression of selected genes of interest across samples after they have undergone log transformation to account for differences in library size.

Differential expression analysis

After carrying out essential sample and experimental quality checks, principal component analysis and hierarchical clustering was performed to assess sample clustering using DEseq2 package on R (version 3.2.3). Next Differential expression was carried out using DESeq2 (version 1.10.1). DEseq assumes a negative binomial distribution model and uses a scaling factor normalisation to account for different sequencing depth between samples and Benjamini-Hochberg method to control the FDR. The thresholds used to detect differentially expressed genes in this thesis were ≥1.5 log (fold change), FDR < 0.1 and $p < 0.05$, unless otherwise stated.

Further gene ontology (GO-term) analysis was performed using the DAVID software online (https://david.ncifcrf.gov/). All genes with base mean >10 were used as the background.

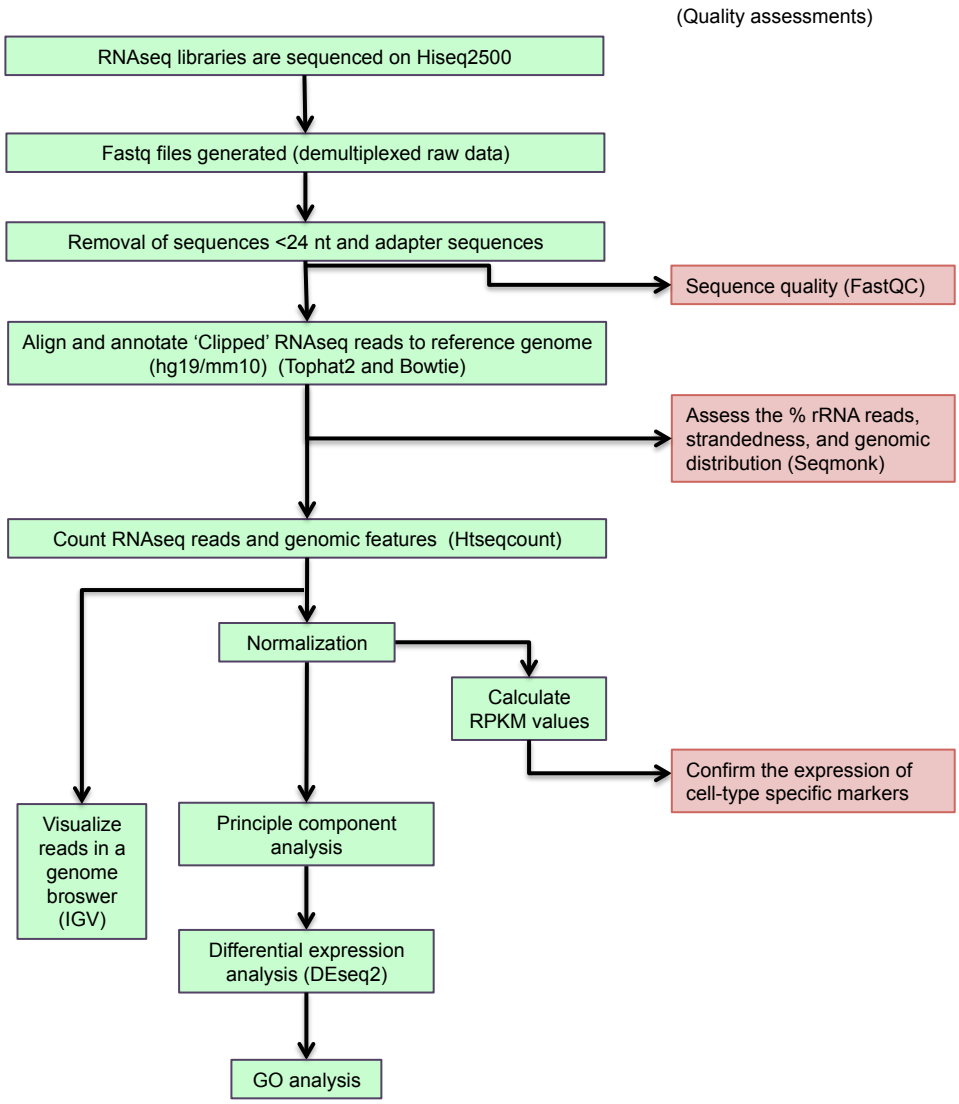


Figure 2.2; RNAseq analysis pipeline

iCLIP

iCLIP libraries have been analysed on the icount server (http://icount.biolab.si), which is set up and maintained by Tomaz Curk (König et al. 2011; Tollervey et al. 2011; Ule et al. 2010). Firstly sequencing data passes several QC/filtering stages and then was mapped (to hg19) and annotated (using ensemble version 60). Crosslink sites were identified and the level of binding at each site quantified. The % binding within different genomic regions eg. Intronic, ncRNA was counted (downloaded from sumtype files). Mapped data was downloaded from cross-link maps tab for individual samples and grouped tab for pooled data (sum of multiple repeats). For this thesis G files were used for all analysis with the exception of repeat element analysis (chapter 4) where B files were used. G files contain single hits in the genome (random barcodes were taken into account). B files include both single and multiple hits in the genome (up to

20 multiple hits are allowed) and the results are weighted (the weight of the sequence is divided by the number of its multiple hits).

Crosslink clusters were identified using peaks analysis, parameters were set to create clusters that are 15nt apart and sites with FDR <0.05 were considered significant.

To calculate the base frequencies of iCLIP sequence reads, the genomic sequence corresponding to the first 10 nucleotides of all reads plus 11 nucleotides of preceding sequence was extracted. 1000 sequences were randomly selected per sample and were presented graphically using Weblogo (version 2.8.2). Further sequence analysis of iCLIP crosslink sites was performed using kmer analysis. Standard kmer analysis was performed by extending the position of the crosslinking site 10 nt in both directions then enriched pentamer z-scores were calculated by taking the true score (total number of times the kmer is counted in defined 20nt region surrounding x-link site) divided by the mean random score (number of times the kmer is counted in the same sequences after random shuffling).

% of binding to repeat elements and motif enrichment analysis was processed by Nejc Haberman. Motif enrichment was performed using dreme (version 4.11.2) using tetramers and pentamers and the two sets of site to compare were used as the background for each other. Finally Crosslink sites between grouped were compared using compare analysis on icount.

Graphics in R

Graphics (heatmaps/correlation plots/bar plots/box plots) were generated in R version 3.2.3, using Bioconductor version 3.1 (BiocInstaller 1.18.5). Ggplot2 (2.1.0) package was used for bar, box and scatterplots.

Materials list

Reagent	Catalogue number	Supplier
4-12%BisTris gel 1.0mmx12wells	NP0322BOX	Life Technologies

4-12%BisTris gel 1.0mmx15wells	NP0316BOX	Life Technologies
4-thiouridine	T4509-100MG	Sigma
4′,6-Diamidino-2-phenylindole dihydrochloride (DAPI)	D9542	Sigma
6%TBE gel 1mmx12wells	EC62652BOX	Life Technologies
6%TBE gel 1mmx15wells	EC62655BOX	Life Technologies
6%TBE UREA gel 1mmx10wells	EC6865BOX	Life Technologies
Accutase	A11105-01	Invitrogen
Agilent RNA 6000 Nano Kit	5067-1511	Agilent
Agilent High Sensitivity DNA Reagents	5067-4627	Agilent
Alexa Fluor® 568 Goat Anti-Mouse IgG1 (γ1)	A-21124	Thermo fisher scientific
Anti-3CB2 antibody (VIM)	3CB2	DSHB
Anti-ATP5B Antibody	ab14730	Abcam
Anti-Choline Acetyltransferase Antibody (CHAT)	AB144P	EMD Millipore
Anti-FOXP1 antibody	AB16645	Abcam
Anti-Human SLC1A3 rabbit antibody (GLAST)	ABIN350309	Antibodies-online
Anti-MAP2 antibody	ab5392	Abcam

Anti-neurofilament H (SMI32)	SMI-32R-500	Cambridge Bioscience
Anti-NFIA antibody	ab41851	Abcam
Anti-Olig2 antibody	AB9610	Millipore
Atto633 adapter	Non-catalog synthesis request (/5rApp/AGATCGGAAGAGCGGTTCAG/3Mod/)	Intergrated DNA Technologies
B27 supplement	17504-044	Thermo-scientific
Compound E	ALX-270-415-M001	Enzo Life sciences
Costar-X filter spin columns	CLS8161-100EA	Sigma
Coverslips (round 13/22mm thick)	631-0158	VWR
Cryovials 2ml (2)	430488	Sigma-Aldrich
Dimethyl sulfoxide	D2650	Sigma
Direct-zol™ RNA MiniPrep, 200 preps, w/ Zymo-Spin™ IIC Columns	R2052	Cambridge Bioscience
Dispase	17105-041	Invitrogen
Dithiothreitol	D9779-10G	Sigma
DMEM-GLUTAMAX	31966-047	GIBCO (Life Technologies)
DMEM/F12	10565018	Thermo-scientific
dNTPs	R0191	Life Technologies

Donkey anti-Goat IgG (H+L) Secondary Antibody, Alexa Fluor® 568 conjugate	A-11057	Thermo fisher scientific
Dorspomorphin	3093/10	R and D systems
Dulbecco's Phosphate-Buffered Saline (DPBS)	14190-250	Life Technologies
Dynabeads® Protein A	10002D	Life Technologies
Dynabeads® Protein G	10004D	Invitrogen
Essential 8 Medium	A1517001	Thermo-scientific
Ethylenediaminetetraaceti c acid (EDTA)	15575-020	Life Technologies
Fluorescence Mounting Medium	S302380-2	Dako
FOXP1	AB16645	Abcam
Geltrex	A1413302	Invitrogen
GeneRead rRNA Depletion Kit (6)	180211	Qiagen
GlycoBlue™ Coprecipitant (15 mg/mL)	AM9516	Life Technology
Goat anti-Chicken IgY (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate	A-11039	Thermo fisher scientific
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor®	A-11004	Thermo fisher scientific

568 conjugate		
Goat anti-Mouse IgG2b Secondary Antibody, Alexa Fluor® 488 conjugate	A-21141	Thermo fisher scientific
Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate	A-11008	Thermo fisher scientific
HoxB4 antibody	I12	DSHB
IBIDI chamber slides	IB-80826	Thistle Scientific
Igepal (NP-40)	I8896	Sigma
Islet1 antibody	40.2D6	DSHB
L-Glutamine	25030-024	Invitrogen
Laminin	L2020	Sigma
Leukemia Inhibitory Factor human	L5283	Sigma
Library Quantification Kit - Illumina/Universal	KK4824	Insight Biotechnology
LookOut Mycoplasma PCR Detection Kit	MP0035-1KT	Sigma
Low Molecular weight DNA Ladder	N3233L	New England Biolabs
Magnesium chloride	63020-1L	Sigma
Maxwell(R) RSC simplyRNA Cells	AS1390	Promega

Maxwell(R) RSC System	AS4500	Promega
Monoclonal Anti-β-Tubulin III	T8578	Sigma
Monoclonal Mouse Anti-Human Glial Fibrillary Acidic Protein (GFAP)	M 0761	Dako
MycoAlert mycoplasma detection kit	LT07-318	Lonza
N2 supplement	17502-048	Thermo-scientific
Neurobasal	12348-017	Thermo-scientific
Non-essential amino acids	11140-050	Life Technologies
PageRuler Prestained NIR Protein Ladder	26635	Thermo-scientific
Pageruler Prestained Protein Ladder	26616	Thermo Scientific
Paraformaldehype (PFA)	2199983	MP Biomedicals
PDI antibody	2446	New England Biolabs
Penicillin Streptomycin	15070063	Invitrogen
phospho-TDP43 anitbody, Ser409/410-1 (pTDP43)	CAC-TIP-PTD-P01	Cosmo Bio Ltd
Pipetting reservoir sterile polystyrene 50mL	11543412	Fisher
ProLong® Gold Antifade Reagent	P36934	Life Technologies
Protease inhibitors	539134-1SET	Calbiochem/Merck
PROTRAN BA85 30cm x	10401196	GE Healthcare Life

3m 1/pk BA membrane		Sciences
Purmorphamine	540220-5MG	Calbiochem (EMD Millipore)
Qubit dsDNA HS Assay Kit	Q32854	Thermo Scientific
Recombinant Human BMP-4	314-BP-050	R & D systems
Recombinant Human FGF-basic	100-18B	Peprotech
Retinoic acid	R2625	Sigma
RNase I	AM2295	Invitrogen
RNase inhibitor	AM2692	Invitrogen
Salubrinal (5mg)	324895	Merck Chemicals ltd
SB431452	1614/10	R and D systems
Sodium arsenite solution	35000-1L-R	Sigma Aldrich
Sodium chloride	S5150-1L	Sigma
Sodium-deoxycholate	D6750-100G	Sigma
Sodiumdodecylsulfate	BP166-500	Fisher Scientific
StemMACS CHIR99021	130-104-172	Macs Miltenyl Biotec
Sucrose	S0389-500G	Sigma
SuperScript® II Reverse Transcriptase	18064014	Thermo-scientific
SuperScript® III Reverse Transcriptase	18080-085	Thermo-scientific
Synaptotagmin 1 antibody	clone 41.1	Synaptic Systems

(Syt1)		
T4 DNA Ligase	M0202S	New England Biolabs
T4 Polynucleotide Kinase	M0201L	New England Biolabs
TBE 5X	J885-1L	Amresco (VWR)
TDP43 antibody	10782-2-AP	Proteintech
Tris HCL	T2194-1L	Sigma
Triton X-100	T8532	Sigma
Trizol LS	10296028	Life Technologies
TruSeq RNA Access Library Prep Kit, Set A/B	RS-301-2001/2	Illumina
TruSeq Stranded mRNA Sample Prep Kit - TruSeq® Stranded mRNA LT - Set A/B	RS-122-2101/2	Illumina
Turbo DNase Kit	AM1907	Life Technology
Tween-20	P1379-1L	Sigma
Whatman™ 1823-010 Grade GF/D Glass Fiber Filter Paper without Binder, Diameter: 1cm, Pore Size: 2.7µm (Pack of 100)	1823-010	Sigma

Chapter 3; Developing the model system

Introduction

There are many different model systems used to study MND. However to date, the translation of MND research has not been fruitful, therefore it would benefit from models that more accurately capture clinical pathophysiology. I have chosen to utilise hiPSC technology, which holds several major advantages to modelling human disease; it's a fully human model system, mutations are expressed at representative pathophysiological doses and, as this is essentially a developmental system, it allows us to recapitulate the earliest pathogenic events occurring during disease manifestation.

To effectively utilize hiPSCs to model MND I had to first develop and refine our culture and differentiation strategies, to make sure they are robust, efficient and validated. As previously discussed the cells most vulnerable in MND are motor neurons (MN) but there is a strong rational to also study the role of glia in the disease pathogenesis. In this chapter I describe the differentiation protocols I have developed for both spinal cord MN and astrocytes (AC).

Prior to the discovery of hiPSCs in 2006, there has been much work carried out on the development of the nervous system, and more specifically on motor neurogenesis, in both mouse and human ESCs (X.-J. Li et al. 2005; Wichterle et al. 2002). The studies, particularly in rodent systems, have provided key insights into some of the processes underlying differentiation and fate restriction. Although there are fundamental differences between species and questions to whether artificially generated stem cells are equivalent to embryonic stem cells, the knowledge gained from these past studies has been since exploited in ES and iPSC.

An overview of motoneurogenesis strategies to date (prior to 2013)

Generally protocols for the differentiation of neuronal cells from pluripotent stem cells are divided into 4 main stages;

1/ neural conversion,

2/ patterning,

3/ propagation

4/ and finally terminal differentiation.

During 2000's the generation of MN from both mouse (Wichterle et al. 2002) and human ESCs (Lee et al. 2007; Li et al. 2005) was described. All followed an embryoid body culture approach for neural induction; briefly this is where neural differentiation is initiated by the withdrawal of feeders and utilizes differential adhesion of NPCs (Zhang et al. 2001). This approach was shown to recapitulate early stages of neural development in vitro, such as the formation of neural tube–like structures (Zhang et al. 2001). Additionally, these early MN approaches all utilized RA and SHH agonist signals in a time-dependent manner to generate region specific motor neuron precursors. After terminal differentiation cells were functionally characterized by looking at synaptic function using electrophysiology techniques and myotube co-culture to assess neuromuscular transmission (Li et al. 2005).

From these earlier studies using RA and SHH signaling to promote motor neurogenesis, the yield of MN generated from mouse ESCs was greater than with human ESC (hESC). In 2008 the efficiency of motor neurogenesis was improved in hESC to generate a nearly homogenous population of ventral spinal progenitor cells and 50% final MN population (li et al. 2008). This was achieved by altering the timing of exposure to SHH signals, which was shown to promote division of olig2 progenitors and subsequently increase post-mitotic MNs. They also speculated that suspension culture was better for differentiation as got rid of 'flat cells' and produced a more homogenous population.

Since, there have been many more studies that have improved our understanding of, and protocols for, motor neurogenesis. Namely, in 2011 Patani et al demonstrated for the first time the generation of MNs in the absence of retinoic acid (Patani et al. 2011). RA signaling plays many roles in MN differentiation and is used reiteratively in development at distinct positions and in time. For example RA promotes the up-regulation of *HOX* gene expression and of Class I *HD* genes involved in D-V neural patterning. This diversifies MN subtypes from MN precursor pools.

Progress in motor neurogenesis approaches 2013-present

In the last few years significant develop has been made to overcome some of the major limitations in MN differentiation to date; including reducing the long culture time and improving poor efficiency and heterogeneity.

Firstly in 2013 Mackenzie Amoroso et al presented an approach to generate predominantly limb innervating MNs in 3weeks (Amoroso et al. 2013). They used a suspension culture system and a cocktail of rationalized signals; Rho-associated kinase inhibitor Y27632 to enhance single cell survival, FGF2 to enhance propagation, SB435142 and LDN193189 (= dual anaplastic lymphoma kinase (ALK) inhibition) for neutralization, RA and SHH for patterning, ascorbic acid to enhance neuronal yield, and growth factors (BDNF, IGF-1, GDNF, CTNF). This overall resulted in a 50% MN yield.

In 2014 two studies presented the use of Wnt signaling to increase efficiency of motor neurogenesis. Hong Chen et al describe a novel protocol achieving 90% motor neurogenesis in 21days, which uses CHIR99021 (a GSK3b antagonist) in addition to the commonly used motor-neuron/neutralization factors; smad inhibition, RA and Pur. Additionally they block cellular proliferation to enhance the generation of post-mitotic neurons (Chen et al. 2014). Then Yves Maury et al presented a large scale differentiation of both spinal cord and cortical MNs in 14days using an embryoid body system (Maury et al. 2014). They analyzed multiple combinations of developmental cues (target pathway, timing and concentration of cues) to reveal the importance of wnt signaling (by CHIR99021). Wnt signaling from an early stage in differentiation (day0) was shown to increased the yield of Olig2 positive precursors (>80%) and reduced the time to generate them.

Finally in 2015 Elizabeth Calder et al also demonstrated an approach to increase MN yield, by early exposure to RA. RA was shown to direct MN specification by suppression of GLI3 and worked independent of SHH signaling (Calder et al. 2015).

Astrogliogenesis differentiation strategies

Compared to Neurogenesis, astrogliogenesis has been relatively poorly studied but their development shares many similarities to that of neurons. Differentiation methods for glia are divided into 4 general phases; neural conversion, patterning, propagation and terminal differentiation. This is similar to neurons and many of the same principles apply.

To date, less attention has been paid to patterning of region-specific AC subtypes, comparatively to their neuronal counterparts. However this is essential given the functional, genetic and structural differences between region specific ACs. In relation to MN biology and disease, there is much interest in generation of spinal cord ACs so several groups have previously published methods to generate spinal cord ACs (Krencik et al. 2011; Roybon et al. 2013). In 2011, Krenick et al showed in hESCs that the regional identity of NPCs, specified by a single morphogen, is maintained through astrogliogenesis. Therefore, the same principles that are previously outlined for neural conversion and patterning can be followed for generating region-specific ACs.

Next NPCs undergo a propagation phase to promote generation of gliogenic precursors (GPCs). Often, to promote gliogenesis, NPCs are propagated in presence of growth factors, namely EGF and FGF, and is further assisted by frequent passaging to reduce cell-cell contact as this promotes neuronal fate (Caldwell et al. 2001; Krencik et al. 2011). ACs can terminally differentiate spontaneously from GPCs after the removal of mitogens, following temporally regulated intrinsic mechanisms (≈180days). But as this can be time consuming developmental insights have been used to target and accelerate astroglioenesis pathways by capitalizing on extrinsic signaling cues. Signals that induce expression of STAT, NF1A transcription factors or Smad signaling have been shown to promote expression of astrocytic genes GFAP, S100B and GLAST. Several in vitro approaches have targeted the JAK-STAT pathways and/or utilized BMP signaling. The JAK-STAT pathway has been activated with neuroregulin or using members of the interleukin IL6 family such as CNTF and LIF (Krencik et al. 2011), resulting in phosphorylation and nuclear localization of STAT3. Using a combination of LIF-mediated JAK-STAT signaling and BMP-mediated Smad signaling (by BMP4) was shown to be more effective in inducing the expression of astrocytic genes and down-regulating the expression of genes typically found in precursor cells, than targeting either pathway individually (Gupta et al. 2012). In addition, using insights from the gliogeneic switch, targeting epigenetic modifications using Aza-Cytidine a DNA methyltransferase inhibitor and Trichostatin-A a histone deacetylase inhibitor, has also been used to accelerate AC differentiation (Majumder et al. 2013).

To summarise the approaches used to generate spinal cord ACs; both utilized RA and SHH signaling to specify a spinal cord identify, expanded NPC populations in suspension culture and in media supplemented with EGF and FGF2 (Krencik et al. 2011; Roybon et al. 2013) and terminal differentiation was promoted by FBS, CNTF or LIF.

Characterization of Neural populations

A crucial step in developing methods to generate specific cell types in vitro is the characterization of cell-type-specific populations. This can be done by assessment of cellular morphology, expression of cell type-specific markers and functional characteristics. I discuss these methods used for cell-specific validation within the results but here I first outline some of the key genetic markers for both MNs and ACs.

Motor neuron markers

Extensive research spanning decades has carefully dissected MN differentiation pathways therefore a set of core cell-type-specific markers have been identified and are commonly used. Neural specification is first identified by the loss of pluripotency markers (Oct4, Nanog) and gain in neural markers such as Pax6, Sox1 and nestin. For MNs the standard markers that are consistently used are HB9 (an early MN marker), CHAT (a mature MN maker), SMI32 and Islet1. In addition there are more general neuronal markers often used in parallel, such as Beta III tubulin (neuronal marker), MAP2 (dendritic marker) and PDS95 (post-mitotic synaptic marker). Further, studies of subtype specification mean that we are able to narrowly pin point the positional identify of MNs through expression of the HOX gene 'postcode'. There are 39 Hox genes, arranged in 4 clusters; HOXA, HOXB, HOXC and HOXD and into 13 paralogue groups HOX1-13 (Philippidou and Dasen 2013). In the hindbrain, Hox genes from paralog groups 1–5 are expressed, while in the spinal cord expression of Hox4-Hox13 is detected. Hox genes 4-11 align with MN pool subtypes, 4 being the most cervical and 11 in the lumbar region.

Astrocytes

The knowledge of reliable markers for the characterisation of AC is much less defined than for motoneurogenesis. AC display huge genetic and morphological diversity making their identification challenging. GPCs are commonly identified by expression of NF1A and/or CD44 and for ACs GFAP expression has generally been considered the gold standard marker, however

there are limitations to its use. It's expression is seen to vary considerably depending on AC subtype diversity, morphology, activation state and during age(Sofroniew and Vinters 2010). This highlights the need to use a range of markers in parallel, such as APQ4, GLAST, ALDH1L1 and GLT1 (a mature/postnatal marker).

Current limitations in hiPSC strategies

Aside from the progress made in neural developmental understanding and refining our differentiation methods in vivo there are several key limitations to our current methods.

The majority of stem cell culture and differentiation strategies have been poorly controlled, extremely labour-intensive and costly. Former stem cell culture strategies routinely used MEFs (and other alternative) feeder layers to support stem cell cultures in their proliferative undifferentiated state. These feeder layers are highly variable, impractical for large-scale culture, allowed some differentiation to occur and would not compatible for clinical application in the future. Therefore there has been a considerable effort to move to feeder-free and xeno-free culture systems. In 2001 Xu et al demonstrated for the first time that hESCs could be cultured without a feeder layer on extracellular matrices (C. Xu et al. 2001), then a further study in 2005 explored what secreted factors from MEFs aided pluripotency (R.-H. Xu et al. 2005).

Further, there is also heterogeneity and inefficiency within the differentiation of stem cells, making methods not suitable for large-scale expansion. As outlined above, a lot of previous work has been carried out on MN differentiation. The above differentiation approaches all differ in terms of which developmental cues were employed for neural conversion and patterning (cues, their concentration and timing), the type of culture method used (suspension/monolayer), length of differentiation protocol and the purity of cells generated. A commonly employed methodology involved the formation of embryoid bodies (EBs). This type of suspension culture approach meant that extracellular signaling was poorly controlled leading to heterogeneous population of cells. A further hold back is due to the length of current differentiation protocols, for example the first protocol outlining the generation of ACs from hESC took 6months, which makes the technology extremely expensive and labour intensive. There are also less transparent reasons that

add to causing heterogeneity and variable maturity of cellular populations, for example cell density hugely influences cellular fate. Finally, another difficulty that I touched on above, is the characterization of differentiated iPSCs, particularly of ACs, due to the lack of reliable markers.

To summarise, the majority of hiPSC approaches to date have involved poorly defined culture conditions, protracted differentiation and low yield. Here I address these issues; firstly by using a monolayer culture approach which enables uniform signaling resulting in a more homogenous population and increased yield of desired cell type. Secondly I use chemically defined culture conditions that can be easily replicated, which improves reliability and accuracy of the methods. Finally I analyze the transcriptional profile throughout astrogliogenesis to better characterize stages of AC development and maturity.

Aims

In this chapter my overarching aim was to establish efficient differentiation strategies for both spinal cord MNs and ACs from hiPSCs.

Here I specifically ask:

1) Can I differentiate spinal cord MNs using solely monolayer culture?
2) Could the differentiation of MNs be accelerated?
3) Could I enhance the enrichment of MNs generated by synchronizing cell cycle exit?
4) Are the Derived MNs functional?
5) Then, can I differentiate spinal cord ACs using solely monolayer culture?
6) Are the derived ACs functional?
7) How do the transcriptional profiles of neurogenic and gliogenic precursors differ?
8) Can I define when the gliogenic switch occurs in our system?
9) What are the underlying molecular pathways that instruct NPC to turn off neurogenesis and turn on gliogenesis?
10) How does the transcriptional signature of our iPSC-derived AC compare to primary human ACs?

Differentiation of spinal cord motor neurons from hiPSCs using a monolayer culture system

Prior to differentiation I cultured hiPSCs on gel-trex coated plates and treated them daily with chemically defined medium that contains extrinsic signals to maintain pluripotency: these signals include FGF2 (activates MAPK and Akt pathways), insulin and TGF-B (smad2/3/4 signaling). Next, following key developmental principles and using insights from previous studies, I have produced directed differentiation strategies for the generation of spinal MNs using, for the first time, a purely monolayer culture system (Figure 3.1.A). This follows 3 main stages (Figure 3.1.A);

1/ Neural conversion

Similar to previous strategies I first directed iPSCs towards the neuro-ectodermal lineage. From the ectoderm germ layer, by default cells head towards a neural fate, however it has been shown that we can actively antagonise pathways that maintain pluripotency to accelerate neural conversion. As described by Chambers et al, simultaneously using two small molecule inhibitors of SMAD signaling, Noggin (BMP antagonist) and SB431542 (Activin/nodal anatagonist), neural conversion is achieved in 10-12days, and it is more efficient (>80%) and homogeneous compared to previous embryoid body or feeder methods (Chambers et al. 2009). Therefore for neural induction I have followed a similar approach, published by Shi et al(Shi, Kirwan, and Livesey 2012), and applied dual smad inhibition, using 10μM SB431542 and 1μM dorsomorphin for 10days, while cells are plated at 100% confluency. This approach promotes the specification of anterior neural precursors.

2/ Patterning (caudalisation and ventralisation of neural precursors)

Following neural conversion I applied developmentally rationalised extrinsic signals to positionally re-specify them to the pMN domain of the spinal cord, which gives rise to motor neuron precursor cells. Following past studies I utilised RA and SHH signalling (X. J. Li et al. 2008; X.-J. Li et al. 2005; Wichterle et al. 2002) and applied gentle trituration to begin to dissociate the neural epithelial sheet, but I did not passage cells into a single cell-suspension at this stage. I used purmorphine, a SHH antagonist which has previously been show to activate/suppress an almost identical set of transcription factors that

are involved in specification of ventral spinal progenitors and MNs, as well as stimulating Gli1 the downstream target of SHH pathway (X. J. Li et al. 2008). Additionally the same study showed that purmorphamine treated cultures expressed Olig2, a marker of motor neuron precursors, several days before expression was detected following SHH treatment. Hence, neural precursors are treated for 8days with 0.5μM retinoic acid to caudalise and 1μM purmorphamine to ventralises the precursor cells, thus mimicking the developmental signaling environment for motor neurogenesis.

3/ Terminal differentiation

Finally motor neuron precursors were left to terminally differentiate on laminin-coated plates for a minimum of 35days in neural maintenance media. During this time I passaged cultures twice to gently triturate clumps of NPCs and gradually break them up, this is important because if the cells are too confluent they will not efficiently differentiate but cell-cell contact is needed to some degree to promote neurogenesis therefore cells should not be plated to sparsely. As shown in the phase contrast images (Figure 3.1.B) derived neurons develop long axonal processes, typical of their distinct polarised morphology. However for precise characterisation of MNs and their subpopulations, assessing morphology is not very reliable and depending on cellular density it can be hard to visualise clearly the full cell.

Notably I did not use any growth factors throughout this method, making the approach cost effective and distinguishable from other strategies.

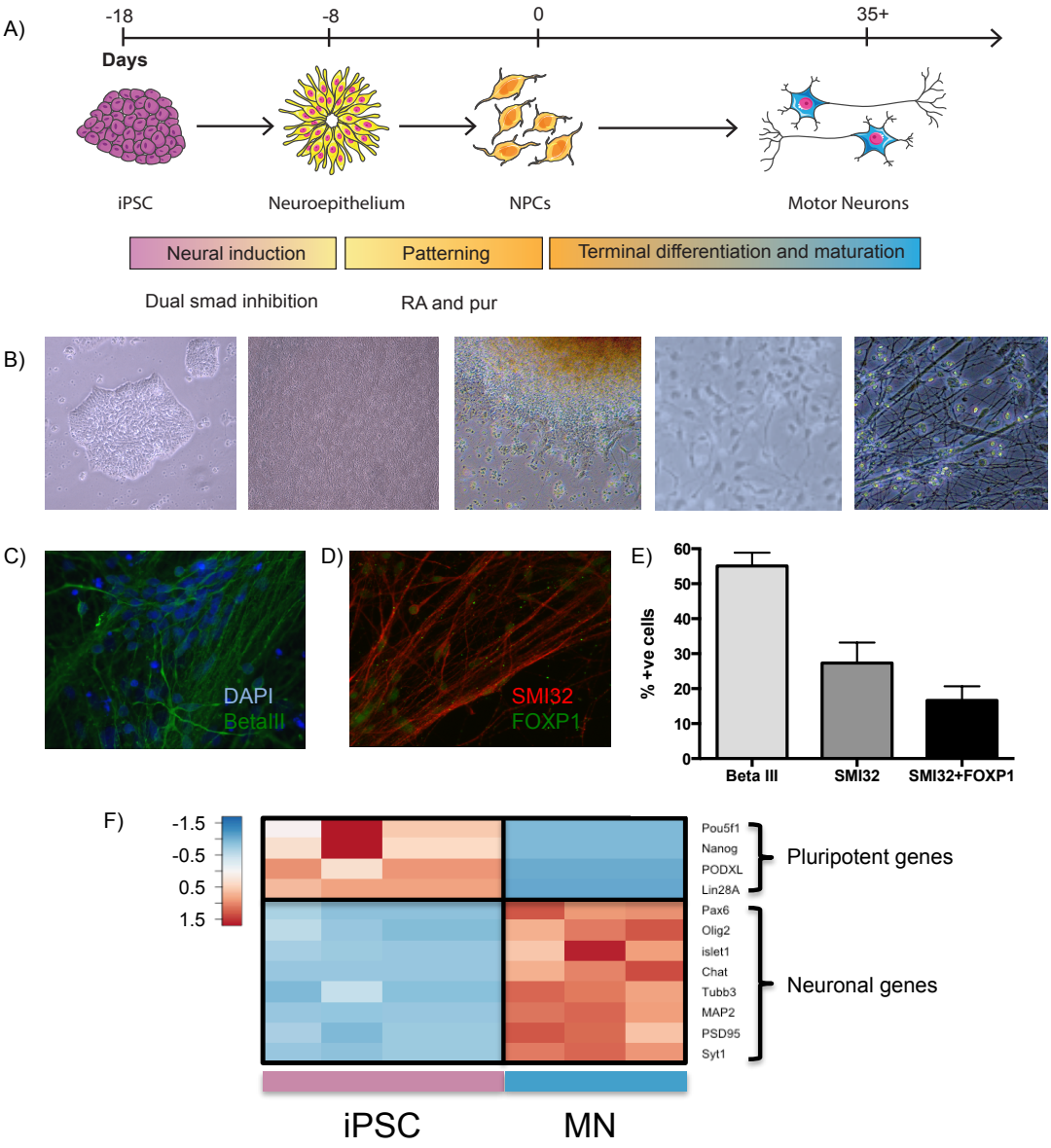


Figure 3.1; Generating MN in monolayer culture

A) Schematic to show the differentiation strategy employed for motor neurogenesis. Stages include; iPSC = induced pluripotent stem cells, NPCs = spinal neural precursors, MN = Motor neurons. B) Representative phase contrast images of 5 stages throughout differentiation. Representative immunocytochemistry images of MNs at Day 35 in culture. C) BetaIII (green) displayed with DAPI D) and SMI32 (red) with FOXP1 (green). E) quantitative immunocytochemistry of selected cellular markers BetaIII (N=3, 500+ cells counted), FOXP1 (N=2, 100+ cells counted) and SMI32 (N=3, 500+ cells counted). Error bars represent mean ± SEM. F) Heatmap showing classic cell-type specific gene expression from iPSC and MN cellular populations (N=3+ technical repeats per time point). Points represent normalized FPKM values that are log2 transformed and mean-centered.

To validate this MN differentiation protocol I carried out quantitative immunocytochemistry (qICC) using BetaIII tubulin (a general neuronal marker), SMI32 (a MN specific marker) and FOXP1 to gain insights into the MN subtype specified (Figure 3.1.C&D). This method produced cultures that are 58% neuronal, of which only 27% were MNs (Figure 3.1.E). Of the MNs (SMI32+ve cells), 59% were FOXP1 positive (Figure 3.1.E), which indicates they are from the lateral motor column (LMC), representing limb-innervating MNs.

In addition to immunocytochemistry validation I performed RNA sequencing of terminal MN cultures and looked at candidate gene expression in MNs vs iPSCs (Figure 3.1.F). This shows that there is successful relative down regulation of pluripotency markers (Pou5F1/OCT4, Nanog, PODXL, Lin28A) and up-regulation of neuronal (MAP2, Pax6, Tubb3/BetaIII), MN specific (Oli2, Islet1, Chat) and mature/synaptic markers (PSD95, Syt1).

Improving efficiency and accelerating motor neuron differentiation

Since producing the above differentiation protocol, several studies have successfully reduced the differentiation time and increased the efficiency of motor neurogenesis (details above). Therefore, to maximize the potential of hiPSC technology and to improve accuracy of downstream analysis by having purer cultures, I asked if we can we accelerate our current 50+day differentiation method and also if we can increase the enrichment of MNs produced (which was previously 27%)?

To adapt our existing protocol, I utilised key concepts from recent publications to produce a novel monolayer strategy for the differentiation of spinal MNs (Figure 3.2.A)(See appendix 8.1 for the detailed protocol). Namely the addition of CHIR99021 (a GSK3b antagonist / Wnt agonist) which if administered early in the differentiation protocol (during neural induction) increased efficiency of motor neurogenesis and the use of compound E which was shown to synchronize differentiation (H. Chen et al. 2014). This protocol uses a unique combination of signals to direct motor neuron specification and is broken into 4 stages (Figure 3.2.A):

1/ Neural conversion

Throughout neural induction CHIR99021 is used in addition to dual SMAD inhibition to aid neural specification. Maury et al have previously shown CHIR99021 increases the percentage of MN precursors specified (Maury, Côme, Piskorowski, Salah-mohellibi, et al. 2014).

2/ Spinal MN Patterning

Secondly, I patterned neural precursors with 0.5μM RA and 1μM purmorphamine for 7days (apposed to 8 as we previously did), and then I additionally treated NPCs for an additional 4 days with 0.1μM purmorphamine only. This was because it had previously been shown that longer SHH signaling is advantageous to generate NPCs specified to the pMN domain as longer exposure to SHH signals lead to sustained olig2 expression and increased the olig2 progenitor population (X. J. Li et al. 2008).

3/ Propagation (optimal)

I next included an optional stage for the propagation of NPCs that could be included when an increased yield of NPCs/MNs is required. To propagate NPCs I cultured NPCs for a maximum of 25days and 4 passages in the presence of 10ng/ul FGF (further details in the protocol found in Appendix.8.1). I show that propagating NPCs for up to 25days doesn't alter the % MNs specified during terminal differentiation using SMI32 staining (Appendix 8.2). Further, I found that propagating NPCs not only increased the yield of MNs but also greatly facilitated cell plating at defined densities, which is of great importance for a variety of assays in including live cell imaging.

4/ Terminal differentiation

Finally I used Compound E (a Notch inhibitor) to accelerate terminal differentiation by synchronizing exit from the cell cycle exit.

As previously, I did not utilize any growth factors throughout the protocol distinguishing our method from all other approaches. This significantly reduces the cost and thus makes the technology more widely accessible.

To validate this new differentiation approach I have used immunocytochemistry to characterize both the percentage of motor neuron precursors and the

percentage of MNs specified. Motor neuron precursors were defined by the presence of Olig2, Hoxb4 and/or Islet1 expression (Figure 3.2.B). After 7days of neural conversion and 7days patterning cultures had 62% Olig2 +ve, 21% HOXB4 +ve and 16% Islet1 +ve cells (Figure 3.2.C). After 17days of terminal MN differentiation culture were stained with two MN specific markers; CHAT and SMI32, and showed cultures were >85% +ve for MNs (86% SMI32+ve and 94% CHAT +ve) (Figure 3.2.D & E).

Therefore, compared to the first protocol I developed for motor neurogenesis and presented above, this adapted protocol has resulted in the acceleration in differentiation of approximately 20 days and a 3-fold increase in enrichment for MNs (up to >86%).

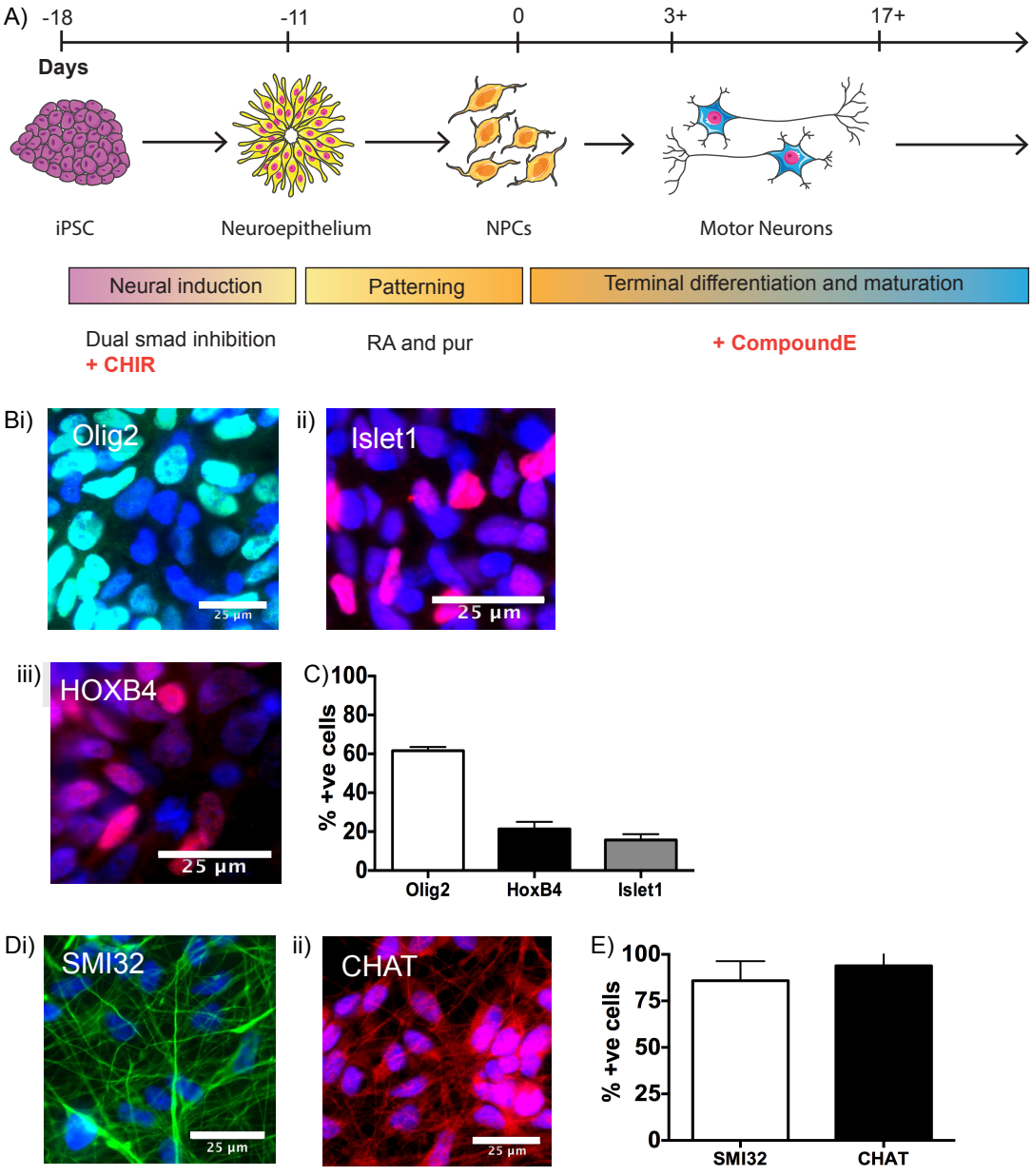


Figure 3.2; Improving efficiency and accelerating motor neurogenesis

A) Schematic to show our improved differentiation strategy for motor neurogenesis. B) Representative immunocytochemistry images of NPCs at D-4. Scale = 25µM, Displayed is i) Olig2 (green), DAPI (blue), ii) Islet1 (red), DAPI (blue) and iii) HOXB4 (red), DAPI (blue) C) quantitative immunocytochemistry of selected cellular markers of spinal NPCs; Olig2, Islet1 and HOXB4. N=3 on independent inductions, Error bars represent mean ± SEM. D) Representative immunocytochemistry images of MNs at D17. Scale = 25µM. Displayed is i) SMI32 (green), DAPI (blue) and ii) CHAT (red), DAPI (blue). E) quantitative immunocytochemistry of selected cellular markers of spinal MNs; SMI32, CHAT. N=3 on independent inductions. Error bars represent mean ± SEM.

RNAseq validation of the accelerated Motor neuron differentiation protocol

In addition to immunocytochemical validation, we conducted RNA sequencing to further confirm the specification of authentic MNs from the accelerated protocol. Libraries were prepared from pluripotent cultures, cultures after NI (D-7), after patterning (NPCs), from derived MNs at D3 (D3MN) and at D17 (D17MN)(Biological n=2 different control hiPSC lines, in technical replicate, overall n=4 per time point), using Truseq stranded mRNA RNAseq kit. All samples passed several stages of quality checks; including RIN scores >7.5, normal library distribution, 90% sequencing reads >Q30, rRNA reads <1%, exonic reads >70% (data not shown).

We first compared the gene expression of pluripotent cells to terminally differentiated MNs (and to iPSC-derived ACs) (Figure 3.8.A). We analyzed the mRNA expression of established cell type specific markers for pluripotency (POU5F1, MYC, NR5A2, Nanog, PODXL and LIN28), of general neuronal lineage (BetaIII tubulin, MAP2, RBFOX3, DLG4, Syt1, STMN2, NCAM1, SLC18A3) and specifically for MNs (Chat, NEFH, Islet1). These cell-type specific markers were able to separately cluster distinct populations of iPSCs and terminally differentiated MNs (and later we show hiPSC-derived ACs too), both confirming the identity of predicted cell types and also demonstrating low/undetectable expression in alternative fates (Figure 3.8.A). Further, we looked at relative candidate gene expression across these cell types and show the RPKM values of 3 pluripotency markers (POU5F1, PODXL, Nanog) and 3 MN markers (ISL1, SCL18A3 and CHAT) in both iPSCs and derived MNs (Figure 3.8.B).

Next I gained insights into MN subtype identity by analyzing the gene expression of HOX genes during MN differentiation. In development, opposing gradients of FGF2/RA signaling results in differential expression of the HOX genes, and consequently different positional identity along the rosto-caudal axis. Based of previous studies, using 0.5μM RA for 7 days we expect to generate progenitors located in the cervical region of the spinal cord, which corresponds to HOX gene expression paralogues 3-5 (Maury et al. 2014; Philippidou and Dasen 2013). Using RPKM values generated from RNAseq data, I showed that at NPC stage most HOX genes are expressed but during

differentiation their expression is refined (Figure 3.3.A). In our culture we generally lose the expression of HOX1 and HOX6-13 but the expression of HOX4 is maintained. This supports our hypothesis that our protocol generates MN positioned approximately in the cervical section of the spinal cord.

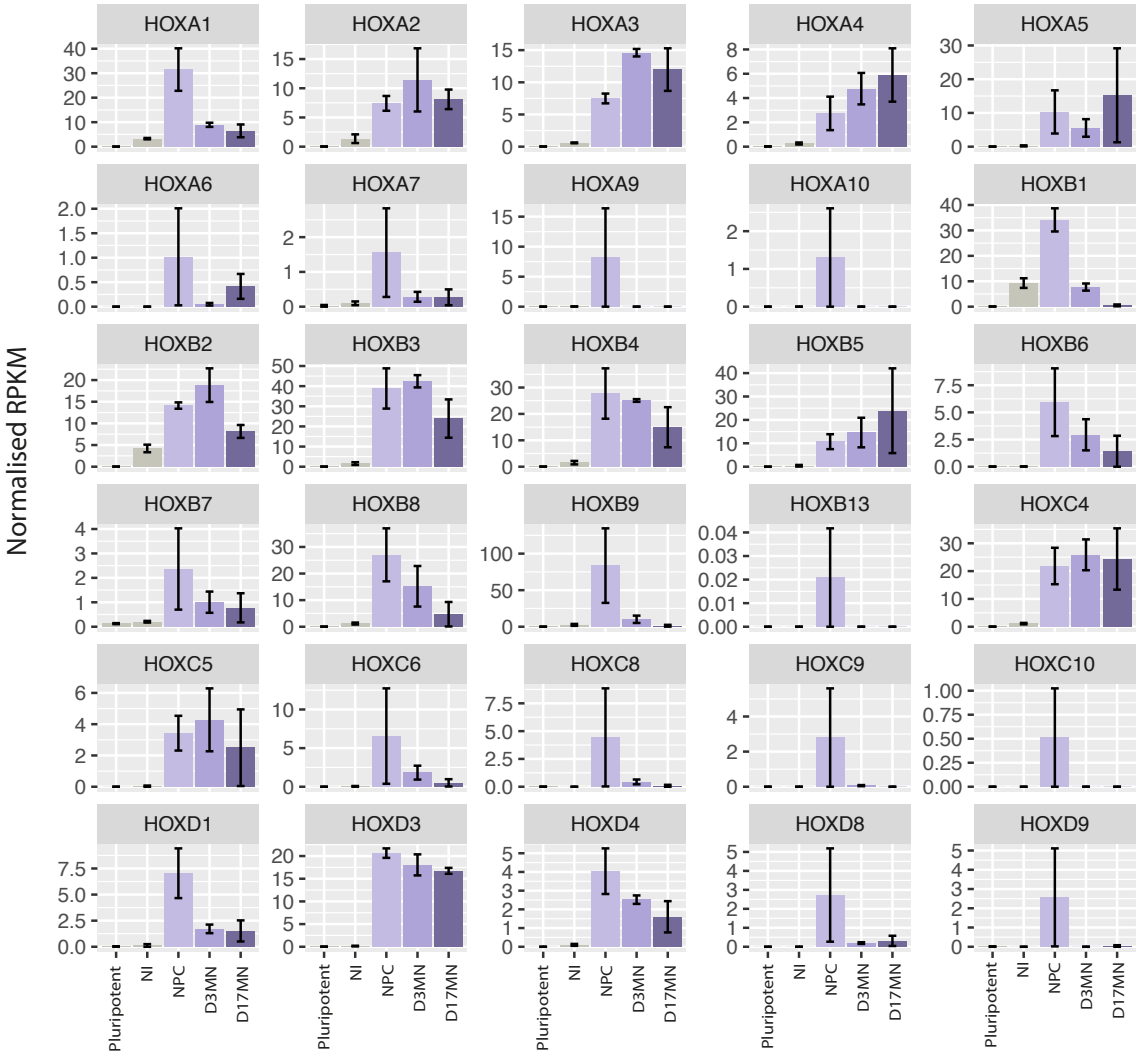


Figure 3.3; Expression of HOX genes throughout motorneurogenesis

Multiple bar plots displaying the relative expression levels of HOX genes at 5 stages in motor neuron differentiation from hiPSCs; pluripotent, NI (Day7/after neural induction), NPC (Day14/after spinal cord patterning), D3MN (immature motor neurons) and D17MN (electrically active Motor neurons). N=3+ for each time point, from 2 biological lines. Error bars represent mean RPKM value ± SEM.

Functional Characterization of hiPSC-derived Motor neurons

The ability of a neuron to carry out its functions depends both upon its structure and its ability to generate and communicate electrical and chemical signals. MNs primary role is to integrate and propagate signals across long distances.

For this, the basic requirements are having 1/ the capacity (receptors) to receive external information, 2/ the required intracellular components to both process the information and communicate this within the subcellular compartments, and 3/ the ability to transmit an electrochemical signal to its target. Here I have collaboratively analyzed multiple parts of this communication network within our iPSC-derived MNs.

Electrophysiology

One of the gold standard methods for functional characterization of neurons is analyzing their electrical activity. After 17days of terminal MN differentiation, Sarah Crisp assessed functional maturation using electrophysiology profiling. Both the passive (n=12 neurons from a minimum of 3 separate inductions across 2 biological lines) and active (n=16 neurons from a minimum of 3 separate inductions across 2 biological lines) properties were analyzed using whole cell patch clamping. Cells had a resting membrane potential of -57.7 ± 1.9 mV (mean ± s.e.m.) and current steps of 4784 ± 726 MΩ were applied to depolarise the cells. In all cells, this input resistance evoked action potentials (rheobase 3.5 ± 0.5 pA); the take-off was -45.0 ± 0.7 mV, the peak was 9.1 ± 2.3 mV and half-width was 4.1 ± 0.4 ms. For 1 of 16 cells only single action potentials were evoked, but the other 15 cells fired repetitively with increasing current injection (Figure 3.4.A). Tetrodotoxin (TTX) is a neurotoxin that binds to sodium channels blocking sodium permeablility and thus the ability to generate action potentials. In response to TTX (1mM) action potentials were reduced in frequency but not eliminated (n = 3) (Figure 3.4.B). Additionally spontaneous firing was recorded in some MNs.

Calcium response

In collaboration with Minee Choi, Zhi Yao and Sonia Gandhi, we used live cell imaging to assess the cytosolic calcium response to physiological calcium stimuli. After 17days of terminal MN differentiation, Minee Choi stimulated cells with glutamate/KCL to confirm the presence of glutamate receptors and voltage dependent calcium channels, which are characteristically present in neurons. She also stimulated cells with ATP, which is predicted to activate selected voltage-gated channels and initiate a calcium response in ACs. At day 17 of terminal differentiation, we confirmed that 98% of our MN cultures responded to KCL and glutamate stimulation, but not to ATP (Figure 3.4.C).

Axonal transport in Motor neurons

The function of neurons, and in particular MNs, relies heavily on their unique morphology, polarization and subcellular compartmentalization into the soma, axonal and dendritic sections. Hence, cellular activities must be coordinated throughout these compartments and often, given the length of MNs, over long distances too. A network of axonal transport machineries and pathways helps achieve this.

Transport occurs at differing speeds and directions (anterograde and retrograde). Anterograde transport make sure organelles, newly synthesized proteins and other messages from the soma reach their peripheral destinations. Then on the other hand retrograde transport ensures extracellular signals, ligands, aging proteins from the distal axon for recycling and degradation and selective organelles reach the cells soma. Slower speeds transport at 1mm/day (0.002-0.01μm/second) ranging up to faster speeds of a couple of hundred mm/day (0.5-4μm/second) (Tang et al. 2013). Although these faster speeds are still much slower than signals transmitted via action potentials.

Some pathogens, for example tetanus toxin, exploit the fast retrograde transport process to invade the nervous system. They enter the distal tips on an axon and travel to the soma by retrograde transport. Here in collaboration with Giulia Tyzack, Alexander Fellows and Gipi Schiavo we have begun to characterize retrograde axonal transport in our iPSC-derived MN cultures. Alexander Fellows treated MN cultures with HcT (Hc binding domain of tetanus neurotoxin TeNT); a ligand that is specifically taken up by neurons and undergoes retrograde transport, and used fluorescent imaging techniques that enable the direct visualization of transport. This enables us to monitor the transport kinetics of axonal signaling endosomes, which are intracellular compartments essential for neuronal differentiation and homeostasis.

We show that iPSC-derived MNs transport at speeds ranging up to ~5μm/s, with an average speed of ~2μm/s (Figure 3.4.E). Additionally we show no difference in the speed of retrograde transport in D3 and D17MN (Figure 3.4.E), which agrees with the recently publication from Nakamura et al where they showed no changes in speed or particle number in hiPSC-derived D5 and

D17 neurons (Nakamura et al. 2015). Further, it has been recently shown that the signaling endosome axonal transport dynamics do not changed between PMN taken from embryonic mice up until neurons taken from mice age 13months (Sleigh and Schiavo 2016). Therefore taken together the studies suggest axonal transport is a process that matures early in development and persists unchanged during 'youth'.

Next we compared transport speed in the iPSC-derived MNs to primary mouse cultures. We show that the signaling endosome axonal transport dynamics are comparable in primary mouse neurons (PMN taken from embyros day 12.5-14.5 and cultured in vitro for 5-7days) to our iPSC-derived MNs (D3/17), both again have an average speed of ~2μm/s and top speeds were shown up to ~5μm/s (Figure 3.4.F). This importantly demonstrates our in vitro differentiation is of comparable maturity to in vivo neural development and that the axonal transport profile we see in the hiPSC-derived MNs is reflective of neurons with matured signaling endosome axonal transport. In addition to gaining insights into neuronal function and maturation status, we have importantly established an assay that can later be exploited to address controversies in the role of axonal transport in disease and/or development.

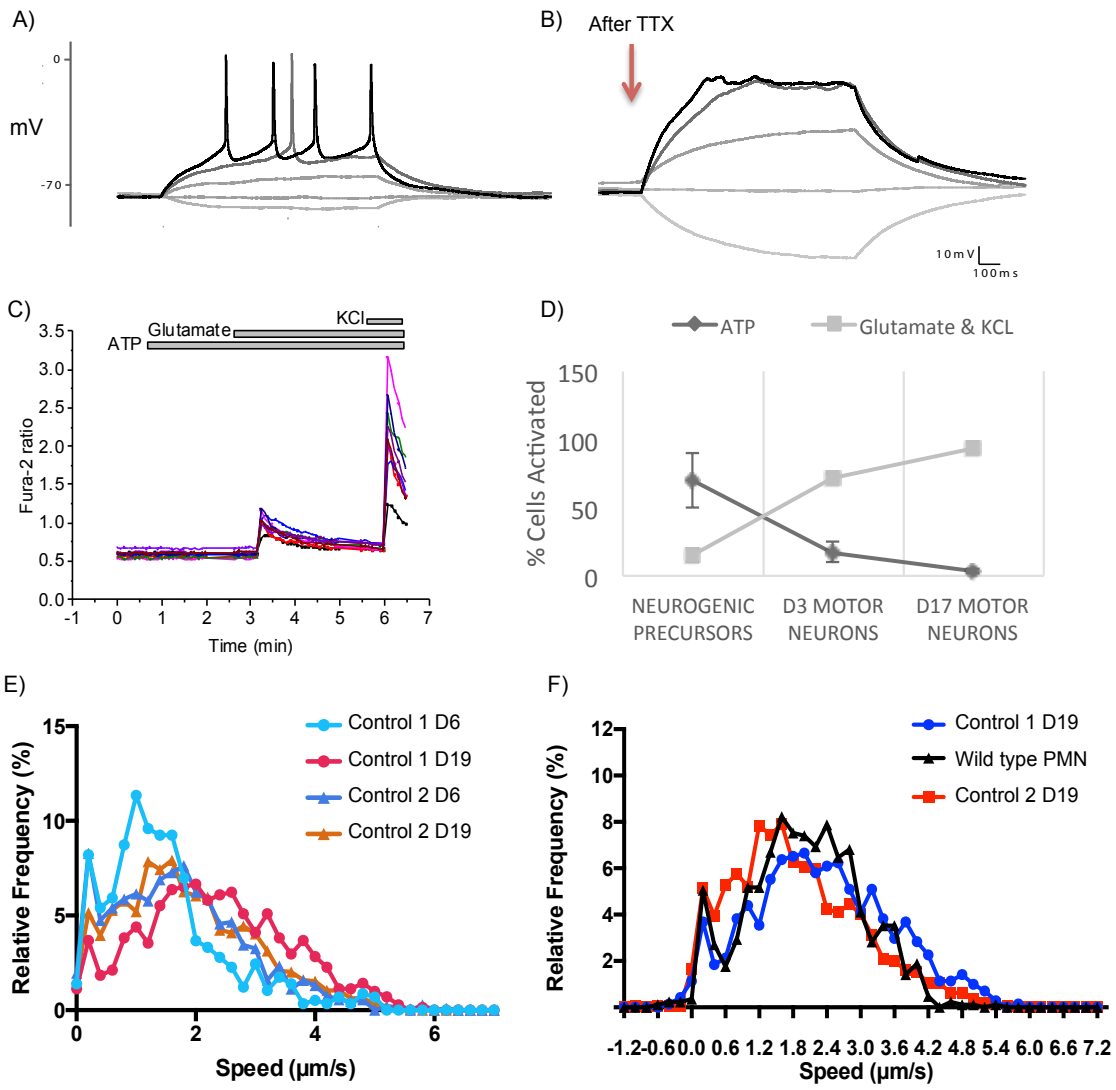


Figure 3.4; Functional validation of motor Neurons

A) iPSC-derived motor neurons generate action potentials at D17 in culture. A representative trace from a cell recorded during current injection (-1, 0, 1, 2, 3 pA) to evoke action potentials. B) Action potentials were blocked by TTX (-5, 0, 5, 10, 15 pA steps). C) Representative Calcium traces from D17MNs in response to ATP, Glutamate and KCL. D) Measurement of Calcium response to ATP, Glutamate and KCL (N=3+ biological replicates). Graph shows the % of cells stimulated at NPCs (n=66 cells), D3MN (n=68 cells) and D17MN (n=125 cells). At D17MNs, 98% of our cultures responded to KCl and glutamate stimulation, but not to ATP. E) Trace to show the speed and frequency of axonal transport in MN at days 6 and 19 in culture. N=2 biological replicates, with 3 technical replicates per time point). F) Trace to show axonal transport of hiPSC-derived MN at D37 compared to primary mouse neurons.

Differentiation of spinal cord Astrocytes from hiPSCs using a monolayer culture system

I next turned to develop a differentiation protocol for regionally defined spinal AC from hiPSCs. (Figure 3.5.A) I have used findings from our lab's previously published methods (Gupta et al. 2012) to guide the development of robust a protocol. In a similar way to neuronal development, AC differentiation can be divided into several stages, first beginning with neuronal conversion and patterning and finishing with terminal differentiation and maturation. As with neuronal specification, AC subtype diversity is specified early on during development during the patterning of neural precursors (Krencik et al. 2011), therefore we can utilize the same approaches as previously descripted for motor neurogenesis for both neural induction and patterning.

After neural conversion and patterning there is an additional step required for the propagation of neural precursors to form gliogenic precursors. This is necessary in order for the cell intrinsic mechanism described above, known as the gliogenic switch, to occur. When using human neural stem cells or fetal tissues differentiation of AC often requires serum, and the capacity to expand is limited. However I propagate NPCs in maintenance media supplemented with 10ng/μl FGF, without the need for serum and can expand NPC populations into large quantities of GPCs. Finally, the synergistic interaction between JAK-STAT and smad signaling have been shown to promote astrocytic differentiation (Gupta et al. 2012) so I have targeted both pathways during the terminal differentiation phase.

To summarise the approach; I carry out neural induction with dual smad inhibition (10μM SB431542 and 1μM Dorsomophin) for 10days, followed by 8days patterning with 0.5μM RA and 1μM Purmorphamine then propagation of spinal motor neural precursors for a minimum of 60days in 10ng/μl FGF. Gliogenic precursors are then terminally differentiated for a minimum of 2weeks in 10ng/μl BMP4 (acts through smad transcription factor family) and 10ng/μl LIF (activates the JAK-STAT pathway).

To validate the protocol I have carried about both immunocytochemistry and RNA sequencing. As highlighted earlier a major issue for AC characterisation is that markers for maturity and specificity are poorly defined. Consequently, I

have used both GLAST general astrocytic marker and GFAP for ICC staining. (Figure 3.5.C) At D28AC our culture is >70% GFAP +ve and >90% GLAST +ve (Figure 3.5.D).

For RNAseq analysis of AC compared to hiPSCs, I have used an array of additional markers of differing maturational stages to show relative up-regulation of astrocytic genes (VIM, NF1A to represent GPCs, and ALDH1L1, Acquaporin4, GLAST and GFAP to represent ACs) and down regulation of pluripotency markers (Pou5f1, Nanog, PODXL, Lin28A and NR5A2) (Figure 3.5.E). It is notable that is changes still occurring between D14 and D28 AC marking an increase of transcriptional maturation. At the later stage of AC terminal differentiation (D28) the expression of VIM and ALDH1L1 appears whilst GFAP, GLAST (SLC1A3) and AQP4 expression is reduced.

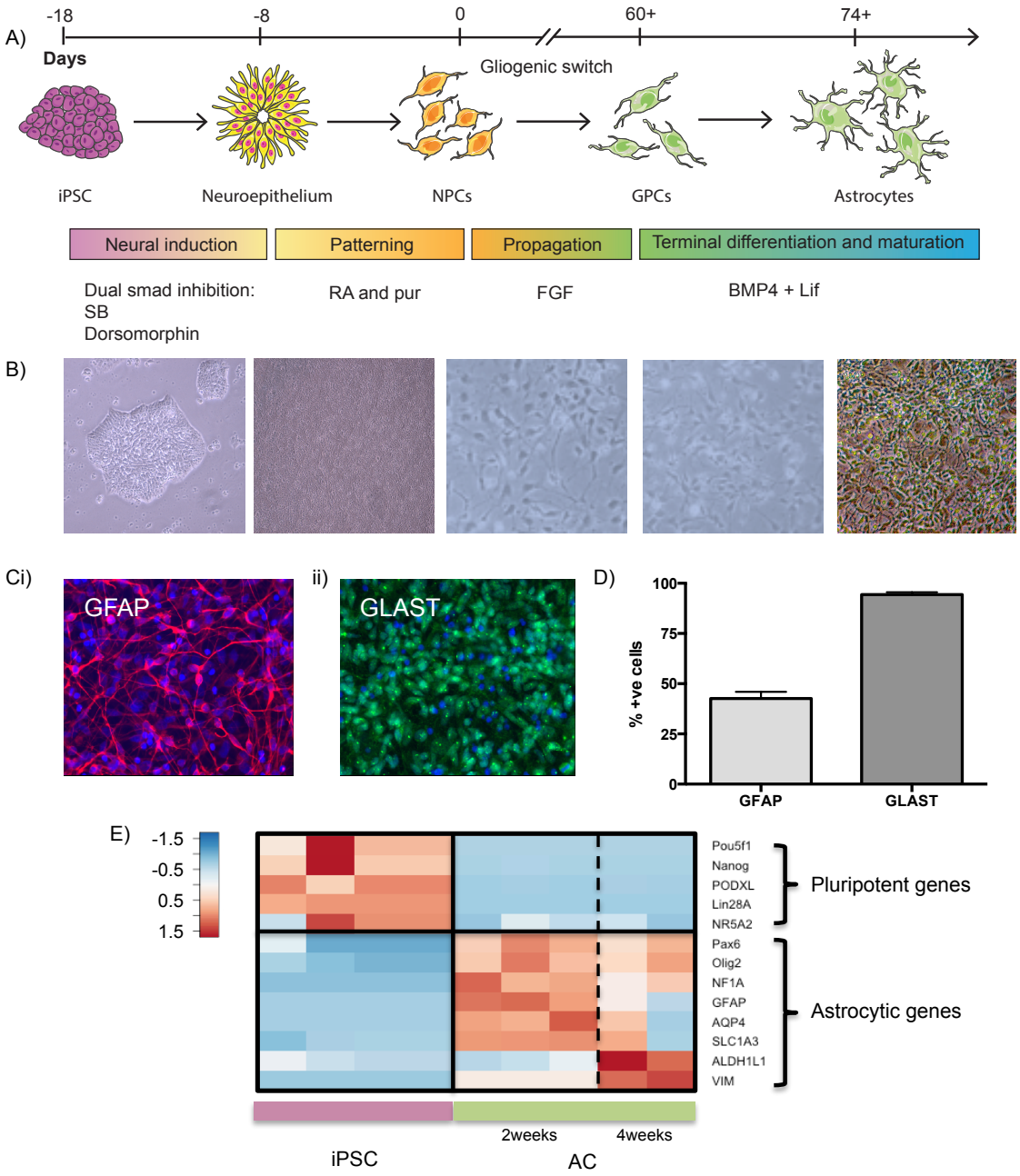


Figure 3.5; Generating AC in monolayer culture

A) Schematic to show the differentiation strategy employed for spinal cord astrogliogenesis. Stages include; iPSC=induced pluripotent stem cells, NPCs = spinal neural precursors, GPCs = Gliogenic precursors, AC = Astrocytes. B) Representative phase contrast images of 5 stages throughout differentiation. C) Representative immunocytochemistry images of i) GFAP (red) and DAPI (blue), and ii) GLAST (green) and DAPI (blue). D) Quantitative immunocytochemistry of selected cellular markers GFAP and GLAST. N=2+ independent inductions. Error bars represent mean ± SEM. E) Heatmap showing classic cell-type specific gene expression from iPSC and two stages of AC cellular populations (technical repeats plotted individually, n=2-4 per time

point). Points represent normalized FPKM values that are log2 transformed and mean-centered.

Astrocyte differentiation after accelerated neuronal precursor specification

Since I have adapted the neural conversion and patterning phases of the MN protocol in order to achieve faster and more efficient progenitor specificity, I asked if spinal AC be made from this adapted protocol. Therefore I have applied the same concepts to our AC differentiation protocol; I use 3 compounds to synchronize neural induction for 7 days followed by 11days patterning (7days with 0.5μM RA and 1μM Purmorphamine plus 4days with 0.1μM Purmorphamine only), propagation and terminal differentiation phases remain as before (Figure 3.6.A).

Following these changes I have validated cellular populations at gliogenic precursor stage and after terminal differentiation to AC (D28). To characterize the presence of gliogenic precursors I used NF1A and 3CB2 staining between day60-70 in FGF (Figure 3.6.B); generally, precursors committed to the AC lineage and immature AC are characterized by the expression of the transcription factor NFIA (Deneen et al. 2006). I found that 99.5% expressed NF1A and 76.1% expressed 3CB2 (Figure 3.6.C), reflecting highly pure population of GPCs. Further, after terminal AC differentiation I show the cultures are 71.5% GFAP +ve and 96.22% GLAST +ve (Figure 3.6.E).

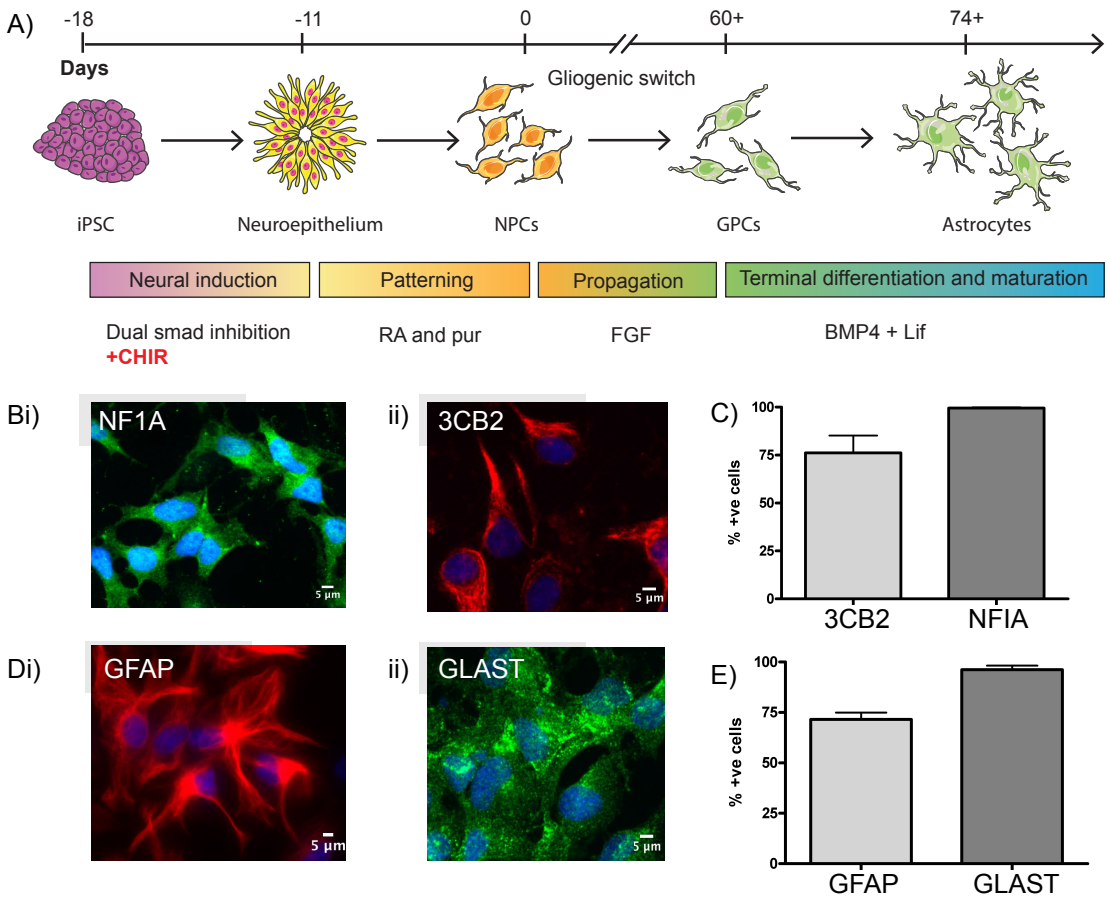


Figure 3.6; Astrogliogenesis from an adapted pMN domain specification protocol
A) Schematic to show our differentiation strategy for astrogliogenesis, after incorporating the first few stages of accelerated MN protocol. B) Representative immunocytochemistry images of i) NFIA (green), DAPI (blue) and ii) 3CB2 (red), DAPI (blue). Scale bar 5μM. C) Quantitative immunocytochemistry of selected cellular markers of GPCs; NFIA and 3CB2. ICC was carried out after propagation for 59days, n=3+ on independent inductions. Error bars represent mean ± SEM. D) Representative immunocytochemistry images of i) GFAP (red), DAPI (blue) and ii) GLAST (green), DAPI (blue). E) Quantitative immunocytochemistry of selected cellular markers of ACs; GFAP and GLAST. ICC was carried out at D14AC, n=3 on independent inductions. Error bars represent mean ± SEM.

Functional characterization of iPS-derived Astrocytes

For functional validation, in collaboration with Minee Choi, Zhi Yao and Sonia Gandhi, we have assessed the cytosolic calcium response to physiological calcium stimuli. Minee Choi stimulated GPC and AC cultures after 14 and 28 days of terminal differentiation with ATP, KCL and glutamate. We confirmed that 98% of cells exhibited calcium waves in response to local application of ATP, but not to KCL and glutamate (Figure 3.7.B), as previously reported (Abramov, Canevari, and Duchen 2003; Guthrie et al. 1999). This suggests our

AC cultures are highly enriched at day 14 and are maintained through to day 28.

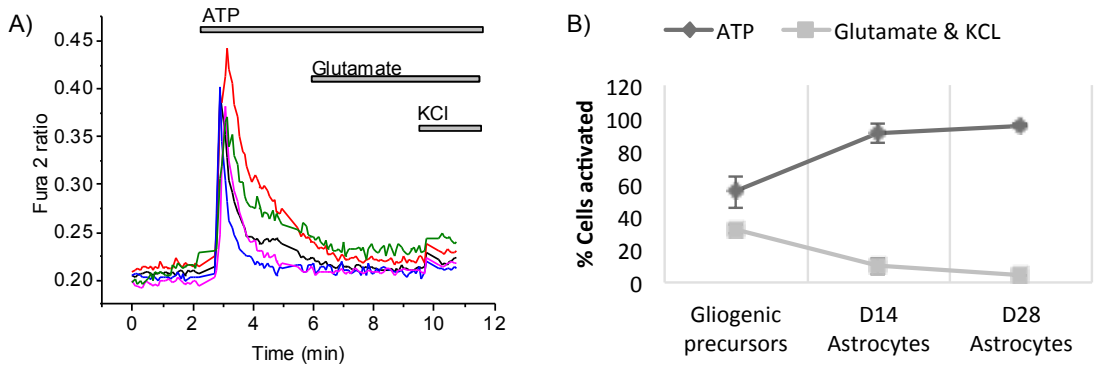


Figure 3.7; Functional characterization of iPSC-derived AC, Calcium response
A) Representative calcium trace of D28AC after treatment with ATP, Glutamate and KCL. B) Measurement of the % of cells stimulated by ATP, Glutamate and KCL (N=3 biological repeats/independent inductions). Graph shows the % of cells stimulated at GPCs (n=73 cells), D14AC (n=107 cells) and D28AC (n=113 cells). AC cultures show >98% cytosolic calcium responsds to ATP but not KCl.

RNAseq validation of the accelerated MN and AC protocols

In addition to immunocytochemical validation, we conducted RNA sequencing to further confirm the differentiation of ACs from the accelerated protocol. I prepared RNAseq libraries from two biological lines in duplicate (n=4) for terminally differentiated ACs (D28) and compared gene expression to pluripotent cultures and terminally differentiated MNs (D17) (Figure 3.8.A). I analyzed the mRNA expression (RPKM values) of established cell type specific markers for pluripotency (POU5F1, Nanog and LIN28), MNs (Chat, NEFH, Islet1, Olig2, Vacht) and ACs (AQP4, GFAP, GLAST). These cell-type specific markers were present in the predicted cell types confirming their identity. Also low/undetectable expression was seen in alternative fates emphasizing the purity of the cultures. Further, I plotted relative candidate gene expression across these cell types (iPSC, MN and AC) (Figure 3.8.B).

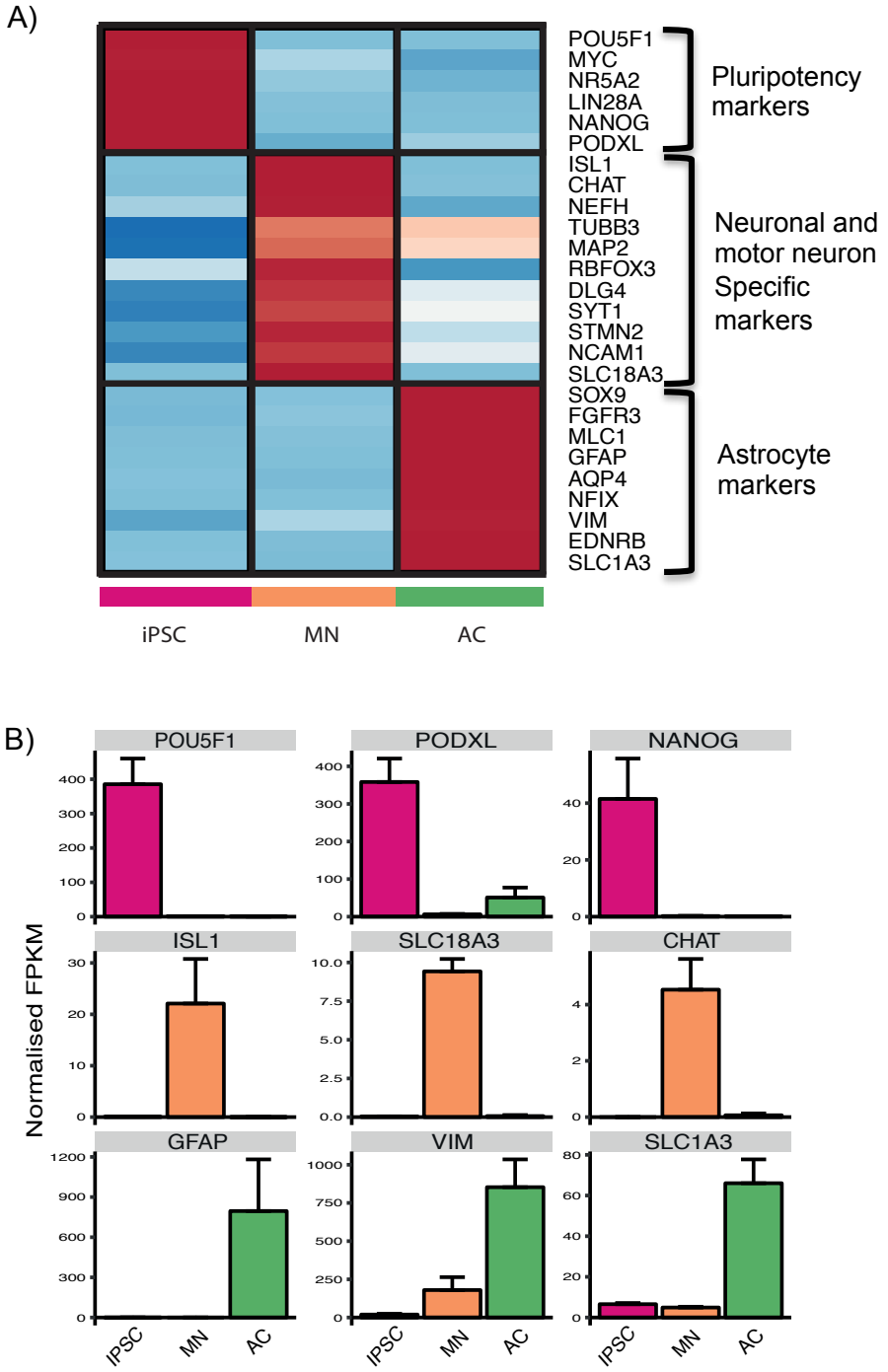


Figure 3.8; RNAseq validation of hiPSC-derived MN and AC

A) Heatmap showing classic cell-type specific gene expression from iPSC, MN and AC cellular populations (n=2, each with 2 technical repeats, per time point). Points represent normalized RPKM values that are log2 transformed and mean-centered. B) Representative normalized RPKM values of cell-type specific genes. Error bars represent mean RPKM values ± SEM.

Transcriptional analysis of astrogliogenesis in hiPSCs

Background

Since there is currently a lack of reliable astrocytic markers, it is undeniably valuable to explore the genetic expression of ACs, in order to try to identify novel and reliable candidate markers for classifying sub populations of ACs - depending on their morphology, activation and maturational status. To date there have been few studies that have explored the global transcriptional profile of ACs, and even fewer have investigated the transcriptional signature throughout astrogliogenesis. However, recently genome wide studies of cell-type-specific populations have become more available due to improved methods of generating purified populations of selected cell types (here the purification of ACs) and advances in genome wide methodologies (discussed separately in the Introduction).

In 2008 microarray analysis of purified cell types from mouse brain neurons, oligodendrocytes and ACs revealed ALDH1L1 as a new candidate AC marker, that has a much broader expression pattern than GFAP (Cahoy et al. 2008). Although this helped to improve our detection of ACs (and distinguishes them from oligodendrocytes and neurons), it did not provide information on astrocytic development or diversity.

In 2013, 5 time-points were selected throughout differentiation of ACs from hPSCs (both hiPSC and hES); NPC, day14, day21, day28 (onset of astrogliogenesis) and day35 and analysed by microarray analysis (Shaltouki et al. 2013). Analysis supported differentiation strategies showing stage-specific lineage gene markers up-regulated in specified cell types and highlight several genes enriched in AC (GFAP, S100B, APQ4, NFIX). Additionally, comparison to the murine AC transcriptional profile allowed confident detection of more AC specific genes (GLAST, ALdhL1, SLC25A18, SLC4A4, ATP1A2). Further, the importance of known regulators and the identification of novel markers and regulators of astrogliogeneis was touched upon (CHD protein 5, NFIA, NFIX, and OMG).

Moreover in 2014, a study again used microarrays to compare iPSC-derived NPCs to primary human fetal ACs. They showed immature ACs express high levels of GFAP and factors from NFI family and also highlighted the importance

of Notch, TGFb, MAPK and growth factor programs in AC development and maintenance (Malik et al. 2014). However, using only two time points taken from two different model systems, means further work is needed to validate the findings.

Lastly in 2015 RNA sequencing was used to compare the transcriptional signatures of fetal and mature, human and mouse primary ACs, using human neurons and oligodendrocytes as control. While similarities were seen between human and mouse populations many differences were detected between gene expression profiling and functional characteristics of human and mouse ACs. They characterised gene expression of human ACs and showed AC exist in at least two distinct developmental stages; immature/fetal proliferative and mature/postnatal (Zhang et al. 2016).

In summary, a mixture of model systems and methods have been used to explore global gene expression in ACs. Two major gaps in the research to date are;
1/ The limited transcriptional characterisation throughout astrogliogenesis.
2/ Comparative transcriptional analysis of iPSC-derived glia to primary human tissue.
Thus, here I have sought to explore both these areas using hiPSC gliogenic derivatives.

The transcriptional signature of neurogenic versus gliogenic precursors

First I have explored the transcriptional differences between neurogenic and gliogenic precursors. It is well established that neuronal precursors are capable of multi-lineage differentiation that is partially temporally regulated. Young NPCs are primarily neurogenic and responsive to morphogenetic instruction, permitting generation of region specific neurons. On the other hand, their late counterparts exhibit a loss in competence to be patterned into regionally defined neuronal subtypes, which coincides with the acquisition of gliogenic potential. Several studies have generated NPCs in vitro and demonstrate their ability to recapitulate key developmental properties, such as their multi-lineage differentiation potential and temporal responsiveness to morphogenic signals permitting the generation of region specific neurons and glia (Bouhon et al. 2006; Elkabetz et al. 2008; Okada et al. 2008; Reubinoff et al. 2001). Within the

lab we have also recapitulated key aspects of human neurodevelopment in hiPSCs, including demonstrating sequential generation of neurogenic then gliogenic precursors and the temporal restriction of NPCs to respond to patterning cues which coincides with the acquisition of gliogenic potential (Unpublished data, not shown). But to further explore differences between these two temporally distinct classes of hiPSC-derived NPCs that differ markedly in their developmental competence, I conducted RNA sequencing. I compared pluripotent cells, early NPCs (which were left propagating <30days) and Late NPCs (progenitors left propagating for <70days) (Figure 3.9).

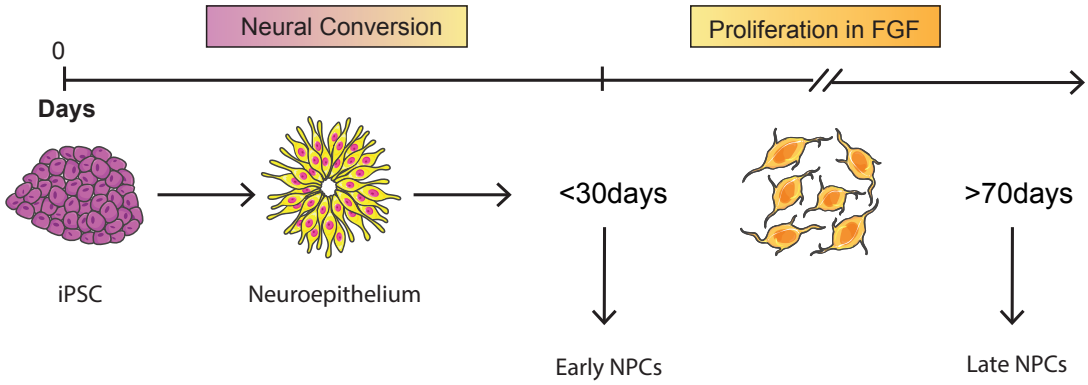


Figure 3.9; Schematic presentation of our strategy to generate Early and Late NPCs

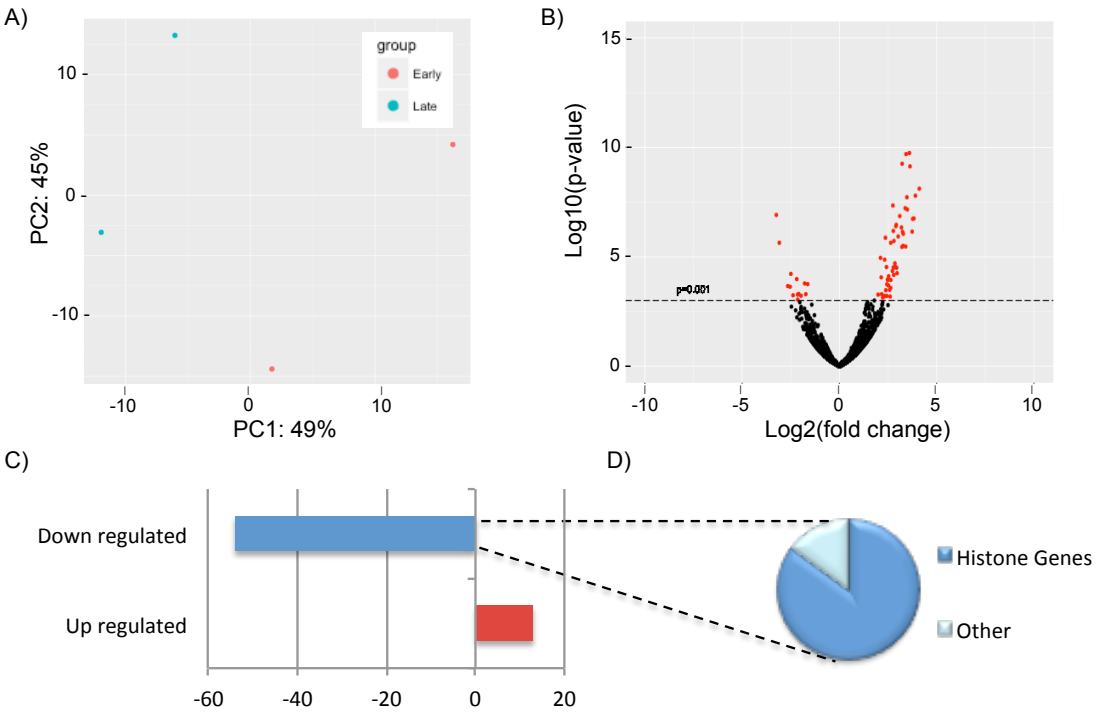
Experiments were carried out using 2 alternative approaches;

1/ First I used NPCs cultured by Rickie Patani from hESCs using a suspension culture system. Early precursors were defined as those propagated for 30days and late NPCs as those left to propagate for 100days. RNA from these samples was extracted using Zymo direct-zol extraction kit and the rRNA was depleted using Qiagen's generead kit. Next I carried out library preparation using the NEXTflex directional RNAseq kit.

2/ Secondly I used NPCs that I had generated from hiPSCs using our monolayer culture system. Using a monolayer system means that signals applied will reach cells in a more homeogenous manner, therefore we predict that the gliogenic switch would occur sooner. For this reason I defined early NPCs as those propagated for 20-30days and late NPCs as those propagated for 70-80 days. RNA was prepared using the automated Promega extraction

machine/kit, then polyA RNA was selected before libraries were prepared using Illumina's Truseq mRNA stranded RNAseq kit.

Firstly, using approach 1, I use principle component analysis (PCA) to describe the variation among samples. PCA identifies principle components in the data, which are directions that highlight the maximum variation in the data. Usually PC1 and PC2 account for the majority of variation in the data (here PC1 and PC2 account for 94% of the variation in the dataset), therefore by plotting PC1 versus PC2 it allows us to visualize how samples are clustered (shows similarities and differences in the data). I have shown separation between early and late NPC population using PCA (Figure 3.10.A). Next I have carried out differential expression analysis between the early and late NPC populations, this is represented using a volcano plot (Figure 3.10.B) and a bar plot to show the number of differentially expressed genes with log(fold change) >1.5 and p-value <0.05; 13 genes are upregulated and 54 genes are down regulated in Late NPCs compared to their early counterparts (Figure 3.10.C). Interestingly I noticed that, of the down-regulated genes in Late NPCs, over 85% of them correspond to histone coding proteins (Figure 3.10.D & E). Further I carried out gene ontology analysis. Figure 3.11 displays a condensed list of the most significant GO terms associated with both up and down regulated genes, generated using revigo. Among the most significant down-regulated biological terms were chromatin silencing, protein complex assembly and methylation, more specifically histone H3-K27 trimethylation, H3-K27 methylation and H3-K4 trimethlyation were highlighted. GO terms associated with molecular function and component parts also implicate a striking convergence on epigenetic changes in early and Late NPCs. Together with previous literature (Allen 2008; Y. Hirabayashi and Gotoh 2010; Majumder et al. 2013), this data suggest epigenetic regulation could play a primary role in the gliogenic switch. However due to heterogeneous culture system and suboptimal RNAseq approach I sought to replicate these results using an optimized approach.



E)

Top 10 Up regulated genes		Top 10 Down regulated genes	
Gene name	Log_2(fold change)	Gene name	Log_2(fold change)
FN1	-3.24	**HIST1H4L**	4.11
AHNAK2	-3.09	**HIST1H4F**	3.91
SCG2	-2.66	IGDCC3	3.83
S100A16	-2.53	**HIST1H2AL**	3.77
HTRA1	-2.49	**HIST1H3H**	3.74
CTSL	-2.38	**HIST1H2AE**	3.63
COL6A3	-2.19	RNU1-67P	3.60
MMP14	-2.15	**HIST1H2BH**	3.49
ASS1	-2.14	**HIST1H2AJ**	3.47
SH3BGRL3	-2.09	**HIST1H2BC**	3.43

Figure 3.10; RNA sequencing between Early-Late NPCs (approach1)

A) Principle component analysis shows Early and Late NPCs are distinct populations, n=2. B) Differential expression of Early and Late NPCs. C) Number of up & down regulated genes with fold change >1.5 and p value <0.05. D) >80% of all down regulated genes are proteins from the histone family. E) Top 10 differentially expressed genes.

	Up regulated			Down regulated		
	GO term	p-value	Enrichment score	GO term	p-value	Enrichment score
Biological process	extracellular matrix disassembly	4.83E-08	43.48	chromatin silencing	9.87E-08	39.92
	response to endogenous stimulus	6.46E-05	7.29	mucosal immune response	2.18E-06	99.8
	glial cell migration	6.24E-05	153.03	protein-DNA complex assembly	4.75E-23	54.79
	collagen metabolic process	5.30E-07	54.65	methylation	5.61E-04	10.11
	extracellular structure organization	3.21E-07	18.07	cellular component organization or biogenesis	1.43E-05	2.18
Molecular function	peptidase regulator activity	2.97E-04	21.66	protein heterodimerization activity	5.02E-15	19.51
	enzyme regulator activity	6.04E-04	6.42	DNA binding	1.95E-10	5.1
	peptidase activator activity	9.36E-04	42.51	chromatin DNA binding	4.29E-05	19.47
Component part	extracellular matrix	3.06E-08	26.69	nucleosome	4.38E-36	121.19
	vesicle	3.00E-06	3.55	macromolecular complex	4.59E-06	2.33
	collagen type VI trimer	1.88E-05	255.06	extracellular region part	1.30E-06	2.92

Figure 3.11; GO term analysis

GO analysis of differentially expressed genes between early and Late NPCs (approach 1) using GOrilla. Redundant terms were condensed using REVIGO.

Shifting to experimental approach 2, I have shown that pluripotent, early and late NPCs cluster as distinct populations in Principal component analysis. There is a greater variation seen between pluripotent cells and neural precursors (≈70%), than between early and late NPC populations (≈22%), as expected. Moreover cell populations express cell type specific markers; I have used Pou5f1, Nanog, PODXL, NR5A2, Lin28A to mark pluripotency, Nestin & Pax6 as early NPC markers and NF1A, VIM for their late counterparts (Figure 3.12).

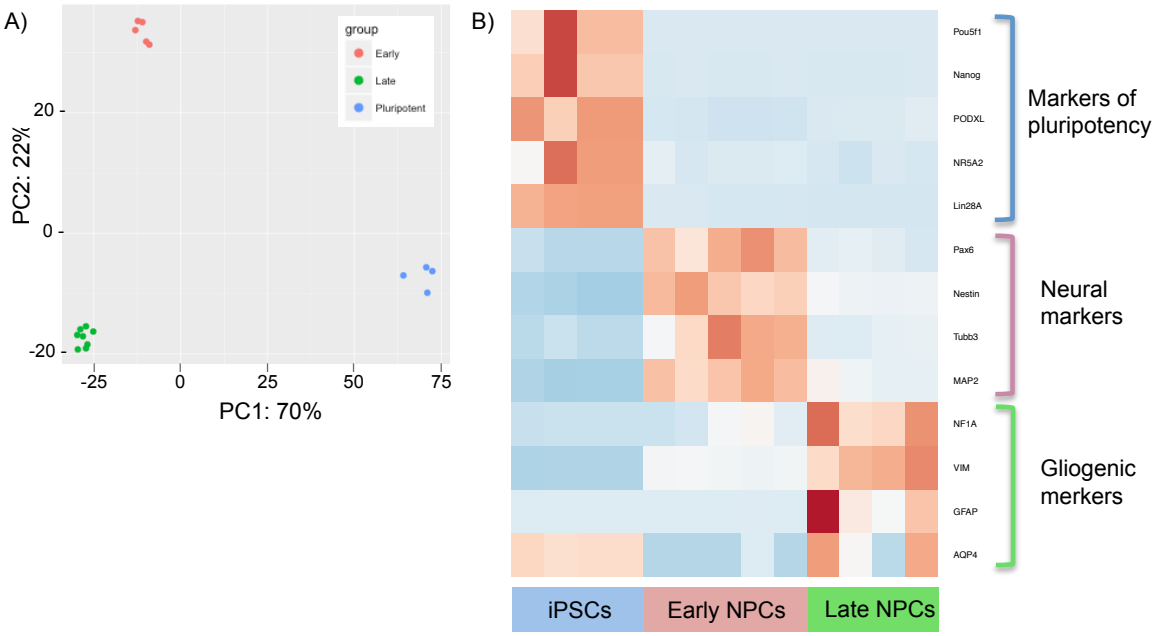


Figure 3.12; RNAseq of Early and Late NPCs, approach 2

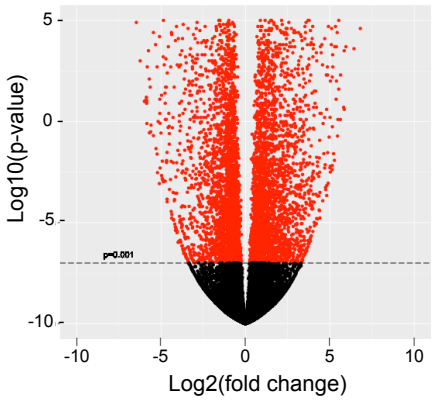
A) Principle component analysis shows Early NPCs, Late NPCs and iPSCs as distinct cellular populations. B) Heatmap showing classic cell-type specific gene expression from iPSC, NPC and GPC cellular populations (n=3 technical repeats per time point). Points represent normalized FPKM values that are log2 transformed and mean-centered.

Next I carried out differential expression analysis between early and late NPCs. I show 1488 down regulated genes and 2050 up regulated genes with a log(fold change) >1.5 and p-value <0.05 in early versus late NPCs (Figure 3.13.A & B). The larger changes observed between early and late NPCs in this data set compared to data generated using approach1 could be for several reasons; approach 2 uses an improved RNAseq methodology (discussed in chapter 4) and NPCs have been derived from an improved differentiation method which is a solely monolayer strategy – this ensures more homogenous signaling in culture therefore generates a less heterogeneous cellular population. Due to the large numbers of changed genes I have performed GO analysis and displayed only a selection of the top terms highlighted (a full list can be found in Appendix 8.3) (Figure 3.13.C). Many of the terms reflect NPCs cell fate potential, for example synaptic regulation, neuronal differentiation, microtubule motor activity and ion transport are down-regulated in Late NPCs, reflecting the loss of neuronal potential. This coincides with the up regulation of Ca+ handling genes (eg. S100B), growth factor binding, cytokine signaling (eg. BMP1,2,4, WNT1, LIF, TGFB1) and AC morphological transformation (integrin

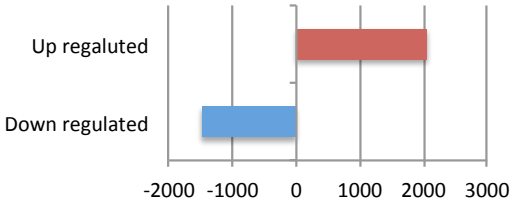
and heparin binding), which are all processes implicated in glial differentiation (Cahoy et al. 2008). Additionally many GO terms highlighted to be up/down regulated in early versus late NPCs are consistent with my previous data set (approach 1) even though many individual genes are not replicated; there is drastic up regulation in genes contributing to the extracellular region (COL6A1 is significantly up-regulated in Late NPCs in both data sets) and down regulated genes involved in DNA binding.

Overall, taking both approaches combined, I have shown distinct transcriptional signatures of early and late NPCs that reflect their developmental potential. Key differences include changes in extracellular matrix composition, chromatin methylation, gene silencing, Ca+ handling and cytokine signaling. However further experiments are required to define the timing and what regulates the switch between early and late NPCs.

A)



B)



C)

		Up regulated			Down regulated	
	GO terms	p-value	Enrichment score	GO terms	p-value	Enrichment score
Biological process	extracellular matrix organization	8.50E-40	3.8	synaptic signaling	9.43E-25	3.29
	regulation of cytokine secretion	1.07E-07	2.99	neuron differentiation	4.05E-13	3.15
	regulation of transforming growth factor beta receptor signaling pathway	3.54E-05	2.77	cellular potassium ion transport	5.97E-11	4.21
	cellular response to interferon-gamma	8.39E-05	4.6	neurotransmitter secretion	1.02E-09	4.16
	cell fate commitment	9.85E-04	2.04	regulation of ion transport	7.98E-09	2.1
	regulation of cell motility	3.35E-21	2.27	spinal cord motor neuron migration	7.18E-05	10.85
	positive regulation of cell migration	1.16E-20	2.76	calcium ion regulated exocytosis	3.41E-04	3.76
Molecular function	calcium ion binding	6.57E-14	2.1	ion channel activity	1.80E-15	3.12
	extracellular matrix structural constituent	3.04E-12	4.87	passive transmembrane transporter activity	3.04E-14	2.91
	cytokine activity	1.01E-11	3.89	DNA binding	2.33E-04	1.25
	growth factor binding	1.71E-10	3.18	potassium channel activity	4.23E-11	4.57
	peptidase regulator activity	1.72E-09	2.9	cGMP binding	3.90E-06	9.3
Component part	extracellular region	1.18E-47	2.97	synapse part	4.28E-19	2.45
	MHC protein complex	3.66E-06	8.04	neuron part	1.56E-17	1.92
	collagen type IV trimer	2.38E-04	8.04	ion channel complex	6.00E-12	3.16

Figure 3.13; Differential expression analysis between Early and Late NPCs from Approach 2

A) Differential expression between Early and Late NPCs. B) Number of up and down regulated genes between Early and Late NPCs. C) GO term analysis (terms with the lowest dispensability displayed).

Timing and regulation of the gliogenic switch

Before I could look further into the transcriptional regulation underlying the gliogenic switch I needed to first narrow down the time frame that it occurs in, in our monolayer culture system. To do this I collected NPCs every 10days throughout propagation in FGF from Day20 until Day80 (in biological n=1 and

technical n=3) and performed RNA sequencing using Illumina's Truseq stranded mRNA kits. Principle component analysis revealed segregation of samples depending on the number of days they have been propagated. I noted an increase in NPC separation between 40-60days of propagation (Figure 3.14.A), which would suggest that the gliogenic switch occurs during this time window.

I repeated RNAseq on NPCs propagated between day 40 and 70 days using an additional biological replicate (biological n=2, technical n=3). Principle component analysis showed that biological replicates are separated by PC2 and that PC1 separated samples based on the duration NPCs were propagated in FGF, which confirms the main variation in NPCs is between those propagated between 40 and 60days (Figure 3.14.B). Using two biological replicates in parallel I noticed that for one line the largest variation is seen between day 40-50 and the other line between day 50-60. This highlighted that the cellular alterations underlying the gliogenic switch could occurs at slightly different times in each patient line. Thus, I next used just one biological line to focus in on 5day windows throughout NPC propagation between 40 and 60days to study the sequence of molecular events throughout the gliogenic switch. I selected 1 line only because I wanted to limit the noise created by different cell lines that undergo the gliogenic switch at slightly different times, but of course it will be of upmost importance to validate any findings on additional biological lines.

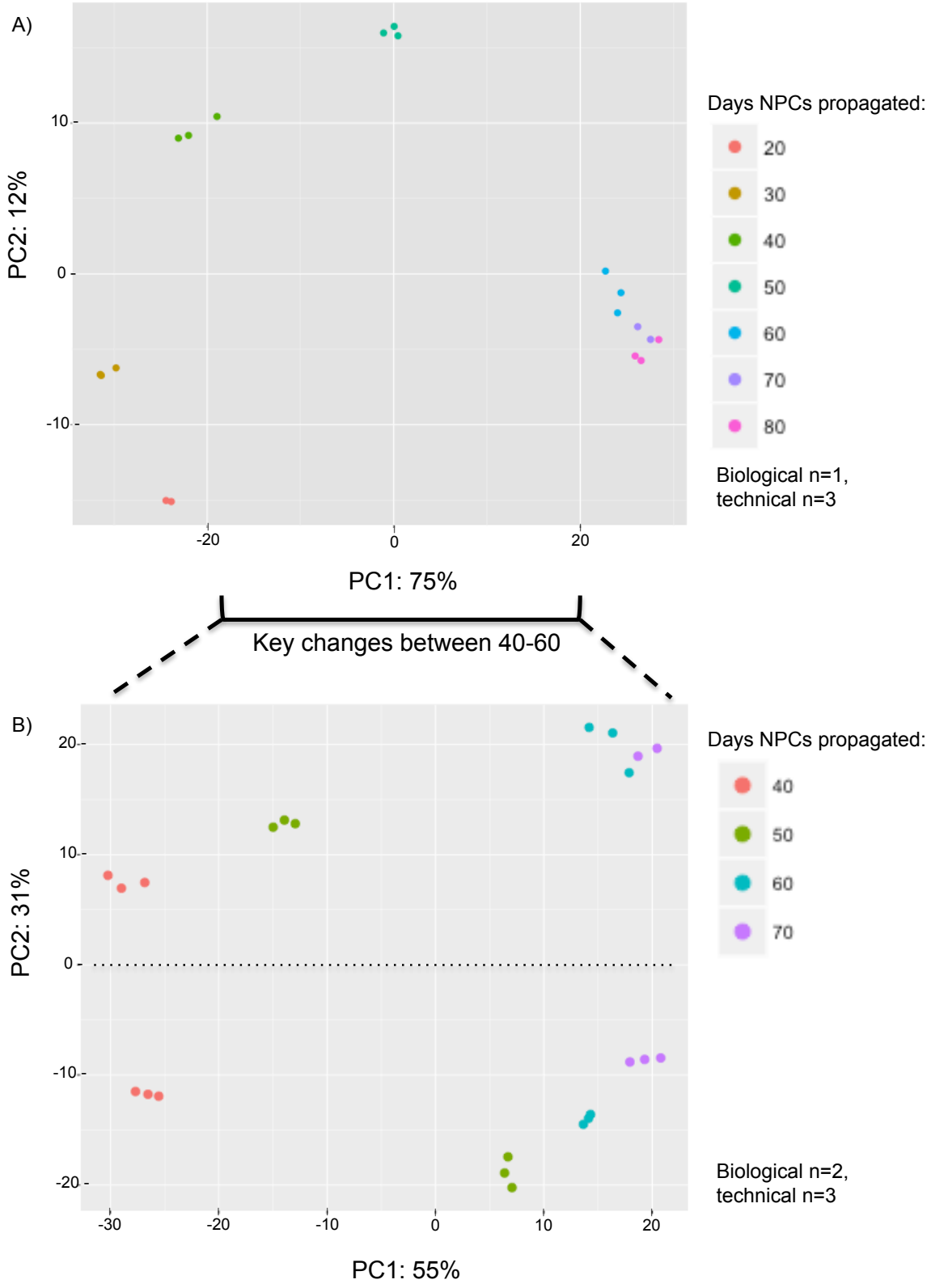


Figure 3.14; Defining when the gliogenic switch occurs in our system

A) Principle component analysis of NPCs propagated in culture for 20-80days (biological n=1, technical n=3). B) Principle component analysis of NPCs propagated in culture for 40-70days (biological n=2, technical n=3)

After carrying out RNAseq on precursors propagated for 40, 45, 50, 55 and 60 days, I have again first carried out principle component analysis to visualize the data separation; increased separation is seen between day 40-55 (Figure 3.15.A). Subsequently I have carried out differential expression analysis, using R package DEseq2, between each 5day time-point (day 40-45, 45-50, 50-55) (Figure 3.15.B). Next I performed GO analysis using differentially expressed genes with log(fold change) >1.5 and p-value <0.05 (Figure 3.16.A). I show there is a time line of transcriptional changes throughout the gliogenic switch (summary shown in Figure 3.15.C & D). As expected transcriptional changes in NPCs entering into gliogenic switch (Day45) begin with a down regulation of genes involved in neuronal differentiation and neural-specific functions, along with down regulation of genes involved in DNA binding and up regulation in genes controlling cellular adhesion, migration, locomotion, peptidase and cytokine activity. Following, at Day50 there is up regulation of genes encoding growth factors but they are quickly down regulated at Day55 along with genes directing cytokine activity and locomotion. Also at Day55 Wnt signaling is up regulated. Moreover, there are continual changes in genes encoding components of the extracellular matrix and space throughout the gliogenic switch, from Day 40-55.

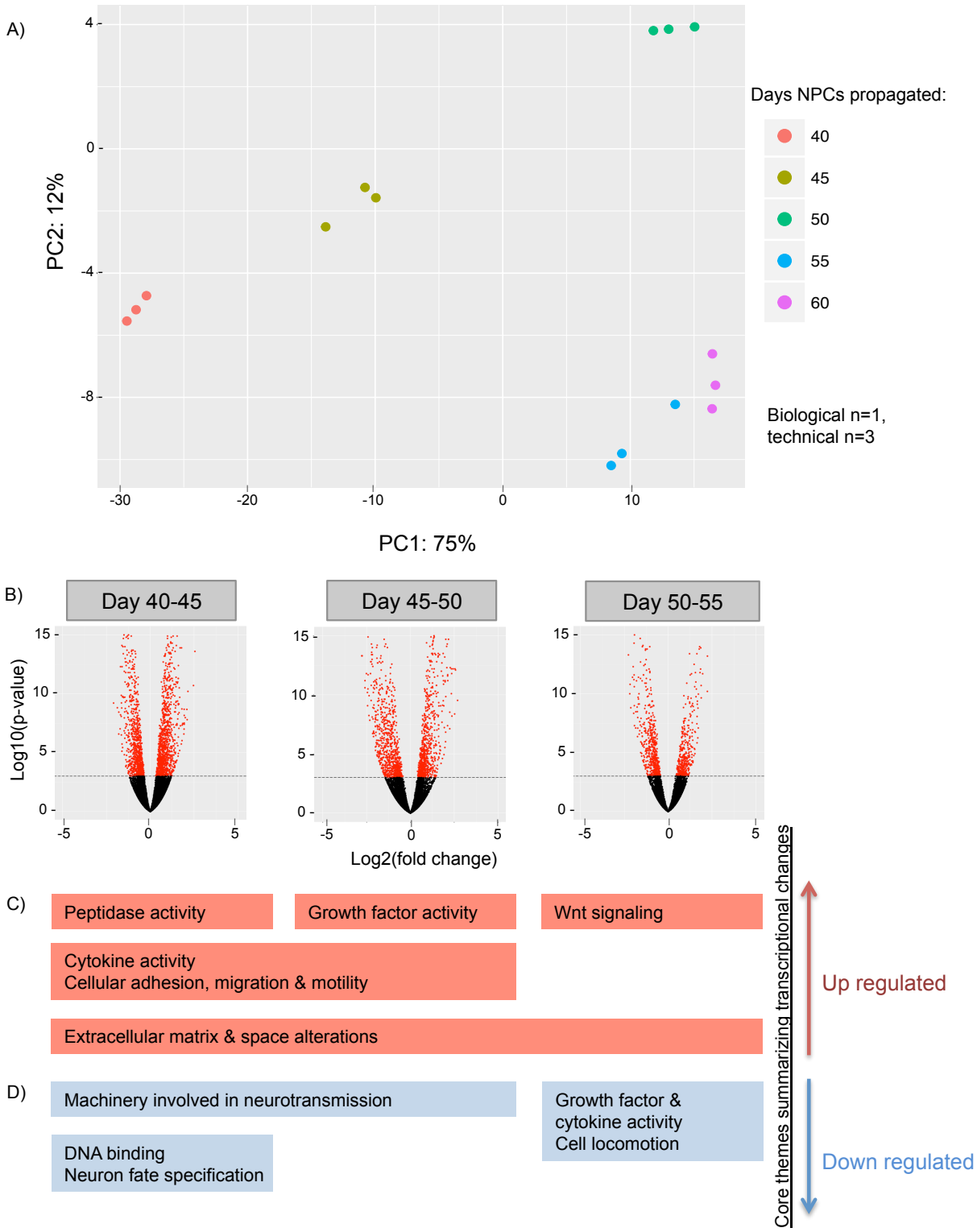


Figure 3.15; Identifying the sequence of transcriptional changes across the gliogenic switch

A) Principle component analysis of NPCs propagated in culture for 40 – 60 days. B) Paired differential expression analysis between NPCs propagated between 40-55days (40-45days, 45-50days and 50-55days). C) and D) Map highlights the key themes associated with transcriptional changes at each time point, categorized into up and down regulated genes respectively.

	D40-45				D45-50				D50-55			
	GO term	Description	P-value	Enrichment score	GO term	Description	P-value	Enrichment score	GO term	Description	P-value	Enrichment score
Up regulated	GO: 0050840	extracellular matrix binding	3.97E-09	28.22	GO: 0008083	growth factor activity	1.10E-10	15.28	GO: 0030247	polysaccharide binding	2.41E-04	83.98
	GO: 0005102	receptor binding	8.15E-06	3.21	GO: 0005125	cytokine activity	1.40E-10	18.08	GO: 0001871	pattern binding	2.41E-04	83.98
	GO: 0005518	collagen binding	9.70E-06	17.21	GO: 0001968	fibronectin binding	5.43E-08	27.38	GO: 0030246	carbohydrate binding	3.59E-04	11.63
	GO: 0005125	cytokine activity	2.61E-05	14.11	GO: 0005520	insulin-like growth factor binding	4.53E-07	29.94	GO: 0042813	Wnt-activated receptor activity	7.94E-04	47.24
	GO: 0004222	metalloendopeptidase activity	6.89E-05	11.56	GO: 0005126	cytokine receptor binding	1.19E-06	8.37				
Down regulated	GO: 0022838	substrate-specific channel activity	2.45E-05	10.08	GO: 0015075	ion transmembrane transporter activity	1.86E-13	4	GO: 0008083	growth factor activity	1.95E-06	15.86
	GO: 0005216	ion channel activity	2.45E-05	10.08	GO: 0005231	excitatory extracellular ligand-gated ion channel activity	2.12E-13	19.37	GO: 0005179	hormone activity	5.47E-06	33.22
	GO: 0022803	passive transmembrane transporter activity	3.79E-05	9.33	GO: 0022891	substrate-specific transmembrane transporter activity	7.74E-13	3.74	GO: 0071855	neuropeptide receptor binding	6.19E-06	74.75
	GO: 0015267	channel activity	3.79E-05	9.33	GO: 0030594	neurotransmitter receptor activity	1.59E-12	16.87	GO: 0005102	receptor binding	7.11E-06	3.52
	GO: 0008066	glutamate receptor activity	3.85E-05	44.66	GO: 0022857	transmembrane transporter activity	2.22E-12	3.54	GO: 0005539	glycosaminoglycan binding	2.60E-05	10.16

Figure 3.16; GO analysis across the gliogenic switch

Displays the top 5 up and down regulated GO terms associated with Molecular function at three time-points during the gliogenic switch; Day 40-45, day 45-50 and day 50-55.

Further, to validate these findings I turned to analyze the transcriptional profile across the gliogenic switch using an additional biological replicate (biological n=2, technical n=6). I selected representative genes from each pathways highlighted as being important in the gliogenic switch as well as genes involved in astrogliogenesis (Ca+ handling, JAK/STAT regulation) and have plotted their relative expression (RPKM values) across 6 time-points during NPC propagation in FGF from 20days to 80days (using 2 biological replicates). This shows a clear transcriptional switch between neural precursors, where genes

involved in neural specification, neural-specific functions, RA metabolism and GTPase activity are up-regulated and gliogenic precursors, where there are large changes in extracellular components and up regulation of genes that control Ca+ handling, JAK/STAT and cytokine signaling (Figure 3.17.A & B and Figure 3.18.F & G). In the middle of the switch (approximately day 50), we observe a transcriptional switch point; some neural pathways are still being down regulated while genes involved in astrogliogenesis are being to be up regulated. But also there are a few pathways that are uniquely up regulated at this 'transition' time point, including a subset of growth factors (IL11 and HBEGF) and genes controlling cellular adhesion (FN1), but their expression is reduced promptly after Day50 (Figure 3.17.F and Figure 3.18.A). Interestingly both genes that control growth factor activity and cellular adhesion are up regulated later on after the gliogenic switch as well as at Day50, however it appears that different genes within these pathways are up regulated at each stage. For example, at Day 80 growth factors IL31RA and LIF are expressed where as IL11 and HBEGF are up regulated at D50 (Figure 3.18.A & G). However we also observe CTGF expression peaks at both Day50 and Day 80 suggesting there is some overlap in genes expressed both during and after the gliogenic switch (Figure 3.18.A).

As expected, many pathways highlighted during this temporal analysis throughout the gliogenic switch overlap with the previously presented data that compared the differentially expressed pathways between extreme early and late NPCs. Pathways identified using both strategies, and using NPCs derived both from hiPSCs and hESCs, include those regulating extracellular matrix changes, peptidase activity and DNA binding (Figure 3.17.D & G). Conversely, several additional pathways that were differentially expressed between the extreme early and Late NPCs, did not appear in the top differentially expressed pathways highlighted during the refined gliogenic switch window (day 40-55). Two such processes were histone regulation and expression of genes associated with MHC protein complexes. This suggests these pathways could play a role in regulating changes that happen immediately before or after the largest alterations underlying gliogenic switch occur. To investigate this I have temporally explored the transcriptional changes of selected genes from these pathways, between early to late NPCs in 10day intervals. This showed that selected genes involved in histone regulation (HIST1H2BH, HIST1H4K and SUV39H1) are down-regulated across the gliogenic switch, but with the main

changes occurring early on (day 20-40) (Figure 3.17.C), and genes associated with the MHC complex (HLA-DRA, HLA-DPA1 and CIITA) are up-regulated shortly after the gliogenic switch (Day60+) (Figure 3.18.C). Both observations are consistent with the previous data that compared transcriptional changes between extreme early and late NPCs. Here the temporal analysis has now provided us additional insights into the timing of these pathways during the gliogenic switch; histone regulation appears to be an early signature whereas changes in the MHC complex occur after the gliogenic switch.

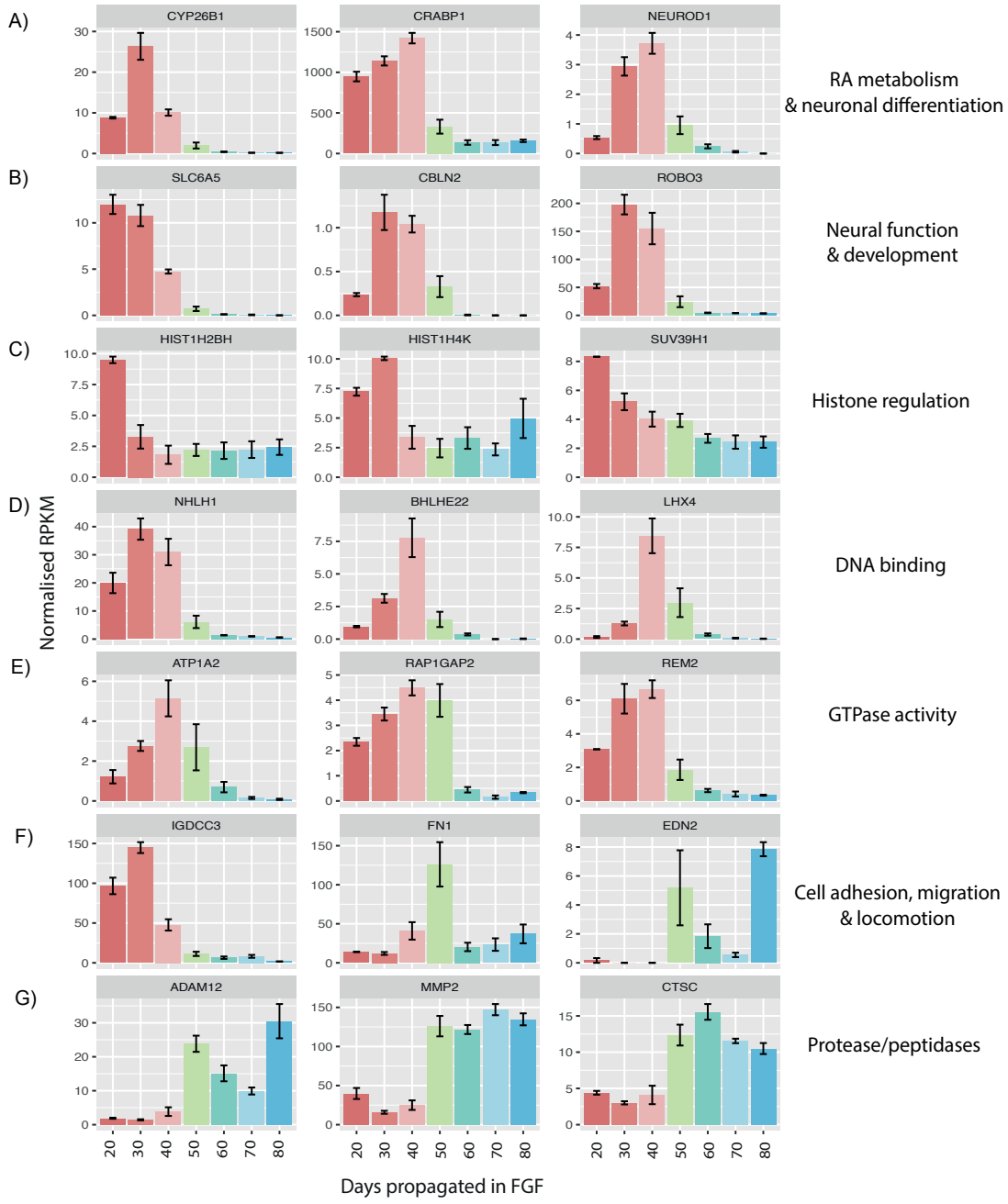


Figure 3.17; mRNA expression of select genes within key pathways involved in the gliogenic switch.

Multiple bar plots to show selected example genes from key pathways highlighted in the gliogenic switch, n= 3+ technical repeats for each timepoint, except D20 n=2. Error bars represent mean ± SEM. The processes highlighted are; A) Retinoic acid (RA) metabolism (CYP26B1 and CRABP1) and neuronal differenitation (NEUROD1). B) Neural function and development (SLC6A5, CBLN2 and ROBO3). C) Histone regulation (HIST1H2BH, HIST1H4K and SUV39H1). D) DNA binding (BHLHE22, NHLH1, LHX4). E) GTPase activity (ATP1A2, RAP1GAP2, REM2). F) Cell adhesion (IGDCC3), migration (FN1) and locomotion (EDN2). G) Protease (CTSC) and peptidases (ADAM12 and MMP2).

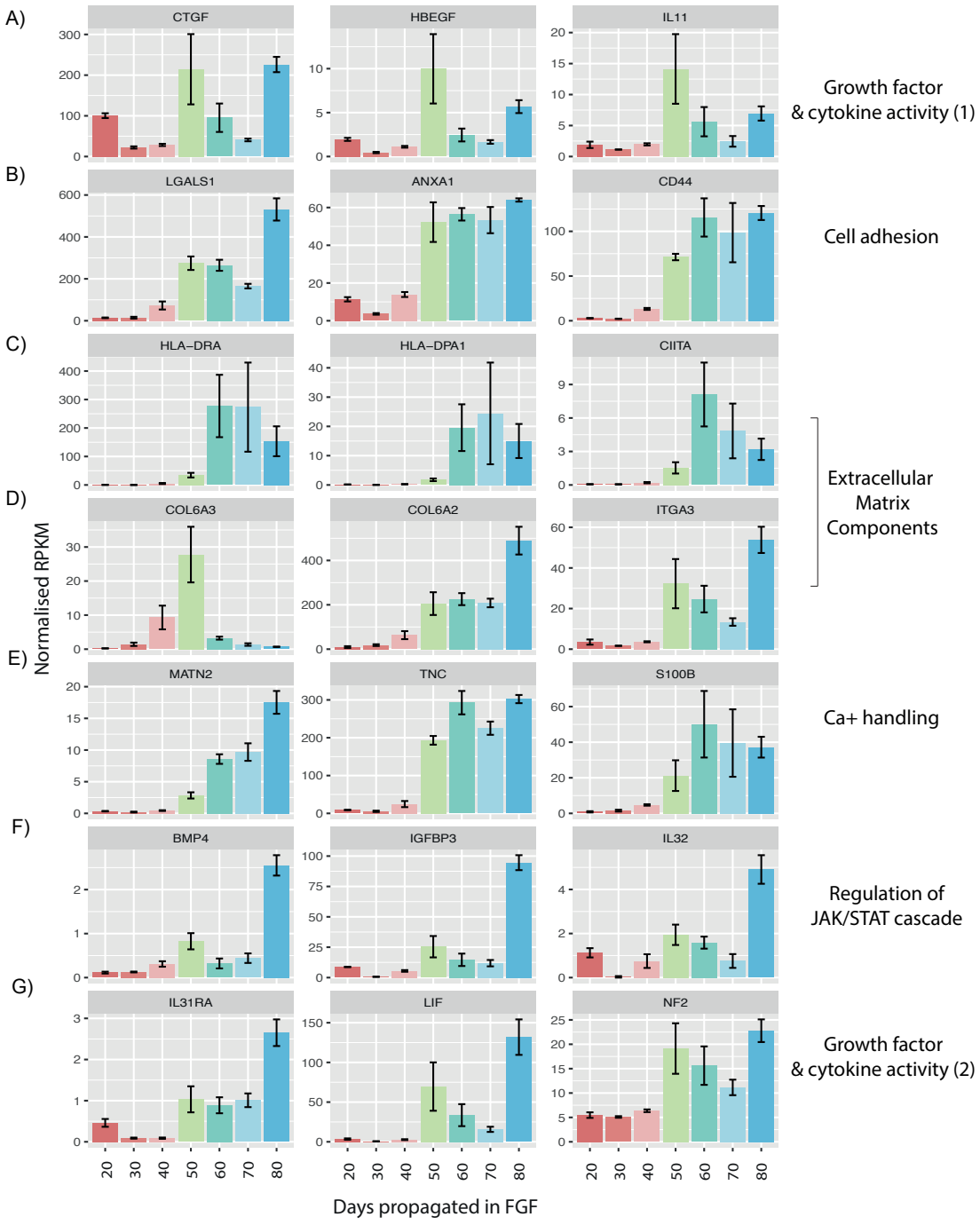


Figure 3.18; mRNA expression of select genes within key pathways involved in the gliogenic switch.

Multiple bar plots to show selected example genes from key pathways highlighted in the gliogenic switch, n= 3+ technical repeats for each timepoint, except D20 n=2. Error bars represent mean ± SEM. The processes highlighted are; A) First induction of growth factor and cytokine activity (CTGF, HBEGF and IL11). B) Cell adhesion (LGALS1, ANXA1 and CD44). C) and D) Extracellular matrix components, including Major histocompatibility complex type II (HLA-DRA, HLA-DPA1, CIITA) and collagen/integrin (COL6A3, COL6A2 and ITGA3). E) Ca+ handling (MATN2, TNC and

S100B). F) Regulation of JAK/STAT cascade (BMP4, IGFBP3 and IL32) and G) A second induction of growth factor and cytokine activity (IL31RA, LIF and NF2).

Overall our approach provides the first in depth temporal transcriptional analysis across the gliogenic switch, which uniquely provides insights into the combinatorial timing of key pathways involved in the precursor fate transition. We observe two distinct phases of neural precursors; neural (NPCs) and Glial competence (GPCs), which is switched at approximately Day50 in our culture system/protocol, but there is some overlap between when neural pathways get switched off and gliogenic pathways on. Interestingly some of the highlighted pathways are observed to play roles at two distinct stages during the gliogenesis, for example growth factor activity was seen to peak at both day 50 and day 80 in our culture system. These findings are consistent with a previous proposed model whereby signals first enable progenitor cells to establish competence for differentiation and then upon a second signal induction progenitors differentiate (e.g. BMP signaling) (Allen 2008). Here I propose the two peaks of growth factor activity we observe control the developmental competence of precursors and subsequently the differentiation of ACs. Further analysis and functional studies are required to define the precise mechanisms that underlie the temporally determined switch between neurogenic and gliogenic precursors.

Lastly, one of the aims from this temporal analysis throughout the gliogenic switch was to guide optimal differentiation strategies for neurons and ACs from hiPSCs. We don't observe the down-regulation of many neural differentiation/development/function genes until Day 40, suggesting that NPCs would be competent to undergo terminal differentiation to MNs until day 40. Further we explored key pathways with act synergistically to promote glial differentiation; neurotrophic cytokines (eg. LIF), JAK/STAT (IL32, LIF and IGFBP3) and BMP signaling (eg. BMP4). We observe that these pathways are up-regulated a while after the gliogenic switch has occurred, at Day 80 in our system. Hence suggesting that the optimal time for terminal AC differentiation would be after 80days of propagation in our culture system.

Comparing the transcriptional profile of hiPSC gliogenic derivatives to primary human fetal and mature purified astrocytes

To date no one has compared the transcriptional signature of fetal and mature human ACs with hiPSC-derived ACs. Therefore here I compare the transcriptional profile of hiPSC-derived GPCs and ACs to human fetal and mature ACs. Evidence, in the form of gene expression and functional characterization to date would suggest that differentiation of ACs from hPSCs only generates immature ACs (Krencik et al. 2011; Roybon et al. 2013), for instance Shaltouki et al in 2013 described the properties and needs of iPSC-derived ACs as similar to fetal ACs (Shaltouki et al. 2013). Therefore we predict that the transcriptional signature of our hiPSC-derived ACs will be closer to fetal immature ACs than to post natal mature ACs.

Here I have compared mRNA expression levels of iPSCs, GPC and AC derived from control iPSCs (n=2 biological lines and technical n=2) to published RNAseq data from human fetal AC, mature AC and to control samples of iPSC-derived MNs, and published data from Cortex, Endothelial, myeloid and oligodendrocytes samples(Zhang et al. 2016). I have used published gene lists for markers of endothelial cells, myeloid cells, oligodendrocytes, fetal and mature ACs (Zhang et al. 2016), in addition to known markers of pluripotency (Pou5f1, Nanog, PODXL, LIN28A & NR5A2), neuronal fate (SYN1, Syt1, VGLUT1, MAP2, NeuN, Tubb3, PSD95 & STMN2), MN specific markers (Islet1, Chat) and additional known genes involved in astrogliogenesis (VIM, MLC1, GFAP, FGFR3). I have plotted the corresponding individual RPKM values on a heatmap to show the relative expression of candidate markers across all samples (Figure 3.19). By plotting individual replicates the diversity/heterogeneity of individual AC populations is highlighted.

I show clear separation of control populations (endothelial, cortex, myeloid, oligodendrocytes, iPSCs and iPSC-derived MN) from all samples from the glial lineage (iPSC-derived GPCs and AC, human Fetal and mature AC). The transcriptional profile between the astrocytic cellular populations also shows separation; firstly mature human ACs are separated from fetal AC and iPSC-derived GPCs and AC by known mature AC markers (including APQ4, ALDH1L1, SLC14A1). iPSC-derived GPCs show up-regulation of all fetal AC markers suggesting a similar developmental status, where as iPSC-derived AC

have only a partial overlap with some fetal markers (TPX2, TNC, VCAM1, PTX3, TNFRSF19, FZR1, ZYX, PXDN) and in addition up-regulate several additional astrocytic markers that are distinct from the mature AC markers (VIM, MLC1, GFAP, FGFR3). This suggests that our iPS-derived ACs are in a transition phase between fetal and mature AC. A previous study in murine brain also finds MCL1 is exclusively detected in glia and an increase in MLC1 expression occurs in the perinatal period (between fetal and postnatal AC phases) (Schmitt et al. 2003), with the highest levels found in gliogenic progenitor populations. This reinforces our iPSC-derived AC still maintain several characteristic of a proliferative immature gliogenic population and have not fully matured. Again suggesting iPSC-derived AC are not fully matured is the expression of FGFR3, which is a specific gliogenic marker of the pMN domain that is up-regulated in proliferative cells and is seen to decrease as GFAP expression increases. This confirms the specification of a spinal cord identity of our iPSC-derived gliogenic derivatives but due to marking only a subtype of AC (pMN domain) and its presence in oligodendrocytes, this marker cannot be used in isolation (Pringle et al. 2003).

Moreover this transcriptional database can also be used to address controversial results in the field, for example the expression of S100B in the gliogenic lineage is so far not clear, many report it as a late mature marker of ACs (Raponi et al. 2007) but others detect its expression in immature and undifferentiated ACs (Shaltouki et al. 2013). Here I see S100B expression is highest in mature ACs suggesting it is in fact a late astrocytic marker, but importantly it cannot be used in isolation to characterize AC populations as it is also expressed in oligodendrocytes.

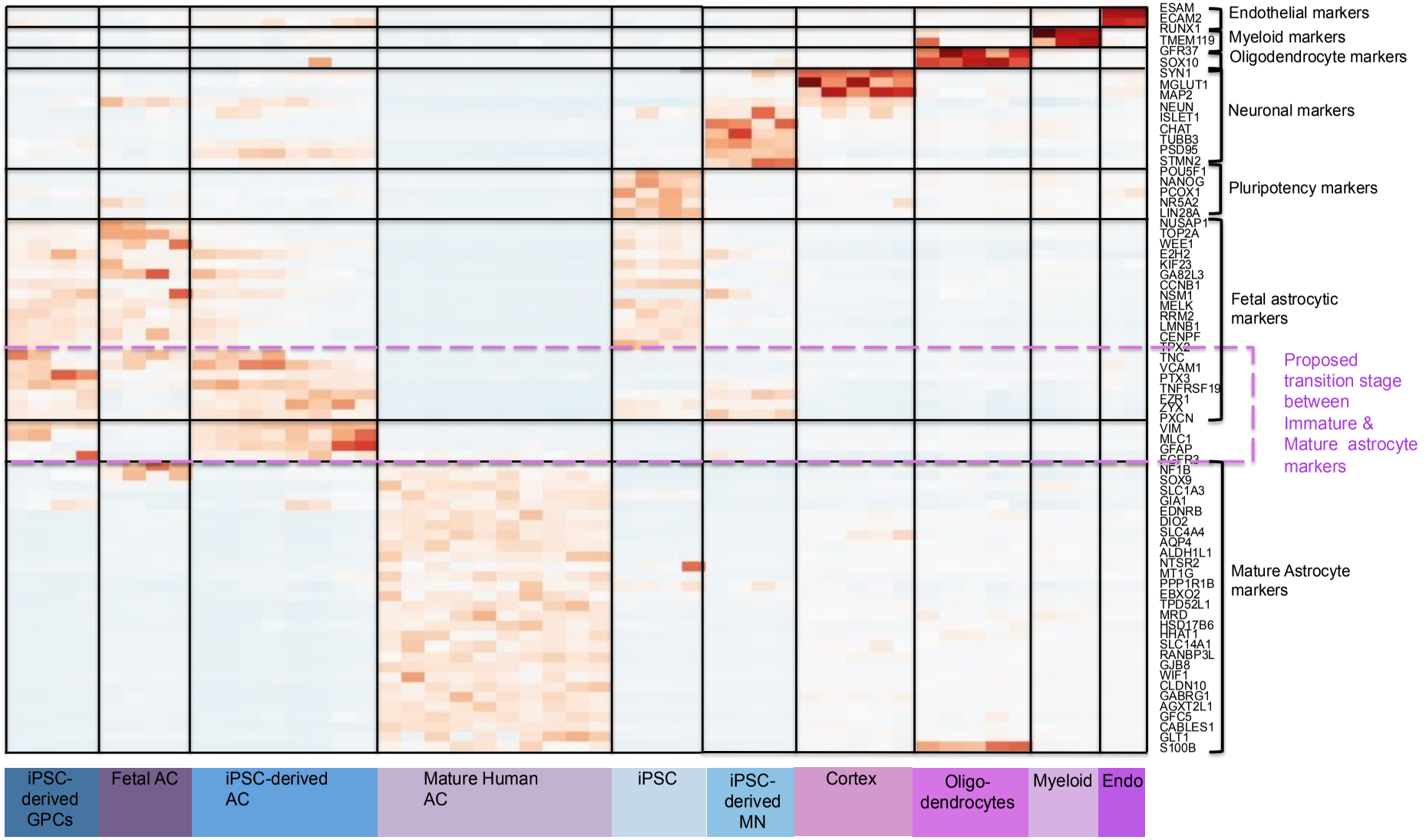

ESAM
ECAM2
RUNX1
TMEM119
GFR37
SOX10
SYN1
MGLUT1
MAP2
NEUN
ISLET1
CHAT
TUBB3
PSD95
STMN2
POU5F1
NANOG
PCOX1
NR5A2
LIN28A
NUSAP1
TOP2A
WEE1
E2H2
KIF23
GA82L3
CCNB1
NSM1
MELK
RRM2
LMNB1
CENPF
TNC
VCAM1
PTX3
TNFRSF19
EZR1
ZYX
PXCN
VIM
MLC1
GFAP
NF1B
SOX9
SLC1A3
GIA1
EDNRB
DIO2
SLC4A4
AQP4
ALDH1L1
NTSR2
MT1G
PPP1R1B
EBXO2
TPD52L1
MRD
HSD17B6
HHAT1
SLC14A1
RANBP3L
GJB8
WIF1
CLDN10
GABRG1
AGXT2L1
GFC5
CABLES1
GLT1
S100B
Endothelial markers
Myeloid markers
Oligodendrocyte markers
Neuronal markers
Pluripotency markers
Fetal astrocytic markers
Proposed transition stage between Immature & Mature astrocyte markers
Mature Astrocyte markers
iPSC-derived GPCs
Fetal AC
iPSC-derived AC
Mature Human AC
iPSC
iPSC-derived MN
Cortex
Oligo-dendrocytes
Myeloid
Endo

Figure 3.19; Comparing the transcriptional signature of hiPSC-derived AC to purified human fetal and mature astrocytes

A) Heatmap to show candidate gene expression of iPSCs and their MN, GPC and AC derivatives, compared to human purified populations of fetal astrocytes, mature astrocytes, oligodendrocytes, cortex, myeloid and endothelial samples. Each block represents 1 technical replicate, with each sample having > n=2. Points represent normalized RPKM values that are log2 transformed and mean-centered.

To conclude these results demonstrate the distinct maturational stages in astrogliogenesis and compare our differentiation strategy in hiPSCs to human development. GPCs show high transcriptional similarity to fetal proliferative ACs, but iPSC-derived AC do not mirror the maturation stage of postnatal mature AC, instead they display a distinct transcriptional signature which we hypothesis is a transitional stage between fetal and post-natal mature ACs.

Overall, I have shown a human model of astrogliogenesis from pluripotent cells, through NPCs followed by GPCs to terminally differentiated ACs. Using highly enriched sample populations I analyzed global gene expression signatures, which build upon previous microarray studies of both murine and human ACs and offer insights into the transcriptional profile of ACs. This database can be used to help identify new genes that are pivotal for astrogliogenesis and could be used to ascertain the purity of AC populations that are differentiated from NSCs. This information will be particularly useful for future studies of human ACs, aiding the large-scale production of homogeneous AC cultures for high throughput drug screens and for understanding neurological disorders resulting from AC dysfunction.

Discussion

The revolutionary discovery of hiPSCs (Takahashi and Yamanaka 2006) has provided an alternative model system to study both development and disease that has remarkable advantages over previous models, including being a fully humanized system and the potential to enable personalized medicine in the future. Here I have;

• Described and functionally validated novel differentiation methods developed for the generation of spinal cord MNs and ACs - that I will later use study the underlying pathogenesis of MND.

• Recapitulated key developmental events in hiPSCs, such as sequential neurogenesis and gliogenesis

• Defined a fine window when the gliogenic switch occurs in our culture system

• Defined a temporal sequence of transcriptional events that occurs as NPCs lose the ability to become neurons and acquire gliogenic potential

• Compared the transcriptional signature of our hiPSC-derived ACs to human samples to assess their developmental stage/maturity.

The generation of spinal cord MNs and ACs

I have built upon the Patani lab's previous published work (Gupta et al. 2012; R Patani et al. 2011) by generating robust directed differentiation strategies for highly enriched, regionally defined and clinically relevant populations of both spinal cord MNs and ACs, in a fully defined, humanized culture system. A major advantage of using a monolayer approach has enabled the consistent generation of highly enriched cellular populations. Importantly, this removes issues of heterologous cell-cell signalling that are frequently observed in suspension culture systems.

Briefly for MN generation I use dual smad inhibition and GSK3B signaling (CHIR99021) to synchronize neural conversion and have altered the timing of patterning with RA and Purmorphamine to increase the generation of spinal neural precursors specified. Using same logic I was able to accelerate terminal differentiation of MNs by treatment with compound E, which promotes synchronized exit from the cell cycle. The same approach for generating spinal neuronal precursors was used for astrogliogenesis, then BMP4 and LIF signaling was utilized to promote terminal differentiation. To summarise the novelty of the protocols I have developed i) they are fully monolayer based, ii) we do not use additional growth factors such as BDNF and GDNF, and iii) we use dorsomorphin instead of LDN during neural induction. All phases of the protocol have been optimized to ensure they are time efficient, generate highly pure populations of cells (which is particularly important for the application/analysis of genomic techniques I use later in this thesis) and that they mirror key developmental processes in vivo.

Separately, I have explored the temporal regulation of early and late NPCs by analyzing the transcriptional signature at selected time points throughout the propagation of NPCs. These findings (discussed more below) guided the

optimization of robust protocols for neurons and glia from hiPSCs. In particular the timing of the gliogenic switch steered the timing of NPC propagation; the maximum time limit for propagation of neurogenic precursors could be defined (40days) and the minimum time NPCs should be propagated before gliogenic precursors can be efficiently terminally differentiated into ACs (60-80days).

Moreover I have comprehensively validated the protocols at transcriptome-wide and functional levels in a stage specific manner. This has provided important insights into the maturity of derived cell types. For functional characterization of derived MNs we have collaboratively carried out several assays in parallel, including electrophysiology profiling, calcium response and axonal transport. It is important to note that although these enable us to investigate neuronal function they have not allowed us to distinguish MN specific functions. For this we will need to carry out co-culture with muscle or inject samples into the developing neuroaxis (eg. of a chicken). However separately, these investigations have resulted in establishing several assays in hiPSC-derived neurons that can later be exploited to address controversies in disease and/or development (although this is not addressed in this thesis). For example both neural synaptic activity and the efficiency of long-range communication caused by defects in axonal transport constitute major pathogenic mechanisms in neurodegenerative disorders (Alami et al. 2014; Devlin et al. 2015; Wainger, Kiskinis, Mellin, Wiskow, Steve, et al. 2014). Deficits in axonal transport have been identified at a pre-symptomatic stage in the SOD1G93A mouse model of MND (Bilsland et al. 2010) but have not yet been followed up in a human clinically relevant system. The early appearance of transport defects in the SOD1G93A mouse suggests axonal transport may play an important role in MND pathogenesis but it is not clear if this is the causative mechanism of MN degeneration or a secondary pathogenic event in MND. This can be elucidated by temporally analyzing hiPSCs throughout differentiation and maturation. Additionally, neural hyperexcitability followed by a progressive loss in synaptic activity has been shown in MND, in particular using hiPSCs carrying mutations in C9ORF72 and *TARDBP* (Devlin et al. 2015). We can now followed up and validate these findings using electrophysiology profiling of MNs derived from additional clinically relevant hiPSC lines.

Finally it is important to highlight there is a huge cellular diversity present in the human motor system, which presents with differential vulnerability in different

neurodegenerative diseases. The majority of iPSC strategies to date, including ours, have focused on deriving MNs generically but comparatively few studies have characterized the motor neuronal subtype diversification (Zirra, Wiethoff, and Patani 2016). I have briefly begun to look at the MN subtype specified by our protocols. HOXB4 staining and RNAseq analysis of the HOX genes confirmed our MNs are specified to the cervical region of the spinal cord. Also FOXP1, a HOX cofactor that also regulates MN subtype diversity, indicates 50% of our cultures are limb innervated MNs of LMC. However much further analysis is needed to fully define the subtypes of MNs derived from this protocol. In the future characterizing and elucidating new efficient protocols for subtype specific MNs will have important implications for more accurate models for studying disease and drug discovery.

Defining the transcriptional landscape of the gliogenic switch and astrogliogenesis

Until recently due to technical limitations, there has been a lack in understanding of genome-wide transcriptional status of iPS cells and their derivatives. In particular the gene expression profiles of ACs have been relatively poorly defined despite their importance in the nervous system and roles in disease. The role of ACs depends upon their maturity, reactivity and regional specificity therefore it is essential we improve our characterization, validation and specification stratergies. Here I have begun to systematically define the transcriptional landscape of astrogliogenesis, which can in the future guide improved protocol development and characterization.

Firstly, I have explored the transcriptional regulation underlying the gliogenic switch by analyzing samples at selected intervals throughout the propagation of NPCs. I have narrowed down the window in which the gliogenic switch occurs in our culture system and begun to dissect the time line of regulatory events underlying the switch. I show a down-regulation of neuronal specific genes occurs (Day40) prior to the up-regulation of astrocytic-specific genes (Day80). Further within the transition period in between these two distinct phases we observe proteases, genes influencing cellular adhesion and migration, selected growth factors and cytokines are all up regulated. However further analysis and functional studies are required to determine the precise mechanisms by which the gliogenic switch is regulated, including dissecting the interaction of both extrinsic and intrinsic mechanisms. Understanding the molecular machinery of

the gliogenic switch will be able to guide the manipulation of relevant cell types (neurons and glia) in culture, enabling us to improve our strategies for generating the respective cell types.

Further I have characterized the equivalent maturity status of our iPSC gliogenic derivatives by comparing their transcriptional signature to human fetal and mature ACs. This is important as the use of hiPSC-derivatives to model late-onset diseases is a wide concern in the field - typically patients do not develop symptoms until later in life, which implicates age is a vital component for disease onset and progression. We observe hiPSC-derived GPCs display a transcriptional signature of comparable maturity to human fetal ACs. Unsurprisingly though hiPSC-derived ACs transcriptional profile did not match that of human mature ACs. Instead they are observed to express some overlap with fetal AC markers (TPX2, TNC, VCAM1, PTX3, TNFRSF19, FZR1, ZYX, PXDN), in addition to several unique AC markers that were not defined in the transcriptional signature of mature ACs (VIM, MLC1, GFAP, FGFR3). This suggests our hiPSC-derived ACs are in a 'transition' phase between fetal and mature ACs. Although hiPSC-derivatives have proved not to be equivalent to mature adult cells, they still hold several advantages over other model systems to study neurodegenerative disease and ideally should be used in conjunction with other model systems (eg. in vivo models and postmortem tissue), to uncover robust disease mechanisms. Importantly hiPSCs overcome critical species differences, mutations are present at pathological doses, in theory any cell type can be generated in unlimited quantities and it is possible to temporally study how diseases evolve and progress.

Chapter 4; Optimizing genomic techniques to study RNA regulation

As discussed in the chapter 1, there is strong evidence to suggest that RNA regulation plays a role in the pathogenesis of MND and I will later explore the contribution of defective RNA regulation in our hiPSC VCP-related MND model system, using RNA sequencing to study transcriptional changes and iCLIP to explore the possibility of altered protein-RNA complexes. However, iPSC technology is both very expensive and time-consuming, so although it can provide a limitless resource of human, patient specific cells, the amount of material available for experiments is restricted. It is therefore crucial to optimize downstream experimental methods first. Here I have comprehensively evaluated multiple RNAseq approaches and iCLIP modifications.

RNA sequencing

There are many different methods to prepare libraries from RNA samples for high-throughput sequencing but all follow a core workflow. Libraries can be created from the total RNA or from a selected subpopulation, for example from mRNA only (to calculate abundance and novel features of the coding transcriptome) or from transcripts of interest (by targeting selective pathways).

Approximately 80% of cellular RNA is ribosomal, which is not translated, instead it functions as a scaffold and/or enzyme of the ribosomes. We chose to remove this abundant RNA before library preparation and deep sequencing to maximize the uniquely mapping reads. There are several methods for removing rRNA before library preparation; the most common are rRNA depletion (which targets selected rRNA molecules for removal but preserves information of non-adenylated, non-coding, and regulatory RNAs) or PolyA selection (capture of all mRNA). Further there are many approaches for RNAseq library preparation but on the whole they follow the same principles; for stranded RNAseq the overall stages are;

1/ Fragmentation – to make RNA optimal size for sequencing

2/ first strand synthesis – uses dNTPs, and Actinomycin D to prevents spurious DNA-dependent synthesis, while allowing RNA-dependent synthesis, improving strand specificity.

3/ second strand synthesis – uses a selective chemical marking (often dUTP) on the second strand

4/ Adenylation and adapter ligation – fragments are adenylated prior to the addition of an adapter. The adapters contain a unique 6 nucleotide sequence/barcode to allow for multiplexing of samples

5/ second strand degradation and PCR amplification - UTP-marked strand is selectively degraded by Uracil-DNA-Glycosylase prior to PCR amplification.

6/ Quantification and sequencing – Libraries are accurately quantified and quality is checked using the bioanalyser before sequencing.

Here, I have tested the preparation of the RNA itself and several cDNA library preparation kits.

RNA preparation

To prepare samples for RNA sequencing, RNA is first extracted then the quantity and quality must be checked. For RNA extraction I compared the Zymo Direct-zol miniprep kit, Qiagen RNeasy extraction kit and the Promega Maxwell automated machine with simplyRNA cells kit. I found that the Zymo Direct-zol gave a consistently higher yield of RNA compared to the Qiagen RNeasy kit (data displayed in Appendix 8.4), however the Promega Maxwell automated machine gave the most consistent RNA yield (data displayed in Appendix 8.4). Samples from each experiment batch were selected at random to run on the bioanalyser to assess RNA quality. For downstream applications only samples with RIN scores >7.5 were used.

Generead rRNA depletion and NEXTflex directional RNAseq approaches

We first decided to deplete rRNA using the qiagen Generead rRNA depletion kit (input of 1-5ug). This works by the hybridization of specific oligonucleotide probes that are complementary to the large (18s, 28s), small (5s, 5.8s), and mitochondrial (12s, 16s) rRNAs. The RNA:DNA hybrid is recognized by an antibody and the resulting antibody-hybrid complex is efficiently captured on a protein G bead. These beads can then be separated from the sample, removing the rRNA in the process. After rRNA depletion, RNA quantity is again checked.

I have tested several approaches for RNAseq library preparation, including unpublished methods produced within the lab and commercial kits. Among others, I trialed the NEXTflex directional RNAseq kit (10-100ng input) (Figure 4.1.A). This kit has two main advantages, firstly it uses random primers used

for first strand synthesis, which reduces bias compared to oligo dT primers and it provides strand specific information (therefore this can be more confidently integrated with iCLIP data later on).

There were many issues I encountered when using the NEXTflex library preparation kit. Firstly I routinely see a strong adapter dimer product (approx. 100 bp), which is stronger when the RNA input is lower (or is lost throughout the library preparation) (Figure 4.1.B). After ligation of the adapter and again after PCR amplification, ampureXP beads were used to purify the library. These work by reversibly binding DNA in the presence of crowding agent - binding is dependent on the concentration of crowding reagent and ratio of beads to DNA, which in turn determines the size of fragments eluted. The NEXTflex protocol recommended using an equal ratio of beads to sample whereas I have tested a higher ratio of beads:product in an attempt to remove the unwanted adapter product. I used a ratio of 4:5 during the clean up step after the adapter ligation to remove the small products (<150bp) and was able to remove the unwanted adapter dimer (Figure 4.1.B). However altering the bead ratios caused the fragment sizes to be larger than that is optimal for sequencing (Figure 4.1.C). Additionally as a result of having low amount of product, we are using a greater cycle number than recommend in PCR amplification (18 versus 15 recommended). Over-amplification can introduce PCR artifacts and is not optimal for sequencing.

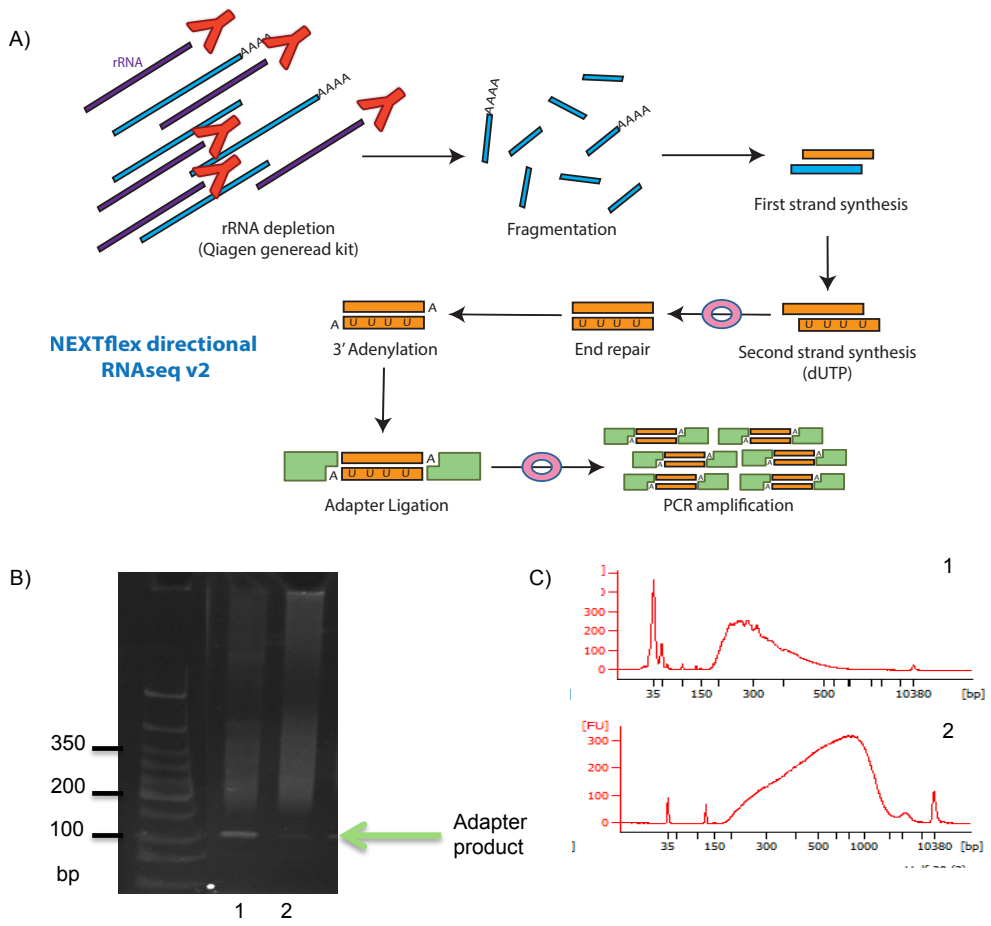


Figure 4.1; Library preparation using rRNA depletion and the NEXTflex RNAseq approach

A) Schematic representation of the NEXTflex RNAseq method. B) Libraries post-PCR amplification. Lane 1 shows a library that was purified post-PCR using the standard NEXTflex protocol. An adapter by-product is highlighted at approximately 100bp. Lane 2 shows a library that was purified post-PCR using an altered approach (decreased ampureXP bead ratio) to remove the adapter by-product. C) Bioanalyser traces of libraries 1 and 2 respectively. Traces show that altering the ampureXP ratio during the final clean up has knock on effect on the library distribution and average fragment sixe.

Once the RNAseq libraries generated using the NEXTflex library preparation kit were sequenced and mapped we used several tests to check the quality of our data. Unfortunately we found further issues with our approach. Firstly there is high proportion of rRNA contamination (Figure 4.2.C) suggesting we need to improve our method of ribosomal RNA removal. I also noticed there was a high percentage of intergenic reads and only a small fraction of exonic (Figure 4.2.B). This is unexpected and suggests there could be a potential DNA contamination in library preparation. In addition we looked at the number of reads found on each the sense and antisense strand (Figure 4.2.A). This again

is unexpected as the reads are mixed between strands (ranging from <5% to >45% on each sense/antisense strand). Finally, looking across all samples (either in the same experiment, same processing batch or taking the total samples in database) I see a huge variability in the quality, mainly in the % of exonic reads and the strandedness.

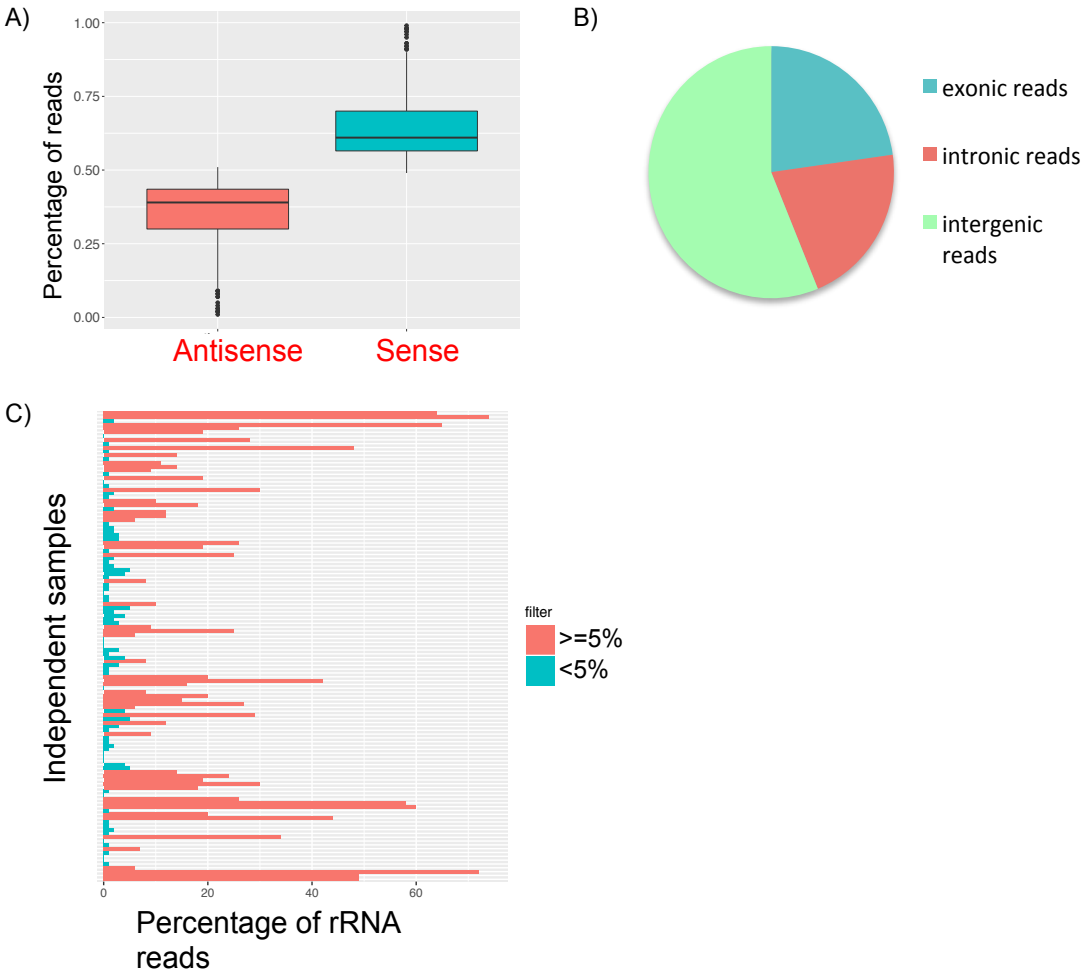


Figure 4.2; Post-sequencing quality checks of libraries produced using the NEXTflex RNAseq method

A) Box plot representation of the average percentage of reads assigned to the sense and antisense strand (taken as an average across 123 samples). B) Pie chart displaying the percentage of reads aligned to exonic, intronic or intergenic regions (numbers calculated as an average across 123 samples). C) Plot displaying percentage of reads aligning to rRNA, Individual samples/replicates are plotted.

All together as a result, we would need to carry out very deep sequencing to get a suitable number of reads for downstream analysis. RNAseq is an important tool for my projects therefore I needed to ensure the method is

reliable, reproducible and high quality. After consideration I did not think this approach was worth pursuing so I turned to exploring alternative methods/kits.

PolyA selection and Illumina Truseq methods

Since Illumina is a leader in sequencing technologies I chose to test out 2 of their RNA sequencing kits; Truseq stranded mRNA and the Truseq access kit (Figure 4.3.A). The two approaches are suited to generating libraries from different amounts of input material (Truseq stranded mRNA requires an input of 0.1-4ug total RNA and the Truseq access kit required 10ng total RNA). The Truseq mRNA kit uses PolyA selection as a method to remove abundant rRNA. This is done using poly-T oligo attached magnetic beads to select mRNA in two rounds of purification. On the other hand the Access kit uses the total RNA to create a template library, from which the coding regions are captured and sequenced using probes to the known coding regions.

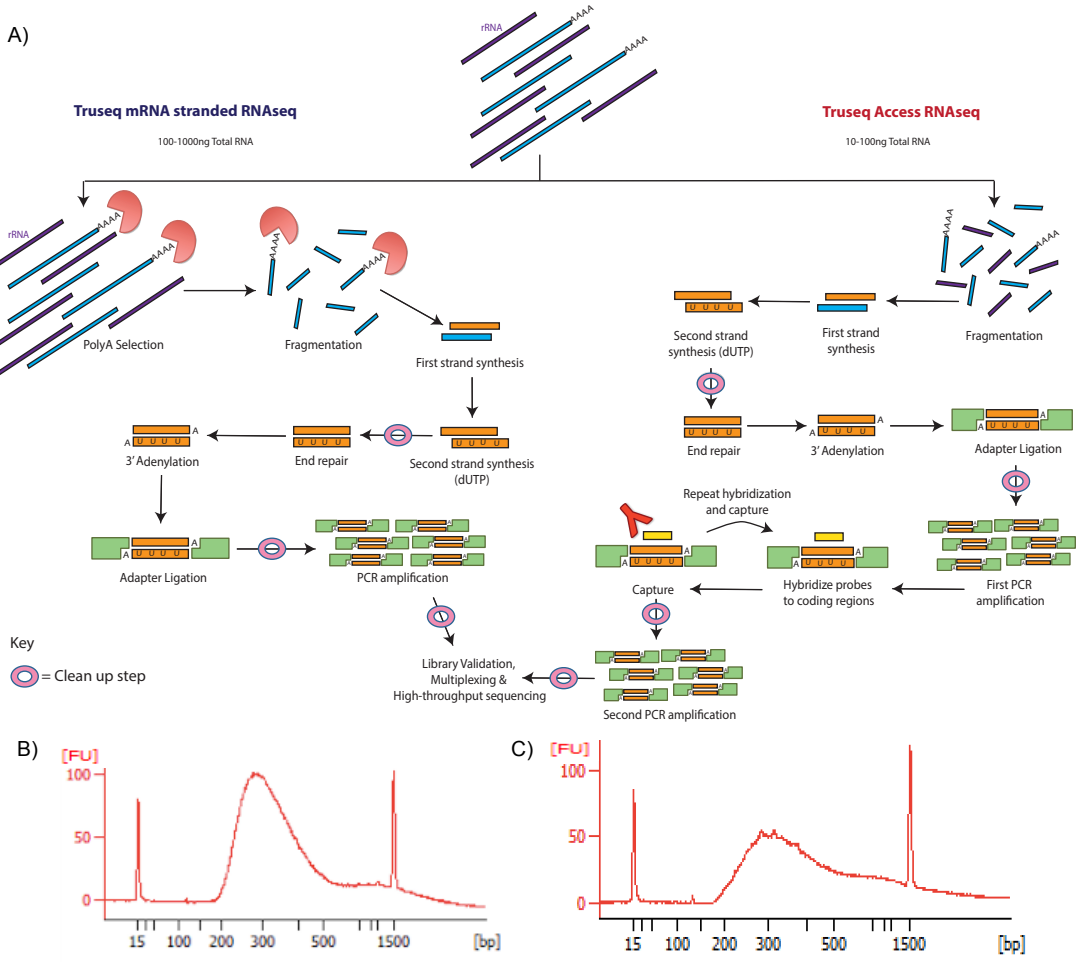


Figure 4.3; Library preparation using PolyA selection and Illumina Truseq RNAseq methods

A) Schematic representation of both the Truseq stranded mRNA and Truseq Access approaches. Representative bioanalyser traces of libraries produced using B) Truseq mRNA and C) Truseq Access kits respectively.

I have tested and produced successful libraries from both Illumina approaches. Both the Truseq mRNA and Access kits were shown to generate libraries of a significantly higher quality than I previously saw with the NEXTflex approach. Firstly during library preparation I see a consistent normal distribution of fragments around 300bp (Figure 4.3.B & C). After sequencing and mapping to hg19 the quality of reads had dramatically improved; the % of strandedness was >90%, there was a high proportion of exonic read >75% and there was no contamination of rRNA (Figure 4.4.A, B & C).

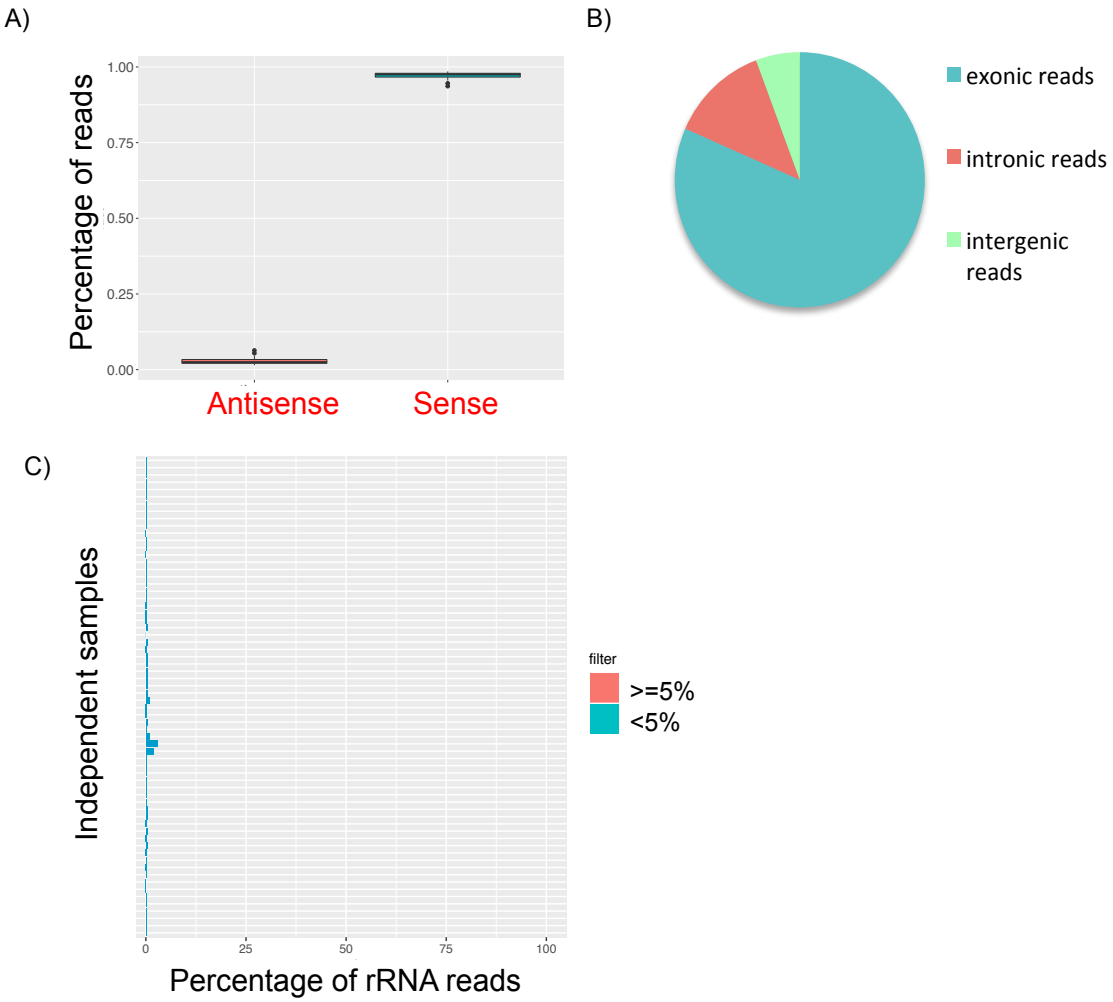


Figure 4.4; Post-sequencing quality checks of libraries produced using the Truseq stranded mRNA RNAseq method

A) Box plot representation of the average percentage of reads assigned to the sense and antisense strand (taken as an average across 66 samples). B) Pie chart displaying

the percentage of reads aligned to exonic, intronic or intergenic regions (numbers calculated as an average across 66 samples). C) Plot displaying percentage of reads aligning to rRNA, Individual samples/replicates are plotted.

For this thesis the Truseq stranded mRNA kit has been used the standard method. A few exceptions that have been generated using the Illumina access kit (where there is very limited input material available) and NEXTflex kit are clearly marked.

Individual nucleotide crosslinking and immunoprecipitation (iCLIP) and non-radioactive modifications

Individual nucleotide crosslinking and immunoprecipitation (iCLIP) has been a powerful method to study protein-RNA interactions. It allows us to study post-transcriptional regulation controlled by RBPs and uncover transcriptome wide maps of the RNA binding sites of an RBP at nucleotide resolution. However the technical aspects of iCLIP remain challenging and are one of the major limitations to its widespread use. A common aspect to all varieties of CLIP is the use of phosphorus-32 (P32) to label the protein-RNA complexes, which restricts application to institutions in which radioactive work is practicable. This was previously the case with northern blotting and in-situ hybridisations too (Cassidy and Jones 2014; S. W. Kim et al. 2010) .

The current iCLIP protocol utilizes antibodies immobilized on magnetic beads to separate the RBP-of-interest from sample lysates. Next an on-bead 3' linker ligation is then undertaken prior to SDS-PAGE analysis (Huppertz et al. 2014). This is also similar for HITS CLIP (Moore et al. 2014), whilst in PAR-CLIP the adapter sequences required for PCR amplification are incorporated after SDS-PAGE analysis (Hafner et al. 2010). Irrespective of the linker ligation, complexes are labeled with P32 prior to SDS-PAGE in all approaches. This way the protein-RNA complexes will be separated according to molecular weight and can be visualized. The resulting autoradiograms can subsequently be used to stringently isolate only the appropriate complexes associated with the RBP-of-interest. Moreover, in theory SDS-PAGE has an additional benefit that it allows ready removal of free RNA that may be stick to the beads during immuno-precipitation, and could contaminate the resulting cDNA libraries, although this has not been experimentally evaluated. Labeled

samples proceed through multiple steps until the radioactive signal is eliminated, and this requires P32-designated apparatus until such steps are reached (Huppertz et al. 2014). Additionally, work with P32 has its own inherent risks and environmental impacts.

Here I have investigated the application of three non-radioactive adaptations to the iCLIP protocol under the supervision of Dr. Chris Sibley. We use the previously CLIP studied heterogeneous nuclear ribonucleoprotein C (hnRNP C) as a model RBP to evaluate these non-radioactive approaches to circumvent the radioactive step during the current iCLIP protocol. The standard iCLIP procedure was followed except a few alterations after immuno-purification and on-bead linker ligation (Figure 4.5).

1/ 'Blind-cutting' of protein-RNA complexes after SDS-PAGE separation.
Previously several RBPs have been studied within the lab such as TIA1, TDP43, FUS and hnRNPC where autoradiographs have confirmed that single RBPs, or dimers of these RBPs, are immuno-purified specifically with no contaminating RBPs (König et al. 2011; Rogelj et al. 2012; Tollervey et al. 2011; Ule et al. 2010). Under such scenarios we hypothesise that it may be possible to purify RBPs in the absence of an autoradiograph through size selection based on molecular weight marker alone. Here no P32-labelling of RNA was undertaken and samples were subjected to SDS-PAGE. No autoradiograph was produced and instead the protein-RNA complexes were cut from the nitrocellulose membrane by using the protein molecular weight ladders as guides for both the lane positioning and for the expected size of hnRNP C (Blind cut).

2/ Elimination of SDS-PAGE analysis altogether – direct Proteinase K (PK) treatment.
If free RNA stuck to the beads is minimual at the point of SDS-PAGE, then we hypothesized purification may be possible to in the absence of SDS-PAGE through proteinase-K digestion of the immuno-precipitated complexes directly off the magnetic beads. Here there was no P32-labelling of RNA, no SDS page and samples were subjected to on-bead proteinase-K digestion and subsequent RNA extraction and reverse transcription as usual.

3/ The use of an Atto663-labelled linker instead of P32 labelling.

Finally we have tested the use of a non-radioactive adapter, so the protein-RNA complexes can still be visualized. The non-radioactive adapter selected needs to be highly sensitive and detect low yields since UV cross-linking is expected to cross-link <1% of the total RBP-RNA contacts under the conditions employed in iCLIP (Ule et al. 2003). Further, it would need to not interfere with the reverse transcription reaction, which uses the adapter as a reverse transcription primer binding site. An atto633-labelled adapter was selected to test, with fluorescence properties 633nm excitation and 657nm emission.

Instead of the standard on-bead ligation of iCLIP linker, the Atto633-labelled linker was ligated, enabling visualisation of protein-RNA complexes using a phosphoimager (Fuji FLA-2000) rather than autoradiograph. The resulting images were used to produce a cutting-mask and guide the isolation of protein-RNA complexes of interest. To aid size determination a near-infrared ladder was used. The iCLIP procedure was rejoined after isolation of the protein-RNA complexes from the membrane.

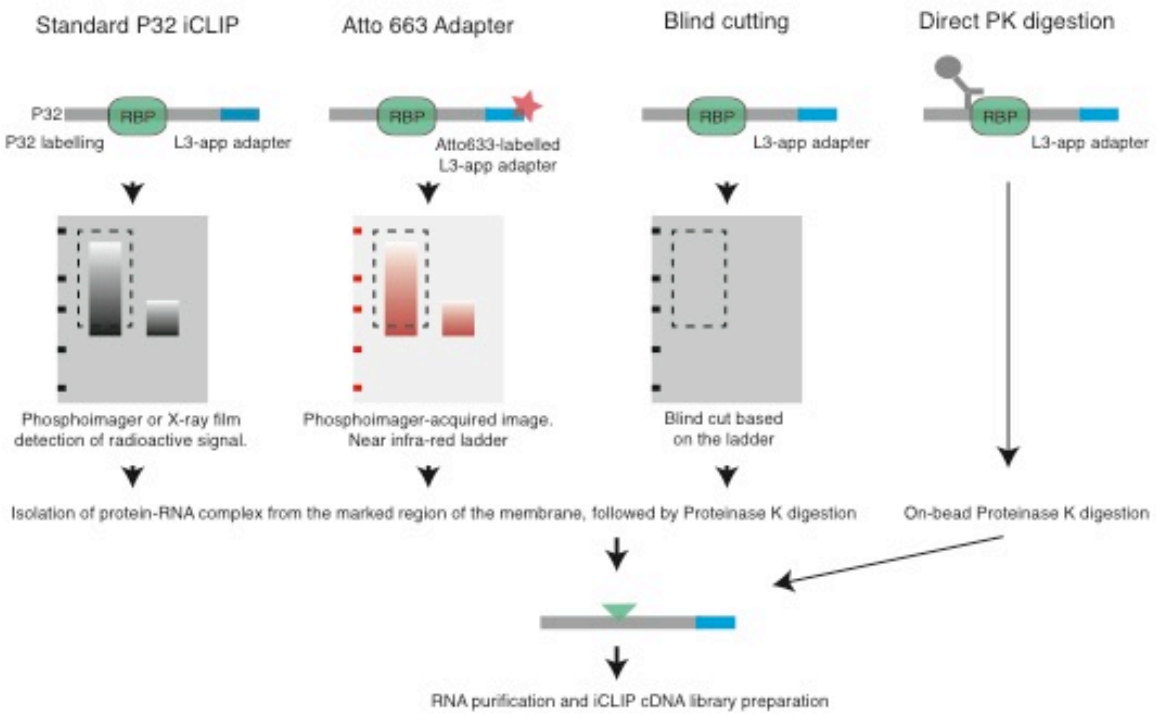


Figure 4.5; Schematic to show 3 alternative non-radioactive iCLIP approaches tested

Comparing the non-radioactive adaptions

Firstly I compared the Atto633-labelled linker and p32 visualisation during library preparation. I found that the Atto633 linker retains the ability to monitor RBP-RNA complexes and the visualisation, on both the autoradiograph and phosphoimage, importantly confirms the RBP contaminant-free nature of the hnRNPC immuno-precipitations (Figure 4.6.A & B).

Next I show the production of cDNA libraries from all 3 alternatives methods and using the standard P32 labelling at similar levels of PCR amplification, demonstrating that all adaptations to the iCLIP method are capable of producing high-throughput sequencing compatible libraries (Figure 4.6.C). Importantly we see that the Atto633 linker did not interfere with downstream preparation of the libraries nor the library complexity.

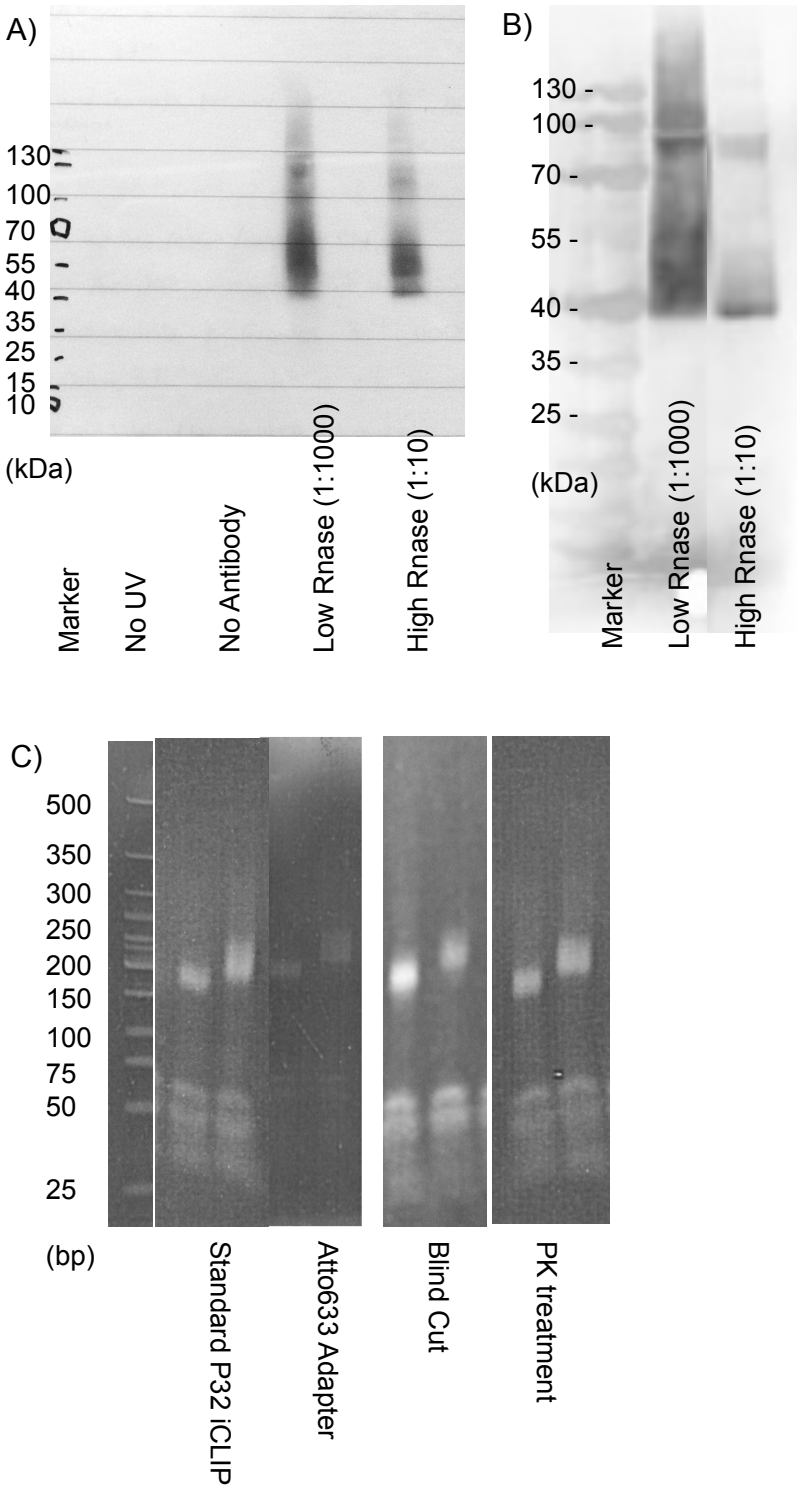


Figure 4.6; Library preparation using the standard iCLIP and 3 alternative non-radioactive iCLIP approaches

A) Autoradiograph after 1hour exposure showing RNA-protein complexes isolated after immunoprecipitation of hnRNPC and RNase treatment following the standard iCLIP method. 2 control samples (prepared using either no UV or no antibody) and 1 sample treated with high RNase are shown. B) Fluro-image of protein-RNA complexes ligated

to a modified adapter containing a flurophore, Atto663. Image is read immediately using a Fuji FLA-2000 phosphoimager. C) Post-PCR amplification of medium and high cDNA fragments prepared from standard iCLIP, Atto633, Blind cut and direct proteinase K approaches.

Following, I have assessed the cross-link clusters (position of each cluster and amount of binding within each cluster) for hnRNPC. Comparison of the cross-link cluster reproducibility of these libraries revealed that 2 of the proposed alternative non-radioactive iCLIP methods (the Atto633 adapter and Blind cutting) had a good correlation to the standard iCLIP protocol and to a previously published hnRNPC library (König et al. 2011) (Figure 4.7). However this did not hold true for the direct PK treatment. We observed that the direct PK digestion approach had the poorest correlation to the standard iCLIP approach (R^2 =0.65 and compared to published iCLIP data R^2 =0.52) (Figure 4.7), a finding we attribute to contaminating free RNA that is not appropriately removed with bead wash-steps alone. Thus, whilst it may be possible to produce iCLIP libraries using this approach, a compromise is made to the quality and resolution of the data. The blind cut based on molecular weight markers produced highly reproducible libraries relative to the standard iCLIP approach (R^2=0.89), demonstrating this approach could be considered when there is no alternative available, however it still remains preferable to visualize RBP-RNA complexes prior purification such that evaluation of other RBP contaminants and the RNase digestion patterns can be appropriately assessed. Finally the Atto633-adapter, produced highly reproducible clusters that mirrored those produced with the standard iCLIP procedure (R^2 =0.84) and to previously published data (R^2 =0.76). Moreover, this approach retains the ability to monitor protein-RNA complex migration patterns during SDS-PAGE, and can be easily adapted with alternative fluorophores to suit institutional capabilities.

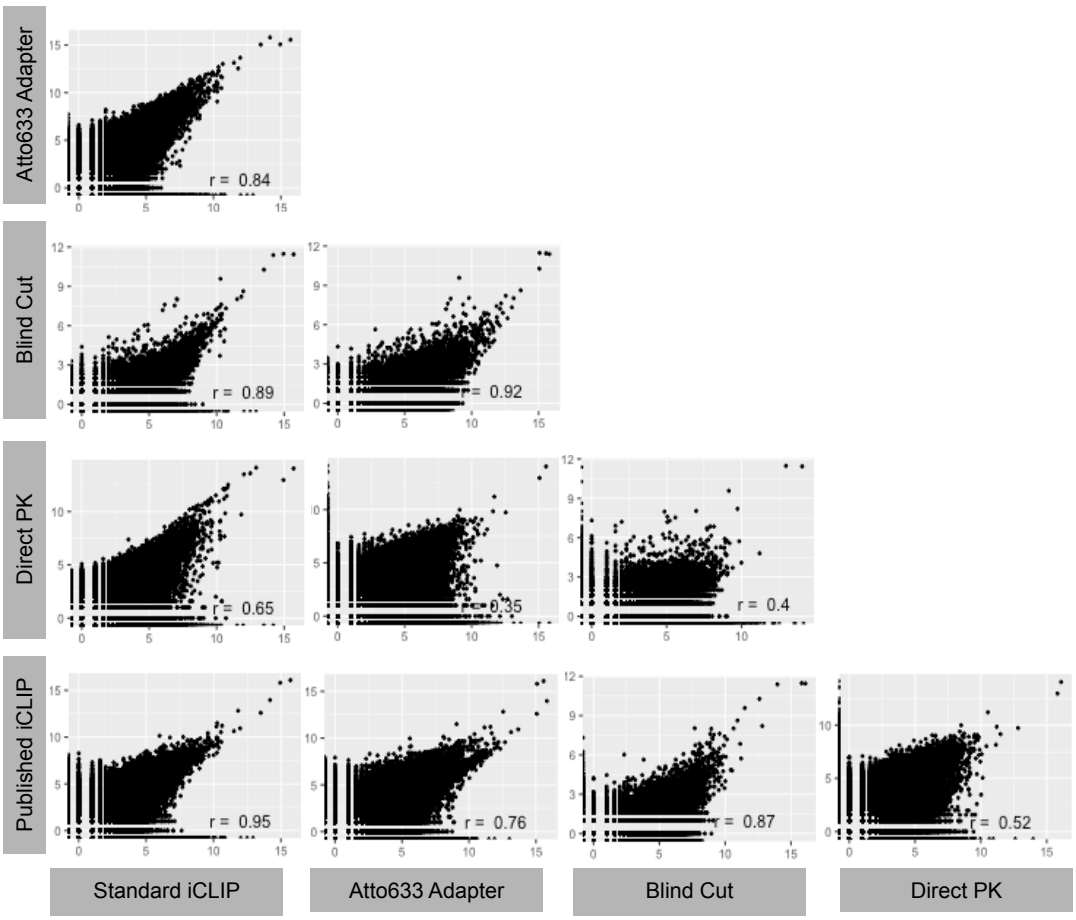


Figure 4.7; Reproducibility of cross-link clusters

Multiple scatterplots display the correlation between hnRNPC cross-link clusters (positions and size) identified using different iCLIP approaches; standard iCLIP, Atto633 adapter, Blind cutting, Direct PK treatment and also to published iCLIP data (Konig et al, 2010). R^2 values shown.

To demonstrate the applicability of this method, we have replicated previously published analysis of hnRNPC that requires high integrity data (König et al. 2011). hnRNPC has previously been characterized to bind to uridine tracts (König et al. 2011). Thus I have sought to test the reproducibility of this by analyzing the sequences immediately surrounding the crosslink site. First using weblogo I have plotted the proportion of each nucleotide located immediately surrounding crosslink site (in a 20nt window) (Figure 4.8.A). Both the Atto633 adapter and blind cut approaches shows T base enrichment similar to we observe with standard iCLIP method and in previously published hnRNPC iCLIP data. Conversely the direct PK treatment method failed to produce a striking T enrichment around hnRNPC's binding sites. Secondly I have performed further sequence analysis by analyzing pentamer enrichments

immediately surrounding hnRNPC's crosslink sites (again in a 20nt region). The top enriched pentamer for all iCLIP approaches is TTTTT (Figure 4.8.B). As before there is high correlation of the pentamer sequence enrichment between the standard iCLIP and Atto633-adapter approach (R^2=0.97), but correlations are relatively poor between direct PK treatment and Blind cut approaches compared to the standard iCLIP approach (R^2=0.18 and R^2= 0.36 respectively) (Figure 4.8.B). The standard iCLIP does not correlate that highly to previously published iCLIP data (R^2=0.34). I hypothesize this is due to increase in library complexity with the latest standard iCLIP data due to improvements in the method over past 5years. Perhaps surprisingly, the pentamer scores generated from the blind cut approach are highly correlated to previously published data (R^2=0.97). I reason that this is due to a reduced library complexity of both the blindcut and previous iCLIP data compared to the standard iCLIP protocol we are currently using.

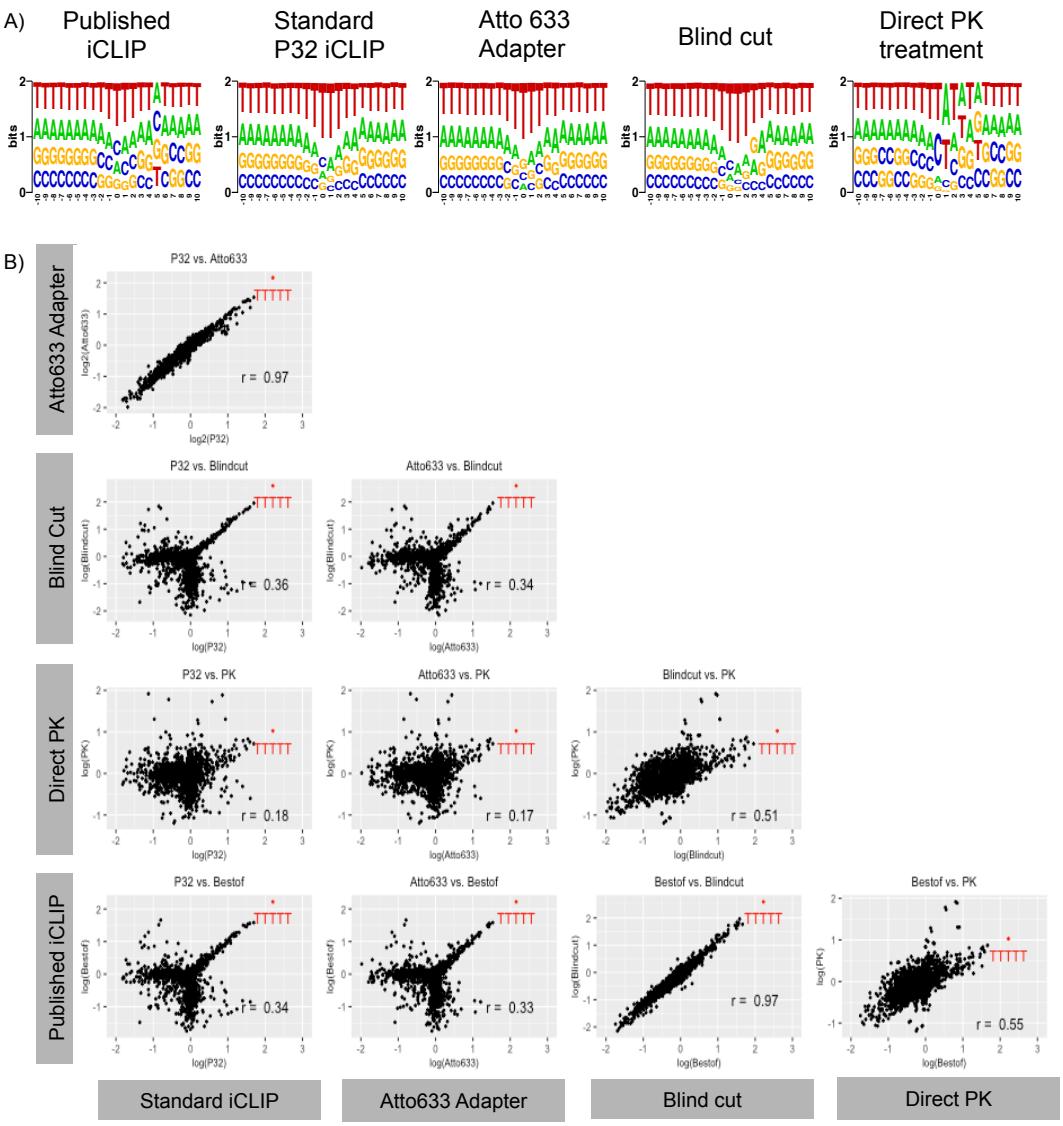


Figure 4.8; Sequence analysis surrounding hnRNPC binding sites

A) Weblogo plots showing the relative nucleotide frequency around the binding sites (position 0) of hnRNP-C. B) Scatterplot plots display the correlation of pentamer enrichment seen within a 20nt window surounding hnRNPC's binding sites, identified using multiple iCLIP approaches; standard iCLIP, Atto633 adapter, Blind cutting, Direct PK treatment and also to published iCLIP data (Konig et al, 2010). R^2 values shown and highest enriched pentamer highlighted in red.

Finally we assessed cross-link positioning for a previously studied gene, CD55 (König et al. 2011). We see crosslink nucleotides identified using the standard iCLIP, non-radioactive Atto633-adapter and Blind cut approaches are comparable to previously published data; they are present along the entire length of CD55 pre-mRNA and accumulate around the alternative exon (Figure 4.9.A, B & C). Direct PK treatment does not reproduce this crosslinking pattern and shows non-specific binding, therefore in agreement with the previous

correlation analysis we conclude direct PK treatment results in inaccurate, compromised data. Further we note, as in previously published data, hnRNPC binds within Alu elements (Zarnack et al. 2013). We show the % of binding to Alu elements is comparable for all alternative approaches and to previously published data (ranging between 24.4% using the Blind cut method and 14.2% using the non-radioactive adapter. The average across methods is 18.1%±3.6) (Figure 4.9.D).

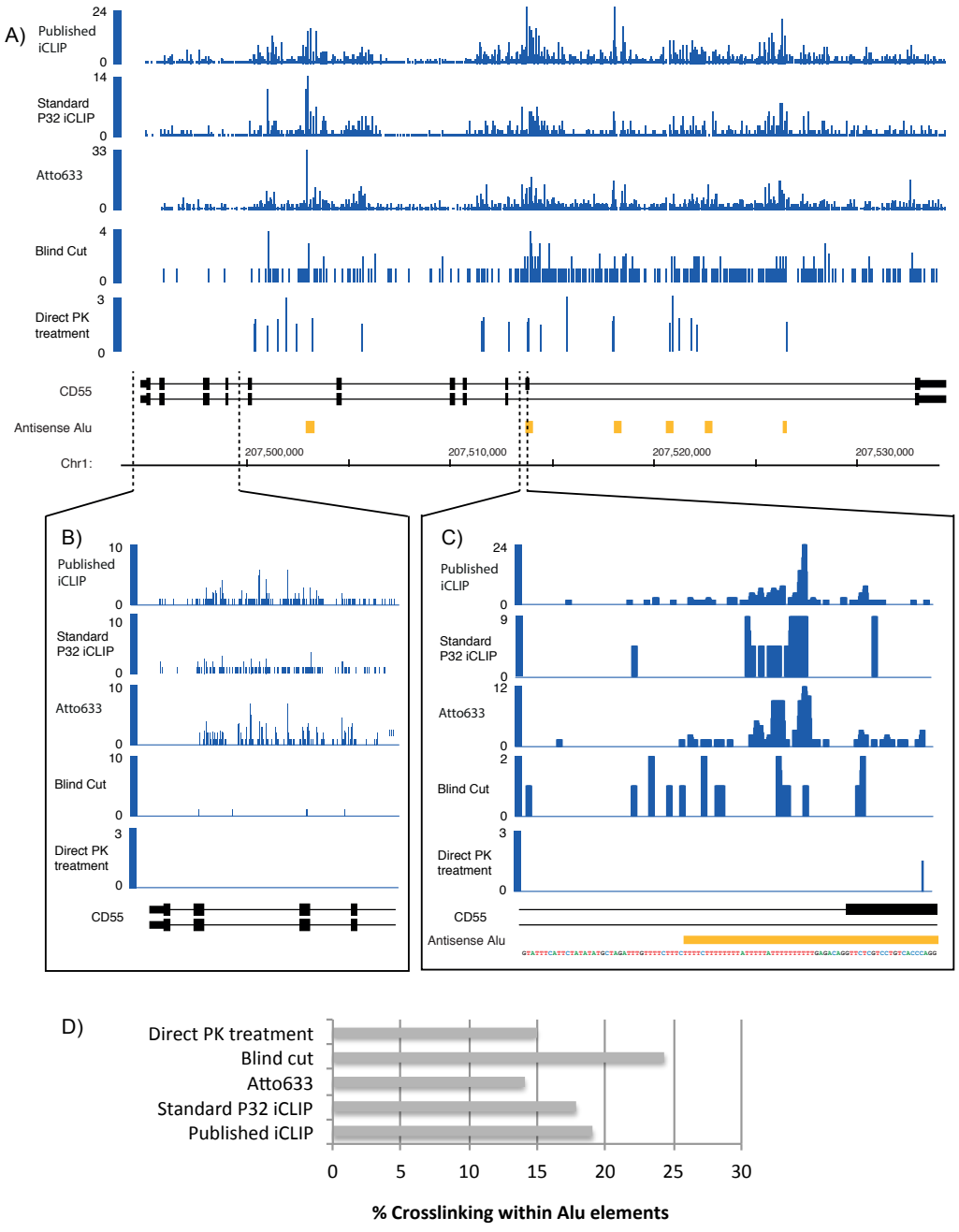


Figure 4.9; Reproducing hnRNPC crosslinking to CD55

A) Global view of hnRNPC cross-linking to CD55; crosslinking nucleotides are present along the entire length of CD55 pre-mRNA. B) Zoomed in crosslinking positions, C) crosslinking accumulation around the alternative exon and within Alu element. D) % binding to Alu elements.

To conclude, the Atto633-labelled linker and blind cut approaches reproduced cross-link positioning and the subsequent findings of previous radiolabelled libraries (binding to U tracts and alu elements). However the Atto633 adapter will be the preferred approach to use because it has the advantage over the Blind cut approach of visualising the protein-RNA complexes.

Discussion

The complexity of gene expression and regulation demands the highest sensitivity for measuring transcripts and detecting changes in abundance and structure. Since the advent of NGS, the use of unbias transcriptome-wide analysis has become increasingly accessible and even over the past decade there has been substantial progress in genomics technologies (including in library construction, sequencing and data handling/analysis). This is reviewed elsewhere (König et al. 2012; Metzker 2010). Here we have evaluated several RNA sequencing and non-radioactive iCLIP approaches to ensure we select the best methods for future projects.

The vast number of recently published RNAseq studies highlights it as the preferred tool for the deepest levels of transcriptomic investigation. A simplified data analysis workflow, coupled with improvements in throughput and cost, makes RNAseq an accessible and affordable solution for transcriptome research. Here I have tested various RNAseq methods side by side to refine a highly reproducible, reliable approach for future studies. The first RNAseq approach trialed was the Qiagen rRNA depletion followed by NEXTflex directional RNAseq kit. This unfortunately proved problematic at various stages; it did not work efficiently with low input material and resulted in high percentage of rRNA and intergenic reads and unclear strand distribution. On the other hand both the Truseq RNAseq approaches used showed consistently good quality data; libraries were composed of optimal size fragments for sequencing, they had the ability to create libraries from low input material, the results generated had a low percentage of rRNA reads, normal library distribution, and high proportion of strandedness and exonic reads. Importantly

we have also set up an analysis pipeline for handling large quantities of RNAseq data systematically and efficiently (see Methods).

Moreover iCLIP has provided a crucial tool for research in protein-RNA interactions, but the wider application has been limited by the need for radiolabelling in order to carefully select protein-RNA complexes of interest. Here I have contributed to the lab's published iCLIP work by developing and comparing several alternative non-radioactive iCLIP approaches. The use of an Atto633-labelled linker has proved successful in reproducing crosslink positioning of the previously characterized RBP, hnRNPC, while also preserving the capability to visualize protein-RNA interactions. hnRNPC is a highly abundant protein therefore further experiments are required to fully evaluate the sensitivity of Atto663 linker, for example testing other less abundant proteins that usually require longer exposure to radioactive isotope in the standard iCLIP approach.

In addition, testing a 'Blind cut' alternative to iCLIP has been of paramount importance in context of recent expansion in ENCODE data. ENCODE is a collaborative initiative to comprehensively annotate all functional sequences in the human genome and provides the data generated in an open resource. One arm of the project collects information on Protein-RNA interactions, which is generated using the eCLIP method (Diehl and Boyle 2016). eCLIP methodology, developed only a couple of months ago, bases its core principles on standard iCLIP approach but utilizes two novel adapters to circumvent an inefficient circularization step (Van Nostrand et al. 2016). This maintained nucleotide resolution of data and resulted in a reduced amplification rate and therefore reduced duplicate reads, an enhanced success rate and a protocol suitable for large-scale use. Nonetheless eCLIP abolishes radiolabelling at the expense of being able to visualize protein-RNA complexes and proceeds with a 'Blind cutting' approach to isolate the protein-RNA complexes of interest. To date it has not be evaluated how this effects quality and complexity of the resulting libraries generated. Here we show that Blind cutting is able to produce high quality data, comparable to the standard and non-radioactive approaches tested side-by-side however this has only been explored for a widely studied protein (hnRNPC) which we know can be cleanly immunoprecipitated. Further comparative tests should be carried out using a variety of RBPs to fully assess the wide application of this approach.

Earlier this year another study also presented a non-radioactive alternative iCLIP approach, irCLIP (Zarnegar et al. 2016). They reported the use of an infra-red dye conjugated and biotinylated ligation adapter in replacement of the iCLIP adapter and P32 labeling. This had comparable ligation efficiency, the infra-red imaging was highly sensitive and reduced the visualization time 10-100 fold. In addition the same study further optimized each stage to produce a simplified protocol requiring far less input material than before. Together, the removal of radio-labeling and the enhanced efficiency of methodology will enable CLIP based technology to be used more widely throughout the scientific community. Given its sensitivity it will facilitate the detection of non-canonical RBPs that are weakly expressed or only bind in limited conditions, which other CLIP methods have struggled to do to date.

To conclude we have established robust, effective and consistent functional genomic methodologies that firstly I will utilize to later address the hypothesis that RNA regulation is defective in VCP-related MND, but also will be of value to the wider scientific community.

Chapter 5; Characterizing the earliest pathogenic events in VCP related MND

So far research on VCP and associated diseases has been done using overexpression, KO or KD models in non-human and/or non-neuronal cells (Buchan et al. 2013; Ju et al. 2009; Rodriguez-Ortiz et al. 2013; Yin et al. 2012). While both animal and cell-based models have provided invaluable insights into MND pathogenesis, such model systems may not precisely capture the human clinical pathophysiological state. Consequently, there is a need for accurate characterization of how VCP mutations affect human motor neurons, in the context of MND. Increasing recognition of glial involvement in MND, through either cell autonomous or non-cell autonomous mechanisms, also raises the question of their contribution in the context of VCP mutations. To begin to address these issues in VCP-related MND research, we employed patient-specific iPSCs and directed differentiation strategies to both spinal cord MNs and ACs. This approach accurately approximates human pathophysiology while bypassing the need for artificial overexpression, knock down or knock out studies.

Aims

Here I will use patient specific VCP-mutant hiPSC and their neural derivatives, to characterize the presence of a range of phenotypes and when these phenotypes arise during the differentiation and maturation of both spinal cord motor neurons (MNs) and astrocytes (ACs). The phenotypes investigated have been selected based on previous studies of VCP'opathies, neurodegeneration and in particular on MND. More specifically I ask;

1. Do VCP neural derivatives show selective vulnerability?
2. Using RNAseq, which transcriptional pathways discriminate VCP mutant from Control?
3. Do we detect selected phenotypes in our VCP model that have previously been described in MND? In particular:
    - Mitochondria dysfunction
    - ER stress
    - Oxidative stress
    - Synaptic defects

4. When is the earliest pathogenic event detected in our hiPSC model of VCP-MND?
5. What is the sequence of events in MNs and ACs? Do gene expression changes arise before cytoplasmic events?
6. Are early events different in MNs and ACs?

The experimental plan

To explore the pathogenic events in our VCP-related MND model, I here considered the role of both ribostasis, using RNA sequencing and cytoplasmic organellar dysfunctional, using an array of imaging technologies. In order to assess the earliest pathogenic events detected I analyzed control and VCP-mutant hiPSC neural derivatives throughout differentiation (Figure 5.1). After pluripotency I selected 4 stages during differentiation and maturation of motor neurons; neutrally inducted precursors (NI), patterned neural precursors (NPCs), early motor neurons (d3 MN) and mature motor neurons (d17 MN), and also 3 stages during astrogliogenesis; Gliogenic precursors (GPC's), 2 and 4 weeks after terminal astrocyte differentiation (d14 or d28 AC).

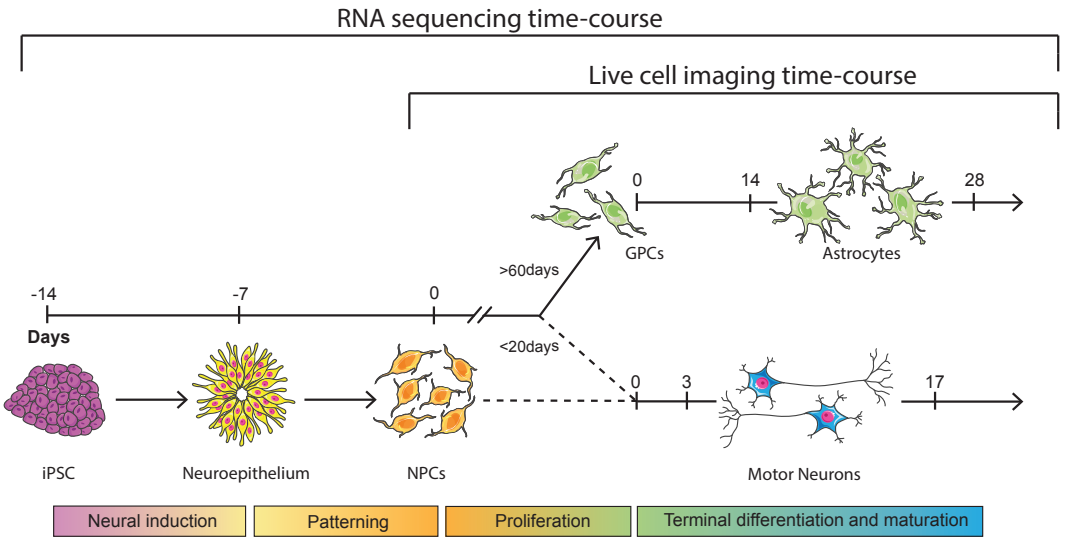


Figure 5.1; Experimental time-course schema

All VCP hiPSC lines and one control line were provided through collaboration with Dr. Selina Wray; she used established reprogramming methods to generated iPSCs from two patients with the VCP mutations R191Q (2 clones) and R155H (2 clones) and from one healthy control. iPSC clones all expressed the pluripotency markers OCT4, Tra1-81 and SSEA4 and exhibited a normal karyotype. VCP mutations were confirmed by Sanger sequencing. An

additional iPSC line (ND41866*C), derived from a matched healthy control, was employed as another comparator from Corriell.

iPSC line	Mutation present	Method of reprogramming	Age of Donor	Age at disease onset	Sex of Donor
Control 1 (JOM)	None	Episomal	78	-	Male
Control 2 (ND41866*C)	None	Retroviral	64	-	Male
VCP mutant 1 (CB1E)	R155H	Episomal	43	40	Female
VCP mutant 2 (CB1D)	R155H	Episomal	43	40	Female
VCP mutant 3 (GL1B)	R191Q	Episomal	42	36	Male
VCP mutant 4 (GL1A)	R191Q	Episomal	42	36	Male

VCP mutant cultures recapitulate cell specific selective vulnerability, a key aspect of MND pathogenesis

We first questioned whether key pathological aspects of MND were recapitulated in our model system. In MND, there is a selective vulnerability of motor neurons to neurodegeneration, therefore we began with a cell type-specific and VCP mutation-dependent cell survival assay across a timecourse of motorneurogenesis and astrogliogenesis.

Firstly this was tested using immunocytochemistry on VCP and control derived MNs at D17. We used the marker caspase3 to depict cellular apoptosis (Figure 5.2.A). Analysis, carried out by Andras Lakatos, showed that number of caspase3 positive cells was increased by over 2 fold in VCP MNs compared to control (Figure 5.2.B). Further, in collaboration with Zhi Yao, Minee Choi and Sonia Gandhi, we optimized and employed an automated high throughput method for assessing cell death based on a PI/Hoechst assay. Cells needed to

be plated as single cells and analysed at approximately 70% confluency (15,000 cells were plated per well of a 96well plate). Zhi Yao stained cultures with Hoechst and PI, which enabled us to respectively count the total number of cells (total number of nuclei) and the number of cells with a loss of membrane integrity, which reflected cellular death. We analysed the percentage of cell death at 3 time points in MN development (NPC, d3 MN, d17 MN) and 3 time points in AC development (GPC, d14 AC, d28 AC). This again showed a selective vunerability of VCP mutant MNs at D17, the basal level of cell death was 22.47%±9.31% in control MNs vs 52.23% ±1.61% in VCP mutant MNs (n= 2 control clones and 3 mutant clones from 2 patients, $p < 0.05$, 2-way ANOVA) (Figure 5.2.D). Additionally we observed a trend of increased cell death in the VCP mutant neuronal cultures at earlier time points (both NPC and d3 MN) although this did not meet statistical significance. Notably, turning our attention to control and VCP mutant ACs or GPCs, we did not find a significant difference in cell death (Figure 5.2.F). Overall we show a cell-type specific and mutation-dependent cell survival vulnerability in VCP MN's but not in ACs.

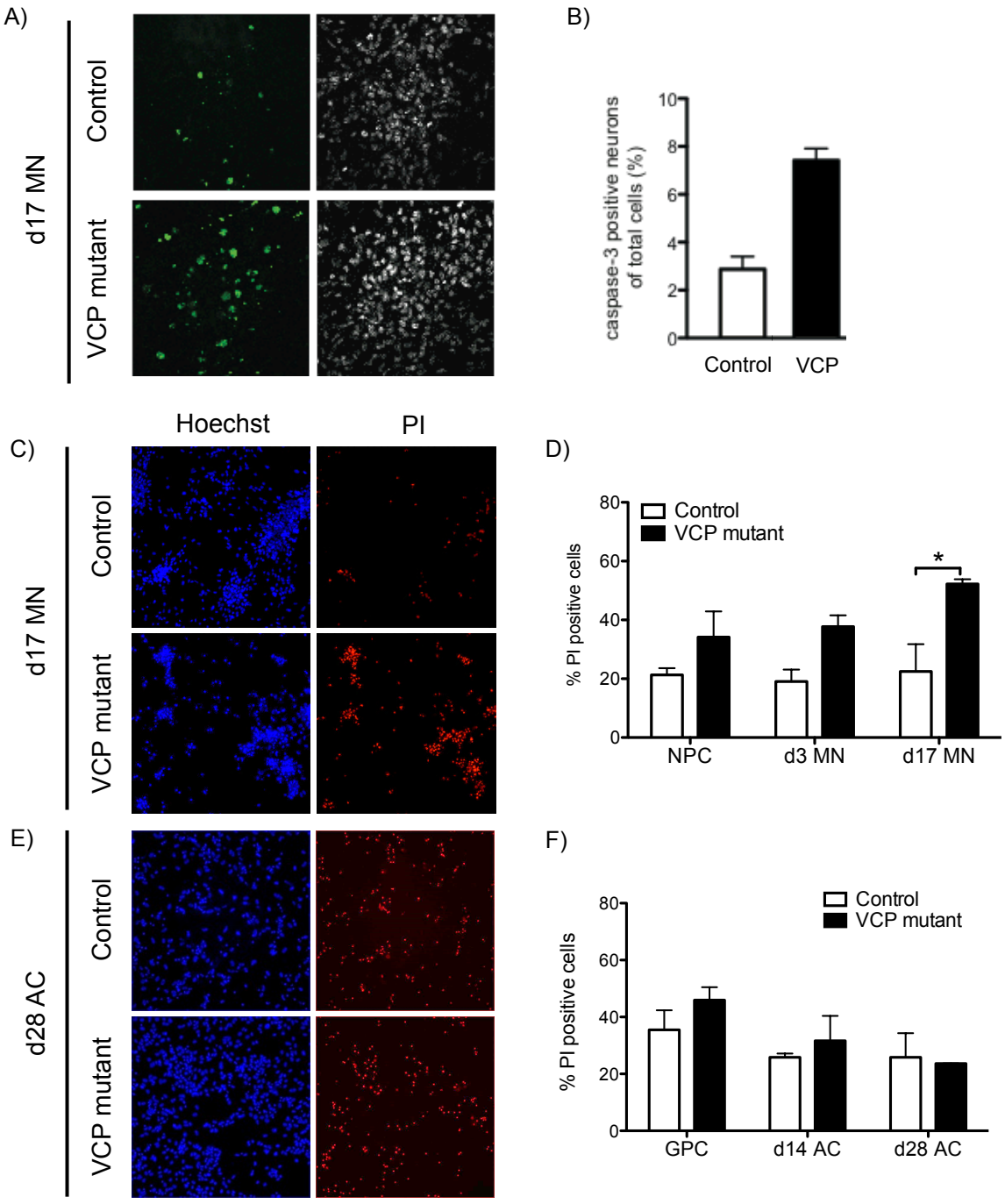


Figure 5.2; VCP mutant MN display a selective vulnerability

A) Representative images of Control and VCP mutant MNs stained with Caspase-3 (green) and DAPI (white). Images acquired on 20x objective. B) Quantification of the number of cells +ve for Caspase-3. C) Representative images from the high-throughput analysis of Hoechst (blue) and PI (red) staining in control and VCP mutant MNs(d17), and E) for control and VCP mutant ACs (d28). Images acquired on 20x objective. D) & F) Quantification of percentage of PI positive cells in control and VCP mutant cells across a motor neuron (NPC, d3MN, d17MN) and astrocyte (GPC, d14AC, d28AC) time-course . N=3, Error bars represent mean ± SEM.

Is RNA regulation disrupted in VCP-related MND?

As rationalized previously, I next explored the hypothesis that RNA regulation is disrupted in VCP-related MND. I have analyzed the transcriptional signature of both control and VCP-mutant iPSC neural derivatives across differentiation using RNA sequencing. Unlike other differentiation methods for hiPSCs, which requires prior purification of cultures, our strategy yielded highly enriched cellular populations consistently and efficiently, enabling us to accurately use RNA sequencing to begin to unravel underlying molecular pathogenic mechanisms in selected cell types.

I generated libraries from Control and VCP mutant samples across the differentiation of motor neurons (iPSC, NI, NPC, d3MN, d17MN) and astrocytes (GPC, d14AC, d28AC), using the Truseq stranded mRNA kit. All samples used went through tight quality checks before analysis (see Methods for further details).

I firstly carried out principle component analysis using both Control and VCP mutant samples throughout motor neuron differentiation (Figure 5.3.A). The segregation pattern supports our differentiation strategy and also highlights that Control and VCP Mutant populations becomes increasingly segregated during neuronal differentiation, and is most prominently seen after terminal differentiation to Motor Neurons. Moreover, I performed differential expression analysis on each stage of motorneurogenesis (iPSC, NI, NPC, d3MN &d17MN) between control and VCP mutant samples (Figure 5.3.B). This also showed that as cells differentiate towards a motor neuron fate there is a greater number of differentially expressed genes and also an increase in significance (p-value) of the differential expression.

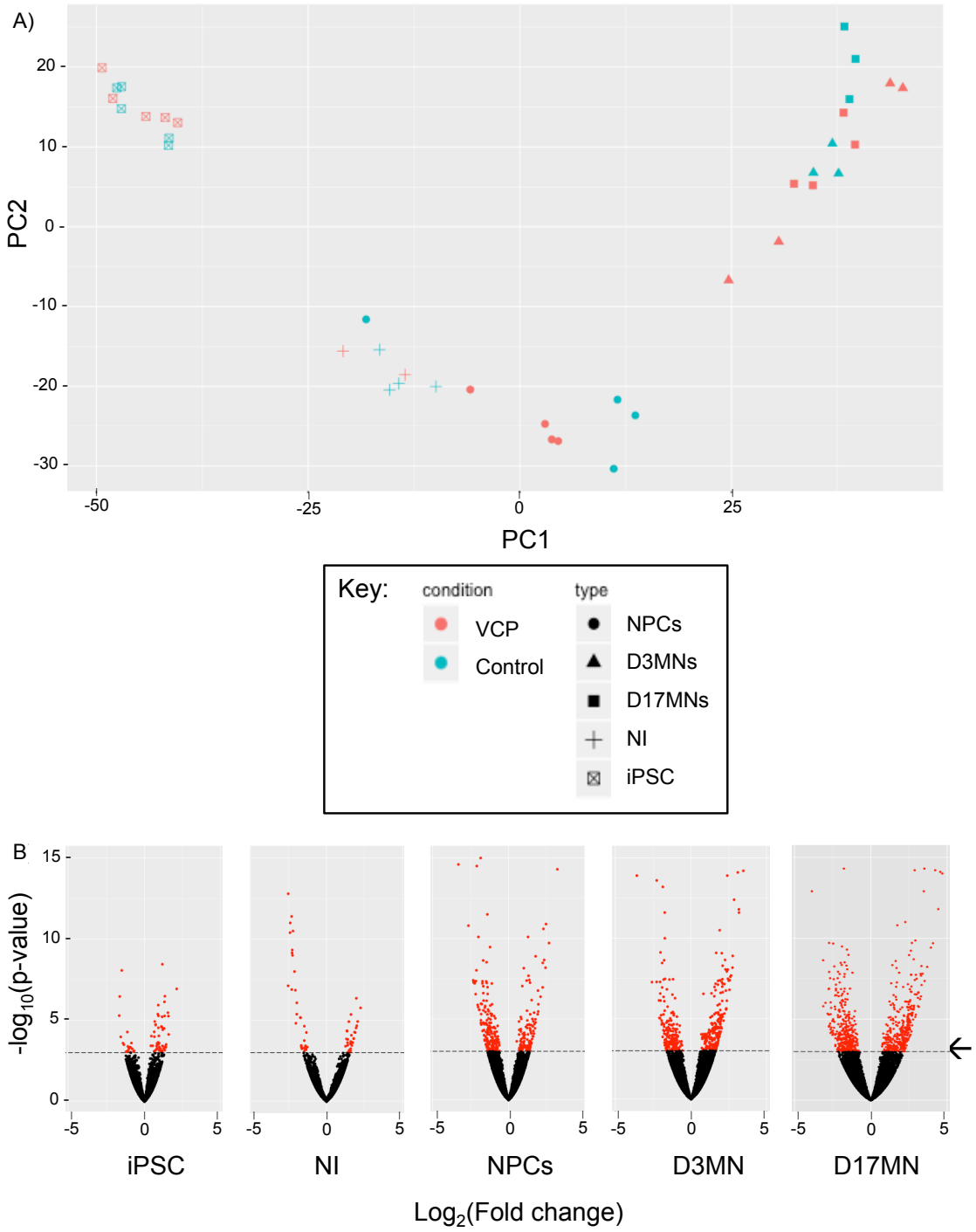


Figure 5.3; Transcriptional profile throughout motor neuron differentiation of Control and VCP mutant cells

A) Principle component analysis separates samples by differentiation stage (Pluripotent, NPC, MN) and after terminal differentiation of MNs we also see separation of control and VCP mutant samples (d3/17MNs). Individual biological replicates plotted. B) Differential expression analysis between control and VCP mutant samples at 5 stages throughout motor neuron differentiation (Pluripotent, NI, NPC, d3MN and

d17MN). Arrow indicates p=0.001. A minimum of biological n=2/technical n=3 for each condition (genotype and differentiation stage) was used.

Focusing on terminal MN's (D17), where the greatest transcriptional changes are seen between VCP mutant and control samples, there are 206 differentially expressed mRNA, 127 up-regulated and 79 down regulated (using a threshold of log(FC)>1.5, and p-adj >0.05). I performed gene ontology analysis on the differentially expressed genes at D17MNs (Figure 5.4.A). Among the terms highlighted there is a strong enrichment of terms related to synaptic functions down regulated in VCP mutant MNs, including ion channels, gated channels and transmembrane transporter activities suggesting there could be a loss of synaptic function in VCP-mutant MNs. On the other hand, cell cycle regulation, microtubule motor activity, and DNA/Protein binding are up-regulated in VCP-mutant MNs.

	Up reg	Enrichment Score	Adjusted p-value		Down reg	Enrichment Score	Adjusted p-value
biological process	cell cycle	42.06	5.13E-49		system process	24.34	1.33E-09
	cell cycle process	32.13	4.24E-55		neurological system process	18.34	1.10E-10
	cell cycle phase	25.87	6.06E-57		cell-cell signaling	16.42	2.21E-10
	mitotic cell cycle	24.52	9.80E-54		multicellular organismal signaling	12.99	8.86E-11
	M phase	15.74	1.24E-50		transmission of nerve impulse	12.68	7.53E-11
	cell division	14.6	2.92E-40		synaptic transmission	11.18	1.33E-09
	organelle fission	12.28	1.31E-46		behavior	7.79	1.20E-05
	M phase of mitotic cell cycle	11.93	1.02E-48		single-organism behavior	5.44	5.02E-06
	mitosis	11.59	1.02E-48		cognition	3.03	5.63E-05
	nuclear division	11.59	1.02E-48		learning or memory	2.73	7.96E-05
molecular function	protein binding	186.34	8.90E-06		substrate-specific transporter activity	14.35	8.68E-05
	DNA binding	57.77	4.23E-06		transmembrane transporter activity	13.23	3.31E-05
	sequence-specific DNA binding transcription factor activity	22.48	3.34E-05		substrate-specific transmembrane transporter activity	12.23	2.48E-05
	sequence-specific DNA binding	13.57	8.03E-06		ion transmembrane transporter activity	11.32	2.48E-05
	chromatin binding	8.74	2.44E-05		channel activity	5.66	8.68E-05
	regulatory region DNA binding	7.18	8.90E-06		metal ion transmembrane transporter activity	5.61	4.11E-05
	regulatory region nucleic acid binding	7.18	8.90E-06		ion channel activity	5.5	8.68E-05
	transcription regulatory region DNA binding	7.05	8.90E-06		gated channel activity	4.5	3.31E-05
	structure-specific DNA binding	4.57	4.23E-06		ion gated channel activity	4.5	3.31E-05
	microtubule motor activity	2.15	3.34E-05		cation channel activity	4.08	8.68E-05
cellular component	microtubule cytoskeleton	25.03	1.20E-14		membrane	114.08	5.96E-11
	chromosome	16.94	2.26E-28		membrane part	84.15	4.17E-09
	chromosomal part	14.28	8.39E-28		intrinsic to membrane	71.14	5.23E-10
	spindle	7.16	2.69E-22		integral to membrane	69.75	6.29E-09
	condensed chromosome	4.81	9.84E-32		cell periphery	56.65	7.19E-14
	chromosome, centromeric region	4.62	2.21E-26		plasma membrane	55.3	4.10E-13
	kinetochore	3.27	3.66E-23		intrinsic to plasma membrane	16	1.58E-07
	spindle pole	2.98	2.52E-13		integral to plasma membrane	15.39	1.58E-07
	condensed chromosome, centromeric region	2.76	4.86E-27		neuron projection	12.88	5.23E-08
	condensed chromosome kinetochore	2.6	7.85E-27		synapse part	7.03	5.26E-08

Figure 5.4; GO analysis in d17MNs

Top 5 GO terms for biological function, molecular function and cellular component associated with differentially expressed genes in control and VCP mutant d17MNs. Analysis carried out on differentially expressed genes that have log(fold change) >1.5 and padj <0.05.

I next turned my attention to the transcriptional profile of VCP mutant ACs. Again I performed principle component analysis for both control and VCP mutant samples throughout AC differentiation, which showed segregation of samples dependent upon differentiation status (pluripotent, NPC and terminally differentiated AC) (Figure 5.5.A). Although we did not detect a cell autonomous selective vulnerability in VCP mutant ACs we did see some changes in RNA levels between control and VCP mutant ACs (Figure 5.5.B). We detected 156 differential expressed genes (p-value <0.05, log(fold change) >1.5); 93 up-regulated and 63 down-regulated mRNAs (Figure 5.5.C).

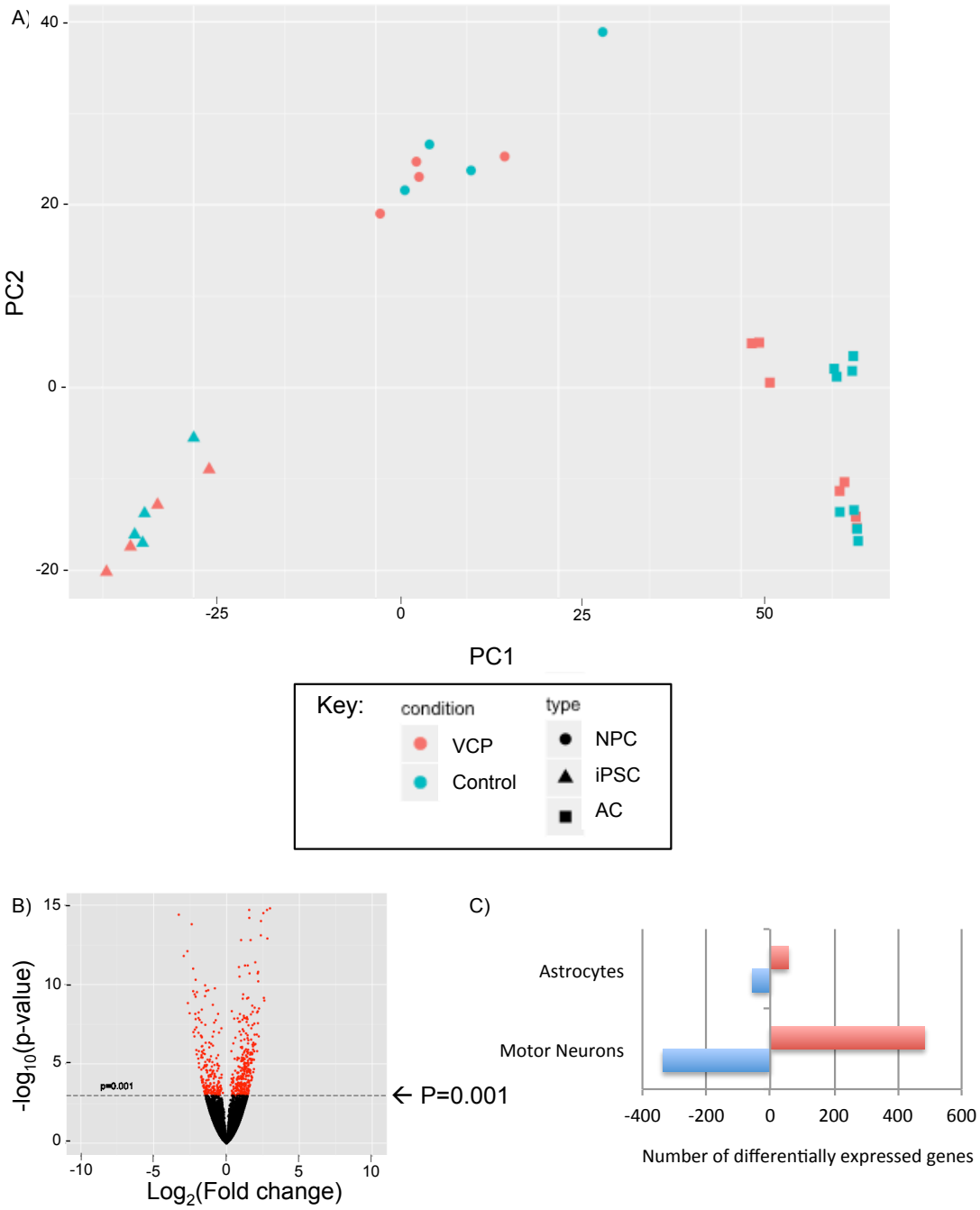


Figure 5.5; RNA sequencing of Control and VCP mutant astrocytes

A) Principle component analysis separates samples by differentiation stage (Pluripotent, NPC, AC). Individual biological replicates plotted. B) Differential expression analysis between control and VCP mutant astrocytes. C) Bar plot to show the number of differentially expressed genes between Control and VCP mutant samples, comparison between MN and AC samples. Analysis carried out using a minimum of biological n=2/technical n=3 for each condition (genotype and differentiation stage).

Following, I performed GO analysis on the differentially expressed genes (log(fold change) >1.5 and p-value <0.05) (Figure 5.6.A). Among the terms associated there is a strong emphasis on genes involved in extracellular matrix and cell-to-cell signaling which raises several hypothesis;

- VCP-mutant ACs play a non-cell autonomous role in MND.

Astrocytes release, among other factors, extracellular matrix proteoglycans and remodeling factors that promote neuronal survival and maturation. In MND it has been shown that for SOD1 mutations, astrocytes release factors that are toxic to MNs. More recently it has also been shown that in disease astrocytes release an increase factors leading to a developmental syndrome (Krencik et al. 2015). Further work needs to be carried out to establish if this could also be a pathogenic mechanism underlying VCP related MND.

- VCP-mutant AC have different reactivity than control AC.

Astrocyte reactivity influences cell-to-cell communications, cytoskeletal changes, extracellular matrix and released factors therefore it is possible that the changes in gene expression we see in VCP-ACs are due to an altered state of reactivity (Pekny and Pekna 2014)(Tyzack et al, 2016, in preparation). In the future more precise AC reactivity should be characterized and then this improved model could provide clarification into the reasons in cell-to-cell signaling/extracellular differences in VCP mutant ACs.

	Up regulated	Enrichment Score	Adjusted p-value		Down regulated	Enrichment Score	Adjusted p-value
biological process	regulation of mammary gland epithelial cell proliferation	0.07	0.0003		antigen processing and presentation of peptide antigen via MHC class II	0.2	3.71E-15
	biological adhesion	3.16	0.0003		antigen processing and presentation of exogenous peptide antigen via MHC class II	0.2	3.71E-15
	cell adhesion	3.15	0.0003		antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	0.21	3.71E-15
	mammary gland epithelial cell proliferation	0.11	0.0005		response to interferon-gamma	0.18	6.24E-14
	regulation of epithelial cell proliferation	0.71	0.0068		regulation of T cell activation	0.29	1.21E-13
	mammary gland epithelium development	0.27	0.008		positive regulation of T cell activation	0.22	2.96E-13
	epithelial cell proliferation	0.82	0.0137		cellular response to interferon-gamma	0.15	5.35E-13
	negative regulation of mammary gland epithelial cell proliferation	0.03	0.0179				
	adult locomotory behavior	0.24	0.0508				
	stem cell proliferation	0.05	0.0508				
Molecular function	cyclin-dependent protein kinase regulator activity	0.08	0.0053		MHC class II receptor activity	0.01	1.90E-15
	cyclin-dependent protein kinase inhibitor activity	0.05	0.0405		peptide antigen binding	0.02	0.0024
	drug transporter activity	0.06	0.051		antigen binding	0.02	0.0031
	protein serine/threonine kinase inhibitor activity	0.08	0.063		peptide binding	0.18	0.0075
	symporter activity	0.31	0.063		amide binding	0.19	0.0075
	L-amino acid transmembrane transporter activity	0.14	0.0879		cytokine receptor activity	0.05	0.0086
	kinase inhibitor activity	0.17	0.0879		signal transducer activity	1.09	0.0206
	active transmembrane transporter activity	0.86	0.0879		molecular transducer activity	1.09	0.0206
	secondary active transmembrane transporter activity	0.47	0.0879		signaling receptor activity	0.7	0.023
	protein kinase inhibitor activity	0.17	0.0879		receptor signaling protein activity	0.14	0.0409
Cellular component	integral to membrane	15.09	0.0028		lysosomal membrane	0.27	1.02E-14
	membrane part	18.18	0.0028		ER to Golgi transport vesicle membrane	0.08	4.79E-14
	intrinsic to membrane	15.37	0.0028		ER to Golgi transport vesicle	0.09	1.14E-13
	extracellular region	4.63	0.0104		vacuolar membrane	0.36	1.51E-13
	membrane	24.67	0.0133		clathrin-coated endocytic vesicle membrane	0.05	3.94E-15
	apical part of cell	0.85	0.0208		clathrin-coated endocytic vesicle	0.06	7.89E-15
	dystrophin-associated glycoprotein complex	0.07	0.0237		trans-Golgi network membrane	0.06	7.89E-15
	apical plasma membrane	0.59	0.0301		MHC class II protein complex	0.03	9.37E-28
	axoneme	0.26	0.249		MHC protein complex	0.06	1.71E-22
	extracellular matrix	1.29	0.3312		integral to lumenal side of endoplasmic reticulum membrane	0.05	3.94E-15

Figure 5.6; GO analysis of differentially expressed genes between Control and VCP mutant ACs

Top 5 GO terms for biological function, molecular function and cellular component associated with differentially expressed genes in control and VCP mutant ACs. Analysis carried out on differentially expressed genes that have log(fold change) >1.5 and padj <0.05.

SYT1 staining in VCP mutant MNs

Since I observe that there are many synaptic associated proteins among the top down-regulated genes in VCP-MNs and additionally because VCP has been previously shown to interact with synaptic protein (H. F. Wang et al. 2011), we collaborated with Andras Lakatos to explore the effect of VCP mutations on MN synapse formation. To address this I immunolabeled both control and VCP mutant terminally differentiated motor neuron cultures (d17) with pre-synaptic marker, synaptotagmin1 (SYT1), and co-immunolabeled with neuronal and motor neuron markers, MAP2 and CHAT respectively (Figure 5.7.A). Next, using confocal microscopy Andras Lakatos identified areas of close acquisition between SYT1 and MAP2 staining, which we denoted as positive puncta on both the soma and dendrites of MNs (Figure 5.7.B & C). Only cells in clusters interconnected by axons, forming a network were analysed. SYT1 density on the soma was significantly reduced by 2-fold in VCP mutant lines (0.038±0.004 and 0.033±0.003 puncta/μm^2; overall $p<0.0001$, Kruskal-Wallis test) when compared to control lines (0.062±0.005 and 0.079±0.007 puncta/μm^2). Similarly, dendritic SYT1 densities were significantly lower in the VCP mutant MN cultures (0.153±0.064 and 0.138±0.058 puncta/μm^2; overall $p<0.0001$, Kruskal-Wallis test) compared to control MNs (0.262±0.117 and 0.305±0.098 puncta/μm^2).

As we have previously shown VCP mutant MNs display increased cell death it is important to determine whether a reduction in SYT1 staining represents early MN dysfunction or loss of inter-connected neurons. To do this Andras Lakatos counted the number of Chat/MAP-2 positive neurons with intact DAPI stained nuclei in the analysed cell clusters (Figure 5.7.D). Also we compared the cellular composition across the controls and mutants samples by calculating the MN ratio (CHAT +ve versus CHAT –ve cells) across each cell line (Figure 5.7.E). We showed that both the number of total neurons and the MN +ve ratios were comparable confirming that a reduction in SYT1 staining in VCP mutant MNs is not attributable to differential loss of VCP MNs. Thus, MNs have the same chance to establish the same number of synaptic connections in both Control and VCP groups. We conclude the reduction in SYT1 staining is not directly caused by the death of the presynaptic neuron in these cell groups, and that there is a pre-synaptic dysfunction before the post-synaptic cell dies. Further work is now required to dissected the causative mechanisms behind

the reduced SYT1 staining observed; what is the trigger or whether this is a developmental defect in VCP-MNs?

Our findings add to other studies that also suggest a role for defective synaptic function in MND. For a while excitotoxicity has been considered to play a role neurodegeneration/MND(Bories et al. 2007; Kuo et al. 2005)(Wainger, Kiskinis, Mellin, Wiskow, Han, et al. 2014). However to define the presence and functional consequences of a synaptic defect in VCP mutant MNs requires further work.

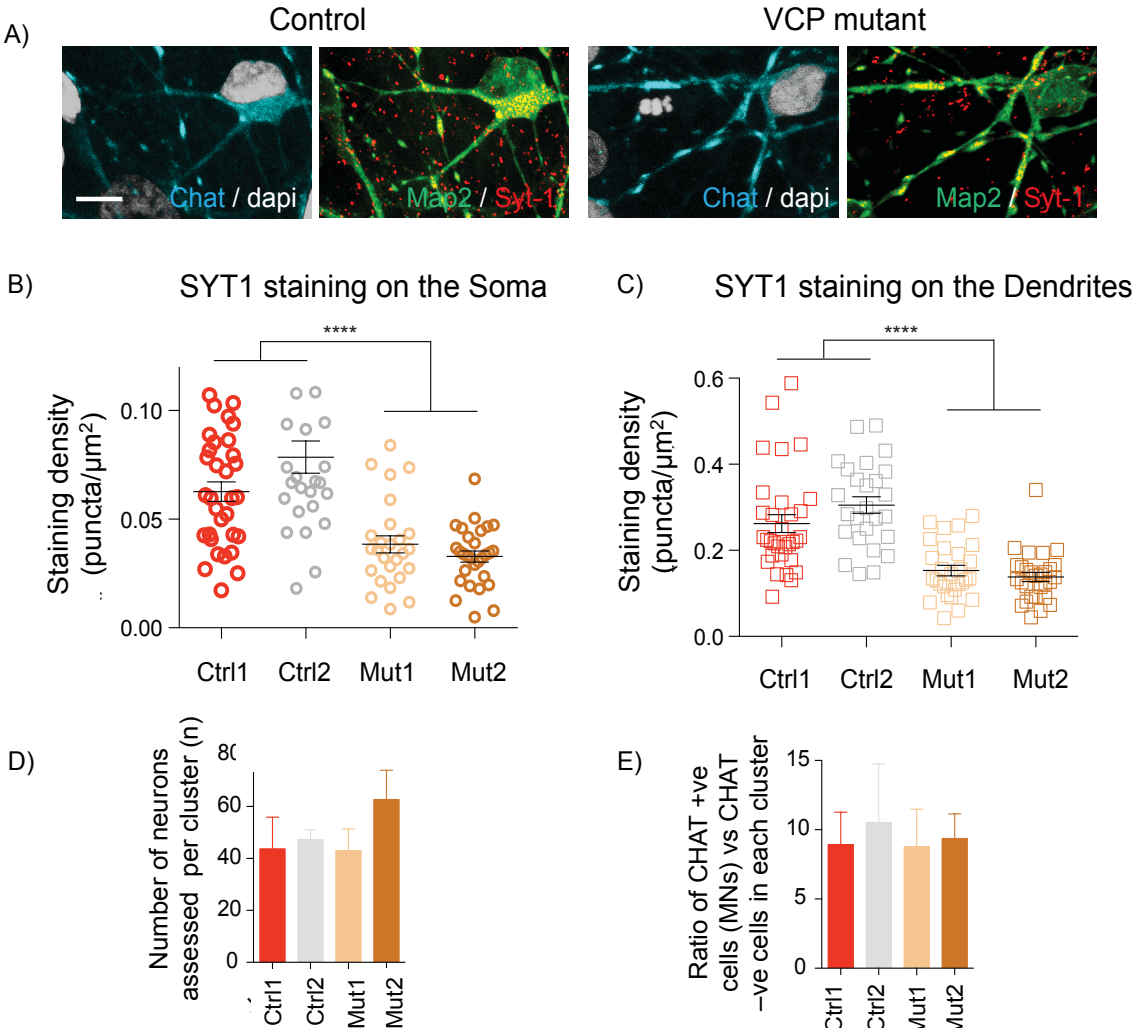


Figure 5.7; Synaptotagmin-1 staining in Control and VCP mutant MNs

A) Maximum projection of Z-stack confocal microscopy images showing synaptic input (*SYT-1 positive puncta are in red*) on Chat (*blue*)/MAP-2 (*green*) positive MNs with intact nuclei (DAPI is white). B) and C) Analysis of SYT1 positive puncta in Control and VCP mutant-MNs on the Soma and the dendrites respectively. Graphs represent mean SYT1 density values defined by puncta / area on the MN soma (N=26-32 cells, Dunn's test) and puncta per length for a main MN dendrite (n=26-32 dendrites, Dunn's test) in two control lines (*Ctrl1 and Ctrl2*) and two VCP mutant lines (*Mut1 and Mut2*). D) Bar

graph demonstrates the mean number of total neurons in the different cell lines. E) Bar plot shows the mean ratio of CHAT positive (MNs) vs CHAT negative cells in each cluster, in the different cell lines. (N=3 cultures, ANOVA). Data represent mean+SEM. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Scale Bar: 25 μm.

Live cell imaging in hiPSC-neural derivatives to explore defective cytoplasmic events in VCP-related MND

Having confirmed a robust cell survival and synaptic phenotypes in VCP-mutant MNs, we next searched for earlier pathogenic events in our VCP model. In collaboration we have looked at the dysregulated cytoplasmic events in VCP-related MND using both ICC and multiple live cell imaging assays. We assess samples across a complementary time-course to the RNAseq experiments and analyze a range of additional phenotypes rationalized by pathways previously implicated in MND (Cozzolino and Carri 2012; D'Amico et al. 2013; Ilieva et al. 2007; Kiskinis et al. 2014; S Sasaki 2010; Weiduschat et al. 2014); we systematically examine, across a time course, the possibility of mitochondrial dysfunction, ER stress, and oxidative stress in VCP-hiPSC derivatives. For these experiments I carried out all iPSC culture and neural differentiation, then mitochondrial function, oxidative stress and ER stress assays were carried out by Minee Choi and Zhi Yao.

Live cell imaging is a powerful method to analyze cellular dynamics in a quick and accurate way. It allows us to study individual cells and follow them across time whilst analyzing particular events. Unlike ICC, there are no artifacts that can be induced due to the fixation process, however careful attention must be played to the health/function of cultures during imaging, as fluorescent illumination can be harmful. Using live cell imaging we are able to look at a range of functions in the cell such as mitochondria viability, activity of the electron transport chain (ETC) and ATP production, ROS production and calcium influx.

Mitochondrial dysfunction

In neurons oxidative phosphorylation is the main source of ATP/energy, so mitochondria being the cells power station play a vital role. Owing to a decline in mitochondrial function in aging it is thought that dysfunctional mitochondria could be potentially an important factor in age related neurodegenerative

diseases, hence there have been many studies analyzing their morphology and functional characteristics in MND pathogenesis. In VCP-mutant patient derived fibroblasts it has previously been shown that there is mitochondrial uncoupling leading to a reduction in ATP production/energy supply(Bartolome et al. 2013). This means the rate of electron transport in the respiratory chain is no longer linked to ATP production. The mitochondria's membrane potential is critical for the production of ATP by respiration and any loss results in depletion of ATP and subsequent cell death.

Mitochondrial membrane potential can be used as a general marker of mitochondrial health and it has previously been show to be reduced in VCP-deficient cellular models and VCP mutant patient cells (Bartolome et al. 2013). Here we have explored whether there this is also seen in VCP mutant iPSC-derived neural derivatives. To do this we have used TMRM, a fluorescent cationic dye that is cell-permeable and is actively taken up by the mitochondria to look at the mitochondrial membrane potential, an indicator of mitochondrial viability telling us about flow of H+ ions across membrane. (Figure 5.8A) We found a reduction in basal levels of mitochondrial membrane potential in VCP mutant d17 MNs compared to control (54.65±6.64% in VCP mutant compared to control 100%, n= 2 control lines and 3 mutant clones, $p< 0.05$) (Figure 5.8B). Interestingly we noted a trend to decreased mitochondrial membrane potential in the VCP mutant NPCs and d3 MNs, but this only reached significance in d17 MN. We also found a decrease in the mitochondrial membrane potential in VCP d14 ACs compared to controls (72.47±2.09% vs 100% control, n=2 control lines and 2 mutant clones, $p<0.05$, 2 tailed t-test). Interestingly this difference had resolved by the d28 AC timepoint, which may reflect the activation of endogenous compensatory mechanisms.

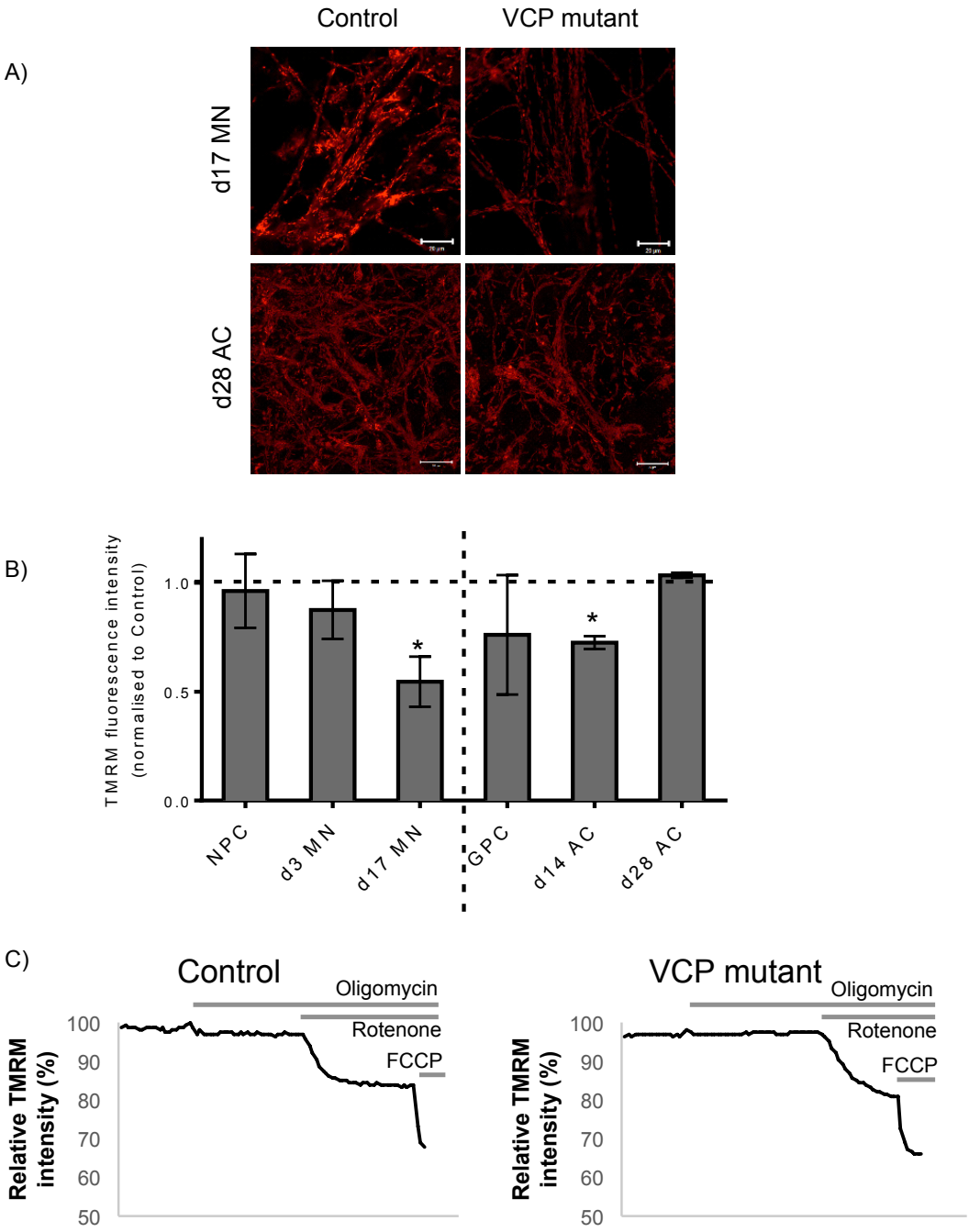


Figure 5.8; Mitochondrial dysfunction in VCP MNs

A) Control and VCP mutant d17 MN and d28 AC loaded with TMRM;. Scale bar = 20µm. B) Histogram showing decreased basal mitochondrial membrane potential in VCP mutant d17 MN and d14 AC compared to control cells. N=3 separate inductions/replicates (from n=2 control cell iPSC lines and n=3+ VCP mutant lines). Error bars represent mean ± SEM. C) Representative traces showing that, in both control and VCP mutant d17 MN, oligomycin did not affect mitochondrial membrane potential; rotenone induced partial depolarization and FCCP induced complete depolarization. Y-axis scale represents the relative TMRM intensity before and after treatment, where 100% is the baseline TMRM intensity for each sample.

To further understand the cause of the reduced mitochondrial membrane potential in VCP mutant d17 MNs, we investigated a possible mechanism by which the mitochondrial membrane potential is maintained, by studying its sensitivity to a range of mitochondrial inhibitors. Zhi Yao and Minee Choi measured fluorescent intensity at baseline and after stimulus; we use three modulators of the electron transport chain (FCCP, oligomycin (complex V inhibitor) and rotenone (complex I inhibitor)). In control d17 MNs, inhibition of complex V by oligomycin did not lead to depolarization, while subsequent inhibition of complex I by rotenone produced a significant reduction in the TMRM signal, indicating a rapid loss of mitochondrial membrane potential (Figure 5.8.C). In contrast to other neurodegenerative models, the inhibition of complex V in VCP mutant d17 MNs did not cause mitochondrial depolarization, whereas inhibition of complex I produced marked reduction of mitochondrial membrane potential (Figure 5.8.C). This result suggests that both VCP mutant MNs maintain mitochondrial membrane potential largely through respiration.

There are several hypotheses why Mitochondria dysfunction could cause the death of MNs, for example by an increased production of ROS leading to oxidative stress and activation of intrinsic apoptotic pathway and instrinic Ca+ buffering to prevent excitotoxicity. Therefore we next turn to investigate oxidative stress in VCP mutant neural derivatives.

Oxidative stress

Oxidative stress is a well-recognized feature in neurodegenerative diseases including in both sporadic and familial MND (Carrì et al. 2015; D'Amico et al. 2013; Gandhi and Abramov 2012; Weiduschat et al. 2014). It occurs due to the imbalance of production of reactive oxidative species and antioxidant defenses. To date oxidative stress has not been implicated in VCP'opathies but, on top of oxidative stress being broadly implicate in MND, VCPs involvement with the mitochondria gives added reason to explore the possibility of oxidative stress playing a role in pathogenesis of VCP-related MND. Mitochondria are a target by oxidative stress (genetic, metabolic and environmental alterations from oxidative insult will make mitochondria less dynamic in responding to needs of the cell) and conversely mitochondria are also the main site of production of ROS. Therefore oxidative stress can result from impairment of mitochondrial function. As previously discussed we detect deficiency in mitochondrial function

in VCP mutant MNs at D17. Thus we next investigate the presence and timing of oxidative stress in VCP-neural derivatives (i.e. pre or post mitochondrial defects?). To study it we have employed live fluorescent imaging of both ROS production and antioxidant levels.

Minee Choi and Zhi Yao evaluated superoxide production by measuring the rate of oxidation of the dihydroethidium dye (DHE) as a ratio of the oxidised over the reduced form (representative traces), normalising the data to percentage values (Figure 5.9.A). We found that the VCP mutant d3 MNs and d17 MNs exhibited significantly higher rates of ROS production compared to control (d3 MN 189.9±30.4% vs 100% control, n=2 control lines and 3 mutant clones, n=3 repeats; d17MNs 187.3±42.7% vs 100% control, n= 2 control clones and 2 mutant clones, n=3 repeats). There was a modest increase in ROS production in the VCP mutant GPCs compared to control (164.3±29.6% vs 100% control, n=2 control clones and 2 mutant clones, n=3 repeats). Notably, there was no significant difference in ROS production between VCP mutant and control in the d14 and d28 ACs (Figure 5.9.B).

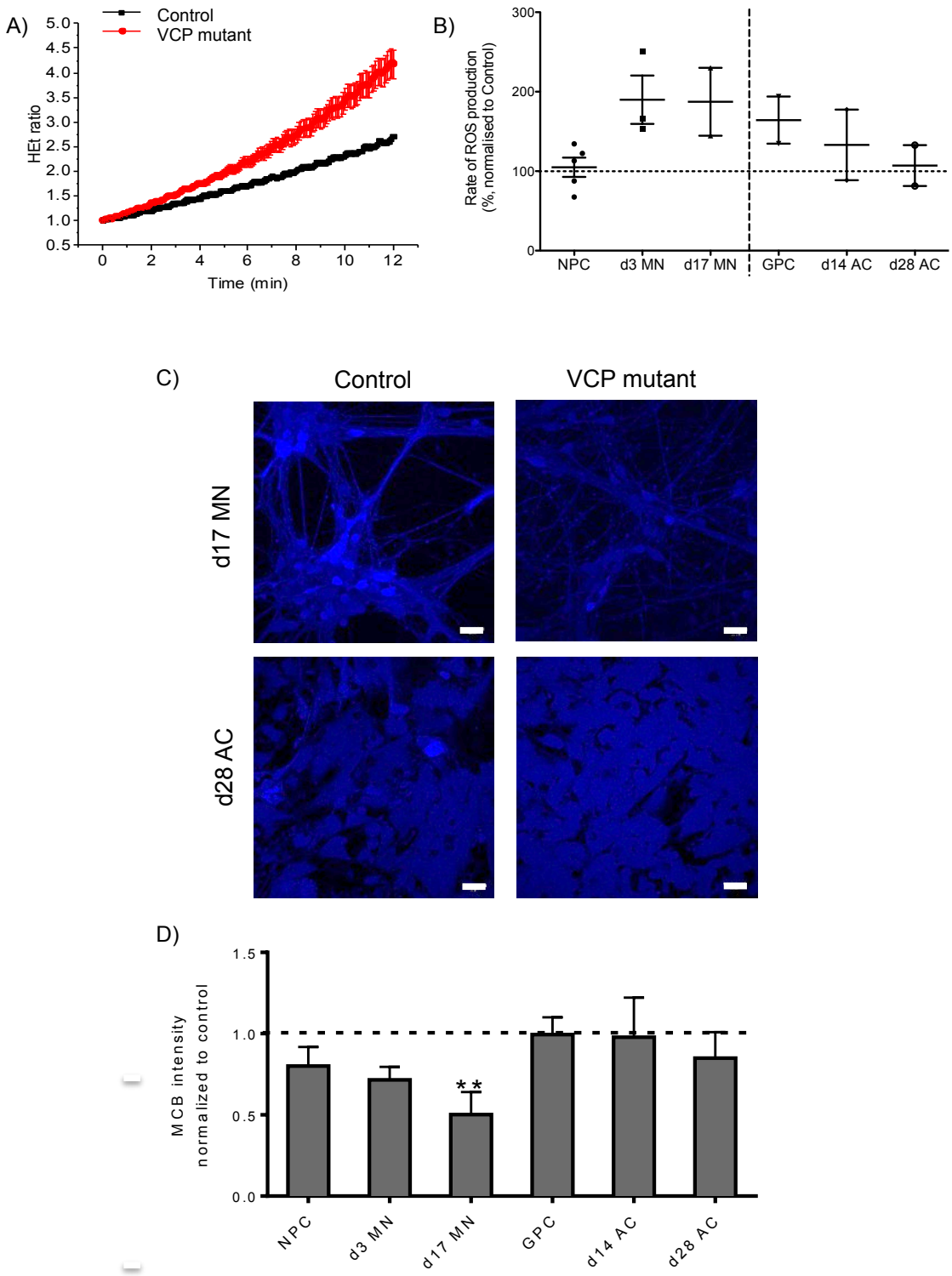


Figure 5.9; Oxidative stress in VCP MNs

A) Representative traces and B) analysis showing increased ROS production rate in VCP mutant cells compared to control. N=3 replicates (from n=2 control cell iPSC lines and n=3+ VCP mutant lines). C) Control and VCP mutant d17 MN and d28 AC loaded with MCB, a glutathione indicator. Scale bar = 20μm. D) Histogram showing decreased GSH level in VCP mutant d17 MN but not NPC, d3 MN, GPC, d14 AC and d28 AC.

N=3+ separate inductions/replicates (from n=2 control cell iPSC lines and n=3+ VCP mutant lines). Error bars represent mean ± SEM.

In order to confirm the presence of oxidative stress, it is necessary to demonstrate an imbalance between the production of ROS and the antioxidant defences. We therefore measured the levels of glutathione using the fluorescent indicator monochlorobiomine (MCB; representative images,Figure 5.9.C). We showed that there is a significant decrease in the levels of glutathione in the VCP mutant d17 MNs only, and not in any other neuronal time point (50.16±8.0% vs 100% control, n = 2 control lines and 3 mutant clones from 2 patients, $p<0.001$) (Figure 5.9.D). Taken together these data implicate an early increase in the generation of ROS particularly in the MNs (at d3 and at d17), and this is associated with oxidative stress due to depletion of glutathione levels by d17 in culture. Therefore we observe oxidative stress alongside mitochondrial dysfunction in VCP-related MND, suggesting they influence each other causing a cascade of damage.

Additionally we investigated glutathione levels in Control and VCP mutant GPCs and ACs. One function of astrocytes is to release glutathione to neurons to protect them & their mitochondria from oxidative stress. We see no significant alteration in glutathione levels at any time point suggesting AC themselves are not under oxidative stress themselves, however if MNs and ACs were co-cultured I hypothesize that we would observe a compensatory glutathione release in astrocytes to protective VCP mutant MNs, unless VCP mutant ACs have dysfunctional cell-cell communication.

ER stress

Finally we ask if ER stress plays a role in VCP-related MND. The pathology of MND (inclusions of misfolded, aggregated & toxic proteins), which is also replicated in VCP'opathies, indicates that a significant disruption to proteostasis has occurred. This is backed by previous research that has demonstrated evidence of ER stress, the role of the unfolded protein response (UPR) and defective protein clearance pathways (eg. Autophagy) in MND pathogenesis (Kiskinis et al. 2014; Moreno et al. 2013; Walker et al. 2013). Further VCP is known to function in ERAD, ER morphogenesis and ER translocation raising the hypothesis that a mutation in VCP will affect its ERs functions making the

ER vulnerable to insult. Together this provides strong rational to probe ER stress in VCP-neural derivatives.

To begin to probe ER stress in VCP mutant neural derivatives we looked at whether VCP mutant neural derivatives exhibited greater sensitivity to an ER stressor than control cells. I treated cultures for 48hours with 0.1-10μM Tunicamycin (an ER stress inducer) and Zhi Yao used the automated high-throughput microscope to calculate percentage cell death. NPCs, MNs, GPCs and ACs were analyzed. Significantly increased cell death is evident in VCP mutant d3 MNs compared to control (43.25±2.72% vs. 79.15±5.24%, n= 2 control lines and 3 mutant lines, > 50,000 cells per condition), but not in ACs or NPCs suggesting a cell-specific vulnerability of MNs (Figure 5.10A). We also note that sensitivity to the ER stressor tunicamyin is initiated early after MN terminal differentiation, whether as the previous phenotypes studied have not shown statistical significance until a more mature MN stage.

Moreover we assess calcium function as an additional marker for ER stress. Calcium dysfunction has been previously been shown to play a role in MND. Increased levels of calcium have been detected both SOD1 mice and sMND patients and have been shown to medicate neuronal toxicity (via TDP43)(Aggad et al. 2014). The two main stores of calcium in the cell are in the ER and the mitochondria. Here we have looked at ER calcium storage across motor neurogenesis using thapsigarin. Thapsigarin causes all ER calcium to be released, which can subsequently be captured by ATP (if not it will be taken up by the mitochondria). Minee Choi measured Ca+ release by evaluating the ratio of Fura2-am bound by Ca+ and free. Examining the ER calcium stores we found significant reduction in ER Ca+ storage VCP mutant d3MNs (63.7±2.4% of control, n= 2 control lines and 3 mutant lines) (Figure 5.10B). This likewise suggests ER stress occurs at an early time-point after the terminal differentiation of MNs (d3 MNs) but further studies are required to understand if Ca+ regulation is involved in pathoegenesis.

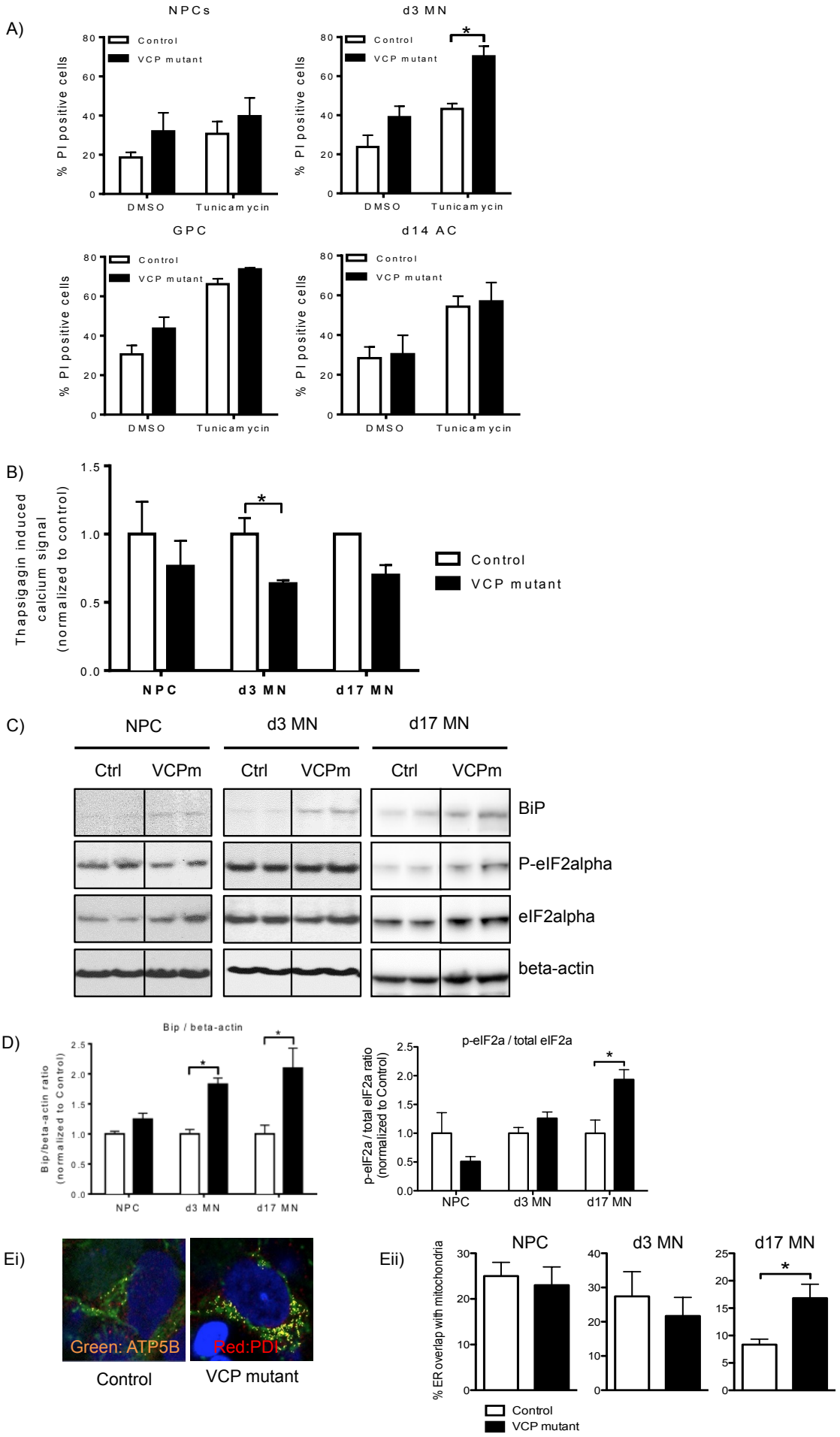

A)
NPCs
d3 MN
GPC
d14 AC
Control
VCP mutant
% PI positive cells
DMSO
Tunicamycin
B)
Thapsigagin induced calcium signal (normalized to control)
NPC
d3 MN
d17 MN
C)
Ctrl
VCPm
BiP
P-eIF2alpha
eIF2alpha
beta-actin
D)
Bip / beta-actin
Bip/beta-actin ratio (normalized to Control)
p-eIF2a / total eIF2a
p-eIF2a / total eIF2a ratio (normalized to Control)
Ei)
Green: ATP5B
Red:PDI
Control
VCP mutant
Eii)
% ER overlap with mitochondria

Figure 5.10; ER stress in immature VCP-neuronal derivatives

A) Histograms showing percentage of PI positive cells in control and VCP mutant NPC, d3 MN, GPC and d14 AC with or without 48 hours treatment with ER stressor Tunicamycin. Error bars represent mean ± SEM. B) Histogram showing thapsigagin induced calcium signal measured in calcium free media (n=2 control and 3 mutant clones). Error bars represent mean ± SEM. C) Western blot images and D) quantification of Bip/beta-actin and phospho-eIF2alpha/total eIF2alpha levels in control and VCP mutant NPC, d3 MN and d17 MN (n= one biological replicate & 3 technical replicates). Error bars represent mean ± SEM. Ei) Representative images show co-localization of mitochondria (labeled by ATP5B, in green) and ER (labeled by PDI, in red) in control and VCP mutant d17 MN. The site of co-localization highlighted in yellow. Eii) Quantification of percentage of ER co-localized with mitochondrial in control and VCP mutant NPC, d3 MN and d17 MN (n= 1+ biological replicates & 3 technical replicates). Error bars represent mean ± SEM.

Next to increase confidence in an early ER phenotype in VCP MNS Zhi Yao measured protein levels of two ER stress markers Bip and p-eIF2alpha across a motor neuron time-course (Figure 5.10.C). In d3 and d17 MNs, we observed an increase in the expression of Bip and p-eIF2alpha in VCP mutant compared to control (Figure 5.10.D), further highlighting ER stress as an early event in VCP-related MND.

Finally ER stress is associated with increased contact points between the ER and mitochondria. Therefore Zhi Yao next quantified the ER-mitochondrial contacts in MN cultures by measuring the co-localisation of a mitochondrial marker (ATP5b) and an ER marker (PDI) by immunocytochemistry (Figure 5.10.E). There was no difference in the ER-mitochondrial contacts between VCP mutant and control NPCs or d3MNs. However we found a significant reduction in the ER-mitochondrial contacts in the control compared to the VCP mutant at d17 MNs (Figure 5.10.F). It is of no surprise that we see a reduction in ER-mitochondria contacts in the Control MNs during terminal differentiation because ER stress is known to play a role early on during neuronal differentiation (Liu et al. 2012; Zhang et al. 2007). However it is of interest that we see a relative increase in ER-mitochondria contacts in the VCP mutant at D17MNs, again implying the presence of ER stress.

Taken together, these data provide evidence of significant ER stress in the VCP mutant MNs, which begins to manifest first biochemically and through

sensitivity to ER stress in d3 MNs, and progresses to an increase in ER-mitochondrial contacts by d17 MNs. Again, no significant ER stress was detected in VCP-ACs, reinforcing a cell-type specific phenotype.

Are transcriptional changes observed before corresponding cytoplasmic events occur?

We next interrogated the RNAseq data set to see if gene expression changes show alterations in pathways highlighted in our VCP model through imaging. It has recently been reported in mutant iPSC-derived MNs that ER stress response is a prominent transcriptional signature(Kiskinis et al. 2014). Therefore have examined gene expression changes in our VCP mutant MN cultures to address whether transcriptional changes precede cytoplasmic evidence of ER stress. We looked at several markers of ER stress including heat shock proteins, chaperonin subunits, protein disulfide isomerases and of the unfolded protein response. We find no clear evidence of transcriptional deregulation in VCP mutant NPCs or d3MNs leading to ER stress, therefore we conclude that this pathway is first activated in the cytoplasm (Figure 5.11.A).

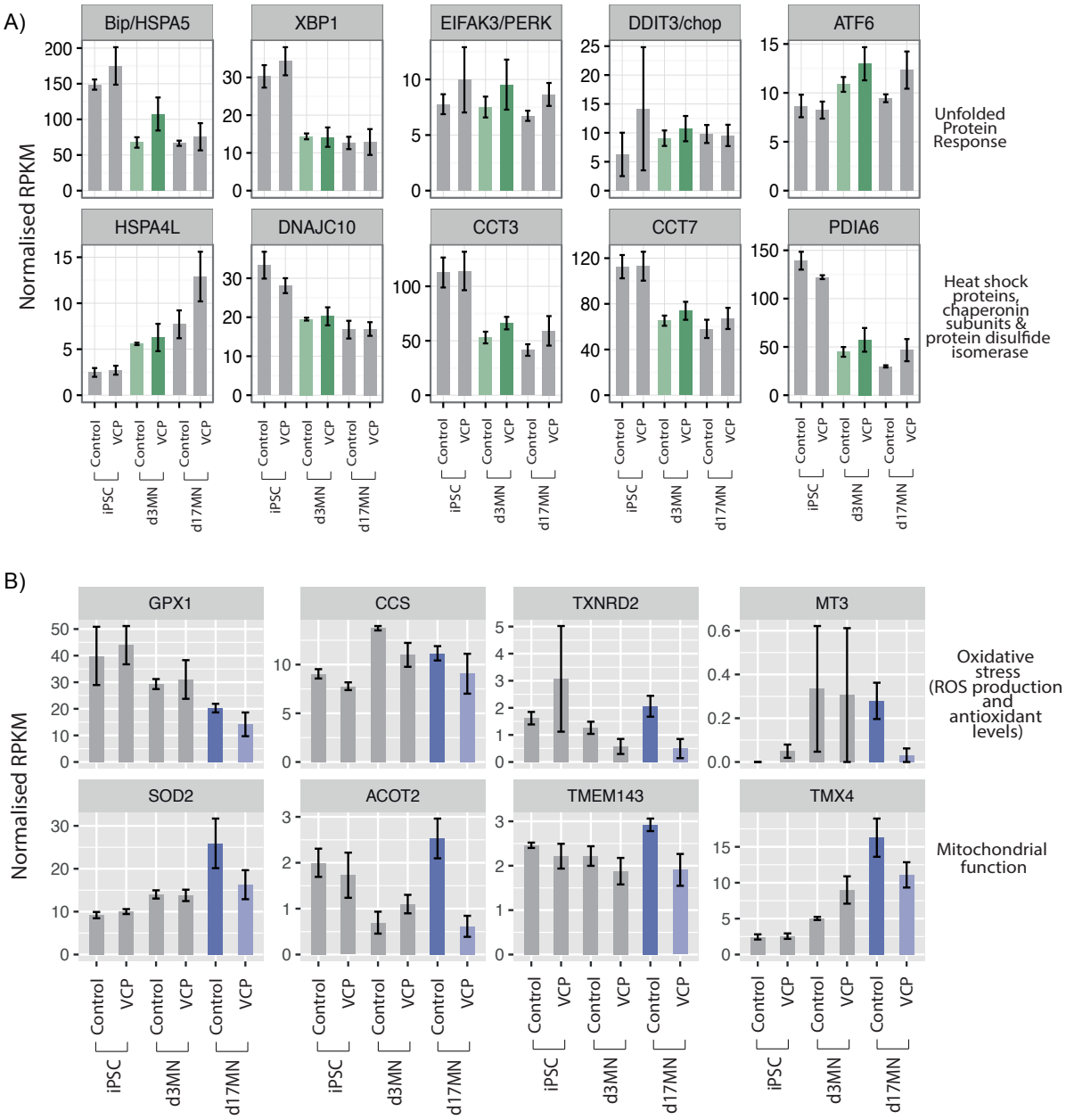


Figure 5.11; Cytoplasmic ER stress response in VCP mutant MNs occurs before gene expression changes but Mitochondrial dysfunction and oxidative stress in VCP MNs are not preceded by coordinate gene expression changes.

A) Gene expression in control and VCP mutant samples analysed across 3 time points in motor neurogenesis (iPSC, D3MN, D17MN). Normalized RPKM values of representative genes associated with ER stress displayed; unfolded response pathway (Bip, XBP1, PERK, Chop, ATF6), heat shock proteins (HSPA4L, DNAJC10), chaperonin subunits(CCT3, CCT7) and protein disulfide isomerases (PDIA6). d3MN is highlighted to represent the onset of cytoplasmic ER stress in VCP mutant MNs. N=3+

technical repeats across a minimum of 2 biological repeats. Error bars represent mean ± SEM. B) Gene expression in control and VCP mutant samples analyzed across 3 time points in motor neurogenesis (iPSC, D3MN, D17MN). RPKM values of representative genes associated with mitochondrial dysfunction (SOD2, ACOT2, TMEM143, TMX4) and oxidative stress (GPX1, MT3, TXNRD2, CCS). d17MN is highlighted to represent the onset of cytoplasmic mitochondrial dysfunction and Oxidative stress in VCP mutant MNs. N=3+ technical repeats across a minimum of 2 biological repeats. Error bars represent mean ± SEM.

Moreover we next analyze the expression of several markers for the later onset phenotypes, including mitochondrial dysfunction, oxidative stress and synaptic defects. Again for mitochondrial dysfunction and oxidative stress we detect no transcriptional differences between control and VCP mutant MNs (Figure 5.11.B). Conversely, we see a clear down regulation of several genes related to synaptic function in VCP mutant MNs including genes encoding voltage gated Ca+ channels, delayed rectifier and inward rectifier K+ channels, Na+ channels, glutamate receptors and genes associated with synaptic density (Figure 5.12). These transcriptional changes are detected at d17MNs suggesting that these are secondary pathological changes in VCP mutant MNs which are indeed orchestrated, at least in part, by their cognate transcriptional programs.

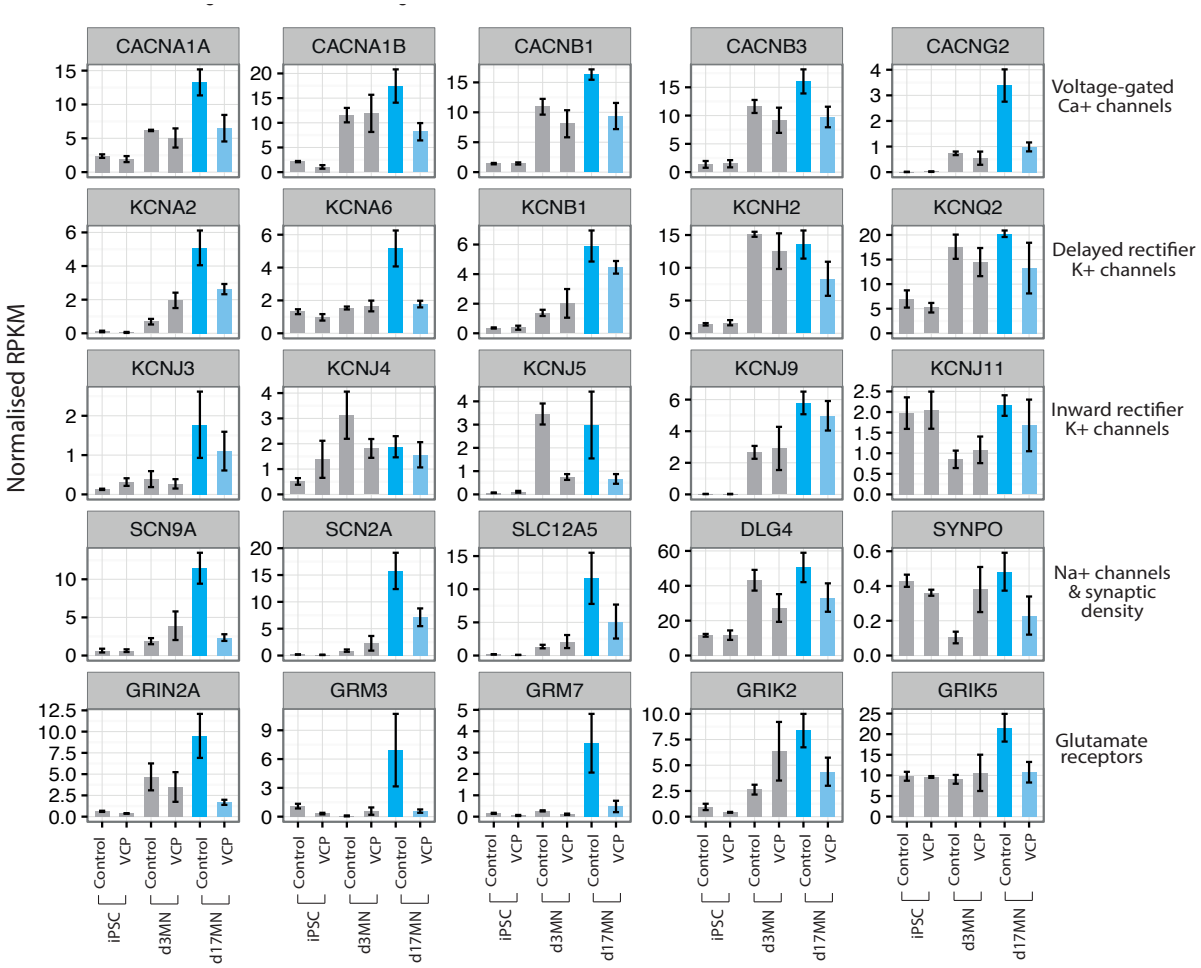


Figure 5.12; The transcriptional signature of VCP mutant MNs reflects synaptic dysfunction.

Gene expression in control and VCP mutant samples analysed across 3 time points in motor neurogenesis (iPSC, D3MN, D17MN). Representative normalised RPKM values of genes encoding selected synaptic components (CACNB1, CACNB3 & CACNG2), K+ delayed rectifier channels (KCNA2, KCNA6, KCNB1, KCNH2 & KCNQ2), K+ inward rectifier channels (KCNJ3, KCNJ4, KCNJ5, KCNJ9 & KCNJ11), Na+ channels (SCN9A, SCN2A & SLC12A5), Synaptic density (DLG4 & SYNPO), Glutamate receptors (GRIN2A, GRM3, GRM7, GRIK2 & GRIK5). d17MN is highlighted to represent the onset of cytoplasmic mitochondrial dysfunction and Oxidative stress in VCP mutant MNs. N=3+ technical repeats across a minimum of 2 biological repeats. Error bars represent mean ± SEM.

Overall we conclude that cytoplasmic ER stress response is a primary event in VCP mutant MNs that precedes transcriptional changes, but that gene expression is seen to contributes to secondary pathogenic events such as synaptic dysfunction. Our findings underscore the importance of our study in

detecting the earliest VCP mutation-related molecular events in ALS using hiPSCs.

Attempts to reverse VCP-mutant phenotype in iPSC-derived MNs by manipulating ER stress pathway

There have been many attempts to reverse the phenotypes seen in MND, including many using pharmacological treatments. Several have demonstrated an improved phenotype, for example UPR manipulators have been explored. Firstly the perk inhibitor (GSK2656157) which reinitiates translation during UPR(Moreno et al. 2013), Salubrinal which promotes translational arrest by inhibiting the phosphorylation of eIF2a(Kiskinis et al. 2014) or TUDCA a molecule that acts upstream of Perk.

Since the earliest molecular event we have detected in VCP mutant MNs is an increase in ER stress we have tried to rescue VCP MNs by targeting the ER stress pathway using a perki inhibitor (GSK2656157) which reinitiates translation during UPR) (Moreno et al. 2013) and Salubrinal which promotes translational arrest(Kiskinis et al. 2014). For each treatment 2 different concentrations were trialed rationalized by previous studies (Perk at 1μM or 10μM and Salubrinal at 50μM or 100μM), and the VCP mutant cell death phenotype monitored (Figure 5.13.A & B). I treated mature MN cultures for 48hours (d15-d17MNs) and cell viability was then analyzed by Zhi Yao using an automated florescence microscope. At this stage we do not see any improvements in cell viability with either treatment, however we cannot rule out Perk or Salubrinal as potentially being viable candidates to modulate the phenotypes described in VCP mutant MNs yet as there are many variables involved and that still to be explored, such as dosage and timing of administration and treatment. I also note we observe a high variability in the data emphasizing the need for increased biological replicates to be used where possible in the future. Moreover we observe a slight trend that treatments increase cell death in the control MNs, which is an important control step that must also be addressed.

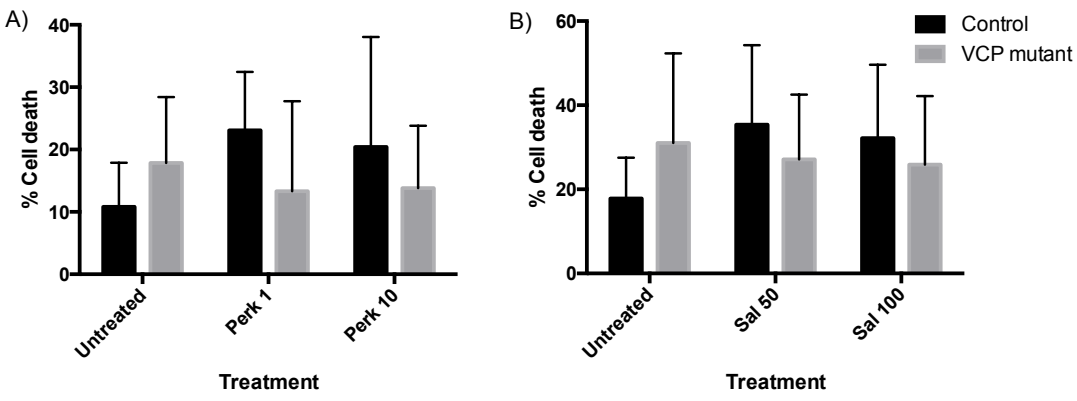


Figure 5.13; Attempt to rescue cell death in VCP MNs by manipulating ER stress pathways

% Cell death of Control and VCP mutant D17MNs shown after 48hours treatment with A) Perk, at 1μM or 10μM and B) Salubrinal, at 50μM or 100μM. N=3+ technical repeats across a minimum of 2 biological repeats. Error bars represent mean ± SEM.

Discussion

Here I have begun to dissect cell specific mechanisms of disease in VCP-related MND, using patient specific hiPSCs differentiated into MNs and ACs. Our efficient differentiation strategies for MN and AC, that I previously optimized, permitted a range of careful cellular phenotyping assays and RNA sequencing technology to systematically identify early VCP mutation-dependent and MN selective phenotypes. We have characterized the presence of a range of phenotypes detected in mutant MNs, including mitochondrial dysfunction, oxidative stress, ER stress, synaptic dysfunction, changes in gene expression levels and cell death.

Analysis throughout the differentiation of motor neurons has so far revealed that there is no differential vulnerability at earlier stages of MN differentiation (prior to NPCs) but upon terminal differentiation a robust phenotype is identified. To date the order of pathogenic events in MND has not been addressed. Using patient specific hiPSCs enabled us to model the timeline of events in VCP-related MND and we demonstrate that ER stress is a primary event in VCP mutant MNs. Further I confirm that it is not preceded by transcriptional deregulation within this pathway. These results build upon several other recent studies that express ER stress is an important event in MND. ER-stress related genes were shown to be upregulated in MNs in fMND mice before symptomatic stage (Saxena, Cabuy, and Caroni 2009) and also

noted in MNs derived from patient specific iPSCs (Kiskinis et al. 2014). In addition several studies link ER stress and TDP43 mislocalisation, the pathological hallmark of MND. Cytoplasmic aggregates of TDP43 are known to induce ER stress in ALS, while ER stress itself has also been proposed to drive cytoplasmic TDP43 mislocalisation (S Sasaki 2010; Walker et al. 2013). At the moment we need further work to establish if TDP34 pathology can be recapitulated in our model of MND (see chapter 6 for further details) and therefore which of TDP43 mislocalization and ER stress is primary.

Following ER stress we detect a deleterious spiral of events, mutation-dependent cell viability, oxidative stress, mitochondrial dysfunction and reduced synaptic density. These pathways when all activated together can reinforce each other, causing a cascade of damage. ER stress has been shown to lead to mitochondria dysfunction and also to promote oxidative stress(Ilieva et al. 2007; Malhotra and Kaufman 2011). The ER is known to cooperate with the mitochondria to maintain cellular homeostasis through both structural and functional interactions. They work together to determine cellular fate under various pathophysiological conditions, for example pro-death pathways are regulated by Ca+ signalling between ER to mitochondria (Malhotra and Kaufman 2011). Here we show ER-mitochondria interactions are increased following ER stress in VCP-mutant D17MNs, which also coincides with when we observe mitochondria dysfunction. Mitochondrial dysfunction is also tightly linked to oxidative stress; a combination of reactive oxygen species (ROS) production and mitochondrial dysfunction can cause widespread protein oxidative damage, which may further overwhelm and impair cellular protein clearance pathways (Sitte et al. 1999). Chronic ER stress can in turn lead to cell death by mitochondria-dependent or independent mechanisms (Lindholm, Wootz, and Korhonen 2006; Malhotra and Kaufman 2011). Further work is needed to unravel the communication between the ER and mitochondria and oxidative stress pathways in VCP-related MND and the mechanism by which this causes cell death.

Moreover I present a VCP mutation-dependent reduction in SYT1 density, which could represent perturbation of pre-synaptic MN terminals, MN autosynapses, interneuronal interactions or reflect post-synaptic pathology in MNs. This observed defect is supported by alterations in gene expression of many synaptic genes, ion channels and glutamate receptors. So far we have

not investigated the functional consequences of reduced synaptic density and altered synaptic gene expression in VCP mutant MNs. Current evidence from other MND studies presents basal MN hyperexcitability in several forms of familial MND (Devlin et al. 2015; Kiskinis et al. 2014). Further studies implicates that excessive glutamatergic synaptic activity in MND leads to an increase in Ca+ influx and cell death. Studies show that glutamatergic activity-induced calcium influx is too much for the mitochondria and ER to handle resulting in ER stress or increased mitochondrial calcium uptake, depolarization and ROS production, therefore further enhances a downward cascade of events (Howland et al. 2002; Shaw et al. 1995). Additionally a link between ER stress and electrical activity has been demonstrated, suggesting that the burden of ER stress makes the VCP mutant MNs more susceptible to synaptic dysfunction (Kiskinis et al. 2014).

Since we observed ER stress as an early pathogenic event in VCP-mutant MNs we began to test if the VCP mutant cell viability phenotype can be reversed by manipulating ER stress pathways. So far we found no improvement in cell viability after treatment with perk or salubrinal but much further testing needs to be carried out, including altering dosage and timing, before conclusions can be drawn. In addition an array of additional modulators will be tested including retigabine (Wainger, Kiskinis, Mellin, Wiskow, Han, et al. 2014), the VCP inhibitor vesnarinione shown to manipulate the NfKb pathway(Hotta et al. 2013), NRF2 activators shown to promote anti-oxidant defences(Gupta et al. 2012) and treatment with stains, since Simvastatin has been shown to promotes mitochondrial function (Chataway 2015). It is important to bear in mind that treatment with such compounds could not only have an effect on the cellular phenotype but also could influence cell-subtype diversity (Dalton, Lyons, and Lomvardas 2013) so this should also be monitored.

Finally we observe that molecular events in VCP mutant motor neurons and astrocytes are different. So far we find no substantial evidence of VCP-mutation related AC pathology, suggesting the VCP-mutant phenotype is specific to MNs. This is consistent with a recent study that also presents a neuronal but not AC mutation-dependent phenotype in mutant SOD1 iPSC-derived MNs (Kiskinis et al. 2014). Conversely, a study exploring *TARDBP* mutations has observed a cell autonomous AC pathology (Serio et al. 2013) in

addition to MN pathology (Bilican et al. 2012), which leads us to propose there could be mutation-specific glial contributions in the pathogenesis of MND. Lastly, our findings do not exclude a non-cell autonomous role for VCP mutant ACs as they could exert deleterious effects on spinal MNs, as in the case of SOD1 ACs (Di Giorgio et al. 2008). We are currently investigating this hypothesis.

As expected, throughout this work we have seen variation between cell lines. This highlights the desirability to use isogenic lines in future studies. We are currently replicating all work with an additional Control and VCP biological lines and are working towards generating isogenic lines by both introducing an R155C VCP mutation into a control line and correcting an R155C mutation in a disease causing line.

To conclude, the in depth phenotyping of VCP-mutant hiPSC lines has provided important insights into the mechanisms underlying VCP-related MND, which has showed considerable over lap with previous studies in other familial and sporadic MND models. Further this work underscores the value of cellular phenotyping at different stages of iPSC lineage restriction to selected cell types of interest (MNs and ACs) to uniquely discriminate primary from secondary pathogenic events in MND.

Chapter 6; Characterizing TDP43 in VCP related MND

Introduction: The RNA binding protein TDP43

TAR DNA-binding protein (TDP43), molecular weight 43kDa, is classified as a heterogeneous ribonucleoprotein and is encoded by the *TARDBP* gene. It is a ubiquitously expressed, highly conserved protein and its features include 2 RNA recognition motifs (RRM), two NLS at N-terminus, NES within RRM2 and a glycine-rich c-terminus (Gendron, Josephs, and Petrucelli 2010). TDP43 is a versatile protein that binds to both DNA and RNA (Buratti and Baralle 2001; Gendron, Josephs, and Petrucelli 2010; Tollervey et al. 2011).

TDP43 functions to control different stages of protein production through its interactions with RNA and DNA. It acts at many stages during an RNA life cycle; including repression of transcription, RNA stability, alternative polyadenylation and alternative splicing (I.J.Huppertz 2015; E. B. Lee, Lee, and Trojanowski 2012). Hence it is no surprise that the majority of TDP43 is localized to the nuclear compartment.

In addition, TDP43 has been found to bind to specific mRNA's 3'UTRs (including its own) (Ayala et al. 2011; Tollervey et al. 2011) and has been shown to shuttle between the nuclear and cytoplasmic compartments (Ayala et al. 2008). Cytoplasmic TDP43 is particularly evident under cellular stress and up to 30% of TDP43 can be localized to the cytoplasmic compartment (Scotter, Chen, and Shaw 2015). Further it has been shown, using a variety of stressors (heat shock, ER, osmotic, and oxidative stress) and cell lines, that TDP43 co-localises selectively with SGs (Colombrita et al. 2009; Dewey et al. 2010; Walker et al. 2013). Also similar to other hnRNP's, TDP43 is involved in the formation of granules by forming protein-RNA complexes that sequester mRNA's unneccessary for survival (Vanderweyde et al. 2013). This function of TDP43 emphasizing the importance of TDP43's role under cellular stress.

The RNA binding properties of TDP43 were first explored in 2001 where two distant RRM domains were described and TDP43 was seen to bind with increasing affinity to UG repeats >6nt long (Buratti and Baralle 2001). A decade later TDP43 binding targets were characterized on a genome wide scale in the

brain using iCLIP (Tollervey et al. 2011). Consistent with previous studies TDP43 showed preferential binding to UG tandem repeats or long clusters of UG-rich motifs and the majority of binding targets were found to be within intronic sites, representing the nuclear localization of TDP43.

Another important function of TDP43, dependent on RNA binding, is to regulate its own protein levels. If TDP43 protein levels are too high they are toxic to cells but a lack of TDP43 is lethal to cells (Kraemer et al. 2010; Tsai et al. 2010). When cellular TDP43 is abundant it binds to its own 3'UTR resulting in alternative polyA site selection, which activates the use of a usually silent intron in its 3'UTR, followed by rapid degradation of TDP43 mRNA (Ayala et al. 2011). A loss of this regulation is implicated in human neuropathology where TDP43 protein levels are seen to be increased (this is discussed further below).

In addition, several pieces of evidence suggest that TDP43 plays a particular role in neuronal cells. Firstly TDP43 has been discovered to regulate the alternative splicing of genes implicated in neuronal development (DLC1, FZ3, KIF2A, KIF1B, AP2, TLE1, TNIK, UNC5C) and neuronal survival (CNTFR, MEF2D, MADD) (Tollervey et al. 2011). On top of this, TDP43 forms cytoplasmic mRNP granules that undergo bidirectional, microtubule-dependent transport in neurons (Alami et al. 2014). This enables the delivery of target mRNA to distal neuronal compartments and facilitates TDP43 to locally regulate RNA processing. Together this permits neuronal plasticity and development (dendritic branching, axonal growth). Finally, aberrant TDP43 mislocalisation and aggregation is at the forefront of many neuronal diseases highlighting the importance of TDP43 functions within neurons (Gendron, Josephs, and Petrucelli 2010; J. O. Johnson et al. 2010; E. B. Lee, Lee, and Trojanowski 2012).

TDP43 in MND

TDP43 pathology is defined by cytoplasmic localization, aggregation, fragmentation and post-translational modifications, namely phosphorylation and ubiquitination (Manuela Neumann et al. 2006; Scotter, Chen, and Shaw 2015). Nuclear inclusions are also sometimes seen (Arai et al. 2006; Manuela Neumann et al. 2006). Collectively this is recognized as the pathological hallmark of MND (detected in >95% of patients), which since discovered in 2006 sparked great interest in TDP43 pathology and has subsequently been

detected in other relative neurodegenerative disorders (Geser et al. 2009). These are collectively known as TDP43 proteinopathies (Liscica et al. 2008). Additionally TDP43 pathology is increasingly seen in normal aging population (over 65) (Geser et al. 2011) suggesting a broader role for TDP43 in aging and disease.

The role of TDP43 pathology in disease pathogenesis is controversial and the features of TDP43 pathology have been extensive examined. Hyperphosphorylation of TDP43 is a recognized featured in disease (Arai et al. 2006; Hasegawa, Arai, and Nonaka 2008; Manuela Neumann et al. 2006) but the role of phosphorylation is not established. One group hypothesized phosphorylation is a protective mechanism but others have shown it is required for toxicity and/or suggests that the phosphorylation negatively impacts TDP43's degradation (H. Y. Li et al. 2011; Liachko, Guthrie, and Kraemer 2011; Y.-J. Zhang et al. 2010). Moreover C-terminal TDP43 fragments are seen to be a major constituent of cytoplasmic inclusions owing to the presence of low complexity domain (responsible for protein interactions) and the removal of NLS enhances cytoplasmic distribution (I.J.Huppertz 2015; Nonaka et al. 2009; Yong-jie Zhang et al. 2009). However TDP43 C-terminal fragmentation has not been shown to directly exert toxicity and is thought to be a secondary effect. Neuronal toxicity is associated with TDP43 mislocalisation, which can be driven by *TARDBP* mutations (Barmada et al. 2010), cellular stress (Colombrita et al. 2009; Dewey et al. 2010) or impaired degradation (van Eersel et al. 2011). Similarly the cleavage of C-terminal TDP43 fragments is influenced by cellular stress, proteasomal inhibition and mutations in C-terminal domain of *TARDBP* (Scotter, Chen, and Shaw 2015). Additional biochemical analysis in disease has shown that aggregated TDP43 is ubiquitinated, making it relatively insoluble (Hasegawa, Arai, and Nonaka 2008).

Even though TDP43 pathology is seen in the majority of patients with MND, mutations in the *TARDBP* gene are only responsible for 2% of cases (Scotter, Chen, and Shaw 2015). All disease-causing mutations discovered in *TARDBP* are found within the highly conserved glycine rich C-terminus (exon 6), with except to one FTLD/MND patient discovered to have the missense mutation A90V, which is located in between the two NLS (Winton et al. 2008). Additional mutations have been found in the non-coding region but these are thought to be benign. Many studies have used the discovery of a disease–causing

mutation hot-spot in TDP43 to serve as a platform for investigating pathobiology for TDP43; C terminus is involved in protein interactions and influences solubility and cellular localization (Ambrogio et al. 2009; Ayala et al. 2008). In addition mutations have been discovered that affect phosphorylation of TDP43; 3 mutations have been found that could abolish phosphorylation sites but on other hand 9 mutations could enhance phosphorylation which in turn could augment TDP43 aggregation (Corrado et al. 2009; Kabashi et al. 2008; Kühnlein et al. 2009).

Although there has been extensive research of TDP43, its the role in disease is still not fully understood. It has been debated whether TDP43 in MND acts via a loss of function mechanism, mediated by nuclear clearance, RNA dysregulation or defective protein clearance, and/or gain of function mediated by aggregates or abnormal cytoplasmic function (E. B. Lee, Lee, and Trojanowski 2012). In a wider context, aggregation is seen across neurodegeneration suggesting that it plays a central role in neurodegenerative disease pathology. However it is still not clear what role inclusions play in disease pathogenesis; 1/ Are they the primary causative toxic species? Or 2/ are they just an innocent bystander (that occurs with another toxic event but themselves are not critical) or 3/ are they a protective mechanism to prevent cellular damage from misfolded/toxic proteins?

TDP43 in VCP-related MND

As touched on earlier, previous models of VCP related MND have recapitulated the pathological hallmark of MND (J. O. Johnson et al. 2010; Ju et al. 2009; M Neumann et al. 2007), suggesting VCP-related MND could share common mechanisms of pathogenesis with other TDP'opathies. In addition another study has explored the genetic interaction of TDP43 and VCP; in drosophila mutant VCP was shown to genetically interact with TDP43 resulting in an enhancement of the pathological phenotype (Ritson et al. 2010). The interaction was enhanced by the presence of a mutation in NLS of TDP43 and abolished by restricting TDP43 expression to the nucleus (Ritson et al. 2010). The findings presented imply that TDP43 is a mediator of toxicity initiated by disease-causing mutations in VCP rather than just an indicator of cytoplasm stress caused by disease but the mechanism underlying the interaction between TDP43 and VCP in disease is not fully understood.

Although key steps have been made, characterizing the role of TDP43 in MND is far from complete. Here I have studied TDP43 in VCP-related MND and begun to unravel the effect of VCP mutation on TDP43 RNA interactions.

Modeling TDP43 proteinopathies

To date there has been an array of TDP43 models generated that have all provided novel insights into the pathogenesis of TDP43-proteinopathies, including MND (Ayala et al. 2008; Bilican et al. 2012; B. S. Johnson et al. 2008, 2009; Kraemer et al. 2010; Serio et al. 2013; Winton et al. 2009). However there is unfortunately a lack of models in the field that have recapitulated fully TDP43 pathology we see in MND; phosphorylation, ubiquitination, cleavage, cytoplasmic mislocalisation and aggregation, which could contribute to the lack of translated research in MND.

Firstly yeast models have been an attractive system to use because they are easily manipulated (genetically). Both purified TDP43 and TDP43 overexpression in yeast results in TDP43 aggregation, which was seen to be dependent on TDP43's C-terminal domain (B. S. Johnson et al. 2008, 2009). However TDP43 C-terminal mutants are not shown to aggregate and remain in the nucleus (B. S. Johnson et al. 2008). This discrepency could be due to the availability of different binding partners influencing the transport system in yeast. Since some features of MND are recapitulated yeast models have been exploited to test modulators of TDP43 toxicity and aggregation, for example Pbp1 (orthologue of ataxin-2) was shown to enhance toxicity providing potential novel therapeutic targets (Elden et al. 2010). Overall the yeast system has provided indisputably valuable mechanistic insights into TDP43 pathology but it us unable to provide complexity introduced by the specialised neuronal cell-signalling networks.

Next studying endogenous or overexpressed TDP43 in mammalian tissue culture models has not recapitulated aggregation as yeast did; TDP43 remains soluble and predominantly localised to the nucleus (Baloh 2011; Winton et al. 2009). But TDP43 has been manipulated to recapitulate pathology; TDP43 mutations have targeted the nuclear localisation signal (Winton et al. 2009), RRM (Ayala et al. 2008), or C-terminal truncation (Nonaka et al. 2009), suggesting that altered TDP43 RNA binding, nuclear localization, or proteolytic cleavage can promote TDP43 aggregation.

Further, numerous animal models have been utilised to study TDP43. Transgenic, virus induced and KO mammalian models have all been generated. One example is the transgenic mouse with A315T mutation that shows ubiquitinated inclusions and c-terminal fragmentation however fails to show cytoplasmic mislocalisation (Wegorzewska et al. 2009). Other rodent models have recaptitulated cytoplasmic mislocalisation but observe insoluble TDP43 is only a minor component of pathological inclusions present (Baloh 2011). Moreover KO models were generated to address the hypothesis that TDP43 acts via a loss of function in disease, however TDP43 KO models proved embryonic lethal (Kraemer et al. 2010). To avoid this conditional KO's were generated but they do not show striking MND symptoms and present with defective body-fat metabolism (Chiang et al. 2010). In summary each model differs and recapitulates different features of typical TDP43-proteinopathies but one common feature across species (yeast, rats, mice, worms, flies and zebrafish) is that TDP43 overexpression is toxic. A more detail comparison of TDP43 animal models are reviewed elsewhere (Baloh 2011; Wegorzewska and Baloh 2011).

Together this highlights the need for novel improved model systems to study MND and TDP43-proteinopathy. hiPSCs provide a unique advantage in that they enable investigation in a human system and that mutations are studied at their pathological dose.

Aims

Here I begin to uncover the role of the RBP TDP43 in the pathogenesis of VCP-related MND, using a humanized clinically relevant model system. Firstly I address the localization of TDP43 in control and VCP-mutant hiPSC neural derivates. Next, I continue to explore the possibility of defective RNA regulation in VCP-related MND by studying Protein-RNA interactions of the candidate RBP TDP43 in hiPSC-derived patient specific MNs. I addressed this by using a technique developed in the Ule lab, Individual nucleotide crosslinking and immunoprecipitation (iCLIP), to explore whether the composition of protein-RNA complexes are changed in VCP-related MND.

Characterizing TDP43 localization in VCP hiPSC neural derivatives

Previous studies using both in-vitro and in-vivo approaches have demonstrated TDP43 pathology in VCP related MND. To explore this pathological hallmark in our VCP mutant cultures, we examined the subcellular localization of TDP43 in patient specific hiPSC-neural derivatives.

Moreover, several studies have linked TDP43 pathology with ER stress (S Sasaki 2010; Walker et al. 2013). Earlier we have shown that ER stress is an early pathogenic event in VCP mutant MNs that manifests shortly after terminal differentiation. Thus I reasoned that ER stress is associated with TDP43 mislocalisation so I have chosen to explore TDP43 pathology at flanking time-points of to this ER phenotype; in NPCs and D3MNs.

TDP43 localisation in VCP hiPSC neural derivatives

TDP43 staining in control NPCs shows 97.1±0.8% localization to the nucleus and 2.9±0.8% in the cytoplasm (n=2, from 1 biological line), compared to 94.8±0.3% within the nucleus and 5.2±0.3% in the cytoplasm in VCP-mutant NPCs (n=4, across 3 biological lines) (Figure 6.1.A & C). (Ungrouped data for each biological line is displayed in Appendix 8.5)

Shortly after terminal differentiation in D3MNs, control cultures display 90.0±1.5% TDP43 localization in the nucleus and 10.0±1.5% in the cytoplasm (n=3 across 1 biological lines). Where-as the percentage of nuclear TDP43 in VCP-mutant cultures is reduced to 75.7±4.6% and the cytoplasmic TDP43 is reciprocally increased to 24.3±4.6% (n=4, across 2 patient lines) (Figure 6.1.B & D). Thus revealing a significant difference in TDP43 localization between control and VCP-mutant cultures in D3MNs but not at the NPC stage.

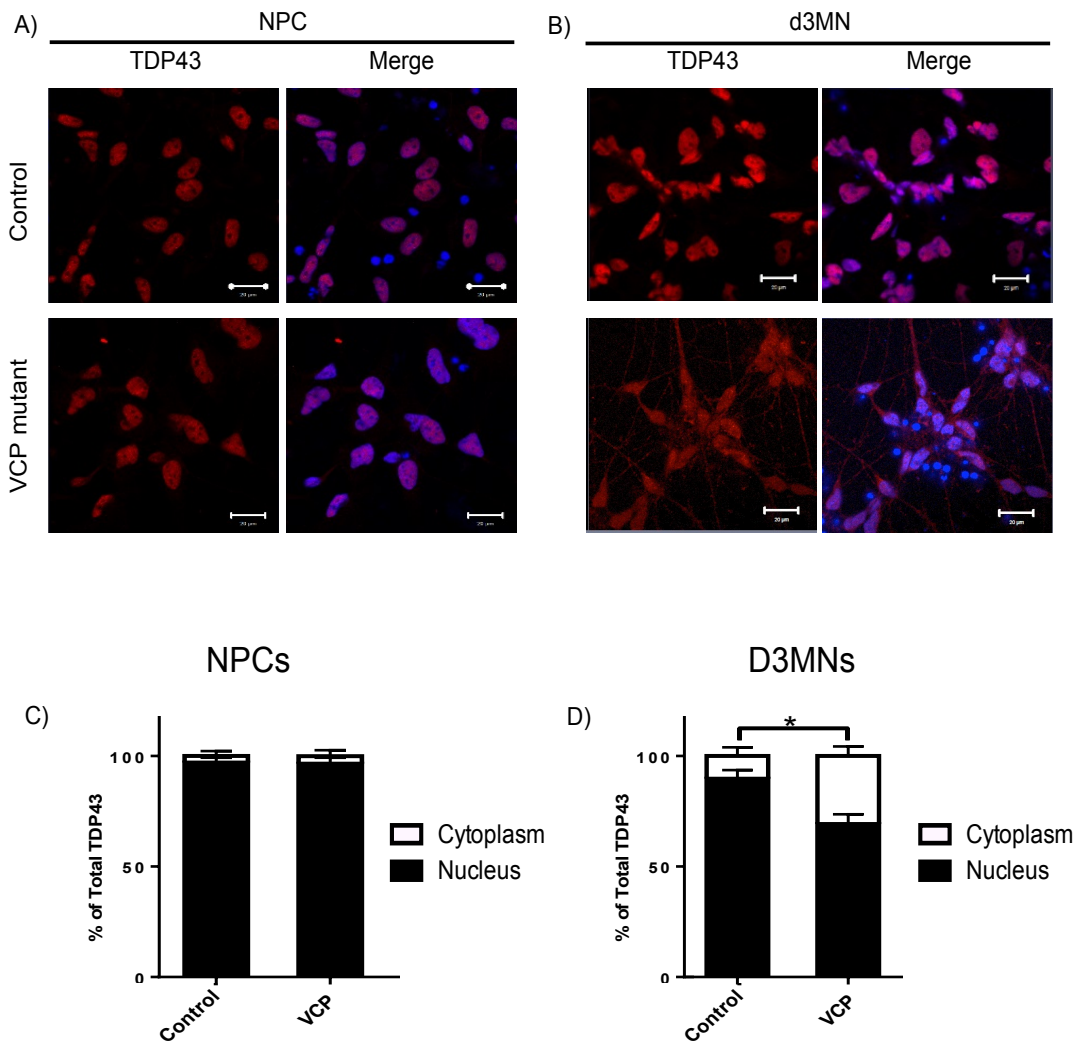


Figure 6.1; TDP43 staining in hiPSC-derived NPCs and MNs

A & B) Representative images of Control and VCP-mutant NPCs and MNs respectively, stained with TDP43 (red) and merged with DAPI staining (blue). Scale bar: 20µm. C & D) quantification of TDP43 staining, displayed as % of TDP43 in the nucleus vs cytoplasm. (N=3 technical replicates from biological n=1 for control lines and n=2 for VCP mutant lines, 50+ cells counted per condition, two-way ANOVA). Error bars represent mean ± SEM. *p<0.05.

TDP43 iCLIP in VCP hiPSC derived Motor neurons

To date, RNA regulation of TDP43 has been studied in the human brain and in the context of FTLD (Tollervey et al. 2011). iCLIP was carried out on human cortical post mortem tissue from 3 healthy control patients and 3 patients with FTLD (another TDP43-proteinopathy). TDP43 mRNA targets were identified on a genome wide scale in human neuronal tissue. This sites were then compared to those found in tissue from FTLD patients; the greatest number of changes

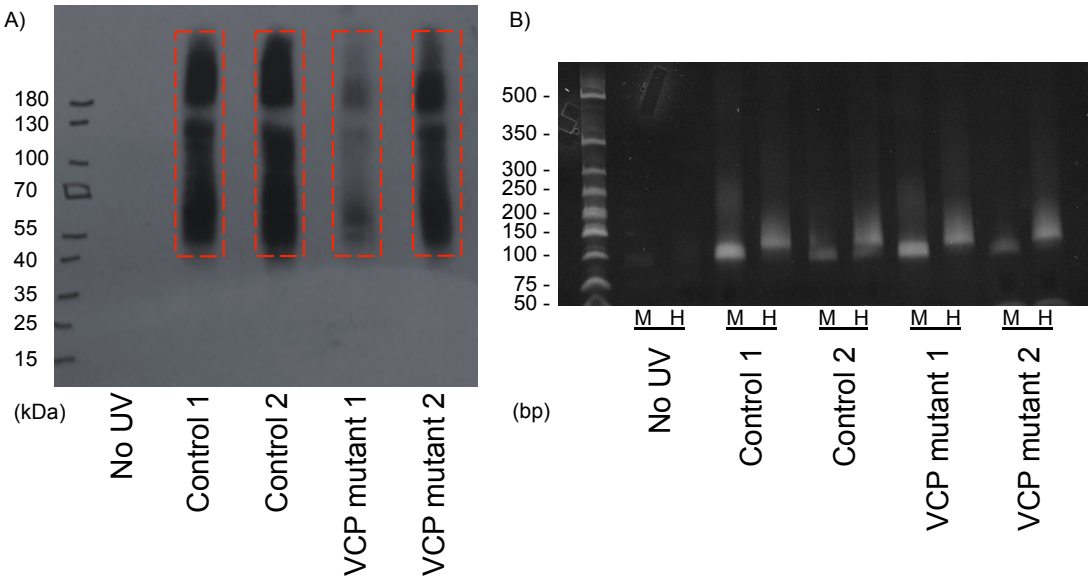
were found within intronic regions, which suggests the largest changes in TDP43 functions in FTLD are in the nucleus. Although this provides key insights into TDP43's role under both control and disease conditions it is difficult to interpret cell specific functions due to a mixed sample population. Further, in combination with microarray experiments in SHSY5Y cells upon TDP43 KD, sites of alternative splicing in reduced TDP43 expression were identified. These results cannot be confidently interpreted in relation to disease yet as it is unresolved if TDP43 acts via GOF or LOF mechanisms in MND.

TDP43's RNA interactions have not been explored selectively in MNs, the cells most vulnerable in MND or on a genome wide scale in VCP-related MND. This is now possible due to optimized and highly pure differentiation protocols using patient specific hiPSCs and advancements in functional genomic techniques (both discussed in previous chapters). Here, I have carried out a focused high-throughput iCLIP experiment using the candidate RBP, TDP43 in human iPSC-derived MNs. Further, to probe the role of TDP43 in MND I have used patient specific VCP-mutant hiPSC-derived MNs to address whether TDP43-RNA complexes are altered in VCP-related MND. I carried out this first experiment on both control and VCP-mutant MNs at D17 MNs, as this is after the onset of a robust cell-type and mutation specific phenotype in VCP mutant MNs (see Chapter 5 for details).

Library preparation

Library preparation for CLIP protocols require a large number of enzymatic steps which all potentially affect the binding site detection. For this reason it is crucial to optimize the experimental conditions first. Within the lab we have recently refined each step of the iCLIP protocol (Huppertz et al. 2014) and on top of this, I have optimized the experiment specific conditions, such as the antibody of use. Here I have firstly used Proteintech TDP43 10782-2-AP antibody as previous studies in the lab (I.J.Huppertz 2015; Tollervey et al. 2011). I next tested the partial RNase digestion conditions. This is importance to test for each specific experiment to ensure enzyme activity has not changed (overdigestion can decrease the number of identified sites). This can be observed on radioactive gel after immunopreciptation, RNase treatment and linker ligation (Figure 6.2.A). Additionally, the visualization of the immunoprecipitated RBP-RNA complexes helped to purify unbound RNA, enabled me to check the size of target RBP-RNA complexes correlates to

expected protein weight (TDP43 to be approx. 43kDa) and it allowed me to see the relative amount of RNA pulled down between conditions (by looking at the signal intensity). After 1 hour I noticed there was less RNA pulled down in the VCP-mutant MN sample 1, but after longer exposure complexes were clearly visualized so I continued with all samples noting a reduction in RNA within the one sample. The size of the dominant TDP43-RNA complex present corresponded to a single TDP43 molecule bound to the RNA. (At approximately 120kDa I observe IgG that blocks visualization; this will not impact complexes purified but causes a separation in the visualization. In future experiments I will use a reducing agent to disrupt the IgGs and make them run individually at 25, and 50 kDa). I isolated protein-RNA complexes, bearing in mind to cut slightly higher on the membrane to ensure all RNA-Protein complexes are extracted due to the IgG interference, and continued with the standard iCLIP library preparation method. PCR amplification was performed for 22 cycles across all samples (Figure 6.2.B).



C)

		Total counts	Unique counts	Ratio
Attempt 1	no UV	9798	1211	8.090834021
	Control_1_1	1016250	17361	58.53637463
	Control_1_2	2100350	12051	174.2884408
	Control_2_1	1738300	11743	148.0286128
	VCP mutant_1_1	418788	7020	59.65641026
	VCP mutant_1_2	455443	2057	221.4112786
	VCP mutant_2_1	869843	3369	258.1902642
Attempt 2	Control_1	2552660	66015	38.66787851
	VCP mutant_1	1903230	134946	14.10364146
	Control_2	2485510	121442	20.46664251
	VCP mutant_2	292377	21379	13.67589691
Attempt 3	no UV	15055	5987	2.514614999
	Control_1	2.59E+06	1.31E+06	1.984901126
	Control_2	2.80E+06	1.52E+06	1.845411585
	VCP mutant_1	1.87E+06	915820	2.043633028
	VCP mutant_2	1.34E+06	600687	2.238603466

Figure 6.2; Optimization of TDP43 iCLIP library preparation from hiPSC-derived MNs

A) Autoradiograph showing protein-RNA complexes isolated after TDP43 immunoprecipitation, taken after 1hour exposure. Region isolated for downstream library preparation is annotated by the red dotted lines. B) Post PCR gel (23cycles used). Medium (75-95nt) and High (95-200nt) fractions were obtained by cutting selected regions from cDNA gel. C) Table displays ratio of unique reads, only experiment attempt 3 has been further analyzed in this thesis.

After sequencing I assess the complexity of the libraries by calculating the ratio of unique reads. My first experimental attempts using hiPSC-derived MNs revealed extremely high ratios of unique reads ranging between 58-258, which reflects poor complexity and would contain many duplicate reads due to over

sequencing (Figure 6.2.C). I hypothesized that this could be partially due to the small sample input from hiPSC-derived MNs (mean input was 0.08mg protein) therefore I repeated the experiment using ten times the input material (0.84mg). This did improve the complexity of the library generated, ratio of unique reads ranged from 14-38 (Figure 6.2.C). However I wanted to further improve the data quality before continuing with analysis. We reasoned that the immunoprecipitation could be improved by using an alternative TDP43 antibody (Sigma TDP43 SAB4200008). This was tested in the lab by Martina Hallegger; immunoprecipitations were carried out on HeK293 cells side by side with the previously used Proteintech antibody. The new sigma antibody revealed a much stronger signal on autograph after 1 hour exposure than previous Proteintech antibody, therefore I attempted iCLIP on hiPSC-derived MN for a third time using increased input material and the Sigma antibody. This resulted in a good library complexity and unique ratio scores between 1.8-2.5 (Figure 6.2.C).

TDP43 binding sites in iPSC-derived MN

Firstly I have defined the clusters bound by TDP43 on a genome wide scale using both control and VCP-mutant samples. I then analyzed the correlation between the clusters bound in all samples (biological replicates and vs control and mutant conditions). Strong correlation was seen between all samples (R^2 values ranged between 0.96 to 0.98) (Figure 6.3).

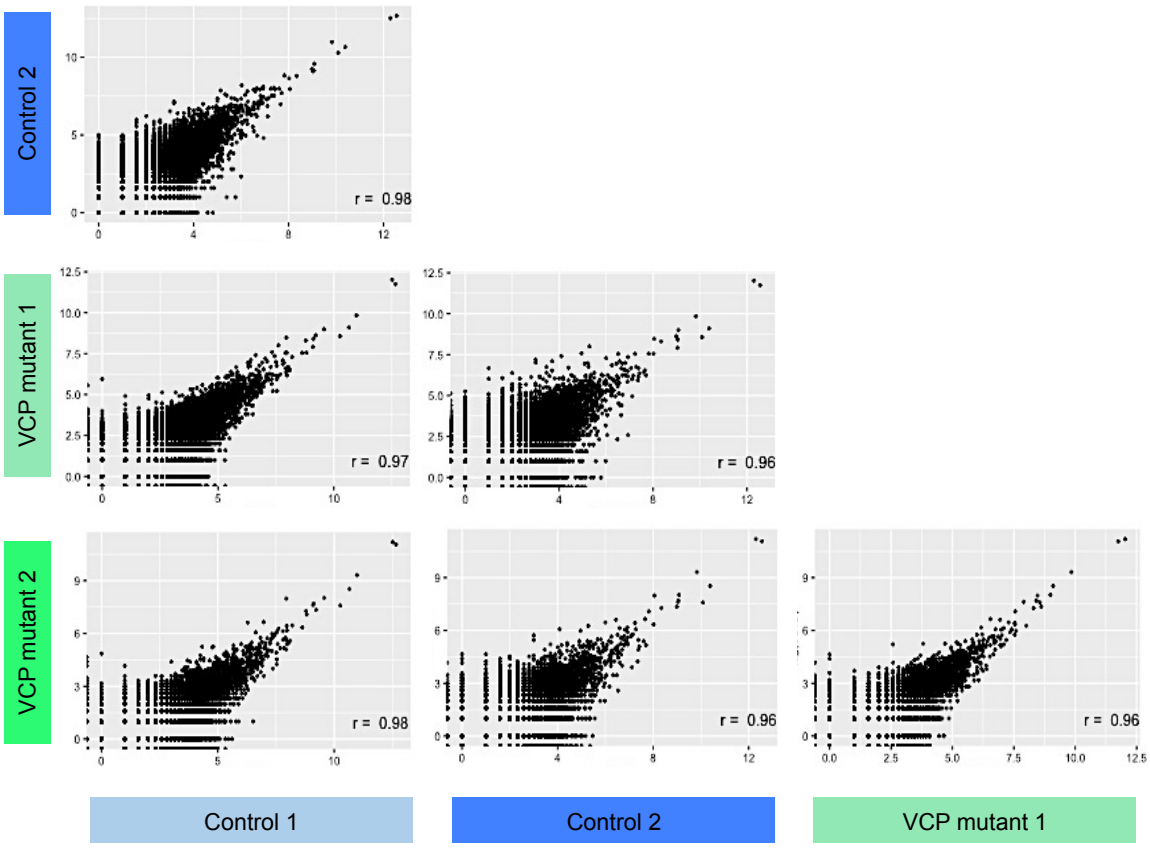


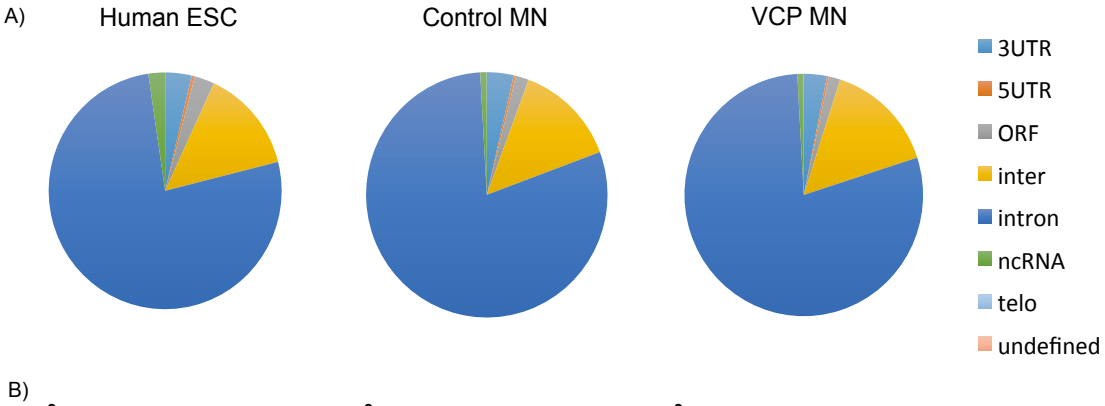
Figure 6.3; Correlation between biological replicates

Clusters were defined based on all samples, then Log(number of reads within each cluster) was plotted for each biological replicate. R^2 value displayed.

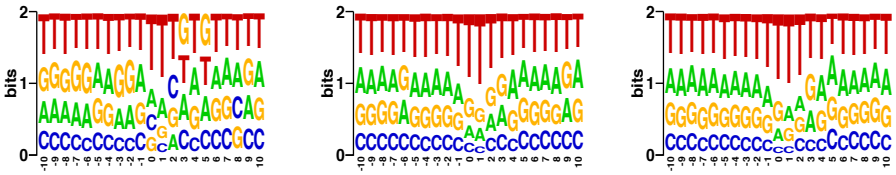
Although we have not seen TDP43 mislocalisation so far in VCP-mutant D3MNs I wanted to explore whether any alteration in protein localization could be detected with an alternative more sensitive technique and at a later stage of MN maturation. Therefore I have evaluated, at the genome-wide level, the % of TDP43 binding to different regions of the genome (3'UTRs, 5'UTRs, ORF, inter and intronic) for both control and VCP mutant groups, and compared my findings to previously published data. This was done using single hits that mapped to genome (Gsumtype files on icount provides the calculation summary of cDNA counts that map to different regions of genome). I find the genomic distribution of TDP43 binding was not altered between the control and VCP mutant MNs and, consistent with previous findings(Tollervey et al. 2011), the highest % of binding was found in introns (Figure 6.4.A). This also confirms the localization of TDP43 is predominantly nuclear (>70%). Since I do not detect TDP43 mislocalization at this point in VCP mutant MNs, this makes the downstream analysis simpler as we do not need to normalize for protein

localization (a change in protein localization would alter the binding targets due to RNA availability).

A)



B)



C)

Pentamer	Enrichment Score		
	hES	Control MNs	VCP mutant MNs
GTGTG	3.41	6.10	6.64
TGTGT	3.05	5.22	5.55
AAAAA	0.73	2.43	1.04
AAAAA	1.01	2.35	1.03
CCTGA	0.88	2.22	0.83
TGCGT	1.89	2.20	2.17
TGTGC	1.40	2.19	1.43
CCTGA	0.85	2.17	0.80
GTATG	1.66	2.17	1.80
GAGCG	1.10	2.15	1.18

Figure 6.4; TDP43-RNA interactions in iPSC-derived MN

A) Genomic distribution of TDP43 binding, B) Weblogo logos displaying the sequence motif around crosslink sites genome wide, and C) Pentamer analysis, in human ESCs (published data), Control hiPSC-derived MNs and VCP-mutant hiPSC-derived MNs.

TDP43 has previously been shown to bind to UG-rich regions in the genome (Buratti and Baralle 2001; Tollervey et al. 2011). To explore whether this finding is consistent in my current data I have first analyzed the proportion of each nucleotide seen surrounding binding sites using weblogo. I have plotted 10nt either side of the crosslink sites and show enrichment of T nucleotides, which is comparable for control, VCP-mutant MNs and to previous TDP43 iCLIP data from hESCs (Figure 6.4.B) (Tollervey et al. 2011). Next I performed further

sequence analysis of TDP43's crosslink sites by analyzing pentamer enrichment around crosslink sites. I captured 30nt either side of each crosslinking site and excluded the region immediately surrounding the site -10 to 10 (due to bias of UV crosslinking nucleotide preferences). I see that in both Control and VCP-mutant MNs TDP43 binds to UG rich sequences (enrichment score = 6.10 for control and 6.64 VCP-mutant MNs) (Figure 6.4.C).

Differential Protein-RNA interactions of TDP43 in VCP mutant MN

Next, I analyzed only the binding sites that were changed between control and VCP mutant MN samples. Using a threshold score of smoothed z>20, there are 261 sites that are more bound in by TDP43 in VCP mutant MN and 463 sites that are less bound (Figure 6.5.A). The top changes in TDP43 binding include; Doublecortin (DCX = a microtubule associated protein involved in neuronal migration), NRG1 (Neuregulin 1 = a membrane glycoprotein involved in cell-cell signaling in growth and development), TSHD7A (thrombospondin type1 domain containing 7A = a neural protein involved in angiogenic patterning) and PHACTR3 (phosphatase and actin regulator 3 = is associate with the nuclear scaffold in proliferating cells), which are more bound in VCP mutant MNs compared to control. On the other side FXYD6 (Phosphohippolin = a transmembrane protein which affects activity of NA,K-ATPase), XPO1 (exportin 1 = a cell-cycle-regulated gene & encodes a protein that mediates leucine-rich NES-dependent protein transport) and STX1B (syntaxin-1B = role in exocytosis of synaptic vesicles) are all less bound by TDP43 in VCP-mutant MNs. It is interesting to note that many of the top changed genes have neuronal specific functions.

Following the detection of differential binding, functional consequences need to be elucidated. I have begun to look at the functional impact of TDP43 binding by combining iCLIP and RNAseq data to explore if differential binding of TDP43 has any effect on RNA expression levels. Here I touch upon two preliminary examples; firstly STX1B is less bound by TDP43 in VCP-mutant MNs and the RNAseq shows a trend for reduced mRNA expression (Figure 6.5.B & C). This suggests that TDP43 acts on STX1B as an enhancer. On the other hand PHACTR3 is more bound by TDP43 but also suggests TDP43 acts in this situation as an enhancer, as our RNAseq data shows a trend towards higher mRNA expression in VCP-mutant cells (Figure 6.5.D & E). Neither of these sites/genes mentioned have significantly different mRNA expression levels in

VCP-mutant MNs compared to control. Further replicates and analysis need to be done to explore the trends observed and to understand the impact of differential TDP43 binding in VCP-mutant MNs.

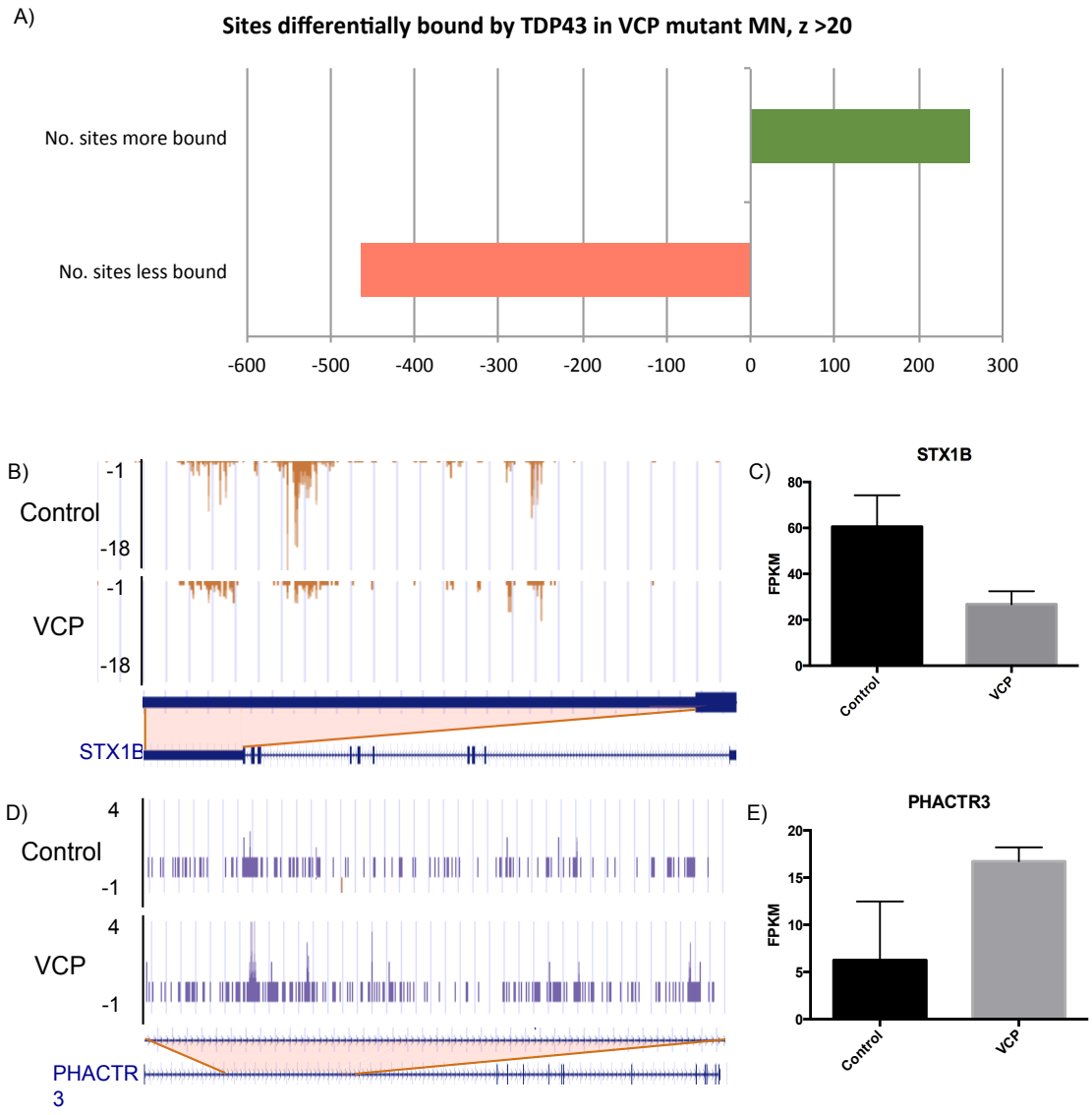
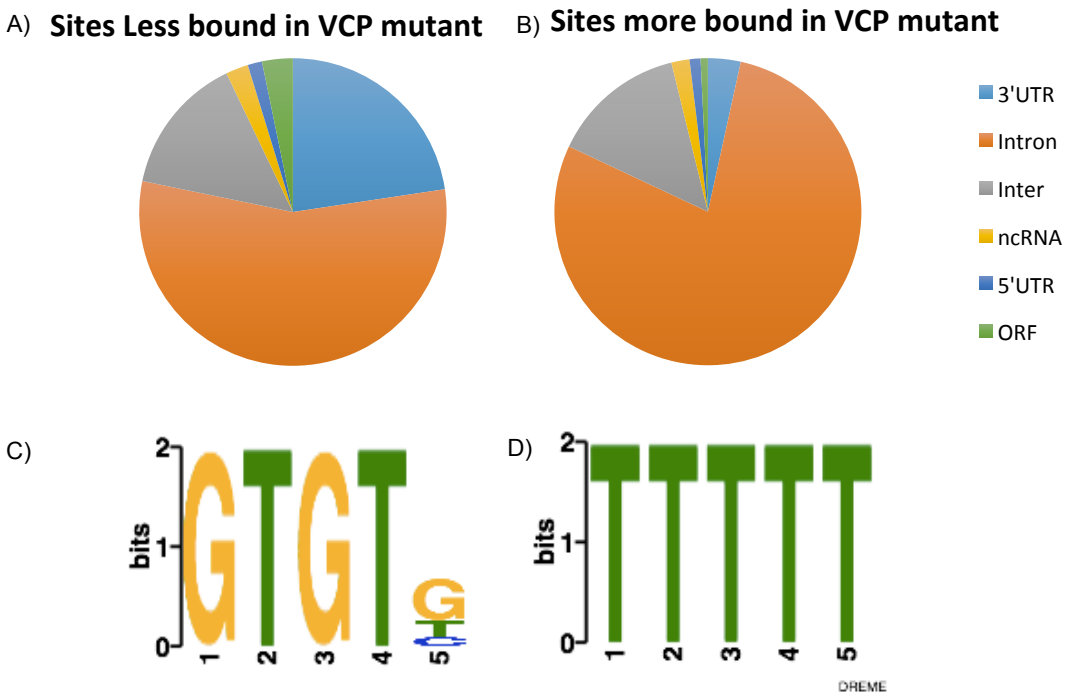


Figure 6.5; Differential TDP43 binding in VCP-mutant MNs

A) Barplot displays the number of sites more and less bound by TDP43 in VCP-mutant MNs. B) Representative trace showing binding of TDP43 to STX1B and C) relative mRNA expression levels of STX1B in Control and VCP-mutant MNs. D) Representative trace showing TDP43 binding to PHACTR3 and E) relative mRNA expression levels of PHACTR3 in Control and VCP mutant MNs. (N=3+ technical repeats across 2 control lines and 3 VCP mutant lines), mean RPKM values plotted with error bars that represent ± SEM.

Genomic localization and Sequence Motif analysis of the differentially bound sites

I have analyzed the genomic distribution of the differentially bound sites in VCP mutant MNs. This interestingly revealed the sites that less bound in the VCP mutant MNs have an enrichment within the 3'UTR (Figure 6.6.A). 3'UTR influences post-transcriptional regulation therefore suggest a level of gene expression regulation by TDP43 is lost in VCP-mutant MNs. The same effect is shown upon TDP43 C-terminal truncation; TDP43 binding is reduced in 3'UTRs and increased in intronic regions (I.J.Huppertz 2015).



E)

Decreased binding in VCP		Increased binding in VCP	
Motif	Enrichment score	Motif	Enrichment score
GTGTB	1.40E-302	TTTTT,	3.80E-294
TGTG	2.9e-324	AGGCT	5.40E-208
KTMYG	1.20E-186	KAGAY	3.90E-189
YGCRT	3.20E-147	GGAGT	5.60E-139
VHTTG	1.00E-126	CAASC	1.10E-134

Figure 6.6; Genomic binding distribution and motif analysis of differential bound TDP43 sites in control and VCP-mutant MNs

A & B) Respective genomic distribution of sites more and less bound by TDP43 in VCP mutant MNs. C & D) Top sequence motif's for the less and more bound sites by TDP43 in VCP-mutant MNs respectively. C) Table displays the top 5 sequence motif's and their correlating enrichment scores for both the more and less bound sites by TDP43 in VCP-mutant MNs

To further explore possible underlying reasons for the change in binding localization, we performed motif analysis on the sites differentially bound. Interestingly we see that the sites that are increasingly bound by TDP43 in the VCP mutant MNs are enriched in U-rich motif's, whereas the sites that are less bound by TDP43 in the VCP mutant are enriched in UG-rich motif's (Figure 6.6.B & C). Within the lab, we have begun to unravel the function of TDP43 binding to U-rich regions. Ina Huppertz has replicated TDP43's binding to U-rich regions in Hela cells and has shown that U rich binding of TDP43 is abolished after truncation of TDP43's LC domain (I.J.Huppertz 2015). This lead us to hypothesize that U-rich binding is specific to TDP43 and that TDP43 binds to U-rich regions with the help of the LC domain, either directly or being forced through its stabilizing interacting partners. The data I have presented here adds to the current hypothesis, implicating that U-rich binding is relevant in MND.

Significant further work will need to be untaken to determine 1/ the consequences of U rich binding, 2/ what interactions in the LC domain influence U rich binding seen? 3/ If this phenomena is broadly important to MND and other TDP43 proteinopathies?

GO analysis of differentially bound sites

In addition I have carried out GO term analysis on the differentially bound sites in VCP-mutant MNs, using all genes with a smoothed z-score >20 (Figure 6.7). I find that the sites that are less bound by TDP43 in VCP mutant MNs are those involved in cellular process regulation – this involves regulating the rate/frequency of processes at the cellular level (not single cell) so for example includes cell-cell communication. On the other hand among the top terms associated with increased TDP43 binding in VCP-mutant MNs is the regulation of neuronal differentiation and development. However at this stage it is not clear the functional implications of these binding sites.

Decreased binding in VCP			Increased binding in VCP		
Top 10 GO terms	p-value	Enrichment score	Top 10 GO terms	p-value	Enrichment score
regulation of cellular process	7.50E-09	1.32	regulation of neuron differentiation	1.93E-08	4.66
negative regulation of biological process	2.28E-08	1.64	regulation of neurogenesis	5.52E-08	4.14
regulation of cell communication	4.33E-08	1.83	regulation of neuron projection development	6.55E-08	5.19
developmental process	5.64E-08	1.62	regulation of cell projection organization	6.93E-08	4.54
biological regulation	6.05E-08	1.27	regulation of nervous system development	7.73E-08	3.87
regulation of biological process	6.51E-08	1.28	regulation of cell development	2.58E-07	3.6
regulation of multicellular organismal development	7.38E-08	2.26	positive regulation of neuron projection development	2.59E-07	6.59
negative regulation of cellular process	8.56E-08	1.64	positive regulation of neuron differentiation	4.03E-07	5.73
positive regulation of cellular process	9.24E-08	1.59	positive regulation of cell projection organization	5.06E-07	5.62
regulation of signaling	1.72E-07	1.8	positive regulation of nervous system development	5.84E-07	4.73

Figure 6.7; GO analysis of differentially bound TDP43 sites in VCP-mutant MNs
Table displaying top 10 GO terms associated with both more and less bound TDP43 sites

Discussion

Here I have firstly investigated the presence of TDP43 pathology in patient specific VCP mutant hiPSCs. I detect a cytoplasmic mislocalisation of TDP43 in D3 VCP mutant MNs, a key pathological hallmark associated with MND. This correlates with the onset of ER stress in VCP mutant MNs and is detected prior to a cell death phenotype, which is observed at D17 (described in chapter 5). The relationship between TPD43 pathology and ER stress has been investigated before; Walker et al describe how ER stress induces the accumulation of TDP43 in the cytoplasm (Walker et al. 2013). Further experiments are required before we can determine whether this sequence of events (ER stress followed by TDP43 aggregation) is true in our model VCP-related MND too. Moreover, previous studies have demonstrated that elevated levels of cytoplasmic TDP43 in MND correlate with cellular toxicity (Wegorzewska and Baloh 2011)(Barmada et al. 2010), I am currently exploring TDP43 pathology in D17MNs too (where we detect a selective vulnerability of VCP mutant MNs).

As previously mentioned many cellular models of TDP43-proteinopathies fail to accurately recapitulate TDP43 pathology. Using hiPSCs we were able to recapitulate TDP43 mislocalisation, a key aspect of TDP43 pathology (Manuela Neumann et al. 2006), specifically in VCP mutant MNs. This emphasizes the

power of hiPSCs to be used as a model system to study MND. However it is of upmost importance for us to explore other aspects of TDP43 pathology in addition to mislocalisation such as ubiquitination, phosphorylation, protein levels and solubility.

Next I have combined hiPSC methodologies with genomics to understand how VCP-mutations affect the composition of TDP43-RNA complexes in iPSC-derived MNs. In agreement with previous work, I show that in hiPSC-derived MNs TDP43 binds to predominantly to intronic regions in the genome and has a preference to UG-rich sequences (Tollervey et al. 2011).

Further I show there are differentially bound sites by TDP43 in Control and VCP-mutant MNs. Many of the sites less bound in VCP-mutant MNs lay within genes implicates in neuronal differentiation and development. This raises the hypothesis that neuronal development is altered in VCP mutant cells. This concept is reinforced by our previous findings (Chapter 5) showing synaptogenesis is defective in VCP mutant MNs and also by VCP's interaction with NF1, a protein that if mutated causes neurodevelopmental disorder neurofibromatosis (Hsueh 2012). Under normal conditions VCP and NF1 cooperate to regulate dendritic spine formation, but in VCP mutations this is abolished (H. F. Wang et al. 2011). Additional studies have also suggested a possible pathophysiological link between development and neurodegeneration (Lule, Ludolph, and Ludolph 2008). However, further work is needed to understand the mechanistic link between neural development and neurodegeneration and also what role the interaction between VCP and TDP43 plays in both processes.

Moreover, the differentially bound sites present a loss of TDP43 binding in sites located in 3'UTRs in VCP mutant MNs suggesting a partial loss of RNA regulation in VCP-related MND. The results also show an increase of TDP43 binding to U-rich sequences in VCP-mutant MNs. The implications of U-rich binding are largely unknown but separately a study has shown that TDP43's preference to U-rich sequences is lost upon truncation of TDP43's LC domain (I.J.Huppertz 2015). The LC domain of TDP43 is also responsible for protein interactions and formation of granules. Granules are present in all neurodegenerative diseases, with TDP43 being a major component, including VCP-related MND (J. O. Johnson et al. 2010; M Neumann et al. 2007). This

suggests that TDP43's function in MND are partially in the formation of granules, implemented via its binding to U-rich sites.

Stress granules are usually highly dynamic and form under cellular stress where RBPs and capped mRNA molecules (stalled translation complexes) aggregate in a process to fine-tune protein expression under stress. VCP itself has been shown to be involved in stress granule dynamics (Buchan et al. 2013), specifically functioning in the disassembly of stress granules. It has been suggested that when mutated this function of VCP is lost leading to persistent disease aggregates (Buchan et al. 2013), although we haven't yet characterized this. I reason that in VCP-related MND, VCP-mutations do not disperse cellular granules, which I believe are partially formed via TDP43 binding to U-rich sequences via its LC domain. All together it leads me to hypothesize TDP43 binding to U-rich sequences are less specific interactions than to UG-rich motifs, and that U-rich binding of TDP43 will be of relevance to MND.

Also to note, TDP43 is highly conserved throughout species but also among human paralogues, which share a 'common' LC/glycine rich region (21% homology), for example with hnRNPA1. Interestingly several paralogues have also been implicated in MND, including hnRNPA2B1 and hnRNPA1 (Benatar et al. 2013; H. J. Kim et al. 2013). Therefore it is reasonable to suggest that other hnRNP's may share similar functions to TDP43, perhaps via their LC domains, and could also play a role in MND. Additionally it has previously been shown that TDP43 binds up to 100 nucleotides away from UG repeats (Tollervey et al. 2011) which further suggests co-operative binding could occur via its glycine rich region.

To conclude this work has provided novel insights into the mechanism underlying RNA regulation in MND and into the characterization of TDP43 in VCP-related MND.

Chapter 7; General Discussion

The pathogenic mechanisms underlying MND are so far not fully characterized, in particular there is very limited understanding of the sequence of pathogenic events. The research I have presented here underscores the value of using hiPSCs to model disease and has furthered our understanding of primary vs. secondary pathogenic events in VCP-related MND.

Use of hiPSC to model neurodegeneration

Over the past decade, the discovery of hiPSCs combined with advancements in neural differentiation strategies has revolutionized our ability to analyze patient specific material and thus model disease. This is of particular benefit to studying MND (and other neurodegenerative disorders) due to the inaccessibility of spinal motor neurons from patients. Additionally it opens up the doors for personalized medicine in the future.

Previously long protocols and heterogeneous cellular yields have discouraged large-scale modeling of disease. Here I have refined differentiation methods for the generation spinal motor neurons and astrocytes that result in a high yield and a near homogeneous cellular population. This has been primarily achieved by the use of monolayer culture approach and compounds to synchronize exit from cell cycle. An additional block in the wide-spread use of hiPSC technology has been the expense. Here we have addressed this by eliminating the use of growth factors, such as BDNF and GDNF, and also by massively shortening our strategy for motoneurogenesis (51+ to 21days). Finally we have ensured our protocols closely mirror key in vivo developmental processes. These advancements provide the opportunity to maximally exploit hiPSC technology in the future. A few laboratories have already exploited hiPSC-derivatives in high-throughput screens; for example a recent study highlights the ability to use hiPSC-derived MNs in high-throughput imaging and drug screens (Burkhardt et al. 2013). They first accessed the presence of TDP43 pathology in 384-well format then secondly screened 1,757 compounds to search for any that would ameliorate the pathology observed (Burkhardt et al. 2013). Our rapid, high yield and reduced cost protocol would benefit this approach.

Further, to date the majority of hiPSC-derived cells have been characterized by a handful of protein markers and function assays, however this doesn't provide comprehensive functional maturity or cell-subtype information. Here we have,

in addition to standard ICC and multiple functional assays, carried out transcriptional profiling to validate our specified cellular population. This enables us to capture a snapshot of many cellular properties at once, for example expression of neurotransmission transporters, region specific genes, ion channels and signals controlling axonal branching, although we loose single cell resolution in this analysis. For selective cell types there is a particular lack of standardized cell-specific markers that encompass subtype specificity and maturational status. This is the case for astrocytes. Therefore here we sought to dissect the transcriptional signature throughout astrogliogenesis, using an unbiased approach, to objectively define novel reliable stage-specific markers. In addition we compared the transcriptional signatures of iPSC gliogenic derivatives to human AC tissue at two different maturational stages (fetal and adult) to gain insights into the maturation status of our iPSC-derivatives. Validation is required before we can confidently define a set of markers to establish cell identity throughout differentiation and maturation, and further assays are required to understand the heterogeneity of astrocytes.

Finally, an important consideration for using hiPSCs to study of adult-onset diseases such as with neurodegenerative conditions, is the maturation status of iPSC-derivatives. Within the relatively short period of time in culture, we and others have observed, cellular phenotypes emerge in patient-derived neurons (H. Chen et al. 2014; Devlin et al. 2015; Wainger, Kiskinis, Mellin, Wiskow, Han, et al. 2014). Could this be due to in vitro induced stress and/or lack of supportive environment in vitro? Neuronal differentiation in vitro may not fully recapitulate neuronal development as it happens in vivo due to the absence of non-neuronal components in culture. Additionally previous studies have demonstrated cellular pathology often appears before symptom onset (sometimes developing over decades) (Bories et al. 2007; Kuo et al. 2005), which opens up the possibility of defects being detected early in development/maturation and thus hiPSCs would be a good system to explore this within. Moreover there have been several attempts to accelerate cellular maturity, namely by either co-culture with other tissue or by administration of specific factors (Miller et al. 2013; Zhang et al. 2016). Overall the meaning of the early phenotypes we detect using hiPSCs compared with those we observe at the end stage of disease (postmortem tissue) and events that occur in adult-onset neurodegeneration, remains unresolved.

Dissecting the pathogenic events in VCP-related MND

To date both defective protein homeostasis and RNA regulation have been implicated to play key roles in MND pathogenesis, however the individual role of each process is controversial. Manipulation of either protein homeostasis or RNA regulation has been shown to cause a forward feedback loop resulting relentless disease progression (Ling, Polymenidou, and Cleveland 2013), but the initiating event in MND remains elusive. The complete in vitro recapitulation of human MND pathology is challenging but in my PhD thesis I have presented to use of patient-specific hiPSC models to dissect the primary pathogenic events and to distinguish the onset of secondary events in VCP-related MND.

VCP usually contributes to a myriad of cellular functions; including protein control, membrane turnover, protein degradation, cell cycle regulation and dendritic formation (Kim et al. 2014; Shih and Hsueh 2016; Vaz, Halder, and Ramadan 2013; Wang et al. 2011; Yamanaka, Sasagawa, and Ogura 2012) but currently the consequences of VCP mutations within human neuronal samples and in the context of MND are unknown/poorly defined. Thus we firstly employed an unbiased approach, RNA sequencing, to uncover pathological processes underlying VCP-related MND at the RNA level. Additionally we separately used a range of analysis techniques to address multiple organelle functions in cells carrying VCP mutations (eg. ER, mitochondrial and synaptic health). Using a combination of hiPSC technologies, RNA sequencing and collaboratively cellular imaging methods, we detect a clear phenotype in VCP mutant iPSC neural derivatives. Analysis throughout the differentiation of motor neurons reveals we first detect a phenotype shortly after terminal differentiation of neuronal precursors; ER stress is detected after 3 days of terminal differentiation and during neuronal maturation (day 3-17) several further striking phenotypes emerge; mitochondrial dysfunction, oxidative stress, synaptic defects and cellular vulnerability.

Importantly, combining RNAseq and imaging analysis across a differentiation time course, we have been able to uncover that defects in cytoplasmic protein homeostasis arise before we detect defective RNA regulation. Further the later transcriptional profile of VCP-mutant MNs complimented a selection of the phenotypes detected through imaging, demonstrating at least some of the later events in VCP-related MND are coordinated by defects in RNA regulation. Together, for the first time these results have begun to untangle the web of

defective protein and RNA regulation in MND; they suggest that in VCP-related MND deregulated proteostasis occurs prior to the onset of defective RNA regulation.

There is further controversy in MND research about the role of astrocytes in disease pathogenesis. For TDP43 mutations a cell autonomous role has been described for AC in MND (Serio et al. 2013). For SOD1 and C9ORF72 mutations instead a non-cell autonomous role of AC has been described, but for TDP43 a non-cell autonomous role is debated (Di Giorgio et al. 2008; K. Meyer et al. 2014; Nagai et al. 2007; Serio et al. 2013; Tong et al. 2013). Overall research to date would suggest that the role of AC in MND may be mutation-specific. The role of VCP-mutant AC in MND has no yet been explored; therefore I have addressed this by probing VCP-mutant hiPSC-derived ACs. So far we have not detected a robust cell autonomous role for VCP-mutant AC in VCP-related MND. This finding accentuates our confidence in the cell-specific vulnerability of VCP-mutant motor neurons we have detected.

Finally we utilised iCLIP methodology to begin to look at underlying mechanisms of the RBP that constitutes the pathological hallmark of MND, TDP43. The involvement of TDP43 in MND is undisputed. However, exactly how TDP43 elicits its effects are not clear; there is evidence supporting both a loss of function and gain of function for TDP43 (E. B. Lee, Lee, and Trojanowski 2012). Further, the onset of TDP43 pathology in MND in relation to the timing of additional phenotypes observed is yet to be disentangled. We have not yet managed to fully characterize TDP43 pathology in our model of VCP-related MND, however we have defined changes in TDP43 RNA-binding in MNs in the presence of VCP-mutations. Our results demonstrate many alterations occur in TDP43 binding to RNA's involved in neuronal functions. It is challenging to combine high-resolution protein–RNA interaction maps (from iCLIP) with other high-throughput data, such as RNAseq, in order to elucidate the functions and dynamics of cellular protein–RNA interactions. Thus further analysis and validation needs be carried out to reveal the functional consequences of this differential TDP43 binding in VCP-mutant MNs.

In addition, our TDP43 iCLIP data in VCP-mutant MNs shows that sequence specific binding is altered in the presence of VCP-mutations in MNs; a slight

decrease of UG-rich motifs coincides with a slight increase of U-rich motifs in the TDP43 iCLIP data produced from VCP-mutant MNs. Previous findings within the lab showed that U-rich binding is lost upon truncation of TDP43's LC domain, which is responsible for protein-protein interactions (I.J.Huppertz 2015). When taken together, this suggests that U rich binding depends on co-operative protein-protein interactions with TDP43 and that these interactions could be of relevance in VCP-related MND. In spite of this, the wider implications of TDP43 binding to U-rich sites in MND are yet to be explored.

A proposed model of pathogenesis in VCP-related MND

Combining my findings so far I propose a model for pathogenesis in VCP-related MND (Figure 7.1). I propose that ER stress, along with TDP43 pathology, are among the earliest pathogenic events, occurring at the onset of terminal MN differentiation. We next simultaneously observe mitochondrial dysfunction and oxidative stress in VCP-mutant MNs, which in turn we propose reinforce each other's activation. Recently it has been shown that TDP43 pathology, particularly pTDP43, accumulates at the mitochondria, triggering mitochondria dysfunction (W. Wang et al. 2016), thus supporting the sequence of events we observe. Separately VCP-mutations cause a synaptic dysfunction, which in part is orchestrated by deregulated RNA. Taken together, this results in selective vulnerability in VCP-mutant MNs. Nevertheless I must reinforce, further investigations need to be carried out before we can present a fully comprehensive model of VCP-related MND with high confidence. This model I propose can be used to design further investigations.

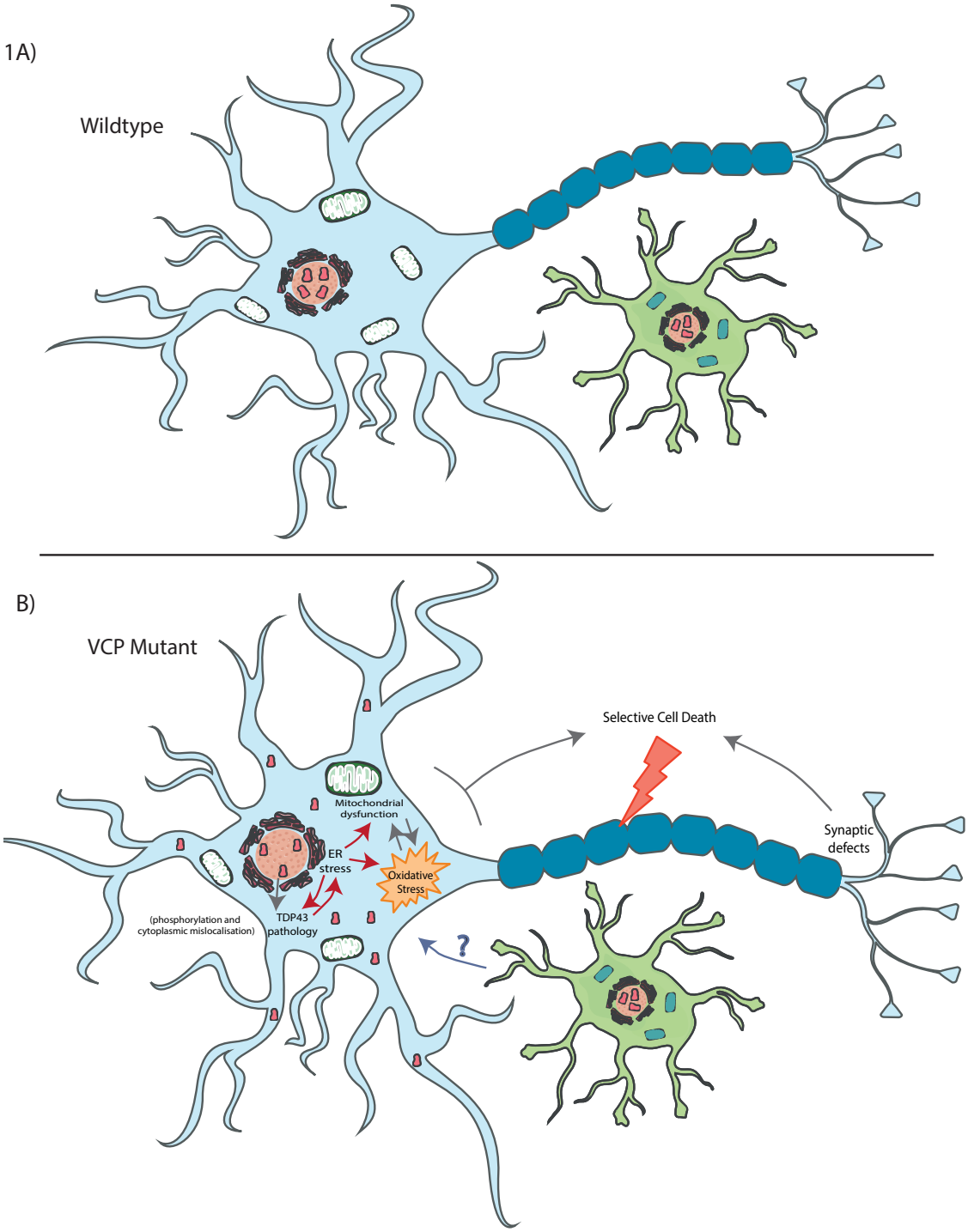


Figure 7.1; Proposed model of pathogenesis in VCP-related MND

A) Control MN and AC. B) Proposed pathogenic pathways contributing to selective vulnerability of VCP-MNs.

Future Plans

In my PhD thesis I have presented many unresolved questions, thus here I propose several ways in which we can address these points in the future.

Overcoming genetic variability when working with patient-specific cells

A considerable challenge in modeling neurodegeneration using hiPSC remains in the heterogeneity of different cell lines. A major source of variation is genetic. This arises firstly from differences in the genetic background between different individuals and secondly can occur between hiPSC clones from the same individual. Here we have found it a challenge to dissect pathogenic events in mutant hiPSCs due to variation seen between cell lines. These problems can be addressed by the use of genome editing technologies, such as CRISPR (**C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats), to produce isogenic lines.

The CRISPR system is a microbial nuclease system that can be implemented in mammalian cells by co-expressing the bacterial Cas9 nuclease along with the guide RNA (Cong et al. 2013; Jiang et al. 2013). This method works by the induction of a double strand break (DSB) in the DNA at the site of target, which are repaired by one of two general repair pathways: non homologous-end joining (NHEJ) – where insertion or deletion often occurs leading to frameshift or a premature stop codon, or Homologous Recombination (HR) where the DSB is repaired using a donor DNA as a template so you can insert your desired sequence/mutation. Isogenic lines can be generated by either introducing a disease-causing mutation into control patients cell lines or by correcting a mutation in a disease-causing cell line. This substantially reduces genetic heterogeneity.

As with most things, this approach has its disadvantages. The majority of neurodegenerative conditions, including MND cases, are sporadic, and research has shown that both the genetic and environmental factors contribute to neurodegeneration. However, the roles and interactions of genetic and environmental elements in disease pathogenesis remain elusive. Since the environmental contributions in neurodegeneration are unclear, only genetic factors are recapitulated in current disease models. Further the genetics underlying sporadic cases remains to be elucidated and it is not possible to

make isogenic controls from patients with sporadic disease. Therefore the approach we currently adopt in the laboratory is to first gain insights into disease pathogenesis by studying models with known disease-causing mutations, then see if findings can be replicated in samples from other disease-causing mutations and sporadic cases.

Rescuing the VCP phenotype

One of our next goals is to rescue our VCP mutant hiPSCs. We will attempt to modulate the phenotype using both pharmacological and genetic approaches. There are many pharmacological treatments that have already been shown to improve phenotypes seen in MND or neurodegenerative disease model; this includes retigabine (Wainger, Kiskinis, Mellin, Wiskow, Han, et al. 2014), the VCP inhibitor vesnarinione shown to manipulate the NfKb pathway(Hotta et al. 2013), UPR manipulators firstly the perk inhibitor (GSK2656157) which reinitiates translation during UPR (Moreno et al. 2013), Salubrinal which promotes translational arrest by inhibiting the phosphorylation of eIF2a (Kiskinis et al. 2014) or TUDCA a molecule that acts upstream of Perk, and treatment with statins, since Simvastatin has been shown to promotes mitochondrial function (Chataway 2015). We find ER stress plays a primary role in VCP-related MND thus we will first test modulators of the ER stress pathway, including Perk inhibitor and Salubrinal.

Separately we will use the CRISPR platform to perturb ER stress, the primary pathogenic pathway in our model system, to try and reverse the primary and secondary phenotypes. It is important to note that pertubation of ER stress pathway could influence cell-subtype diversity in addition to effecting the cellular phenotype (Dalton, Lyons, and Lomvardas 2013). Thus we will assess both subtype selection and cellular health, including cellular viability and mitochondria function, using cellular imaging and RNAseq as readouts.

Further dissecting the underlying molecular mechanisms in VCP-related MND

Although we have demonstrated the sequential onset of pathogenic events in VCP-related MND, further studies are required to definitively show the causative sequence of events. After 17 days of terminal motoneurogenesis we detect striking mitochondria dysfunction, ER stress, oxidative stress, selective vulnerability and synaptic defects. All mentioned pathways are intricately linked

and often work to reinforce each other (Barmada et al. 2010; Carrì et al. 2015; Colombrita et al. 2009; Ilieva et al. 2007; Malhotra and Kaufman 2011; Walker et al. 2013), therefore it is no surprise we detect them in parallel. In order to unravel the precise course of disease progression we will analyze events in across shorter windows of time, eg. every 24hours throughout terminal motoneurogenesis. Further we can induce each phenotype individually (by adding pharmacological stressors or using genome editing technology) and analyze the phenotypic consequences on the other interlinked pathways.

Non-cell autonomous mechanisms of injury in VCP-related MND

Although traditionally neurodegenerative conditions, such as MND, have been viewed as neuron-specific it is increasingly recognised that glia may contribute to disease pathogenesis. Here we have not detected a robust cell autonomous role for astrocytes in VCP-related MND, however non-cell autonomous mechanisms of injury in VCP-related MND remain unevaluated. To assess this we are analysing both control and mutant, MN and AC co-cultures and also using transwell culture supports, which enable a continuous exchange of soluble factors between AC and MNs.

Common mechanisms in MND

After fully characterising and establishing a pathogenic model for MND caused by VCP mutations it is important to next explore if the underlying mechanisms are mutation (VCP) or disease (MND) specific. Within the lab we are addressing this by analysing hiPSC-derivatives from patients carrying mutations in SOD1 (a cytoplasmic antioxidant, where the main function is to convert superoxide, which is highly reactive, to hydrogen peroxide/oxygen), and FUS (FUS is an RBP located in the nucleus and is involved in many stages of RNA processing). Both SOD1 and FUS mutations do not recapitulate the typical TDP43 pathology seen in >95% of MND cases, but it would be interesting to see if there are any shared pathways or highlighted events that could be generically targeted (eg. UPR/autophagy). This can then be tested by pharmacological rescue or genetic manipulation in these additional mutant lines. Further we are currently carrying out transcriptional analysis of post mortem samples from sporadic and C9ORF72-mutant MND patients to again see if any common pathways are highlighted across different genetic models of MND.

Chapter 8; Appendix

8.1 Enriched monolayer spinal motor neurogenesis protocol

Preparation TIMING ~ 1day

1. For iPSC culture, coat cell culture plates with 150µg/ml Geltrex (diluted in DMEM media) and leave to incubate at 37°C for 1 hour. Pre-coated plates can be used for up to 7days if kept sterile at 4°C.
2. Thaw a vial of iPSCs by rotating quickly between your hands and adding media (pre-warmed to room temperature) drop-wise. Transfer the suspension to a 15ml falcon and dilute in 10ml media. Spin for 1min at 1000g at room temperature. Resuspend the iPSCs in 2ml Essential 8 media and add directly into 1well of a pre-coated geltrex 6well plate. Shake the plate side-to-side and back-and-forth to evenly distribute the cells. Leave the plate at 37°C overnight (Cell attachment can be checked after leaving to settle for a minimum of 20minutes).

TIP – Can add Y27632 into essential 8 medium to improve cell survival during thawing.

iPSC cultures TIMING ~ 1week

3. The following day after thawing check the iPSC colonies have settled and are healthy. Carry out a full media change and remove the Y27632 supplement if added.
4. Continue feeding iPSC cultures daily with 2ml fresh Essential 8 media to expand iPSC colonies.
5. Cultures should be passaged prior to contacting with each other. Passage iPSC on a 6well plate by removing the media, adding 1ml EDTA and leaving to incubate for 5minutes at 37°C. After 5minutes check colonies have begun to lift under the microscope, then aspirate EDTA and collect cells using a p1000 in 1ml essential 8 media. Cells can be split across several wells of a 6well plate and expanded as much as required.

TIP – Good quality cultures at this stage are essential in order to maximize success in differentiation - ensure there is minimal cell death and iPSC colonies are pure (no contaminating cells present).

6. The day before you start neural induction, cells need passaged to achieve 100% confluency onto fresh geltrex-coated plates.

TIP – This is usually achieved by passaging 3 semi-confluent wells of a 6well plate into 1 well.

Neural Induction TIMING ~ 7days

7. Before beginning neural induction check iPSCs are 100% confluent, but that they are not growing on top of each other.
8. Feed daily with 2ml neural induction media for a total of 7days. During this time a neuroepithlial sheet will form.
9. After 7days of neural induction the neuroepithelial layer should be lifted and re-plated. First prepare laminin-coated plates at dilution 10-20µg/ml and incubate at 37°C for 4hours.
10. Lift the neuroepithlial sheet by adding 350uL dispase into each well containing 2ml neural induction media and leave to incubate at 37°C. Check the plate after 10minutes and then every 1-2 minutes for the neuroepithlial sheet to lift.
11. Once lifted, collect the sheet using a p1000 and add into a 15ml falcon containing 10ml neural maintenance media. Spin for 2 minutes at 500g.
12. Aspirate media and carefully add a fresh 10ml neural maintenance media to wash the cells (dispase is not inactivated by media so needs to be thoroughly washed away). Spin again for 2minutes at 500g.

Patterning TIMING ~ 11days

13. Gently re-suspend the cells in spinal cord patterning media, as to minimize breaking up cellular clumps, and dry plate onto laminin coated plates. (Dry plating helps with cellular attachment.)

TIP – Usually 1 well of neural induction can be placed into 2wells after washing. Cells should not be plated too sparsely or be broken up too much at this stage.

14. 1hour after dry plating, add 2ml spinal cord patterning media to each well.
15. For 7days feed daily with 2ml spinal cord patterning media. RA and Purmorphamine will act respectively to caudalise and ventralise neural precursors to the pMN domain.
16. After 7days passage neural precursors. Usually this can be done simply by using a p1000 to collect neural precursors but if needed EDTA can also be used to lift precursors from the plate. Add collected cells to a 15ml falcon and spin for 1minute at 1000g.

17. Re-suspend the pellet in secondary patterning media and gently triturate before dry plating onto fresh laminin-coated plates (laminin coating at 10-20µg/ml in DPBS). Again after approximately 1hour add 2ml media to each well.
18. Feed daily for an additional 3days with secondary patterning media.

Propagation TIMING ~ 0-30days

TIP – To increase yield and facilitate easier plating of neurons at a single cell level, precursors should be propagated in FGF for a maximum of 30day. This stage is not essential and can skip directly to terminal differentiation.

19. Passage NPCs by removing the media, adding 1ml EDTA and leaving to incubate for 5minutes at 37°C. After 5minutes aspirate EDTA and collect cells using a p1000 in 1ml propagation media.
20. Plate precursors on geltrex coated plates.
21. Feed 3 times weekly with propagation media and passage weekly. Gradually break up precursors further each time they are passaged.

Terminal differentiation and maturation TIMING ~ 3+days

TIP – Replate before terminal differentiation to the cell density and plates required for analysis. Neurons do not passage well after terminal differentiation of 5+ days.

22. Prepare Laminin coated plates (20-40µg/ml in DPBS) as required for downstream analysis (eg. 96well format, coverslips etc).
23. Passage NPCs by removing media, add 1ml accutase and incubate for 3minutes at 37°C.
24. Check the cells are lifting under the microscope, then collect them using a p1000 and neural maintenance media. Add into a falcon with 10ml neural maintenance media and spin for 1minute at 1000g.
25. Re-suspend the precursors in terminal differentiation media and plate onto laminin-coated plates.
26. Feed 3 times weekly during terminal differentiation.

Table 8.1; Troubleshooting

Step	Problem	Possible reasons	Solution
5	Contaminating	• Influence of cell-	To clean up a mixed

	cells present in between iPSC colonies	cell contact (Ensure cells are passaged at the correct time to avoid this in future - i.e. don't let cultures over grow) • Cells are not receiving the correct signals (Can double check all reagents are in date, active and that the media has been made up correctly).	culture where the majority of cells are pluripotent colonies - passage harshly and after a couple of passages only pluripotent cells should remain. To clean up pluripotent colonies that have been "taken over" by another cell type – pick pluripotent colonies and replate.
7	iPSCs have not reached 100% confluency	• Too few cells plated	If small gaps are observed between cells, feed with 4ml essential 8 media and check confluency again after 24hours. If 100% confluent then begin neural induction as per step 8. If there are large gaps between cells then passage again for neural induction and plate more cells than previous attempt.
10	Neuroepithlial sheet lifting early	• Beginning with over confluent cells before	Advise to start again

		induction • Toxicity possible induced by CHIR99021	
13-15	NPCs die off	• Neuroepithlial sheet broken up too harshly or left in dispase too long	Start over again from neural induction but do not break the cells up or plate too sparcely at stage 13. They need cell-to-cell signaling at this stage.
15	Cells will not attach		Feed the precursors in suspension; gently aspirate half the media and replace with fresh. Replate as described after 7days of patterning. Attachment should be more efficient once the cells are broken up further. Additionally you can increase concentration of laminin used to 1:50 to improve attachment.
26	MNs begin to lift from the plate during terminal differentiation	• Laminin-coating dried out or not concentrated enough • Excessive movement of the plate can disturb the MNs	As long as it is not too late (i.e. the majority of each MN is attached to the plate but you notice the edges are peeling) spike additional laminin into the media

			to improve adherence. Also carry out half media changes to reduce impact of feeding on the MNs.

8.2 Enrichment of MNs before and after propagation of NPCs

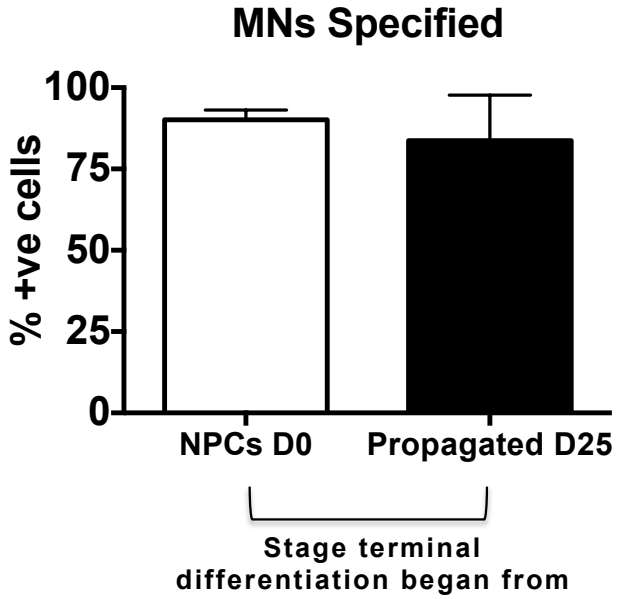


Figure 8.2; The percentage of MNs specified before and after the propagation of NPCs.

Quantitative ICC of SMI32 shows the percentage of MNs specified after 17days of terminal differentiation starting with either NPCs (D0) or NPCs that have been propagated for 25days prior to terminal differentiation.

8.3 GO analysis for differentially expressed genes between Early and Late NPCs

Table 8.1

Full list of GO terms associated with differentially expressed genes highlighted between early and Late NPCs with Log(Fold change) >1.5, p-value <0.05 and enrichment score>2.

Up regulated in Late NPCs				Down regulated in Late NPCs			
GO term	Description	P-value	Enrichment score	GO term	Description	P-value	Enrichment score
GO:0004872	receptor activity	4.51E-25	2.32	GO:0099600	transmembrane receptor activity	1.77E-18	2.65
GO:0060089	molecular transducer activity	4.51E-25	2.32	GO:0038023	signaling receptor activity	2.43E-17	2.49
GO:0004871	signal transducer activity	3.77E-21	2.08	GO:0004888	transmembrane signaling receptor activity	8.89E-17	2.6
GO:0038023	signaling receptor activity	7.25E-21	2.4	GO:0022836	gated channel activity	1.26E-16	3.57
GO:0099600	transmembrane receptor activity	2.44E-20	2.46	GO:0022838	substrate-specific channel activity	7.03E-16	3.15
GO:0004888	transmembrane signaling receptor activity	1.61E-19	2.47	GO:0005261	cation channel activity	1.55E-15	3.59
GO:0005509	calcium ion binding	6.57E-14	2.1	GO:0005216	ion channel activity	1.80E-15	3.12
GO:0005201	extracellular matrix structural constituent	3.04E-12	4.87	GO:0004872	receptor activity	2.14E-15	2.18
GO:0005125	cytokine activity	1.01E-11	3.89	GO:0060089	molecular transducer activity	2.14E-15	2.18
GO:0019838	growth factor binding	1.71E-10	3.18	GO:0022803	passive transmembrane transporter activity	3.04E-14	2.91
GO:0061134	peptidase regulator activity	1.72E-09	2.9	GO:0015267	channel activity	3.04E-14	2.91
GO:0030246	carbohydrate binding	3.69E-09	2.52	GO:0004930	G-protein coupled receptor activity	1.64E-12	3.28
GO:0030414	peptidase inhibitor activity	7.90E-09	3.24	GO:0046873	metal ion transmembrane transporter activity	8.35E-12	2.71
GO:0005539	glycosaminoglycan binding	1.41E-08	2.71	GO:0015079	potassium ion transmembrane transporter activity	3.01E-11	4.19
GO:0004866	endopeptidase inhibitor activity	1.80E-08	3.22	GO:0005267	potassium channel activity	4.23E-11	4.57
GO:0005518	collagen binding	2.64E-08	3.67	GO:0015075	ion transmembrane transporter activity	2.07E-10	2.03
GO:0005126	cytokine receptor binding	5.45E-08	2.58	GO:0022843	voltage-gated cation channel activity	2.33E-10	4.27
GO:0050840	extracellular matrix binding	7.75E-08	3.91	GO:0022834	ligand-gated channel activity	2.58E-10	4.11
GO:0061135	endopeptidase regulator activity	1.50E-07	2.95	GO:0015276	ligand-gated ion channel activity	2.58E-10	4.11
GO:0008083	growth factor activity	1.70E-07	3	GO:0022891	substrate-specific transmembrane transporter activity	3.18E-10	1.96
GO:0002020	protease binding	3.28E-07	2.91	GO:0022832	voltage-gated channel activity	3.47E-10	3.62
GO:0004930	G-protein coupled receptor activity	4.53E-07	2.3	GO:0005244	voltage-gated ion channel activity	3.47E-10	3.62
GO:0008201	heparin binding	4.86E-07	2.68	GO:0005249	voltage-gated potassium channel activity	3.64E-10	5.15

Up regulated in Late NPCs continued..				Down regulated in Late NPCs continued..			
GO term	Description	P-value	Enrichment score	GO term	Description	P-value	Enrichment score
GO:0005178	integrin binding	9.79E-07	2.84	GO:0008324	cation transmembrane transporter activity	5.36E-10	2.2
GO:0001664	G-protein coupled receptor binding	1.42E-06	2.29	GO:0015077	monovalent inorganic cation transmembrane transporter activity	4.27E-09	2.58
GO:0005161	platelet-derived growth factor receptor binding	3.10E-06	5.57	GO:0022890	inorganic cation transmembrane transporter activity	1.38E-08	2.2
GO:0048407	platelet-derived growth factor binding	3.23E-06	7.04	GO:0008188	neuropeptide receptor activity	3.01E-08	7.75
GO:0004896	cytokine receptor activity	4.13E-06	3.75	GO:0005231	excitatory extracellular ligand-gated ion channel activity	3.43E-08	5.43
GO:0004867	serine-type endopeptidase inhibitor activity	4.47E-06	3.55	GO:0005230	extracellular ligand-gated ion channel activity	4.84E-08	4.69
GO:0019955	cytokine binding	6.42E-06	3.18	GO:0030594	neurotransmitter receptor activity	5.63E-08	4.93
GO:0044548	S100 protein binding	6.53E-06	5.85	GO:0005179	hormone activity	1.89E-07	5.22
GO:0004222	metalloendopeptidase activity	8.05E-06	2.68	GO:0071855	neuropeptide receptor binding	2.36E-07	7.51
GO:0003823	antigen binding	9.83E-06	3.73	GO:0001077	transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	2.30E-06	2.39
GO:0019199	transmembrane receptor protein kinase activity	1.55E-05	2.73	GO:0030553	cGMP binding	3.90E-06	9.3
GO:0004714	transmembrane receptor protein tyrosine kinase activity	1.85E-05	2.97	GO:0001640	adenylate cyclase inhibiting G-protein coupled glutamate receptor activity	6.59E-06	10.85
GO:0008237	metallopeptidase activity	1.96E-05	2.19	GO:0008528	G-protein coupled peptide receptor activity	9.89E-06	3.92
GO:0001968	fibronectin binding	2.01E-05	4.02	GO:0001653	peptide receptor activity	1.40E-05	3.81
GO:0050839	cell adhesion molecule binding	2.49E-05	2.11	GO:0005251	delayed rectifier potassium channel activity	1.81E-05	5.79
GO:1901618	organic hydroxy compound transmembrane transporter activity	2.72E-05	4.23	GO:0008066	glutamate receptor activity	3.05E-05	4.88
GO:0017171	serine hydrolase activity	3.54E-05	2.6	GO:0015464	acetylcholine receptor activity	3.98E-05	7.23
GO:0005544	calcium-dependent phospholipid binding	3.78E-05	3.37	GO:0001228	transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding	4.64E-05	2.01
GO:0004620	phospholipase activity	4.76E-05	2.79	GO:0015081	sodium ion transmembrane transporter activity	4.65E-05	2.88

Up regulated in Late NPCs continued..			
GO term	Description	P-value	Enrichment score
GO:0015291	secondary active transmembrane transporter activity	5.41E-05	2.06
GO:0042056	chemoattractant activity	8.39E-05	4.6
GO:0004252	serine-type endopeptidase activity	9.02E-05	2.87
GO:0008236	serine-type peptidase activity	9.43E-05	2.51
GO:0016641	oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as acceptor	9.46E-05	5.12
GO:0005160	transforming growth factor beta receptor binding	1.37E-04	3.4
GO:0004859	phospholipase inhibitor activity	1.59E-04	6.7
GO:0005534	galactose binding	2.38E-04	8.04
GO:0038024	cargo receptor activity	2.68E-04	3.02
GO:0048306	calcium-dependent protein binding	3.54E-04	2.68
GO:0016298	lipase activity	3.87E-04	2.4
GO:0008227	G-protein coupled amine receptor activity	4.98E-04	5.75
GO:0055102	lipase inhibitor activity	4.98E-04	5.75
GO:0042605	peptide antigen binding	4.98E-04	5.75
GO:0008028	monocarboxylic acid transmembrane transporter activity	5.00E-04	3.45
GO:0016638	oxidoreductase activity, acting on the CH-NH2 group of donors	7.00E-04	4.02
GO:0050431	transforming growth factor beta binding	7.00E-04	4.02
GO:0001618	virus receptor activity	9.24E-04	2.68
GO:0008528	G-protein coupled peptide receptor activity	9.24E-04	2.68

Down regulated in Late NPCs continued..			
GO term	Description	P-value	Enrichment score
GO:0034617	tetrahydrobiopterin binding	7.18E-05	10.85
GO:0003777	microtubule motor activity	7.34E-05	2.89
GO:0004114	3',5'-cyclic-nucleotide phosphodiesterase activity	9.56E-05	4.82
GO:0022824	transmitter-gated ion channel activity	9.56E-05	4.82
GO:0022835	transmitter-gated channel activity	9.56E-05	4.82
GO:0005246	calcium channel regulator activity	1.15E-04	4.25
GO:0004889	acetylcholine-activated cation-selective channel activity	1.18E-04	7.75
GO:0004112	cyclic-nucleotide phosphodiesterase activity	1.52E-04	4.57
GO:0022841	potassium ion leak channel activity	2.91E-04	6.78
GO:0016594	glycine binding	2.91E-04	6.78
GO:0005184	neuropeptide hormone activity	3.32E-04	8.68
GO:0030551	cyclic nucleotide binding	5.00E-04	3.95
GO:0004497	monooxygenase activity	5.76E-04	3.06
GO:0042165	neurotransmitter binding	5.87E-04	5.01
GO:0001972	retinoic acid binding	6.05E-04	6.03
GO:0043138	3'-5' DNA helicase activity	6.05E-04	6.03
GO:0022842	narrow pore channel activity	6.05E-04	6.03
GO:0022840	leak channel activity	6.05E-04	6.03
GO:0042166	acetylcholine binding	6.05E-04	6.03
GO:0070405	ammonium ion binding	7.05E-04	3.77
GO:0005272	sodium channel activity	7.09E-04	4.22
GO:0031406	carboxylic acid binding	7.70E-04	2.03
GO:0004517	nitric-oxide synthase activity	7.81E-04	10.85
GO:0004118	cGMP-stimulated cyclic-nucleotide phosphodiesterase activity	7.81E-04	10.85
GO:0008401	retinoic acid 4-hydroxylase activity	7.81E-04	10.85
GO:0005222	intracellular cAMP activated cation channel activity	7.81E-04	10.85
GO:0043177	organic acid binding	8.65E-04	2.01
GO:0047555	3',5'-cyclic-GMP phosphodiesterase activity	9.24E-04	7.23

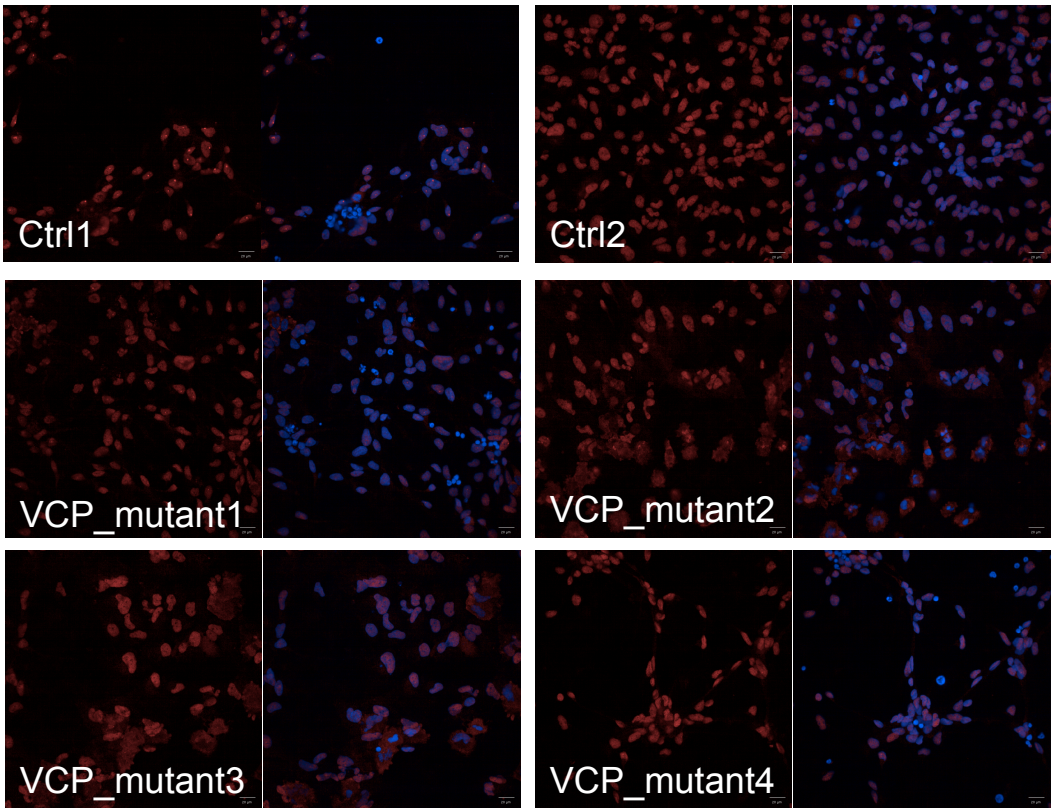
8.4 Comparing the RNA yield from multiple RNA extraction methods

Table 8.2

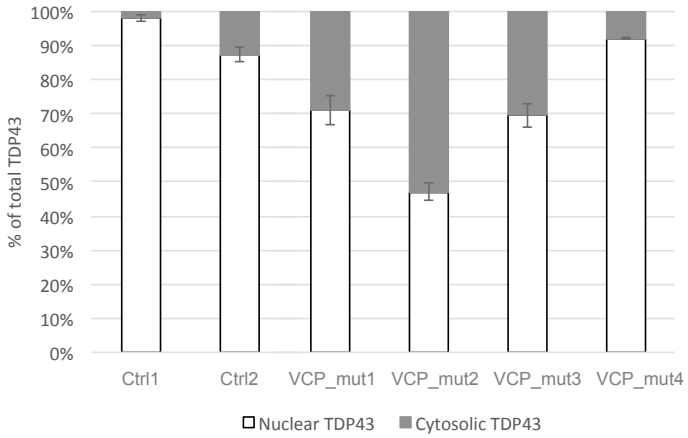
Randomly selected samples	Extraction method (ug extracted per sample)		
	Qiagen Rneasy	Zymo Direct-zol miniprep	Promega Maxwell simplyRNA cells
1	7.47	9.06	19.645
2	6.99	13.68	19.705
3	4.41	11.31	16.81
4	9.03	7.98	16.58
5	2.22	7.62	17.68
6	4.14	8.79	11.14
7	7.65	11.454	21.435
8	5.97	17.544	21.58
9	9.66	21.969	18.205
10	6.87	17.853	18.155

8.5 TDP43 localization in Control and VCP mutant D3MNs

A)



B)



C)

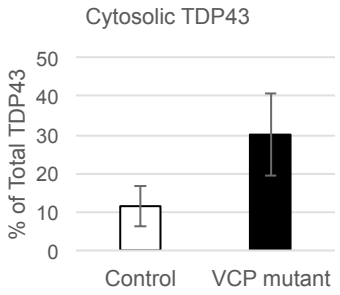


Figure 8.3; Ungrouped Control and VCP mutant TDP43 staining for individual cell lines in d3MNs

A) Representative images of Control and VCP-mutant d3MNs stained with TDP43 (red) and merged with DAPI staining (blue). B) Quantification of TDP43 staining, displayed as % of TDP43 in the nucleus vs cytoplasm. C) Percentage of TDP43 staining in the cytoplasm for Control and VCP mutant d3MN grouped data (Biological n=2 for control lines and n=4 for VCP mutant lines, two-way ANOVA). Error bars represent mean ± SEM.

References

1. Abramov, Andrey Y, Laura Canevari, and Michael R Duchen. 2003. “Changes in Intracellular Calcium and Glutathione in Astrocytes as the Primary Mechanism of Amyloid Neurotoxicity.” *The Journal of Neuroscience* 23(12): 5088–95.
2. Aggad, Dina, Julie Vérièpe, Arnaud Tauffenberger, and J Alex Parker. 2014. “TDP-43 Toxicity Proceeds via Calcium Dysregulation and Necrosis in Aging Caenorhabditis Elegans Motor Neurons.” *The Journal of Neuroscience* 34(36): 12093–103.
3. Alami, Nael H et al. 2014. “Axonal Transport of TDP-43 mRNA Granules in Neurons Is Impaired by ALS-Causing Mutations.” 81(3): 536–43.
4. Allen, Nicholas D. 2008. “Temporal and Epigenetic Regulation of Neurodevelopmental Plasticity.” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 363(February 2007): 23–38.
5. Almeida, Sandra et al. 2013. “Modeling Key Pathological Features of Frontotemporal Dementia with C9ORF72 Repeat Expansion in iPSC-Derived Human Neurons.” *Acta Neuropathologica* 126(3): 385–99.
6. Ambrogio, Andrea et al. 2009. “Functional Mapping of the Interaction between TDP-43 and hnRNP A2 in Vivo.” *Nucleic Acids Research* 37(12): 4116–26.
7. Amoroso, Mackenzie W et al. 2013. “Accelerated High-Yield Generation of Limb-Innervating Motor Neurons from Human Stem Cells.” *The Journal of Neuroscience* 33(2): 574–86.
8. Arai, Tetsuaki 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): 602–11.
9. Ayala, Youhna M et al. 2011. “TDP-43 Regulates Its mRNA Levels through a Negative Feedback Loop.” *The EMBO journal* 30(2): 277–88.
10. Ayala, Youhna M. et al. 2008. “Structural Determinants of the Cellular Localization and Shuttling of TDP-43.” *Journal of Cell Science* 121(22): 3778–85.
11. Baloh, Robert H. 2011. “TDP-43: The Relationship between Protein Aggregation and Neurodegeneration in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration.” *The FEBS journal* 278(19): 3539–49.
12. Baltz, Alexander G. et al. 2012. “The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts.” *Molecular Cell* 46(5): 674–90.
13. Barmada, Sami 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): 639–49.
14. Baron, Desiree M et al. 2013. “Amyotrophic Lateral Sclerosis-Linked FUS/TLS Alters Stress Granule Assembly and Dynamics.” *Molecular neurodegeneration* 8(1): 30.
15. Bartolome, Fernando et al. 2013. “Pathogenic VCP Mutations Induce

Mitochondrial Uncoupling and Reduced ATP Levels." *Neuron* 78: 57–64.
16. Bembich, Sara et al. 2014. "Predominance of Spliceosomal Complex Formation over Polyadenylation Site Selection in TDP-43 Autoregulation." *Nucleic acids research* 42(5): 3362–71.
17. Benatar, Michael et al. 2013. "Motor Neuron Involvement in Multisystem Proteinopathy: Implications for ALS." *Neurology* 80(20): 1874–80.
18. 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* 109(15): 5803–8.
19. Bilsland, Lynsey G et al. 2010. "Deficits in Axonal Transport Precede ALS Symptoms in Vivo." 107(47): 1–6.
20. Blokhuis, Anna M. et al. 2013. "Protein Aggregation in Amyotrophic Lateral Sclerosis." *Acta Neuropathologica* 125: 777–94.
21. Bories, Cyril, Julien Amendola, Boris Lamotte D'Incamps, and Jacques Durand. 2007. "Early Electrophysiological Abnormalities in Lumbar Motoneurons in a Transgenic Mouse Model of Amyotrophic Lateral Sclerosis." *European Journal of Neuroscience* 25(2): 451–59.
22. Bouhon, Isabelle a et al. 2006. "Embryonic Stem Cell-Derived Neural Progenitors Display Temporal Restriction to Neural Patterning." *Stem cells* 24(8): 1908–13.
23. Bramham, Clive R, and David G Wells. 2007. "Dendritic mRNA : Transport , Translation and Function." 8.
24. Buchan, J. Ross, Regina Maria Kolaitis, J. Paul Taylor, and Roy Parker. 2013. "XEukaryotic Stress Granules Are Cleared by Autophagy and Cdc48/VCP Function." *Cell* 153(7): 1461–74.
25. Budini, Mauricio, and Emanuele Buratti. 2011. "TDP-43 Autoregulation: Implications for Disease." *Journal of Molecular Neuroscience* 45(March): 473–79.
26. Buratti, Emanuele, and Francisco E. Baralle. 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): 36337–43.
27. Burkhardt, Matthew F. et al. 2013. "A Cellular Model for Sporadic ALS Using Patient-Derived Induced Pluripotent Stem Cells." *Molecular and Cellular Neuroscience* 56: 355–64.
28. Cahoy, JD et al. 2008. "A Transcriptome Database for Astrocytes, Neurons, and Oligodendrocytes: A New Resource for Understanding Brain Development and Function." *Culture* 28(1): 264–78.
29. Calder, Elizabeth L et al. 2015. "Retinoic Acid-Mediated Regulation of GLI3 Enables Efficient Motoneuron Derivation from Human ESCs in the Absence of Extrinsic SHH Activation." *The Journal of Neuroscience* 35(33): 11462–81.
30. Caldwell, Maeve A et al. 2001. "Growth Factors Regulate the Survival and Fate of Cells Derived from Human Neurospheres." *Nature biotechnology* 19(5): 475–79.
31. Camus, Anne, Aitana Perea-Gomez, Anne Moreau, and Jérôme Collignon. 2006. "Absence of Nodal Signaling Promotes Precocious Neural

Differentiation in the Mouse Embryo." *Developmental Biology* 295(2): 743–55.
32. Carrì, Maria Teresa, Cristiana Valle, Francesca Bozzo, and Mauro Cozzolino. 2015. "Oxidative Stress and Mitochondrial Damage: Importance in Non-SOD1 ALS." *Frontiers in cellular neuroscience* 9(February): 41.
33. Cassidy, Andrew, and Julia Jones. 2014. "Developments in in Situ Hybridisation." *Methods* 70(1): 39–45.
34. Chambers, Stuart M et al. 2009. "Highly Efficient Neural Conversion of Human ES and iPS Cells by Dual Inhibition of SMAD Signaling." *Nat Biotechnol* 27(3): 275–80.
35. Chaolin Zhang and Robert B. Darnell. 2011. "Mapping in Vivo Protein-RNA Interactions at Single-Nucleotide Resolution from HITS-CLIP Data." *Nature biotechnology* 29(7): 607–14.
36. Chataway, Jeremy. 2015. "Simvastatin in Patients with Progressive Multiple sclerosis–Authors' Reply." *The Lancet* 384(9947): 952–53.
37. Chen, Hong et al. 2014. "Modeling ALS with iPSCs Reveals That Mutant SOD1 Misregulates Neurofilament Balance in Motor Neurons." *Cell Stem Cell* 14(6): 796–809.
38. Chen, Sheng, Pavani Sayana, Xiaojie Zhang, and Weidong Le. 2013. "Genetics of Amyotrophic Lateral Sclerosis: An Update." *Molecular neurodegeneration* 8(1): 28.
39. Chiang, Po-min et al. 2010. "Deletion of TDP-43 down-Regulates Tbc1d1, a Gene Linked to Obesity, and Alters Body Fat Metabolism." *Proceedings of the National Academy of Sciences of the United States of America* 107(37): 16320–24.
40. Colombrita, Claudia et al. 2009. "TDP-43 Is Recruited to Stress Granules in Conditions of Oxidative Insult." *Journal of Neurochemistry* 111(4): 1051–61.
41. Cong, L et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *science reports* (February): 819–24.
42. Corrado, Lucia et al. 2009. "High Frequency of TARDBP Gene Mutations in Italian Patients with Amyotrophic Lateral Sclerosis." *Human Mutation* 30(4): 688–94.
43. Cozzolino, Mauro, and Maria Teresa Carri. 2012. "Mitochondrial Dysfunction in ALS." *Progress in Neurobiology* 97(2): 54–66.
44. D'Amico, Emanuele, Pam Factor-Litvak, Regina M Santella, and Hiroshi Mitsumoto. 2013. "Clinical Perspective on Oxidative Stress in Sporadic Amyotrophic Lateral Sclerosis." *Free radical biology & medicine* 65: 509–27.
45. Dalton, Ryan P., David B. Lyons, and Stavros Lomvardas. 2013. "Co-Opting the Unfolded Protein Response to Elicit Olfactory Receptor Feedback." *Cell* 155(2): 321–32.
46. Darnell, Robert B. 2013. "RNA Protein Interaction in Neurons."
47. Defects, Reveal Age-related Nucleocytoplasmic et al. 2015. "Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Article Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects." : 1–14.
48. Dejesus-hernandez, Mariely et al. 2011. "Expanded GGGGCC Hexanucleotide Repeat in Non-Coding Region of C9ORF72 Causes

Chromosome 9p-Linked Frontotemporal Dementia and Amyotrophic Lateral Sclerosis." *Neuron* 72(2): 245–56.
49. Deneen, Benjamin et al. 2006. "The Transcription Factor NFIA Controls the Onset of Gliogenesis in the Developing Spinal Cord." *Neuron* 52(6): 953–68.
50. Devlin, Anna-Claire et al. 2015. "Human iPSC-Derived Motoneurons Harbouring TARDBP or C9ORF72 ALS Mutations Are Dysfunctional despite Maintaining Viability." *Nature Communications* 6: 1–12.
51. Dewey, Colleen M. et al. 2010. "TDP-43 Is Directed to Stress Granules by Sorbitol, a Novel Physiological Osmotic and Oxidative Stressor." *Molecular and Cellular Biology* 31(5): 1098–1108.
52. Diehl, Adam G., and Alan P. Boyle. 2016. "Deciphering ENCODE." *Trends in Genetics* 32(4): 238–49.
53. Doyle, Michael, and Michael A Kiebler. 2011. "Mechanisms of Dendritic mRNA Transport and Its Role in Synaptic Tagging." *The EMBO journal* 30(17): 3540–52.
54. van Eersel, Janet et al. 2011. "Cytoplasmic Accumulation and Aggregation of TDP-43 upon Proteasome Inhibition in Cultured Neurons." *PLoS ONE* 6(7).
55. Elden, Andrew C et al. 2010. "Ataxin-2 Intermediate-Length Polyglutamine Expansions Are Associated with Increased Risk for ALS." *Nature* 466(August).
56. Elkabetz, Yechiel et al. 2008. "Human ES Cell-Derived Neural Rosettes Reveal a Functionally Distinct Early Neural Stem Cell Stage." *Genes and Development* 22: 152–65.
57. Esmaeili, Mohammad A. et al. 2013. "Premature Death of TDP-43 (A315T) Transgenic Mice due to Gastrointestinal Complications prior to Development of Full Neurological Symptoms of Amyotrophic Lateral Sclerosis." *International Journal of Experimental Pathology* 94(1): 56–64.
58. Fan, Guoping et al. 2005. "DNA Methylation Controls the Timing of Astrogliogenesis through Regulation of JAK-STAT Signaling." *Development (Cambridge, England)* 132: 3345–56.
59. Ferraiuolo, Laura et al. 2011. "Molecular Pathways of Motor Neuron Injury in Amyotrophic Lateral Sclerosis." *Nature Reviews Neurology* 7(11): 616–30.
60. Fonslow, Bryan R et al. 2012. "The Importance of Preclinial Timing - a Potential Reason for the Disconnect between Mouse Studies and Human Clincal Trials in ALS." 10(1): 54–56.
61. Freeman, Marc R. 2010. "Specification and Morphogenesis of Astrocytes." *Science (New York, N.Y.)* 330(6005): 774–78.
62. Fu, Xinrong, Christine Ng, Daorong Feng, and Chun Liang. 2003. "Cdc48p Is Required for the Cell Cycle Commitment Point at Start via Degradation of the G1-CDK Inhibitor Far1p." *Journal of Cell Biology* 163(1): 21–26.
63. Gandhi, Sonia, and Andrey Y. Abramov. 2012. "Mechanism of Oxidative Stress in Neurodegeneration." *Oxidative Medicine and Cellular Longevity* 2012.
64. Gendron, T.F., K.A. Josephs, and L Petrucelli. 2010. "Review: Transactive Response DNA-Binding Protein 43 (TDP-43): Mechanisms of Neurodegeneration." *Neuropathol Appl Neurobiol.* 36(2): 97–112.
65. Geser, Felix et al. 2009. "Amyotrophic Lateral Sclerosis, Frontotemporal

Dementia and beyond: The TDP-43 Diseases." 256(8): 1205–14.
66. ———. 2011. "NIH Public Access." *Pathology* 67(10): 1238–50.
67. Di Giorgio, Francesco Paolo et al. 2007. "Non-Cell Autonomous Effect of Glia on Motor Neurons in an Embryonic Stem Cell-Based ALS Model." *Nature neuroscience* 10(5): 608–14.
68. Di Giorgio, Francesco Paolo, Gabriella L. Boulting, Samuel Bobrowicz, and Kevin C. Eggan. 2008. "Human Embryonic Stem Cell-Derived Motor Neurons Are Sensitive to the Toxic Effect of Glial Cells Carrying an ALS-Causing Mutation." *Cell Stem Cell* 3(6): 637–48.
69. Glisovic, Tina, Jennifer L Bachorik, Jeongsik Yong, and Gideon Dreyfuss. 2008. "RNA-Binding Proteins and Post-Transcriptional Gene Regulation." 582(14): 1977–86.
70. Grunz, Horst, and Lothar Tacke. 1989. "Neural Differentiation of Xenopus Zueuis Ectoderm Takes Place after Disaggregation and Delayed Reaggregation without Inducer." 28: 211–17.
71. Gupta, K et al. 2012. "Human Embryonic Stem Cell Derived Astrocytes Mediate Non-Cell-Autonomous Neuroprotection through Endogenous and Drug-Induced Mechanisms." *Cell Death and Differentiation* 19(5): 779–87.
72. Guthrie, P B et al. 1999. "ATP Released from Astrocytes Mediates Glial Calcium Waves." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 19(2): 520–28.
73. Hafner, Markus et al. 2010. "Transcriptome-Wide Identification of RNA-Binding Protein and microRNA Target Sites by PAR-CLIP." *Cell* 141(1): 129–41.
74. Haidet-Phillips, Amanda M et al. 2011. "Astrocytes from Familial and Sporadic ALS Patients Are Toxic to Motor Neurons." *Nature biotechnology* 29(9): 824–28.
75. Hardiman, Orla, Leonard H. van den Berg, and Matthew C. Kiernan. 2011. "Clinical Diagnosis and Management of Amyotrophic Lateral Sclerosis." *Nature Reviews Neurology* 7(11): 639–49.
76. Hasegawa, M, T Arai, and T Nonaka. 2008. "Phosphorylated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis." *Annals of Neurology* 64(1): 60–70.
77. Hatada, Izuho et al. 2008. "Astrocyte-Specific Genes Are Generally Demethylated in Neural Precursor Cells prior to Astrocytic Differentiation." *PLoS ONE* 3(9): 1–9.
78. Hemmati-brivanlou, Ali, and Douglas A Melton. 1994. "Inhibition of Activin Receptor Signaling Promotes Neuralization in Xenopus." 77: 273–81.
79. Hirabayashi, M et al. 2001. "VCP/p97 in Abnormal Protein Aggregates, Cytoplasmic Vacuoles, and Cell Death, Phenotypes Relevant to Neurodegeneration." *Cell death and differentiation* 8(10): 977–84.
80. Hirabayashi, Y, and Y Gotoh. 2010. "Epigenetic Control of Neural Precursor Cell Fate during Development." *Nature Reviews Neuroscience* 11(6): 377–88.
81. Hotta, Kentaro et al. 2013. "Vesnarinone Suppresses TNFα mRNA Expression by Inhibiting Valosin-Containing Protein." *Molecular pharmacology* 83(May): 930–38.
82. Howland, David 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): 1604–9.
83. Hsueh, Yi-Ping. 2012. "From Neurodevelopment to Neurodegeneration: The Interaction of Neurofibromin and Valosin-Containing protein/p97 in Regulation of Dendritic Spine Formation." *Journal of biomedical science* 19(1): 33.
84. Huppertz, Ina et al. 2014. "iCLIP: Protein-RNA Interactions at Nucleotide Resolution." *Methods* 65(3): 274–87.
85. I.J.Huppertz. 2015. "Understanding the Dynamics of Ribonucleoprotein Assembly through iCLIP."
86. Ilieva, Ekaterina V. et al. 2007. "Oxidative and Endoplasmic Reticulum Stress Interplay in Sporadic Amyotrophic Lateral Sclerosis." *Brain* 130(12): 3111–23.
87. Ingre, Caroline et al. 2015. "Risk Factors for Amyotrophic Lateral Sclerosis." *Clinical Epidemiology* 7: 181–93.
88. Jiang, Wenyan et al. 2013. "Articles RNA-Guided Editing of Bacterial Genomes Using CRISPR-Cas Systems." *Nature Biotechnology* 31(3): 233–39.
89. Johnson, Brian S, J Michael McCaffery, Susan Lindquist, and Aaron D Gitler. 2008. "A Yeast TDP-43 Proteinopathy Model: Exploring the Molecular Determinants of TDP-43 Aggregation and Cellular Toxicity." *Proceedings of the National Academy of Sciences of the United States of America* 105(17): 6439–44.
90. Johnson, Brian S. et al. 2009. "TDP-43 Is Intrinsically Aggregation-Prone, and Amyotrophic Lateral Sclerosis-Linked Mutations Accelerate Aggregation and Increase Toxicity." *Journal of Biological Chemistry* 284(30): 20329–39.
91. Johnson, Janel O. et al. 2010. "Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS." *Neuron* 68: 857–64.
92. Johnson, Rory, Wendy Noble, Gian Gaetano Tartaglia, and Noel J. Buckley. 2012. "Neurodegeneration as an RNA Disorder." *Progress in Neurobiology* 99: 293–315.
93. Ju, Jeong-Sun et al. 2009. "Valosin-Containing Protein (VCP) Is Required for Autophagy and Is Disrupted in VCP Disease." *The Journal of Cell Biology* 187 (6): 875–88.
94. Kabashi, Edor et al. 2008. "TARDBP Mutations in Individuals with Sporadic and Familial Amyotrophic Lateral Sclerosis." *Nature Genetics* 40(5): 572–74.
95. Kapeli, Katannya, and Gene W. Yeo. 2012. "Genome-Wide Approaches to Dissect the Roles of RNA Binding Proteins in Translational Control: Implications for Neurological Diseases." *Frontiers in Neuroscience* 6(OCT): 1–15.
96. Kasai, Mana, Kiyotoshi Satoh, and Tetsu Akiyama. 2005. "Wnt Signaling Regulates the Sequential Onset of Neurogenesis and Gliogenesis via Induction of BMPs." *Genes to Cells* 10(8): 777–83.
97. Kiebler, Michael A., Peter Scheiffele, and Jernej Ule. 2013. "What, Where, and When: The Importance of Post-Transcriptional Regulation in the Brain." *Frontiers in Neuroscience* 7(192).
98. Kim, Hong Joo et al. 2013. "Mutations in Prion-like Domains in hnRNPA2B1

and hnRNPA1 Cause Multisystem Proteinopathy and ALS." *Nature* 495(7442): 467–73.
99. Kim, Nam Chul et al. 2014. "VCP Is Essential for Mitochondrial Quality Control by PINK1/ Parkin and This Function Is Impaired by VCP Mutations." 78(1): 65–80.
100. Kim, Sang Woo et al. 2010. "A Sensitive Non-Radioactive Northern Blot Method to Detect Small RNAs." *Nucleic Acids Research* 38(7): 1–7.
101. Kiskinis, Evangelos et al. 2014. "Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1." *Cell Stem Cell* 14(6): 781–95.
102. König, Julian et al. 2011. "iCLIP Reveals the Function of hnRNP Particles in Splicing at Individual Nucleotide Resolution." 17(7): 909–15.
103. König, Julian, Kathi Zarnack, Nicholas M. Luscombe, and Jernej Ule. 2012. "Protein–RNA Interactions: New Genomic Technologies and Perspectives." *Nature reviews. Genetics* 13(February): 3501–11.
104. Koppers, Max et al. 2012. "VCP Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis." *Neurobiology of Aging* 33(4): 837.e7–837.e13.
105. Kraemer, Brian C. et al. 2010. "Loss of Murine TDP-43 Disrupts Motor Function and Plays an Essential Role in Embryogenesis." *Acta Neuropathologica* 119(4): 409–19.
106. Krencik, Robert et al. 2011. "Specification of Transplantable Astroglial Subtypes from Human Pluripotent Stem Cells." *Nature biotechnology* 29(6): 528–34.
107. ———. 2015. "Dysregulation of Astrocyte Extracellular Signaling in Costello Syndrome." 7(286): 1–24.
108. Kühnlein, Peter et al. 2009. "Two German Kindreds with Familial Amyotrophic Lateral Sclerosis due to TARDBP Mutations." 65(9): 1185–89.
109. Kuo, J J, T Siddique, R Fu, and C J Heckman. 2005. "Increased Persistent Na(+) Current and Its Effect on Excitability in Motoneurones Cultured from Mutant SOD1 Mice." *The Journal of physiology* 563(Pt 3): 843–54.
110. Kwiatkowski, T J et al. 2009. "Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis." *Science* 323 (5918): 1205–8.
111. Lagier-Tourenne, Clotilde et al. 2012. "Divergent Roles of ALS-Linked Proteins FUS/TLS and TDP-43 Intersect in Processing Long Pre-mRNAs." *Nature neuroscience* 15(11): 1488–97.
112. Lagier-tourenne, Clotilde, Magdalini Polymenidou, and Don W Cleveland. 2010. "TDP-43 and FUS / TLS : Emerging Roles in RNA Processing and Neurodegeneration." 19(1): 46–64.
113. Lee, Edward B., Virginia M.-Y. Lee, and John Q. Trojanowski. 2012. "Gains or Losses: Molecular Mechanisms of TDP43-Mediated Neurodegeneration." *Nature Reviews Neuroscience* 13(1): 38–50.
114. Lee, Hyojin et al. 2007. "Directed Differentiation and Transplantation of Human Embryonic Stem Cell-Derived Motoneurons." *Stem cells* 25: 1931–39.
115. Li, Huei Ying et al. 2011. "Hyperphosphorylation as a Defense Mechanism to Reduce TDP-43 Aggregation." *PLoS ONE* 6(8).

116. Li, X.J. et al. 2008. “Directed Differentiation of Ventral Spinal Progenitors and Motor Neurons from Human Embryonic Stem Cells by Small Molecules.” *Stem Cells* 26(4): 886–93.
117. Li, Xue-Jun et al. 2005. “Specification of Motoneurons from Human Embryonic Stem Cells.” *Nature biotechnology* 23(2): 215–21.
118. Liachko, Nicole F., Chris R. Guthrie, and Brian C. Kraemer. 2011. “Phosphorylation Promotes Neurotoxicity in a C. Elegans Model of TDP-43 Proteinopathy Nicole.” 4(164): 16208–19.
119. Licatalosi, Donny D et al. 2008. “HITS-CLIP Yields Genome-Wide Insights into Brain Alternative RNA Processing.” *Nature* 456(7221): 464–69.
120. Lindholm, D, H Wootz, and L Korhonen. 2006. “ER Stress and Neurodegenerative Diseases.” *Cell death and differentiation* 13(3): 385–92.
121. Ling, Shuo Chien, Magdalini Polymenidou, and Don W. Cleveland. 2013. “Converging Mechanisms in Als and FTD: Disrupted RNA and Protein Homeostasis.” *Neuron* 79(3): 416–38.
122. Liscica, R. M. et al. 2008. “ALS and FTLD: Two Faces of TDP-43 Proteinopathy.” *European Journal of Neuroscience* 15(8): 772–80.
123. Liu, Lihua et al. 2012. “ER Stress Response during the Differentiation of H9 Cells Induced by Retinoic Acid.” *Biochemical and Biophysical Research Communications* 417(2): 738–43.
124. Lule, D., A. C. Ludolph, and A. G. Ludolph. 2008. “Neurodevelopmental and Neurodegenerative Diseases - Is There a Pathophysiological Link? Attention-Deficit/hyperactivity Disorder and Amyotrophic Lateral Sclerosis as Examples.” *Medical Hypotheses* 70(6): 1133–38.
125. Lykke-Andersen, Jens, and Eric J. Bennett. 2014. “Protecting the Proteome: Eukaryotic Cotranslational Quality Control Pathways.” *Journal of Cell Biology* 204(4): 467–76.
126. Majumder, Anirban et al. 2013. “Inhibition of DNA Methyltransferases and Histone Deacetylases Induces Astrocytic Differentiation of Neural Progenitors.” *Stem Cell Research* 11(1): 574–86.
127. Malhotra, J D, and R J Kaufman. 2011. “ER Stress and Its Functional Link to Mitochondria: Role in Cell Survival and Death.” *Cold Spring Harb Perspect Biol* 3(9): a004424.
128. Malik, Nasir et al. 2014. “Comparison of the Gene Expression Profiles of Human Fetal Cortical Astrocytes with Pluripotent Stem Cell Derived Neural Stem Cells Identifies Human Astrocyte Markers and Signaling Pathways and Transcription Factors Active in Human Astrocytes.” *PLoS ONE* 9(5).
129. Maragakis, Nicholas J, and Jeffrey D Rothstein. 2006. “Mechanisms of Disease: Astrocytes in Neurodegenerative Disease.” *Nature clinical practice. Neurology* 2(12): 679–89.
130. Maury, Yves, Julien Côme, Rebecca A Piskorowski, Nouzha Salah-mohellibi, et al. 2014. “Combinatorial Analysis of Developmental Cues Efficiently Converts Human Pluripotent Stem Cells into Multiple Neuronal Subtypes.” (November): 1–10.
131. Maury, Yves, Julien Côme, Rebecca a Piskorowski, Nouzha Salah-Mohellibi, et al. 2014. “Combinatorial Analysis of Developmental Cues Efficiently Converts Human Pluripotent Stem Cells into Multiple Neuronal Subtypes.” *Nat Biotechnol* 33(November): 89–96.

132. Metzker, Michael L. 2010. “Sequencing Technologies - the next Generation.” *Nature reviews. Genetics* 11(1): 31–46.
133. Meyer, Hemmo, and Conrad C Weihl. 2014. “The VCP/p97 System at a Glance: Connecting Cellular Function to Disease Pathogenesis.” *Journal of cell science* 127: 3877–83.
134. Meyer, Kathrin et al. 2014. “Direct Conversion of Patient Fibroblasts Demonstrates Non-Cell Autonomous Toxicity of Astrocytes to Motor Neurons in Familial and Sporadic ALS.” *Proceedings of the National Academy of Sciences of the United States of America* 111(2): 829–32.
135. Miller, Justine D. et al. 2013. “Human iPSC-Based Modeling of Late-Onset Disease via Progerin-Induced Aging.” *Cell Stem Cell* 13(6): 691–705.
136. Mizuno, Yuji, Seiji Hori, Akira Kakizuka, and Koichi Okamoto. 2003. “Vacuole-Creating Protein in Neurodegenerative Diseases in Humans.” *Neuroscience Letters* 343(2): 77–80.
137. Modic, Miha, Jernej Ule, and Christopher R. Sibley. 2013. “CLIPing the Brain: Studies of Protein-RNA Interactions Important for Neurodegenerative Disorders.” *Molecular and Cellular Neuroscience* 56: 429–35.
138. Molliex, Amandine et al. 2015. “Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization.” *Cell* 163(1): 123–33.
139. Moore, Michael J et al. 2014. “Mapping Argonaute and Conventional RNA-Binding Protein Interactions with RNA at Single-Nucleotide Resolution Using HITS-CLIP and CIMS Analysis.” *Nat. Protocols* 9(2): 263–93.
140. Moreno, Julie a et al. 2013. “Oral Treatment Targeting the Unfolded Protein Response Prevents Neurodegeneration and Clinical Disease in Prion-Infected Mice.” *Science translational medicine* 5(206): 206ra138.
141. Morrison, S J et al. 2000. “Transient Notch Activation Initiates an Irreversible Switch from Neurogenesis to Gliogenesis by Neural Crest Stem Cells.” *Cell* 101: 499–510.
142. Nagai, Makiko et al. 2007. “Astrocytes Expressing ALS-Linked Mutated SOD1 Release Factors Selectively Toxic to Motor Neurons.” *Nature neuroscience* 10: 615–22.
143. Nakamura, Haruko et al. 2015. “Quantitative Analysis of Intraneuronal Transport in Human iPS Neurons.” *Journal of Pharmacological Science* 128(4): 170–78.
144. Neumann, M et al. 2007. “TDP-43 in the Ubiquitin Pathology of Frontotemporal Dementia with VCP Gene Mutations.” *J Neuropathol Exp Neurol.* 66(2): 152–57.
145. Neumann, Manuela et al. 2006. “Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis.” *Science* 314 (5796): 130–33.
146. Nonaka, Takashi et al. 2009. “Truncation and Pathogenic Mutations Facilitate the Formation of Intracellular Aggregates of TDP-43.” *Human Molecular Genetics* 18(18): 3353–64.
147. Nonhoff, Ute et al. 2007. “Ataxin-2 Interacts with the DEAD/H-Box RNA Helicase DDX6 and Interferes with P-Bodies and Stress Granules.” *Molecular biology of the cell* 18(December): 1385–96.
148. Van Nostrand, Eric L et al. 2016. “Robust Transcriptome-Wide Discovery of

RNA-Binding Protein Binding Sites with Enhanced CLIP (eCLIP).” *Nature methods* 13(November 2015): 1–9.
149. Okada, Yohei et al. 2008. “Spatiotemporal Recapitulation of Central Nervous System Development by Murine Embryonic Stem Cell-Derived Neural Stem/Progenitor Cells.” *Stem Cells* 26(12): 3086–98.
150. Okita, Keisuke et al. 2011. “A More Efficient Method to Generate Integration-Free Human iPS Cells.” *Nature methods* 8(5): 409–12.
151. Pasinelli, Piera, and Robert H Brown. 2006. “Molecular Biology of Amyotrophic Lateral Sclerosis: Insights from Genetics.” *Nature reviews. Neuroscience* 7(September): 710–23.
152. Patani, R et al. 2011. “Retinoid-Independent Motor Neurogenesis from Human Embryonic Stem Cells Reveals a Medial Columnar Ground State.” *Nature communications* 2: 214.
153. Patani, Rickie, Patrick A. Lewis, et al. 2012. “Investigating the Utility of Human Embryonic Stem Cell-Derived Neurons to Model Ageing and Neurodegenerative Disease Using Whole-Genome Gene Expression and Splicing Analysis.” *Journal of Neurochemistry* (122): 738–51.
154. Patani, Rickie, Christopher R. Sibley, Siddharthan Chandran, and Jernej Ule. 2012. “Using Human Pluripotent Stem Cells to Study Post-Transcriptional Mechanisms of Neurodegenerative Diseases.” *Brain Research* 1462: 129–38.
155. Pekny, Milos, and Marcela Pekna. 2014. “Astrocyte Reactivity and Reactive Astrogliosis: Costs and Benefits.” *Physiological reviews* 94(4): 1077–98.
156. Philippidou, Polyxeni, and Jeremy S. Dasen. 2013. “Hox Genes: Choreographers in Neural Development, Architects of Circuit Organization.” *Neuron* 80(1): 12–34.
157. Pringle, Nigel P et al. 2003. “Fgfr3 Expression by Astrocytes and Their Precursors: Evidence That Astrocytes and Oligodendrocytes Originate in Distinct Neuroepithelial Domains.” *Development (Cambridge, England)* 130(1): 93–102.
158. Pun, San et al. 2006. “Selective Vulnerability and Pruning of Phasic Motoneuron Axons in Motoneuron Disease Alleviated by CNTF.” *Nature neuroscience* 9(3): 408–19.
159. Raponi, Eric et al. 2007. “S100B Expression Defines a State in Which GFAP- Expressing Cells Lose Their Neural Stem Cell Potential and Acquire a More Mature Developmental Stage.” *Glia* 55(14): 1416–25.
160. Renton, Alan E et al. 2011. “A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD Alan.” *Neuron* 72(2): 257–68.
161. Reubinoff, Benjamin E et al. 2001. “Neural Progenitors from Human Embryonic Stem Cells.” *Nature Biotechnology* 19(12): 1134–40.
162. Ritson, Gillian P et al. 2010. “TDP-43 Mediates Degeneration in a Novel Drosophila Model of Disease Caused by Mutations in VCP/p97.” *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30(22): 7729–39.
163. Rodriguez-Ortiz, Carlos J. et al. 2013. “Neuronal-Specific Overexpression of a Mutant Valosin-Containing Protein Associated with Ibmpfd Promotes Aberrant Ubiquitin and TDP-43 Accumulation and Cognitive Dysfunction in

Transgenic Mice." *American Journal of Pathology* 183(2): 504–15.
164. Rogelj, Boris et al. 2012. "Widespread Binding of FUS along Nascent RNA Regulates Alternative Splicing in the Brain." *Scientific Reports* 2: 1–10.
165. Rossi, D et al. 2008. "Focal Degeneration of Astrocytes in Amyotrophic Lateral Sclerosis." *Cell death and differentiation* 15(11): 1691–1700.
166. Rothstein, J D et al. 1995. "Selective Loss of Glial Glutamate Transporter GLT-1 in Amyotrophic Lateral Sclerosis." *Annals of neurology* 38(1): 73–84.
167. Rothstein, Jeffrey D et al. 2005. "Beta-Lactam Antibiotics Offer Neuroprotection by Increasing Glutamate Transporter Expression." *Nature* 433(7021): 73–77.
168. Rothstein, Jeffrey D. et al. 1996. "Knockout of Glutamate Transporters Reveals a Major Role for Astroglial Transport in Excitotoxicity and Clearance of Glutamate." *Neuron* 16(3): 675–86.
169. Roybon, Laurent et al. 2013. "Human Stem Cell-Derived Spinal Cord Astrocytes with Defined Mature or Reactive Phenotypes." *Cell Reports* 4(5): 1035–48.
170. Rumpf, Sebastian et al. 2014. "Drosophila Valosin-Containing Protein Is Required for Dendrite Pruning through a Regulatory Role in mRNA Metabolism." *Proceedings of the National Academy of Sciences of the United States of America* 111(20): 7331–36.
171. Sasaki, S. 2010. "Endoplasmic Reticulum Stress in Motor Neurons of the Spinal Cord in Sporadic Amyotrophic Lateral Sclerosis." *J Neuropathol Exp Neurol* 69(4): 346–55.
172. Sasaki, Shoichi, and Makoto Iwata. 1995. "Synaptic Loss in the Proximal Axon of Anterior Horn Neurons in Motor Neuron Disease." *Acta Neuropathologica* 90(2): 170–75.
173. Sasaki, Shoichi, and Shoichi Maruyama. 1994. "Synapse Loss in Anterior Horn Neurons in Amyotrophic Lateral Sclerosis." *Acta Neuropathologica* 88(3): 222–27.
174. Saxena, Smita, Erik Cabuy, and Pico Caroni. 2009. "A Role for Motoneuron Subtype-Selective ER Stress in Disease Manifestations of FALS Mice." *Nature neuroscience* 12(5): 627–36.
175. Schmitt, Angelika et al. 2003. "The Brain-Specific Protein MLC1 Implicated in Megalencephalic Leukoencephalopathy with Subcortical Cysts Is Expressed in Glial Cells in the Murine Brain." *Glia* 44(3): 283–95.
176. Scotter, Emma L., Han Jou Chen, and Christopher E. Shaw. 2015. "TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets." *Neurotherapeutics* 12(2): 352–63.
177. Serio, Andrea et al. 2013. "Astrocyte Pathology and the Absence of Non-Cell Autonomy in an Induced Pluripotent Stem Cell Model of TDP-43 Proteinopathy." *Proceedings of the National Academy of Sciences of the United States of America* 110(12): 4697–4702.
178. Shaltouki, Atossa et al. 2013. "Efficient Generation of Astrocytes from Human Pluripotent Stem Cells in Defined Conditions." *Stem Cells* 31(5): 941–52.
179. Shaw, Pamela J. et al. 1995. "CSF and Plasma Amino Acid Levels in Motor Neuron Disease: Elevation of CSF Glutamate in a Subset of Patients." *Neurodegeneration* 4(2): 209–16.

180. Shi, Yichen, Peter Kirwan, and Frederick J Livesey. 2012. “Directed Differentiation of Human Pluripotent Stem Cells to Cerebral Cortex Neurons and Neural Networks.” *Nat. Protocols* 7(10): 1836–46.
181. Shih, Yu-Tzu, and Yi-Ping Hsueh. 2016. “VCP and ATL1 Regulate Endoplasmic Reticulum and Protein Synthesis for Dendritic Spine Formation.” *Nature Communications* 7(May 2015): 11020.
182. Sitte, Nicolle, Michael Huber, Tilman Grune, and Axel Ladhoff. 1999. “Proteasome Inhibition by Lipofuscin / Ceroid during Postmitotic Aging of Fibroblasts.” *The FASEB Journal*.
183. Sleigh, James, and Giampietro Schiavo. 2016. “Older but Not Slower: Aging Does Not Alter Axonal Transport Dynamics of Signalling Endosomes in Vivo   ” *Matters* (July).
184. Sofroniew, Michael V., and Harry V. Vinters. 2010. “Astrocytes: Biology and Pathology.” *Acta Neuropathologica* 119(1): 7–35.
185. Soto, Claudio. 2013. “Protein Misfolding in Neurodegenerative Diseases : The Key Pending Questions.” : 19–22.
186. Spriggs, Keith a., Martin Bushell, and Anne E. Willis. 2010. “Translational Regulation of Gene Expression during Conditions of Cell Stress.” *Molecular Cell* 40(2): 228–37.
187. Sreedharan, Jemeen et al. 2008. “TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis.” *Science* 319 (5870): 1668–72.
188. Stern, Claudio D. 2005. “Neural Induction: Old Problem, New Findings, yet More Questions.” *Development (Cambridge, England)* 132: 2007–21.
189. Takahashi, Kazutoshi, and Shinya Yamanaka. 2006. “Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.” *Cell* 126: 663–76.
190. Tang, Yong et al. 2013. “Fast Vesicle Transport Is Required for the Slow Axonal Transport of Synapsin.” *Journal of Neuroscience* 33(39): 15362–75.
191. Tollervey, James R et al. 2011. “UKPMC Funders Group Characterising the RNA Targets and Position-Dependent Splicing Regulation by TDP-43 ; Implications for Neurodegenerative Diseases.” *October* 14(4): 452–58.
192. Tong, Jianbin et al. 2013. “Expression of ALS-Linked TDP-43 Mutant in Astrocytes Causes Non-Cell-Autonomous Motor Neuron Death in Rats.” *The EMBO journal* 32(13): 1917–26.
193. Tsai, Kuen-Jer et al. 2010. “Elevated Expression of TDP-43 in the Forebrain of Mice Is Sufficient to Cause Neurological and Pathological Phenotypes Mimicking FTLD-U.” *The Journal of experimental medicine* 207(8): 1661–73.
194. Turner, Martin R et al. 2013. “Controversies and Priorities in Amyotrophic Lateral Sclerosis.” *The Lancet Neurology* 12(3): 310–22.
195. Tyzack, Giulia E et al. 2014. “Astrocyte Response to Motor Neuron Injury Promotes Structural Synaptic Plasticity via STAT3-Regulated TSP-1 Expression.” *Nature communications* 5: 4294.
196. Ule, Jernej et al. 2003. “CLIP Identi Es Nova-Regulated RNA Networks in the Brain.” *Science* 302(November): 1212–15.
197. ———. 2010. “iCLIP Predicts the Dual Splicing Effects of TIA-RNA Interactions.” *PLoS Biology* 8(10).
198. Ule, Jernej, Kirk Jensen, Aldo Mele, and Robert B. Darnell. 2005. “CLIP: A

Method for Identifying Protein-RNA Interaction Sites in Living Cells." *Methods* 37(4): 376–86.
199. Vance, Caroline et al. 2009. "Mutations in FUS , an RNA Processing Protein , Cause Familial Amyotrophic Lateral Sclerosis Type 6." *Science* 323(5918): 1208–11.
200. Vanderweyde, Tara, Katie Youmans, Liqun Liu-Yesucevitz, and Benjamin Wolozin. 2013. "Role of Stress Granules and RNA-Binding Proteins in Neurodegeneration: A Mini-Review." *Gerontology* 59: 524–33.
201. Vaz, Bruno, Swagata Halder, and Kristijan Ramadan. 2013. "Role of p97/VCP (Cdc48) in Genome Stability." *Frontiers in genetics* 4(April): 60.
202. Voigt, Aaron et al. 2010. "TDP-43-Mediated Neuron Loss In Vivo Requires RNA-Binding Activity." *PLoS ONE* 5(8).
203. Wainger, Brian J, Evangelos Kiskinis, Cassidy Mellin, Ole Wiskow, S W Steve, et al. 2014. "Intrinsic Membrane Hyperexcitability of ALS Patient-Derived Motor Neurons." *Cell Reports* 7(1): 1–11.
204. Wainger, Brian J., Evangelos Kiskinis, Cassidy Mellin, Ole Wiskow, Steve S W Han, et al. 2014. "Intrinsic Membrane Hyperexcitability of Amyotrophic Lateral Sclerosis Patient-Derived Motor Neurons." *Cell Reports* 7(1): 1–11.
205. Walker, Adam K. et al. 2013. "ALS-Associated TDP-43 Induces Endoplasmic Reticulum Stress, Which Drives Cytoplasmic TDP-43 Accumulation and Stress Granule Formation." *PLoS ONE* 8(11): 1–12.
206. Wang, Hsiao Fang et al. 2011. "Valosin-Containing Protein and Neurofibromin Interact to Regulate Dendritic Spine Density." *Journal of Clinical Investigation* 121(12): 4820–37.
207. Wang, Wenzhang et al. 2016. "The Inhibition of TDP-43 Mitochondrial Localization Blocks Its Neuronal Toxicity." *Nature Medicine* (May).
208. Wegorzewska, Iga et al. 2009. "TDP-43 Mutant Transgenic Mice Develop Features of ALS and Frontotemporal Lobar Degeneration." *Proceedings of the National Academy of Sciences of the United States of America* 106(44): 18809–14.
209. Wegorzewska, Iga, and Robert H. Baloh. 2011. "TDP-43-Based Animal Models of Neurodegeneration: New Insights into ALS Pathology and Pathophysiology." *Neurodegenerative Diseases* 8(4): 262–74.
210. Weiduschat, N. et al. 2014. "Motor Cortex Glutathione Deficit in ALS Measured in Vivo with the J-Editing Technique." *Neuroscience Letters* 570: 102–7.
211. Weihl, Conrad C. 2011. "Another VCP Interactor: NF Is Enough." *Journal of Clinical Investigation* 121(12): 4627–30.
212. Wen, Shu, Hong Li, and Jia Liu. 2009. "Dynamic Signaling for Neural Stem Cell Fate Determination." *Cell adhesion & migration* 3(1): 107–17.
213. Wichterle, H, I Lieberam, J a Porter, and T M Jessell. 2002. "Directed Differentiation of Embryonic Stem Cells into Motor Neurons." *Cell* 110: 385–97.
214. Winton, Matthew J. et al. 2008. "A90V TDP-43 Variant Results in the Aberrant Localization of TDP-43 in Vitro." *FEBS Letters* 582(15): 2252–56.
215. ———. 2009. "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): 13302–9.

216. Xu, Chunhui et al. 2001. “Feeder-Free Growth of Undifferentiated Human Embryonic Stem Cells.” 19(October): 971–74.
217. Xu, Ren-He et al. 2005. “Basic FGF and Suppression of BMP Signaling Sustain Undifferentiated Proliferation of Human ES Cells.” *Nature methods* 2(3): 185–90.
218. Yamanaka, Kunitoshi, Yohei Sasagawa, and Teru Ogura. 2012. “Recent Advances in p97/VCP/Cdc48 Cellular Functions.” *Biochimica et Biophysica Acta - Molecular Cell Research* 1823(1): 130–37.
219. Yerbury, Justin J. et al. 2016. “Walking the Tightrope: Proteostasis and Neurodegenerative Disease.” *Journal of Neurochemistry*: 1–17.
220. Yi, Ling et al. 2012. “Altered Intracellular Localization and Valosin-Containing Protein (p97 VCP) Interaction Underlie ATP7A-Related Distal Motor Neuropathy.” *Human Molecular Genetics* 21(8): 1794–1807.
221. Yin, H Z et al. 2012. “Slow Development of ALS-like Spinal Cord Pathology in Mutant Valosin-Containing Protein Gene Knock-in Mice.” *Cell death & disease* 3: e374.
222. Zarnack, Kathi et al. 2013. “Direct Competition between hnRNP C and U2AF65 Protects the Transcriptome from the Exonization of Alu Elements.” *Cell* 152(3): 453–66.
223. Zarnegar, Brian J et al. 2016. “irCLIP Platform for Efficient Characterization of protein–RNA Interactions.” *Nature methods* (April 2016).
224. Zhang, Su-Chun et al. 2001. “In Vitro Differentiation of Transplantable Neural Precursors from Human Embryonic Stem Cells.” *Nature Biotechnology* 19(12): 1129–33.
225. Zhang, Xiaochu, Eva Szabo, Marek Michalak, and Michal Opas. 2007. “Endoplasmic Reticulum Stress during the Embryonic Development of the Central Nervous System in the Mouse.” 25: 455–63.
226. Zhang, Ye et al. 2016. “Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse.” *Neuron* 89: 37–53.
227. Zhang, Yong-jie et al. 2009. “Aberrant Cleavage of TDP-43 Enhances Aggregation and Cellular Toxicity.” 106(18).
228. Zhang, Yong-Jie et al. 2010. “Phosphorylation Regulates Proteasomal-Mediated Degradation and Solubility of TAR DNA Binding Protein-43 C-Terminal Fragments.” *Molecular Neurodegeneration* 5(33): 1–13.
229. Zhang, Zhen-Ning et al. 2016. “Layered Hydrogels Accelerate iPSC-Derived Neuronal Maturation and Reveal Migration Defects Caused by MeCP2 Dysfunction.” *Proceedings of the National Academy of Sciences* 113(12): 201521255.
230. Zhou, Zhi Dong, Udhaya Kumari, Zhi Cheng Xiao, and Eng King Tan. 2010. “Notch as a Molecular Switch in Neural Stem Cells.” *IUBMB Life* 62(8): 618–23.
231. Zirra, Alexandra, Sarah Wiethoff, and Rickie Patani. 2016. “Neural Conversion and Patterning of Human Pluripotent Stem Cells : A Developmental Perspective System : An Overview.” 2.