Investigating the functional consequences of widespread RNA dysregulation in TDP-43 related Amyotrophic Lateral Sclerosis cell models

Jennifer Elizabeth Dodd

Thesis submitted for Doctor of Philosophy (PhD)

Supervisors: Dr Guillaume Hautbergue and Professor Dame Pamela Shaw

Sheffield Institute for Translational Neuroscience (SITraN)

Department of Neuroscience

University of Sheffield

Date of submission: 5th January 2017

Abstract

Dysregulation of RNA metabolism constitutes a major determinant in most fatal neurodegenerative diseases including amyotrophic lateral sclerosis (ALS). Many research groups have characterised hundreds to thousands of abnormal transcript levels/isoforms in cell and animal models of ALS, as well as in post-mortem brains from patients. However, these studies did not (i) allow discrimination of pathophysiological alterations from those due to subsequent dysregulations and (ii) did not investigate whether altered mRNA production is reflected at the protein level. This study tested current translatome methodologies, and further developed a high-throughput methodology, GRASPS (Genome-wide RNA Analysis of Stalled Protein Synthesis), to identify selectively enriched translating mRNA molecules based on stringent precipitation of ribosomes prior to next generation RNA sequencing. High purity of GRASPS-purified ribosomes was validated by mass spectrometry analysis. The nuclear loss and cytoplasmic mislocalisation of TDP-43 protein referred to, as TDP-43 proteinopathy, is a hallmark of ALS. TDP-43 is a DNA/RNA binding protein involved in multiple regulatory steps of gene expression. In disease, mutations in *TARDBP*, the gene encoding TDP-43, result in the production of thousands of abnormal levels/isoforms of transcripts however the functional consequences remain unknown. Therefore, GRASPS was applied to human TDP-43 ALS-inducible cell models. The first human TDP-43 ALS translatome was in parallel compared to whole-cell and cytoplasmic transcriptomes. Strikingly, in contrast to transcriptomes, the translatome findings highlighted novel altered biological processes that were highly enriched and directly relevant to neurodegeneration. These included altered nuclear export of mRNA and impaired neuronal differentiation. Functional validation showed that TDP-43 proteinopathy requires a neuronal context and is associated to the abnormal nuclear loss and cytoplasmic accumulation of bulk mRNA.

Acknowledgements

Firstly, I would like to thank my supervisors Dr Guillaume Hautbergue and Professor Dame Pamela Shaw for their invaluable support and inspiration during my PhD. I would also like to thank the Moody Endowment, SITraN and the Motor Neurone Disease Association (MNDA) for funding.

Also a big thank you to everyone in RNA Biology: Dr Matthew Walsh, Simeon Mihaylov, Dr Lydia Castelli and Dr Ya-Hui Lin.

Another big thanks to everyone in SITraN. In particular, Dr Adrian Higginbottom and Dr Matthew Stopford for their support with the cell model generation, Matthew Wyles, Catherine Gelsthorpe, Dr Paul Heath, Dr Robin Highley, Prof Paul Ince and Dr Johnathan Cooper-Knock.

I would also like to thank Christianne Fowler from the Centre for Genomic Research, Liverpool for the assistance with the next generation RNA sequencing and Dr Mark Dickman from the Department of Chemical Engineering at the Life Science Interface, University of Sheffield, for the mass spectrometry. Also Dr Marta Milo, Dr Luisa Cutillo, Ilaria Granata and Mario Guarracino for the Bioinformatics analysis.

Also a big thank you to all my new friends and colleagues at the University of Liverpool for helping with the last minute stresses completing this work.

And last, but by no means least, a massive thank you to my fantastic parents, to Ellie and Luke and to my family and friends who have put up with me throughout my PhD!

Table of contents

List of figures

List of tables

Abbreviations

AD	Alzheimer's disease
ALS	amyotrophic lateral sclerosis
BB	Bunina bodies
CHX	cycloheximide
CLIP	crosslinking immunoprecipitation
CMV	cytomegalovirus
CNS	central nervous system
CTF	C-terminal fragment
DAVID	Database for Annotation, Visualisation and Integrated Discovery
DEG	differentially expressed gene/s
DEPC	diethyl pyrocarbonate
dNTPs	deoxynucleotide
ECL	enhanced chemiluminescence
eEF2	eukaryotic translation elongation factor 2
eIF4a	eukaryotic initiation factor 4A
EJC	exon junction complex
FC	fold change
FRT	Flp recombinase target
FTD	frontotemporal dementia
FTLD-U	frontotemporal lobar degeneration with ubiquitinated inclusions
GFP	green fluorescent protein
GOI	gene of interest
GRASPS	Genome-wide RNA Analysis of Stalled Protein Synthesis
h	hour(s)
HCI	hyaline conglomerate inclusion
HD	Huntington's disease
hnRNP	heterogeneous RNP
HRP	horseradish peroxide
HSP	hereditary spastic paraplegia
IF	immunofluorescence
iNPC	induced neuronal progenitor cells
IP	immunoprecipitation
iPSC	induced pluripotent stem cells
kbp	kilobase pair
KIF	kinesin superfamily of proteins
min	minute
miRNA	micro RNA
MN	motor neuron
MND	motor neuron disease
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NCI	neuronal cytoplasmic inclusions
NES	nuclear export signal
NGO	no-go decay
NGS	Next Generation Sequencing
NLS	nuclear localisation signal
NMD	nonsense mediated decay
NNI	neuronal intranuclear inclusions
NPC	nuclear pore complex

NSD	non-stop decay
nt	nucleotide
PBS	phosphate buffered saline
PD	Parkinson's disease
PLD	prion-like domain
PLS	primary lateral sclerosis
PMA	progressive muscular atrophy
Poly(A)$^+$	polyadenylated
PTC	premature termination codon
p-TDP-43	phosphorylated TDP-43
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RAN	repeat associated non-ATG
RBP	RNA binding protein
RIP	RNA immunoprecipitation
RISC	RNA induced silencing complex
RT	room temperature
RNAPII	RNA polymerase II
RNP	ribonucleoprotein
ROS	reactive oxygen species
RPL	ribosomal protein
RRM	RNA recognition motif
rRNA	ribosomal RNA
tRNA	transfer RNA
SCA	spinocerebellar ataxia
sec	second(s)
SG	stress granule
siRNA	short interfering RNA
SMA	spinal muscular atrophy
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
TBST	Tris Buffered Saline with Tween-20
Tet	Tetracycline
TetR	Tetracycline resistance repressor elements
TO	Tetracycline operator
TRAP	Translating ribosome affinity purification
TREX	Transcription/export complex
UBI	ubiquitylated inclusion
UPS	ubiquitin-proteasome system
V	volts
v/v	volume/volume
WB	western blot
WT	wild-type
w/v	weight/volume
YFP	yellow fluorescent protein

1. Introduction

Motor neuron diseases are a spectrum of diseases causing the degeneration of motor neurons and the human motor system. The motor system is a group of cells that control voluntary muscle actions for general body movement. The spectrum of motor neuron diseases includes hereditary spastic paraplegias (HSP), spinal muscular atrophy (SMA) and motor neuron disease (MND). MND can be further categorised into amyotrophic lateral sclerosis (ALS), primary lateral sclerosis (PLS) and progressive muscular atrophy (PMA).

ALS is a fatal neurodegenerative disease of the human motor system caused by progressive injury and cell death of lower motor neurons in the spinal cord and brainstem, and upper motor neurons in the motor cortex. The annual incidence rate of ALS is 4-6 per 100,000, with an average age of onset between 50 and 60 years (Shaw, 2005). Rare juvenile variants of ALS, also exist, including ALS2 (Hadano et al., 2001; Y. Yang et al., 2001).

ALS is a devastating disease generally resulting in death within 3-5 years from the onset of symptoms. The causes of motor neuron injury in ALS are still being elucidated. Dysregulation of RNA synthesis, processing and axonal transport have more recently been reported to play an important pathophysiological component in disease and are a common feature of many neurodegenerative diseases (Walsh et al., 2015). However, the biological consequences of widespread dysregulation of mRNA metabolism on protein levels have not yet been investigated. For instance, it is unknown whether abnormally spliced transcripts get exported in the cytoplasm and how altered levels of transcripts impact upon the ALS translatome.

1.1 RNA and disease

1.1.1 Eukaryotic expression of genes

In eukaryotes, gene expression is the result of several processes starting with transcription in the nucleus and ending with translation in the cytoplasm. Gene expression depends upon RNA (messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and microRNA (miRNA)) to allow the correct synthesis of proteins. Multiple other species of non-coding RNAs are also believed to play regulatory roles. This is essential for the healthy development and survival of an organism. By tightly regulating the production of RNA a cell can change the expression of its genes at any given time. Control of gene expression is achieved using a series of quality control pathways, RNA surveillance and regulated decay pathways (Luna et al., 2008; Lebreton et al., 2008; Rodríguez-Navarro et al., 2011; Kervestin et al., 2012).

Within a cell, RNA binding proteins (RBP) and RNA bind directly, to form ribonucleoprotein complexes (RNPs). The composition of RNPs reflect both the function and fate of the corresponding RNA (Dreyfuss et al., 1993; Glisovic et al., 2008). The messenger RNP (mRNP) complex controls the multi-component processes involved in gene expression; these include the RNA polymerase II (RNAPII) transcription of pre-mRNA, pre-mRNA processing, mRNA nuclear export and translation of mRNA into proteins (Luna et al., 2008; M. Schmid et al., 2008).

The biogenesis of mRNA occurs in the nucleus in a number of separate, extensively coupled steps (Maniatis et al., 2002). Transcription is the first step in the RNP complex lifecycle. Pre-mRNA are transcribed by RNAPII from genomic DNA (Hsin et al., 2012). As the pre-mRNA emerges from RNAPII it undergoes co-transcriptional RNA processing. The 5' end is capped with a 7-methyl guanine nucleotide, by three

enzymes. This 5' methyl guanosine cap at the 5' end of the mRNA helps distinguish it from other RNA molecules and enables it to be properly processed and exported, as RNAPII processed RNA represent the major RNA type capped in this way (Hsin et al., 2012). Nascent pre-mRNA are also cleaved and polyadenylated at the 3' end (Tian et al., 2013; Elkon et al., 2013).

The vast majority of exons form the protein coding regions of genes. Exons are interrupted by long non-coding sequences termed introns. Pre-mRNA is spliced into its mature form, mRNA, by the fusion of exons and excision of introns. Post-splicing, the spliceosome directs a set of proteins to bind at the position previously occupied by the intron, these proteins make up the exon junction complex (EJC) and mark a successful splicing event (Dreyfuss et al., 2002). This is a surveillance mechanism, which triggers the decay of incorrectly spliced RNA. When mutations occur in a sequence critical for splicing, a new pattern of splicing is followed in a process called exon skipping. The spliceosome picks out the next best pattern of splicing, and usually an exon is skipped.

The spliceosome is a dynamic machine, containing many loosely associated proteins which may only be required in response to a given signal, enabling splicing of a wide variety of mRNA introns (Wahl et al., 2009). The majority of exons are constitutive and are always spliced into the mRNA product. However, some pre-mRNAs have multiple splicing positions giving rise to a family of related proteins from a single gene (Black, 2003). Alternative splicing adds to the diversity of gene expression, through the excision of different introns from the same transcript. On average one gene gives 10 alternatively spliced isoforms. It is a complex genetic switch, used to affect the quantitative control of gene expression and enhance the functional diversity of proteins from a given gene (Lopez, 1998).

Correctly processed mRNAs are identified by the presence of the 5' cap, the poly-A 3' tail and the EJC, this produces an export competent RNP, which can be exported into the cytoplasm via a nuclear pore complex (NPC) (Köhler et al., 2007; Carmody et al., 2009; Rodríguez-Navarro et al., 2011). The transcription-export (TREX) complex, which interacts with elongating RNAPII, couples the co-transcriptional processing of pre-mRNA with the nuclear export of mature mRNA acting as a binding platform for the remodelling of the export receptor NXF1 for the subsequent transport of mRNA through the NPC (Viphakone et al., 2012).

The targeting of mRNAs to specific cellular compartments is particularly important for neurons and allows the transport of mRNA to dendrites. There are at least three types of RNA-containing granule in mature neurons. Transport granules maintain mRNA in a transitionally silent state until they are at the site of protein synthesis; stress granules capture mRNA until it is required to be translated, for instance in response to neuronal injury, and P-bodies, also known as degradative granules, which degrade mRNA. The kinesin superfamily of proteins (KIFs) enables the transport of granules along microtubules in neurons (Kanai et al., 2004; Hirokawa, 2006).

The translation of the genetic code carried by mRNA into a protein occurs at the ribosome. The ribosome is made up from ribosomal proteins and rRNA (Yonath, 2009). tRNA translates the genetic code, bringing together the amino acids, to produce the correct amino acid sequence. Its cloverleaf shape enables it to do this with an amino acid attachment site at one end and the anticodon end at the other. The anticodon is able to recognise the codons on the mRNA. One molecule of mRNA is usually translated by multiple ribosomes.

Post-transcriptional controls regulate this process with tight surveillance and quality control pathways (Lebreton et al., 2008; M. Schmid et al., 2008; Kervestin et al., 2012). The degradation of RNA in the nucleus and cytoplasm is carried out by the exosome (Luna et al., 2008; Lebreton et al., 2008) or in P-bodies. Three pathways operate during translation to prevent the accumulation of aberrant mRNA; nonsense-mediated decay (NMD); non-stop decay (NSD) and no-go decay (NGD) (Kervestin et al., 2012). NMD is activated when mRNA reaches a premature termination codon (PTC), a stop codon located before the EJC. The resulting mRNA is then defective and is exported. NSD targets mRNAs, which lack a stop codon and NGD targets mRNA bound by ribosomes, which are stuck in elongation.

miRNA and short interfering RNA (siRNA) regulate translation and mRNA degradation in a pathway referred to as RNA interference (RNAi). These RNA are assembled with proteins to form a RNA induced silencing complex (RISC), which interacts with mRNA targets via complementary nucleotide binding (Gregory et al., 2005; Valencia-Sanchez et al., 2006). Once bound to targets endonucleolytic cleavage or translational repression of mRNA occurs.

The proteasome is responsible for regulatory and quality control proteolytic, protein degradation. Proteins marked by ubiquitin or phosphates are degraded in specific-mediated protein degradation to regulate the cell cycle. It also destroys potentially harmful proteins, which have not folded correctly or fold slowly, using ATP-dependent protease enzymes, in the cytosol and nucleus.

1.1.2 Alterations in RNA metabolism

The various RNA metabolic processes driving and regulating the expression of genes do not happen in a simple cascade of events, but rather occur concurrently at least in the nucleoplasm. The complexity and number of molecules involved in each

step allow for the accurate production and translation of mRNA. This complexity, although essential, allows for mutations and the dysregulation of RNA processing. Mutations that affect either the RNA or RBPs within the RNP complex can cause diseases. Alterations of transcription and pre-mRNA splicing (Cooper et al., 2009; Singh, 2012) as well as dysregulation of protein synthesis (Le Quesne et al., 2010) have been associated with several neurodegenerative diseases including Huntington's disease (HD), Alzheimer's disease (AD), SMA and ALS. A summary of diseases associated with alterations in RNA and RNPs is shown in Table 1.1.

1.1.3 RNA-mediated neurodegeneration

Neurodegeneration is characterised by the progressive loss of neuronal function and neurons. Altered RNA metabolism is a key pathophysiological mechanism causing several neurodegenerative diseases including HD, spinocerebellar ataxias (SCAs), SMA and ALS (Walsh et al., 2015).

Genetic mutations occur in both coding and non-coding regions of genes whose products do not always contribute to the gene expression process. Several pathogenic mechanisms causing RNA toxicity exist, these include the formation of RNA foci and sequestration of protein factors, altered production of coding and non-coding transcripts or the accumulation of RNA/protein aggregates (Belzil et al., 2013).

1.1.3.1 RNA toxic gain of function: the protein sequestration hypothesis

The expression of RNA from a mutated allele can be pathogenic and is observed in microsatellite expansion disorders. Microsatellite expansion disorders are caused by the expansion of a short, repeated trinucleotide sequence in transcribed DNA and are commonly associated with neurodegeneration (Ashley et al., 1995; Paulson et al., 1996; Renoux et al., 2012).

Table 1.1 Summary of diseases associated with RNA and RNPs

Table of diseases with mutations in gene products involved in RNA processing or RNP formation and appropriate literature reference.

Disease	Gene/mutation	Function	Reference
ALS	TAR DNA Binding protein (TARDBP)	Splicing, transcription	(Sreedharan et al., 2008; Kabashi et al., 2008)
	Fused in sarcoma (FUS)	Splicing, transcription	(Kwiatkowski et al., 2009; Vance et al., 2009)
	C9ORF72	Unknown, non-coding RNA	(DeJesus-Hernandez et al., 2011; Renton et al., 2011)
Autism	7q22-q33 locus breakpoint	Non-coding RNA	(Muhle et al., 2004)
Cancer	Serine/arginine-rich splicing factor 1 (SFRS1)	Splicing, translation, export	(Karni et al., 2007)
	miRNA-372, miRNA-373	RNA interference	(Voorhoeve et al., 2006)
Charcot-Marie-Tooth disease	Glycl-tRNA synthetase (GRS)	Translation	(Antonellis et al., 2003)
	Tyrosyl-tRNA synthetase (YRS)	Translation	(Jordanova et al., 2006)
Fragile X syndrome	Fragile X mental retardation 1 (FMR1)	Translation/mRNA localisation	(Hinds et al., 1993)
Hereditary spastic paraplegia (HSP)	Spastic paraplegia 7 (SPG7)	Ribosome biogenesis	(Casari et al., 1998)
Myotonic dystrophy type 1 (DM1)	DMPK	Protein kinase	(Brook et al., 1992)
Myotonic dystrophy type 2 (DM2)	CCHC-type zinc finger, nucleic acid binding protein (CNBP) or zinc finger binding protein 9 (ZNF9)	Unclear, RNA/DNA binding	(Liquori et al., 2001)
Huntington disease-like 2 (HDL2)	Junctophilin-3 (JPH3)	Ion channel function	(Holmes et al., 2001)
Spinal Muscular Atrophy (SMA)	Survival of motor neuron 2 (SMN2)	Splicing	(Lefebvre et al., 1995)
Spinocerebellar ataxia type 8 (SCA8)	Ataxin 8/Ataxin 8 opposite strand (ATXN8/ATXN8OS)	Unknown, non-coding RNA	(Koob et al., 1999)

Myotonic dystrophy type 1 (DM1) is an example of a non-coding repeat expansion disorder. Myotonic dystrophy is an autosomal dominant neuromuscular disease, characterised by specifically distributed and progressive muscle weakness and wasting. DM1 is the most common cause of adult muscular dystrophy and is caused by a CTG trinucleotide repeat in the 3' UTR region of a gene encoding myotonic dystrophy protein kinase (DMPK) (Brook et al., 1992). Mutant DMPK mRNA contains

CUG repeats and is retained in nuclear foci, cellular hallmarks of DM1. CUG repeats form hairpin-like structures, which sequester several splicing factors, including muscle blind-like 1 (MBL1) and CUG-binding protein 1 (CUG-BP1) (Timchenko et al., 1996; Miller et al., 2000). Sequestration of these splicing factors causes specific dysregulation of alternative splicing of other pre-mRNA, such as bridging integrator-1 (BIN1), causing muscle weakness associated with DM1 (Fugier et al., 2011).

Nuclear foci are molecular hallmarks of neurodegenerative disease and are formed by the aggregation of expanded repeat containing RNA and RBP. The effects of this include the inhibition of transcription, loss of function of protein products and toxicity (from both the mutant transcript and encoded protein) dysregulating cell processes and causing disease (Wojciechowska et al., 2011). It has been suggested that RNA foci cause disease via the sequestration of RBP. The sequestered RBP are then absent causing errors and the dysregulation of other RNA transcripts, altering mRNA splicing and protein isoform frequency; this is known as the 'sequestration hypothesis'.

1.1.3.2 RNA loss of function

SMA is an autosomal recessive, neuromuscular disorder, characterised by the loss of lower motor neurons in the spinal cord, resulting in progressive paralysis with muscle atrophy. SMA is caused by mutations in the survival of motor neuron 1 (*SMN1*) gene (Lefebvre et al., 1995). The severity of SMA is determined by a nearly identical copy of the *SMN1*, the survival of motor neuron 2 (*SMN2*) gene that makes extra copies of SMN protein. The SMN protein is involved in the biogenesis of snRNP complexes, which make up the spliceosome (Neuenkirchen et al. 2008). *SMN1* and *SMN2* differ functionally due to a single nucleotide mutation in exon 7 of *SMN2* which results in frequent exon 7 skipping, resulting in a truncated and unstable protein (Wirth et al., 2006). *SMN1* produces the most functional SMN

protein, with only 10 % of SMN protein produced by *SMN2* being full length and functional (Wirth et al., 2006). An increased number of copies of the *SMN2* gene are associated with less severe symptoms. Mutations in *SMN1* result in lower levels of SMN protein. Lower levels of SMN cause the death of motor neurons and a reduction in nerve impulses between the brain and muscles, resulting in the symptomatic weakness and impaired movement experienced by patients with SMA. The low levels of SMN protein alter the stoichiometry of snRNA, resulting in splicing defects over a wide range of gene transcripts (Zhenxi Zhang et al., 2008) and mutations in RBP. Mutations in vesicle associated membrane protein associated protein B (*VAPB*) have been associated with an adult onset form of SMA (Nishimura et al., 2004) but it is unclear how *VAPB* mutations cause SMA.

The implication of aberrant RNA and mutations in RNA-binding proteins is recurrent when investigating neurodegeneration (Belzil et al., 2013; Walsh et al., 2015), showing the significance of RNA metabolism in normal development. Since the discovery of RNA-binding protein TDP-43 which is mislocalised as a pathological feature in most cases of ALS and a repeat expansion in *C9ORF72* being identified as the most common known genetic cause of ALS (DeJesus-Hernandez et al., 2011; Renton et al., 2011) there is a growing focus on the role of dysregulated RNA metabolism in the pathogenesis of ALS.

1.2 Amyotrophic lateral sclerosis (ALS)

1.2.1 Clinical features

ALS is a disorder defined by degeneration of both the upper and lower motor neurons (UMN and LMN). Slower progressing variants of the disease exist affecting only UMN and LMN (PLS and PMA respectively) (Kiernan et al., 2011). There is selective sparing of the motor neurons controlling eye movement and the pelvic floor

muscles, so a person suffering from ALS remains able to move their eyes and remains continent (Mannen et al., 1977; Brockington et al., 2013).

ALS primarily affects the motor system, but is described as a multisystem neurodegenerative disease, as areas other than the motor area can be affected. It is estimated that 50 % of ALS patients also have frontotemporal-type cognitive deficits (Massman et al., 1996; Ringholz et al., 2005). Frontotemporal lobar degeneration (FTLD) is the atrophy of the frontal and/or temporal lobes that results in abnormalities in behaviour, language and personality. Approximately 15 % of patients with FTLD go on to develop ALS (Burrell et al., 2016; Mann et al., 2017).

1.2.2 Epidemiology and genetics

The lifetime risk of developing ALS is 1 in 400. Men are more commonly affected than women, with a male: female ratio of approximately 2:1. There are 6 in 100,000 people suffering from ALS in most countries worldwide, making the prevalence of ALS globally uniform. However, geographic clusters of ALS have been observed on the island of Guam, the Kii peninsula of Japan and other areas in the West Pacific. These clusters have highlighted the environmental factors acting alongside genetics, with the implication that the neurotoxin, β-methylalimo-L-alanine (BMAA) from cycad seeds was involved in ALS after bioaccumulation in the food chain (Ince et al., 2005).

Sporadic ALS (sALS) accounts for 90-95 % of ALS cases. There is no known genetic model for sALS. Familial ALS (fALS) accounts for the remaining 5-10 % of all ALS cases and is generally considered to classify those with a first or second degree relative with ALS (Byrne et al., 2011). There is a Mendelian pattern of inheritance, which is usually autosomal dominant, though autosomal recessive and X-linked cases have also been observed. fALS is genetically heterogeneous and there have

been 19 genes and loci of major effect that have been recorded. These are shown in in Table 1.2).

The most common mutations observed in approximately 70 % fALS are mutations in *SOD1* (20 %), *TARDBP* (4 %), *FUS* (4 %) and *C9ORF72* (40-50 %). Mutations in *SOD1*, *TARDBP* and *FUS* have also been observed at low levels in cases of sALS (M. D. Alexander et al., 2002; Kirby et al., 2010; Chiò et al., 2011). Mutations in *C9ORF72* have been reported at much higher levels, accounting for approximately 10 % of apparently sporadic cases of ALS (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Al-Chalabi et al., 2012). These genes and their mutations in ALS will be discussed in further detail below.

1.2.2.1 SOD1

In 1993 the first ALS gene was described (Rosen et al., 1993). Mutations in copper/zinc binding superoxide dismutase (*SOD1*) are found in 20 % of fALS cases and 5 % of sALS cases. More than 150 different mutations of *SOD1* have been reported ("http://alsod.iop.kcl.ac.uk"). The SOD1 protein is made up from 154 amino acids; this means that mutations have been observed in almost all amino acid positions in the primary sequence, suggesting that unfolded, toxic, oligomeric conformations of the mutant proteins result from these mutations. The majority of mutations are missense, although there are some insertions and deletions. *SOD1* is located on chromosome 21q22.1 (Siddique et al., 1991). SOD1 is a soluble, cytoplasmic enzyme, which binds copper and zinc to destroy naturally occurring, harmful superoxide radicals. It converts superoxide radicals into molecular oxygen and hydrogen peroxide.

Table 1.2 Genetics of ALS Gene locus, chromosomal locus and genes implicated in ALS. The inheritance and implicated disease mechanisms along with the appropriate literature references are shown. (AD = autosomal dominant, AR = autosomal recessive, XD = X-linked dominant)

Gene locus	Chromosomal locus	Gene	Inheritance	Implicated disease mechanisms	References
ALS1	21q22.11	Superoxide dismutase 1 (*SOD1*)	AD/AR	Oxidative stress	(Siddique et al., 1991; Rosen et al., 1993)
ALS2	2q33	Alsin (*ALS2*)	AR	Endosomal trafficking	(Hadano et al., 2001; Y. Yang et al., 2001)
ALS3	18q21	Unknown	AD	Unknown	(Hand et al., 2002)
ALS4	9q34	Senataxin (*SETX*)	AD	RNA metabolism	(Chen et al., 2004)
ALS5	15q15-q22	Unknown	AR	DNA damage repair, axon growth	(Hentati et al., 1994)
ALS6	16p11.2	Fused in sarcoma/translated in liposarcoma (*FUS/TLS*)	AD/AR	RNA metabolism	(Kwiatkowski et al., 2009; Vance et al., 2009)
ALS7	20ptel-p13	Unknown	AD/AR	Unknown	(Parkinson et al., 2006)
ALS8	20q13.3	Vesicle associated membrane protein (*VAMP*) – associated protein B (*VAPB*)	AD	ER stress	(Nishimura et al., 2004)
ALS9	14q11.2	Angiogenin (*ANG*)	AD	RNA metabolism	(Greenway et al., 2006)
ALS10	1p36.2	TAR DNA binding protein (*TARDBP*)	AD	RNA metabolism	(Sreedharan et al., 2008; Kabashi et al., 2008)
ALS11	6q21	FIG4 (*FIG4*)	AD	Endosomal trafficking	(Chow et al., 2009)
ALS12	10p15-p14	Optineurin (*OPTN*)	AD/AR	Autophagy	(Maruyama et al., 2010)
ALS13	12q23-q24.1	Ataxin 2 (*ATXN2*)	AD	RNA metabolism	(Elden et al., 2010)
ALS14	9p13	Valosin-containing protein (*VCP*)	AD	Autophagy	(Johnson et al., 2010)
ALS15	Xp11.21	Ubiquilin 2 (*UBQLN2*)	XD	UPS, autophagy	(Deng et al., 2011)
ALS16	9p13	Sigma non-opioid intracellular receptor 1 (*SIGMAR1*)	AD	UPS, autophagy	(Luty et al., 2010; Al-Saif et al., 2011)
ALS17	3p12.1	Chromatin modifying protein 2B (*CHMP2B*)	AD	Endosomal trafficking	(Parkinson et al., 2006)
ALS18	17p13.3	Profilin 1 (*PFN1*)	AD	Cytoskeleton	(C.-H. Wu et al., 2012)
ALS19	2q33.3-q34	V-erb-b2 avian erythroblastic leukaemia viral oncogene homolog 4 (*ERBB4*)	AD	Neuronal development	(Takahashi et al., 2013)
ALS20	12q13.1	Heterogeneous nuclear ribonucleoprotein A1 (*HNRNPA1)*	AD	RNA metabolism	(Hong Joo Kim et al., 2013)
ALS21	5q31.2	Matrin 3 (*MATR3*)	AD	RNA metabolism	(Johnson et al., 2014)

ALS22	2q35	Tubulin alpha-4A chain (*TUBA4A*)	AD	Cytoskeleton	(Smith et al., 2014)
ALS-FTD1	9q21-q22	Chromosome 9 open reading frame 72 (*C9ORF72*)	AD	RNA metabolism, autophagy	(DeJesus-Hernandez et al., 2011; Renton et al., 2011)
ALS-FTD2	9p13.2-p21.3	Coiled-coil-helix-coiled-coil-helix domain containing 10 (*CHCHD10*)	AD	Mitochondrial maintenance	(Bannwarth et al., 2014)
ALS-FTD3	5q35	Sequestosome 1 (*SQSTM1*)	AD	Autophagy	(Fecto et al., 2011)
ALS-FTD4	12q14.2	TANK binding kinase 1 (*TBK1*)	AD	Autophagy, neuroinflammation	(Cirulli et al., 2015; Freischmidt et al., 2015)

Mice deficient in *SOD1* do not develop motor degeneration, suggesting that it is not loss of function, but a toxic gain of function that contributes to ALS (Reaume et al., 1996). The exact mechanism by which mutant *SOD1* causes motor neuron damage has not yet been recognized, however, it is implicated in several different mechanisms including oxidative stress, mitochondrial dysfunction, excitotoxicity, protein aggregation and neuroinflammation (Shaw, 2005).

1.2.2.2 TARDBP

TAR DNA binding protein (*TARDBP*) encodes TAR DNA binding protein of 43 kDa (TDP-43). TDP-43 is an RNA and DNA binding protein and functions as a RNP in the regulation of transcription and alternative pre-mRNA splicing. In healthy neurons it is located predominantly in the nucleus, but shuttles continuously between the nucleus and the cytoplasm in a transcription-dependent manner (Ayala et al. 2008).

There are 39 known mutations in *TARDBP* associated with ALS ("http://alsod.iop.kcl.ac.uk"). The majority are missense mutations in exon 6, which encodes the Gly rich and C-terminal of TDP-43, involved in protein interactions and nuclear translocation. Mutations in *TARDBP* have been found in patients with fALS (3-4 %) and sALS (1 %) (Sreedharan et al., 2008; Kabashi et al., 2008; Kirby et al., 2010).

1.2.2.3 FUS

Fused in sarcoma or translocation in liposarcoma (*FUS/TLS*) encodes another RBP called FUS and mutations in *FUS/TLS* have been found in both patients with fALS and sALS (Kwiatkowski et al., 2009; Vance et al., 2009; Corrado et al., 2010). FUS is both an RNA and DNA binding protein found predominantly in nucleus; but shuttles between the nucleus and cytoplasm (Zinszner et al., 1997).

FUS is a member of the TET family of proteins alongside Ewing's sarcoma (EWS) and the TATA-binding protein-associated factor 15 (TAF15). This family of proteins all bind nucleic acids and have a function in RNAPII-mediated transcription (Uranishi et al., 2001), pre-mRNA splicing, repression of RNA polymerase III transcription, RNA transport in neurons (A. Y. Tan et al., 2009) and a role in the DNA repair pathway (Baechtold et al., 1999). FUS is involved in transcription, RNA splicing (L. Yang et al., 1998) and mRNA nucleocytoplasmic shuttling (Zinszner et al., 1997).

FUS contains a RRM, N-terminal Ser, Gly, Gln and Tyr-rich region and a C-terminal RGG containing region (Morohoshi et al., 1998). Mutations in FUS occur in 5 % fALS (Kwiatkowski et al., 2009) and in less than 1 % of sALS. The majority of mutations occur in the conserved C-terminus and nuclear localisation signal (NLS). Mutations in the NLS disrupt transportin-dependent nuclear import, which results in the cytoplasmic redistribution of FUS and its incorporation into stress granules (Dormann et al., 2010). The majority of mutations in ALS are missense, but insertion/deletion of 2 glycines in the glycine-rich region (of 10 glycines) have also been observed. Deletions in the C-terminus prevent interaction with SR splicing factors and may affect their role in RNA splicing (L. Yang et al., 1998). Progressive neurodegeneration in an age- and dose-dependent manner is observed in mice overexpressing human wild-type FUS (Mitchell et al., 2013).

1.2.2.4 Other RNA binding proteins

Mutations have also been found in genes encoding other RBP such as angiogenin (ANG) and the RNA/DNA helicase senataxin (SETX). ANG is a member of the pancreatic ribonuclease A superfamily and induces neovascularisation and the growth of tumours, evident from its elevated expression in several cancers. ANG is expressed in the nucleus and cytoplasm of motor neurons and is involved in neurite extension and survival of motor neurons. ANG acts as a neuroprotective factor towards motor neurons that are under oxidative stress (Subramanian et al., 2008).

Loss of function mutations in *ANG* have been found in fALS and sALS (Greenway et al., 2006; D. Wu et al., 2007). This is supported with some variants of ALS mutant ANG causing reduction or loss of ribonucleolytic activity and some variants directly affecting the catalytic centre (Crabtree et al., 2007). Mutations in *ANG* prevent its neuroprotective function and are toxic to motor neurons (Subramanian et al., 2008). The exact mechanism by which *ANG* mutations induce motor neuron disease is not fully understood.

Vascular endothelial growth factor (*VEGF*) encodes another angiogenic factor associated with ALS, which performs a similar function in the development of neurons (Lambrechts et al., 2003). VEGF is neuroprotective to motor neurons under hypoxia by stimulating angiogenesis after the binding of hypoxia response factors to a response element in VEGF. Deletion of the hypoxia response element in the *VEGF* promoter of mice, reduces hypoxic expression and causes progressive motor neuron degeneration, reminiscent of ALS (Oosthuyse et al., 2001).

SETX encodes a protein of unknown function, however, it contains a DNA/RNA helicase domain, which has strong homology to two genes, which encode proteins involved in RNA processing. This suggests that mutations in *SETX* may cause neurodegeneration by the dysfunction of helicases or RNA processing errors (Chen et al., 2004). Mutations in *SETX* are usually associated with an autosomal dominant juvenile form of ALS (ALS4). Autosomal recessive mutations in *SETX* cause an unrelated disorder, ataxia oculomotor apraxia type 2 (AOA2) (Chen et al., 2006). The mechanisms by which ALS4 or AOA2 mutations cause disease have not yet been elucidated.

hnRNPA1 is a ribonucleoprotein, which can bind to RNA. Mutations in hnRNPA1 occur in prion-like domains and hnRNPA1 along with other RBP with prion-like domains, such as FUS and TDP-43, have been shown to accumulate in disease pathology (Hong Joo Kim et al., 2013). Similarly, RNA binding proteins MATR3 and Atatxin-2 has also been shown to interact in complexes with TDP-43 (Elden et al., 2010; Johnson et al., 2014). Muatations in genes encoding RNA binding proteins and their accumulation in disease pathology heavily implicates RNA metabolism in disease pathogenesis.

1.2.2.5 C9ORF72

In 2011, a GGGGCC hexanucleotide repeat expansion in *C9ORF72* was identified that is strongly associated with ALS and FTD (DeJesus-Hernandez et al., 2011; Renton et al., 2011). It has since been found to be the most common known genetic cause of ALS causing 40-50 % of familial (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Cooper-Knock, Hewitt, et al., 2012) and 7-10 % of sporadic cases (Majounie et al., 2012).

C9ORF72 encodes a protein of the DENN family and has recently been implicated in the initiation of autophagy (Webster et al., 2016), a process known to be altered in ALS and more generally in neurodegeneration. The DENN family of proteins are also regulators of membrane trafficking events (Dapeng Zhang et al., 2012). The hexanucleotide repeat is found in the intron of *C9ORF72* so it is not present in the amino acid sequence of the protein, but in the pre-mRNA molecules. The mechanism by which the *C9ORF72* expansion causes ALS has not yet been elucidated.

1.2.3 Neuropathological features

The accumulation of proteinaceous aggregates in the cytoplasm or nucleus of the central nervous system (CNS) is a hallmark of many neurodegenerative diseases including HD, AD and Parkinson's disease (PD) (Forman et al., 2004).

In cells, ubiquitin covalently attaches to proteins, marking them for destruction by the proteasome. Ubiquitinated inclusions (UBIs) are most commonly observed inclusions in ALS, appearing in both familial and sporadic cases (Leigh et al. 1991). The appearance of these inclusions takes either one or a combination of two forms; rounded and compact (Lewy-body like) or filamentous (skein-like).

In 2006, TDP-43 was found to be the major component of cytoplasmic and ubiquitinated inclusions associated with ALS and frontotemporal lobar dementia with ubiquitinated inclusions (FTLD-U) (Neumann et al., 2006; Arai et al., 2006). TDP-43 inclusions have been identified in 97 % of sporadic and familial ALS cases (Mackenzie et al., 2007; Maekawa et al., 2009). TDP-43 inclusions are not present in cases with *SOD1* or *FUS* mutations (Mackenzie et al., 2007; C.-F. Tan et al., 2007; Vance et al., 2009). The mislocalisation of TDP-43 and the identification of mutations in *TARDBP*, provide a link between sporadic and inherited ALS, as cytoplasmic aggregates of TDP-43 are a prominent feature of sporadic ALS. *TARDBP* mutations have also been observed in patients with FTLD, both with and without motor neuron disease (Gitcho et al., 2009; Benajiba et al., 2009; Borroni et al., 2009; Ince et al., 2011). The histopathological changes that occur to TDP-43 in neurodegenerative disease are referred to as 'TDP-43 proteinopathies' (Kwong et al., 2007).

In ALS and FTD, TDP-43 is cleared from the nucleus and is mislocalised being deposited and aggregated in the cytoplasm (Neumann et al., 2006). Pathological TDP-43 is hyperphosphorylated (p-TDP-43), ubiquitinated and cleaved to generate C-terminal fragments; these are all major components of intranuclear and cytoplasmic inclusions (Neumann et al., 2006). The cleavage of C-terminal fragments results in the mislocalisation to the cytoplasm and formation of toxic aggregates (Nonaka, Arai, et al., 2009; Yong-Jie Zhang et al., 2009; Nonaka,

Kametani, et al., 2009). It is unclear which features of TDP-43 proteinopathy are responsible for disease: is it the loss of normal function of TDP-43 with clearance from the nucleus or a toxic gain of function due to aggregation and transformation of TDP-43 in the cytoplasm?

Neuronal cytoplasmic inclusions (NCIs) and neuronal intranuclear inclusions (NIIs) are hallmarks of ALS and FTLD. Typically in ALS these are characterised by ubiquitin, p62 and p-TDP-43 positive staining (A. King et al., 2011; Al-Sarraj et al., 2011; Al-Chalabi et al., 2012). TDP-43 positive inclusions are found in motor neurons of the motor cortex; brain stem and spinal cord in cases of sALS and *TARDBP*/*C9ORF72* related ALS, but not in *SOD1* or *FUS* subtypes of ALS.

Aggregates of FUS are observed in the cytoplasm of patients with a *FUS* mutation (Kwiatkowski et al., 2009; Vance et al., 2009). In patients with a *FUS* mutation, TDP-43-positive inclusions are not observed. This implies that the neurodegenerative mechanisms set in motion by *FUS* mutation are independent of TDP-43 (Vance et al., 2009). FUS is also the major component of nuclear protein aggregates in other neurodegenerative diseases such as FTLD (Dormann et al., 2010), however, *FUS* mutations do not cause FTLD.

Bunina bodies (BBs) are not ubiquitinated, but also serve as a hallmark for ALS. BBs are small eosinophilic inclusions, first observed in patients with familial ALS (Bunina, 1962). Since their discovery, cystatin C (Okamoto et al., 1993) and transferrin (Mizuno, Amari, Takatama, Aizawa, Mihara, and Okamoto, 2006a) have been identified as present in BBs, but TDP-43 is not present (C.-F. Tan et al., 2007). BBs appear both in normal looking and degenerating motor neurons, suggesting a role in the initial stages of motor neuron degeneration (Okamoto et al., 2008). The role of BBs in ALS has not yet been elucidated.

Hyaline conglomerate inclusions (HCIs) are large, phosphorylated and non-phosphorylated neurofilament accumulations observed in spinal cord motor neurons (Hirano et al., 1967; Wood et al., 2003). HCIs have been identified in other diseases involving the lower motor neurons such as multiple system atrophy and also in healthy controls, suggesting a less specific role for HCIs in ALS (Sobue et al., 1990).

C9ORF72 ALS cases exhibit additional pathologies due to the unconventional repeat associated translation of the C9ORF72 transcript. Firstly, nuclear RNA foci are observed in patient CNS tissue in the frontal cortex, hippocampus and cerebellum (DeJesus-Hernandez et al., 2011; Gendron et al., 2013). These foci are formed from the $(G_4C_2)_n$ repeat expansion being transcribed in both the sense and antisense directions. These foci are also observed in astrocytes, microglia and oligodendrocytes (Mizielinska et al., 2013) and fibroblast and induced pluripotent stem cell (iPSC) derived motor neurons from C9ORF72/FTD patients (Lagier-Tourenne et al., 2013; Sareen et al., 2013). The majority of foci are nuclear in localisation, but rare cytoplasmic foci have also been observed (Mizielinska et al., 2013; Cooper-Knock et al., 2014). These foci colocalise with various RNA binding proteins including SRSF2, AlyRef, heterogeneous ribonucleoprotein (hnRNP) A1 and A3, SC35 and Pur-α (Mori, Lammich, et al., 2013; Sareen et al., 2013; Y.-B. Lee et al., 2013; Haeusler et al., 2014; Cooper-Knock et al., 2014). In motor neurons from patients with C9ORF72-ALS, there is a correlation between TDP-43 nuclear loss and antisense RNA foci, which is not observed with sense foci (Cooper-Knock et al., 2015). Sense and antisense $(G_4C_2)_n$ repeat transcripts also undergo unconventional repeat associated non-ATG (RAN) translation to form dipeptide repeat proteins (Ash et al., 2013; Mori, Weng, et al., 2013). DPR proteins are a major component of p62 positive, TDP-43 negative aggregates in C9ORF72 ALS/FTD patients (Mann et al., 2013).

1.2.4 Pathogenesis

The underlying cause and selectivity for motor neuron death in ALS is unknown. The disease has no global or geographical pattern and no definite link to environmental toxins. There are complex interactions between genetic and molecular pathways implicated in ALS, suggesting that the pathophysiological mechanism underlying the development of ALS is multifactorial. There is also complex communication between neurons and surrounding cells such as microglia and astrocytes.

1.2.4.1 Oxidative damage

Cellular reactive oxygen species (ROS) are the by-product of aerobic respiration. The major source of oxygen consumption in cells occurs in mitochondria for ATP production. Mitochondria leak electrons, resulting in the incomplete reduction of molecular oxygen to give superoxide radicals and hydrogen peroxide. Superoxide radicals and hydrogen peroxide undergo further reactions to produce highly reactive peroxynitrite and hydroxyl radicals, which are damaging to the cell. Oxidative stress occurs due to an imbalance between the production of ROS and the ability to repair or remove the damage caused by them. Oxidative stress is capable of causing damage to proteins, lipids DNA and RNA, altering the protein conformation, cell membrane dynamics and causing alterations in DNA and RNA species (Barber et al., 2010; Kong et al., 2010).

Oxidative stress has been implicated in ALS, due to the identification of a mutation in *SOD1*, a free radical scavenger. *SOD1* knockout mice models do not develop motor neuron degeneration, so the *SOD1* mutation does not cause a loss of function, but a toxic gain of function, generating free radicals that eventually lead to cell injury and cell death (Joyce et al., 2011). Immunohistochemistry on the end products of lipid peroxidation and protein oxidation in spinal cord tissue of patients with sALS and fALS have confirmed enhanced levels, suggesting that spinal cord motor neuron and glial cells are exposed to oxidative stress in ALS (Shibata et al., 2001).

ROS production increases with age, primarily due to an increase in the leakage of electrons from mitochondria. This could partially explain the onset of ALS later in life, with the median age of onset being 47 years (Kiernan et al., 2011). This suggests the possibility of a 'threshold effect' that the mutations found in ALS do not cause the disease until a later age, due to the fact that later in life mitochondria become dysfunctional.

Oxidative stress can also damage other cellular processes such as RNA molecules and therefore RNA metabolism. Under the same oxidative conditions, RNA is more susceptible to damage than DNA (Li et al., 2006). RNA is more susceptible to damage due to its single stranded nature and the fact that no oxidation repair systems have been found for RNA (Li et al., 2006). The oxidation of mRNA has been identified in patients and transgenic mice expressing fALS SOD1 mutations (Chang et al., 2008). Oxidation of mRNA occurs at the presymptomatic stages, so is not a consequence of motor neurons dying. The most highly oxidised RNA species were those encoding proteins involved in protein biosynthesis, mitochondrial electron transport and ALS associated Dynactin 1, VAMP and neurofilament subunits. Where Oxidised mRNAs are translated to form full length, non-functional proteins or full length, functional proteins with reduced activity (M. Tanaka et al., 2007). There is a direct correlation between the extent of mRNA oxidation and the frequency of errors in translation. Translation errors can result in the synthesis of defective proteins.

1.2.4.2 Protein misfolding and aggregation

Misfolding of mutant proteins and protein aggregation is a commonly occurring theme in neurodegenerative diseases (Forman et al., 2004). There is continued debate over the significance of intracellular aggregates and the role they play in disease pathogenesis. It is unclear whether aggregates directly cause toxicity or whether they are harmless by-products of neurodegeneration. The possibility that

aggregations are beneficial to the cell has also been considered, with the sequestration of abnormal proteins that could potentially harm the cell.

Recently, it has been shown that cytoplasmic aggregates disrupt the nucleocytoplasmic transport of RNA and proteins (Woerner et al., 2016). The inclusions observed in ALS are ubiquitinated, suggesting impairment of the ubiquitin proteasome system (UPS) and autophagy degradation systems. Furthermore, there are ALS genes involved in endosomal trafficking (*ALS2*, *VAPB*, *CHMP2B*, *FIG4*), the UPS (*UBQLN2*, *SQSTM1*, *SIGMAR1*) and autophagy (*OPTN*, *VCP*, *TBK1, C9ORF72*), which have been identified (See Table 1.2).

1.2.4.3 Neuronal excitotoxicity

Excitotoxicity is the process by which neurons are damaged and killed by the excessive stimulation of glutamate receptors. Glutamate is a neurotransmitter in the vertebrate nervous system, which is stored in vesicles and released at synapses. Excess activation and binding to the postsynaptic receptor by glutamate causes activation of calcium dependent enzymatic pathways and excessive release of free radicals, which can damage intracellular organelles. Glutamate receptor expression is controlled by astrocytes, non-neuronal glial cells in the brain and spinal cord. Astrocytes are therefore possibly also implicated in the modulation of excitotoxicity. There is evidence of altered glutamate levels in the tissues of ALS patients and impaired glutamate metabolism in motor neuronal models of fALS (D'Alessandro et al., 2011).

Riluzole is the only disease modifying therapy for patients with ALS, extending survival by 3-6 months (Kiernan et al., 2011). It inhibits glutamate release possibly by blocking of sodium channels. It is approved in most countries, but not all due to its high cost for only a 'modest' effect. Riluzole has been proven to have a neuroprotective action upon motor neurons and glia when applied early after an

excitotoxic stimulus (Cifra et al., 2011), further supporting the role of excitotoxicity in the pathogenesis of ALS.

1.2.4.4 Neuroinflammation and non-cell autonomous toxicity

Neuroinflammation is characterised by microgliosis, the activation and proliferation of microglia, and the accumulation of T-lymphocytes. As well as occurring as a result of neuronal injury, the inflammatory pathway can also affect the survival of motor neurons, either having neuroprotective or deleterious effects (Lewis et al., 2012; Evans et al., 2013). Astrocytes and microglia may contribute to neurodegeneration via insufficient release of neurotrophic factors, responsible for the growth, survival and maintenance of neurons. Microglia also release neurotoxic mediators as part of their normal function, but the levels of these decrease once a task is complete.

Astrocytes derived from SOD1, C9ORF72 and sporadic ALS fibroblasts have been shown to be toxic to neurons in co-culture (Meyer et al., 2014). Non-cell autonomous toxicity has also been described in SOD1 ALS mouse models (Nagai et al., 2007). During inflammation microglia remain active for extended periods and large amounts of mediators such as cytokines, chemokines, proteases and amyloid precursor protein are produced. Increased inflammatory mediators and activated astrocytes and microglia have been observed in the spinal cord of patients with ALS (Henkel et al., 2004) and SOD1 mouse models (Clement et al., 2003).

Inflammation also generates oxidative stress with activated microglia providing a source of oxygen free radicals. This links neuroinflammation to the oxidative damage mechanism of neurodegeneration, and further highlights the overlap between the pathogenic mechanisms underlying ALS.

1.2.4.5 Axonal transport defects

Due to the large size and long axonal compartment of motor neurons, they have a high-energy demand and depend on a robust cytoskeleton for axonal transport.

Pathological hallmarks of ALS include the accumulation and abnormal assembly of neurofilaments, identified by the immunoreactivity of neurofilament epitopes within ubiquitinated inclusions of surviving motor neurons (De Vos et al., 2008). Neurofilaments are the major component of the neuronal cytoskeleton and their main functions are structural support of the axon and axonal transport. Neurofilaments transport proteins and neurotransmitters throughout the neuron via two types of transport proteins: dynein and kinesin.

Dynein is involved in retrograde axonal transport (towards cell body) and requires a dynactin complex to transport vesicles along microtubules. Mutations in the P150 subunit of dynactin (*DCTN1*) have been associated with motor neuron disease (Puls et al., 2003). This mutation distorts the folding of the microtubule-binding domain in dynactin. Kinesin is involved in anterograde axonal transport (away from the cell body). Mutations in the gene encoding the kinesin-1 protein are associated with HSP, a neurodegenerative disorder characterised by progressive weakness and spasticity of the legs (E. Reid et al., 2002).

VAPB is also mutated in a rare form of ALS and adult onset SMA (Nishimura et al., 2004). VAPB is associated with vesicle trafficking within neurons and is thought to have a role in membrane transport particularly in the endoplasmic reticulum (ER) (Skehel et al., 2000). Membrane bound vacuoles in axons and dendrites of mutant *SOD1* (*G37R*) transgenic have also been reported (Wong et al., 1995).

Defective mitochondrial transportation along axons has been described as contributing to motor neuron degeneration in ALS models (De Vos et al., 2007; Mórotz et al., 2012). This links defective axonal transport to mitochondrial dysfunction.

1.2.4.6 Mitochondrial dysfunction

Mitochondria are vital for cell survival due to their role in the production of cellular ATP and many other biosynthetic intermediates. They are also responsible for intracellular calcium buffering and contribute to cellular stress responses triggering apoptosis and autophagy. Mitochondrial dysfunction has also been implicated in ALS. Human patients with ALS and mice over expressing the mutant SOD1 and TDP-43 proteins, were both found to have morphological and functional defects in mitochondria (Shi et al., 2010; Magrané et al., 2014).

Due to the high-energy demand of neurons, alteration to the production of ATP causes damage to motor neurons. Functional defects in mitochondria could result in an increase of ROS, linking mitochondrial dysfunction to oxidative stress in the pathogenesis of ALS.

1.2.4.7 Dysregulated RNA metabolism

As discussed previously, dysregulated RNA metabolism has been strongly implicated in ALS due to the number of ALS genes involved in RNA processing. Aggregates of RNA binding protein TDP-43 are a pathological hallmark in most cases of ALS. The repeat expansion in *C9ORF72* has been implicated as a direct cause of RNA toxicity and the production of toxic dipeptide repeat (DPR) proteins. The potential areas where RNA metabolism is altered and where ALS gene mutations and disease processes occur are summarised in Figure 1.1.

1.3 TDP-43 and ALS

1.3.1 TDP-43 structure and localisation

Encoded by *TARDBP*, TAR DNA binding protein of 43 kDa (TDP-43) is a ubiquitously expressed DNA/RNA-binding protein (Ou et al., 1995). TDP-43 contains 2 RNA recognition motifs (RRMs) (Buratti and Baralle, 2001), a bipartite NLS,

nuclear export signal (NES) (Winton et al., 2008), a glycine-rich C-terminal domain (Buratti et al., 2005) and prion-like domain (PLD) (Fuentealba et al., 2010).

In order to maintain protein homeostasis, many RBPs utilise alternative splicing-mediated NMD and highly conserved intronic regions to auto-regulate their own expression levels (Buratti et al., 2011; Y. Zhou et al., 2013). FUS is able to auto-regulate its expression by regulation of the alternative splicing of exon 7 (Y. Zhou et al., 2013). When FUS levels are high, there is an increase in binding to exon 7 and its flanking introns, promoting exon 7 skipping and NMD to reduce excess FUS protein. To increase FUS production when levels of FUS are lower, inclusion of exon 7 would be favoured.

TDP-43 uses a different mechanism to regulate the stability of its own mRNA and auto-regulate TDP-43 protein levels (Ayala et al., 2011; Polymenidou et al., 2011). TDP-43 is able to regulate its own expression by binding to a highly conserved region in the 3' UTR of *TARDBP* pre-mRNA, the TDP-43 binding region (TDPBR) which is critical for auto-regulation (Sephton et al., 2011; Ayala et al., 2011; Polymenidou et al., 2011). The C-terminal is also required for self regulation (Ayala et al., 2011). When TDP-43 is over-expressed, increased binding to the TDPBR promotes mRNA instability, by transcriptional stalling of RNAPII and polyadenylation site switching, which results in the NMD of alternatively spliced and post-transcriptionally modified transcripts (Avendaño-Vázquez et al., 2012; D'Alton et al., 2015).

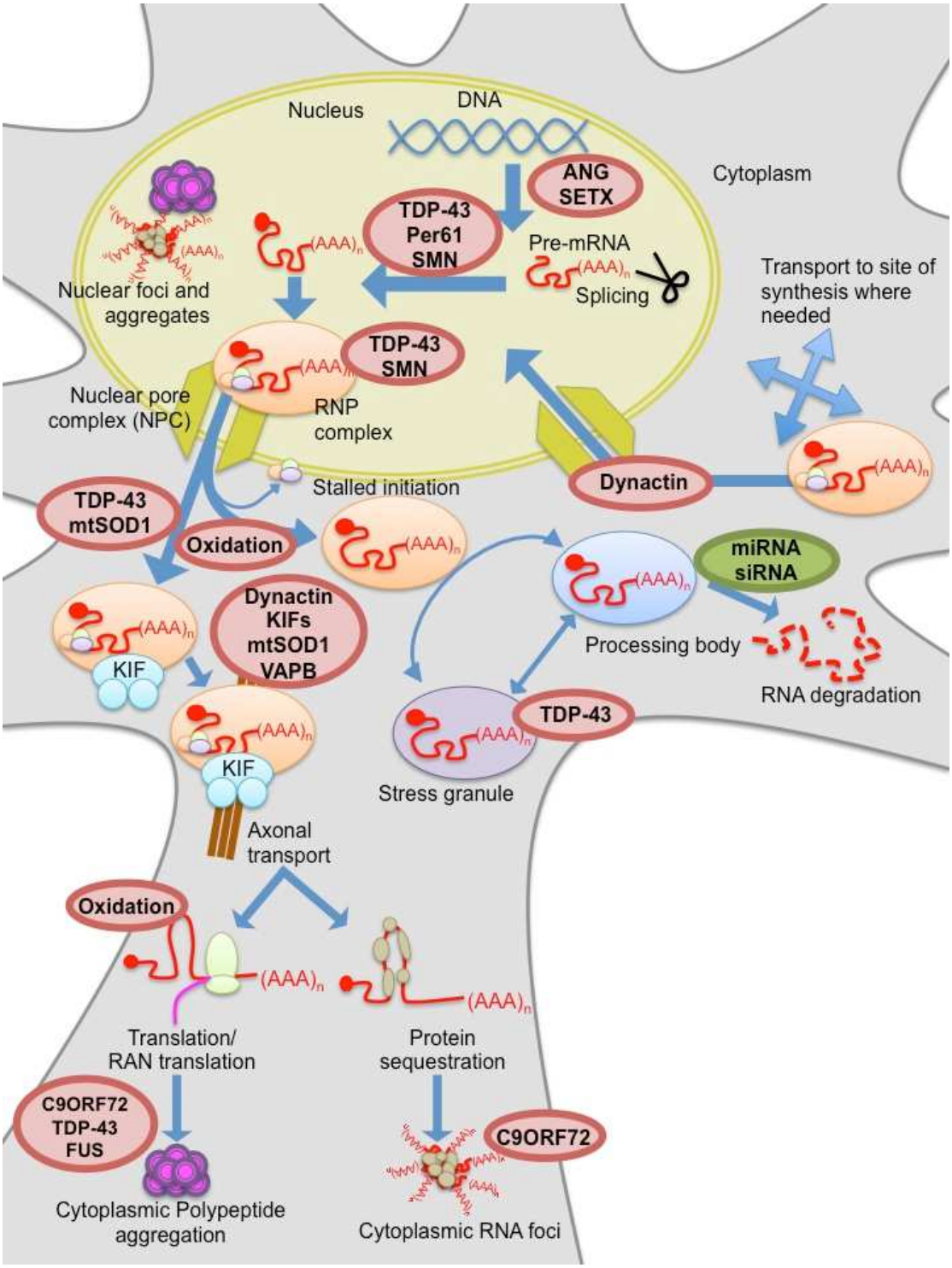


Figure 1.1 RNA metabolism and the potential areas where ALS gene mutations or disease processes can occur

After transcription, pre-mRNA undergoes several modifications (5' capping, addition of 3' polyA tail and splicing). The mature mRNA is then incorporated into a RNP complex for export into the cytoplasm via NPC. Translation then occurs at the ribosome. In the cytoplasm further regulation of gene expression occurs involving siRNA and miRNA (NMD, NSD, NGD and RNAi). mRNAs are either translated or held in stress granules until required. Aberrant mRNAs are degraded in processing bodies (P-bodies). Nuclear and cytoplasmic polypeptide aggregation and RNA foci can be observed. Potential sites of disease with implicated gene alterations have been identified in red.

Many RBPs also have other cellular factors, which cross-regulate their expression, binding to their RNA to modulate splicing and stability. It has been shown that Hu Antigen R (HuR) binds to the 3' UTR of TDP-43 and regulates its expression by translation efficiency, not via RNA stability (Lu et al., 2014).

Auto-regulation of TDP-43 and FUS protein levels to tightly control their function, suggests that unbalancing this regulation may underpin the mechanism of neurodegeneration in ALS. Indeed, animal models expressing TDP-43 which lacked the auto-regulation sequence have been shown to develop a more severe neurodegeneration than those expressing wild-type (WT) or mutant TDP-43 including the auto-regulation sequence (Ling et al., 2013).

The majority of TDP-43 is located in the nucleus, but is capable of nucleocytoplasmic shuttling (I.-F. Wang et al., 2008; Winton et al., 2008; Ayala et al., 2008). Cytoplasmic TDP-43 has been identified associated with RNA granules (Elvira et al., 2006), hnRNPs, mitochondria and stress granules. In healthy neurons it is located predominantly in the nucleus, but shuttles continuously between the nucleus and the cytoplasm during transcription (Ayala et al., 2008). The nucleocytoplasmic transport of TDP-43 is highly dependent upon transcription and any disruption to the interaction of TDP-43 to target mRNA or alterations to transcription, results in the cytoplasmic accumulation of TDP-43 (Ayala et al., 2008).

1.3.2 TDP-43 proteinopathy and disease pathology

TDP-43 proteinopathy is defined by the nuclear loss, cytoplasmic mislocalistion, fragmentation and aggregation of TDP-43. The cellular distribution of TDP-43 protein in post mortem motor neurons in health and disease are shown in Figure 1.2. A key question in ALS research is which part of the proteinopathies described is responsible for the development of disease. Is it a loss of function due to the

mislocalisation of protein preventing normal function of TDP-43 protein, gain of function due to the toxicity of TDP-43 aggregates, or due to a combination of both?

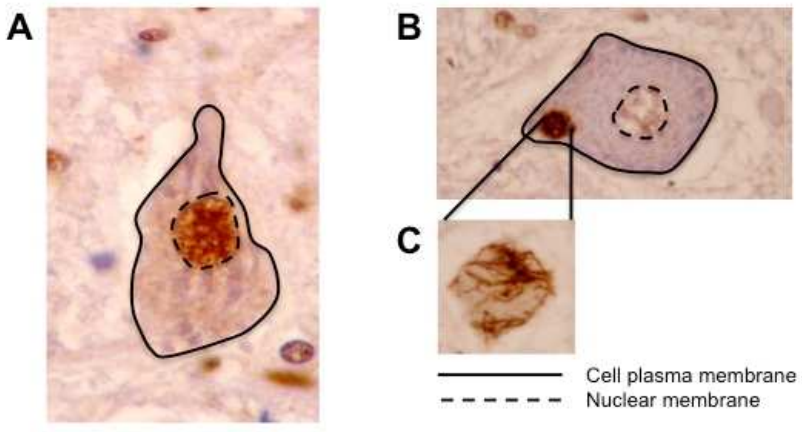


Figure 1.2 TDP-43 proteinopathy
(A) TDP-43 is localised to the nucleus in healthy motor neurons. In disease motor neurons TDP-43 forms aggregates. There are several types of aggregates (i) skein-like, (ii) Lewy body like and (iii) compact. (B) TDP-43 is mislocalised from the nucleus to the cytoplasm (C) Cytoplasmic TDP-43 aggregate in the cytoplasm of a motor neuron. Pictures kindly provided by Highley JR and Ince PG, Neuropathology group in SITraN.

The redistribution of TDP-43 is an early event in ALS with hyperphosphorylation and ubiquitination of TDP-43 in the cytoplasm occurring later (Giordana et al., 2010). Acetylation of TDP-43 has also been identified as promoting the formation of insoluble, hyperphosphorylated inclusions in ALS (Cohen et al., 2015). The spread of transformed TDP-43 protein from cortical and spinal cord motor neurons to other cortical regions can be used to stage ALS pathological progression, highlighting the significance of TDP-43 protein to the pathogenesis of ALS (Brettschneider et al., 2013).

Pathological TDP-43 is ubiquitinated, hyperphosphorylated and N-terminally cleaved to generate toxic, insoluble C-terminal fragments (CTFs) (Neumann et al., 2006; Hasegawa et al., 2008; Yong-Jie Zhang et al., 2009). TDP-43 is cleaved by caspases to form 25 kDa and 35 kDa fragments referred to as TDP-25 and TDP-35 (Yong-Jie Zhang et al., 2007). CTFs mislocalise to the cytoplasm due to the lack a NLS and aggregate due to the presence of the prion-like domain (Winton et al., 2008; Igaz et

al., 2009; Fuentealba et al., 2010). The cleavage of TDP-43 is enhanced by C-terminal mutations (Rutherford et al., 2008; Sreedharan et al., 2008), cellular stress (Dormann et al., 2009; Suzuki et al., 2011) and proteasomal inhibition (C.-C. Huang et al., 2014).

Both CTFs form inclusions that are able to sequester full-length TDP-43 (Nonaka, Arai, et al., 2009; C. Yang et al., 2010; Che et al., 2011). TDP-35 colocalises with stress granules (SGs) (Nishimoto et al., 2010; Liu-Yesucevitz et al., 2014) and when aggregated in inclusions causes alterations to RNA processing (Che et al., 2011). TDP-35 is able to bind to RNA and this is beneficial to the formation of inclusions and responsible for the sequestration of full length TDP-43 by TDP-35 (Che et al., 2015). TDP-25 mediated toxicity occurs by increasing formation of inclusions, impaired proteasomal degradation, ER-stress mediated apoptosis and impaired neurite outgrowth (C. Yang et al., 2010; Suzuki et al., 2011; Caccamo et al., 2015).

The cleavage of TDP-43 is not required prior to aggregation of TDP-43 (Dormann et al., 2009). Furthermore, the nuclear loss of TDP-43 can be induced by fragmented TDP-43 without the formation of aggregates (C. Yang et al., 2010). Caspase cleavage of TDP-43 to TDP-35 and TDP-25 is an intermediate step in the degradation of TDP-43 under normal conditions (C.-C. Huang et al., 2014). This has led to the suggestion of a ‘two-hit’ hypothesis for inclusion formation where a second deleterious event, such as RNA depletion or disrupted microtubule transport, is required to form inclusions following the production of CTFs (Pesiridis et al., 2011; Arnold et al., 2013).

It is interesting to note that TDP-43-mediated neuron loss does not require mutation, mislocalisation or fragmentation of TDP-43, but requires intact RNA binding capacity (Voigt et al., 2010). This suggests that TDP-43 aggregation does not directly exert

toxicity, but may be toxic due to a loss of function caused by altered RNA binding activity of TDP-43.

The chaperone system, ubiquitin proteasome system (UPS) and autophagy work to remove aggregated TDP-43 (Urushitani et al., 2010; Brady et al., 2011). TDP-43 stabilises the expression of HDAC6, a protein involved in cytoskeleton trafficking and autophagy-mediated degradation of misfolded proteins (Fiesel et al., 2010). Inefficient clearance of aggregates may result in the accumulation of TDP-43 in the cytoplasm in disease.

TDP-43 proteinopathy is observed in all fALS, excluding SOD1 and FUS ALS (Keller et al., 2012), suggesting that these cases cause ALS via different disease mechanisms from those displaying TDP-43 proteinopathy. However, most reports of SOD1 fALS predate the discovery of TDP-43 inclusions in ALS, so the presence of TDP-43 inclusions may have been overlooked (Ince et al., 2011). It has also been shown that mtSOD1 protein, but not WT SOD1 protein, interacts with TDP-43 protein in the cytoplasm and can induce mild biochemical alteration of TDP-43 in cell culture and SOD1 mouse models (Higashi et al., 2010). Recent studies have also identified changes to TDP-43 metabolism in mtSOD1 patient fibroblasts (Sabatelli et al., 2015). This suggests there may be an interaction between mutant SOD1 and TDP-43 and that there are more similarities in the disease mechanisms than previously thought.

1.3.3 TDP-43 and RNA processing

TDP-43 functions as a ribonucleoprotein in the regulation of transcription and pre-mRNA splicing and has over 4000 binding targets. TDP-43 within the nervous system binds to >6000 pre-mRNAs, affects the levels of approximately 600 mRNAs and alters the splicing patterns of around 950 mRNAs (Polymenidou et al., 2011).

Therefore TDP-43 is involved in a diverse range of cellular processes and all aspects of mRNA transcription, translation and post-translational processing.

TDP-43 is involved in transcriptional repression and can prevent pre-mRNA transcription from DNA. This was first noted in the regulation of the HIV-1 gene expression (Ou et al., 1995). It has also been shown that mouse sperm acrosomal protein SP-10 is repressed by TDP-43 (Acharya et al., 2006; Abhyankar et al., 2007; Lalmansingh et al., 2011).

Following transcription, TDP-43 is recruited to pre-mRNA binding to UG repeats at 3' or 5' splice sites to direct the spliceosome to targets for alternative splicing. It can mediate exon exclusion, inclusion, alternative exon usage and alternative splice sites (Arnold et al., 2013; De Conti et al., 2015; Colombrita et al., 2015). For example, TDP-43 silences splicing in cystic fibrosis transmembrane regulator (CFTR) exon 9 (Buratti, Dork, et al., 2001), apolipoprotein AII (ApoAII) exon 3, eukaryotic termination factor 1 (ETF1) and retinoid X receptor gamma (RXRG) (Mercado et al., 2005) and enhances splicing in breast cancer 1 (BRCA1)-mutated substrate (Passoni et al., 2012) and S6 kinase 1 Aly/REF-like target (SKAR) (Fiesel et al., 2012). Abnormal splicing of TDP-43 mRNA has been shown to generate a pathogenic low molecular weight species of TDP-43 which causes motor neuron death (Xiao et al., 2015).

Abnormalities of TDP-43 in ALS are reflected by changes to RNA processing and in particular the aberrant splicing of mRNA (Xiao et al., 2011). Recently, it has been shown that aberrantly spliced mRNA species are translated. A mutant of the protein SKAR, a component of the exon junction complex (EJC), translated from an aberrant splice variant was also recently reported to increase global translation upon TDP-43 knock-down (Fiesel et al., 2012). This further suggests that aberrantly spliced mRNA can be exported, translated and may acquire new biological functions.

TDP-43 is also involved in miRNA production and regulation. TDP-43 (and FUS also) associates with Drosha and Dicer in the microprocessor complex (Gregory et al., 2004; Fukuda et al., 2007). Mature miRNAs are generated from primary miRNA (pri-miRNA) complexes, which are cleaved by Drosha to produce pre-miRNA. These pre-miRNA are transported to the cytoplasm by Exportin-5 where Dicer cleaves them into mature miRNA. Mature miRNA involved in silencing are incorporated into the RISC complex where they bind to a miRNA recognition element in the 3' UTR of the target mRNA to repress or degrade these mRNA. TDP-43 facilitates the production of pre-miRNA in the nucleus by its association with Drosha and binding to pri-miRNA. In the cytoplasm TDP-43 interacts with Dicer and binds to the terminal loop of pre-miRNA to promote their processing (Kawahara et al., 2012). TDP-43 facilitates the post-transcriptional processing of a small subset of miRNAs including those required for neuronal outgrowth (Kawahara et al., 2012). TDP-43 can also negatively regulate certain miRNA by affecting their availability for incorporation into the RISC complex (I. N. King et al., 2014).

1.3.3.1 TDP-43 in RNP complexes

TDP-43 is a component of ribonucleoprotein (RNP) complexes (also referred to as RNA granules) that are required for the regulation of RNA metabolism. RNP complexes are made up of ribosomal subunits, translational factors, RBP, helicases and decay enzymes. They are bundled into RNP complexes in order to coordinate the expression of nascent mRNA and direct protein synthesis in response to cellular needs. There are three types of RNP complexes in neurons: transport granules, stress granules and processing bodies (P-bodies). Proteomic analysis has identified TDP-43 as having strong interactions with translation machinery and extensive interactions with proteins involved in the regulation of RNA metabolism (Freibaum et al., 2010).

Transport granules (tRNP) are thought to contain mRNA in a translationally silent state until it has been transported to the site of protein synthesis either along neurites or to synapse terminals (Hirokawa, 2006). TDP-43, along with other RNA binding proteins such as heterogeneous nuclear ribonucleoprotein A2 (hnRNPA2), Staufen proteins, Purine-Rich Element Binding Protein alpha (Purα), and motor proteins, such as kinesins, are components of RNP complexes and involved in the transport of mRNA to dendrites (Kanai et al., 2004; Kiebler et al., 2006; I.-F. Wang et al., 2008). In neurons, TDP-43 is a key component in dendritic RNA transport granules and the regulation of local protein synthesis in dendrites to maintain neuronal plasticity (Diaper et al., 2013; Alami et al., 2014; Liu-Yesucevitz et al., 2014).

Stress granules (SG) are cytoplasmic complexes consisting of RBP, RNA and stalled translation initiation complexes (Anderson et al., 2008). They are formed upon cellular stress in a reversible manner and following stress mRNA is either exchanged with RNP, made translationally active or exported to P-bodies for degradation. P-bodies contain factors involved in RNA degradation and RNA silencing. There is a dynamic exchange between SG and P-bodies and RNA transcripts held in P-bodies can either be returned for translation or degraded (R. Parker et al., 2007).

In response to stress, TDP-43 (and TDP-35) associate with stalled ribosomes in SG, playing a role in mRNA stability and cell survival (Colombrita et al., 2009; Nishimoto et al., 2010; Dewey et al., 2011; Higashi et al., 2013). However, TDP-43 is not required for SG formation (Colombrita et al., 2009). Cells with mutant TDP-43 react differently in response to stress, compared to cells with WT TDP-43, suggesting that mutations in TDP-43 may affect SG dynamics, (Dewey et al., 2011). It has been shown in *Drosophila* that the remodelling of TDP-43 RNA granules can reduce the toxicity of TDP-43 by restoring translation (Coyne et al., 2015). TDP-43 recruited to SG in response to stress may be responsible for the formation of irreversible protein

aggregates as observed in TDP-43 proteinopathies (S. J. Parker et al., 2012; Hyung-Jun Kim et al., 2014).

As a member of the hnRNP family, TDP-43 is aided by and forms complexes with a variety of other hnRNPs (Buratti et al., 2005; 2010; M. Romano et al., 2014). TDP-43 interacting with other hnRNPs affects a variety of the functions of TDP-43, including its binding to other hnRNPs, the aggregation of TDP-43 and regulation of the splicing/alternative splicing of other hnRNPs (Buratti et al., 2005; Highley et al., 2014; Lauranzano et al., 2015). hnRNP K binds to TDP-43, creating SG in response to cellular stress and loss of hnRNP K from SGs can prevent accumulation of TDP-43 (Moujalled et al., 2015). This suggests that other hnRNPs may also be involved in TDP-43 proteinopathy.

1.3.4 Other roles and functions of TDP-43

1.3.4.1 TDP-43 in mitochondrial dynamics and function

Altered mitochondrial morphology and function has been extensively studied in ALS, particularly in SOD1 ALS (Cozzolino et al., 2013). TDP-43 may also be directly involved in mitochondrial dysfunction. Mutant TDP-43 impairs mitochondrial dynamics and function in motor neurons via enhanced localisation in mitochondria (W. Wang et al., 2013). TDP-43 protein accumulated in the mitochondria of neurons has been observed in ALS/FTD patients (W. Wang et al., 2016). Mitochondrial phenotypes and alterations to mitochondrial distribution and morphology have been described in TDP-43 mouse models (Xu et al., 2010; Shan et al., 2010; Xu et al., 2011). TDP-43 and its fragments can also enhance mitochondrial dysfunction and activate mitophagy, the clearance of damaged mitochondria (Hong et al., 2012).

1.3.4.2 TDP-43 and development

TDP-43 is developmentally regulated and is essential for early embryonic development (Sephton et al., 2010). Lack of TDP-43 leads to embryonic lethality and

motor dysfunction in homozygous and heterozygous knock out mice (L.-S. Wu et al., 2010; Kraemer et al., 2010). Depletion of TDP-43 in *Drosophila* caused altered locomotive function and abnormal neuromuscular junctions (NMJ) (Feiguin et al., 2009). Zebrafish models also show defective development of motor neurons and a motor dysfunction upon knock-down of TDP-43 (Kabashi et al., 2010). TDP-43 also has splicing activities in genes essential for neuronal development and CNS integrity (Rogelj et al., 2012; Lagier-Tourenne et al., 2012). This suggests that alterations to regulation and function of TDP-43 in development could also play a role in later pathogenic degeneration that occurs in disease.

1.3.4.3 Axonal growth and synaptic transmission

TDP-43 is specifically localised to motor neuron axons and regulates dendrite and axon outgrowth and cytoskeleton integrity (Fallini et al., 2012; Tripathi et al., 2014). Depletion of TDP-43 inhibits neurite growth and induces death by dysregulation of Rho GTPases (Iguchi et al., 2009). TDP-43 is also required by glia in the regulation of morphology and function of neuromuscular junctions (Estes et al., 2013; G. Romano et al., 2015). The fact that TDP-43 levels can affect the growth of axons and neurites suggests that TDP-43 plays a role in the processing and axonal fate of mRNAs.

1.3.5 Modelling TDP-43 ALS

In order to elucidate the physiological role of TDP-43 in the nervous system both *in vitro* cell models and *in vivo* animal models have been generated to try and understand the pathogenic mechanisms of mutant TDP-43 in ALS. To do this investigators have either taken a gain-of-function approach overexpressing TDP-43 protein or a loss-of-function approach by knockdown or conditional knockdown of the TDP-43 protein. The recent advances in inducible systems allow for a more controlled expression of TDP-43.

Modelling TDP-43 proteinopathy has remained a challenge for over a decade since its discovery. There are therefore several considerations to take into account when designing a TDP-43 related disease model. The overall expression of TDP-43 wild type and ALS mutant protein needs to remain at comparable physiological expression levels, where total levels of wild type and pathogenic TDP-43 proteins appear similar to healthy controls. TDP-43 is able to auto-regulate its own levels (Ayala et al., 2011; Polymenidou et al., 2011), therefore chromosomal integration of a transgene into a model can result in the down-regulation of endogenous TDP-43 levels (Igaz et al., 2009). Over-expression and knock-down of TDP-43 also produce a variety of differing side effects and have very different consequences with respect to gene expression and splicing (Hazelett et al., 2012). This can cause issues in the identification of disease related effects, which may be masked by over-expression effects. Although over-expression of TDP-43 is inherently toxic, there are differential differences between different models of WT TDP-43 and mutant TDP-43, suggesting that individual mutations themselves also need to be taken into account (Estes et al., 2011). The time of the induction of TDP-43 WT or mutant TDP-43 in the development of an organism can also have an effect on the neuronal sensitivity and so is also worth consideration (Cannon et al., 2012).

1.3.6 Cellular models of TDP-43 ALS

Various components of TDP-43 pathology have been modelled by transient transfection into immortalised cell lines (Winton et al., 2008; Igaz et al., 2009). However, transient transfection of TDP-43 results in huge over expression of TDP-43, which is not a true reflection of disease. Therefore inducible cell lines have been generated in other studies to overcome this problem where gene expression can be controlled to produce a more moderate expression of TDP-43. Cells derived from fibroblasts from patients with TDP-43 mutations have also been generated. A summary of cellular models of TDP-43 ALS is shown in Table 1.3.

Table 1.3 Cellular models of TDP-43 ALS

Different cellular models of TDP-43 ALS are described. The type of model, TDP-43 localisation in model, main findings are shown, alongside the relevant literature reference.

Model	Type	TDP-43 localisation	Main findings	Reference
Tetracycline inducible				
HEK Flp (Invitrogen)	WT TDP-43, TDP-43 G298S, TDP-43 Q331K, TDP-43 M337V	Predominantly nuclear	Mutations in TDP-43 increase stability and complex formation with FUS	(Ling et al., 2010)
HEK Flp (Invitrogen)	FLAG TDP-43 WT, FLAG TDP-43 A90V	Not investigated		(Zhijun Zhang et al., 2013)
HEK Flp (Invitrogen)	12 repeats of Q/N region of TDP-43 339-369 linked to EGFP (EGFP-12xQ/N)	Aggregates sequester endogenous TDP-43 causing nuclear depletion	N terminus sequence of TDP-43 enhances interaction of TDP-43 with aggregates and their insolubility	(Budini et al., 2014)
Doxycycline inducible				
M17	TDP-43 Q331K	Data not shown	Inhibition of Dbr1 suppresses TDP-43 toxicity in cell line	(Armakola et al., 2012)
HEK, SH-SY5Y	TDP-43, HA-tagged and EGFP tagged TDP-43 WT, TDP-43 ΔNLS, TDP-43 CTF	TDP-43 WT predominantly nuclear, TDP-43 ΔNLS nuclear and cytoplasmic, TDP-43 CTF cytoplasmic	Soluble TDP-43 degraded by UPS, TDP-43 aggregate clearance requires autophagy	(Scotter et al., 2014)
PC12	TDP-43 WT-GFP	Diffuse nuclear	Sodium arsenite induced aggregation response from cell lines, novel compounds found to inhibit this	(Boyd et al., 2014)
TDP-43 ALS patient derived				
iPSCs from fibroblasts	TDP-43 M337V	Predominantly nuclear	M337V specific susceptibility to antagonism of PI3K signalling	(Bilican et al., 2012)
iPSCs from fibroblasts	TDP-43 Q343R, TDP-43 M337V, TDP-43 G298S	Nuclear and cytoplasmic, cytoplasmic aggregates	Anacardic acid reverses ALS phenotype	(Egawa et al., 2012)
iPSCs from fibroblasts	TDP-43 M337V, TDP-43 A90V	Nuclear, mislocalised in response to stress (staurosporine)	Downregulation of MicroRNA-9 when derived from patients with TDP-43 mutations	(Zhijun Zhang et al., 2013)
iPSCs from fibroblasts	Fibroblasts from sALS and fALS	Predominantly nuclear, nuclear aggregates	iPSC MNs show TDP-43 aggregates and are suitable for use in drug screening	(Burkhardt et al., 2013)
Co-cultures Muscle-MN	Mouse embryonic stem cells stably expressing TDP-43 WT or TDP-43 A315T	Predominantly nuclear	Non-cell autonomous effects on MNs caused by muscle cells expressing human TDP-43 A315T	(Wächter et al., 2015)

The main limitation with current cellular models of TDP-43 is the inability to consistently reproduce the TDP-43 proteinopathy that is observed in disease. TDP-43 mislocalisation can be achieved in neuronal cells with the addition of stressors. Issues with the toxic over-expression of TDP-43 are overcome by inducible cell model systems. However, currently the majority of these cell lines are in non-neuronal cell types and the neuronal cell types available still do not reproduce TDP-43 proteinopathy. Despite being derived from patient fibroblasts iPSC MN cultures also do not produce consistent TDP-43 pathology, with differences between cultures derived from the same patient (Egawa et al., 2012).

1.3.7 Animal models of TDP-43 ALS

A variety of *in vivo* TDP-43 models exist, these include rodent (Tsao et al., 2012; McGoldrick et al., 2013), *Drosophila melanogaster* (M. Romano et al., 2012; Casci et al., 2015), Zebrafish (Kabashi, Brustein, et al., 2011) and *Caenorhabditis elegans* (Therrien et al., 2014). Non-human primate models have also been described (Uchida et al., 2012; Jackson et al., 2015).

1.3.7.1 Rodent models

Knock out of *Tardbp* resulted in rodents which died early in embryogenesis, providing evidence for the essential role of TDP-43 in development (L.-S. Wu et al., 2010; Sephton et al., 2010; Kraemer et al., 2010). Therefore conditional knock out mice were generated which caused loss of body fat and rapid death (Chiang et al., 2010). This identified Tbc1d1, a protein linked with obesity, as being downregulated upon *Tardbp* knockout and established a role for TDP-43 in fat metabolism. Specific knock out of *Tardbp* Hb9-positive spinal cord motor neurons in mice, similarly, resulted in mice with lower bodyweight as well as progressive, male-dominant ALS phenotypes including motor neuron loss, muscle weakness/atrophy and motor dysfunction (L.-S. Wu et al., 2012)

Several groups have used the mouse prion protein promoter (mPrP) (Borchelt et al., 1996) to drive protein expression of human TDP-43 WT, TDP-43 ALS mutations A315T/M337V/Q331K, in the CNS of mice (Wegorzewska et al., 2009; Xu et al., 2010; Stallings et al., 2010; Xu et al., 2011; Y. Guo et al., 2012; Arnold et al., 2013). Mice expressing high levels of TDP-43 in the spinal cord and brain had motor deficits and early lethality. However, there was no TDP-43 proteinopathy observed and no difference in the phenotypes of mice expressing comparable levels or either TDP-43 WT or mutant TDP-43.

Modified murine Thy1 promoters (Thy1.2) have been used to drive the expression in TDP-43 mouse lines almost exclusively in neurons from one week of age in order to overcome any issues with over-expression during development (Wils et al., 2010; Shan et al., 2010). Other promoters have been used to target TDP-43 expression to particular areas, such Ca2+/calmodulin-dependent kinase II (CaMKII) promoter to target the forebrain (Tsai et al., 2010).

In order to try and overcome toxicity due to high levels of TDP-43 protein, bacterial artificial chromosome (BAC) and endogenous size of promoters (Swarup et al., 2011; Mutihac et al., 2015), as well as inducible systems (Cannon et al., 2012; Stribl et al., 2014; Alfieri et al., 2014; Walker et al., 2015) have been used to generate mice with more moderate expression of TDP-43. Again, these models displayed age-related development of motor and cognitive dysfunction, with some features of TDP-43 pathology. In some of the models, partial reversal of motor and cognitive deficits was achieved by short term suppression expression in the later stages of disease (Alfieri et al., 2014; Walker et al., 2015).

Several rat models of TDP-43 ALS have also been generated (H. Zhou et al., 2010; C. Huang et al., 2012). Constitutive expression of mutant TDP-43 M337V, but not

WT TDP-43, caused widespread neurodegeneration that predominantly affected the motor system and cytoplasmic phosphorylated/fragmented TDP-43 protein (H. Zhou et al., 2010). Tetracycline inducible expression of mutant human TDP-43 in rat models has been achieved using neurofilament heavy chain promoters and choline acetyltransferase (ChAT) promoters (C. Huang et al., 2012). Similarly, partially reversible motor neuron degeneration was observed with some ubiquitination aggregation and cytoplasmic mislocalisation of TDP-43.

In summary, the onset of pathology and presentation of inclusions and aggregate formation is not uniform amongst the different rodent models. Despite replicating some aspects of the ALS pathology, TDP-43 proteinopathy defined by the nuclear loss and cytoplasmic mislocalisation is not observed in rodent models. High levels of TDP-43 expression are also produced, meaning these models are all over-expression models. Phenotypes observed can depend upon the promoter, time of induction or region of expression. An additional side effect of gastrointestinal problems in some mouse models is also observed following TDP-43 expression (Wegorzewska et al., 2009; Y. Guo et al., 2012; Herdewyn et al., 2014; Hatzipetros et al., 2014).

1.3.7.2 Drosophila melanogaster

Drosophila are one of the most widely used model organisms due to their short life span and ease of genetic manipulation (M. Romano et al., 2012; Casci et al., 2015). *Drosophila melanogaster* is used as it has the highest species homology with humans and has been used thoroughly to model neurodegeneration. *Drosophila* have a TDP-43 ortholog TBPH which has the structure and *in vitro* functions, such as repression of splicing of specific exons, similar to the human protein (Ayala et al., 2005; M. Romano et al., 2014). The loss of TBPH during development is also lethal (Feiguin et al., 2009).

Flies deficient in TBPH show reduced life span and anatomical defects at the NMJ (Feiguin et al., 2009; Diaper et al., 2013). These phenotypes can be rescued in some neurons, including motor neurons, by expression of human TDP-43 (Feiguin et al., 2009). This suggests that a loss of function of TDP-43, rather than toxic aggregation may cause ALS and that impaired synaptic transmission and NMJ defects may be the early events in TDP-43 ALS pathogenesis.

Overexpressed and mutated TBPH or human TDP-43 results in movement defects and paralysis, locomotive deficits and impaired NMJ function (Cragnaz et al., 2014; G. Romano et al., 2015; Cragnaz et al., 2015). Dose and age dependent eye degeneration specific to fly models is also observed in *Drosophila* overexpressing or expressing mutant TBPH/TDP-43 (Estes et al., 2013; Cragnaz et al., 2014). These eye phenotypes are also accompanied by axonal aggregation of TDP-43 providing further evidence of an association between proteinopathy and disease (Estes et al., 2013).

Loss and gain of TBPH alters the mRNA expression of NMJ glutamate transporters excitatory amino acid transporter (EAAT) 1 and EAAT2 (Diaper et al., 2013). TBPH expression in motor neurons and glial cells have opposing synaptic phenotypes but cause comparable locomotive defects (Estes et al., 2013). The effects on gene expression and splicing are very different following over-expression or depletion of TBPH in the CNS of *Drosophila* (Hazelett et al., 2012). However gene enrichment analysis again identified TDP-43 as having a role in the control of genes involved in synaptic release and synaptic transmission.

TBPH and human TDP-43 aggregation in *Drosophila* has been modelled (Cragnaz et al., 2014; Budini et al., 2014; Cragnaz et al., 2015). These models use EGFP-12x Q/N constructs to express a Q/N repeat insertion in the CTD which promotes

inclusion formation, but is not observed in human pathology (Budini et al., 2012). Using this model system and expressing TBPH in the *Drosophila* eye has shown that neurotoxicty due to TBPH over-expression can be prevented by its incorporation into aggregates, suggesting a neuroprotective role of aggregation in modulating levels of TBPH (Cragnaz et al., 2014). Age-related reductions in TDP-43/TBPH mRNA levels which precede the onset of locomotor defects have also been observed using these models (Cragnaz et al., 2015).

Other studies using *Drosophila* have shown that levels of HDAC6 mRNA and protein is decreased upon TDP-43 silencing and is accompanied by an increase in HDAC6 substrate acetyl-tubulin (Fiesel et al., 2010). This is of interest in relation to TDP-43 proteinopathy, as HDAC6 has also been linked to the pathogenesis of AD and is associated with two protein degradation pathways: the ubiquitin proteaosome system and autophagy (Pandey et al., 2007).

TDP-43 has also been shown to suppress CGG repeat induced toxicity in a *Drosophila* model of FTXAS via interactions with hnRNP A2/B1 (He et al., 2014). These data suggest a convergence in the mechanisms behind repeat expansions and RNA binding proteins in neurodegeneration.

1.3.7.3 Zebrafish

Zebrafish have several key advantages as a vertebrate model organism such as their external development and optical transparency. Several neurodegenerative diseases have already been modelled in the zebrafish (Bandmann et al., 2010). Zebrafish contain tardbp, and a paralogue tardbpl (TAR DNA binding protein-like). Therefore, zebrafish do not represent the best system for modelling TDP-43 mediated neurodegeneration due to the presence of two tardbp-like genes. *tardbp* mutants do not show a phenotype due to compensation by a unique splice variant of *tardbpl* (B. Schmid et al., 2013). The loss of both results in locomotor swimming defects,

impaired axon growth and alterations to synaptic transmissions (Kabashi, Bercier, et al., 2011; B. Schmid et al., 2013; Hewamadduma et al., 2013). Similar behaviours were observed in zebrafish expressing mutant TDP-43 (G348C) or FUS (Armstrong and Drapeau, 2013a; 2013b).

1.3.7.4 Caenorhabditis elegans

Caenorhabditis elegans (*C. elegans*) are useful models to study neurodegenerative diseases owing to their small size, transparency and short life cycle of 3 days from egg to adult at 25 °C (A. G. Alexander et al., 2014). *C. elegans* also have a well characterised nervous system making them ideal to study neuronal toxicity (Therrien et al., 2014). Furthermore, the RNA toxicity produced by non-coding mutations in *C. elegans* reflects the toxicity observed in mammals (L.-C. Wang et al., 2011).

Several groups have generated *C. elegans* TDP-43 models which have phenotypes that include motility defects (Ash et al., 2010; Liachko et al., 2010) accompanied by hyperphosphorylation, truncation and ubiquitination of TDP-43 accumulation (Liachko et al., 2010). Some of these models also showed phenotypes which became more severe with age (Liachko et al., 2010).

C. elegans also has an ortholog of TDP-43, TDP-1, which is a primarily nuclear protein, which is expressed in most tissues including neurons (Ayala et al., 2005). Mutant TDP-1 animals show similar phenotypes including slow development and locomotor defects (Liachko et al., 2010; Tao Zhang et al., 2012).

1.3.8 Alteration of gene expression in TDP-43 ALS

Comprehensive transcriptomic studies of ALS have been carried out in human post mortem tissue, animal models and cell model systems (F. Tanaka et al., 2011; Cooper-Knock, Kirby, et al., 2012; Heath et al., 2013).

In TDP-43 transgenic mouse models, up to one third of the transcriptome is altered, with specific alteration due to the Q331K ALS mutation (Arnold et al., 2013). Transcriptomic analysis of TDP-43 has identified alteration of levels and/or the splicing of genes involved in RNA processing, neurotrophic factor synthesis and synaptic function (Arnold et al., 2013; C. Huang et al., 2014; Highley et al., 2014).

TDP-43 has over 4000 RNA binding targets, many of which play a critical role in neuronal development and plasticity (Sephton et al., 2011; Tollervey et al., 2011). Individual-nucleotide crosslinking and immunoprecipitation (iCLIP) coupled with high throughput RNA sequencing (CLIP-seq), have allowed the mapping of TDP-43 binding sites in RNA from both rodents (Sephton et al., 2011; Tollervey et al., 2011) and humans (Polymenidou et al., 2011; Tollervey et al., 2011; Lagier-Tourenne et al., 2012). Recent studies of TDP-43 RNA targets and TDP-43 transcriptomes are summarised in Table 1.4.

1.4 Genome-wide expression profiling

Gene expression and its tight regulation are vital to maintain homeostasis of cells. Genome-wide expression profiling has been a widely used tool to investigate the transcriptome of prokaryotes and eukaryotes in ALS either by microarray or more recently by mRNA sequencing (RNA-seq) (Cooper-Knock, Kirby, et al., 2012; Heath et al., 2013).

The analysis of transcript abundance may not constitute the best method to study gene expression. This is because mRNA transcript levels are not necessarily linked to the levels of their corresponding proteins, due to cellular compensatory mechanisms. An increase in the steady state level of mRNA can be associated with a decrease in the protein level as the cell tries to counteract the down-regulation of a

protein by increasing the stability or amount of mRNA (Domínguez-Sánchez et al., 2011).

The direct analysis of translation would provide a more accurate measure of gene expression. The identification of mRNAs directly undergoing protein synthesis should better reflect protein expression changes and the directionality of altered biological processes in health and disease.

Table 1.4 Studies investigating TDP-43 RNA targets and TDP-43 transcriptomes

Studies investigating RNA targets of TDP-43 and the TDP-43 transcriptomes. The samples, platform for the study, strategy and main findings of each study are listed beside the literature reference. (U2aF65: U2 small nuclear RNA auxiliary factor 2)

Study	Samples	Platform	Strategy	Main findings
(Sephton et al., 2011)	Rat cortical neurons	RIP and RNA sequencing (RIP-seq)	RNA sequencing. Read density in exons and introns compared and classed targets as exonic, intronic or dual targets.	TDP-43 RNA targets enriched for synaptic function, RNA metabolism and neuronal development
(Tollervey et al., 2011)	Human cortical tissue from post mortem samples from 3 healthy controls and 3 sporadic FTLD-TDP patients; human embryonic stem cells; SH-SY5Y	CLIP the RNA sequencing (CLIP-seq)	RNA sequencing. Identify TDP-43 targets. Knock down TDP-43 in SH-SY5Y and characterise changes in alternative splicing	Alternative mRNA isoforms regulated byTDP-43 encode proteins that regulate neuronal development or implicated in neurodegenerative disease.
(Xiao et al., 2011)	SH-SY5Y	UV-CLIP	UV-CLIP followed by conventional cloning to identify targets	TDP-43 binds mainly to intronic regions. Alternative splicing of TDP-43 targets in ALS vs control lumbar spinal cord.
(Polymenidou et al., 2011)	Adult mouse brain	CLIP-seq	Antisense oligonucleotide knock down of TDP-43 in mouse striatum.	TDP-43 auto-regulates its synthesis. Knock down of TDP-43 causes alternative splicing of targets. RNA that have depleted levels following TDP-43 knock down are involved in synaptic activity.
(Lagier-Tourenne et al., 2012)	Mouse and human brain	CLIP-seq	Antisense oligonucleotide knock down of FUS or FUS and TDP-43	Depletion of FUS and/TDP-43 affects levels or splicing of genes with long introns that encode genes essential for neuronal integrity.
(Rogelj et al., 2012)	E18 mouse brain	iCLIP of TDP-43, FUS and control U2AF65	Transcriptome-wide binding map	TDP-43 (and FUS) bind distinct sites and regulate distinct alternative exons. TDP-43 (and FUS) regulate splicing of genes enriched in neuronal development and neurodegenerative diseases.
(Arnold et al., 2013)	Transgenic mice expressing TDP-43 WT or TDP-43 Q331K in CNS. RNA extracted from cortex and spinal cord in 2 month-old transgenic and non-transgenic mice	Splicing sensitive microarrays	Compared with non-transgenic, absolute separation score < 0.3, $p < 0.05$	Differentially spliced exons in Q331K but not WT enriched for known TDP-43 binding sites and synaptic function
(C. Huang et al., 2014)	Astrocytes from (GFAP)-tTa/TRE-TDP-43 M337V double transgenic rats. RNA extracted after 3/4/6 days of induction of mtTDP-43	Microarray	Compared with baseline FC > 2 at one or more time points and progressive change in FC over induction time points	Induction of TDP-43 M337V expression altered expression of secreted factors, neurotrophic factor expression increased, neurotoxic factor expression decreased
(Highley et al., 2014)	RNA extracted from fibroblasts derived from ALS patients with TDP-43 and SOD1 patients, sporadic ALS patients and controls (n = 3-6 from each group)	Microarray	Compared with controls, ANCOVA $p < 0.01$	Functional enrichment in differentially expressed/spliced genes for categories related to RNA processing in mtTDP-43. Alternative splicing most abundant in mtTDP-43 ALS compared to other ALS cases

1.4.1 Current methodologies to investigate the translatome

The term translatome has been used to define mRNAs that are directly undergoing protein synthesis. Current methodologies to investigate the translatome identify mRNAs that are associated with the ribosome and from this infer the translatome.

A variety of methods using polysome separation, ribosome profiling and immunoprecipitation have been described to study translation (Kapeli et al., 2012; Ingolia, 2014).

1.4.1.1 Polysome profiling

Early, qualitative studies of global translation were carried out by comparing ribosome-bound mRNA with the total amount of mRNA within a cell. Polysome profiling later improved the quantification by separating mRNA on the basis of the number of ribosomes bound to them. Polysomes are complexes formed of single mRNA molecules associated with up to 5-10 ribosomes. The first reported translatome was generated using polysome profiling, which involves the separation of polysomes through sedimentation in a sucrose gradient (Johannes et al., 1999; Arava et al., 2003). Extraction of mRNA from polysomes can then be identified using qRT-PCR, microarrays or more recently using next generation RNA sequencing.

The use of polysome profiling as a high throughput method is hindered due to the technical difficulties associated with polysome fractionation, which requires delicate manipulation of the sucrose gradient, ultracentrifugation and the collection of many fractions per sample, which can cause issues with reproducibility and large numbers of samples to analyse. Moreover, high molecular weight RNP complexes, named pseudo-polysomes, were shown to contaminate polysomal fractions (Thermann et al., 2007).

1.4.1.2 Ribosome profiling

The position of the ribosome on a transcript can be determined as ribosomes protect a discrete footprint of approximately 30 nucleotides on its mRNA template from nuclease digestion (Steitz, 1969). Cycloheximide treatment causes ribosomes to physically enclose 28-30 nucleotides of a transcript, shielding them from nuclease digestion (Wolin et al., 1988). This allows for the mapping of ribosomes on mRNA by foot printing. Before the advances with microarray and RNA sequencing, fragments would be analysed by direct internal radiolabelling of mRNA being translated or by primer extension or both. Despite the limitations of only being able to analyse footprints from single transcripts *in vitro*, this method enabled the first mapping of the locations of translation initiation sites and allowed the study of ribosome dynamics during translation, revealing the sites where translation is stalled before translocation (Steitz, 1969; Wolin et al., 1988).

More recently, ribosome profiling or foot printing has allowed genome-wide mapping *in vivo* of ribosomes onto RNA molecules at a near nucleotide resolution (Ingolia et al., 2009). Ribosome profiling uses a limited RNAse digestion prior to sucrose gradient isolation of ribosome-protected RNA fragments (ribosome/monosome footprints) before identification of RNA fragments by next generation RNA sequencing.

This method has been used to expand knowledge of ribosome biochemistry and cellular translational activity. Examples include the study of translation elongation rates (Ingolia et al., 2011), identification of novel translated sequences (Ingolia et al., 2009; 2014), the role of miRNA in decreasing mRNA levels (H. Guo et al., 2010) and translational reprogramming (Liu et al., 2014).

An updated method of ribosome profiling has been described which does not require the purification of ribosomes (D. W. Reid et al., 2015). It instead extracts RNA after nuclease digestion, separating ribosome-protected fragments by size using gel electrophoresis. Ribosome protected fragments are then cut from the gel and the RNA extracted for RNA sequencing.

The application of this method, as with polysome profiling, is hindered due to the technical difficulty of the protocol, which again includes the delicate manipulation of sucrose gradient sedimentation. Size exclusion chromatography has been used to replace ultracentrifugation in commercially available kits (Freeberg et al., 2013). However, this approach is much more costly than traditional ribosome profiling and does not reduce the amount of time required for the protocol. It has been used to delineate translational activities in a cell, but not quantification of translated transcripts between two cell lines (control/normaliser versus experimental/disease). A common issue comes from the fact that the particularly short sequences of ribosome-protected RNA fragments (approximately 25-30 nucleotides) prevent paired-end sequencing and present added challenges for mapping reads to repeated sequences and alternatively spliced transcripts (Ingolia, 2014). Moreover, this technique will also generate RNA sequences that are not necessarily translated as it isolates monosomes after a limited RNAse digest (so includes both polysomes and 80S initiating complex).

1.4.1.3 Ribosome affinity purification

Ribosome affinity purification, does not rely on the sucrose gradient separation of ribosome species and further allows cell-type-specific purification of tagged ribosome subunits (Halbeisen et al., 2009). It was firstly applied to the tandem-affinity purification of PUMILIO under the control of an ovary-specific promoter in *Drosophila* (Gerber et al., 2006).

Translating ribosome affinity purification (TRAP) was further developed using restricted neuronal expression of the GFP-tagged ribosomal protein L10a (Heiman et al., 2008; Kulicke et al., 2014). Similarly, TRAP was used to delineate the translational activity of YFP-tagged RPL10a ribosomes in subcellular domains of Purkinje neurons (Kratz et al., 2014) and specific muscle cells in *Drosophila* embryos (Bertin et al., 2015).

Ribosome affinity purification offers the best high-throughput prospects compared to polysome profiling and ribosome profiling. However, it involves the over-expression of tagged ribosomal subunits, which could potentially affect the assembly of ribosomes. It also requires the generation of transgenic cell lines or animal models for the investigation of each cell type of interest.

Ribotag uses a very similar strategy to TRAP and involves inserting a tag onto a ribosomal subunit (Sanz et al., 2009). Expression of this tag requires Cre recombination and therefore Ribotag mice can be crossed with neuron-specific Cre-recombinase expressing mice before immunoprecipitation to separate mRNA from neuronal cell populations expressing Cre recombinase for example. This method has an advantage over TRAP given that Ribotag mice can be crossed with any Cre recombinase expressing mouse line, rather than the generation of a new mouse line for each cell type of interest. However, it does still require the tagging of ribosomal proteins and generation of transgenic animals.

Ribosome affinity purification using TRAP and Ribotag technologies offer the highest of the high-throughput prospects compared to polysome profiling and ribosome profiling.

1.4.2 Challenges facing translatome technologies

The current technologies are limited due to the lack of discrimination between actively translating mRNA and non-translating, ribosome-associated mRNA (mRNA indirectly bound to ribosomes or bound to non-translating mRNA). This is reflected in the number of publications related to the identification of translatomes (<100 in *PubMed*) in comparison to the number of publications related to the interrogation of transcriptomes (>27,000 in *PubMed*).

It is further hindered by the fact that some proteins such as TDP-43 can interact with stalled ribosomes, potentially contaminating ribosomes with non-translating mRNA which is being transported by TDP-43 (Higashi et al., 2013). Pseudopolysomes have also been shown to contaminate polysomes, highlighting the need for investigating multiple sucrose gradients (Thermann et al., 2007). Further development of these methodologies will allow for more accurate investigation of the translatome.

1.5 Summary and overall aims of the project

All sufferers of ALS experience the same symptoms and face the same bleak prognosis, despite differing genetic backgrounds. Mutations in the four major genes of interest in ALS can all be linked to the dysregulation of protein synthesis. The molecular mechanisms underlying ALS mutations are not well characterised. The nuclear export of aberrant pre-mRNA and translation of aberrant pre-mRNA at genome-wide level has not been investigated. Characterisation of pre-mRNA, which are exported into the cytoplasm and translated into proteins, may identify factors whose synthesis is dysregulated. This novel research would provide an insight into the molecular mechanisms involved in ALS and could reveal possible targets for therapeutic manipulation.

TDP-43 proteinopathy is observed in the majority of ALS cases. Transcriptomic studies in post mortem material and cell or animal models have identified thousands of altered events in TDP-43 related ALS. However, from these events it is not clear which events are casual and which may be the consequence of an earlier dysregulation. Therefore, analysis of the TDP-43 ALS translatome may elucidate early events in the pathogenesis of ALS and could identify new disease pathways and therapeutic targets.

The overall aims of this project were to:

- Test current translatome methodologies in order to develop a method to purify mRNAs preferentially undergoing active translation
- Select and build an appropriate TDP-43 ALS cell model with Tetracycline inducible expression of TDP-43 WT and TDP-43 Q331K
- Characterise the cell model for TDP-43 mediated toxicity and proteinopathy
- Use this cell model to create a TDP-43 ALS translatome
- Compare the TDP-43 ALS total transcriptome, cytoplasmic transcriptome and translatome

2. Materials and Methods

All chemicals, reagents and primers were from Sigma unless otherwise stated.

2.1 Molecular Cloning

2.1.1 Strains, cell lines and plasmids

2.1.1.1 Strains and cell lines

Bacterial strain

Escherichia coli (*E. coli*) DH5-alpha (DH5-α) (Genotype: fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17)

Mammalian cell lines

HEK 293T: 293tsA1609neo

Human embryonic kidney cells 293T is a highly transfectable derivative of the 293 cell line into which the temperature sensitive gene for SV40 T-antigen was inserted.

HEK293 FRT: HEK293 FRT cell line (commercially available from Invitrogen) was created by transfecting HEK 293 cells with pFRT/lacZeo.

NSC-34: Mouse motor neuron-like hybrid cell line derived by (Cashman et al., 1992) by fusing the aminopterin-sensitive neuroblastoma N18TG2 with motor neuron-enriched embryonic spinal cord cells.

NSC-34 FRT: Constructed in SITraN by Dr Adrian Higginbottom by transfecting NSC-34 cells with pFRT/lacZeo.

2.1.1.2 Plasmids

All plasmids used in this study and their sources are shown in Table 2.1.

Table 2.1 Plasmid source and manipulations

Plasmid name	Description	Source
pcDNA5/FRT/TO/His	6 x HIS Tag inserted into *Hind*III and *Bam*HI restriction sites	Addgene
pcDNA5/FRT/TO-Empty	Negative control 'Sham'	Invitrogen
pPGKFLPobpA	Expression of recombinase	Addgene
pcDNA5 FRT/TO/GFP AH	pcDNA5/FTR/TO/GFP (Addgene 19444) has been cut with *Bam*HI and *Xho*I, end filled and ligated to create a stop codon at the 3' end of GFP	Dr Adrian Higginbottom, SITraN
pcDNA5 FRT/TO/*GFP*	*D37D M79V M89V* directed mutagenesis of pcDNA5/FRT/TO/GFP AH	This study
pcDNA5 FRT/TO/*M1V*	*M1V D37D M79V M89V M155V* directed mutagenesis of pcDNA5/FRT/TO/GFP	This study
pcDNA5 FRT/3xFLAG RPL10a	Expression of 3xFLAG tagged RPL10a	This study
pcDNA5/FRT/TO/*TARDBP* WT	Expression of untagged wild type TDP-43	Dr Adrian Higginbottom, SITraN
pcDNA5/FRT/TO/*TARDBP* A315T	Expression of untagged mutant TDP-43 A315T	Dr Adrian Higginbottom, SITraN
pcDNA5/FRT/TO/*TARDBP* M337V	Expression of untagged mutant TDP-43 M337V	Dr Adrian Higginbottom, SITraN
pcDNA5/FRT/TO/*TARDBP* Q331K	Expression of untagged mutant TDP-43 Q331K	Dr Adrian Higginbottom, SITraN

2.1.2 Molecular cloning

DNA fragments were amplified by PCR, digested with restriction enzymes, gel-purified and cloned into a dephosphorylated and restricted vector. Oligonucleotides were designed as outlined in Figure 2.1.

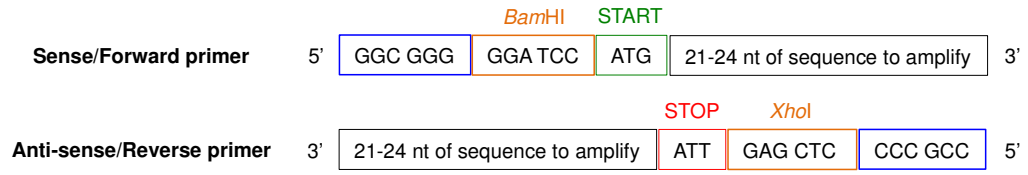


Figure 2.1 Strategy for oligonucleotide design
GC-rich sequence (outlined in blue) allows efficient binding and cleavage by restriction enzymes. Restriction sites *Bam*H1 and *Xho*I are shown as examples (outlined in orange). START (outlined in green) and STOP (outlined in red) codons are shown. 21-24 nucleotides (nt) of sequence were selected to amplify (outlined in black).

PCR samples contained 100 ng template DNA, 100 ng forward primer, 100 ng reverse primer, 2.5 mM 10 X dNTPs (Bioline), 5 µL 10 X Accuzyme buffer (Bioline), 36 µL dH_2O and 1 µL Accuzyme (Bioline). PCR mixtures were held at 94 °C for 2

min, followed by 25 amplification cycles of 30 sec at 94 °C, 30 sec at 54 °C and 1 min at 72 °C, before a final step of 10 min at 72 °C.

2.1.3 Quickchange site directed mutagenesis

PCR samples contained 50 ng plasmid DNA, 125 ng forward primer, 125 ng reverse primer, 2.5 mM 10 X dNTPs (Bioline), 5 µL 10 X PfuTurbo buffer (Stratgene), 4 µL DMSO (Dimethyl sulfoxide), 1 µL PfuTurbo DNA polymerase 3U (Stratagene) and the volume made up to 50 µL with dH_2O. Quickchange primers used are shown in Table 2.2. PCR mixtures were held at 95 °C for 30 sec, followed by 26 amplification cycles of 30 sec at 95 °C, 30 sec at 53 °C and 15 min at 68 °C, before a final step of 10 minutes at 68 °C. PCR product was then digested by DpnI (10 U, Roche) for 1 hour at 37 °C. The DNA was then transformed into *E.coli* DH5-α.

Table 2.2 Quickchange primers used for site directed mutagenesis
Blue indicates change to nucleotide sequence.

Primer	F	R
GFP QC M1V	aaa ctt aag ctt gcc acc gtg gtg agc aag ggc gag g	c ctc gcc ctt gct cac cac ggt ggc aag ctt aag ttt
GFP QC D37D	c ggc gag ggc gag ggc gac gcc acc tac ggc aag ctg	cag ctt gcc gta ggt ggc gtc gcc ctc gcc ctc gcc g
GFP QC M89V	gac ttc ttc aag tcc gcc gtg ccc gaa ggc tac gtc c	g gac gta gcc ttc ggg cac ggc gga ctt gaa gaa gtc
GFP QC M155V	agc cac aac gtc tat atc gtg gcc gac aag cag aag a	t ctt ctg ctt gtc ggc cac gat ata gac gtt gtg gct
GFP QC M79V	agc cgc tac ccc gac cac gtg aag cag cac gac ttc t	a gaa gtc gtg ctg ctt cac gtg gtc ggg gta gcg gct
TARDBP Q331K	gcc gcc cag gca gca cta aag agc agt tgg ggt atg a	t cat acc cca act gct ctt tag tgc tgc ctg ggc ggc

2.1.4 Agarose gel electrophoresis

DNA fragments produced by PCR or restriction digest were resolved on agarose (Bioline) gels at 80 V for 40 min or 1 h, respectively, using a Generuler™ DNA ladder mix (Thermo Scientific). The percentage agarose in the gel was determined using the size of the DNA fragments. 1-2 % w/v agarose was chosen for for DNA products shorter than 1 kbp and 0.8 % w/v for DNA products greater than 1 kbp.

2.1.5 DNA gel extraction

DNA fragments observed by preparative agarose gel electrophoresis were cut and the DNA extracted from the gel using the QIAquick® Gel Extraction Kit (Qiagen), following the manufacturer's instructions.

2.1.6 Restriction digests

Restriction digests were carried out using fast digest restriction enzymes (Thermo Scientific) and 10 X fast digest green restriction buffer (Thermo Scientific) in dH_2O at 37 °C for 2 h.

2.1.7 Ligation

T4 DNA ligase (Roche) and 10 X ligase buffer (Roche) were used to ligate the DNA insert and plasmids. The ligation was left at 16 °C overnight.

2.1.8 Competent cell preparation

50 mL LB broth (10 g Bacto tryptone, 10 g NaCl, 5 g yeast extract fill to 1 L with dH_2O, pH 7.0, autoclaved) was inoculated with *E.coli* DH5-α and left growing at 37 °C overnight. 400 mL LB broth was inoculated with the culture grown overnight to optical density of 600 nm (OD_{600}) =0.05 and grown at 37 °C until an OD_{600} = 0.48 was reached. The cells were centrifuged for 10 min at 17 000*g*, 4 °C. The pellet was then resuspended in TfbI buffer (15% v/v Glycerol, 30 mM KAc, 100 mM $RbCl_2$, 10 mM $CaCl_2$, 50 mM $MnCl_2$; pH 5.8) and left on ice for 10 min. The cells were centrifuged for 10 min at 17 000*g*, 4 °C. The pellet was resuspended in TfbII buffer (10 mM MOPS, 75 mM $CaCl_2$, 100 mM $RbCl_2$, 15% v/v Glycerol; pH 6.5) and left on ice for 10 min. The cells were then transferred into pre-chilled microcentrifuge tubes and frozen in liquid nitrogen. The competency of the cells to take up plasmids was tested before use using a known amount of plasmid.

2.1.9 Plasmid transformation

Plasmids were transformed into competent *E.coli* DH5-α ($RbCl_2$). 10 µL of ligation or 100 ng of plasmid DNA were chilled on ice with 100 µL competent *E.coli* DH5-α ($RbCl_2$) before heat shocking the mixture for 1 min at 37 °C, before returning to ice. 900 µL LB was added to the *E.coli* and left for a recovery period of 45 min at 37 °C. The *E.coli* was then spread onto LB agar plate (1.5% agar in LB broth) containing the appropriate antibiotic(s) and incubated at 37 °C overnight.

2.1.10 Plasmid purification

Plasmids isolated from transformed bacterial colonies were checked for successful integration of the DNA fragment/insert by performing a miniprep.

Bacterial colonies from transformation were inoculated in 5 mL LB broth and grown overnight at 37 °C overnight. Cultures were centrifuged for 1 min at 17 000*g* and the pellet resuspended in 200 µL Solution A (50 mM Glucose, 25 mM Tris at pH 8.0, 10 mM EDTA, 0.1 mg/ml RNase A) before adding 200 µL Solution B (0.2 M NaOH, 1% SDS) and inverting to mix. The solution was incubated for 5 min at RT before adding 300 µL Solution C (3 M KAc, 11.5 % v/v Glacial Acetic Acid) and inverting ten times. The lysate was centrifuge for 10 min at 17 000*g* and the supernatant transferred to fresh microcentrifuge tube before adding 700 µL isopropanol. Tubes were incubated for 10 min at RT before centrifugation for 10 min at 17 000*g*. The pellet was air dried before resuspending in 50 µL dH_2O. Restriction digests using 5 µL of each sample were carried out and analysed on an agarose gel to check for the correct size of both vector and DNA insert.

After establishing which colonies contained inserts plasmids were purified. *E.coli* colonies were picked and grown in 5 mL or 100 mL of LB broth containing the appropriate antibiotic(s) at 37 °C on a shaker overnight before purification using the

QIAprep Spin Miniprep and QIAGEN Plasmid Plus Midi Kit according to the manufacturer's instructions. A NanoDrop-1000 spectrophotometer (Thermo Scientific) was used to record the DNA concentration of the plasmids. A double digest of 1 µg of DNA was carried out and analysed on an agarose gel to confirm the presence of the plasmid.

2.1.11 Sequencing

1 µg of DNA was sequenced using the Sanger method using BigDye Terminator v3.1 (Applied Biosystems) according to the manufacturer's instructions. Sequencing primers used are shown in Table 2.3. A sequencing PCR program of 45 cycles of 30 s at 95 °C, 15 s at 50 °C and 4 min at 60 °C was used. The PCR product was precipitated and sent to Source Bioscience for analysis. The DNA chromatogram was returned in *abi* format and was analysed using Finch TV software (Geospiza).

Table 2.3 Sequencing primers used

Primer	Sequence
BGH reverse	TAGAAGGCACAGTCGAGG
CMV forward (pcDNA5)	CGCAAATGGGCGGTAGGCGTG
GFP seq forward	CGAGAAGCGCGATCACATGGTC

2.2 Tissue culture

HEK FRT and Flp cell models generated (Invitrogen) and NSC-34 (Neil Cashman) and NSC-34 FRT and Flp cell models generated (Adrian Higginbottom) were cultured in Dulbecco's modified eagle medium (DMEM, Sigma) containing 10% Tetracycline-free FBS (Biosera) and 50 U/mL penicillin/streptomycin (Lonza) with the appropriate antibiotic selection and induction conditions. Cells were cultured in the presence of Blasticidin S (Calbiochem) and either Zeocin (Invitrogen) or Hygromycin B (Invitrogen) and induced with Tetracycline (Invitrogen). See Table 2.4 for antibiotic selection and Tetracycline induction conditions.

HEK FRT and Flp cell models were trypsinised using 1 mL trypsin (Lonza) per 10 cm^2 and quenched using DMEM containing 10 % Tetracycline-free FBS.

NSC-34 FRT and Flp cell models were removed from the plate by resuspending them in fresh media using a stripette.

Table 2.4 Antibiotic selection and induction conditions for inducible Flp cell models

Conditions	Stock concentration	Working concentration
HEK FRT (before transfection)		
Blasticidin	10 mg/mL	15 µg/mL
Zeocin	100 mg/mL	100 µg/mL
HEK Flp GOI (after transfection)		
Blasticidin	10 mg/mL	15 µg/mL
Hygromycin	50 mg/mL	100µg/mL
HEK Flp Induction		
Tetracycline	10 mg/mL	10 µg/mL
NSC-34 FRT (before transfection)		
Blasticidin	10 mg/mL	2.5 µg/mL
Zeocin	100 mg/mL	20 µg/mL
NSC-34 Flp GOI (after transfection)		
Blasticidin	10 mg/mL	2.5 µg/mL
Hygromycin	50 mg/mL	100 µg/mL
NSC-34 Flp Induction		
Tetracycline	10 mg/mL	1 µg/mL

2.3 Construction of cell models

2.3.1 Construction of HEK inducible cell models

The Flp In T-REx kit (Invitrogen) was used to build HEK Flp inducible cell models. Dr Adrian Higginbottom at SITraN had already constructed HEK Flp TDP-43 WT and TDP-43 Q331K cell lines.

10 cm plates of HEK FRT cells were cotransfected with a 6:4 ratio of pPGKFLPobpA :pcDNA5FRT/TO plasmid DNA using 1 mL Opti-MEM® Reduced Serum Medium (Life Technologies) and 50 µg PEI (poly(ethyleneimine); Sigma). The transfection mixture was added to a 10 cm plate containing 10 mL of culture medium containing Blasticidin, but without Zeocin. 24 hours after transfection the medium was removed and replaced with fresh medium containing Blasticidin. 48 hours after transfection the

10 cm plate was trypsinised. The trypsinised cells were plated into three 10 cm plates in the presence of Blasticidin and Hygromycin (selection medium). The medium was changed every 3 days with selection medium until foci were identified. Hygromycin resistant foci were expanded before verifying that the pcDNA5/FRT/TO construct had integrated into the FRT site by testing each clone for Zeocin sensitivity.

2.3.2 Construction of NSC-34 inducible cell model

The FRT cassette was integrated into NSC-34 cells (Neil Cashman) by Dr Adrian Higginbottom (SITraN).

10 cm plates of NSC-34 FRT cells were cotransfected with a 6:4 ratio of pPGKFLPobpA :pcDNA5FRT/TO plasmid DNA using 3 mL Opti-MEM (Opti-MEM® Reduced Serum Medium (Life Technologies) and 15 µL Lipofectamine® 2000 (Invitrogen). The medium was removed and the transfection reagent added directly onto the cells for a period of 6 h. The plates were rocked every hour to ensure hydration. After 6 h, medium containing Blasticidin was added to the transfection mixture. 48 h after transfection the cells were split onto three 10 cm plates in the presence of Blasticidin and Hygromycin (selection medium). The medium was changed every 2 days with selection medium until foci were identified. Hygromycin resistant foci were expanded before verifying that the pcDNA5/FRT/TO construct had integrated into the FRT site by testing each clone for Zeocin sensitivity.

2.3.3 Verification of FRT recombination

Hygromycin resistant cells were used to seed a 24 well plate. Cells were then grown in medium containing either Zeocin and Blasticidin (pre-transfection FRT medium) or Hygromycin and Blasticidin (selection medium) and the effect on their growth observed. Cells successfully grown in Hygromycin and Blasticidin were expanded and western blots carried out in order to confirm expression of the integrated transgene.

2.4 Standard molecular biology techniques

All primary antibodies used in this study and their sources are shown in Table 2.5.

Table 2.5 Primary antibodies used

Antibody	Purification	Dilution	Source
eIF4a	Rabbit, polyclonal	WB: 1:1000	Cell Signalling 2490
eEF2	Rabbit, polyclonal	WB: 1:10,000	Merck Millipore 2199716
FLAG	Mouse, monoclonal	WB: 1:2000	Sigma F1804 clone M2
GFP	Rabbit, polyclonal	WB: 1:1000 IF: 1:1000	Thermo Fisher Scientific A11122
p62	Rabbit, polyclonal	IF: 1:500	MBL Life Science PM045
Ribosomal protein 10a (RPL10a)	Mouse, monoclonal	WB: 1:500	Santa Cruz JK-16
Ribosomal protein L26 (RPL26)	Rabbit, polyclonal	WB: 1:500	Sigma R0655
Ribosomal protein L29 (RPL29)	Rabbit	WB: 1:2000	Stuart Wilson
TDP-43	Rabbit, polyclonal	WB: 1:1000 IF: 1:1000	Proteintech 12892-1-AP
TDP-43 (human)	Mouse, monoclonal	WB: 1:3000 IF: 1:2000	Abnova H00023435-M01 clone 2E2-D3
TDP-43 (phospho)	Mouse, monoclonal	IF: 1:1000	Cosmo pS409/410
Tubulin (α)	Mouse, monoclonal	WB: 1:10,000 IF: 1:2000	Sigma T9026 clone DM1A
Tubulin (α)	Rabbit, monoclonal	IF: 1:2000	Cell Signalling 11H10

2.4.1 Western blotting

Cells were washed in PBS (Phosphate Buffered Saline; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and lysed in lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 0.5 % Triton X-100, 1 mM EDTA) containing 2 mM phenylmethanesulfonyl fluoride (PMSF) (Sigma) and SIGMAFAST™ Protease Inhibitor Cocktail tablets, EDTA free (Sigma) according to manufacturer's instructions. Protein concentration was determined using Bradford reagent (Biorad) and the OD_{595nm} read using S1200 Diode array spectrophotometer.

SDS-polyacrylamide gels were prepared by pouring 4.5 mL of the desired percentage resolving gel and allowing it to set before pouring a 5 % stacking gel

above it and inserting a 1.0 mm 15-well comb. Gels were made according to the compositions shown in Table 2.6. Gels were made using the Mini-PROTEAN Tetra Cell Casting Stand and clamp system (Biorad).

Lysates were mixed with 4 X SDS gel loading buffer (240 mM Tris-HCl pH 6.8, 277 mM SDS, 0.04 % w/v bromophenol blue, 40 % v/v glycerol, 5 % v/v β-mercaptoethanol) and boiled at 95 °C for 5 min to denature the proteins prior to loading the samples onto the gel. Gels were set up in the Mini-PROTEAN Tetra Vertical Electrophoresis cell (Biorad) and samples loaded next to 2 µL BLUeye Prestained Protein Ladder (Geneflow). The gels were run at 150 V for 1.5 h or until the dye front had reached the bottom of the gel. Gels were transferred to reinforced nitrocellulose membranes (GE Healthcare Life Sciences) using a semi-dry transfer apparatus (Biometra) and transfer buffer (47.9 mM Tris, 38.6 mM glycine, 1.38 mM SDS, 20 % v/v Methanol) soaked Whatman paper. The membranes were blocked in 5 % w/v milk (Marvel) in Tris Buffered Saline with Tween-20 (TBST) (137 mM NaCl, 20 mM Tris, 0.2 % v/v Tween-20, pH 7.6) for 1 h on a roller. The membranes were incubated with primary antibody in 5 % milk in TBST for 1 h on a roller. The membranes were then washed five times in TBST at RT for 25 min. The membranes were then incubated in secondary antibody conjugated to horseradish peroxidase (HRP). Anti-mouse-HRP and anti-rabbit HRP (Promega) were used at a dilution of 1:10,000 in 5% w/v milk in TBST at RT for 1 h on a roller. The membranes were then washed five times in TBST at RT for 25 min. The membrane was then incubated with enhanced chemiluminescence (ECL) solution 1 (100 mM Tris-HCl pH 8.5, 2.5 mM Luminol, 400 µM p-Coumaric acid in dH_2O) and solution 2 (100 mM Tris-HCl pH 8.5, 6.1 µL H_2O_2 in dH_2O) in a 1:1 ratio for 1 min before imaging using a G:BOX (Syngene).

Table 2.6 Composition of stacking and resolving SDS polyacrylamide gels

	5 % stacking	12 % resolving	15 % resolving
dH_2O			
30 % w/v bis-acrylamide (Geneflow)	1.7 mL	4 mL	5 mL
Resolving buffer (1.5 M Trizma®, 13.9 mM SDS, pH 8.8, filtered)	N/A	2.5 mL	2.5 mL
Stacking buffer (0.5 M Trizma®, 13.9 mM SDS, pH 6.8, filtered)	2.5 mL	N/A	N/A
10 % Ammonium persulfate (APS)	50 µL	50 µL	50 µL
N, N, N', N'-Tetramethyl-ethylenediamine (TEMED)	20 µL	20 µL	10 µL

2.4.2 Immunofluorescence

Coverslips were coated in 0.5 mg/mL gelatine (BDH Biochemical) in dH_2O for 30 min at 37 °C or overnight at 4 °C for the immunofluorescence of NSC-34 Flp cell models.

For cells induced with Tetracycline for over 48 h, the media was changed refreshing the Tetracycline every 48 h and the cells were grown in 10 cm dishes until 5 days prior to experiment date, when they were placed onto coverslips.

Cells grown on coverslips were washed in PBS and fixed with 4 % paraformaldehyde (PFA) containing 0.2 % Triton-X for permeablisation for 20-30 min. Cells were blocked in 2 % BSA (Sigma) in PBS for 20 min at room temperature. Cells were incubated in primary antibody diluted with 2 % w/v bovine serum albumin (BSA) in PBS for 1 h at RT. The cells were then washed with PBS and incubated with secondary antibody diluted with 2 % w/v BSA in PBS for 30 min at RT. The following fluorophore-labelled secondary antibodies were used at a concentration of 1:10, 000: Alexa Fluor 488 nm (green) anti-rabbit (Abcam, ab150077) and anti-mouse (Abcam, ab150113), Alexa Fluor 594 nm (red) anti-rabbit (Abcam, ab150080) and anti-mouse (Abcam, ab150116). The cells were then incubated with Hoechst bizBenzimide H (Sigma 33258) (blue) diluted with 2 % w/v BSA in PBS for 10 min at RT to visualise nuclei. The coverslips were then mounted onto glass slides using Dako Fluorescent

mounting medium (Dako). The slides were observed using fluorescence microscopy (Nikon).

2.4.3 RNA Fluorescent In Situ Hybridiation (FISH)

Coverslips were coated in 0.5 mg/mL gelatine (BDH Biochemical) in dH_2O for 30 min at 37 °C or overnight at 4 °C for the immunofluorescence of NSC-34 Flp cell models.

For cells induced with Tetracycline for over 48 h, the media was changed refreshing the Tetracycline every 48 h and the cells were grown in 10 cm dishes until 5 days prior to experiment date, when they were placed onto coverslips.

Cells were treated by adding 5 µg/mL Actinomycin D directly to the media for 1.5 h prior to fixing cells. Cells were washed 0.1 % Diethylpyrocarbonate (DEPC) (Applichem) PBS to inhibit RNase activity. Cells were fixed with 4 % PFA containing 0.2 % Triton-X for permeablisation for 20-30 min. Cells were then washed with DEPC PBS. Coverslips were incubated in hybridisation buffer (20 % formamide, 2 X sodium saline citrate (SSC), 10 % dextran sulphate, 1 % BSA) containing 50 µL/mL single stranded DNA and 1 µg/µL Cy3 Oligo(dT) for 2-3 h at 37 °C in a sealed container containing hybridisation buffer soaked filter paper. Coverslips were washed with DEPC PBS before immunofluorescence staining and mounting as described in section 2.4.2.

2.4.4 Growth curve

For HEK Flp cell models 1 x 10^6 cells were plated in 10 cm dishes in the presence or absence of Tetracycline. At three day intervals the cells were trypsinised and resuspended in 5 mL of media. The cells were then counted using a haemocytometer and the total number of cells in each plate calculated. After counting, 1 x 10^6 cells were returned to the 10 cm plate and given fresh media

containing Tetracycline where required. This was repeated for the duration of the growth curve.

For NSC-34 Flp cell models 2 x 10^6 cells were plated in 10 cm dishes in the presence or absence of Tetracycline. At three-day intervals the media was removed and the cells resuspended in 5 mL of media. The cells were then counted using a haemocytometer and the total number of cells in each plate calculated. After counting, 2 x 10^6 cells were returned to the 10 cm plate and given fresh media containing Tetracycline where required. This was repeated for the duration of the growth curve.

2.4.5 Cell proliferation/metabolic assay MTT

Cells were grown in a 24 well tissue culture plate (Greiner) for 24 h before induction with Tetracycline. 4-6 wells per cell line per independent experiment were used. For cells induced with Tetracycline for over 48 h, the media was changed refreshing the Tetracycline every 48 h and the cells were grown in 10 cm dishes until 5 days prior to experiment date, when they were placed into 24 well plates. Thiazolyl Blue Tetrazolium Bromide (MTT) was added to each well to a final concentration of 0.5 mg/mL and the plate was returned to the incubator for 30 min for HEK cells and 1 h for NSC-34 cells. The cells were then lysed in a 1:1 volume of MTT lysis buffer (20 % w/v SDS, 50 % v/v DMF, pH 7.4) and left for 1 h at RT on an orbital shaker. The absorbance at 595 nm of each well was recorded using a PHERAstar FS plate reader (BMG labtech).

2.5 RNA preparation

All solutions for RNA work were made up with water treated with 0.1 % DEPC and PBS for RNA work was also treated with 0.1 % DEPC to inactivate any RNase enzymes in water. Water and PBS were treated with 0.1 % DEPC and then autoclaved before use to inactivate the DEPC.

2.5.1 RNA extractions for sequencing

Cells were harvested from 10 cm plates. Cells were grown for 24 hours before induction with Tetracycline. Cells were lysed 24 h (HEK Flp *GFP* and *M1V*) or 48 h (HEK Flp Sham, TDP-43 WT and TDP-43 Q331K) post Tetracycline induction when an approximate confluency of 70-80 % had been achieved.

2.5.1.1 Whole cell RNA extraction

Total fractions were collected directly in 400 µL of Reporter lysis buffer (Promega) containing 0.16 U/µL Ribosafe RNase inhibitors (Bioline), 2 mM PMSF and SIGMAFAST™ Protease Inhibitor Cocktail tablets, EDTA free according to manufacturer's instructions and lysed for 10 min on ice before centrifugation at 17 000*g*, 5 min, 4 °C. The supernatant was then added to Trizol (Thermo Fischer) and RNA was extracted using Direct Zol RNA Miniprep Plus (Zymo Research).

2.5.1.2 Cytoplasmic RNA extraction

HEK Flp cell models were typsinised and quenched before centrifugation at 400*g* for 5 min to give a cell pellet for cytoplasmic fractionation. Cell pellets were quickly washed with hypotonic lysis buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT). Cell pellets were then lysed in 400 µL hypotonic lysis buffer containing 0.16 U/µL Ribosafe RNase inhibitors (Bioline), 2 mM PMSF (Sigma) and SIGMAFAST™ Protease Inhibitor Cocktail tablets, EDTA free (Sigma) according to manufacturer's instructions. Cells were lysed gently using a cut P1000 tip ensuring that no physical force was exerted on the cell pellet. The lysate then underwent differential centrifugation (3 min at 1500*g*, 4 °C then 8 min at 3500*g*, 4 °C and then 1 min at 17 000*g*, 4 °C) transferring the supernatant to a fresh tube after each centrifugation. The resulting supernatant is the cytoplasmic fraction. The supernatant was then added to Trizol (Thermo Fischer) and RNA was extracted using Direct Zol RNA Miniprep Plus (Zymo Research). Anti-SSRP1 western blotting was done on cytoplasmic lysates to confirm the absence of nuclear leakage.

2.5.2 Trizol RNA extraction

RNA was extracted by adding Trizol (Thermo Fisher) directly to the cells and scraping cells from tissue culture dish or by adding three volumes Trizol to one volume of cell lysate. Cells were left at RT for 10 min before adding chloroform (one-fifth of volume of lysate and Trizol). Samples were shaken vigorously for 20 sec before being left at RT for 5 min. Extractions then underwent centrifugation for 10 min at 11 600*g*, 4 °C. The aqueous layer was taken and added to 1 µL 5 mg/mL Glycogen (Ambion) and 300 mM NaAc before adding an equal volume of isopropanol and storing at - 20 °C overnight. The precipitating RNA then underwent centrifugation for 20 min at 17 000*g*, 4 °C before carefully removing the supernatant from the RNA pellet. The RNA pellet was dried at RT for 20 min. This method of RNA extraction was used for the extraction of total RNA for qRT-PCR.

2.5.3 DNase treatment

RNA pellets were fully resuspended in DEPC H_2O before DNase I treatment (Roche) incubating RNA at 37 °C for 30 min before heat inactivation at 70 °C for 10 min.

2.5.4 Reverse transcription

2 µg of RNA was used for cDNA synthesis using poly(dN)$_6$ random priming described by the manufacturer (Bioscript, Bioline). Samples were incubated at 25 °C for 10 min followed by 42 °C for 60 min before terminating the reaction by incubating at 85 °C for 10 min and chilling at 10 °C before storage.

2.5.5 mRNA purification

NEB Next® Poly(A)$^+$ mRNA Magnetic Isolation Module (New England BioLabs Inc.) was used to purify mRNA based upon on the coupling of oligo d(T)$_{25}$ to 1 µm paramagnetic beads which were then used as the solid support for the direct binding of poly(A)$^+$ RNA. Beads were separated from the supernatant using a magnetic rack.

15 µL of Oligo $d(T)_{25}$ beads were placed in 0.2 mL PCR tube and washed twice with 100 µL of 2 X RNA binding buffer to remove the supernatant. A 50 µL volume of 2 X RNA binding buffer was then added to the beads. Total RNA was diluted with DEPC water to a final volume of 50 µL and added to the magnetic beads in RNA binding buffer. Samples were incubated at 65 °C for 5 min and placed on ice for 2 min to denature RNA and facilitate the binding of $poly(A)^+$ RNA to the beads. Samples were incubated at RT for 5 min to allow RNA to bind to the beads. The tubes were then placed on the magnetic rack for 2 min to separate $poly(A)^+$ RNA bound to the beads from the solution. The supernatant was taken and kept. At this point the protocol was optimised and an additional binding step carried out in order to increase the yield of ribosomal $poly(A)^+$ RNA. The beads were washed twice with 200 µL wash buffer, pipetting the volume up and down 6 times to ensure mixed thoroughly and placing on the rack for 2 min to ensure proper separation of unbound RNA. The beads were then stored on ice. The saved supernatant (from the initial binding) was then incubated at 65 °C for 5 min and placing on ice for 2 min. The binding of $poly(A)^+$ RNA was then repeated as above with the beads stored on ice. The beads were then washed twice with 200 µL wash buffer, pipetting the volume up and down 6 times to ensure mixed thoroughly and placing on the rack for 2 min to ensure proper separation of unbound RNA. After ensuring total removal of wash buffer, 50 µL elution buffer was added to the beads and mixed well by pipetting. Tubes were placed at 80 °C for 2 min then placed at RT immediately to elute the $poly(A)^+$ RNA from the beads. 50 µL of 2 X RNA binding buffer was added to each sample pipetting the volume up and down 6 times to ensure mixed thoroughly to allow RNA to bind to the beads again. Samples were incubated at RT for 5 min with agitation every few min during the incubation. The tubes were allowed to stand the magnetic rack for 2 min before removing the supernatant and washing the beads twice with 200 µL wash buffer, pipetting the volume up and down 6 times to ensure mixed thoroughly and placing on the rack for 2 min to ensure proper separation of unbound RNA. The

mRNA was eluted for the final time by adding 20 µL of elution buffer to the beads and incubating at 80 °C for 2 min before immediately putting the samples onto the magnetic rack for 2 min. Purified mRNA was transferred to a clean nuclease-free PCR tube before the yield and size distribution was assessed using a RNA Pico Chip (Agilent Technlogies, Inc.).

2.5.6 RNA yield and quality assessment

RNA yield and purity was determined using the NanoDrop™ 1000 (Thermo Fisher). The absorbance ratio at 260 nm and 280 nm was used as a measure of purity in nucleic acid extractions. For RNA a ratio of approximately 2 is considered 'pure' for RNA. The absorbance ratio at 260 nm and 230 nm is used as a secondary measure of nucleic acid purity and is usually higher between 2.0 and 2.2.

Yield and size distribution of RNA and mRNA was determined using Nanochip or Picochip respectively and an Agilent 2100 Bioanalyser (Agilent Technologies, Inc.). This assesses the size distribution of the 18S and 28S rRNA peaks, producing an electropherogram and assigning a RIN number where 1 is the most degraded profile and 10 is the most intact.

2.5.7 Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Extracted RNA was treated with DNAse I (Roche) at 37 °C and heat inactivated at 70 °C for 5 min. cDNA was obtained after reverse transcription polymerase chain reactions (RT-PCR) using approximately 2 µg RNA with random primers and the Bioscript Reverse Transcription kit (Bioline), according to manufacturer's instructions. Resulting cDNA was diluted before use - ribosomal RNA was diluted 1:1000, all other RNA diluted 1:3 before qPCR.

Primer standard curves were generated by performing a 4-fold serial dilution of HEK and/or NSC-34 cDNA using DEPC H_2O. Reaction mixtures were prepared using 1 µL of 5 µM primer mix, 5 µL 2 X Brilliant III Ultra Fast SYBR master mix (Agilent), and 3 µL nuclease-free water. 9 µL of reaction mixture and 1 µL of diluted cDNA added to each well of a 96 well PCR plate, which was then sealed using plastic caps. Target DNA amplification was performed by thermal cycling: initial denaturation at 95 °C for 10 min, followed by 45 cycles of 95 °C for 30 s and 60 °C for 30 s, then 72 °C for 1 min in a PCR detection system (Stratagene Mx3000p).

Cycle threshold (Ct) values were plotted on the y-axis against cDNA concentration on the x-axis. Using the equation for the line [y = m*log(x) + b] (where m is the slope of the line) the efficiency can be calculated using the equation PCR efficiency = 10(-1/slope) – 1. The efficiency corresponds to the proportion of template molecules that are doubled every cycle. A 100 % efficient reaction will have a slope of -3.322. The R squared (R^2) value is an indicator of the quality of the fit of the standard curve to the standard data points plotted. The value will always be between 0 and 1. The closer the R^2 value is to 1 the better the fit of the line. qRT-PCR primer sequences, the R^2 value and efficiencies are shown in Table 2.7. Multiple *TARDBP* primers were tested and those shown in Table 2.7 were found to be best, despite the efficiency being greater than 100 % when tested with both mouse and human cDNA.

Reaction mixtures were prepared using 1 µL of 5 µM primer mix, 5 µL 2 X Brilliant III Ultra Fast SYBR master mix (Agilent), and 3 µL nuclease-free water. 9 µL of reaction mixture and 1 µL of diluted cDNA added to each well of a 96 well PCR plate, which was then sealed using plastic caps. Target DNA amplification was performed by thermal cycling: initial denaturation at 95 °C for 10 min, followed by 45 cycles of 95 °C for 30 s and 60 °C for 30 s, then 72 °C for 1 min in a PCR detection system

(Stratagene Mx3000p). Relative quantification was determined using the ΔCt method and normalized to U1 snRNA/β-actin mRNA/18S rRNA.

Table 2.7 qRT-PCT primer sequences and optimisation

Primers	Primer sequence 5' > 3'	Tm	% GC	Human R^2	Slope	Efficiency %	Mouse R^2	Slope	Efficiency %
GFP	F' AAGCAGCACGACTTCTTCAAG R' TTCAGCTCGATGCGGTTCA	60.27 59.71	47.62 52.63	0.999	-3.407	96.6	Not tested		
Human TARDBP	F' TGGGATGAACTTTGGTGCGT R' TTTGGCTCCCTCTGCATG	60.18 59.89	50.00 50.00	0.986	-2.549	146.8	0.976	-2.257	177.4
U1 snRNA	F' CCATGATCACGAAGGTGGTT R' ATGCAGTCGAGTTTCCCACA	59.21 60.10	50.00 50.00	0.999	-3.23	104	0.996	-3.189	105.9
18S	F' CGGACATCTAAGGGCATCAC R' GTGGAGCGATTTGTCTGGTT	59.00 58.60	55.00 50.00	0.999	-3.223	104.3	0.998	-3.281	101.7
β-actin	F' TCCCCCAACTTGAGATGTATGAAG R' AACTGGTCTCAAGTCAGTGTACAGG	62.10 59.60	46.00 48.00	0.997	-3.457	94.7	0.992	-3.374	97.9

2.6 Translating ribosome affinity purification (TRAP)

The protocol was followed from (Kulicke et al., 2014) using a FLAG-tagged RPL10a instead of the GFP-tagged RPL10a described.

Anti-FLAG M2 affinity gel beads were incubated in TRAP lysis buffer (20 mM HEPES KOH pH 7.4, 150 mM KCl, 10 mM $MgCl_2$, 1 % v/v NP-40 in DEPC H_2O) containing 1 % BSA and 5 µL 10 mg/mL ssDNA at 4 °C on a wheel at 10 rpm overnight.

For each TRAP experiment two T175 flasks of cells were grown for 24 h before transfection with 30 µg of plasmid DNA and 105 µg PEI in 1 mL Opti-MEM Reduced Serum Medium (Life Technologies) per T175 flask. Approximately 6-8 h post transfection the cells were induced with Tetracycline. 24 h post transfection and approximately 12 h post Tetracycline induction the cells were incubated with 100 µg/mL cycloheximide (CHX) for 15 min on ice. The cells were then washed with DEPC PBS containing 100 µg/mL CHX. Each flask was lysed in 1 mL of ice cold TRAP lysis buffer (20 mM HEPES KOH pH 7.4, 150 mM KCl, 10 mM $MgCl_2$, 1 % v/v NP-40 in DEPC H_2O) containing 5 µL 10 mg/mL ssDNA, 100 µg/mL CHX, 0.16 U/µL Ribosafe RNase inhibitors (Bioline), 0.5 mM DTT, 2 mM PMSF and SIGMAFAST™ Protease Inhibitor Cocktail tablets, EDTA free according to manufacturer's instructions. Cells were incubated for 10 min on ice before removing cells using a cell scraper and transferring to ice-cold tubes. The lysate was drawn through a 25G needle five times before differential centrifugation (5 min at 2400*g* rpm then 5 min at 9600*g*, 4 °C) transferring the supernatant to a fresh tube after each centrifugation. The protein concentration was measured using Bradford reagent (Biorad) and the OD_{595nm} read using S1200 Diode array spectrophotometer. The samples were diluted to a maximum protein concentration of 2 mg/mL with the TRAP lysis buffer. The

same volume of lysate was loaded on to the pre-blocked FLAG beads and left to bind at 4 °C on a wheel at 10 rpm overnight.

The beads were collected by centrifugation for 1 min at 400*g*, 4 °C and the unbound supernatant stored on ice for western blot. The beads were washed four times in 900 µL high salt buffer (20 mM HEPES KOH pH 7.4, 350 mM KCl, 10 mM $MgCl_2$, 1 % v/v NP-40 in DEPC H_2O) containing 0.5 mM DTT and 100 µg/mL CHX. The high salt buffer was removed and the beads stored at room temperature. The beads were resuspended in 125 µL of EZ-RNA Total extraction buffer (Biological Industries) and incubated at room temperature for 1 h on a wheel at 10 rpm. The supernatant was removed from the beads and 100 µL taken for RNA extraction using the EZ-RNA Total extraction kit (Biological Industries) according to manufacturer's instructions. The remaining sample was used for western blot analysis.

2.7 RNA Immunoprecipitation (RIP)

40 µL Protein G Sepharose™ 4 Fast Flow beads (GE Healthcare) or Anti-FLAG M2 affinity gel beads were washed with RIP lysis buffer (20 mM Tris, 140 mM KCl, 5 mM $MgCl_2$, 0.5 M DTT, pH 8.0). Beads were pelleted by centrifugation for 1 min at 400 *g*. Protein G Sepharose beads were then blocked overnight at 4 °C on a wheel at 10 rpm with RIP lysis buffer containing 5 µg antibody, 1 % Triton X-100, 1 % BSA and 5 µL/mL ssDNA. Anti-FLAG M2 affinity gel beads were then blocked overnight at 4 °C on a wheel at 10 rpm with RIP lysis buffer containing 1 % Triton X-100, 1 % BSA and 5 µL/mL ssDNA.

Cells were plated for 24 hours before induction with Tetracyline and RIP was carried out 24 h post-induction. Cells were treated with 100 µg/mL CHX for 1 h prior to lysis. Four 10 cm plates of cells were used for each RIP. 10 cm plates of cells were washed in DEPC PBS and lysed in 500 µL RIP lysis buffer containing 0.16 U/µL

Ribosafe RNase inhibitors (Bioline), 2 mM PMSF and SIGMAFAST™ Protease Inhibitor Cocktail tablets, EDTA free according to manufacturer's instructions and 100 µg/mL CHX. The lysates were combined and drawn through a 25G needle five times and left for 10 min on ice before centrifugation for 10 min at 17 000*g*, 4 °C. The supernatant was removed and 250 µL taken for RNA extraction using Trizol and 50 µL for western blot (input). The resulting supernatant was taken and the protein concentration determined by Bradford assay. No more than 2 mg/mL protein in equal volumes were loaded onto the beads and incubated at 4 °C on a wheel at 10 rpm for 3 h. The beads were washed five times with 500 µL RIP lysis buffer before removing all the buffer using a gel loading tip. 125 µL RIP elution buffer (20 mM Tris, 140 mM KCl, 5 mM $MgCl_2$, 0.5 M DTT, 10 mM EDTA, 10 mM DTT, 1 % SDS pH 8.0) containing 0.16 U/µL Ribosafe RNase inhibitors (Bioline) was added to the beads and left to elute for 2 h on a wheel at 10 rpm at RT. The eluate was then divided into 100 µL for Trizol RNA extraction and 25 µL for western blot. Equal amounts of input and eluted were loaded onto SDS PAGE gels for western blot analysis.

2.8 Purification of translating ribosomes using the GRASPS™ methodology developed in this study

Cells were harvested from 10 cm plates. Cells were grown for 24 h before induction with Tetracycline. Cells were lysed 24 h (HEK Flp *GFP* and *M1V*) or 48 h (HEK Flp Sham, TDP-43 WT and TDP-43 Q331K) post Tetracycline induction when an approximate confluency of 70-80 % had been achieved.

All GRASPS buffers were treated with 0.1 % DEPC. Cells were removed from the plate in culture media using a cell scraper before centrifugation for 5 min at 400*g* to pellet cells. Cell pellets were quickly washed with GRASPS buffer A1 pre-optimisation (250 mM sucrose, 250 mM KCl, 5mM $MgCl_2$, 50 mM Tris-HCl pH 7.4) or GRASPS buffer A post-optimisation (250 mM sucrose, 5 mM KCl, 50 mM Tris-HCl

pH 7.4). The cell pellet size was measured and resuspended sequentially in three volumes equivalent to the size of the pellet of of GRASPS buffer A1 pre-optimisation (250 mM sucrose, 250 mM KCl, 5mM $MgCl_2$, 50 mM Tris-HCl pH 7.4) or GRASPS buffer A post-optimisation (250 mM sucrose, 5 mM KCl, 50 mM Tris-HCl pH 7.4) containing 0.16 U/µL Ribosafe RNase inhibitor (Bioline), 2 mM PMSF (Phenylmethanesulfonyl fluoride) and SIGMAFAST Protease Inhibitor Cocktail tablets, EDTA free according to manufacturer's instructions. NP-40 was added to a final concentration of 0.7 % v/v from a cold 10 % NP-40 in DEPC dH_2O stock solution. The lysate was left on ice for 5 min then pipetted up and down twice before leaving for a further 5-10 min on ice. The lysate was transferred to one well of a 6 well plated and UV irradiated at 0.3 J/cm^2. The nuclear pellet was discarded after centrifugation for 10 min at 750*g*, 4 °C and the supernatant carefully transferred to a prechilled tube before centrifugation for 10 min at 12 500*g*, 4 °C to pellet the mitochondrial fraction. The supernatant, the post-mitochondrial fraction, was transferred to a fresh prechilled tube and 4 M KCl was added to give a final concentration of 0.5 M KCl. The post-mitochondrial fraction was then stored on ice.

1 mL sucrose cushion solution (1 M sucrose, 5 mM $MgCl_2$, 50 mM Tris-HCl pH 7.4) was added to a clean, cold TLA100 centrifuge tube. The 0.5 M KCl post-mitochondrial fraction was made up to a volume of 1 mL using GRASPS buffer B (250 mM sucrose, 500 mM KCl, 50 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$). 900 µL of the 0.5 M post-mitochondrial fraction was loaded onto the sucrose cushion carefully before centrifugation for 2 h at 250 000*g*, 4 °C.

Post-centrifugation the supernatant was removed carefully. The translucent ribosomal pellet was washed carefully by adding 150 µL of cold dH_2O and removing immediately. The pellet was then resuspended in 250 µL of ribosomal resuspension buffer on ice (50 mM HEPES pH 7.9, 150 mM NaCl, 1 mM DTT, 1 mM EDTA) and

transferred to a clean tube. Proteinase K was added to a final concentration of 100 µg/mL and 0.16 U/µL Ribosafe RNase inhibitor (Bioline) was added. The mixture was then pulse vortexed and incubated at 37 °C for 30 min. 10 mM EDTA and 50 mM NaAc were added and the mixture vortexed. 750 µL of Trizol-LS (Invitrogen) was added before RNA was extracted using Direct Zol RNA Miniprep Plus (Zymo Research).

2.9 Next Generation RNA sequencing and Bioinformatics analysis

Total RNA extracted from whole-cell or cytoplasmic fractions and poly-adenylated RNA isolated from GRASPS-purified ribosomes were sequenced on *Illumina HiSeq 2500* by the Centre for Genomic Research at the University of Liverpool.

Briefly, RNASeq libraries were prepared using the ScriptSeq™ Complete Gold Kit protocol to deplete the total RNA for rRNA (for whole-cell and cytoplasmic transcriptomes) and prepare strand-specific libraries from rRNA-depleted RNA samples. For GRASPS samples, the same library kit was used, but no rRNA depletion was performed. Indexed libraries were subjected to paired-end sequencing (approximately 2x 100 bp) multiplexed as three libraries per lane of the Illumina HiSeq platform to generate data of in excess of 120 millions clusters/lane. Trimming of sequencing reads and initial quality check prior to the generation of fastq files were performed.

The raw Fastq files are trimmed for the presence of Illumina adapter sequences using Cutadapt version 1.2.1 (Martin, 2011). The option -O 3 was used, so the 3' end of any reads, which match the adapter sequence for 3 bp or more are trimmed. The reads are further trimmed using Sickle version 1.200 (https://github.com/najoshi/sickle) with a minimum window quality score of 20. Reads shorter than 10 bp after trimming were removed. If only one of a read pair passed

this filter, it is included in the R0 file. The output files from Cutadapt and Sickle are available on link on attached CD-ROM (See C1. Cutadapt and Sickle output files). Statistics were generated using fastq-stats from EAUtils (https://expressionanalysis.github.io/ea-utils/).

Dr Marta Milo, at the University of Sheffield, carried out the bioinformatics analysis. An overview of the analysis is outlined in Figure 2.2.

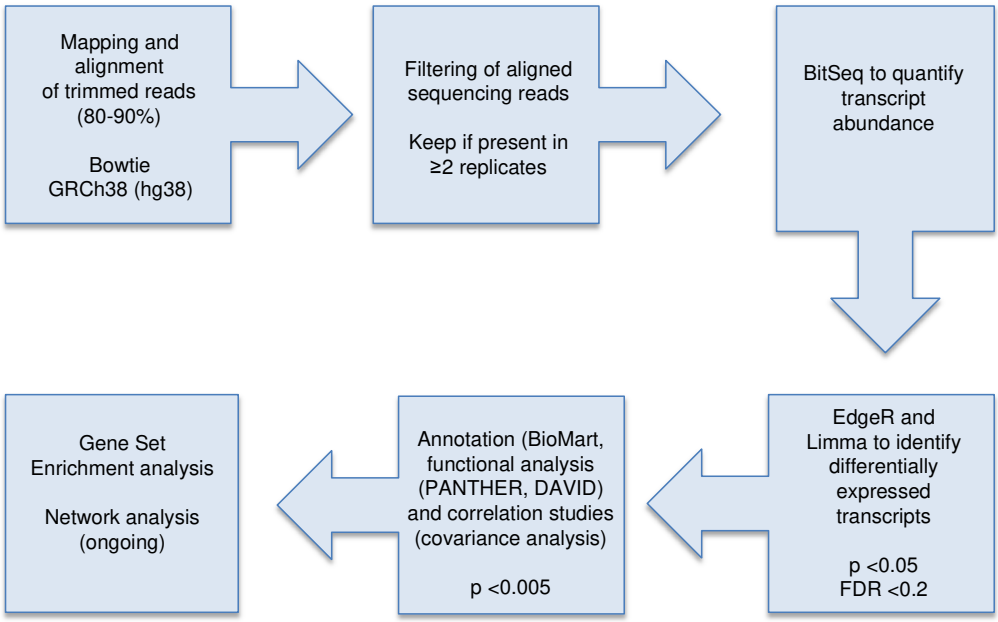


Figure 2.2 Outline of the Bioinformatics analysis performed on RNA seq data sets
Trimmed reads were mapped to the genome using Bowtie software, before being filtered. Bitseq was used to quantify transcript abundance before identification of differentially expressed transcripts. Annotation, function analysis and correlation studies were then performed. Gene enrichment analysis was then carried out. TDP-43 Q331K transcriptome and translatome changes were normalised to the Sham control.

Reads per sample were mapped to the human genome (~60x coverage) using Bowtie (Langmead et al., 2009) and transcripts quantified using Bitseq (Glaus et al., 2012) and Limma (Ritchie et al., 2015). BioMart was used to convert transcripts into Ensembl transcript identities (Kasprzyk, 2011).

Enrichment analysis was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (D. W. Huang et al., 2008; 2009) and the

PANTHER overrepresentation test (Mi et al., 2013). TDP-43 Q331K transcriptome and translatome changes were normalised to the isogenic induced Sham control line.

The differentially expressed gene (DEG) lists and lists generated from the functional enrichment analysis using DAVID, PANTHER and targets of TDP-43 are in an excel spread sheet on the CD-ROM attached (see C2. DEG lists).

2.10 Statistical analysis methods

T-tests were used to assess whether the means of two groups were significantly different from each other. The means of three or more samples were compared with one-way analysis of variance (One-way ANOVA) with Tukey's multiple comparison tests. The influence of two different indendepent variables on the means of three or more samples were compared using two-way analysis of variance (Two-way ANOVA) with Tukey's multiple comparison tests. All statistical analyses were carried out using GraphPad Prism 6.0.

3. Development of a novel high throughput methodology to identify selectively enriched translating mRNA: Genome-wide RNA analysis of stalled protein synthesis (GRASPS)

3.1 Introduction

Dysregulation of gene expression is common feature in ALS with alterations to the transcriptome in cell models, animal models and in CNS tissue from MND cases (F. Tanaka et al., 2011; Cooper-Knock, Hewitt, et al., 2012; Heath et al., 2013). However, these studies did not (i) allow discrimination of pathophysiological alterations from those due to downstream dysregulations and (ii) did not investigate whether altered mRNA production is reflected at the protein level. Recent studies have also identified that transcript levels are not necessarily linked to the levels of corresponding proteins, due to cell compensatory mechanisms (Domínguez-Sánchez et al., 2011).

Therefore identification of mRNA actively being translated in the ribosomes should reflect protein expression changes and the directionality of altered biological processes. Current translatome technologies, such as ribosome affinity purification, allow the study of mRNA that is associated with the ribosome. However, no study has specifically tested whether or not the mRNA isolated using these methods is actively undergoing translation. This is of consideration for this work, as TDP-43 has been shown to interact with stalled ribosomes (Higashi et al., 2013) and therefore could potentially contaminate ribosomes with indirectly bound mRNA, which is not being translated but being transported by TDP-43.

A novel high throughput translatome technology named GRASPS (Genome-wide RNA Analysis of Stalled Protein Synthesis) was developed to try and address some of the limitations of affinity-based methods. The GRASPS strategy aimed to reproducibly purify ribosomes without the need for chemical inihibitors or density/size/affinity purification steps. The GRASPS method involves the biochemical precipitation of UV-crosslinked ribosome-bound RNAs by ultracentrifugation. The steps and timeline for the TRAP and GRASPS technologies are summarised in Figure 3.1.

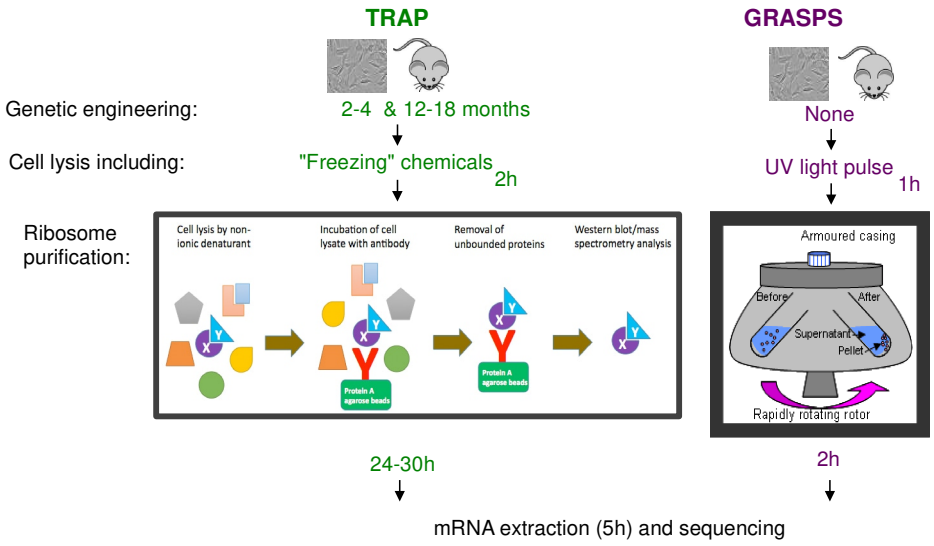


Figure 3.1 Summary of the steps and timeline for the TRAP and GRASPS technologies. The genetic engineering, cell lysis, ribosome purification and subsequent analysis are outlined alongside the timelines for the steps involved in each technlology.

The aim of the work described in this chapter is to explore current translatome technologies in order to identify a method that would allow the identification of mRNA actively undergoing protein synthesis. In order to do this, a cell model would need to be built which would allow the distinction between mRNAs that are being actively translated into protein and mRNAs that are not being translated into protein. This cell model would then be used to explore current translatome technologies in order to

develop a reliable method to identify mRNAs being actively translated into protein at the ribosome.

3.2 Results

3.2.1 Generating inducible GFP reporter cell models

In order to test methodologies used to specifically isolate mRNAs that are undergoing translation, two GFP reporter cell models expressing either translated or non-AUG translated transcripts were built. Site-directed mutagenesis of the pcDNA5 FRT/TO GFP AH plasmid was used to generate constructs to build the cell models. ATG triplet sequences were first mutated, in all frames, in over two thirds of the GFP nucleotide sequence (465 nucleotides). These included a silent *D37D* mutation and in-frame methionine to valine mutations *M79V* and *M89V*. This construct is functional and will be referred to as *GFP*. It contains the original start codon and Kozak sequence (Kozak, 1987), but mutations of the next three ATG triplet sequences. The mutations in the ATG triplet sequence would ensure that truncated versions of GFP could not be produced from a start codon later in the nucleotide sequence. The start codon of the pcDNA5 FRT/TO/*GFP* construct was mutated in order to generate a second construct. This *GFP M1V* construct does not encode a protein and will be referred to as *M1V*. This transcript lacks all potential start codons in all frames within 655 out of 739 nucleotides of the GFP coding sequence (*M1V, D37D, M79V, M89V*). The pcDNA5 FRT/TO *GFP* and *M1V* constructs generated are shown in Figure 3.2. The constructs were sequenced and an example of the sequencing trace from the *M1V* plasmid is shown in Appendix 1.

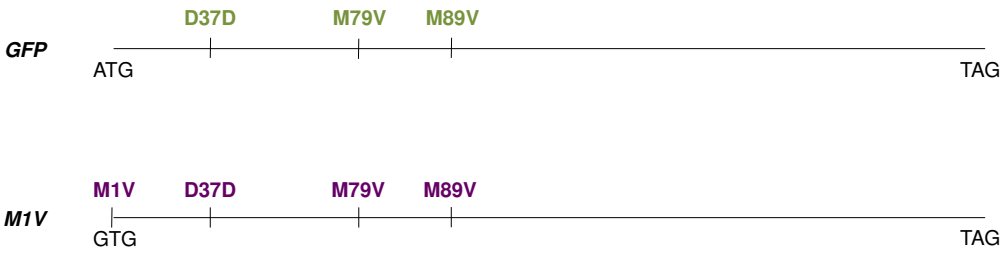


Figure 3.2 pcDNA5 FRT/TO GFP plasmids constructed for GFP reporter cell model

Site-directed mutagenesis of pcDNA5 FRT/TO/GFP AH was used to insert a silent *D37D* mutation and in-frame methionine to valine mutations *M79V* and *M89V* in the *GFP* plasmid. The *GFP* plasmid was then used to generate the *M1V* plasmid by further mutating the start codon *M1V*. Both constructs contain the Kozak sequence.

The cell model generated using the *GFP* construct (HEK Flp *GFP*) would be used as an inducible control for a translating mRNA reporter as it would produce GFP and should be enriched in the mRNA fraction isolated from actively translating ribosomes. The second model generated using the *M1V* construct will have no start codon and therefore will not produce GFP. The unique difference between the two isogenic cell models is the presence or absence of a start codon. The Kozak sequence has been preserved in both reporter transcripts. The *M1V* cell line (HEK Flp *M1V*) would be used to generate mRNA, which are not expected to be conventionally translated by the ribosome since they lack AUG start codons. This model could therefore be used as a test to whether AUG-translating mRNAs could selectively be isolated.

The expression of GFP from the plasmids was tested by transfection into HEK cells (Figure 3.3). Transfection of the pcDNA5 FRT/TO/*GFP* plasmid produces GFP, whereas no GFP is observed after transfection of the pcDNA5 FRT/TO/*M1V* plasmid.

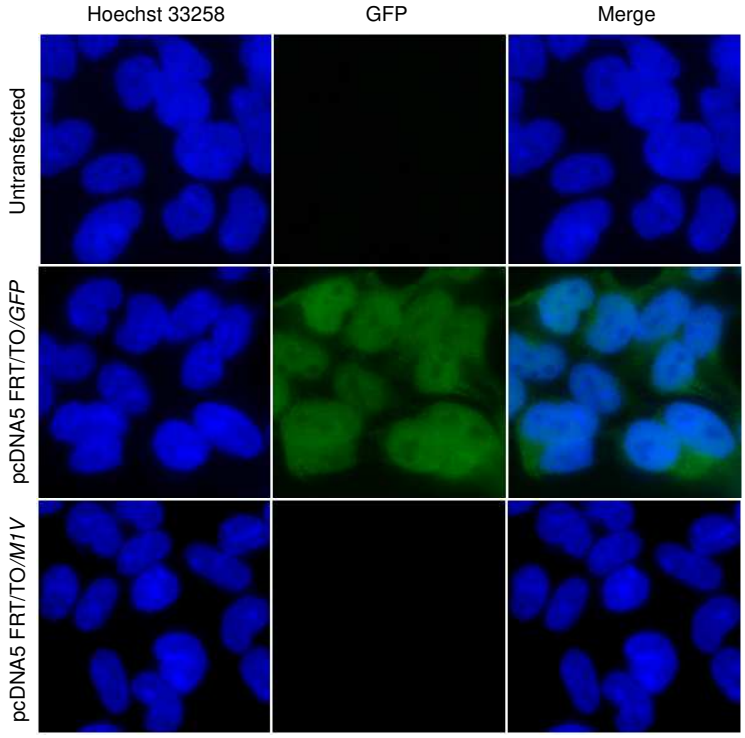


Figure 3.3 GFP expression from pcDNA5 FRT/TO/*GFP* and pcDNA5 FRT/TO/*M1V* constructs
GFP and *M1V* plasmids were transfected into HEK cells. Cells were stained with GFP (green) and Hoechst (blue). GFP is produced following transfection of pcDNA5 FRT/TO/*GFP*, but no GFP protein is produced in the untransfected HEK cells or following transfection of pcDNA5 FRT/TO/*M1V*.

To construct inducible GFP cell models the Invitrogen Flp-In T-REx kit was used (Figure 3.4). This system generates isogenic, stable cell lines with Tetracycline-inducible expression of the gene of interest (GOI).

Figure 3.4 Strategy for building stable cell line using the Flp In T-REx system
pcDNA5/FRT/TO vector is cotransfected with a recombinase plasmid and host cell line. Homologous recombination occurs at the FRT site, inserting the gene of interest into the host cell line. This changes the antibiotic resistance of the cells from Zeocin to Hygromycin. Expression of the inserted gene is controlled by the addition of Tetracycline.

Antibiotic resistance was used to identify which cells had undergone homologous recombination at the Flp recombinase target (FRT) site. HEK FRT cells grow in medium containing Blasticidin and Zeocin. Cells are resistant to Zeocin due to the expression of the lacZ-Zeocin resistance gene under the control of the SV40

promoter and due to the presence of the ATG initiation codon. The pcDNA5/FRT/TO vectors contain a Hygromycin resistance gene, which lacks a promoter and ATG initiation codon. After transfection and recombination at the FRT site, the SV40 promoter and ATG initiation codon are brought into frame with the Hygromycin resistance gene. The Flp recombinase mediated recombination at the FRT site is the only way by which Hygromycin resistance can be acquired. Therefore, cells, which are sensitive to Zeocin and resistant to Hygromycin, can be selected as cells that have undergone homologous recombination at the FRT site.

The FRT cells contain the Tetracycline resistance repressor elements (TetR) and a Blasticidin resistance gene. The Blasticidin selects for cells, which will express the Tetracycline repressor protein. Maintaining cells in Blasticidin in culture ensures that the cells will be under the control of the Tetracycline operator (TO).

The pcDNA5/FRT/TO vectors contain the GOI under the control of Tetracycline-regulated hybrid human cytomegalovirus (CMV)/$TetO_2$ promoter. Tetracycline resistance repressor elements (TetR) express a Tetracycline repressor protein, which binds the TO, inhibiting expression from the CMV promoter. Tetracycline induces transcription by binding to Tetracycline repressor protein, causing it to undergo a conformational change. This change makes the repressor protein unable to bind the TO, resulting in expression of the GOI.

In order to ensure a totally isogenic population was achieved transfected cells were left to grow until visible colonies of cells were formed from a single transfected cell. Single transfected cells were allowed to grow until a colony of cells visible without the use of microscope was observed. Twelve individual colonies were screened from each cell model. Each colony, grown from a single transfected cell, was picked and grown in a 24 well plate. To confirm homologous recombination colonies were

divided in two and grown in medium containing either Hygromycin and Blasticidin or Zeocin and Blasticidin (Figure 3.5 A). Cells that had undergone homologous recombination grew quickly in the medium containing Hygromycin and Blasticidin, and a change in the colour (from red to orange/yellow) of the media can be observed. These cells did not grow in medium containing Zeocin and Blasticidin and no colour change of the media is observed. Additional confirmation of successful

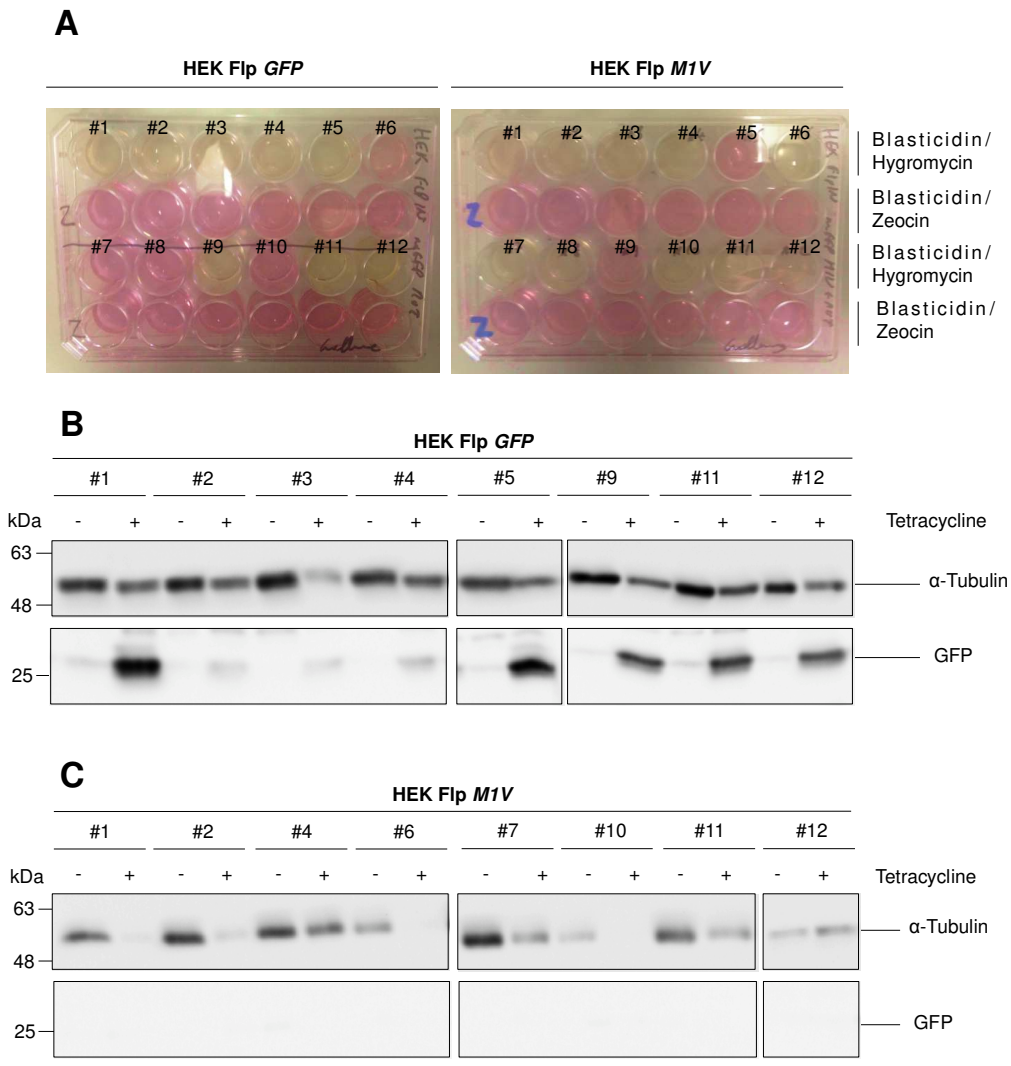


Figure 3.5 Selection of Hygromycin resistant and Tetracycline inducible clones(A) 12 colonies were grown until confluent in a 24 well before dividing into two wells of a 24 well plate, one containing Blasticidin/Hygromycin media and one containing Blasticidin/Zeocin media. A change in the colour of the media was observed for cells that were Hygromycin resistant and continued to grow. (B) Hygromycin resistant clones from HEK Flp *GFP* were grown for 48h with and without 10 µg/mL Tetracycline before lysis and western blot probing for GFP. Clone number 9 was selected. (C) Hygromycin resistant clones from HEK Flp *M1V* were grown for 48h with and without 10 µg/mL Tetracycline before lysis and western blot probing for GFP. Clone number 4 was selected.

model building was carried out by testing the expression of the gene of interest following induction with Tetracycline (Figure 3.5 B and C). Clones were selected based upon the expression of the GFP and 'leakiness' of the cell model i.e. how much GFP is observed in the non-induced sample. HEK Flp *GFP* clone number 9 was selected for use based upon these principles. No GFP protein is observed in the HEK Flp *M1V* clones in the presence or absence of Tetracycline. Therefore clone number 4 was selected, as no GFP was present in either the presence or absence of Tetracycline.

Immunofluorescence studies were carried out in order to confirm the production of GFP from the cell models (Figure 3.6). The HEK Flp *GFP* cell model produced visible GFP protein when grown in the presence of Tetracycline, but not in the absence of Tetracycline. In contrast, no GFP protein was produced in the HEK Flp *M1V* cell model either in the presence or absence of Tetracycline. This is expected as the M1V transcripts lack AUG start codons within most of the GFP sequence in the HEK Flp *M1V* cell model.

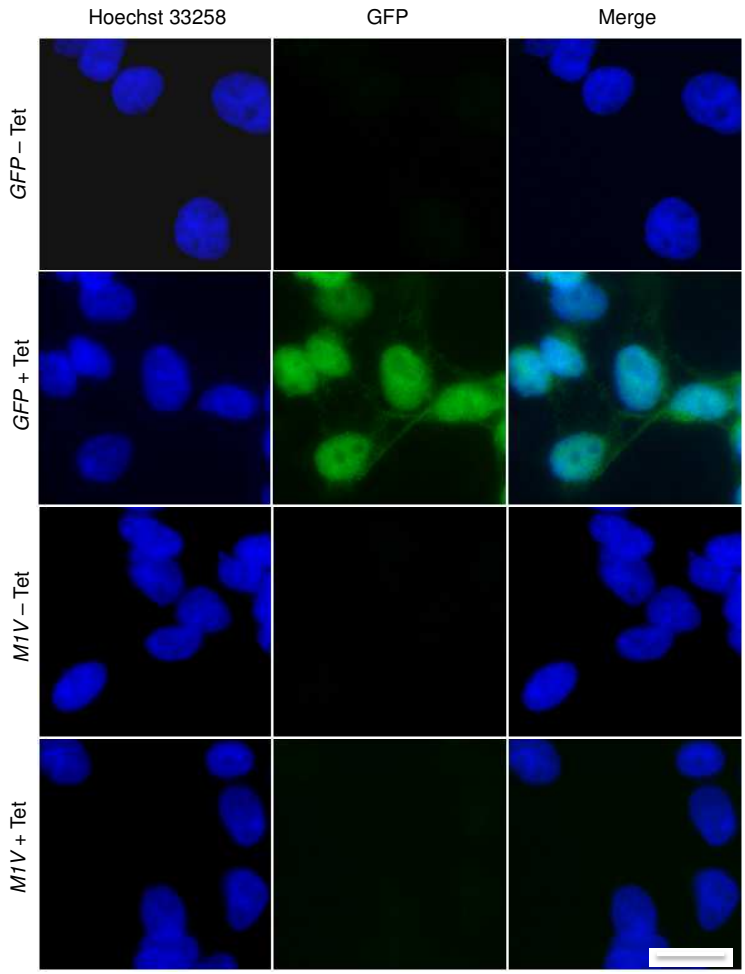


Figure 3.6 Immunofluorescence of GFP protein in HEK Flp *GFP* and *M1V* Cells were grown on coverslips for 24 h before the addition of 10 µg/mL Tetracycline for 24 h. Cells were stained with GFP (green) and Hoechst (blue). GFP is observed in the HEK Flp *GFP* cells grown in the presence of Tetracycline for 24 h. No GFP is detected in the HEK Flp *M1V* cells grown in the presence of Tetracycline for 24 h. Scale bar = 20 µm.

The expression of GFP from the cell model was quantified by western blot (Figure 3.7). In the absence of Tetracycline approximately 7 % expression of GFP is observed compared to HEK Flp *GFP* cells grown in the presence of GFP. Again, no GFP protein is produced by the HEK Flp *M1V* cell model either in the presence or absence of Tetracycline. This is in agreement with the immunofluorescence shown in Figure 3.6. Some weaker expression level of the GFP protein was detected in HEK Flp *GFP* cells grown in the absence of Tetracycline. This is probably due to the leakiness of the promoter. This does not cause a problem for these studies, as the

amount produced upon on the addition of Tetracycline is significantly greater than that produced in the absence of Tetracycline.

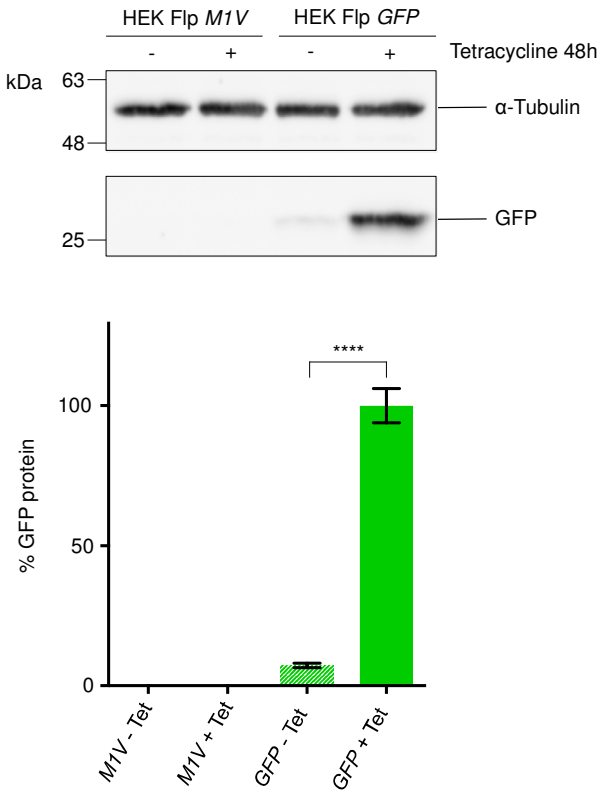


Figure 3.7 GFP protein levels in HEK Flp *GFP* and *M1V* cell models
Cells were plated for 24 h before a 48 h induction in the presence or absence of 10 µg/mL Tetracycline prior to western blot analysis. Quantification was normalised to HEK Flp *GFP* + Tet. In the absence of Tetracycline approximately 7 % GFP expression is observed in the HEK Flp *GFP* cell model. In the presence and absence of Tetracycline no GFP protein is observed in the HEK Flp *M1V* cell model. (Mean ± SEM; One-way ANOVA with Tukey's multiple comparison, ****: $p < 0.0001$; N = 3).

qRT-PCR primers were designed to recognise the GFP transcripts from both the HEK Flp *GFP* and *M1V* cell model. It was very important that the primers were able to detect transcripts from the HEK Flp *GFP* and *M1V* cell models equally. This was to compare the GFP transcript levels in the cell models and to distinguish in future experiments between *GFP* translated transcripts and *M1V* transcripts. The primers were designed against regions in which the transcripts were identical and the primers were tested with cDNA from both cell models (Figure 3.8). The primers were able to detect GFP transcripts from both cell models equally with efficiencies of close to 100

% and R^2 values of 0.99. Due to the design of the qPCR primers and mutation of AUG start codons in the first 655 nucleotides of the *GFP M1V* sequence, any qPCR products or transcripts detected in purified ribosomes from the HEK Flp *M1V* cell model will not correspond to AUG-driven initiation of translation.

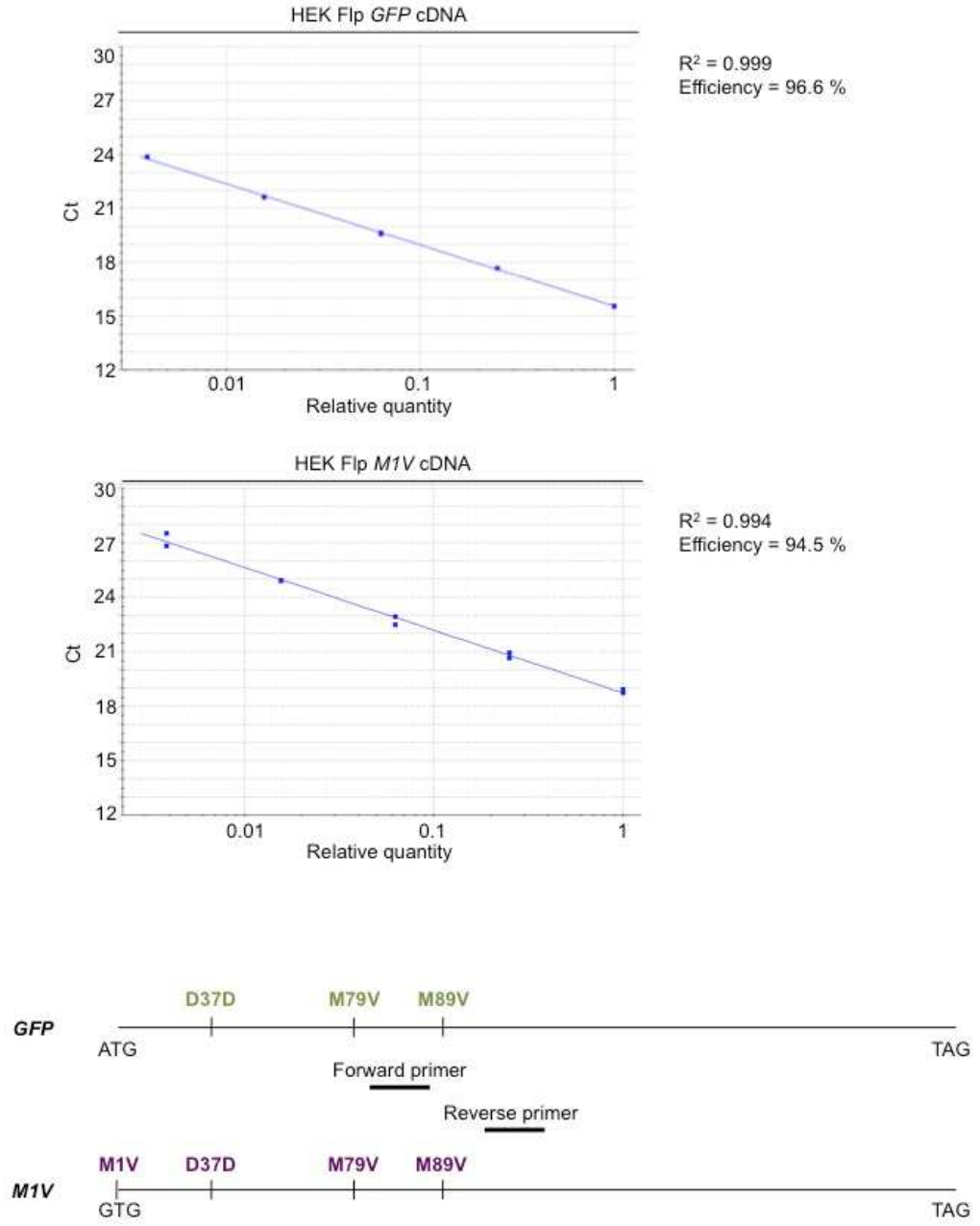


Figure 3.8 GFP qRT-PCR primer design and standard curves
Standard curves were generated using serial dilutions of cDNA from both the HEK Flp *GFP* and HEK Flp *M1V* cell models grown in the presence of 10 µg/mL Tetracycline for 24 h. The primers were able to detect GFP transcripts from both the HEK Flp *GFP* and HEK Flp *M1V* cell models with efficiencies close to 100 % and R^2 values of 0.99. Primers were designed against regions that were common to both the HEK Flp *GFP* and HEK Flp *M1V* cell lines. Mutations are indicated. (ATG: start codon; TAG: stop codon.)

GFP and M1V transcript levels were determined in HEK Flp *GFP* cells grown in the presence and absence of Tetracycline and in HEK Flp *M1V* cells grown in the presence of Tetracycline (Figure 3.9). Upon the addition of Tetracycline there is a 20-fold increase in the GFP mRNA in the HEK Flp *GFP* cell model. Over multiple experiments, there was no statistically significant difference in the GFP mRNA levels between HEK Flp *GFP* and HEK Flp *M1V* cells grown in the presence of Tetracycline indicating that both GFP and M1V reporter transcripts are similarly expressed and that substituting AUG start codons does not significantly affect RNA biogenesis and stability. This was important as it showed that the lack of GFP protein produced from the HEK Flp *M1V* cell model was not due to a difference in the steady state level expression of the M1V transcripts produced.

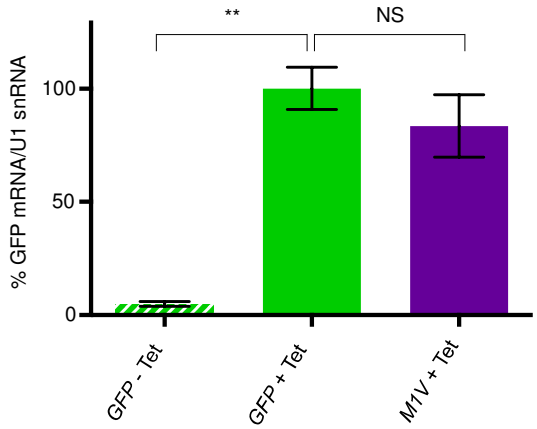


Figure 3.9 GFP transcript levels in HEK Flp *GFP* and *M1V* cell linesCells were plated for 24 h before the addition of 10 µg/mL Tetracycline for 24 h. Quantification was normalised to HEK Flp *GFP* + Tet. A 20-fold increase in GFP mRNA was observed in the HEK Flp *GFP* cell line upon the addition of Tetracycline. No significant difference in GFP mRNA levels was observed between HEK Flp *GFP* + Tet and HEK Flp GFP *M1V* + Tet. (Mean ± SEM; 2way ANOVA, Tukey's Test, **: $p < 0.01$, NS: non-significant; N = 3)

The HEK Flp *GFP* cell model produces GFP transcripts containing a start codon that are translated into GFP protein. The HEK Flp *M1V* cell models produces similar levels of M1V transcripts, but does not produce GFP protein. This made these cell models suitable tools to discriminate between translating and non-AUG translating mRNAs.

3.2.2 Testing current translatome technologies

It has been observed that translation can occur at other non-conventional start codons in genome-wide translatome experiments involving ribosome profiling using specific inhibitors (S. Lee, Liu, Lee, Huang, Shen, and Qian, 2012b). The aim of these experiments was to use the GFP reporter cell models to test whether the current translatome technologies based upon affinity purification were able to discriminate between translating and unconventional non-AUG translating trancripts.

RIP was performed using antibodies against ribosomal proteins (L10a and L26). This was using the principle that immunoprecipitation could be used to isolate mRNA associated with ribosomal proteins and that these mRNAs would be undergoing active protein synthesis. The feasibility of using ribosomal protein L10a and L26 antibodies in RNA immimmunoprecipitation (RIP) was explored using HEK cell lysates (Figure 3.10). Both antibodies bound non-specifically to control beads and therefore were not suitable for use in RIP.

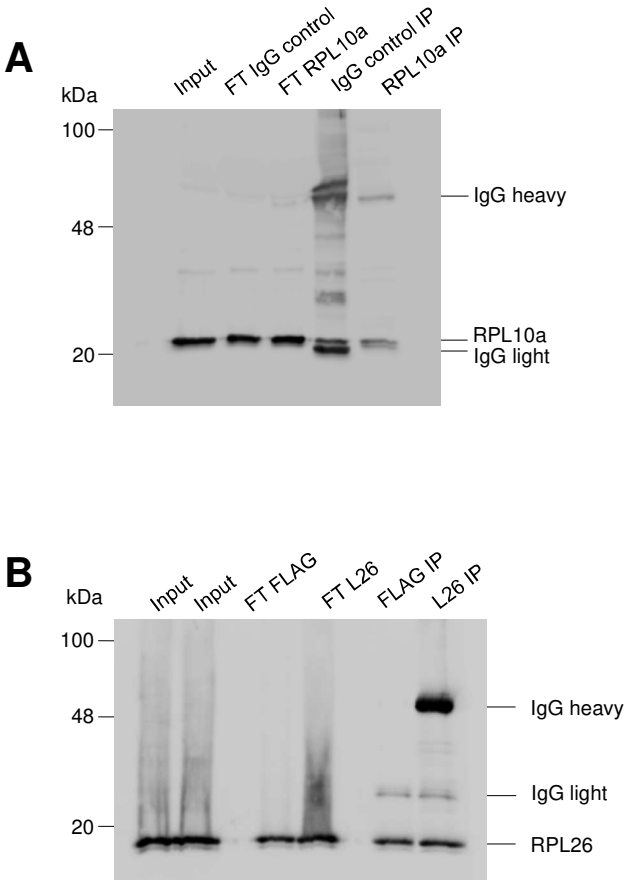


Figure 3.10 RIP using RPL10a and RPL26
RIP using HEK cell lysates. (A) RIP using uncoated Protein G beads (IgG control) or RPL10a coated Protein G beads (RPL10a). RPL10a bound non-specifically to control IgG. (RPL10a molecular weight = 25 kDa) (B) RIP using FLAG coated (FLAG) or RPL26 coated Protein G beads (RPL26). RPL26 bound non-specifically to FLAG control beads. (RPL26 molecular weight = 17 kDa). (FT = flow through – supernatant removed from beads post binding)

Another methodology for purifying ribosomes was therefore used. The TRAP protocol described by (Kulicke et al., 2014) was used to purify mRNA bound to FLAG tagged RPL10a from HEK Flp *GFP* and HEK Flp *M1V* cells induced with Tetracycline (Figure 3.11). Cells were transfected with either FLAG or FLAG RPL10a and the TRAP carried out.

RPL10a protein is enriched in the eluted fraction of FLAG RPL10a but not in the FLAG control. This shows that FLAG RPL10a binds specifically to the FLAG beads.

RNA was extracted from eluted fractions; these RNA are bound to FLAG RPL10a and therefore it is assumed that these RNA that are being translated. There was no significant difference between the numbers of GFP or M1V transcripts in the eluted fractions of the HEK Flp *GFP* cells compared to the HEK Flp *M1V* cells. This unexpected result indicates that TRAP does not discriminate between cognate GFP-translating mRNAs and potentially non-AUG translating M1V transcripts.

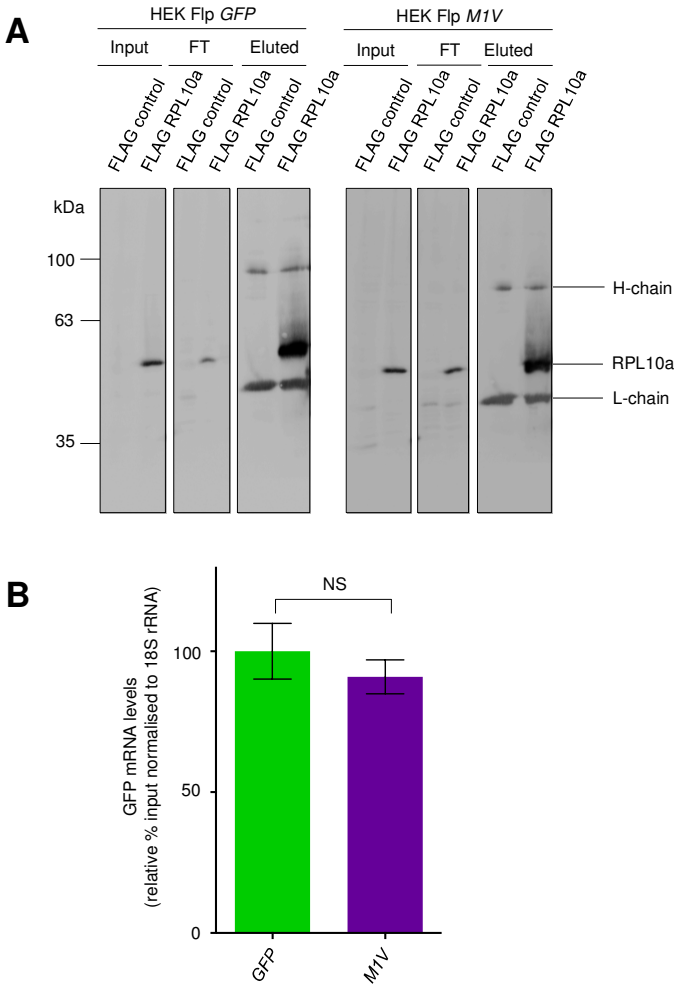


Figure 3.11 TRAP of HEK Flp *GFP* and *M1V* cell lines using FLAG RPL10a
Cells were induced for 12-15 h with 10 µg/mL Tetracycline before TRAP purification. (A) Western blot shows presence of RPL10a in input and flow through (FT) in both HEK Flp *GFP* and HEK Flp *M1V* cell lines in FLAG RPL10a but not in FLAG controls. RPL10a is enriched in eluted fractions from both cell lines in FLAG RPL10a but not in FLAG controls. (B) qRT-PCR of eluted fractions. Values were normalised to HEK Flp *GFP* and no significant difference in the GFP transcript level between the cell lines was observed. (Mean ±SEM; t-test, NS = non-significant; N = 3 RT reactions)

3.2.3 Optimising a new purification protocol for selective enrichment of actively AUG-driven translating ribosomes

Instead of tagging ribosomes or use RIP assays, an alternative strategy for the selective enrichment of actively translating ribosomes was designed and developed. This involved precipitating endogenous ribosomes using differential centrifugation in stringent buffer conditions that would allow for selective enrichment of actively driven AUG-translating ribosomes. Differential centrifugation allowed the removal of different cellular components such as the nuclei and mitochondria. The stringent buffer conditions would allow dissociation of accessory ribosome-associated RNA binding proteins, such as initiation and other regulatory factors, that would contaminate isolated ribosomal mRNA with other non-AUG translating RNA.

Several approaches were tested based upon the biochemical precipitation of ribosomes and use of UV-irradiation and inhibitors. UV-irradiation was used to covalently crosslink RNA:protein interactions, preventing their dissociation to ensure RNA could subsequently be extracted from ribosomal fractions. The use of cycloheximide (CHX), an inhibitor of translation elongation, was also used, to investigate whether inhibiting translation before purifying ribosomes would enhance the specific purification of actively translating mRNA.

Live cells were lysed before UV-irradiation prior to NP-40 lysis. Ribosomes were biochemically purified using differential centrifugation and ultra centrifugation. Proteinase K treatment was carried out to remove ribosomal proteins before Trizol RNA extraction. Extracted RNA was converted to cDNA using reverse transcription. Semi-quantitative RT-PCR was then carried out and the PCR products run on agarose gels to determine any differences in ribosomal GFP transcripts between samples.

In the development of this method extracted RNA was DNase treated and used for semi-quantitative RT-PCR. The first experiment was carried out on induced and non-induced HEK Flp *GFP* cells and the GFP transcript levels were compared (Figure 3.12). A 5-cycle difference is observed in the presence of Tetracycline, this corresponds to an approximate 32-fold enrichment upon on the addition of Tetracycline. This showed that the biochemical purification of ribosomes and subsequent RNA extraction from ribosomes precipitated could be used to detect changes in ribosome-associated RNA.

A 32-fold enrichment on the addition of Tetracycline means approximately 3 % expression in the non-induced cells due to leakiness of the promoter. This correlates well with the previous experiments where 5 % expression was observed in the non-induced cells by qRT-PCR and 7 % expression observed in the non-induced cells quantified from western blots. It also showed that this method could be used to detect the difference in GFP transcript level between induced and non-induced cells.

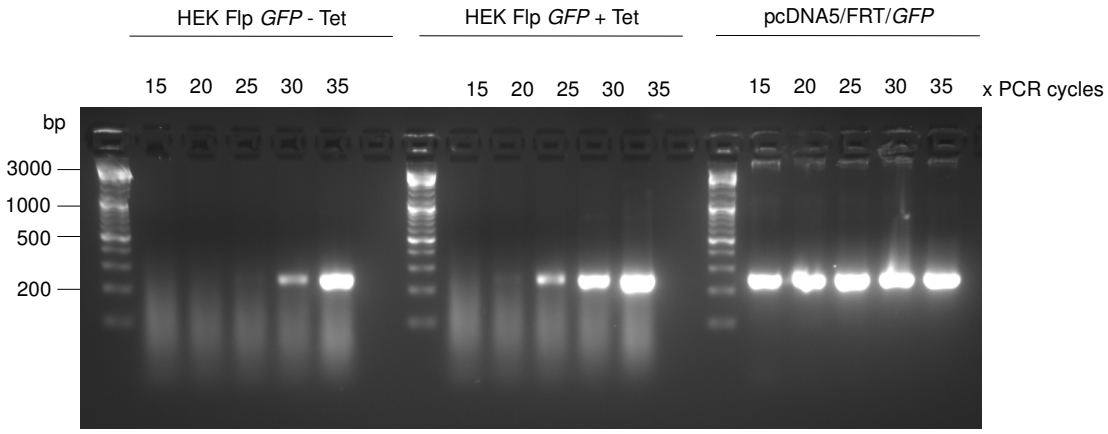


Figure 3.12 Biochemical purification of ribosomes in HEK Flp *GFP* cells
HEK Flp GFP cells were UV-irradiated prior to lysis in GRASPS lysis buffer A1. Semi-quantitative RT-PCR was performed on samples from HEK Flp *GFP* cells grown for a total of 48 h, in the presence or absence of 10 µg/mL Tetracycline for 24 h. There is a difference of at least 5 cycles between induced and non-induced cells. This corresponds to approximately 32-fold enrichment upon the addition of Tetracycline. pcDNA5/FRT/*GFP* plasmid was used as a control for semi-quantitative RT-PCR.

This method was then used on both HEK Flp *GFP* and HEK Flp *M1V* cells to determine if a difference between the detection of GFP and M1V mRNAs was possible (Figure 3.13).

In the ribosomal fractions, a 5-cycle difference was observed upon the addition of Tetracycline in the HEK Flp *GFP* cell line and HEK Flp *M1V* cell line. GFP transcripts are detected 5 cycles later in the HEK Flp *M1V* compared to the HEK Flp *GFP* cells when cells are grown in the presence of Tetracycline. This indicates a 32-fold difference in the amount of GFP reporter mRNAs detected in the purified ribosomes between the two cell lines. The post-ribosomal fraction corresponds to RNA extracted from the supernatant, which is removed after the ribosomes have been pelleted by ultra-centrifugation. In the post-ribosomal fraction there are less GFP transcripts than in the ribosomal fraction, indicating enrichment of ribosomes with GFP reporter transcripts in the GRASPS purified fractions from both cell models. This indicates successful purification of ribosomes has been achieved using this method. The no RT controls for both fractions appear over 10 cycles later.

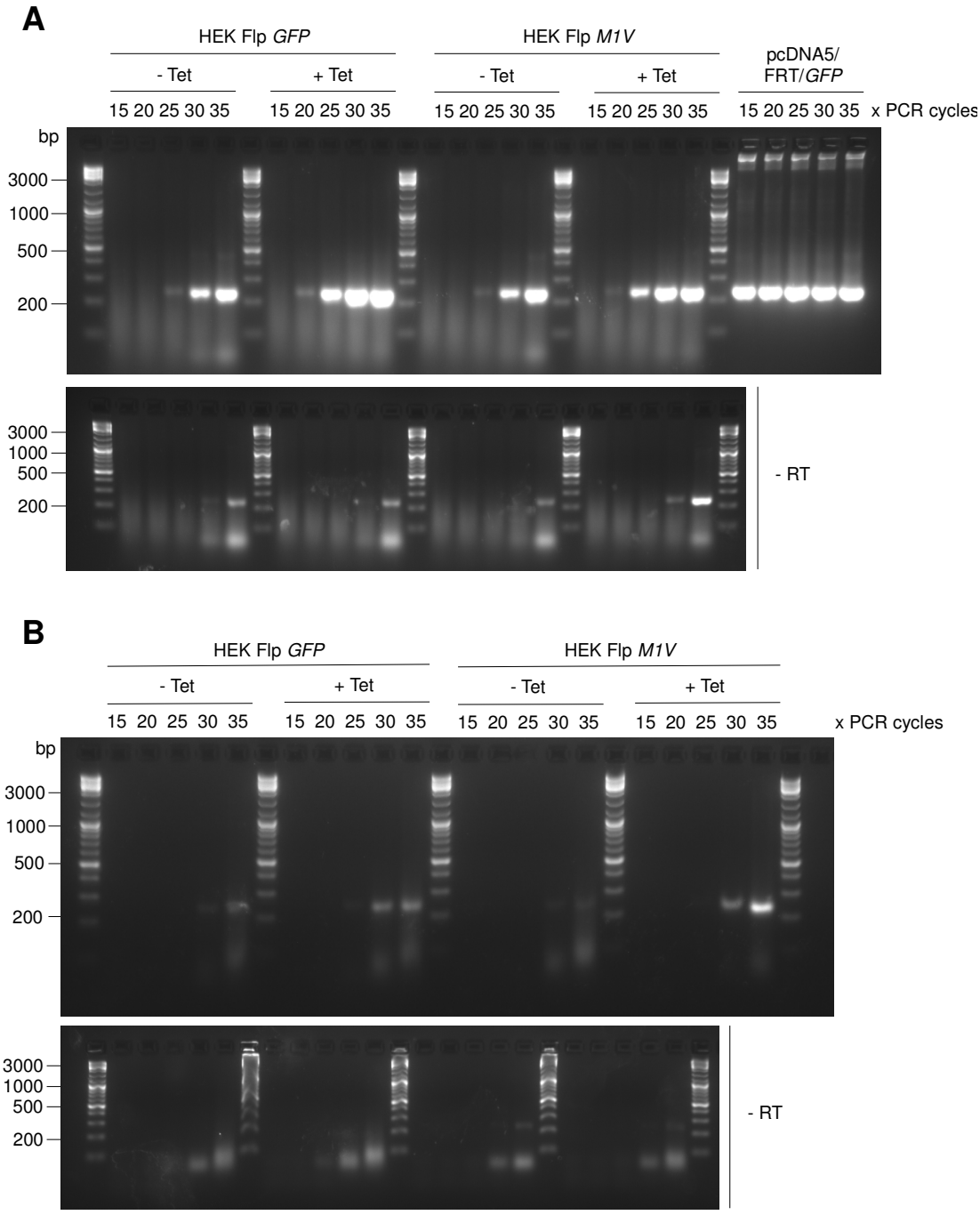


Figure 3.13 Biochemical purification of ribosomes in HEK Flp *GFP* and *M1V* cells

Cells were UV-irradiated prior to lysis in GRASPS lysis buffer A1. Semi-quantitative RT-PCR was performed on samples from HEK Flp *GFP* and HEK Flp *M1V* cells grown for a total of 48 h, in the presence or absence of 10 µg/mL Tetracycline for 24 h. pcDNA5/FRT/*GFP* plasmid was used as a control for semi-quantitative RT-PCR. (A) RNA extracted from ribosome. There is a 5-cycle difference in the point at which GFP transcripts are observed between induced HEK Flp *GFP* and *M1V* cells. (B) RNA extracted from post-ribosomal fraction (supernatant post ultra-centrifugation). There is a visible reduction in the amount of GFP present in the post-ribosomal fraction in both HEK Flp *GFP* and *M1V* cell lines.

In order to confirm purification of ribosomes, protein extracts from different stages of the GRASPS purification were analysed on Coomassie gels and by western blotting (Figure 3.13). The Coomassie gels showed a distribution of higher molecular weight proteins (45 kDa and above) in the post-mitochondrial and post-ribosomal fractions with an enrichment of low and medium molecular weight proteins (less than 50 kDa) in the ribosomal fraction. Lower molecular weight, ribosomal proteins were enriched in the ribosomal fraction and lost from the post-ribosomal fraction, showing that GRASPS had successfully purified ribosomal proteins. The distribution of proteins between the different fractions is the same for both HEK and NSC-34 lysates, showing it is applicable to different cell types.

Western blot analysis showed the ribosomal protein L26 (RPL26) and ribosomal protein L29 (RPL29) were present in the cytoplasmic, post-mitochondrial fractions and ribosomal fractions, but lost from the post-ribosomal fraction. Ribosome-associated accessory factors eIF4a and eEF2 were lost from the ribosomal fraction. This shows the clean purification of ribosomes, with the removal of RNA-binding accessory factors eIF4a and eEF2. This is important as it shows that GRASPS is able to purify ribosomes with the removal of accessory factors. These accessory factors are able to bind RNA alone, meaning that if they were purified with ribosomes in the GRASPS purification they could contaminate the purified sample with non-translating mRNA. These data show successful purification of ribosomes without accessory factors using the GRASPS method.

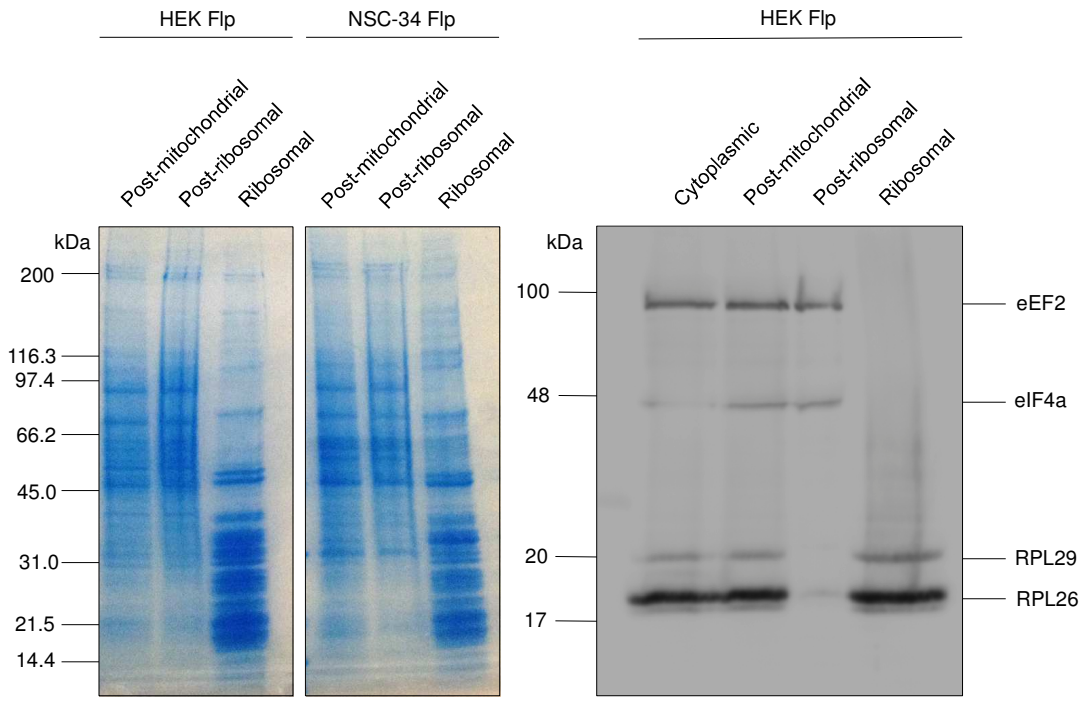
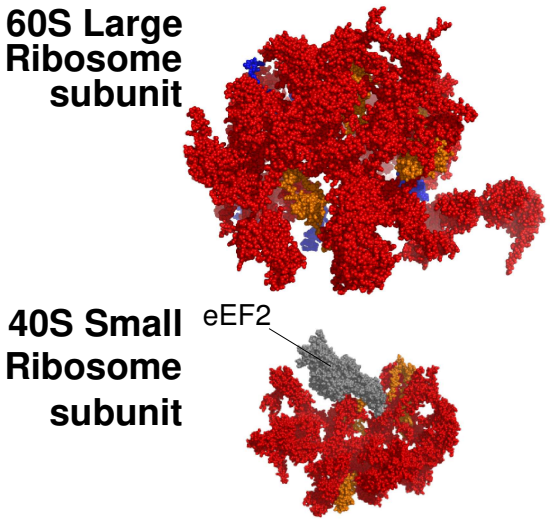


Figure 3.14 Confirming purification of ribosomes
Different fractions from the ribosomal purification of HEK Flp and NSC-34 Flp cell models were run on Coomassie gels. In the post-mitochondrial and post-ribosomal fraction there are proteins of high and low molecular weights. In the ribosomal fraction there is an enrichment of low and medium molecular weight proteins, which would include low molecular weight ribosomal proteins. HEK Flp cell fractions were analysed by western blot. In the cytoplasmic and post-mitochondrial fractions ribosomal proteins (RPL26 and RPL29) and accessory factors (eEF2 and eIF4a) can be identified. In the post-ribosomal fraction only accessory factors are present, as the ribosomal proteins have been pelleted in the ribosomal fraction.

Ribosomal proteins purified were separated by SDS-PAGE and identified by mass spectrometry. The mass spectrometry was performed by Prof Mark Dickman (Department of Chemical Engineering at the Life Science Interface, University of Sheffield). The Mascot algorithm was used to compare the molecular weights from peptides obtained from the digestion of proteins purified in the GRASPS sample to the known protein database (PDB) using the *Mascot* program and give hits a probabilistic score for protein identification by counting the number of peptides that match with known proteins in the database. The higher the score, the more peptides that are matched to the database, and therefore the higher the confidence in the protein identification. Proteins were positively identified if the Mascot score >25 and if

at least two different peptides were characterised for each protein. The score given and location on the ribosome subunit is shown in Figure 3.15.



	Mascot score >50 shown in RED	Mascot score 25<50 shown in ORANGE	Manually identified shown in BLUE
RPL:	0, A1, A2, 3, 4, 5, 6, 7, 7A, 8, 9, 10L, 10A, 11, 12, 13, 13A, 14, 15, 17, 18, 19, 21, 22, 23, 23A, 24, 26, 27, 27A, 28, 29, 31, 34, 35, 35A, 36, 37A, 38	18A, 32, 36A	30, 37, 39, 40, 41
RPS:	2 = 4, 3, 3A, 4 X-isoform, 5, 5A, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 15A, 16, 17L, 18, 19, 20, 21, 23, 24, 25, 26, 28, GBLP = RACK1, PAIRB	27, 27A, 29, 30	

Figure 3.15 Mass spectrometry on purified ribosomes Mass spectrometry of purified ribosomal proteins separated by SDS-PAGE. Subunits shown in red have been identified with high confidence with a Mascot score of >50 and at least two different peptides used for identification. Subunits shown in orange have been identified with lower confidence with a Mascot score of 25<50 and using one or two peptides for identification. Subunits shown in blue were missing from Mascot analysis due to a score of less than 25, but were identified by manually looking for these proteins in the list of those with a Mascot score of less than 25. Purified proteins were mapped onto the 80S ribosome structure. (RPS: ribosomal protein small subunit; RPL: ribosomal protein large subunit.)

From the filtered hits and only considering those with Mascot scores greater than 50, 92 human proteins were identified. Of these 84 % were identified as ribosome-associated proteins. Of the 81 ribosome protein subunits in the high resolution Cryo-EM structure of the human 80S ribosome (Anger et al., 2013), 69 ribosome protein subunits were identified in the GRASPS purified proteins. In addition to these 69 ribosome protein subunits, 8 other ribosome-associated proteins were also identified in the GRASPS sample (PABP1, PABP4, SYMC, SYRC, LARP4, eIF4A3, IF6, MCA3). From these ribosome-associated proteins, 7 are also published in the

structure of the human 80S ribosome (Anger et al., 2013). This shows that GRASPS is able to purify fully associated ribosomes with high coverage of the published 80S ribosome subunit.

mRNA-associated proteins hnRNP C, hnRNP F, hnRNP H1, hnRNP M and hnRNP R were identified and accounted for 5 % of GRASPS purified protein. False positives or DNA-binding proteins such as ACTB, YBOX1, YBOX3, H2A1B, H12, NUCL, TBA-1B (tubulin), TCPB (Chaperone), TIF-1B (DNA-binding), VIME were identified and these accounted for 11 % of the GRASPS purified proteins. Contamination by known false positives and DNA-binding proteins or mRNA associated proteins accounted for only 16 % of the proteins identified in the GRASPS purified proteins. This shows that GRASPS is stringent, with low levels of contaminating proteins.

The proteins identified by mass spectrometry did not include eEF2, despite its association with the ribosome. These data show successful purification of ribosomes without accessory factors (at least eIF4A which binds mRNA directly or eEF2) using the GRASPS method.

In order to try and remove non-AUG translated transcripts from GRASPS purified samples the use of inhibitors, composition of lysis buffers and timing of UV irradiation were explored.

CHX is a protein synthesis inhibitor, which works by interfering with the translocation step blocking translational elongation. CHX was added to the growth media of cells for 2 h prior to GRASPS purification, with the idea that this may stall ribosomes on translating mRNA and allow the detection of actively translating mRNA specifically. CHX was added to the culture media for 2 h prior to ribosome purification (Figure 3.16). The addition of CHX did not have any affect upon the GFP transcript level at

the ribosome in either the HEK Flp *GFP* or HEK Flp *M1V* cell line and non-AUG translated M1V transcripts were still detected in the HEK Flp *M1V* GRASPS purified samples.

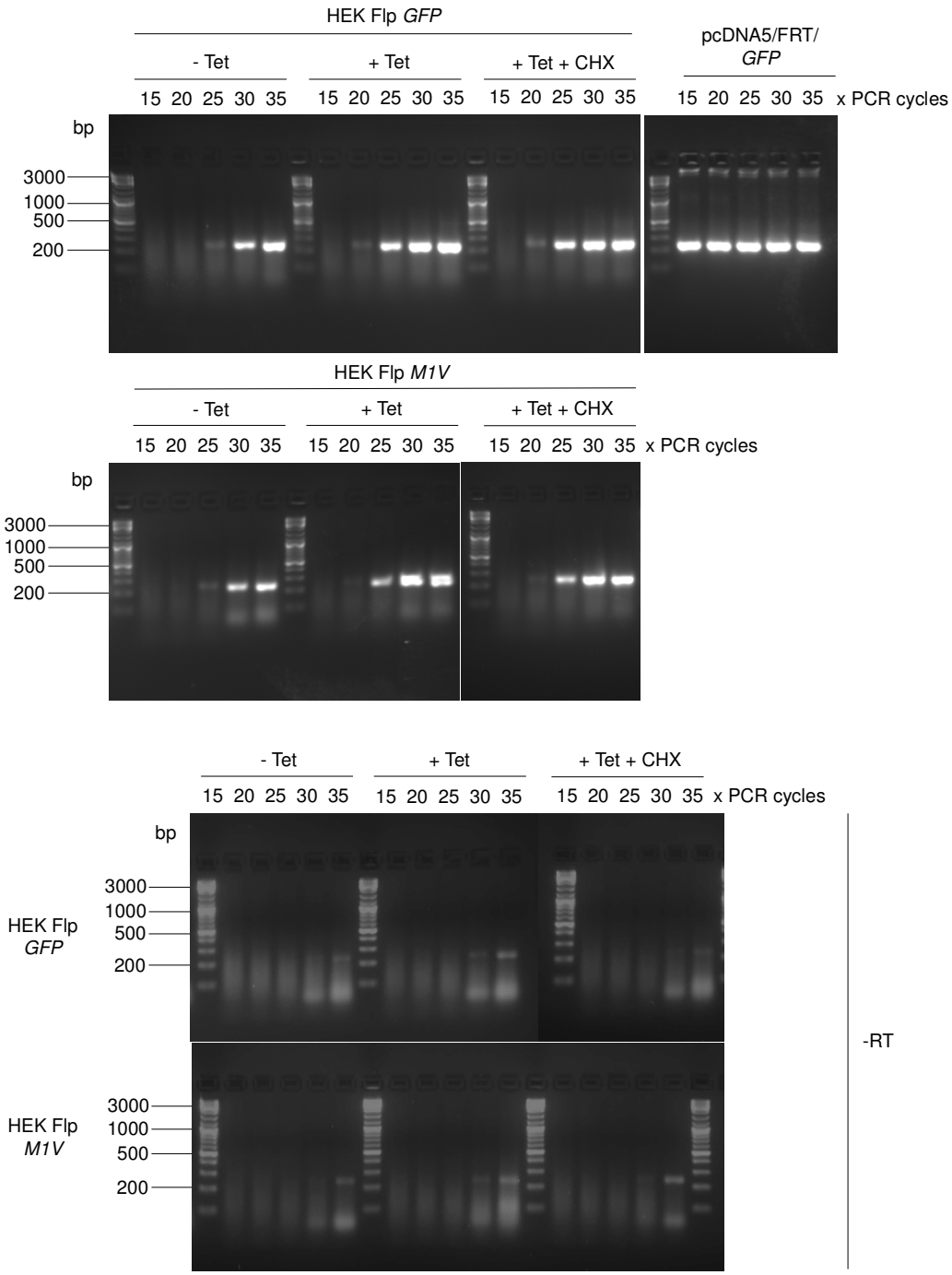


Figure 3.16 Biochemical purification of ribosomes with CHX
Cells were treated with 100 µg/mL CHX in the growth media for 2 h prior to experiment. Cells were UV-irradiated prior to lysis in GRASPS buffer A1. Treatment with CHX did not affect the GFP transcript level at the ribosome in either cell line.

The point in which cells were UV-irradiated was moved to after the NP-40 lysis before the cellular fractionation, in order to see whether crosslinking after lysis in stringent buffer condition (250 mM sucrose, 250 mM KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl pH 7.4) would allow for dissociation of non-AUG translating M1V transcripts from the ribosomes (Figure 3.17). The timing modification in the UV-irradiation did not change the overall result as association of non-AUG transcripts to ribosomes were still detected in the HEK Flp *M1V* GRASPS purified samples. However, the appearance of both GFP and M1V transcripts from the GRASPS purified samples from HEK Flp *GFP* and *M1V* cells appeared 5 cycles later, indicating an improved signal to noise reduction compared to UV-irradiating prior to NP-40 lysis.

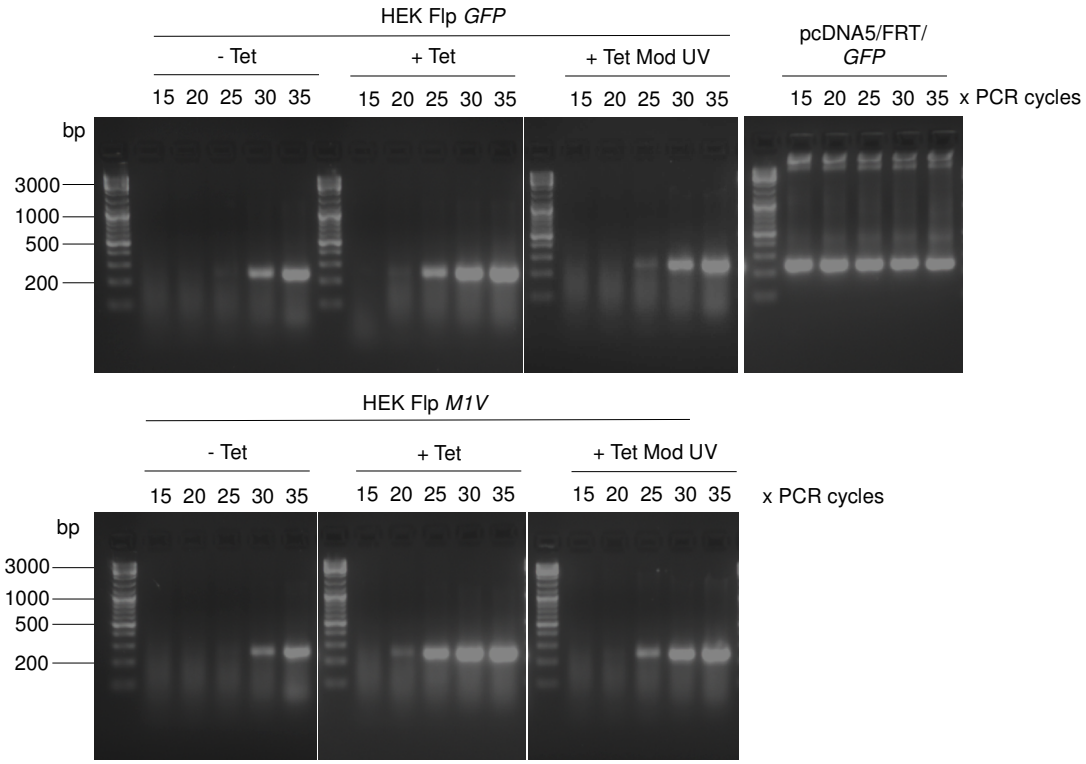


Figure 3.17 Biochemical purification of ribosomes with modified UV-irradiation
Cells were lysed in GRASPS buffer A1 and UV-irradiated as usual and following NP-40 lysis (Mod UV). The delay in crosslinking shifted the appearance of the GFP transcripts by 5 cycles in both the HEK Flp *GFP* and HEK Flp *M1V* cell lines.

In addition to modification in applying the UV-irradiation step, changes to the composition of lysis buffer were also explored (Figure 3.18). This strategy was to change the ionic environment of the ribosomes and destabilise interactions between weakly associated or non-AUG translated mRNAs and the ribosomes. The removal of $MgCl_2$ and addition of EDTA had no effect on the number of GFP transcripts in the ribosome.

However, when GRASPS was performed using the hypotonic GRASPS lysis buffer A (5 mM KCl and without $MgCl_2$) and with UV-irradiation after cell lysis, a difference between the number of GFP and M1V transcripts in HEK Flp *GFP* compared to *M1V* was observed. This reduction in salt concentration may disrupt the weaker hydrophobic interactions of the large and small ribosome subunits to non-AUG translating M1V mRNAs in the HEK Flp *M1V* cell model, causing a reduction in the amount of M1V transcripts detected in the ribosomes. The other buffer compositions had no effect upon the detection of GFP between the two cell models. GRASPS will therefore be performed using GRASPS lysis buffer A and UV-irradiation post NP-40 lysis.

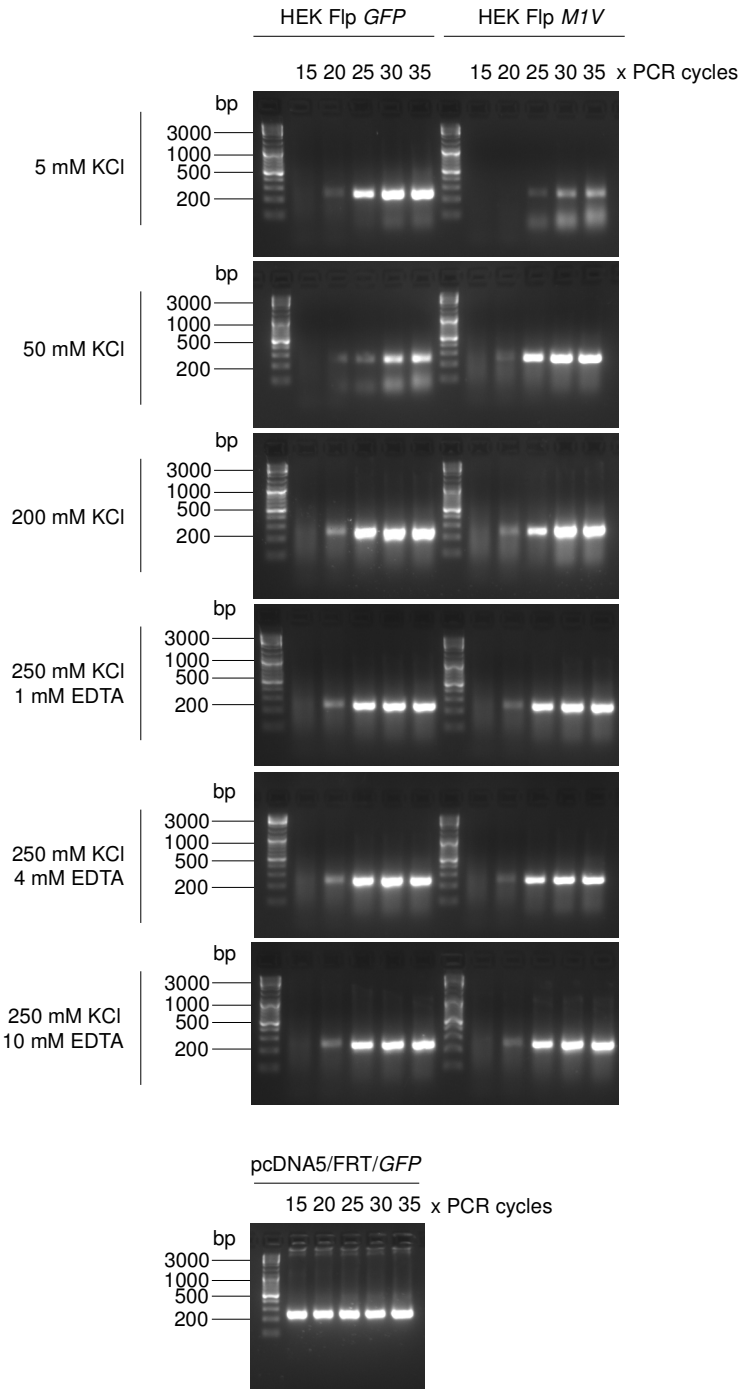


Figure 3.18 Biochemical purification of ribosomes with different lysis buffer compositions

The composition of GRASPS lysis buffer A1 was changed. Cells were lysed in a buffer without $MgCl_2$ with various KCl concentrations and with/without the addition of EDTA. Cells were UV-irradiated following lysis with NP-40. When cells were lysed in buffer without $MgCl_2$, with 5 mM KCl and without EDTA the amount of GFP transcripts in the ribosome in the HEK Flp *M1V* cell model is reduced compared to the number of GFP transcripts in the ribosome in the HEK Flp *GFP* cell model. The removal of $MgCl_2$ and addition of EDTA into the lysis buffer had no effect on the GFP transcripts in the ribosome in either cell model.

3.2.4 Comparison of the proportion of actively translating reporter mRNAs isolated from various translatome methodologies

The GRASPS methodology both using GRASPS buffer A1 (250 mM KCl and 5 mM $MgCl_2$) and GRASPS buffer A (5 mM KCl and without $MgCl_2$) was compared to the current TRAP methodology using the HEK Flp *GFP* and *M1V* cell models (Figure 3.19). For each method, the percentage of GFP transcripts was normalised to the input and 18S rRNA and is shown in the graph as a percentage of GFP transcripts compared the GFP transcripts from the HEK Flp *GFP* + Tet.

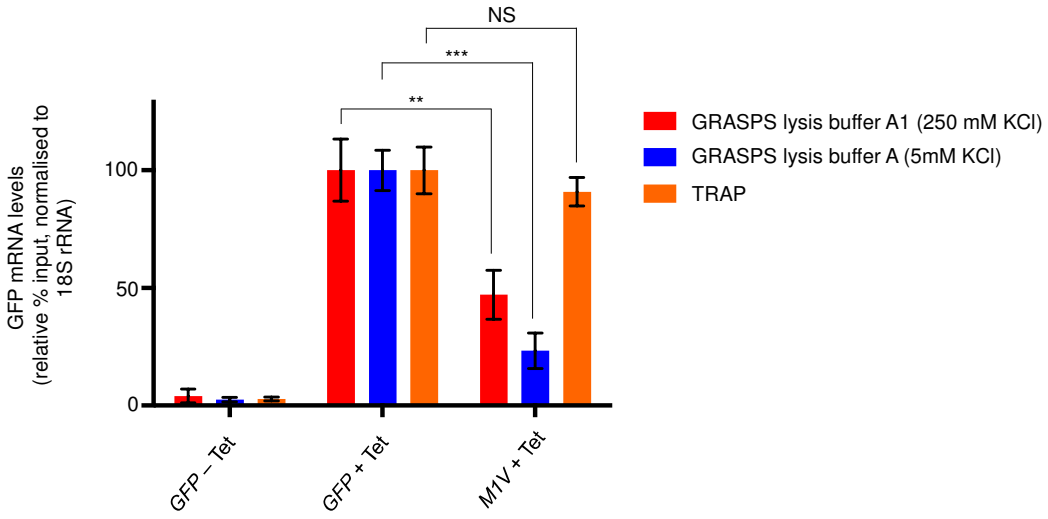


Figure 3.19 Comparison of TRAP and GRASPS
qRT-PCR analysis of Anti-FLAG-RPL10a TRAP and GRASPS purified RNA in HEK Flp *GFP* and *M1V* cell lines grown in the presence and absence of 10 µg/mL Tetracycline. GFP levels are normalised to 18S rRNA and shown as relative percentage of GFP for each method. There is a marked reduction (~80%) of non-AUG translating *M1V* mRNA in GRASPS- purified ribosomes lysed in GRASPS buffer A in contrast to TRAP. (Mean ±SEM; Two-way ANOVA with Tukey's multiple comparison, **: $p < 0.01$, ***: $p < 0.001$, NS = non-significant; N = GRASPS: 3, TRAP: 3 RT reactions)

When cells were lysed using GRASPS buffer A1 (250 mM KCl and 5 mM $MgCl_2$) and the ribosomes purified, 47% of *M1V* non-AUG translated mRNA was identified compared to AUG translated *GFP* mRNA. When GRASPS buffer A (5 mM KCl and no $MgCl_2$) was used to purify ribosomes 23 % of *M1V* mRNA was identified compared to AUG translated *GFP* mRNA. Therefore, the lysis of cells in GRASPS buffer A is the best method to selectively enrich for AUG-translated mRNAs in this

reporter system. There was a significant difference in the amount of GFP AUG-driven translation of mRNAs identified compared to non-AUG M1V translating mRNAs using GRASPS when the cells were lysed in GRASPS buffer A1 and GRASPS buffer A.

When TRAP was used to purify RPL10a associated transcripts 90 % of *M1V* transcripts remained identified compared to the AUG-translating *GFP* mRNA. TRAP therefore does not appear to discrimate between active translation of the cognate GFP-encoding mRNAs and non-AUG translated M1V reporter transcripts.

3.3 Discussion

The word 'translatome' is widely used to define ribosome associated mRNA, with the general assumption that this mRNA is undergoing translation. Methodologies have been described as identifying the translatome, without checking that mRNA purified is in fact undergoing translation. In order to monitor actively translating mRNA molecules HEK GFP reporter cell lines were built.

The translating ribosome affinity purification (TRAP) method (Heiman et al., 2008; Kulicke et al., 2014) was used to purify translating mRNA from the HEK Flp *GFP* and *M1V* cell lines. A FLAG-tagged RPL10a, rather than a GFP-tagged RPL10a, was used in these experiments. This modification was designed with the idea that as the FLAG tag is smaller than GFP, it would therefore provide less steric hindrance to the ribosome subunits. RPL10a is a good subunit to tag as it is on the outside of the ribosomal subunit and is therefore accessible to tag and has been widely used in TRAP purifications (Heiman et al., 2008; Kulicke et al., 2014; Kratz et al., 2014; Bertin et al., 2015).

RPL10a was enriched in eluted fractions from both the *GFP* and the *M1V* cell line showing successful immunoprecipitation using the FLAG tag. There was no

significant difference between the eluted GFP and M1V mRNA transcript levels from the HEK Flp *GFP* and *M1V* cell lines (Figure 3.11). Surprisingly, this showed that TRAP allowed for the purification of non-conventionally M1V translating mRNAs and did not specifically enrich for AUG-translating mRNA.

Non-AUG translating M1V mRNAs asssociated to ribosomes were also detected in GRASPS purified samples from the HEK Flp M1V cell model. The amount purified in GRASPS was approximately 75 % less than in TRAP purified samples. This showed that the GRASP purification of ribosomes was, to a lesser extent, also purifying both translating and non-AUG translating mRNA. This mRNA did not correspond to translation initiated from AUG start codons from most of the sequence of GFP (655 of 730 nucleotides of the coding GFP sequence did not contain any conventional AUG start codons), as the primers were designed to recognise an identical region in both *GFP* and *M1V* in the first two thirds of the sequence.

A method to purify ribosomes from human cell lines using a simple fractionation cell lysis and ultracentrifugation to pellet ribosomes has been described (Belin et al., 2010). Using the steps from this methodology and adding additional steps, a new method referred to as GRASPS for the Genome-wide RNA analysis of stalled protein synthesis was developed in this study. This method uses UV-irradiation to crosslink RNA to ribosomes preventing the need to use inhibitors of translation. Cell fractionation using stringent buffer conditions to remove mitochondria and nuclei is performed before loading the post-mitochondrial fraction (the cytoplasm) onto a sucrose cushion. Ultracentrifugation was used to pellet ribosomes preventing the need to use affinity-based purifications. Pelleted ribosomes were then washed and Proteinase K treated to remove ribosomal protein before RNA extraction DNase treatment and subsequent mRNA purification using magnetic beads.

In summary, the Flp In system has been used to build GFP reporter cell models to allow the study of translating and non-AUG translating mRNA. The GRASPS method has been developed to selectively enrich for actively AUG-driven GFP translating mRNA. The GRASPS method removes significantly more non-AUG translating mRNAs compared to the TRAP methodology and selectively enriches for actively translating mRNA, making it a more appropriate method to use to study the translatome.

4. Identification of TDP-43 Q331K ALS transcriptomes and translatome

4.1 Introduction

Alteration of mRNA processing has been highlighted as a major dysregulated pathway in ALS. TDP-43 proteinopathy characterised by the nuclear loss and cytoplasmic accumulation of TDP-43 in motor neurons occurs in approximately 95 % of ALS, both familial and sporadic cases. TDP-43 is implicated in multiple aspects of RNA metabolism including transcriptional regulation, alternative splicing, miRNA processing, axonal transport of mRNA and potential regulation of translation by association with stalled ribosomes in stress conditions. TDP-43 has been reported to bind 4,352 coding transcripts primarily involved in neuronal functions in rat cortical neurons (Sephton et al., 2011) and hundreds of aberrant RNA binding and splicing events have been identified in TDP-43 related ALS (Tollervey et al., 2011; Xiao et al., 2011; Rogelj et al., 2012; Lagier-Tourenne et al., 2012; Highley et al., 2014; D'Alton et al., 2015). A third of transcriptomes are altered in TDP-43 WT and ALS mutation Q331K transgenic mouse models of ALS, with specific alterations due to the Q331K mutation (Arnold et al., 2013). These observations suggest that cellular protein synthesis is dysregulated in TDP-43 ALS and that aberrantly processed mRNAs may be translated into uncharacterised mutant proteins.

However, the majority of these studies highlighted the widespread RNA dysregulation mediated by TDP-43 mutations in ALS in transcriptomes, thereby investigating the entire RNA population within a cell. Investigation of individual cellular compartments or RNA that is being actively translated may better represent protein expression changes and also the directionality of altered biological

processes. Identification of actively translating mRNA in TDP-43 ALS could identify novel proteins and pathways involved in the pathogenesis of ALS. For these reasons TDP-43 ALS provides an interesting background in which to apply the GRASPS methodology.

Furthermore, thousands of RNA alterations have been described in TDP-43 ALS and from these it is hard to identify which alterations are causal and which may be the result of an upstream dysregulation. Inducible cell models allow the expression of TDP-43 to be turned 'on' and 'off' with the addition of Tetracycline. Studies can therefore be done immediately following induction with Tetracycline in order to study the early events in the pathogenesis of TDP-43 ALS.

The Invitrogen Flp In system had been previously used to generate Tetracycline-inducible HEK Flp TDP-43 WT and TDP-43 Q331K cell models in SITraN (Dr Adrian Higginbottom, unpublished data), but these models had not yet been characterised. Characterisation of these models was undertaken prior to identification of transcriptomes and GRASPS-translatomes. The specific ALS mutation Q331K was selected for this study, as within SITraN, research has already been carried out using TDP-43 WT and TDP-43 Q331K transgenic mouse models (Arnold et al., 2013). Also there are post-mortem samples from a variety of both sporadic and familial ALS cases which exhibit TDP-43 proteinopathy. These would provide useful tools to validate any findings from cell models.

The aims of the work described in this chapter were to characterise the behaviour of HEK Flp TDP-43 WT and TDP-43 Q331K cell lines before using cell fractionation and the GRASPS methodology to compare whole cell transcriptomes, cytoplasmic transcriptomes and GRASPS purified translatomes from HEK Flp Sham and TDP-43 Q331K ALS cell models.

4.2 Results

4.2.1 Building and characterisation of non-neuronal TDP-43 ALS inducible cell models

Dr Adrian Higginbottom had already constructed HEK Flp In Sham, TDP-43 WT and TDP-43 Q331K cell models in SITraN. The Sham cell line is a model constructed by transfecting pcDNA5FRT/TO-Empty plasmid into the cell model to serve as an isogenic negative control for experiments. This model has undergone the same construction as the TDP-43 cell models, but does not contain a GOI.

TDP-43 has been shown to bind the 3' UTR of its own mRNA to repress its expression levels and tightly maintain its expression level (Ayala et al., 2011). The integrated transgene does not contain a 3' UTR so that the transgene protein can be expressed and hopefully bind endogenous TDP-43 mRNA in the regulatory feedback loop, in order to avoid over-expression, thus creating a model which is not an over-expression model and allowing for the control of TDP-43 expression using Tetracycline.

The tetracyline inducible expression of TDP-43 protein in the HEK Flp Sham, TDP-43 WT and TDP-43 Q331K was compared over a 48 h period (Figure 4.1). Upon the addition of Tetracycline there is an increase in the amount of total TDP-43 protein in the HEK Flp TDP-43 WT cell model by 6.2 times and an increase in the HEK Flp TDP-43 Q331K cell model of 3.6 times, compared to endogenous TDP-43 levels in the HEK Flp Sham cell model. This showed that TDP-43 expression could be induced by Tetracycline in the HEK cell model system. Endogenous TDP-43 levels are shown in the HEK Flp Sham. There is approximately twice as much TDP-43 is produced in the HEK Flp TDP-43 WT compared to in the HEK Flp TDP-43 Q331K.

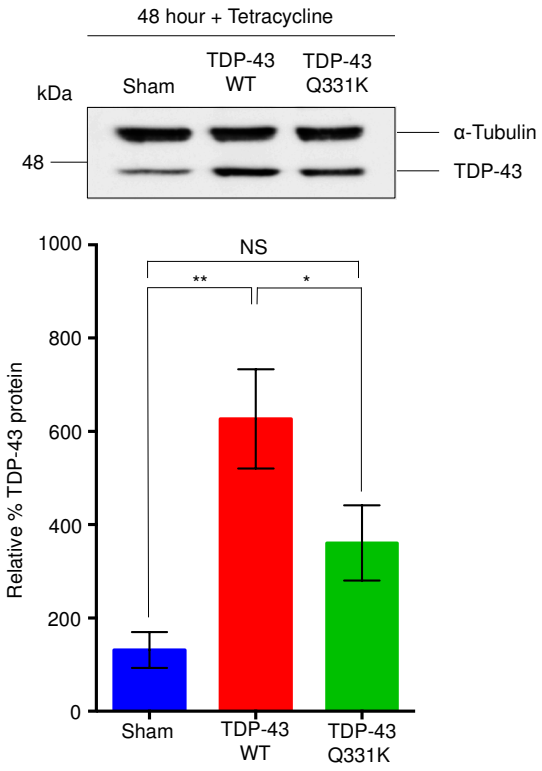


Figure 4.1 Tetracycline induction of TDP-43 protein in HEK Flp cell models
Western blot analysis of HEK Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated for 24 h before a 48 h induction with 10 µg/mL Tetracycline. Graph represents western blot quantification when normalised to α-Tubulin and then to HEK Flp Sham. (Mean ± SEM; One-way ANOVA with Tukey's multiple comparison, *: $p < 0.1$, **: $p < 0.01$, NS = non-significant; N = 4).

TDP-43 transcript levels were then investigated by qRT-PCR (Figure 4.2). Two independent normalisers were used, due to the widespread dysregulation of RNA metabolism in TDP-43 mediated neurodegeneration and the results were very similar. When normalised to U1 snRNA there was a 3.9-fold increase in the amount of TDP-43 mRNA in the HEK Flp TDP-43 WT and a 2.2-fold increase in the amount of TDP-43 mRNA in the HEK Flp TDP-43 Q331K, when compared to HEK Flp Sham. When normalised to β-actin there was a 4.9-fold increase in the amount of TDP-43 mRNA in the HEK Flp TDP-43 WT and a 2.5-fold increase in the amount of TDP-43 mRNA in the HEK Flp TDP-43 Q331K, when compared to HEK Flp Sham. The

difference in the transcript level of TDP-43 WT compared to TDP-43 Q331K is consistent with the differences observed at protein level (Figure 4.1).

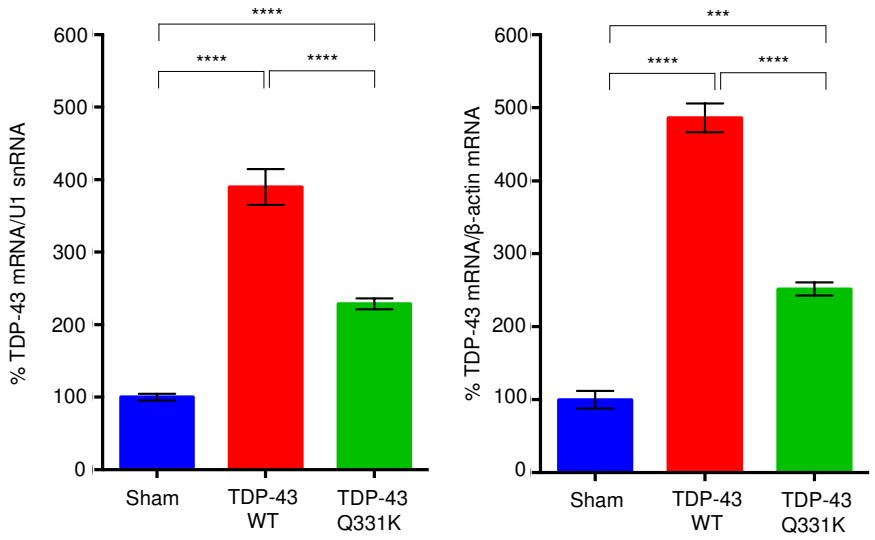


Figure 4.2 TDP-43 transcript levels in HEK Flp cell models
qRT-PCR analysis of HEK Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated for 24 h before a 48 h induction with 10 µg/mL Tetracycline. Quantification was performed using the ΔCt method before normalisation to HEK Flp Sham. (Mean ± SEM; 1way ANOVA with Tukey's multiple comparison, ***: $p < 0.001$, ****: $p < 0.0001$; N = 4).

The growth of HEK Flp Sham, TDP-43 WT and TDP-43 Q331K cells in the presence or absence of Tetracycline was recorded over a fifteen-day period (Figure 4.3). In the absence of Tetracycline the cells increased in number over the fifteen-day period and there were no significant differences in the growth of cells between the different cell lines. In the presence of Tetracycline there was a significant reduction in the growth of HEK Flp TDP-43 WT and HEK Flp TDP-43 Q331K compared to HEK Flp Sham. This reduction was significant by day 9 for the HEK Flp TDP-43 WT and day 12 for the HEK Flp TDP-43 Q331K. This shows that an increase in TDP-43 WT or mutant has a toxic effect. In the presence of Tetracycline the HEK Flp Sham showed no reduction in growth and the cells continued to increase in number over the fifteen-day period.

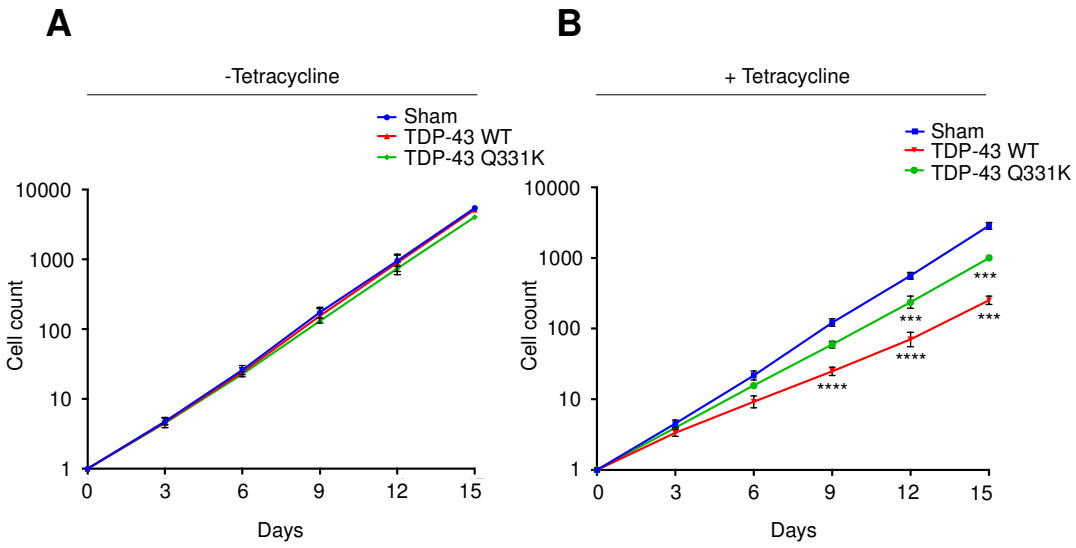


Figure 4.3 Growth curve of HEK Flp cell models Growth of HEK Flp Sham, TDP-43 WT and TDP-43 Q331K cell lines over a fifteen-day period (A) Growth in the absence of Tetracycline (B) Growth in the presence of 10 µg/mL Tetracycline. Each cell line is compared to the HEK Flp Sham cell line and only significant differences are labelled. (Mean ± SEM; 2way ANOVA with Tukey's multiple comparison, ***: $p < 0.001$, ****: $p < 0.0001$; N = 3).

MTT cell proliferation assays were carried out in the absence and presence of Tetracycline for 48 h, 7 days and 14 days (Figure 4.4). As expected, there is no significant difference in proliferation of all cell lines grown in the absence of Tetracycline. The proliferation of all cells decreases over time when grown in Tetracycline. There is a significant reduction in the HEK Flp TDP-43 WT compared to HEK Flp Sham after 48 h in the presence of Tetracycline, which continues at 7 days and 14 days. There is no significant reduction in the proliferation of HEK Flp TDP-43 Q331K compared to HEK Flp Sham when grown in the presence of Tetracycline.

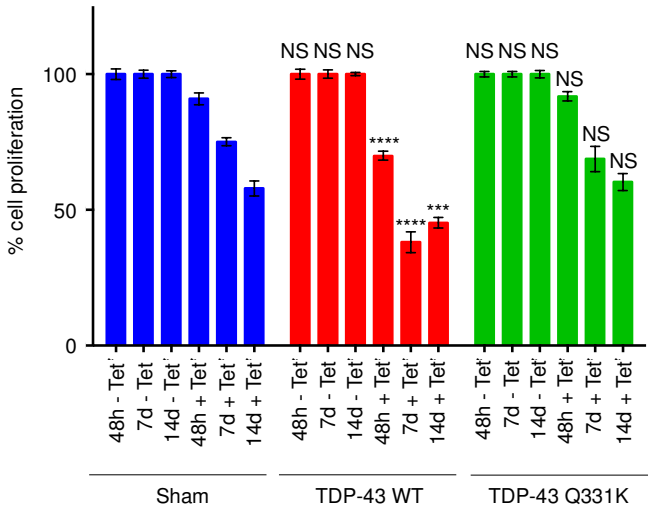


Figure 4.4 Cell proliferation of HEK Flp cell models
MTT assays on HEK Flp Sham, TDP-43 WT and TDP-43 Q331K cell lines grown in the presence (+ Tet) or absence (-Tet) of 10 µg/mL Tetracycline for 48 h, 7 days (7d) and 14 days (14d). Each bar of TDP-43 WT and TDP-43 Q331K is compared to corresponding bar from Sham cell line. (Mean ± SEM; 2way ANOVA with Tukey's multiple comparison, ***: $p < 0.001$, ****: $p < 0.0001$, NS: non-signifcant; N (independent experiments, 6 wells per experiment) = 3).

HEK Flp TDP-43 cell models were successfully characterised. Upon the addition of Tetracycline the cell models exhibit an increase in the amount of TDP-43 protein produced, which causes a reduction in the cell growth and cell proliferation.

4.2.2 Cell fractionation and sample preparation for Next Generation Sequencing

As described in Chapter 3, the GRASPS method allows for the selectively enriched purification of AUG translating mRNA. In order to compare the translatome to the cytoplasmic transcriptome, cellular fractionation was optimised in order to obtain cytoplasmic lysates for RNA extraction.

Cellular fractionation was performed on HEK Flp Sham and TDP-43 Q331K cells and cytoplasmic RNA was extracted. Successful cytoplasmic fractionation was confirmed by comparing total and cytoplasmic lysates for the presence of the nuclear protein SSRP1 (Figure 4.5). Nuclear SSRP1 is present in total fractions, but minimal in

cytoplasmic fractions indicating successful cytoplasmic fractionation without significant nuclear contamination.

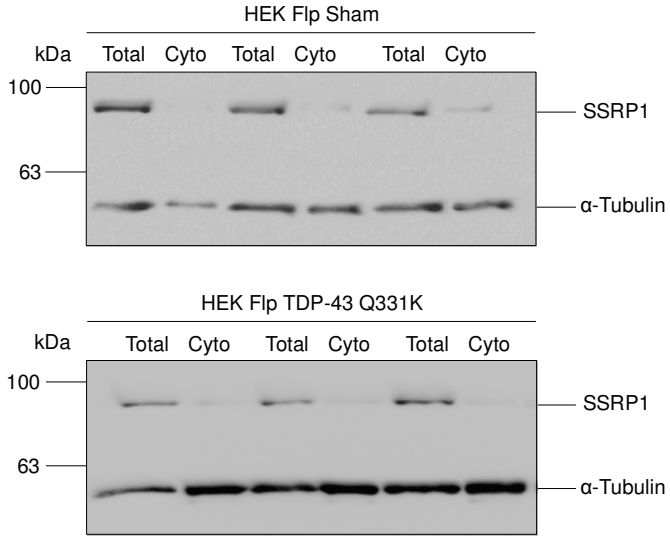


Figure 4.5 Cytoplasmic fractionation of HEK Flp Sham and TDP-43 Q331K
Western blot to assess nuclear contamination with nuclear protein SSRP1 in cytoplasmic fractions obtained from HEK Flp Sham and TDP-43 Q331K cells grown in the presence of 10 µg/mL tetracyline for 48 h. SSRP1 protein bands can be observed in total fractions (Total) but are minimal in cytoplasmic (Cyto) fractions indicating a clean cytoplasmic fractionation.

In order to study TDP-43 Q331K transcriptomes and the translatome Next Generation RNA sequencing experiments were designed in order to compare the whole cell transcriptome, cytoplasmic transcriptome and translatome from HEK Flp Sham and HEK Flp TDP-43 Q331K cells following a 48 h induction with Tetracycline. The experimental design is shown in Table 4.1.

Induced HEK Flp Sham cells would be used as a control, as these cells had undergone the same recombination events in construction of the cell line and the same treatment with Tetracycline, thus removing any changes, which may be the result of cell line manipulation or due to induction with Tetracycline. A 48 h timepoint was selected to analyse the early events in RNA metabolism dysregulation following induction with Tetracycline.

Table 4.1 RNA sequencing experimental design Total RNA, cytoplasmic RNA and GRASPS purified mRNA were extracted from HEK Flp Sham and TDP-43 Q331K cells following a 48 h induction with Tetracycline. (rep = replicates).

Whole cell transcriptome (WCT)	Cytoplasmic transcriptome (CyT)	GRASPS translatome (GRASPS)
Sham control (3 rep) **TDP-43 Q331K (3 rep)** **Total = 6 total RNA samples**	Sham control (3 rep) TDP-43 Q331K (3 rep) Total = 6 total RNA samples	Sham control (3 rep) TDP-43 Q331K (3 rep) Total = 6 poly(A)$^+$ RNA samples

RNA was extracted in independent triplicates from: whole cell extracts to generate the whole cell transcriptomes, cytoplasmic lysates to generate the cytoplasmic transcriptomes and using the GRASPS methodology to generate the translatomes.

The yield and quality of RNA extracted was assessed before sending the samples for RNA sequencing. The Agilent 2100 Bioanalyser and RNA 6000 Nano and Pico kits assess RNA quality using electrophoretic separation on microfabricated chips to separate and detect RNA samples using laser induced fluorescence detection. There is a decrease in the 18S to 28S ribosomal RNA (rRNA) band ratio as degradation proceeds. The RNA integrity number standardises whole-cell RNA quality assessment by taking into account the whole electropherogram trace and calculating an RNA integrity number (RIN). RIN allows for the classification of eukaryotic total RNA using a scale from 1-10, where 1 indicates complete degradation and 10 indicates completely intact RNA. Sample electropherograms from both intact and varying degrees of degradation are shown in Figure 4.6. Intact RNA samples show three distinct peaks from the ladder, 18S and 28S ribosomal peaks. There is clear separation between the 18S and 28S ribosomal peaks. More degraded RNA samples show additional peaks and less clear separation of the 18S and 28S ribosomal peaks.

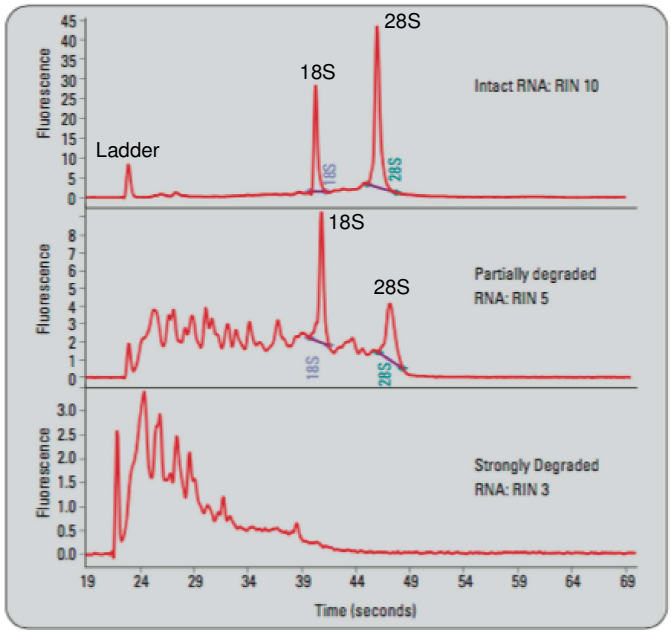


Figure 4.6 Sample electropherograms from Agilent 2100 Bioanalyser manual
The molecular weight marker (ladder) and peaks corresponding to the 18S and 28S ribosomal RNA have been labelled on the example electropherograms, which show intact, partially degraded and strongly degraded RNA samples. Time in seconds is plotted along the x-axis with RNA concentrations as defined by arbitrary units of fluorescence [FU] plotted along the y-axis.

Nucleic acids have absorbance maxima at 260 nm and proteins at 280 nm. The ratio of absorbance at these wavelengths ($A_{260/280}$) is used as an indication of RNA purity. A ratio of approximately 2.0 is considered for 'pure' RNA. An absorbance at 230 nm is considered to be the result of contamination and therefore the ratio between absorbance at 260 nm and 230 nm ($A_{260/230}$) are also calculated. This value is often higher than the $A_{260/280}$ and between 2.0 and 2.2.

The yield and quality of RNA for whole cell and cytoplasmic RNA extracted are shown in Table 4.1. All extracted RNA has a RIN value of over 8, showing that RNA samples are intact and not degraded. The absorbance ratios $A_{260/280}$ are all approximately 2.0 and the $A_{260/230}$ are also approximately 2.0. This shows high nucleic acid purity and low level of protein contamination within the samples.

Table 4.2 Quality and yield of total and cytoplasmic RNA

RNA was extracted from HEK Flp In Sham and TDP-43 Q331K cells following induction with 10 µg/mL Tetracycline for 48 h. The RNA concentration was calculated using a Nanodrop 1000 spectrophotometer. The yield was calculated by multiplying the total volume extracted with the concentration. The quality was assessed using an Agilent 2100 Bioanalyser to measure RNA integrity. RNA integrity numbers (RIN) were scored based upon a scale from 1-10, where 0 corresponds to completely degraded RNA and 10 corresponds to completely intact RNA.

Sample	RNA concentration ng/µL	RNA yield µg	$A_{260/280}$	$A_{260/230}$	RIN
Sham Total 1	1503.60	30.07	2.06	2.12	8.80
Sham Total 2	1733.10	34.66	2.05	2.12	9.40
Sham Total 3	1734.10	34.68	2.06	2.10	9.40
Q331K Total 1	524.20	10.48	2.04	1.99	9.70
Q331K Total 2	1074.50	21.49	2.05	2.08	9.40
Q331K Total 3	783.60	15.67	2.07	2.12	9.20
Sham Cyto 1	121.9	2.68	2.12	2.09	8.00
Sham Cyto 2	145.4	2.12	1.91	1.91	9.20
Sham Cyto 3	168.5	3.03	2.00	2.00	9.10
Q331K Cyto 1	99.3	1.79	1.96	1.96	9.00
Q331K Cyto 2	92.3	1.66	1.99	1.99	9.30
Q331K Cyto 3	106.3	1.91	1.96	1.96	9.10

A visual assessment of RNA integrity is shown in electropherograms produced by the Agilent 2100 Bioanalyser electropherograms. Electropherograms for total RNA, cytoplasmic RNA and GRASPS purified mRNA are shown in Figure 4.7. Total and cytoplasmic RNA are all very similar and show distinct separation of the 18S and 28S rRNA peaks. There is small peak following the ladder that corresponds to small RNAs. There are very few peaks due to degradation products between the small RNA, 18S and 28S rRNA peaks. This indicates the samples contain intact RNA with very little degradation. The electropherogram for GRASPS purified mRNA, show a different profile to the total and cytoplasmic RNA. This is likely to be because the GRASPS purified samples consist of purified poly-adenylated mRNA, whereas total and cytoplasmic samples contained total RNA. The electropherograms produced by the GRASPS purified mRNA show a ladder at the start followed by a wide bell shaped curve and no distinct 18S and 28S rRNA peaks. This curve is likely due to the enrichment of mRNA and depletion of rRNA. The height of the curve is much lower in fluorescence units for the third replicate of HEK Flp TDP-43 Q331K

GRASPS purified mRNA as the sample was loaded onto a nanochip, rather than a picochip like the other GRASPS mRNA samples.

The yield and quality of GRASPS purified total RNA and mRNA is shown in Table 4.3. Only GRASPS purified mRNA was analysed using the Agilent 2100 Bioanalyser. No RIN values can be calculated for mRNA as the RIN values are generated using the entire electropherogram and the ratio of 18S and 28S rRNA peaks. The absorbance ratios $A_{260/280}$ are all approximately 2.0 and the $A_{260/230}$ are also approximately 2.0. This shows high nucleic acid purity and does not indicate contamination within the samples.

Table 4.3 Quality and yield of GRASPS total RNA and purified mRNA

RNA was extracted from HEK Flp In Sham and TDP-43 Q331K cells following induction with 10 µg/mL Tetracycline for 48 h. The RNA concentration was calculated using a Nanodrop 1000 spectrophotometer before (Total RNA conc) and after mRNA purification (mRNA conc). The yield was calculated by multiplying the total volume extracted with the concentration.

	Total RNA conc ng/µL	Total RNA yield µg	Total RNA $A_{260/280}$	Total RNA $A_{260/230}$	mRNA conc ng/µL	mRNA yield ng
Sham 1	3638.8	109.2	2.00	2.10	8.9	151.3
Sham 2	3952.8	118.6	2.00	2.10	10.8	183.6
Sham 3	3767.6	112.1	2.00	2.10	13.9	236.3
Q331K 1	3363.9	100.9	2.00	2.10	11.1	188.7
Q331K 2	3499.2	105.0	2.00	2.10	13.1	222.7
Q331K 3	2588.7	57.0	2.02	1.61	45.2	768.4

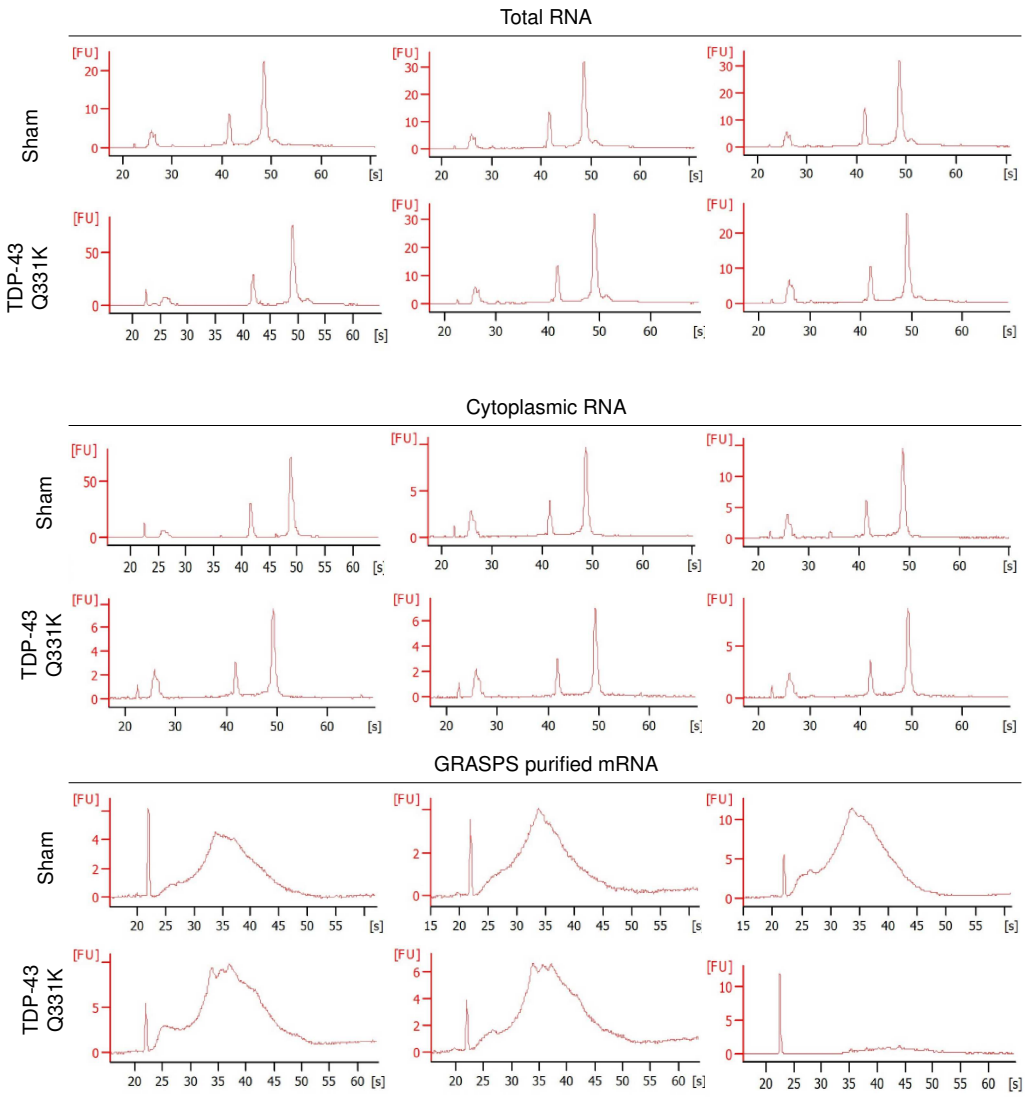


Figure 4.7 RNA Integrity of HEK Flp Sham and TDP-43 Q331K for RNA NGS Electropherograms produced by Agilent 2100 Bioanalyser of RNA extracted from HEK Flp Sham or TDP-43 Q331K after induction with 10 µg/mL tetracyline for 48 h. Time in seconds is plotted along the x-axis with RNA concentrations as defined by arbitrary units of fluorescence [FU] plotted along the y-axis. Total and cytoplasmic RNA samples were analysed using a nanochip, whereas GRASPS purified mRNA samples were analysed using a picochip. Scale and distribution is different for GRASPS purified mRNA sample 3, this is due to sample being analysed using a nanochip rather than a picochip.

4.2.3 Next Generation RNA sequencing

Total RNA extracted from whole-cell or cytoplasmic fractions and poly-adenylated RNA isolated from GRASPS-purified ribosomes were sequenced on *Illumina HiSeq 2500* at the Centre for Genomic Research at the University of Liverpool. Briefly, RNASeq libraries were prepared using the ScriptSeq™ Complete Gold Kit protocol to deplete RNA for rRNA (for whole-cell and cytoplasmic transcriptomes) and prepare

strand-specific libraries from rRNA-depleted RNA samples. For GRASPS samples, the same library kit was used, but no rRNA depletion was performed. Indexed libraries were subjected to paired-end sequencing (approximately 2x 100 bp) multiplexed as three libraries per lane of the Illumina HiSeq platform to generate data of in excess of 120 millions clusters/lane. Trimming of sequencing reads and initial quality check prior to the generation of fastq files were performed by the Centre for Genomic Research.

The raw Fastq files are trimmed for the presence of Illumina adapter sequences using Cutadapt version 1.2.1 (Martin, 2011). The option -O 3 was used, so the 3' end of any reads, which match the adapter sequence for 3 bp or more are trimmed. The reads are further trimmed using Sickle version 1.200 (https://github.com/najoshi/sickle) with a minimum window quality score of 20. Reads shorter than 10 bp after trimming were removed. If only one of a read pair passed this filter, it is included in the R0 file. The output files from Cutadapt and Sickle are available on link on attached CD-ROM (See C1. Cutadapt and Sickle output files). Statistics were generated using fastq-stats from EAUtils (https://expressionanalysis.github.io/ea-utils/).

The initial report from the Centre is attached below (Figure 4.8). The bar chart in Figure 4.8 A shows the number of paired reads at 60 million reads per sample or above. There is a very small number of singlet reads and the number of reads discarded due to adapter contamination or poor quality is also very small. This shows that the RNA samples were high quality and the majority of reads could be read. It is common for a small read to consist mostly of adapter-derived sequence. The reads range between 60-100 nucleotide reads, showing that more than adapter-derived sequence is being read. The boxplots in Figure 4.8 B show the distribution of trimmed reads for forward, reverse and singlet reads.

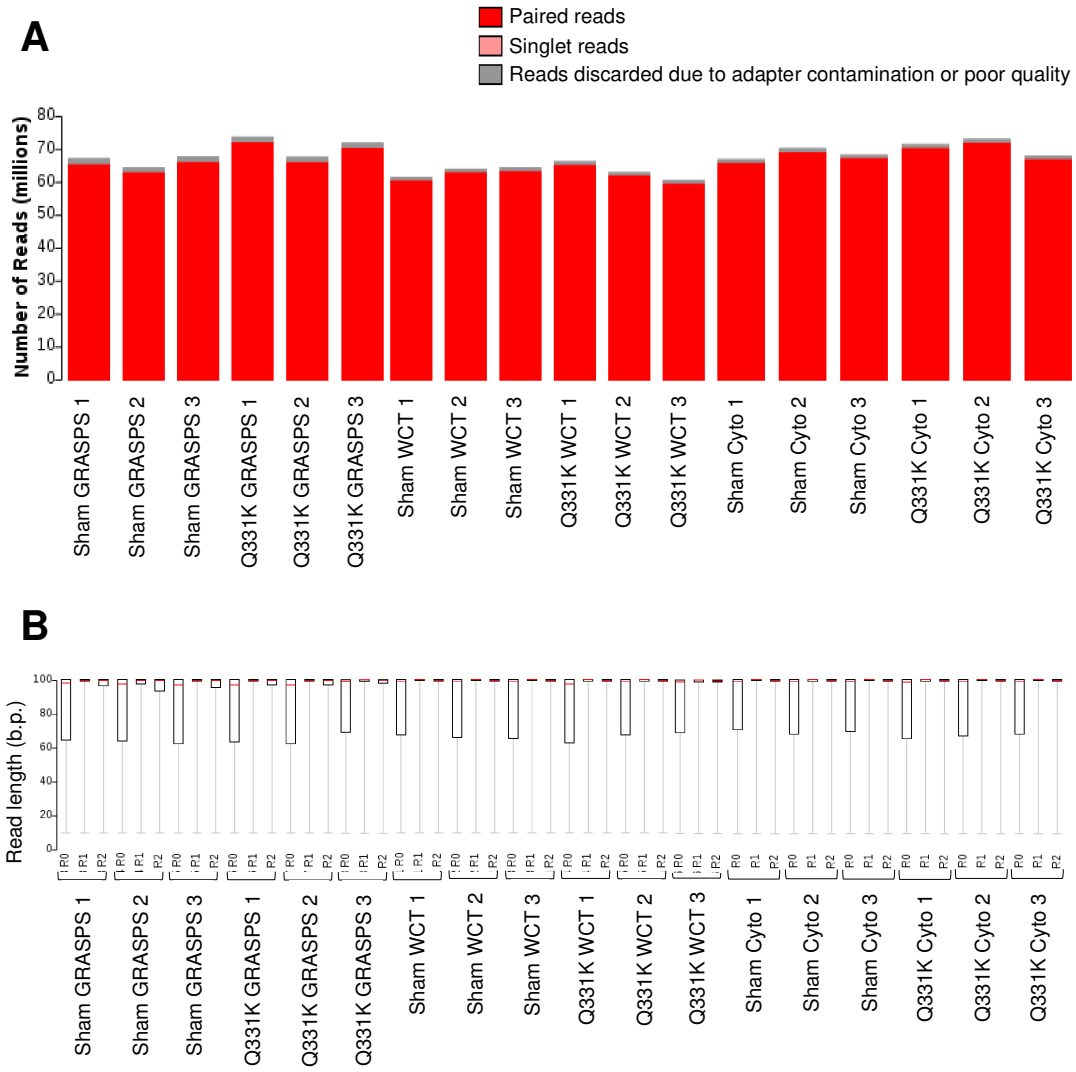


Figure 4.8 RNA sequencing report from Centre for Genomic Research at the University of Liverpool

(A) Diagram shows the total number of reads obtained per sample. (B) Box plot shows the distribution of trimmed read lengths for the forward (R1), reverse (R2) and singlet (R0) reads. Red line indicates median length. Box indicates interquartile range. Whiskers indicate minimum and maximum read lengths. All read lengths are over 60 nucleotide reads indicating that more than the adaptor sequence is being read. (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of the Centre for Genomic Research, University of Liverpool]

4.2.4 Bioinformatics analysis overview of RNA seq data sets

Typically, 60-70 millions reads per sample were mapped to the human genome (~60x coverage) using Bowtie (Langmead et al., 2009) and transcripts quantified using Bitseq (Glaus et al., 2012) and Limma (Ritchie et al., 2015). BioMart was used

to convert transcripts into Ensembl transcript identities (Kasprzyk, 2011). The percentage of sequence mapped to the genome is shown in Table 4.4. On average for the samples, 88 % of sequence reads could be aligned to the human genome. These results show successful and excellent quality mapping at genome wide level for the transcriptomes and translatome.

Table 4.4 Table to show the percentage of the mapping of sequence reads to the genome
(WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

Sample	Replicate	Mapping of sequence reads to the genome %	Average alignment %
WCT	Sham 1	92.69	92.88
	Sham 2	92.93	
	Sham 3	93.01	
	Q331K 1	90.06	91.90
	Q331K 2	92.96	
	Q331K 3	92.69	
CyT	Sham 1	90.58	92.17
	Sham 2	93.05	
	Sham 3	92.88	
	Q331K 1	92.77	92.96
	Q331K 2	92.97	
	Q331K 3	93.15	
GRASPS	Sham 1	80.35	77.79
	Sham 2	79.11	
	Sham 3	73.92	
	Q331K 1	81.78	81.85
	Q331K 2	78.95	
	Q331K 3	84.81	

4.2.5 Differential gene expression into HEK TDP-43 ALS inducible cell models

The size of the transcriptome and translatome before looking at gene differential expression was compared. All transcripts were mapped to the genome and coding genes were identified and counted only once, regardless of number of transcripts identified. Coding genes identified from the whole cell transcriptome (WCT), cytoplasmic transcriptome (CyT) and GRASPS translatome (GRASPS) were plotted in a Venn diagram (Figure 4.9). There were 11,280 genes found to be common between the whole cell transcriptome, cytoplasmic transcriptome and translatome.

This analysis includes all transcripts before any filtering and shows that at the start of the analysis all three categories contain the same genes. There is no bias on the genome, with all three groups containing on average the same number of identified genes indicating the uniform transcript coverage across data sets.

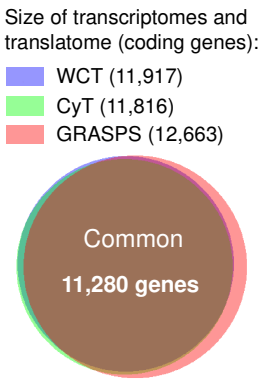


Figure 4.9 Coding genes identified from WCT, CyT and GRASPS translatome
All transcripts mapped to the genome were included in the analysis and each coding gene identified was only counted once. Blue indicates genes identified from whole cell transcriptome (WCT), green indicates genes identified from cytoplasmic transcriptome (CyT) and red indicates genes identified from the GRASPS translatome (GRASPS).

An overview of the bioinformatics analysis carried out is shown in Figure 2.2. Bayesian inference of transcripts from sequencing data (BitSeq) is used to estimate the transcript expression levels (Glaus et al., 2012). Transcript expression levels are quantified, taking into account the biological variance between replicates for each condition, to generate a mean expression level for each transcript per condition. This is very important for RNA sequencing and differential expression analysis. During eukaryotic transcription genes can be spliced into different transcripts, which share parts of their sequence. RNA sequencing is sequencing the transcripts of genes and can distinguish between individual transcripts of genes. Therefore to estimate transcript expression levels, probabilistic methods have to be used to estimate transcript levels for reads aligning to a shared subsequence. BitSeq uses a probabilistic model to generate transcript expression in order to analyse expression

changes between conditions. Following BitSeq, the transcript list has been filtered to only include those read in more than two replicates and each transcript has been quantified to give a whole count. The whole count is an absolute count estimated by BitSeq using variational Bayes Methods and probability distribution. It is an absolute count as it reflects the amount of expression, rather than a saturated value from fluorescence intensities, as in a microarray.

LIMMA and *edgeR* analyses were performed to convert the read counts into log2 counts per million (logCPM) (Ritchie et al., 2015). LIMMA converts the number of counts to the log scale and estimates the mean variance. This is done by normalising the number of counts to the total number of counts. Transcripts were filtered using three criteria: p value of less than 0.05, a false discovery rate (FDR) of less than 0.2 and a fold change of 2 or more. The p value of 0.05 was chosen so that it is relevant to all samples (WCT, CyT and GRASPS). For GRASPS a lower p value could be selected, however, this lower p value would not also be applicable to WCT and CyT. The FDR chosen means there is less than 20 % likelihood of a false discovery. The p value and FDR value were selected, as these are the generally accepted values in RNA sequencing analysis. These criteria make this a stringent and robust analysis.

In order to determine which genes were differentially expressed, genes were filtered using three criteria: p value of less than 0.05, a false discovery rate (FDR) of less than 0.2 and a fold change of 2 or more. TDP-43 Q331K transcriptome and translatome changes were normalised to the isogenic induced Sham control line.

MA-plots were generated to compare differentially expressed transcripts from WCT, CyT and GRASPS in TDP-43 Q331K ALS (Figure 4.10). The number $\log_2$ fold change of a transcript is plotted against the average $\log_2$ counts per million. The scatter plot is bigger for GRASPS translatome compared to the WCT and CyT,

showing that the GRASPS translatome reads more differentially expressed transcripts than the WCT and CyT. The log_2 FC range is greater for the GRASPS translatome, showing that differentially expressed transcripts have greater fold changes in the GRASPS translatome compared to WCT and CyT. The average log_2 CPM is also greater in the GRASPS translatome compared to the WCT and CyT, indicating that the transcripts identified in the GRASPS translatome have a higher count number.

Venn diagrams of DEG within each category (WCT, CyT and GRASPS) were produced to depict the number of genes up- and down-regulated that are unique or common to each category at different fold changes (Figure 4.12 A). The GRASPS translatome was able to identify the most DEG at FC of 2, 3 and 8.

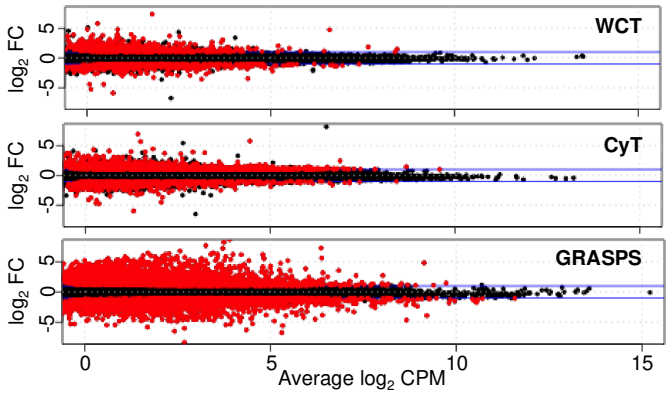


Figure 4.10 MA plots to show differentially expressed transcripts from WCT, CyT and GRASPS in TDP-43 Q331K ALS
The log_2 fold change (FC) of a transcript is plotted against the average log_2 counts per million (CPM). Transcripts with a FC less than 2 are shown in black and are considered non-significant. Differentially expressed transcripts are shown in red and have a fold change greater than 2. ($p < 0.05$, $log_2FC \geq 1$) (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

DAVID looks at the gene ontology enrichment by assessing the number of genes involved in a biological process. The more enriched a process, the greater number of alterations to genes in that pathway and subsequently the greater the enrichment

score. The result of greater enrichment suggests that this process plays a role in disease.

Gene ontology analysis was carried out using the functional annotation chart module DAVID to assess the significance of biological process enrichment scores. This groups genes associated with a given functional category and counts the number of differentially expressed genes in the WCT, CyT and GRASPS translatome within that category assigning them a fold enrichment score. The gene ontology analysis of differentially expressed genes using DAVID in the TDP-43 Q331K WCT, CyT and GRASPS translatome are listed in Appendix 2 Tables A1-A3. The list generated from the GRASPS translatome is greater (at least 3-fold) than the WCT and CyT, showing that GRASPS method provides more information.

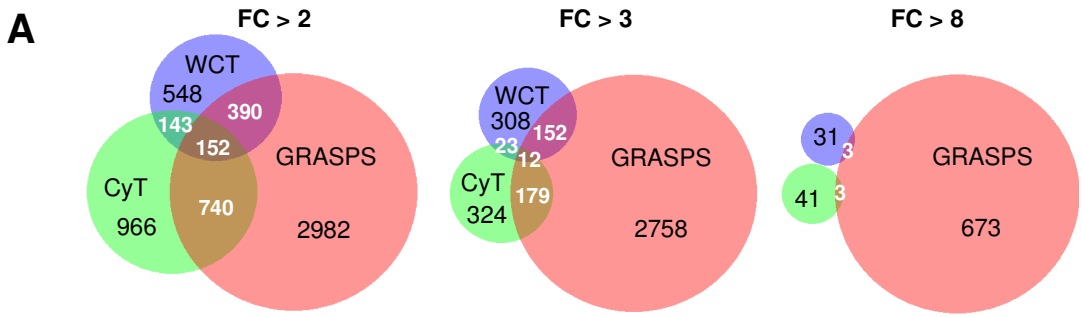
Enriched biological processes identified by DAVID were further grouped into broader categories based upon biological processes. The gene ontology terms from the DAVID analysis and genes involved in each pathway are listed in Appendix 2 Tables A4-A6. In total 16 broader categories were identified for the WCT, 18 for the CyT and 42 for the GRASPS translatome.

These broader categories and the enrichment score from the gene ontology term with the highest enrichment score were put into a table, listing the enrichment scores and categories identified at a fold change of 2 ($\log_2$ FC >1), approximately 3 ($\log_2$ FC > 1.5) and 8 ($\log_2$ FC > 3) (Figure 4.12 B). At a fold change of 2, the same functional pathways (protein transport/localisation, protein degradation/ubiquitin, RNA splicing/processing, DNA damage/repair and cell cycle) are enriched in WCT, CyT and GRASPS. The CyT identifies 3 pathways that are not enriched in WCT (cell death/apoptosis, phosphorylation and RNAPII transcription). These pathways are also enriched in the GRASPS translatome. An additional five pathways are enriched

in the GRASPS translatome (nucleocytoplasmic transport, hexose metabolic process, aeriobic respiration, tRNA aminoacylation and neuron differentiation).

At a fold change of 3, only one category is identified in the WCT (protein transport/localisation). At a fold change of 3 the RNA splicing/processing category is no longer enriched in the CyT, but all other categories previously enriched at a fold change of 2 remain. The GRASPS translatome is the only method used which is able to identify functionally enriched pathways at a fold change of 8.

Venn diagrams for some of of the DEG in the pathways identified using DAVID at a fold change of greater than 2 were constructed (Figure 4.12 C). These diagrams show in all of the pathways displayed that GRASPS identifies more DEG than the WCT and CyT. This shows that the GRASPS translatome is not biased to a single pathway, as the GRASPS method provides more data in all of the pathways.



B

Enrichment scores for:	FC > 2			FC > 3			FC > 8		
Protein transport/ localisation	4.1	5.9	4.9	1.8	1.5	5.3	-	-	-
Protein degradation/ ubiquitin	3.3	4.0	14.4	-	4.7	12.9	-	-	2.6
RNA splicing/ processing	3.2	1.4	4.2	-	-	3.5	-	-	4.7
DNA damage/ repair	3.8	5.6	10.0	-	1.6	4.6	-	-	2.8
Cell cycle	1.5	6.0	8.9	-	8.3	9.3	-	-	3.1
Cell death/ apoptosis	-	3.1	6.1	-	2.7	4.8	-	-	3.6
Phosphorylation	-	2.4	3.8	-	2.1	4.5	-	-	3.3
RNAPII transcription	-	1.8	2.4	-	2.6	2.4	-	-	1.6
Nucleocytoplasmic transport	-	-	1.9	-	-	1.6	-	-	1.6
Hexose metabolic process	-	-	2.5	-	-	1.9	-	-	-
Aerobic respiration	-	-	2.0	-	-	2.6	-	-	-
tRNA aminocylation	-	-	1.7	-	-	3.5	-	-	-
Neuron differentiation	-	-	1.9	-	-	-	-	-	-

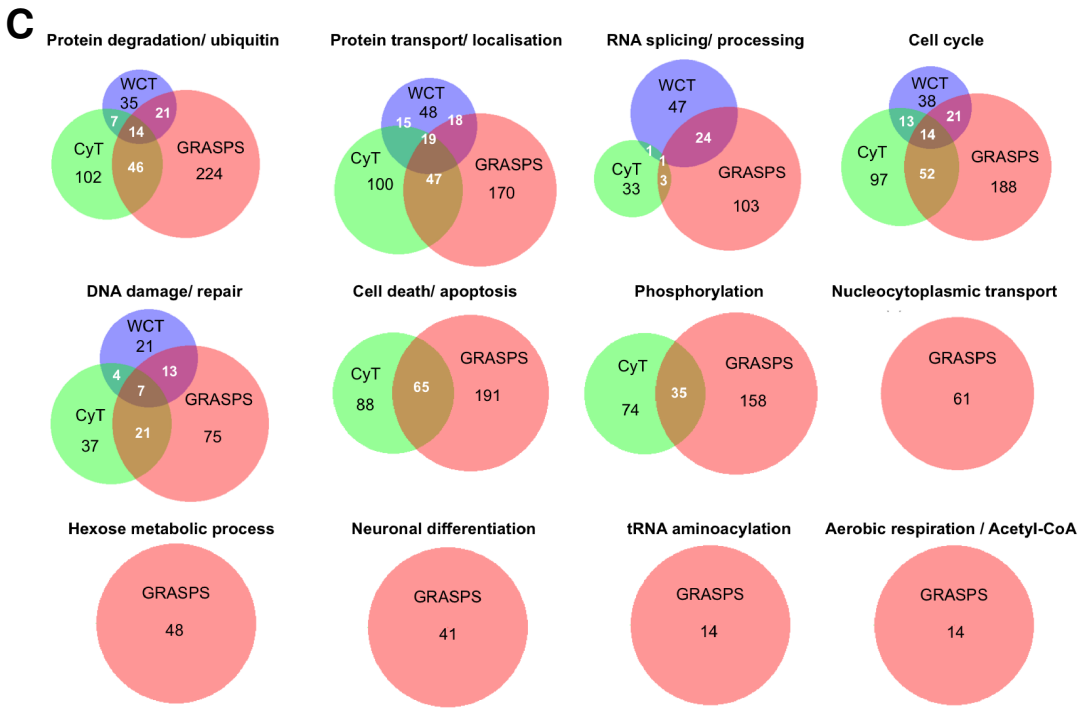


Figure 4.11 Differential gene expression and functional clustering using DAVID
(A) Venn diagrams to depict the total number of up- and down-regulated DEG that are unique or common to each category at fold change (FC) of 2, FC 3 and FC 8. ($p < 0.005$) (B) Functional clustering using DAVID. Gene ontology terms were further categorised into broader groups and awarded a gene enrichment score based upon the gene ontology term with the highest enrichment score within that group at FC 2 ($\log_2 FC > 1$), FC 3 ($\log_2 FC > 1.5$) and FC 8 ($\log_2 FC > 3$). ($p < 0.005$) (C) Venn diagrams displaying some of the DEG grouped by the pathways identified by DAVID. (FDR < 0.2, FC >2, p< 0.05). Throughout the figure whole cell transcriptome (WCT) is shown in blue; cytoplasmic transcriptome (CyT) is shown in green; GRASPS translatome (GRASPS) is shown in red. The broader grouping and complete gene lists for all categories are shown in Appendix 2 Table A4, A5 and A6 for WCT, CyT and GRASPS translatome respectively.

All the genes involved in the nuclear export of mRNA based on the literature and their relative fold change from WCT, CyT and GRASPS translatome were illustrated in a Pheatmap (Figure 4.13). This pathway was selected as altered nucleocytoplasmic transport has been reported with cytoplasmic protein aggregation like that observed in TDP-43 proteinopathy (Woerner et al., 2016) and also linked to C9ORF72-ALS (Jovičić et al., 2015; Freibaum et al., 2015; Ke Zhang et al., 2015).

The majority of the transcripts in the WCT and CyT are shown in yellow indicating no change was detected in the WCT or CyT. However, the GRASPS translatome was able to identify both up and down regulated genes with much larger fold changes between -16 and 16. Moreover, any few detected changes identified in the WCT and CyT are much smaller fold changes compared to the GRASPS translatome. The fold change for WCT, CyT and GRASPS rarely correlates. The most upregulated transcript in the GRASPS translatome in this pathway is karyopherin (importin) beta subunit 1 (KPNB1) with a fold change of over 16 and exportin 1 (XPO1) is the most downregulated gene with a fold change of -16.

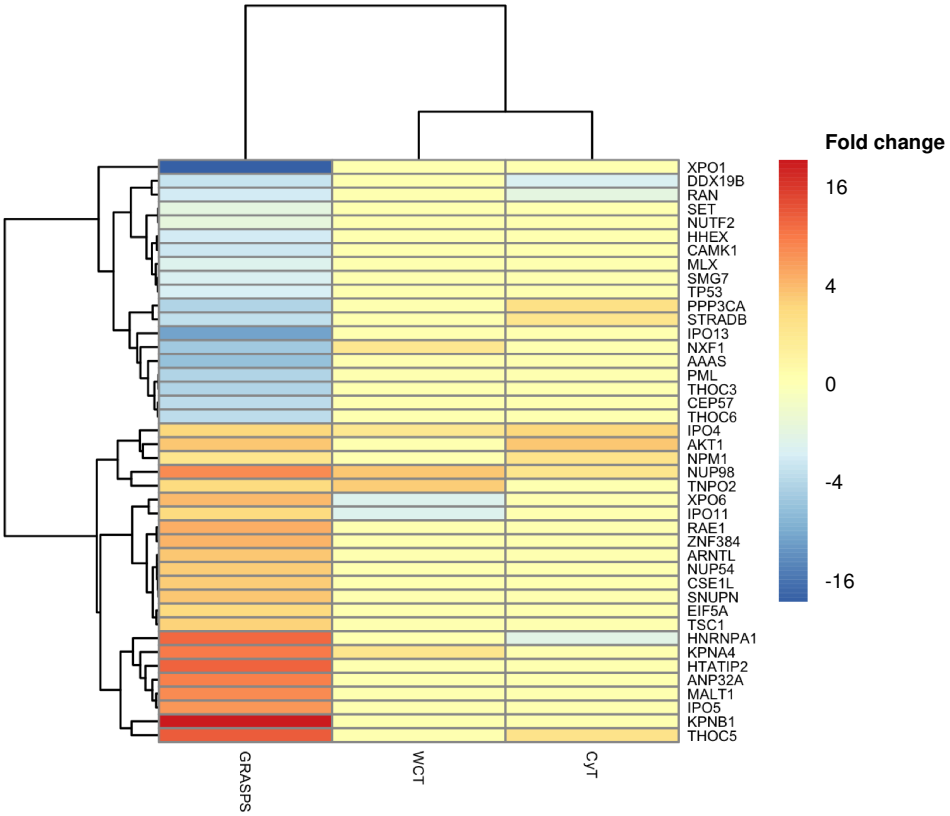


Figure 4.12 Pheatmap of DEG identified by functional clustering using DAVID involved in nuclear and cytoplasmic transport from WCT, CyT and GRASPS
Functional clustering of GRASPS translatome using DAVID identified the nuclear and cytoplasmic transport pathway as being significantly enriched. The DEG from this identified pathway and their relative fold change from the WCT, CyT and GRASPS translatome are shown in the Pheatmap. The colour indicates the fold change, which is shown on the scale bar. (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

A similar Pheatmap was also generated using the DEG list for genes involved in neuronal differentiation and their fold changes from WCT, CyT and GRASPS (Figure 4.13). This pathway was selected as ALS is a disease of motor neurons and so it is likely that induction of TDP-43 could have an effect on neuronal differentiation.

The majority of the transcripts in the WCT and CyT are shown in yellow indicating no change is detected. The GRASPS translatome is able to identify both up and downregulated genes with much larger fold changes between - 256 and 64. The most up regulated transcript in the GRASPS translatome is actin-related protein 3

(ACTR3) with a fold change greater than 16 and cellular retinoic acid-binding protein 1 (CRABP1) is the most downregulated transcript with a fold change of - 256.

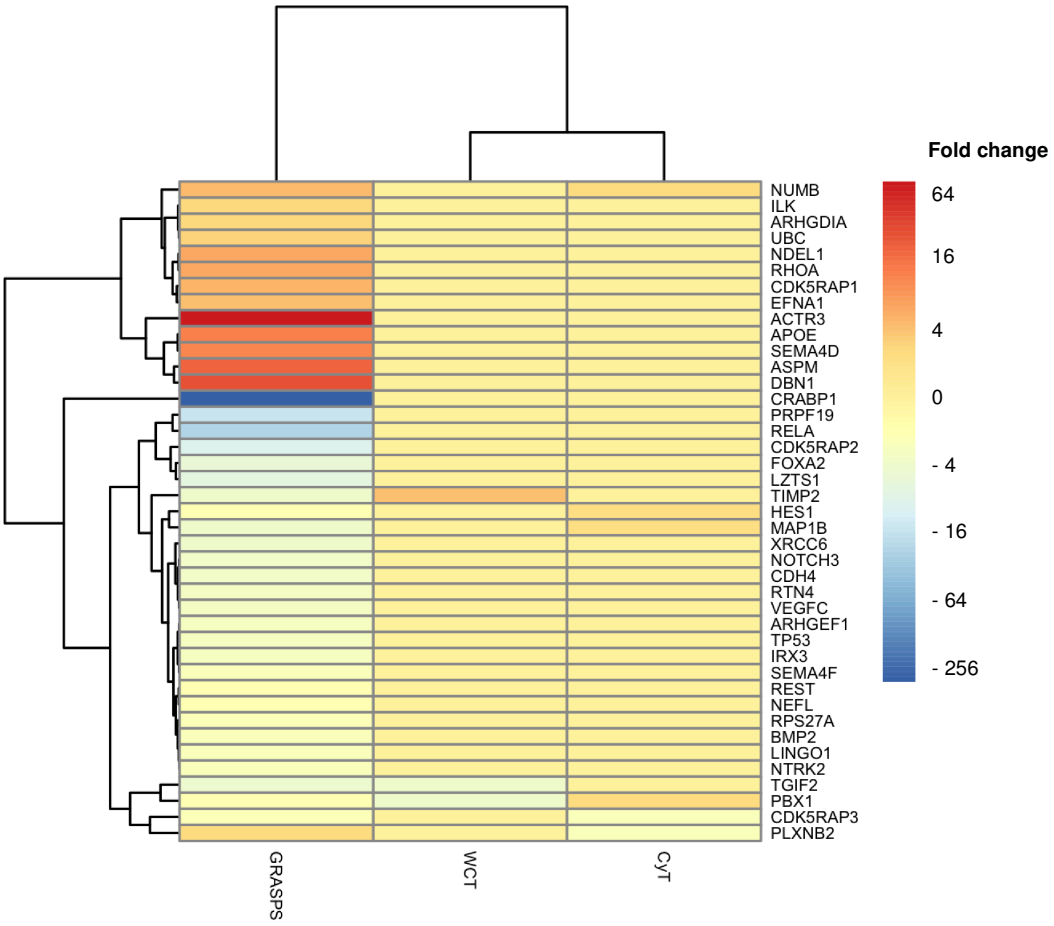


Figure 4.13 Pheatmap of DEG involved in neuronal differentiation from WCT, CyT and GRASPS

The DEG from this identified pathway and their relative fold change from the WCT, CyT and GRASPS translatome are shown in the Pheatmap. The colour indicates the fold change, which is shown on the scale bar. (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

The PANTHER overrepresentation test (version 10.0 released 2015-05-15) from the Gene Ontology (GO) database (http://www.pantherdb.org/) was also used to statistically assess significant enrichment of biological processes (PANTHER GO-Slim Biological Process) and validate the DAVID analysis.

PANTHER is an overrepresentation test so compares the biological process under investigation to an ‘expected’ number of genes which you would expect to be present

at random on the basis of a reference list. The reference is the human genome and the expected value is based upon the number of genes involved in a particular process compared to the number of genes in the human genome. If the number of genes involved in a process under investigation is higher than this expected number, it is overrepresented and if lower underrepresented. The p value is then used to determine whether it is significant.

The gene ontology analysis of differentially expressed genes using PANTHER in the TDP-43 Q331K WCT, CyT and GRASPS translatome are listed in Appendix 2 Tables A7-A9. From the DEG list, PANTHER identified 34 biological processes in the WCT, 49 in CyT and 69 in the GRASPS translatome that showed statistically significant enrichment. The GRASPS translatome identifies a greater number of biological processes compared to WCT and CyT.

The directionality of differentially expressed genes, which were common between the WCT, CyT and GRASPS translatome were compared in a heat map (Figure 4.14). Strikingly, few changes correlate between the transcriptome and translatome.

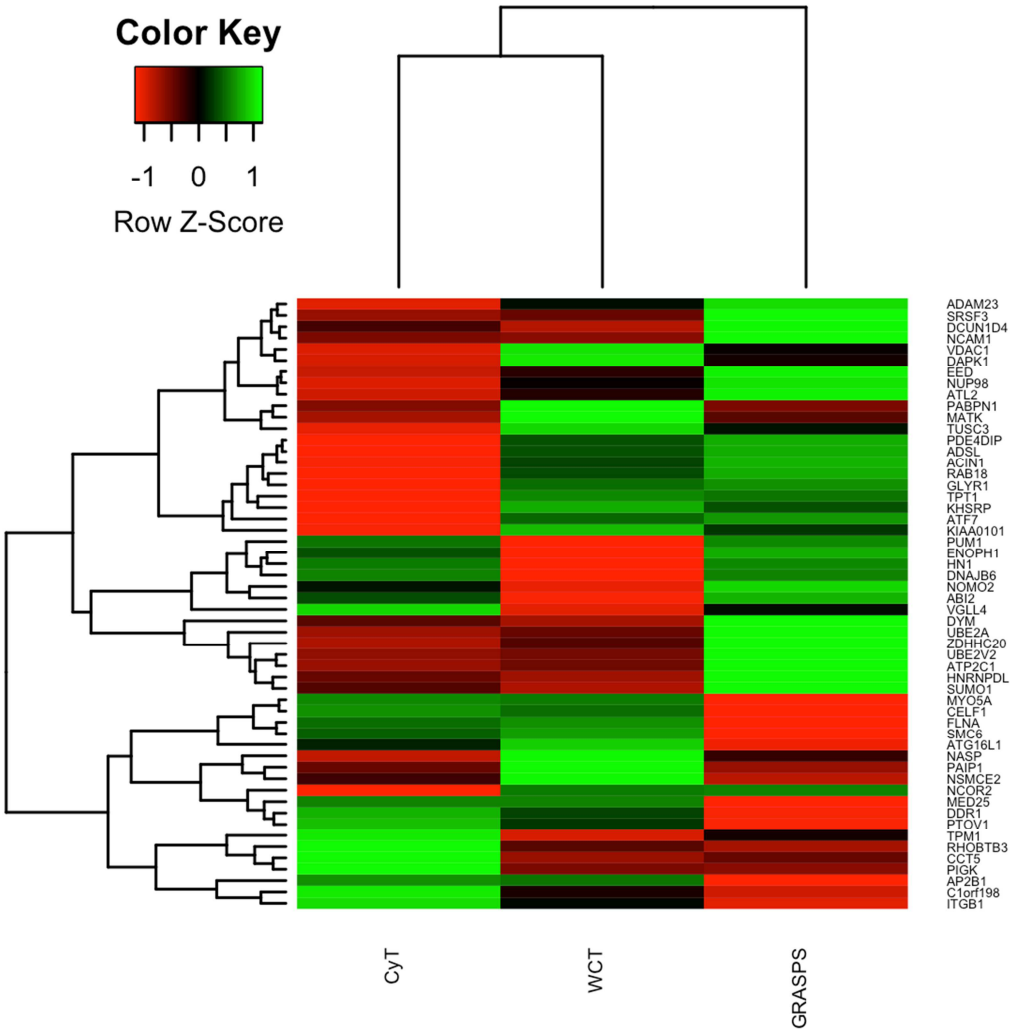


Figure 4.14 Heat map to show differentially expressed transcripts that are common between WCT, CyT and GRASPS translatome

Differentially expressed genes which are common between WCT, CyT and GRASPS translatome with $\log_2$ FC >1. ($p < 0.005$). Each row corresponds to a gene. Green indicates up regulation of a gene. Red indicates down regulation of a gene. (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

Since TDP-43 binds directly several thousands transcripts, we were interested in identifying the differentially expressed targets. Biomart was used to identify human analogues from the targets of TDP-43 identified in rat (Sephton et al., 2011). The list of human orthologue targets of TDP-43 from Biomart is shown in a spreadsheet on attached CD-ROM (see C3. Human orthologue targets of TDP-43 targets). TDP-43 targets found in the broad categories made of gene ontologies identified with the DAVID are shown in red in Appendix 2 Tables A4-A6.

The DEG that are targets of TDP-43 found in the WCT, CyT and GRASPS translatome were illustrated in a Venn diagram (Figure 4.16). Approximately 30 % of all differentially expressed genes identified in the WCT, CyT and GRASPS translatome are targets of TDP-43 despite the fact that the total number of DEG in each of these groups is different: 384 targets out of 1233 DEG for WCT, 662 targets out of 2001 DEG for CyT and 1377 targets out of 4264 DEG for the GRASPS translatome.

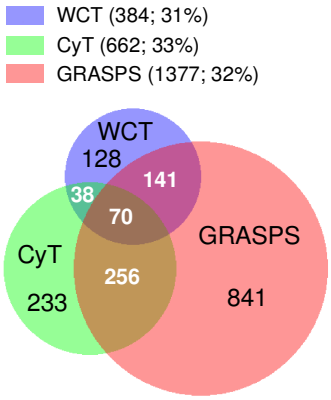


Figure 4.15 Targets of TDP-43 from WCT, CyT and GRASPS translatome
DEG that are targets of TDP-43 from the WCT, CyT and GRASPS translatome with a FC > 2. The whole cell transcriptome (WCT) is shown in blue; cytoplasmic transcriptome (CyT) is shown in green; GRASPS translatome (GRASPS) is shown in red. The number of targets and the percentage of TDP-43 targets found from the total number of DEG is shown in brackets. There were 384 targets out of 1233 DEG for WCT, 662 targets out of 2001 DEG for CyT and 1377 targets out of 4264 DEG for the GRASPS translatome.

Differentially expressed transcripts that are targets of TDP-43, which were common between the WCT, CyT and GRASPS are shown in Appendix 2 table A10 and illustrated in a heat map in Figure 4.16. The directionality of expression is not always the same for the WCT, CyT and GRASPS translatome.

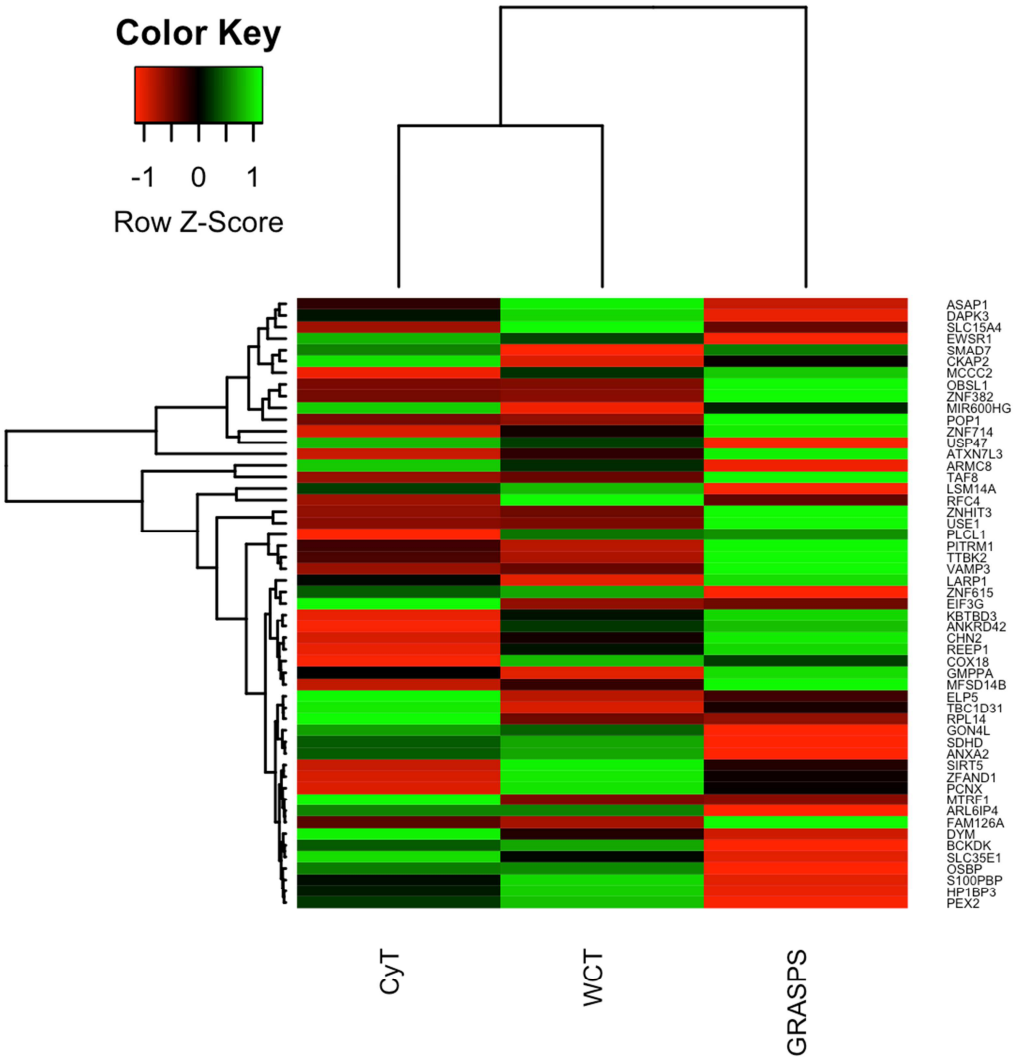


Figure 4.16 Heat map to show differentially expressed genes that are targets of TDP-43 that are common between WCT, CyT and GRASPS translatome
Differentially expressed genes that are target of TDP-43, which are common between WCT, CyT and GRASPS translatome with $\log_2$ FC >1. ($p < 0.005$) (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome) [Figure courtesy of Dr Marta Milo, University of Sheffield]

4.2.6 Correlation studies (WCT vs GRASPS and CyT vs. GRASPS)

For genome-wide investigation, correlation matrices were generated using the differentially expressed transcripts comparing WCT with GRASPS translatome (Figure 4.18 A) and comparing CyT with GRASPS translatome (Figure 4.18 B).

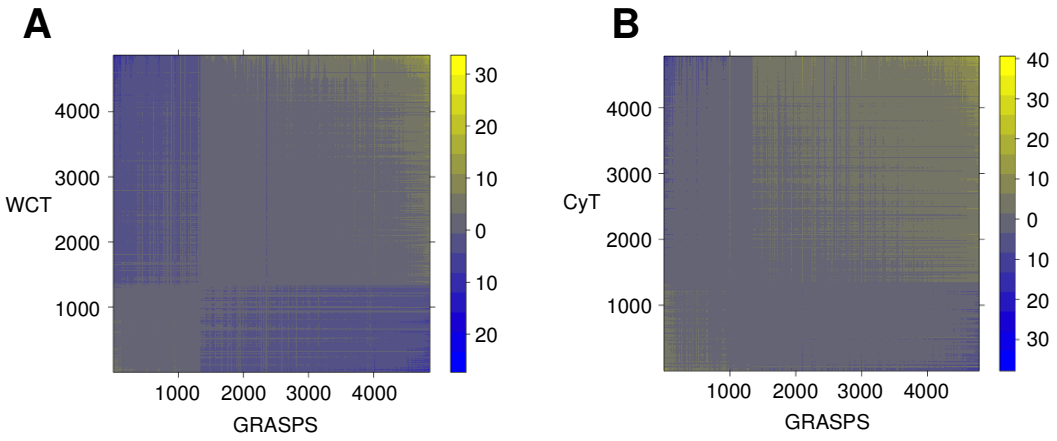


Figure 4.17 Correlation matrices comparing differentially expressed transcripts
Yellow indicates correlation in directionality of expression (up/downregulation of transcript in both data sets). Blue indicates anti-correlation in directionality of expression (upregulation of transcript in one data set and down regulation in the other data set). (A) Comparison of differentially expressed transcripts in WCT with GRASPS translatome. (B) Comparison of differentially expressed transcripts in CyT with GRASPS translatome. (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome). [Figure courtesy of Dr Marta Milo, University of Sheffield]

When comparing the correlation between whole cell transcriptome and GRASPS translatome there are large areas of blue and grey and approximately one third of the matrix is blue showing they are not correlated (Figure 4.18 A). When comparing the correlation between the cytoplasmic transcriptome and translatome there is less blue and more yellow, indicating more correlation between the directionality of differentially expressed transcripts in each group (Figure 4.18 B).

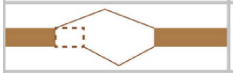
4.2.7 Isoforms analysis

An additional analysis was performed by other bioinformatician collaborators, Ilaria Granata and Mario Guarracino, who have specific expertise in the reconstruction of transcript isoforms for the analysis of alternative splicing. This was an independent analysis where a different pipeline was used to generate a DEG list before splicing analysis was carried out.

The splicing analysis identified abnormal splicing isoforms, which appeared in the HEK Flp TDP-43 Q331K, but not the HEK Flp Sham transcriptomes and translatomes. The type and numbers of abnormal splicing events are shown in Table 4.5. There are approximately 28 times more abnormal splicing events in the WCT and CyT compared to the GRASPS translatome, suggesting that most abnormally spliced transcripts do not get translated into proteins.

Table 4.5 Abnormal splicing events in TDP-43 Q331K

Abnormal splicing events identified from the transcriptomes (WCT, CyT) and GRASPS translatome which are present in the HEK Flp TDP-43 Q331K, but not in HEK Flp Sham. Both gene level and isoform levels are differentially expressed (WCT = whole cell transcriptome; CyT = cytoplasmic transcriptome; GRASPS = GRASPS translatome)

	Event type	WCT	CyT	GRASPS
	Exon skipping/inclusion (ESI)	180	223	6
	Multiple exon skipping/inclusion (MESI)	0	1	1
	Intron skipping/inclusion (ISI)	41	59	0
	Alternative 5’ splice site (A5)	34	50	5
	Alternative 3’ splice site (A3)	186	197	6
	Alternative transcription start site (ATSS)	147	150	4
	Alternative transcription termination site (ATTS)	129	169	3
	Mutually exclusive exons (MEE)	106	145	0

The 33 abnormal splicing events identified in the GRASPS translatome were further investigated, as they would potentially be translated into abnormal proteins (25 listed in Table 4.5 and 8 further events where the gene level was not differentially expressed but the isoform was). The abnormal isoforms identified, their gene names and links to neurodegeneration are outlined in Table 4.6. There were 33 abnormal splicing events reported in the translatome. *IFRD1*, *TBK1* and *PADI2* have already been linked to neurodegeneration and all have altered splicing events identified in the GRASPS translatome.

Table 4.6 Abnormal splicing events in the TDP-43 Q331K translatome

Abnormal splicing events and the type of splicing event is described for each gene. The function of the gene product is outlined. Any links to neurodegeneation and corresponding references are listed. (ATSS= alternative transcription start site; ATTS = alternative transcription termination site; A3 = alternative 3' splice site; A5 = alternative 5' splice site; ESI = exon skipping/inclusion; MESI = multiple exon skipping/inclusion; ISI = intron skipping/inclusion; MEE = mutually exclusive exons)

Gene	Number of splicing events	Type of splicing events	Gene product function	Link to neurodegeneration	References
Peptidyl arginine deiminase 2 (*PADI2*)	1	ATSS	Peptidyl arginine deiminase, post translational deamination of proteins	Enzyme may play a role in the onset and progression of neurodegenerative diseases such as Alzheimer's disease, multiple sclerosis (MS) and is a possible biomarker of inflammatory diseases.	(Nicholas et al., 2004; Ishigami et al., 2005; Jang et al., 2008; 2010)
Zinc finger and BTB domain containing 33 (*ZBTB33*)	1	Generation	Transcriptional regulator with DNA binding specificity		
Heparan sulfate 6-O-sulfotransferase 2 (*HS6ST2*)	2	A3, MESI,	Heparan sulphate proteoglycan – component of cell surface, extracellular matric and basement membranes		
Fibulin 1 (*FBLN1*)	3	ATSS, A5, Generation	Glycoprotein incorporated into fibrillar extracellular matrix		
TANK binding kinase 1 (*TBK1*)	2	A5, A3	Kinase with role in regulating inflammatory response, can mediate NF-kappa-B	Associated with frontotemporal dementia and ALS4	(Freischmidt et al., 2015)
C-Maf inducing protein (*CMIP*)	1	ATSS	Role in T-cell signalling pathway		

Transmembrane protein 123 (*TMEM123*)	1	ATSS	Cell surface receptor that mediates cell death		
Zinc finger protein 558 (*ZNF558*)	1	A5	Transcriptional regulation, nucleic acid binding		
Glia maturation factor beta (*GMFB*)	3	ATTS, A3, A5,	Inducing differentiation of brain cells, neural regeneration and inhibition or proliferation of tumour cells		
Fatty acyl-CoA reductase 1 (*FAR1*)	3	Generation, A5, A3,	Reduction of fatty acids		
DLG associated protein 5 (*DLGAP5*)	2	ATTS, Generation	Mitotoic phosphoprotein regulated by ubiquitin-proteasome pathway		
Solute carrier family 25 member 1 (*SLC25A1*)	1	Generation	Mitochondrial carrier protein		
Nuclear protein 1, transcriptional regulator (*NUPR1*)	1	A3	Chromatin binding protein, transcriptional regulator		
High mobility group AT-hook 2 (*HMGA2*)	1	Generation	Cell cycle regulation		
Zinc finger protein 614 (*ZNF614*)	1	Generation	Transcriptional regulation, nucleic acid binding		
Sorting Nexin 4 (*SNX4*)	1	Generation	Intracellular trafficking		
Spindle apparatus coiled-coil protein 1 (*SPDL1*)	4	Generation, MESI, A3, A5	Mitotic spindle formation and chromosome segregation		
Interferon related developmental regulator 1 (*IFRD1*)	4	Generation, A3, ATSS, ATTS	Transcriptional co-activator/repressor, differentiation in development	Mutations associated with sensory/motor neuropathy with ataxia	(Brkanac et al., 2009)

4.3 Discussion

Growth curve and cell proliferation assays on HEK Flp TDP-43 cell lines showed that an increase in TDP-43 WT or mutant has a toxic effect (Figure 4.3, Figure 4.4). This is consistent with studies in mice that have shown over-expression of human WT or mutant TDP-43 is toxic (Barmada et al., 2010; Stallings et al., 2010). The reduction in growth occurs three days earlier in HEK Flp TDP-43 WT compared to TDP-43 Q331K. This may be because there is more TDP-43 produced in the HEK Flp TDP-43 WT model (Figure 4.1, Figure 4.2), so the toxic result of this is observed earlier. The difference in the protein produced from the two cell models may due to a difference in stability of the TDP-43 WT and TDP-43 Q331K transcripts.

The percentage of mapped reads for the transcriptome and translatome is high. This shows a good depth of reads and that the RNA was of a good quality. The similarity in the mapping between the two cell lines is due to the cell lines being isogenic. The percentage of mapping of sequence reads to the genome is lower for the GRASPS translatome compared the WCT and CyT. This is likely due to the fact that the GRASPS translatome identifies more data than transcriptomic studies and therefore may be enriched for unknown isoforms, which cannot be mapped fully. The mapping is also based upon previous transcriptomic mapping.

The GRASPS translatome identifies a greater number of differentially expressed transcripts and those differentially expressed transcripts have a higher fold changes compared to WCT and CyT (Figure 4.10). The average $\log_2$ CPM is much higher for GRASPS compared to WCT and CyT. This is likely because the GRASPS translatome is enriched for mRNA and more mRNAs of the same

sequence will give an increased number of single sequencing reads. RNA isolated for WCT and CyT are generated from much more diverse RNA populations giving rise to increase signal to noise ratio in the predicted identification of differentially expressed transcripts which would not be detected as significantly and statistically enriched compared to the control samples. The WCT and CyT contain a bigger, more diverse population of RNA, which will include pre-mRNA, non-coding RNA and RNA that is being held in stress granules. The bigger population of RNAs means that there is more noise and therefore it is harder to find signal.

The DAVID analysis of the GRASPS translatome has identified enriched biological pathways with a fold change greater than 8 (Figure 4.11). Biological processes containing DEG with a fold change greater than 8 are likely to be more physiologically relevant than those with smaller fold changes such as 1.5-2 more commonly found in transcriptome profiling, but this would need further investigation. The higher enrichment provided by GRASPS is also due to the fact that there are subtantially more DEGs identified in each of the cellular pathways.

There is not much correlation between the directional changes of DEGs and DEGs in WCT and CyT. This is likely due to the fact that the WCT contains both nuclear and cytoplasmic fractions and that some of the nuclear mRNAs identified as DEGs are still retained in the nucleus and have not been yet exported as mature mRNAs in the cytoplasm. In agreement with this idea, the CyT and GRASPS translatome show a better correlation in the direction of predicted changes. This is expected as the cytoplasm contains mature mRNAs and a fraction of which will become direct substrates for translation into proteins. The correlation is yet far from 1:1 as many cytoplasmic mRNAs might be

trapped in stress granules for example. The GRASPS translatome is enriched for mRNA, so there is less noise and this gives the GRASPS translatome a greater statistical power.

Nucleocytoplasmic transport was a novel pathway significantly enriched by the functional clustering annotation DAVID in GRASPS. The differentially expressed genes involved in this pathway are listed in Appendix 2 Table A6. This list includes several components of the TREX complex including DDX39B, THO subunits and NXF1, several nucleoporins and cytoplasmic RNA helicase DDX19b/DBP5. There are also four nuclear export adaptors included in this list: SRSF1, SRSF3, SRSF7 and CHTOP. The genes involved in this pathway and their relative fold change in WCT, CyT and GRASPS were illustrated in a Pheatmap (Figure 4.13). The majority of genes in the WCT and CyT are shown in yellow, indicating no change in differential expression. This also explains why the pathway had not been detected for the WCT and CyT in the DAVID analysis, as there are only a small number of changes identified. However, in the GRASPS translatome many changes are identified and the fold changes are much greater than those observed in the transcriptomes. Again the lack of correlation between fold changes observed in the transcriptomes compared to the translatome can be seen.

The list of enriched biological processes generated using PANTHER (Appendix 2 Tables A7-A9) identifies the same processes as generated independently using DAVID. This shows that the analyses are reliable, as two independent measures produce the same set of enriched biological processes from the lists of DEG. The GRASPS translatome, again, generated more biological processes compared to the transcriptomes.

The main difference in the two different approaches is that DAVID identifies enriched processes based upon the number of differentially expressed genes identified for a specific process, whereas PANTHER identifies processes that are enriched significantly by comparing the number of differentially expressed genes involved in the pathway compared to the number of genes in a pathway that would be expected at random. The similarity in the results using two different methods of analysis, gives increased confidence in the results of the analyses.

The TDP-43 target transcripts with a FC greater than 2 and which were common between the whole cell transcriptome, cytoplasmic transcriptome and translatome were compared (Figure 4.16, Appendix 2 Table A10) and illustrated in a heat map (Figure 4.17). The GRASPS translatome identifies 1377 differentially expressed genes that are targets of TDP-43 compared to 662 and 384 genes for the cytoplasmic transcriptome and whole cell transcriptome respectively. The GRASPS translatome therefore provides more detail than the transcriptomes. It also identifies targets of TDP-43, which are being translated; these could be the early dysregulation events as they are direct targets of TDP-43.

The directionality of expression in DEG common between the WCT, CyT and GRASPS translatome (Figure 4.14) and of DEG, which are TDP-43 targets, common between the WCT, CyT and GRASPS translatome (Figure 4.17) was compared. The directionality of expression is not always the same for the WCT, CyT and GRASPS translatome. For example, pumilio RNA Binding family member 1 (PUM1) is a target of TDP-43, which is downregulated in the WCT, but upregulated in the CyT and GRASPS translatome. This is evidence that the changes observed in the transcriptome do not correlate with what is being

translated at the ribosome, further emphasising the need to investigate the translatome in order to decipher disease pathways.

Similarly, Filamin A (FLNA) is a TDP-43 target, which is up regulated in the WCT and CyT, but downregulated in the GRASPS translatome. These differences are consistent with cell compensatory mechanisms. An up regulation of a transcript in the whole cell transcriptome and cytoplasmic transcriptome may be due to a decrease in protein levels, therefore a downregulation is observed in the GRASPS translatome, or vice versa.

The correlation in directionality of differentially expressed transcripts between the transcriptome and translatome was compared in covariance matrices (Figure 4.18). The WCT and GRASPS translatome are are not correlated (Figure 4.18 A). This is due to the total RNA in the whole cell sample containing many RNAs, which will not be translated. The transcriptome and translatome are not correlated. As shown in the heat maps, the anticorrelation between the transcriptome and translatome, is evidence of cell compensatory mechanisms in action.

When comparing the CyT and translatome there is more correlation between the directionality of differentially expressed transcripts in each group (Figure 4.18 B). This is because the cytoplasmic transcriptome will not contain pre-mRNA or mRNA stalled in nuclear speckles but will contain mRNA exported for translation. However, it will not fully correlate, as it will also contain non-translated mRNA in stress granules or mRNA that will be degraded. As expected, there is some correlation between the cytoplasmic transcriptome and the GRASPS translatome.

The splicing analysis of the WCT, CyT and GRASPS translatome revealed a higher proportion of abnormal splicing events in the WCT and CyT compared to the GRASPS translatome (Table 4.5). These data show that abnormally spliced transcripts are exported into the cytoplasm but the majority are not associated with ribosomes and are unlikely to be translated. The large decrease in the number of abnormal splicing events in the translatome suggests an unknown mechanism may occur in the cytoplasm, which prevents abnormally spliced transcripts in the cytoplasm from reaching the ribosome to be translated.

The 33 abnormal splicing events in the translatome were further investigated and the potential role of the gene products in neurodegeneration was explored (Table 4.6). Of the 33 abnormal splicing events, 7 of these corresponded to gene products, which have already been reported in neurodegeneration (*PADI2*, *TBK1*, *IFRD1*). Notable from these is *TBK1* as mutations in *TBK1* have already been reported as causing familial ALS and frontotemporal dementia (Freischmidt et al., 2015).

The HEK Flp TDP-43 Q331K ALS translatome identified cellular retinoic acid binding protein 1 (*CRABP1*) as the most downregulated transcript. CRABP1 is a specific binding protein involved in retinoic acid differentiation and proliferation processes. Neuron differentiation is also an enriched dysregulated pathway identified from the ALS translatome. In order to investigate the retinoic acid differentiation HEK Flp In Sham and TDP-43 Q331K cells were converted into iNPCs before inducing neuronal differentiation with retinoic acid. This work is currently being carried out by Mr Simeon Mihaylov. A modified protocol based upon (Janghwan Kim et al., 2011) and (Meyer et al., 2014) was used to convert HEK Flp In cells into iNPCs, omitting the lentiviral transduction due to HEK cells already being immortalised. The cells were then subjected to a 10-day retinoic

acid differentiation in the absence of Tetracycline and 48 h into the differentiation iNPCs derived from the HEK Flp TDP-43 Q331K cells show fewer, shorter projections compared to iNPCs derived from HEK Flp Sham cells. Later in the protocol, more projections occur in the iNPCs derived from HEK Flp TDP-43 Q331K, showing delayed neuronal differentiation. This difference in response to the retinoic acid differentiation is observed in iNPCs grown in the absence of Tetracycline. As shown and discussed previously, some inserted transgene expression is observed in the absence of Tetracycline due to the leakiness of the promoter. This change is the result of a small amount of TDP-43 Q331K expression. When the iNPCs are grown in the presence of Tetracycline and subjected to differentiation, the iNPCs from the HEK Flp TDP-43 Q331K again make few projections throughout the entire protocol. This shows that the production of TDP-43 Q331K has a negative effect on neuronal differentiation due to the downregulation of CRABP1 and represents functional validation of the GRASPS finding.

Furthermore, CRABP1 levels in HEK Flp Sham and TDP-43 Q331K are being compared by qRT-PCR following a 48 h induction with Tetracycline by Dr Ya-Hui Lin. There is a significant difference in the expression levels of CRABP1 in HEK Flp Sham cells compared to HEK Flp TDP-43 Q331K cells. Approximately one hundred times less CRABP1 transcripts are observed in the HEK Flp TDP-43 Q331K cells compared to Sham.

Preliminary data investigating gene expressions in whole cell, cytoplasmic and nuclear fractions from HEK Flp In cell lines has been obtained Dr Ya-Hui Lin. These data shows a nuclear reduction and cytoplasmic accumulation of TDP-43 targets in the HEK Flp TDP-43 Q331K cell lines. This is consistent with the

disruption of nuclear export and the GRASPS translatome analysis reporting altered mRNA transport/export transcripts.

5. Investigating TDP-43 proteinopathy and potential dysregulation of mRNA nuclear export in TDP-43 ALS

5.1 Introduction

This study and others, have shown that HEK Flp TDP-43 cell models do not display the characteristic TDP-43 proteinopathy that is observed in diseased motor neurons (Ling et al., 2010). This could be due to the fact that HEK cells are not neuronal and TDP-43 proteinopathy is only observed in motor neurons. Therefore the Flp In system was integrated into motor neuron-like NSC-34 cell line within SITraN. This cell line could be used to generate neuronal motor neuron-like cell models with Tetracycline-inducible TDP-43 WT and TDP-43 Q331K. The future advantages of having both a non-neuronal and a neuronal-like cell model of TDP-43 ALS would be from comparing the two, as this could reveal what makes motor neurons (unlike other cell types) susceptible to disease.

The aim of the work described in this chapter was to construct a motor neuron-like TDP-43 ALS cell model in which to potentially investigate TDP-43 proteinopathy. This cell model could then be used to investigate and validate any dysregulated pathways highlighted from the HEK Flp TDP-43 Q331K translatome such as potential alteration of the mRNA nuclear export pathway.

5.2 Results

5.2.1 Investigating the cellular distribution of TDP-43 in HEK Flp In cell models

In order to validate the data from Ling et al. (Ling et al., 2010), immunofluorescence studies were carried out in order to determine the cellular localisation of TDP-43 protein within the previously described HEK Flp cell models (Figure 5.1). Cells grown in the absence of Tetracycline all show similar amounts of TDP-43 protein, which is located predominantly in the nucleus as expected for TDP-43. Upon the addition of Tetracycline, there is an increase in the amount of TDP-43 protein in the HEK Flp TDP-43 WT and HEK Flp TDP-43 Q331K cell lines compared to HEK Flp Sham due to expression of TDP-43 from the integrated transgene. The addition of Tetracycline induction and length of induction did not change the localisation of TDP-43, with TDP-43 remaining predominantly nuclear in all cell lines at both 48 h and 14 days.

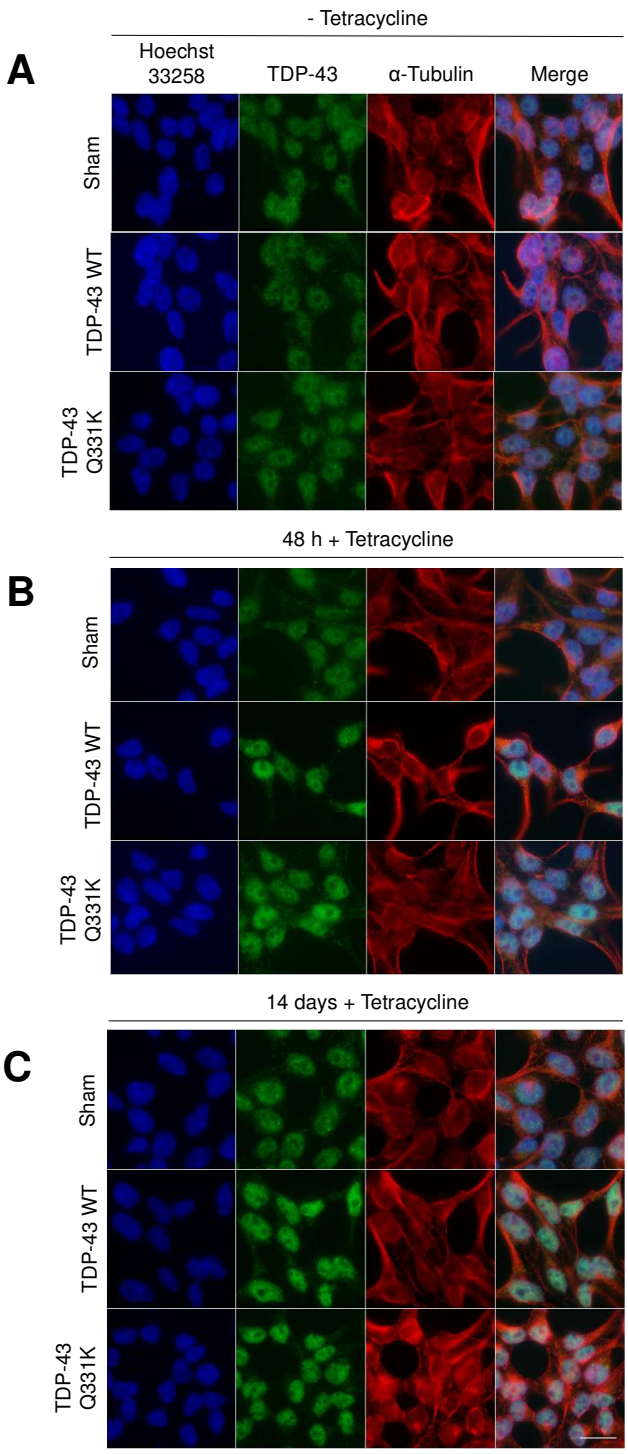


Figure 5.1 TDP-43 localisation in HEK Flp cell models
Immunofluorescence of HEK Flp Sham, TDP-43 WT and TDP-43 Q331K cells to investigate TDP-43 localisaton. Cells were plated and left 24 h before the addition of 10 µg/mL Tetracycline. Cells were stained with TDP-43 (green), α-Tubulin (red) and Hoechst (blue). (A) HEK Flp cells grown in the absence of Tetracycline. TDP-43 staining shows endogenous TDP-43. (B) HEK Flp cells grown in the presence of Tetracycline for 48 hours. All cell lines show predominantly nuclear localisation of TDP-43. (C) HEK Flp In cells in the presence of Tetracycline for 14 days. All cell lines show predominantly nuclear localisation of TDP-43. Scale bar = 20 µm.

5.2.2 Building and characterisation of neuronal TDP-43 ALS inducible cell models

The Flp In system was integrated into motor neuron-like NSC-34 cells by Dr Adrian Higginbottom in SITraN (Figure 5.2). NSC-34 cells were generated by fusion of motor neuron enriched from embryonic mouse spinal cord with mouse neuroblastoma cells and resemble developing motor neurons (Cashman et al., 1992).

The FRT cassette was first transfected into the NSC-34 cell line. Successful integration would give NSC-34 cells resistance to Zeocin. Zeocin resistant clones were collected and stored. Southern blotting was used to confirm integration of a single FRT site. This is important, as only one FRT site is required so that integration will occur at the same, single site in every model built. Therefore only clones with one FRT site were taken forward. Integration of the FRT site could have occurred in a transcriptionally silent part of the genome and in order to test this, a GFP plasmid was transfected into the NSC-34 cells with a single FRT site. Successful integration of GFP would give cells resistance to Hygromycin and the addition of Tetracycline to these Hygromycin-resistant cells would result in expression of GFP, indicating integration of the FRT into a transcriptionally active site. Cells with a single FRT site, which had been shown to be transcriptionally active, were therefore taken forward as the parental FRT cell lines.

To construct a Tetracycline-inducible cell model, the Tetracycline repressor needed to be added. Parental FRT lines were transfected with the Tetracycline repressor. Successful integration of the repressor element gives cells resistance to Blasticidin. Zeocin and Blasticidin resistant clones were therefore collected

and stored. The Tetracycline repression of these cells was then tested. Again, this was done by transfection of a GFP plasmid into Zeocin and Blasticidin resistant clones. Integration of GFP at the FRT site changes the antibiotic resistance from zeocin resistant to Hygromycin resistant. Hygromycin resistant clones were then grown in the presence and absence of Tetracycline. Cells that expressed GFP in the presence of Tetracycline and showed repression of GFP expression in the absence of Tetracycline were deemed suitable clones. A Zeocin and Blasticidin resistant clone, that had been shown to be Tetracycline inducible and repressable was then selected as the NSC-34 parent FRT TREx cell line (Dr Adrian Higginbottom). This cell line would be used to generate all NSC-34 Flp In cell models from hereon.

NSC-34 Flp Sham, NSC-34 Flp TDP-43 WT and NSC-34 Flp TDP-43 Q331K cell lines were generated, selecting Hygromycin resistant clones and testing the induction with Tetracyline in the same way as previously described for the construction of the HEK Flp GFP cell models (Chapter 3).

NSC-34 Flp TDP-43 models were built containing either human TDP-43 WT or ALS mutations (A315T, M337V and Q331K) (Figure 5.3). As previously mentioned, TDP-43 is able to control its expression level by binding the 3' UTR of its own mRNA. Therefore the integrated transgene does not contain a 3' UTR, so that the human protein can be expressed and hopefully bind endogenous mouse TDP-43 mRNA in the regulatory feedback loop, in order to avoid over-expression. Upon the addition of Tetracycline, there is an increase in the amount of total TDP-43 protein in all NSC-34 Flp TDP-43 cell lines. This is due to the addition of human TDP-43 from the transgene to the

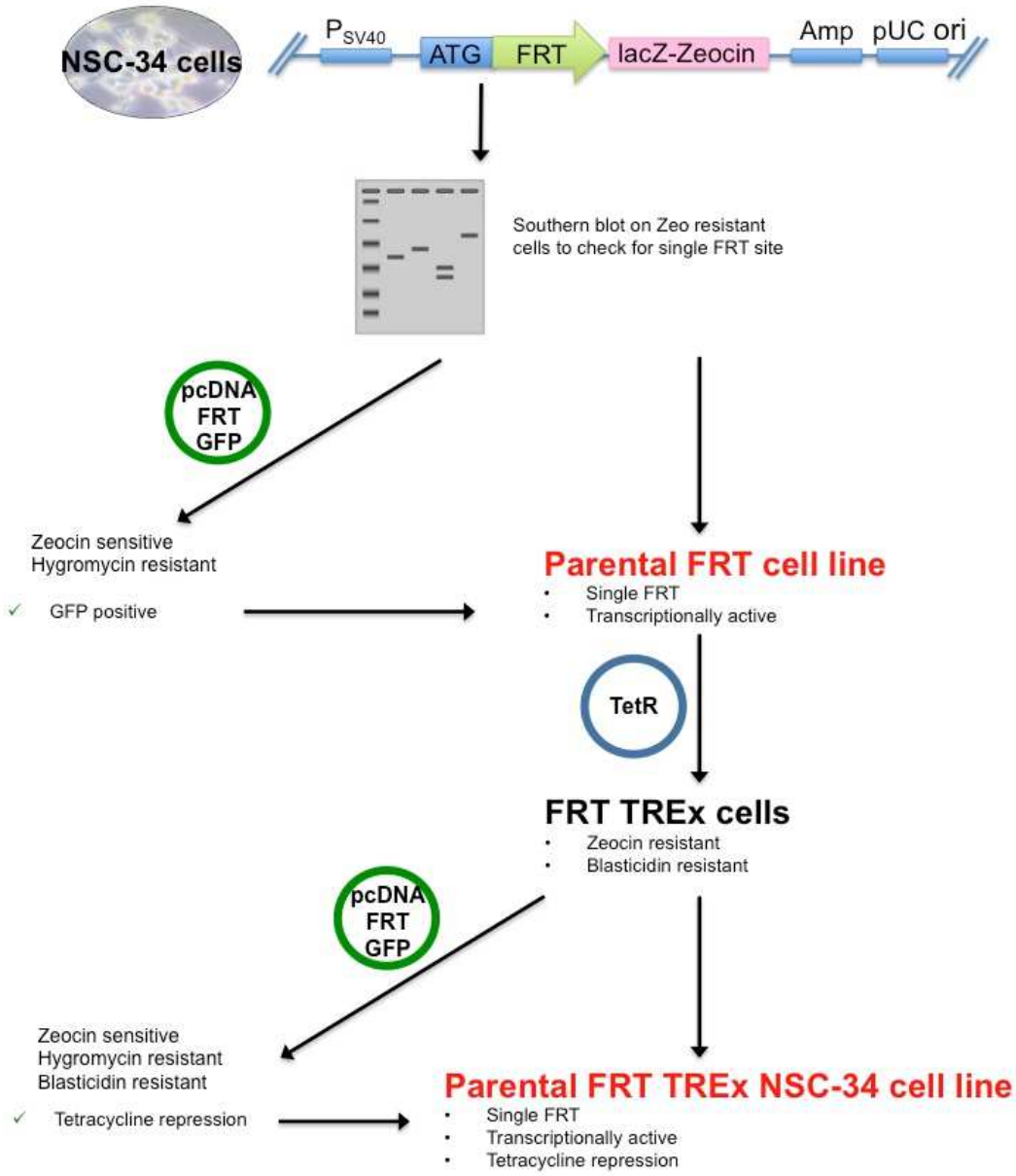


Figure 5.2 Strategy for building inducible NSC-34 cell line using the Flp In T-REx system

The FRT site was integrated into the NSC-34 cell line and southern blotting used to confirm integration of a single FRT site. The transcriptional activity of the FRT site was then tested by transfection of a GFP plasmid. Cells with single FRT sites, which were transcriptionally active, were taken and the Tetracycline repressor inserted. Tetracycline repression and induction were tested by transfection of a GFP plasmid. A clone with a single FRT in a transcriptionally active site, that also had Tetracycline repression, was selected as the parental FRT T-REx NSC-34 cell line.

endogeneous levels of mouse TDP-43 protein. For total TDP-43, TDP-43 is present in the cells grown in the absence of Tetracycline. This is due to the antibody recognising the endogenous mouse TDP-43. When using a human specific TDP-43 antibody, no TDP-43 is observed in the cells grown in the

absence of Tetracycline, due to repression of the human transgene. However, upon the addition of Tetracycline, human TDP-43 WT and mutants are expressed from the transgene.

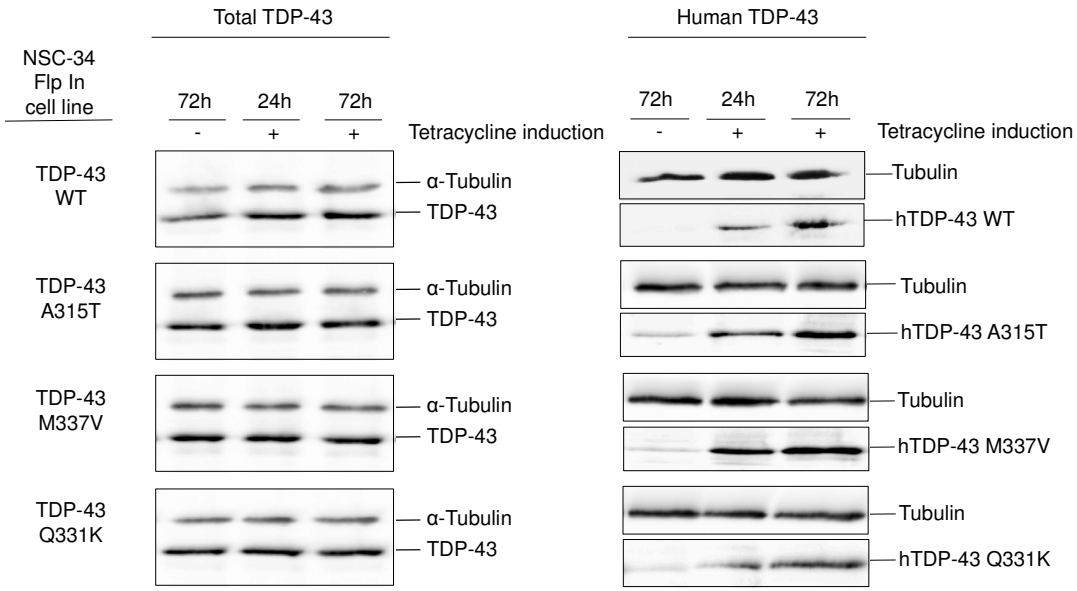


Figure 5.3 NSC-34 Flp TDP-43 cell models
Cells were grown for 24 h before the addition of 10 µg/mL Tetracycline for 24 h and 72 h. Western blot analysis probing for total TDP-43 (left) and human TDP-43 (right). (hTDP-43 = human TDP-43)

Human TDP-43 transcript levels were measured in the presence and absence of Tetracycline in the NSC-34 Flp Sham, NSC-34 Flp TDP-43 WT and NSC-34 Flp TDP-43 Q331K cell lines (Figure 5.4). No human TDP-43 mRNA is present in the NSC-34 Flp Sham cell lines. Upon the addition of Tetracycline there is an increase in human TDP-43 mRNA in the NSC-34 Flp TDP-43 WT and TDP-43 Q331K cell lines. When cells are grown in Tetracycline there is 2.6 times more human TDP-43 mRNA produced in the NSC-34 Flp TDP-43 WT compared to NSC-34 Flp TDP-43 Q331K.

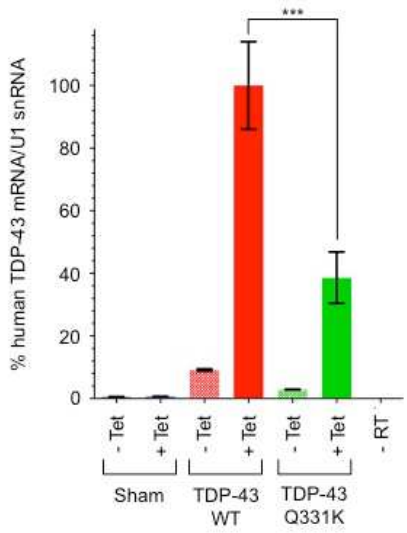


Figure 5.4 Human TDP-43 transcript levels in NSC-34 Flp cell models
qRT-PCR analysis of NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated for 24 h before a 48 h induction with 10 µg/mL Tetracycline (+ Tet) or without Tetracycline (- Tet). Quantification was performed using the ΔCt method before normalisation to NSC-34 Flp TDP-43 WT + Tet. There was 2.6 times more human TDP-43 mRNA in the NSC-34 TDP-43 WT cells compared to TDP-43 Q331K. (Mean ± SEM; 2way ANOVA with Tukey's multiple comparison, ***: $p < 0.001$; N = 3)

The cell proliferation in the presence of Tetracycline was recorded at 48 h and 16 days post induction (Figure 5.5). After a 48 h induction with Tetracycline there was a significant reduction in the proliferation of NSC-34 Flp TDP-43 WT and TDP-43 Q331K compared to NSC-34 Flp Sham, by approximately 25 %. There was no significant difference in the proliferation of NSC-34 Flp TDP-43 WT compared to NSC-34 Flp TDP-43 Q331K. After a 16 day induction with Tetracycline there was a significant reduction in the proliferation of NSC-34 Flp TDP-43 WT compared to NSC-34 Flp Sham, by approximately 80 % and a significant reduction in the proliferation of NSC-34 Flp TDP-43 Q331K compared to NSC-34 Flp Sham, by approximately 70 %. There was a significant but small reduction in the proliferation of NSC-34 Flp TDP-43 WT compared to NSC-34 Flp TDP-43 Q331K.

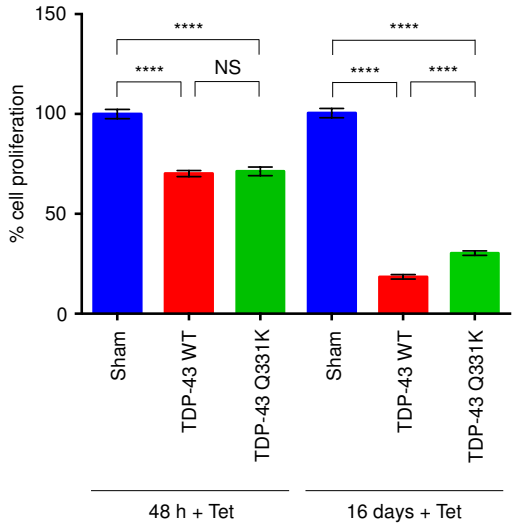


Figure 5.5 Cell proliferation of NSC-34 Flp cell models
MTT assays on NSC-34 Flp cell lines grown in the presence of 10 µg/mL Tetracycline for 48 h and 16 days. Each bar is normalised to Sham for each timepoint. Each bar of TDP-43 WT and TDP-43 Q331K is compared to corresponding bar from Sham cell line for each timepoint separately. (Mean ± SEM; 1way ANOVA with Tukey's multiple comparison, ****: $p < 0.0001$, NS: non-significant; N (independent experiments, 6 wells per experiment): 48 h=3; 16 days=2).

Immunofluorescence studies were carried out in order to determine the cellular localisation of human TDP-43 protein within the NSC-34 Flp cell models following induction of Tetracycline for different lengths of time (Figure 5.6). No human TDP-43 was observed in the NSC-34 Flp Sham cell line following induction with Tetracycline. Human TDP-43 has a predominantly nuclear localisation in NSC-34 Flp TDP-43 WT and TDP-43 Q331K. The length of induction does not affect the localisation of human TDP-43 in the NSC-34 Flp TDP-43 cell lines. There was more human TDP-43 in the NSC-34 Flp TDP-43 WT compared the NSC-34 Flp TDP-43 Q331K, following induction with Tetracycline at all three time points.

In order to try and reduce the difference in expression of human TDP-43 between the NSC-34 Flp TDP-43 WT and TDP-43 Q331K cell models, the protein levels were quantified following induction with different concentrations of Tetracycline (Figure 5.7). The amount of total (human and mouse) TDP-43

remained the same at all Tetracycline concentrations indicating that TDP-43 was not overexpressed following induction with Tetracycline (Figure 5.7 A). There was approximately three times more human TDP-43 protein, produced from the transgene, in the NSC-34 TDP-43 WT compared to NSC-34 Flp TDP-43 Q331K at all Tetracycline concentrations (Figure 5.7 B).

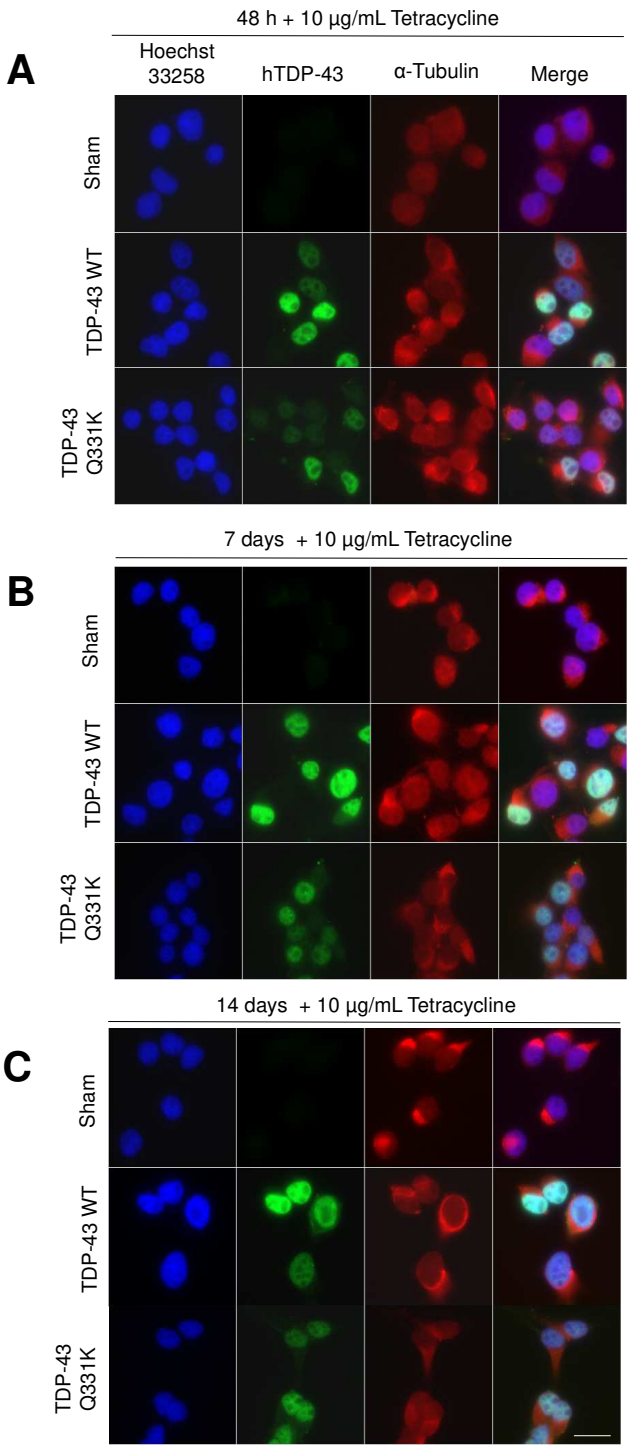


Figure 5.6 Human TDP-43 localisation in NSC-34 Flp cell models
Human TDP-43 localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K grown in the presence of 10 µg/mL Tetracycline for different periods of time. Cells were stained with human TDP-43 (green), α-Tubulin (red) and Hoechst (blue). No human TDP-43 was observed in the NSC-34 Flp Sham cell lines. Human TDP-43 was predominantly nuclear in NSC-34 Flp In TDP-43 WT and NSC-34 Flp In TDP-43 Q331K cell lines grown in Tetracycline for 48 h, 7 days and 14 days. Scale bar = 20 µm.

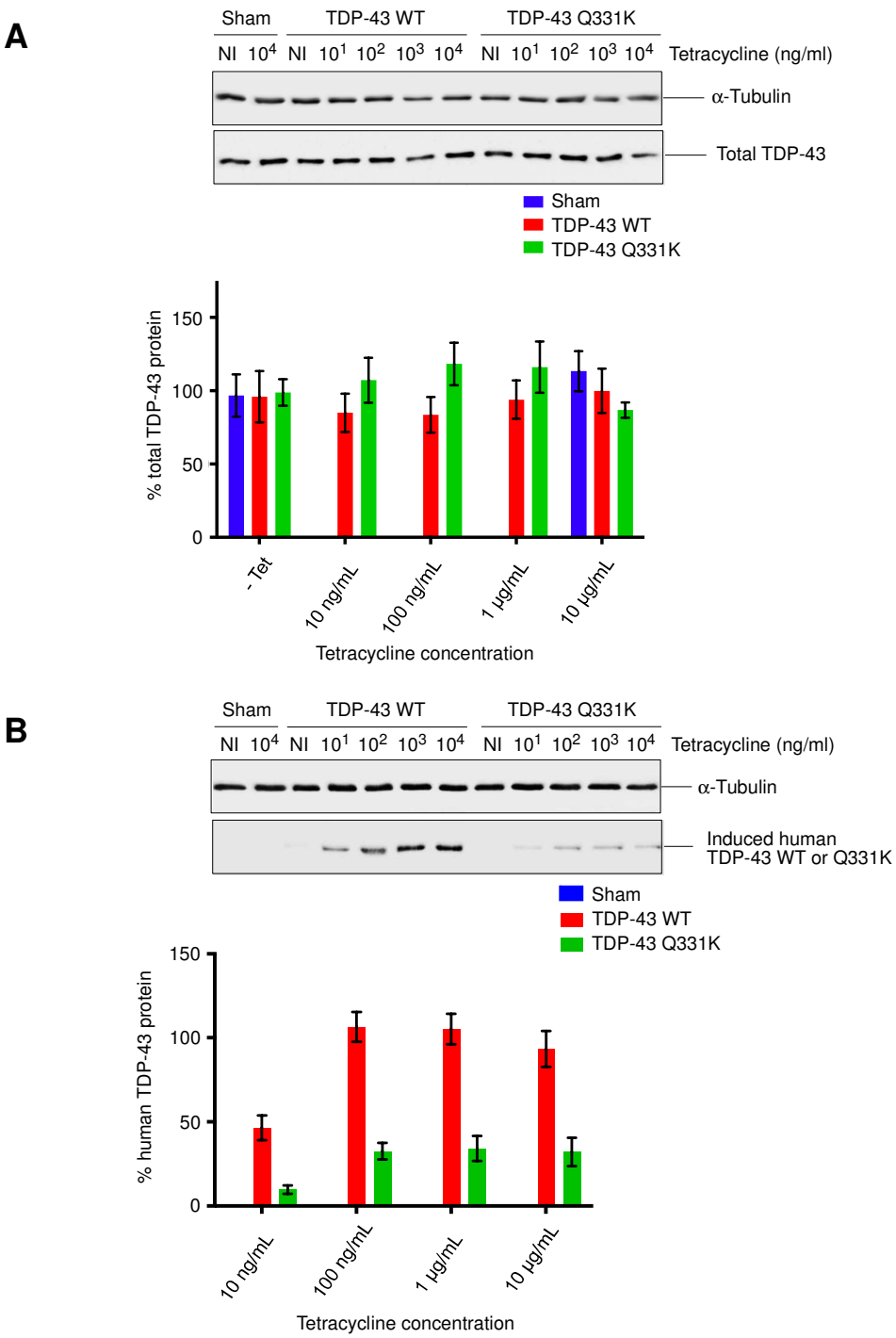


Figure 5.7 Tetracycline concentration on TDP-43 protein levels in NSC-34 Flp cell models

Western blot of NSC-34 Flp cell lines grown in the absence or presence of various concentrations of Tetracycline for 48 h. Graph shows quantification from western blot image. Each bar is normalised to NSC-34 Flp TDP-43 WT at 10 µg/mL Tetracycline. The amount of total TDP-43 remained the same at all concentrations, showing that the cell model is not an over-expression model. There was approximately 3 times more human TDP-43 expressed in the NSC-34 Flp In TDP-43 WT cells compared to the NSC-34 Flp In TDP-43 Q331K cells at all Tetracycline concentrations. The amount of Tetracycline used does not increase the amount of human TDP-43 protein produced. (Mean ± SEM; N (Total TDP-43): Sham = 3 TDP-43 WT and Q331K = 6; N (Human TDP-43) = at least 4)

The cell proliferation following induction with various concentrations of Tetracycline was investigated (Figure 5.8). There was no significant difference in cell proliferation at any of the Tetracycline concentrations tested, although a trend towards slightly reduced cell proliferation was observed for the TDP-43 Q331K cell line at all Tetracycline concentrations.

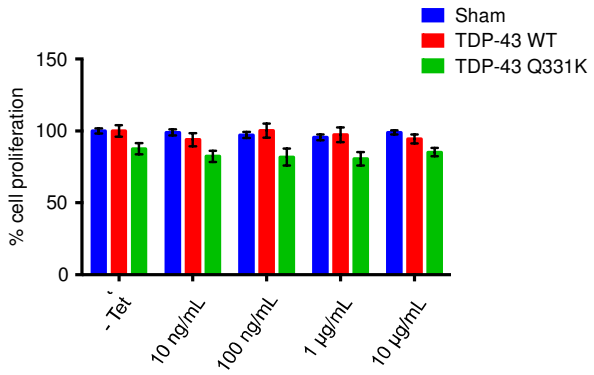


Figure 5.8 The effects of Tetracycline concentration on cell proliferation in NSC-34 Flp cell models

MTT assays were carried out on NSC-34 Flp cell lines grown in the absence (-Tet) or presence (+ Tet) of Tetracycline for 48 h. Each concentration was normalized to –Tet. Each bar of TDP-43 WT and TDP-43 Q331K was compared to corresponding bar from Sham cell line. No significant differences in cell proliferation were observed. (Mean ± SEM; Two-way ANOVA with Tukey's multiple comparison; N (independent experiments, 4 wells per experiment)= 4).

Inducing NSC-34 Flp TDP-43 WT cells with 10 ng/mL Tetracycline and NSC-34 TDP-43 Q331K cells with 100 ng/mL would give a more similar level of human TDP-43 protein expression (Figure 5.7 B). A growth curve was carried out using these Tetracycline concentrations, to see what effect expressing similar levels of human TDP-43 had upon the growth of the WT and mutant cell lines (Figure 5.9). The growth of NSC-34 Flp Sham, TDP-43 WT and TDP-43 Q331K cells in the presence and absence of Tetracycline was recorded over a twenty one-day period. There was no significant difference in the growth of cells grown in the absence or presence of Tetracycline until day 18. At day 18, whether grown in the absence or presence of Tetracycline, there was a significant reduction in the growth of NSC-34 Flp TDP-43 WT compared to Sham. At day 21, whether grown in the absence or presence of Tetracycline, there was a significant

reduction in the growth of both NSC-34 Flp TDP-43 WT and Q331K compared to NSC-34 Flp Sham.

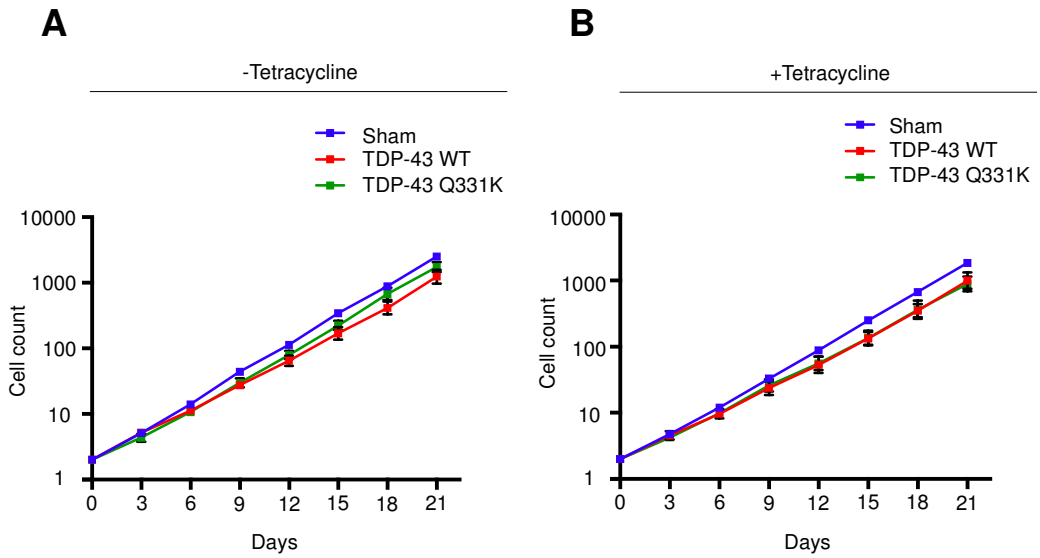


Figure 5.9 Growth curve of NSC-34 Flp cell models when expressing similar levels of human TDP-43 protein Growth of NSC-34 Flp cell lines over a twenty one-day period (A) Growth in the absence of Tetracycline (B) Growth in the prescence of 100 ng/mL Tetracycline for NSC-34 Flp In Sham and TDP-43 Q331K cell lines and 10 ng/mL Tetracycline for NSC-34 Flp TDP-43 WT cell line. Each cell line is compared to the NSC-34 Flp Sham. - Tet: Day 18 Sham vs WT* Day 21 Sham vs WT**** Day 21 Sham vs Q331K****. +Tet: Day 18 Sham vs WT* Day 21 Sham vs WT**** Sham vs Q331K****. (Mean ± SEM; 2way ANOVA with Tukey's multiple comparison, *: $p < 0.1$, ****: $p < 0.0001$; N = 3).

From hereon the concentration of Tetracycline used to induce all cell lines was changed from 10 µg/mL Tetracycline to 1 µg/mL Tetracycline.

Immunofluorescence studies were carried out in order to determine the cellular localisation of human TDP-43 protein within the NSC-34 Flp cell models following induction with 1 µg/mL Tetracycline for 48 h and 14 days (Figure 5.10). No human TDP-43 was observed in the NSC-34 Flp Sham. Human TDP-43 was predominantly nuclear in NSC-34 Flp TDP-43 WT and TDP-43 Q331K cells. The length of Tetracycline induction did not have an effect on the localisation of human TDP-43.

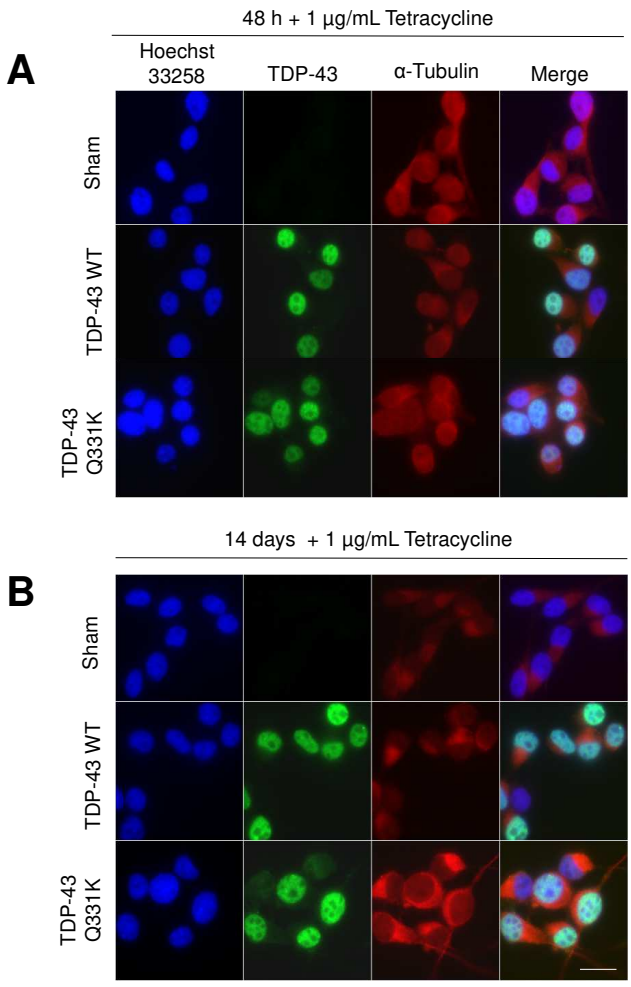


Figure 5.10 Human TDP-43 localisation in NSC-34 Flp models following induction with 1 µg/mL Tetracycline

Human TDP-43 localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K grown in the presence of 1 µg/mL Tetracycline for 48 h (A) and 14 days (B). Cells were stained with human TDP-43 (green), α-Tubulin (red) and Hoechst (blue). No human TDP-43 was observed in the NSC-34 Flp Sham cell lines. Human TDP-43 was predominantly nuclear in NSC-34 Flp TDP-43 WT and NSC-34 Flp In TDP-43 Q331K cell lines. Scale bar = 20 µm.

Previous studies have reported TDP-43 mislocalisation following cellular stress (Barmada et al., 2010). The serum in growth media was reduced to 2.5 % in order to cause cellular stress and the total TDP-43 protein localisation was investigated (Figure 5.11). TDP-43 was predominantly nuclear in all cell lines. There was an increase in the amount of cytoplasmic TDP-43 observed in the NSC-34 Flp TDP-43 Q331K cell line compared to the NSC-34 Flp TDP-43 WT and Sham. Following a 9-day induction with Tetracycline, there was an increase in the amount of cytoplasmic TDP-43, but the TDP-43 protein remained

predominantly nuclear in the NSC-34 Flp TDP-43 Q331K cell line. This does not fully recapitulate the TDP-43 proteinopathy in patient motor neurons where nuclear loss is also observed.

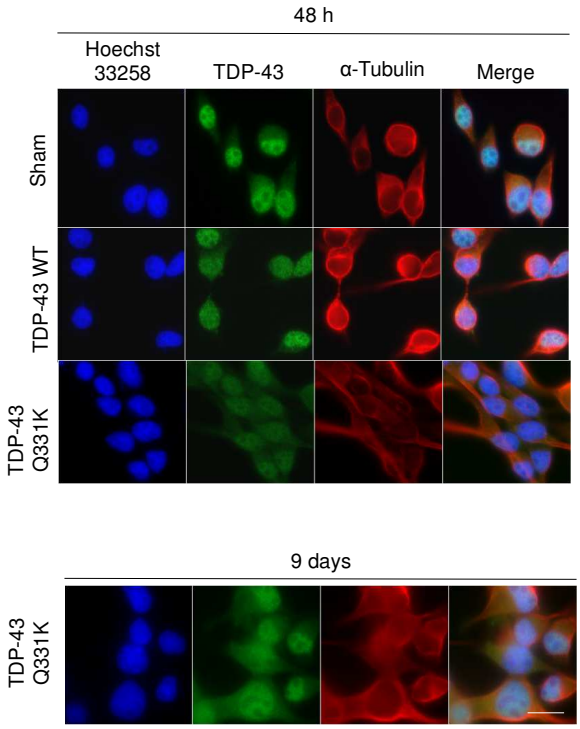


Figure 5.11 Reduction of serum on total TDP-43 localisation in NSC-34 Flp cell models
Total TDP-43 localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cell were grown in the presence of 1 µg/mL Tetracycline and 2.5 % serum media for 48 h and 9 days. Cells were stained with total TDP-43 (green), α-Tubulin (red) and Hoechst (blue). TDP-43 was predominantly nuclear in all cell lines. There was an increase in cytoplasmic TDP-43 observed in the NSC-34 Flp In TDP-43 Q331K cell line. Scale bar = 20 µm.

A previous study indicated that TDP-43 WT shuttles continuously between the nucleus and the cytoplasm in a transcription dependent manner (Ayala et al., 2008). Moreover, previous experiments in RNA Biology, SITraN, showed that TDP-43 ALS mutant proteins (A315T, Q331K, M337V and Y374X) are less efficiently recruited to poly(A)$^+$ mRNA, suggesting potential impairment in the co-transcriptional processing of transcripts. In agreement with this, the GRASPS analysis highlighted that the mRNA nuclear export pathway which functions through the TREX complex is altered in the TDP-43 ALS mutant Q331K. Therefore, a transcription inhibitor, such as Actinomycin D, may exacerbate the

TDP-43 proteinopathy. Human TDP-43 protein localisation and poly(A)$^+$ mRNA localisation were investigated following Actinomycin D treatment with cells induced with Tetracycline for 48 h and 7 days (Figure 5.12). No human TDP-43 was observed in the NSC-34 Flp Sham cell line in the presence or absence of Actinomycin D. In the absence and presence of Actinomycin D following a 48 h induction with Tetracycline, human TDP-43 and poly(A)$^+$ mRNA were predominantly nuclear in both the NSC-34 Flp TDP-43 WT and TDP-43 Q331K. After treatment with Actinomycin D following a 7 day induction with Tetracycline human TDP-43 and poly(A)$^+$ mRNA are predominantly nuclear in the NSC-34 Flp TDP-43 WT, but were both lost from the nucleus and accumulated in the cytoplasm in NSC-34 Flp TDP-43 Q331K cell line.

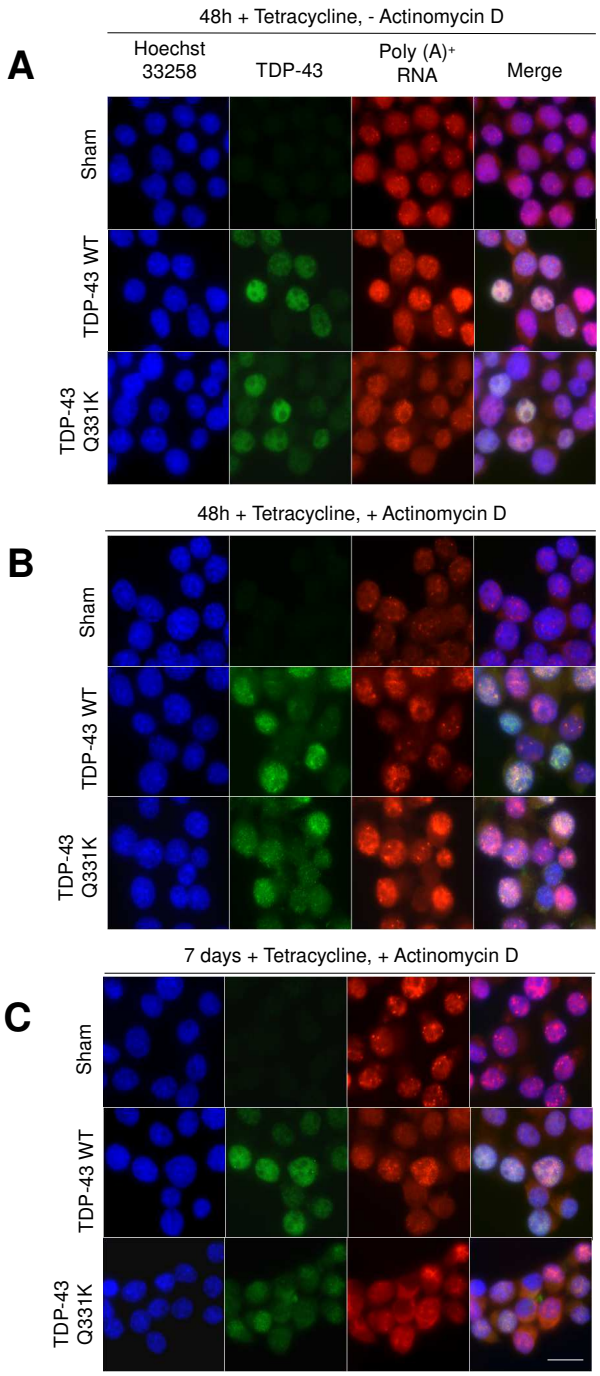


Figure 5.12 Human TDP-43 and poly(A)$^+$ mRNA localisation in NSC-34 Flp cell models following treatment with Actinomycin D

Human TDP-43 localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated and left 24 h before the addition of 1 µg/mL Tetracycline. Cells were stained with total TDP-43 (green), poly(A)$^+$ mRNA (red) and Hoechst (blue). (A) NSC-34 Flp In cells grown in the prescence of Tetracycline for 48 h and the absence of Actinomycin D. Human TDP-43 was absent in the Sham cell line and was predominantly nuclear in the TDP-43 WT and TDP-43 Q331K cell line. (B) NSC-34 Flp In cells in the presence of Tetracycline for 48 h and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. TDP-43 WT and TDP-43 Q331K cell lines have predominantly nuclear TDP-43 and poly(A) $^+$ RNA (C) NSC-34 Flp In cells in the presence of Tetracycline for 7 days and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. TDP-43 WT cells have predominantly nuclear human TDP-43 protein. Nuclear loss of human TDP-43 can be seen in TDP-43 Q331K cells. Scale bar = 20 µm.

This was repeated using the antibody recognising total TDP-43 (Figure 5.13). In the presence and absence of Actinomycin D, cells induced for 48 h with Tetracycline showed predominantly nuclear TDP-43 protein and poly(A)$^+$ mRNA. NSC-34 Flp Sham and TDP-43 WT cells induced for 7 days and treated with Actinomycin D, again showed predominantly nuclear TDP-43 and poly(A)$^+$ mRNA. NSC-34 Flp TDP-43 Q331K cells induced for 7 days and treated with Actinomycin D showed nuclear loss and cytoplasmic accumulation of TDP-43 and poly(A)$^+$ mRNA.

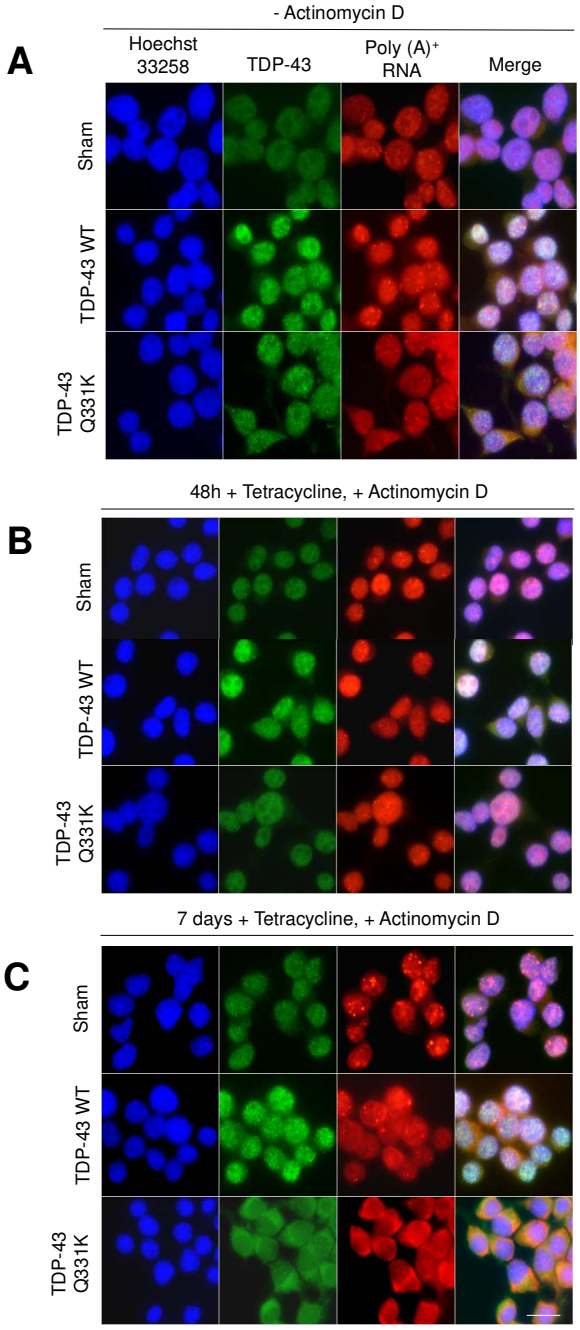


Figure 5.13 TDP-43 and poly(A)$^+$ mRNA localisation in NSC-34 Flp cell models following treatment with Actinomycin D

Total TDP-43 localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated and left 24 h before the addition of 1 µg/mL Tetracycline. Cells were stained with total TDP-43 (green), poly(A)$^+$ mRNA (red) and Hoechst (blue). (A) NSC-34 Flp In cells grown in the prescence of Tetracycline and the absence of Actinomycin D. TDP-43 was predominantly nuclear. (B) NSC-34 Flp In cells in the presence of Tetracycline for 48 hours and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. All cells had predominantly nuclear TDP-43 and poly(A)$^+$ RNA (C) NSC-34 Flp In cells in the presence of Tetracycline for 7 days and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. NSC-34 Flp In Sham and TDP-43 WT cells had predominantly nuclear TDP-43 and poly(A)$^+$ RNA. NSC-34 Flp In TDP-43 Q331K cells had nuclear loss and cytoplasmic accumulation of TDP-43 and poly(A)$^+$ RNA. Scale bar = 20 µm.

The nuclear loss of TDP-43 protein, nuclear loss of poly(A)$^+$ mRNA and nuclear loss of both TDP-43 protein and poly(A)$^+$ mRNA following treatment with Actinomycin D was quantified for cells that had been induced with Tetracycline for 48 h (Figure 5.14). There was no significant difference between the number of cells with both loss of TDP-43 and poly(A)$^+$ mRNA in the NSC-34 Flp TDP-43 Q331K compared to NSC-34 Flp Sham and TDP-43 WT.

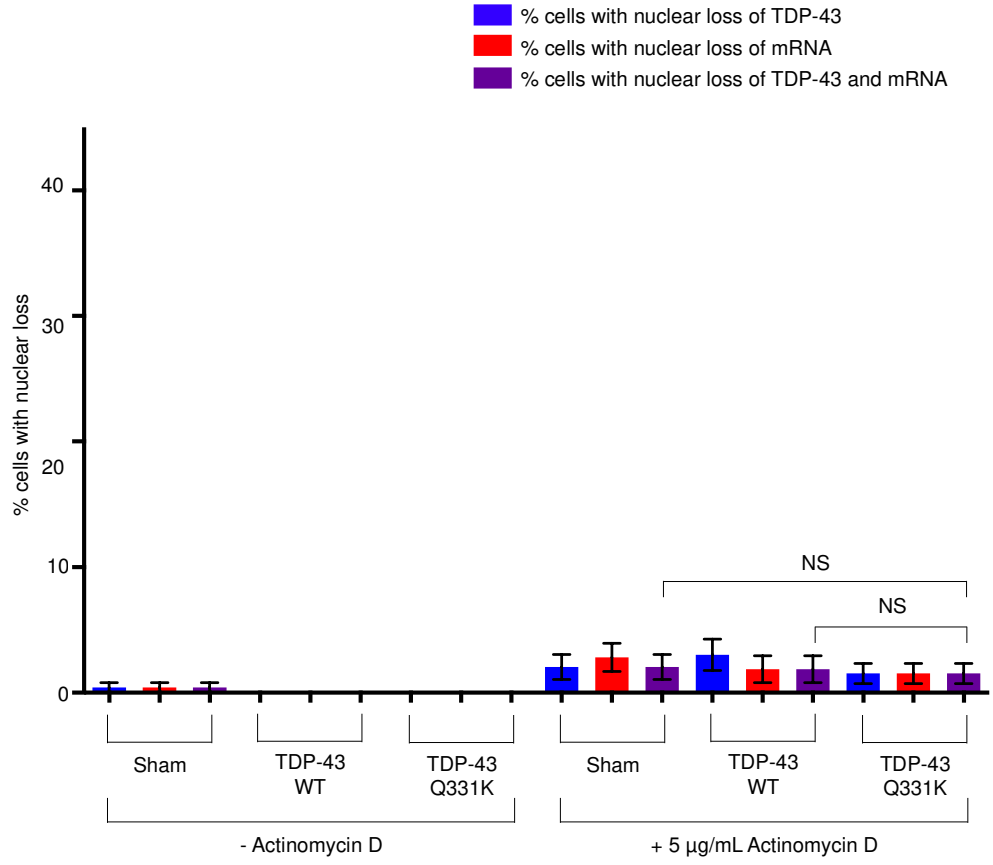


Figure 5.14 Nuclear loss of TDP-43 and poly(A)$^+$ mRNA in NSC-34 Flp cell models induced for 48 h following Actinomycin D treatment Cells were plated onto coverslips for 24 h before induction with 1 µg/mL Tetracycline for 48 h. Cells were untreated/treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. 20 fields of view were taken for each condition and the cells with nuclear loss of total TDP-43, poly(A)$^+$ mRNA and both TDP-43 and poly(A)$^+$ mRNA were counted. The number of cells with loss of TDP-43, poly(A)$^+$ mRNA and both TDP-43 and poly(A)$^+$ mRNA were expressed as a percentage of the total number of cells per field of view. (Mean ± SEM, Two-way ANOVA with Tukey's multiple comparison test, NS: non-significant; N (cells Sham/TDP-43 WT/TDP-43 Q331K) = -ActinomycinD: 282/285/296, + 5 µg/mL Actinomycin D: 248/274/296)

The nuclear loss of TDP-43 protein, nuclear loss of poly(A)$^+$ mRNA and nuclear loss of both TDP-43 protein and poly(A)$^+$ mRNA following treatment with two

concentrations of Actinomycin D was quantified for cells that had been induced with Tetracycline for 7 days (Figure 5.15). Following treatment with 5 µg/mL Actinomycin D 27.8 % of NSC-34 Flp TDP-43 Q331K showed nuclear loss of both TDP-43 and poly(A)$^+$ mRNA. There was no significant difference in the number of cells in NSC-34 Flp TDP-43 Q331K with nuclear loss of TDP-43 compared to cells with nuclear loss of both TDP-43 and poly(A)$^+$ mRNA, suggesting that loss of TDP-43 is accompanied by nuclear loss of poly(A)$^+$ mRNA. This was significantly greater than in the NSC-34 Flp Sham and TDP-43 WT cells, where less than 2 % of cells showed nuclear loss of both TDP-43 and poly(A)$^+$ mRNA.

Following treatment with 2 µg/mL Actinomycin D only 4 % of NSC-34 Flp TDP-43 Q331K showed nuclear loss of both TDP-43 and poly(A)$^+$ mRNA. There was no significant difference compared to NSC-34 Flp Sham and TDP-43 WT cells, where less than 2 % of cells showed nuclear loss of both TDP-43 and poly(A)$^+$ mRNA. Again, there was no significant difference in the number of cells with nuclear loss of TDP-43 compared to cells with nuclear loss of both TDP-43 and poly(A)$^+$ mRNA, suggesting that loss of TDP-43 is accompanied by nuclear loss of poly(A)$^+$ mRNA.

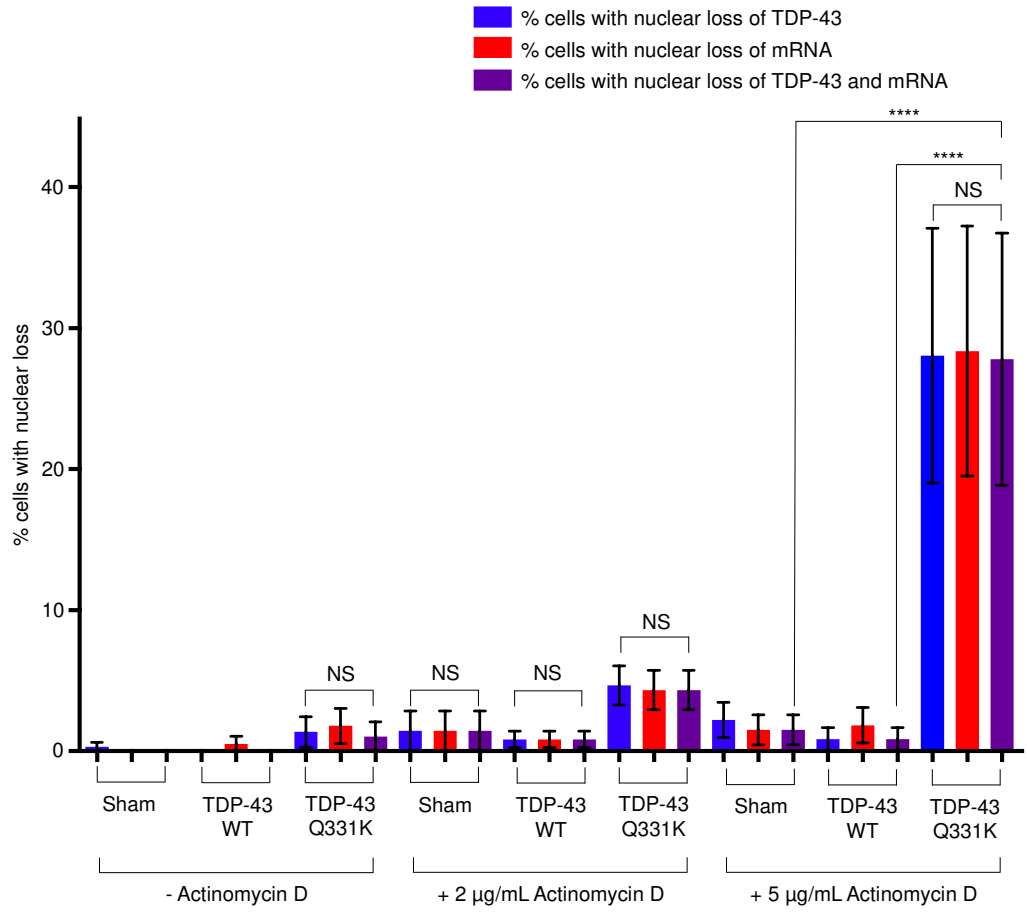


Figure 5.15 Nuclear loss of TDP-43 and poly(A)$^+$ mRNA in NSC-34 Flp cell models induced for 7 days following Actinomycin D treatment

Cells were plated for 24 h before induction with 1 µg/mL Tetracycline. Cells were induced for 48 h before plating onto coverslips. The media was changed every 48 h to refresh Tetracycline. Cells were untreated/treated with 2 µg/mL Actinomycin D or 5 µg/mL Actinomycin D for 1.5 h prior to fixing. 20 fields of view were taken for each condition and the cells with nuclear loss of total TDP-43, poly(A)$^+$ mRNA and both TDP-43 and poly(A)$^+$ mRNA were counted. The number of cells with loss of TDP-43, poly(A)$^+$ mRNA and both TDP-43 and poly(A)$^+$ mRNA were expressed as a percentage of the total number of cells per field of view. (Mean ± SEM, Two-way ANOVA with Tukey's multiple comparison test, ****: $p < 0.0001$, NS: non-significant; N (cells Sham/TDP-43 WT/TDP-43 Q331K) = -ActinomycinD: 281/292/345, + 2 µg/mL Actinomycin D: 203/213/341, + 5 µg/mL Actinomycin D: 294/341/356)

The effect of Actinomycin D treatment on total TDP-43 and human TDP-43 protein levels was investigated (Figure 5.16). The total TDP-43 protein levels remained unchanged, with no significant difference in TDP-43 levels between the cell lines. There was a significant increase in the amount of human TDP-43 in the NSC-34 Flp TDP-43 WT compared to the NSC-34 Flp TDP-43 Q331K, by 5.5 times.

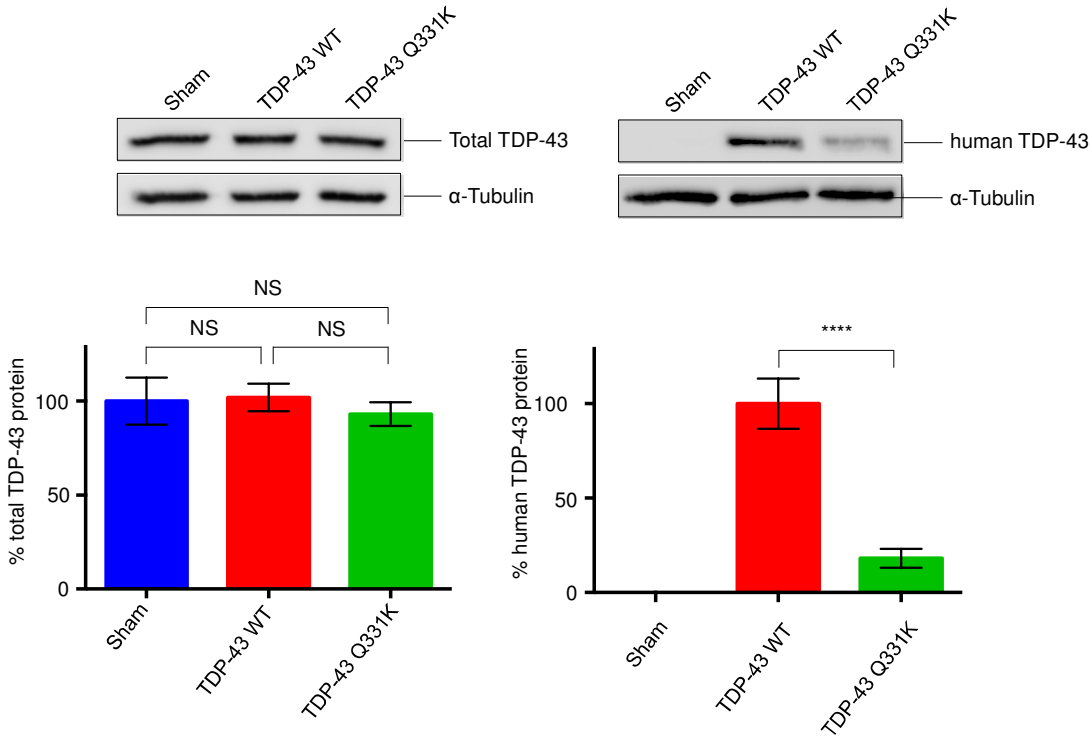


Figure 5.16 The effect of Actinomycin D treatment on total TDP-43 and human TDP-43 protein levels in NSC-34 Flp cell models

Western blot analysis of NSC-34 Flp cell lines grown in the presence 1 µg/mL of Tetracycline for 7 days and treated with 5 µg/mL Actinomycin D for 1.5 h prior to lysis. Graph shows quantification from western blot image. Total TDP-43 - each bar was normalised to NSC-34 Flp Sham (Mean ± SEM; One-way ANOVA with Tukey's multiple comparison, NS: non-significant; N = 3). Human TDP-43 - each bar was normalised to NSC-34 Flp TDP-43 WT (Mean ± SEM; One-way ANOVA with Tukey's multiple comparison, ****: $p < 0.0001$; N = 5)

Cell proliferation was investigated following treatment with Actinomycin D (Figure 5.17). There was no significant difference in the proliferation of NSC-34 Flp Sham and NSC-34 Flp TDP-43 WT cells which had been treated with Actinomycin D following induction with Tetracycline for 48 h. There was a significant but very small reduction in the proliferation of NSC-34 Flp TDP-43 Q331K treated with Actinomycin D following induction with Tetracycline for 48. There was no significant difference in the cell proliferation of any of the cell lines following treatment with Actinomycin D following a 7-day induction with Tetracycline.

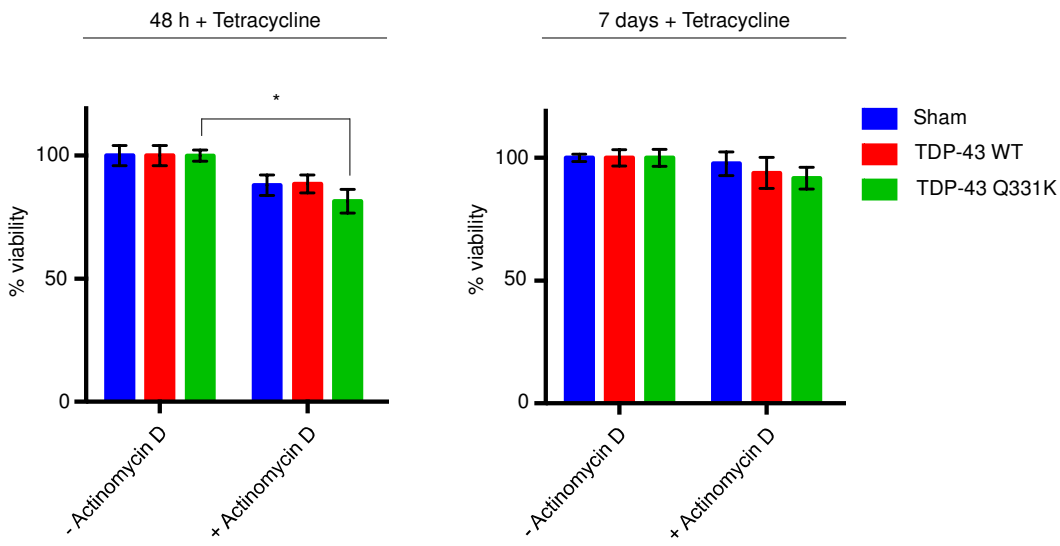


Figure 5.17 The effect of Actinomycin D on cell proliferation in NSC-34 Flp cell models

MTT assays on NSC-34 Flp cell lines grown in the presence of 1 µg/mL Tetracycline for 48 h (A) and 7 days (B). Cells were treated with DMSO (-Actinomycin D) or with 5 µg/mL Actinomycin D (+ Actinomycin D) for 1.5 h before lysis. Each bar was normalised to – Actinomycin D within each cell line for each time point (48 h or 7 days) separately. No significant differences were observed between NSC-34 Flp Sham and TDP-43 WT cells treated with or without Actinomycin D at 48 h. There was a significant difference between NSC-34 Flp TDP-43 Q331K cells treated with or without Actinomycin D at 48 h. No significant differences were observed between cells treated with or without Actinomycin D at 7 days. (Mean ± SEM; 2way ANOVA with Tukey's multiple comparison, *: $p < 0.1$; N (independent experiments, 6 wells per experiment) = 3).

TDP-43 inclusions have been shown to contain phosphorylated and ubiquitinated fragments of TDP-43 and co-localise with the inhibited-autophagy marker p62 (Neumann et al., 2006; Arai et al., 2006). Phosphorylated TDP-43 and p62 co-localisation was investigated in cells that had been treated with Actinomycin D following induction with Tetracycline for 48 h (P) and 7 days (Figure 5.19). After a 48 h induction with Tetracycline, pTDP-43 is predominantly nuclear, with some co-localisation with p62 in the cytoplasm, in both untreated and treated cells.

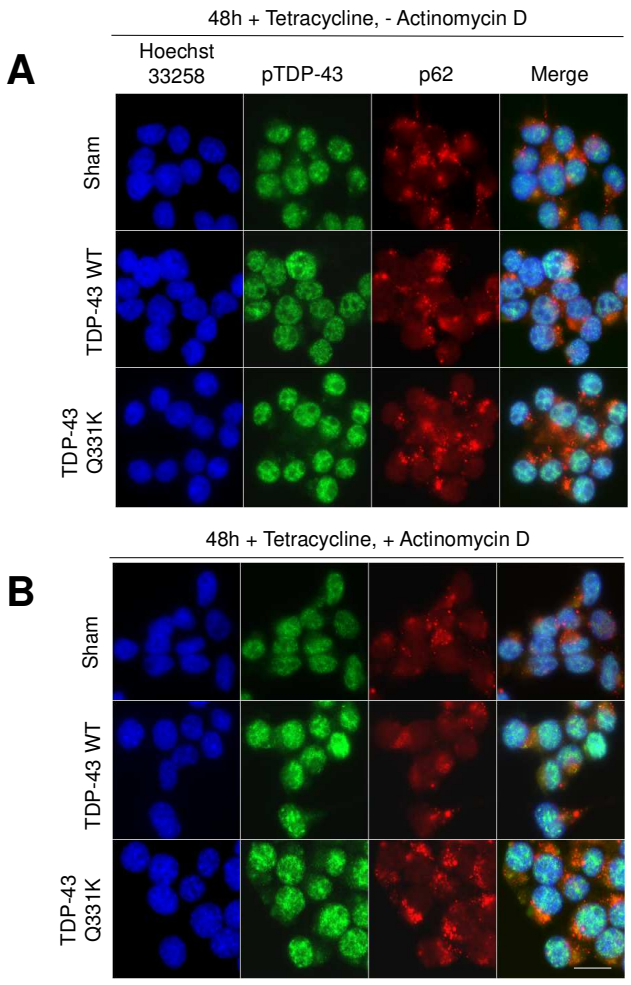


Figure 5.18 Phosphorlyated TDP-43 and p62 co-localisation after treatment with Actionmycin D in NSC-34 Flp cell models induced for 48 h
Phosphorylated TDP-43 (pTDP-43) and p62 co-localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated and left 24 h before the addition of 1 µg/mL Tetracycline. Cells were stained with pTDP-43 (green), p62 (red) and Hoechst (blue). (A) NSC-34 Flp In cells grown in the prescence of Tetracycline for 48 h and the absence of Actinomycin D. (B) NSC-34 Flp In cells in the presence of Tetracycline for 48 h and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. pTDP-43 was predominantly nuclear with some co-localisation with p62 in the cytoplasm in untreated and treated cells. Scale bar = 20 µm.

After a 7-day induction with Tetracycline, pTDP-43 is predominantly nuclear with some co-localisation with p62 in the cytoplasm in cells not treated with Actinomycin D. When treated with Actinomycin D the pTDP-43 is predominantly nuclear with some co-localisation with p62 in the cytoplasm in NSC-34 Flp Sham and TDP-43 WT. However, in NSC-34 Flp TDP-43 Q331K, pTDP-43 is lost from the nucleus and accumulates in the cytoplasm. This accumulation of pTDP-43 in the cytoplasm correlates with an increased co-localisation with p62.

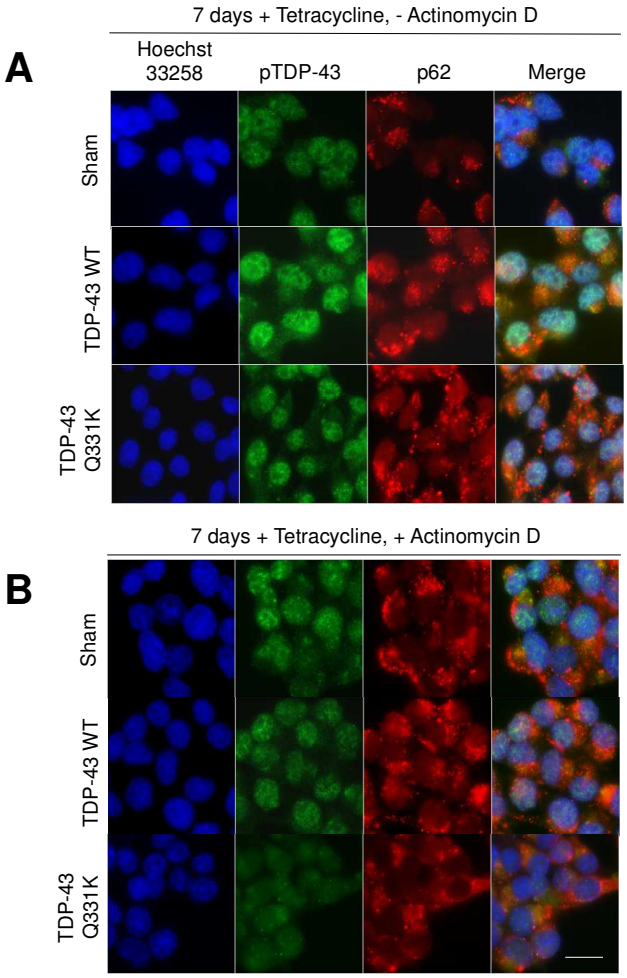


Figure 5.19 Phosphorylated TDP-43 and p62 co-localisation after treatment with Actionmycin D in NSC-34 Flp cell models induced for 7 days
Phosphorylated TDP-43 (pTDP-43) and p62 co-localisation in NSC Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated and left 24 h before the addition of 1 µg/mL Tetracycline. Cells were stained with pTDP-43 (green), p62 (red) and Hoechst (blue). (A) NSC-34 Flp In cells grown in the prescence of Tetracycline for 7 days and the absence of Actinomycin D. (B) NSC-34 Flp In cells in the presence of Tetracycline for 7 days and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. pTDP-43 was predominantly nuclear with some co-localisation with p62 in the cytoplasm in untreated and treated cells. Scale bar = 20 µm.

TDP-43 proteinopathy was not previously detected in the HEK Flp cells and so the effect of Actinomycin D on the HEK Flp TDP-43 cell lines was studied. HEK Flp Sham, TDP-43 WT and TDP-43 Q331K cells were treated with Actinomycin D to investigate the TDP-43 and poly(A)$^+$ mRNA localisation (Figure 5.20). In the absence and presence of Tetracycline, the TDP-43 protein and poly(A)$^+$ mRNA is predominantly nuclear when treated with Actinomycin D. The length of

Tetracycline induction has no effect upon the TDP-43 and poly(A)$^+$ mRNA localisation when treated with Actinomycin D.

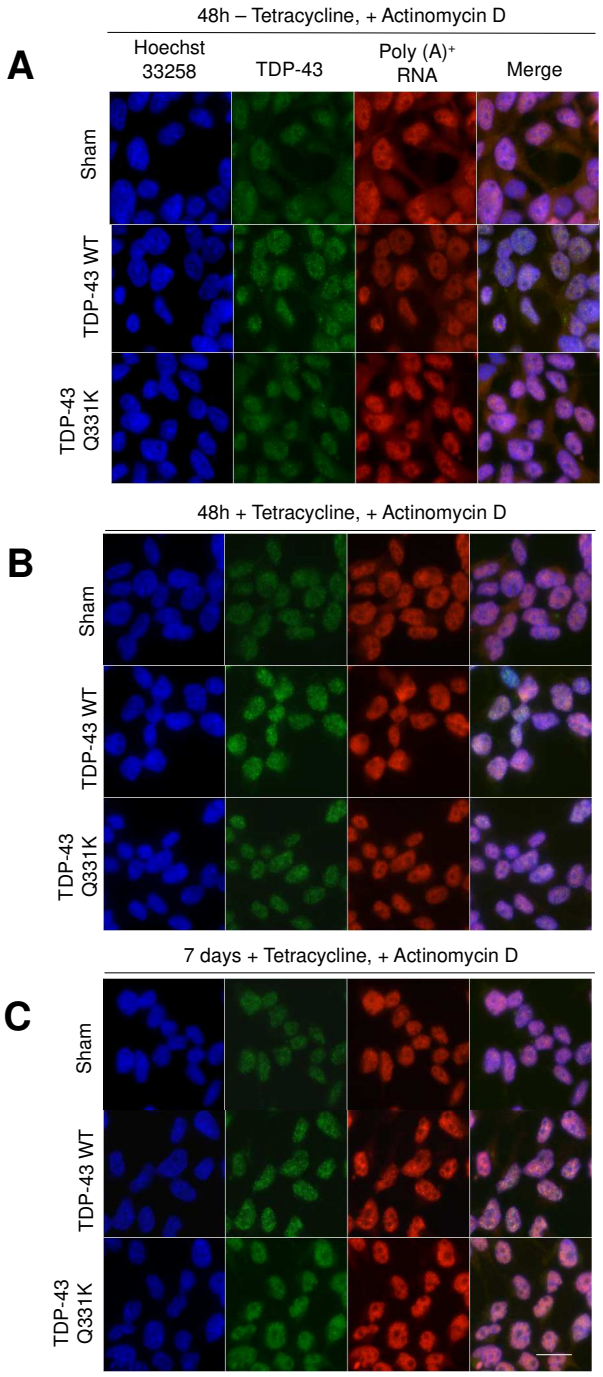


Figure 5.20 TDP-43 and poly(A)$^+$ mRNA localisation in HEK Flp cell models following treatment with Actinomycin D

TDP-43 localisation in HEK Flp Sham, TDP-43 WT and TDP-43 Q331K. Cells were plated and left 24 h before the addition of 10 µg/mL Tetracycline. Cells were stained with total TDP-43 (green), poly(A)$^+$ mRNA (red) and Hoechst (blue). (A) HEK Flp In cells grown in the absence of Tetracycline for 48 h and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. (B) HEK Flp In cells in the presence of Tetracycline for 48 h and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. (C) HEK Flp In cells in the presence of Tetracycline for 7 days and treated with 5 µg/mL Actinomycin D for 1.5 h prior to fixing. In the absence or presence of Tetracycline for 48 h and 7 day, HEK Flp cells show predominantly nuclear TDP-43 protein and predominantly nuclear poly(A)$^+$ mRNA when treated with Actinomycin D. Scale bar = 20 µm.

5.3 Discussion

TDP-43 was predominantly nuclear in all HEK cell lines following induction with Tetracycline (Figure 5.1). Increasing the length of induction from 48 h to 14 days had no effect upon the localisation of TDP-43. These results are consistent with other HEK Flp TDP-43 ALS cell models which do not exhibit TDP-43 proteinopathy (Ling et al., 2010).

However, this is in contrast to motor neurons from patients and the TDP-43 A315T mouse model of ALS, which show a mislocalisation of TDP-43 from the nucleus, which accumulates and forms aggregates in the cytoplasm (Neumann et al., 2006; Stallings et al., 2010). The fact that the HEK Flp TDP-43 Q331K cell model does not produce proteinopathy could be due to a number of reasons. TDP-43 proteinopathy is only observed in the CNS of ALS patients and not in other disease cell types. HEK cells are not neuronal and are human embryonic kidney cells, so simply may not exhibit proteinopathy because of a cell-specific resistance to this pathological change. Another difference between HEK cells and motor neurons, is cell division. Motor neurons do not divide, whereas HEK cells are actively dividing cells. Immunofluorescence studies were carried out following 48 h and 14-day induction with Tetracycline and ALS is an age of onset disease, perhaps by 14 days the cells are not aged sufficiently to see proteinopathy or cell division prevents the cells from ageing sufficiently to observe proteinopathy. Another explanation could come from the immunofluorescence studies being carried out under basal conditions. In ALS it has been shown that motor neurons are under oxidative stress and the chemical environment in ALS will be very different from the cell culture conditions deployed in this study. The addition of a cellular insult such as hydrogen peroxide (Richardson et al., 2013) or cumene hydroperoxide exposure (Argüelles et al., 2013) to induce oxidative stress may provide a more similar

chemical environment to that seen in ALS and this could cause the TDP-43 proteinopathy in cells to occur.

No neuronal Flp In system is commercially available. Therefore, the Flp In system was integrated into motor neuron-like NSC-34 cells (Figure 5.2) and these cells were used to generate NSC-34 Sham, TDP-43 WT and TDP-43 Q331K cell lines. Other TDP-43 mutation cell lines were generated but not used in this study.

NSC-34 cells are murine, hybrid cell line generated by the fusion of neuroblastoma with motor neuron enriched spinal cord cells. The advantage of these cells is that they exhibit properties of primary motor neurons such as extending processes, synthesising and storing acetylcholine, expressing neurofilament proteins and supporting action potentials (Cashman et al., 1992). However, unlike motor neurons, due to the fusion with neuroblastoma cells, they divide and grow in culture. The advantage of this is that it enables the large amounts of material required for biochemical studies to be produced in a short time. Integration of the FRT into the NSC-34 cells took approximately 12 months to achieve. Generation of subsequent cell models from the NSC-34 FRT (or HEK FRT) cells took approximately 2-3 months. This is relatively quick compared the length of time it would take to create transgenic mouse models.

NSC-34 Flp TDP-43 WT and TDP-43 ALS mutations showed induction of human TDP-43 on the addition of Tetracycline, showing successful integration of the Tetracycline-inducible system into NSC-34 cells (Figure 5.3). There was very little or no human TDP-43 present in the absence of Tetracycline. This is because NSC-34 cells are murine and the integrated transgenes are human.

Any human TDP-43 present in the non-induced cell lines is due to the leakiness of the promoter.

The total TDP-43 shows the amount of both endogenous mouse TDP-43 and human TDP-43 expressed from the transgene. Upon the addition of Tetracycline the total level of TDP-43 remains constant. This is due to the auto-regulation by TDP-43. TDP-43 auto-regulates its own production to maintain stable cellular protein levels by binding to the 3' UTR of its mRNA before exosome-mediated mRNA degradation (Ayala et al., 2011). The inserted transgene does not contain the 3' UTR, in order to fully control expression from the transgene via the use of Tetracycline. Endogenous TDP-43 transcripts do contain the 3' UTR and can auto-regulate TDP-43 levels. This autoregulation of TDP-43 is likely to be the reason that no increase in total TDP-43 was observed. However, an increase in unregulated inserted human TDP-43 levels can be seen upon the addition of Tetracycline as the integrated human transgene lacks the 3' UTR.

The increase in TDP-43 upon the addition of Tetracycline in both the HEK and NSC-34 cell models is only moderate. This is important, as it shows that the models are not over-expression models and therefore much better reflect the levels of autosomal dominant inheritance of TDP-43 observed in autosomal dominant TDP-43-related ALS. Over-expression or transient transfection of TDP-43 would not be as useful, as such high levels are not observed in disease and the observed effects of this are most likely to be due to the high levels of TDP-43 and not due to expression of mutant TDP-43.

Human TDP-43 transcript levels were calculated in the NSC-34 cell models in the absence and presence of Tetracycline (Figure 5.4). No human TDP-43 transcripts were observed in the NSC-34 Flp Sham cell model. This is because

the NSC-34 model is a murine model without any human transgene integrated. Human TDP-43 transcripts are however observed in the non-induced NSC-34 Flp TDP-43 WT and TDP-43 Q331K, this is likely due to the leakiness of the promoter. There are 2.6 times the number of TDP-43 transcripts in the NSC-34 Flp TDP-43 WT compared to the TDP-43 Q331K. This difference in expression was also observed in the HEK Flp In cell lines. This difference in transcript level may be due to differences in the stability of the TDP-43 WT transcripts compared to those encoding TDP-43 Q331K.

Higher expression levels in the HEK Flp TDP-43 WT cell model caused a decrease in cell proliferation compared the HEK Flp TDP-43 Q331K (Figure 4.4). This decrease in cell proliferation due to higher expression of TDP-43 WT compared to TDP-43 Q331K was also observed in NSC-34 cells at 16 days post Tetracycline induction (Figure 5.5). At 48 h post induction with Tetracycline, there was a significant difference in cell proliferation between the NSC-34 Sham and NSC-34 Flp TDP-43 cell models. However, no significant difference was observed in between the NSC-34 Flp TDP-43 WT and TDP-43 Q331K, despite the difference in expression levels of human TDP-43.

In different concentrations of Tetracycline, the total TDP-43 levels remained unchanged (Figure 5.7). This could be due to TDP-43 auto-regulating its levels, as discussed previously. This is also further evidence that the models are not over-expression models. At all Tetracycline concentrations recorded there was approximately 3 times more human TDP-43 protein produced in the NSC-34 Flp TDP-43 WT compared to TDP-43 Q331K. However, when using 10 ng/mL Tetracycline in the NSC-34 Flp TDP-43 WT and 100 ng/mL Tetracycline in the NSC-34 Flp TDP-43 Q331K similar levels of human TDP-43 were observed.

Therefore, using different concentrations of Tetracycline to induce cells, may give more similar human TDP-43 protein levels.

MTT assays using the range of Tetracycline showed that the concentration of Tetracycline used had no effect upon the cell proliferation of NSC-34 Flp Sham compared to NSC-34 Flp TDP-43 cell models (Figure 5.8). This was important to check, as perhaps higher concentrations of Tetracycline may have had an effect on the cell proliferation compared to lower Tetracycline concentrations. This showed that the Tetracycline did not have an effect upon cell proliferation at the concentrations explored and that any differences observed using that concentration would be due to the production of human TDP-43 and not due to the addition of Tetracycline. (Figure 5.5). This result is in contrast to the previous cell proliferation assays carried out at 10 µg/mL where a significant reduction in cell proliferation (by 30 %) was observed in NSC-34 Flp TDP-43 WT and TDP-43 Q331K compared to NSC-34 Flp Sham following induction with Tetracycline for 48 h (Figure 5.5) were normalised to the NSC-34 Flp Sham cell line. This heavily relies upon exactly the same the number of cells being plated for each cell line. The Tetracycline concentration MTT assay (Figure 5.8) normalised to cells without Tetracycline induction and the proliferation at each concentration compared. Normalisation to non-induced cells accounts for any differences in cell number, as the same number of cells would be plated for induced and non-induced at the start of the experiment.

This highlights the some of the weaknesses with MTT assays. MTT assays are dependent upon many different factors. They require the same number of cells to be seeded in each well in order to compare different time points or Tetracycline concentrations. External environmental factors can also influence the MTT assays. Any fluctuations in temperature and carbon dioxide levels due

to the frequent opening of incubators can affect the cell proliferation. The MTT assay is a good quick screen in order to determine any differences in the proliferation of cells by measuring the activity of mitochondria. However, they may not represent the best choice of assay for a disease where mitochondrial alteration has been reported, such as ALS (Ferraiuolo et al., 2011). For these reasons the growth curve, albeit more time consuming and laborious, is a more accurate assay to assess cell growth over a time period following the induction of cells with Tetracycline. However, the external environment as discussed for the MTT assay needs to be well controlled for a growth curve also, as these factors could affect the growth of cells, regardless of induction with Tetracycline.

Growth curves were carried out using 10 ng/mL Tetracycline in the NSC-34 Flp TDP-43 WT and 100 ng/mL Tetracycline in the NSC-34 Flp TDP-43 Q331K so upon on the addition of Tetracycline similar amounts of human TDP-43 was produced (Figure 5.9). The NSC-34 Flp Sham growth curve was carried out using the highest Tetracycline concentration. No significant difference in the growth of NSC-34 Flp TDP-43 WT compared to Sham was observed until day 18; this difference was also significant at day 21. No significant difference in the growth of NSC-34 Flp TDP-43 Q331K compared to Sham was observed until day 21. There was no difference in the growth of NSC-34 Flp TDP-43 WT compared to TDP-43 Q331K. These results suggest that the expression of human TDP-43, WT or mutant, needs to be for a longer period of time before any difference in cell growth is observed. This is not unexpected, as ALS is a late age of onset disease. The reduction in growth observed in the NSC-34 Flp TDP-43 WT cell model could be due to the higher levels of TDP-43 expression, as regulation of TDP-43 levels is important and perhaps higher levels (even though they are much lower than over-expression levels) may still have an effect upon cell growth.

The normalisation of human TDP-43 expression did not show any ALS-specific difference in cell growth. For these reasons it was decided that the concentration of Tetracycline to be used on all cell lines from hereon would be 1 µg/mL rather than 10 µg/mL. This was because experiments had shown that at both concentrations almost identical amounts of human TDP-43 were expressed. The use of a lower concentration would be beneficial in reducing any effects the Tetracycline may have upon the experiments performed.

The nuclear loss and cytoplasmic accumulation and aggregation of TDP-43, referred to as TDP-43 proteinopathy, are characteristic hallmarks of ALS. Transient transfection or over-expression of TDP-43 have been widely use to cause the cytoplasmic aggregation of TDP-43 in cells, but these involve much higher levels of TDP-43 than what would be involved in the human disease. A Tetracycline-inducible model with lower levels of TDP-43, which reproduced TDP-43 proteinopathy, would be a much more disease relevant model.

No nuclear loss of human TDP-43 was observed following induction with Tetracycline for 7 and 14 days with 1 µg/mL or 10 µg/mL Tetracycline (Figure 5.6, Figure 5.10 respectively). Not all cells are expressing human TDP-43 at the point the cells are fixed; so looking specifically at the human TDP-43 expression may not be the best way to identify TDP-43 proteinopathy. Therefore an antibody recognising both human and mouse TDP-43 was used to investigate total TDP-43 proteinopathy.

Serum depletion of NSC-34 cells has been shown to promote expression of functional glutamate receptors as cells mature into a form, which bear phenotypic resemblance to motor neurones (Eggett et al., 2000; C. Yang et al.,

2010; Maier et al., 2013; Kanjilal et al., 2014). Serum withdrawal from the NSC-34 cell models may induce differentiation resulting in more phenotypic resemblance to motor neurons, which may in turn result in TDP-43 proteinopathy when human mutant TDP-43 is expressed. The serum in the culture media was reduced to a quarter of its normal concentration and total TDP-43 proteinopathy was investigated after 48 h and 9-day induction (Figure 5.11). After 48 h no differences in TDP-43 localisation were observed, the TDP-43 was predominantly nuclear in all cell lines. Serum withdrawal caused an increase in the number of processes, but was also quite toxic to the cells. After 9 days withdrawal and induction with Tetracycline, an increase in cytoplasmic TDP-43 was observed, but no nuclear loss of TDP-43 was seen. Serum withdrawal was toxic and the cells were less viable after serum withdrawal. In order to study RNA dysregulation in a TDP-43 model of ALS we would not want to induce proteinopathy and extract RNA from cells which are not growing well in culture, as these cells would show large alterations to RNA metabolism due to the high stress levels the cells are under in addition to any ALS specific alterations. This would not be ideal to study RNA dysregulation following such a toxic insult to the cells. Therefore, differentiation of the NSC-34 cell models was not continued.

As discussed previously, the environment may have an effect upon TDP-43 proteinopathy. The culture of NSC-34 cells is slightly different to that of HEK cells, with their neuronal properties making them more sensitive to sudden environmental changes. Hence the changes to culture and immunofluorescence protocols for NSC-34 cells, with lower concentrations of selection antibiotics for the culture of cells and gelatine coating of coverslips prior to immunofluorescence studies. The NSC-34 cell models were exposed to different stresses such as low seeding density, and being removed from the incubator

and kept at room temperature for periods of time each day (Data not shown). This was with the idea that if the cells were stressed, they may be more likely to show TDP-43 proteinopathy. However, these additional 'stresses' had no effect upon the localisation of TDP-43 within the cell models.

It is also worth noting that the NSC-34 Flp TDP-43 Q331K cell line when cultured had a different morphology to the NSC-34 Flp Sham and TDP-43 WT cell lines. The cell line grows less uniformly with small clusters of cells with very defined projections. This difference could be observed in the absence of Tetracycline also. There is 'leakiness' from the promoter, as previously discussed, therefore slight induction of mutant TDP-43 could have an effect upon the cell morphology. This would be worth further investigation.

Analysis of the TDP-43 ALS translatome highlighted mRNA transport/export as an enriched dysregulated pathway. Other work within the laboratory has demonstrated that TDP-43 is part of the TREX complex and interacts with the export adaptor NXF1 (Data not shown). Furthermore, mRNP capture assays also showed TDP-43 ALS mutant proteins have reduced binding to poly-adenylated RNA (Data not shown). Altered recruitment of TDP-43 proteins to mRNPs would likely slow the rate of co-transcriptional processing of mRNA, which over a long duration of disease could lead to additional alterations such as R-loops and DNA damage which have been reported in ALS (Haeusler et al., 2014). The DNA damage/response pathways were significantly enriched in the RNA sequencing analysis (Chapter 5). In ALS there is a long disease progression. Slowing transcription using Actinomycin D, an inhibitor of transcription, dramatically enhanced the TDP-43 proteinopathy.

Fluorescence in-situ hybridisation of poly $(A)^+$ mRNA was carried out on the NSC-34 cell models following induction with Tetracycline. Cells were treated with Actinomycin D, an antibiotic that inhibits transcription by binding to DNA at the transcription initiation complex and inhibiting elongation of the RNA chain by RNA polymerase. Human TDP-43 was lost from the nucleus and accumulated in the cytoplasm in NSC-34 Flp TDP-43 Q331K cells following a 7-day induction, but remained predominantly nuclear in NSC-34 Flp Sham and TDP-43 WT (Figure 5.12). No nuclear loss of human TDP-43 was observed at 48 h in any of the cell lines. The number of cells with nuclear loss of poly $(A)^+$ mRNA and TDP-43 was not quantified for either timepoint when using the human specific antibody, due to the lower protein levels when compared to total TDP-43. However, it is clear from the images that when there is nuclear loss of human TDP-43 there is also nuclear loss of poly$(A)^+$ mRNA, as observed when looking at total TDP-43.

Fluorescence in-situ hybridisation of poly$(A)^+$ mRNA was repeated using the total TDP-43 antibody on the NSC-34 cell models following induction with Tetracycline (Figure 5.13). NSC-34 Flp TDP-43 Q331K cells that had been induced for 7 days and treated with Actinomycin D showed nuclear loss of total TDP-43 and poly$(A)^+$ mRNA, which accumulated in the cytoplasm. This was not observed in NSC-34 Flp Sham or TDP-43 WT cells which had been induced for 7 days or in any cells which had been induced for 48 h and treated with Actinomycin D or induced for 48 h without Actinomycin D treatment. This shows that it is both endogenous TDP-43 and human TDP-43 expressed from the transgene, which is mislocalised from the nucleus to the cytoplasm after treatment with Actinomycin D.

The number of cells with nuclear loss of poly(A)$^+$ mRNA, TDP-43 and both poly(A)$^+$ mRNA and TDP-43 was quantified for cells, which had been treated with Actinomycin D after a 48 h and 7-day induction with Tetracycline (Figure 5.15, Figure 5.15). After a 48 h induction there was no nuclear loss of TDP-43 and poly(A)$^+$ mRNA in the NSC-34 Flp TDP-43 Q331K compared to the NSC-34 Flp Sham and TDP-43 WT. Approximately 30 % of NSC-34 Flp TDP-43 Q331K cells showed nuclear loss of poly(A)$^+$ mRNA and TDP-43 after a 7-day induction with Tetracycline. There was no significant difference between the number of cells with nuclear loss of poly(A)$^+$ mRNA and nuclear loss of TDP-43, indicating that the nuclear loss of TDP-43 was accompanied by nuclear loss of poly(A)$^+$ mRNA, or vice-versa. As previously discussed, TDP-43 binds RNA and has many roles in the regulation of gene expression. A possible explanation for the cytoplasmic accumulation of poly(A)$^+$ mRNA when TDP-43 is lost from the nucleus could be that RNA bound to TDP-43 is also mislocalised upon the mislocalisation of TDP-43.

As Actinomycin D inhibits transcription, the protein levels of TDP-43 may be affected and this could be the reason for the nuclear loss of TDP-43 protein. Following treatment with Actinomycin D after a 7-day induction with Tetracycline, the total levels of TDP-43 and levels of human TDP-43 were not changed (Figure 5.16). This showed the difference in TDP-43 localisation in NSC-34 Flp TDP-43 Q331K cells was not due to any change in the protein levels following treatment with Actinomycin D.

Treatment with Actinomycin D following a 7-day induction with Tetracycline did not change the cell proliferation of the NSC-34 Flp cell lines (Figure 5.17). This showed that Actinomycin D treatment was not having any toxic effects upon cell proliferation. This, along with unchanged protein levels, showed that inhibition of

transcription caused a Q331K specific nuclear loss of poly(A)$^+$ mRNA and TDP-43 protein in the NSC-34 Flp cell lines.

After treatment with Actinomycin D following a 48 h induction with Tetracycline, there was a significant reduction in the proliferation of the NSC-34 Flp TDP-43 Q331K compared to Sham. This is surprising as after a 48 h induction no nuclear loss of TDP-43 was observed and following a longer induction with Tetracycline no reduction in proliferation is seen. There was a reduction in proliferation in all cell lines after treatment with Actinomycin D compared to untreated cells at 48 h. This was a small reduction of approximately 12 % in NSC-34 Flp Sham and TDP-43 WT compared to 19 % in NSC-34 Flp TDP-43 Q331K.

Observations from NSC-34 Flp cells which have been treated with Actinomycin D is that when the level of TDP-43 is similar in the nucleus and the cytoplasm, the poly(A)+ mRNA is also evenly distributed between the two cellular compartments. Some cells with some nuclear poly(A)+ mRNA and TDP-43 staining also have a cytoplasmic accumulation of poly(A)$^+$ mRNA and TDP-43. This suggests that the cytoplasmic accumulation of TDP-43 occurs prior to the nuclear loss of TDP-43. Similar levels of TDP-43 and poly(A)+ mRNA staining are also seen in NSC-34 Flp Sham and TDP-43 WT cells, and proteinopathy does not occur as frequently as observed in the NSC-34 Flp TDP-43 Q331K, suggesting there is some Q331K specific alteration which prevents normal protein localisation returning in the NSC-34 Flp TDP-43 Q331K cell line. This suggests that the cytoplasmic mislocalisation of TDP-43 occurs prior to the nuclear loss of TDP-43. Understanding the order of these events is important in understanding the precise molecular mechanisms underpinning TDP-43 proteinopathy in disease.

In disease, mislocalised TDP-43 forms aggregates which are hyperphosphorylated, ubiquitinated and cleaved to generate C-terminal fragments (Neumann et al., 2006; Arai et al., 2006). These aggregates are also positive for p62 (Mizuno, Amari, Takatama, Aizawa, Mihara, and Okamoto, 2006b). p62 plays a role in protein degradation via the UPS and autophagy pathway (Seibenhener et al., 2004) and aberrant protein degradation has been implicated in the pathogenesis of ALS. p62 has been implicated as having a protective role in pathological conditions due to its upregulation in cultured neuronal cells during initiation of apoptosis and proteasomal inhibition (Kuusisto et al., 2001). Therefore the accumulation of pTDP-43 and p62 in the NSC-34 Flp cell models was investigated to see whether there was an increase in phosphorylation of TDP-43 and whether p62 was upregulated, when TDP-43 was lost from the nucleus.

After a 48 h induction with Tetracycline and no Actinomycin D treatment, pTDP-43 was predominantly nuclear and similar cytoplasmic p62 levels are observed (Figure 5.18 A). After treatment with Actinomycin D there is no change in pTDP-43 and p62 localisation and expression levels (Figure 5.18 B). From previous experiments after a 48 h induction with Tetracycline and treatment with Actinomycin D no significant nuclear loss and cytoplasmic accumulation of TDP-43 was observed. Therefore without the nuclear loss and cytoplasmic accumulation of TDP-43, it is not expected that cytoplasmic aggregates of pTDP-43 and p62 would be observed.

After a 7-day induction with Tetracycline and no treatment with Actinomycin D, pTDP-43 was predominantly nuclear and similar cytoplasmic p62 levels were observed (Figure 5.19 A). However following treatment with Actinomycin D,

pTDP-43 is lost from the nucleus and accumulates in the cytoplasm in NSC-34 Flp TDP-43 Q331K cells (Figure 5.19 B).

At both time points following induction with Tetracycline, there was a co-localisation of cytoplasmic pTDP-43 and p62. However, no clear aggregates are observed. This may be because the cells are dividing and therefore aggregates do not have time to build up before the cell divides. ALS is a late age of onset disease and these aggregates are observed in post-mortem tissue. Therefore aggregates will have had a long period to build up in diseased motor neurons, whereas cells are dividing in the cell model.

Having established a method of producing TDP-43 proteinopathy in the NSC-34 Flp cell model, Actinomycin D treatment was performed on the HEK Flp cell models (Figure 5.20). Following treatment with Actinomycin D after a 48 h and 7-day induction with Tetracycline, TDP-43 remained predominantly nuclear. This indicates that a neuronal context is also likely to be important for TDP-43 proteinopathy to occur. This result is consistent with disease, as TDP-43 proteinopathy is only observed in the CNS of ALS patients and not in other cell types.

The nuclear loss of poly(A)$^+$ mRNA also occurs when TDP-43 proteinopathy is observed. This suggests an alteration to mRNA transport and export in TDP-43 ALS, a pathway that was also identified from the ALS translatome as potentially playing a role in disease (Chapter 4).

6. Discussion and Conclusions

ALS is a devastating neurodegenerative disease characterised by the loss of both upper and lower motor neurons (Kiernan et al., 2011). The pathogenesis of ALS is multifactorial and in the past decade altered RNA metabolism has been recognised as an important component of pathophysiology (Walsh et al., 2015). TDP-43 proteinopathy is a hallmark of ALS and is observed in 95 % of ALS cases both familial and sporadic (Keller et al., 2012). TDP-43 neurodegeneration involves widespread RNA dysregulation affecting up to a third of the transcriptome (Arnold et al., 2013). However, the key functional consequences of this dysregulation are unknown at protein level.

This project aimed to identify the functional consequences of RNA dysregulation in ALS by developing a new technology that would allow identification of mRNA molecules undergoing active translation at genome wide level, therefore producing a TDP-43 ALS translatome. This was to be done by generating GFP reporter cell lines in which to test different translatome methodologies, in order to develop a method that would preferentially isolate actively translating mRNA. This technology would then be applied to a specifically engineered human TDP-43 ALS-inducible model.

6.1 Advantages and limitations of inducible cell models

Cell models have been used to study disease mechanisms in ALS for many years (Veyrat-Durebex et al., 2013). Cell models are valuable tools in ALS research due to the nature of the disease. This is for a number of reasons. The ideal way to investigate the molecular events in ALS would be to obtain samples whilst a patient is alive, to better understand the spread of pathogenic events over time. However, pathological samples of brain and spinal cord cannot be

taken whilst the patient is alive and therefore any pathological samples used are post-mortem. RNA obtained from post-mortem samples is of low quality due to degradation. In addition, the study of surviving neurons from ALS patients may also not be the best choice when trying to establish the molecular events that have caused the death of other motor neurons. Particularly when surviving neurons will have accumulated hundreds if not thousands of gene expression dysregulations as a cause(s) and consequence of ALS. Motor neurons from animal models are also difficult to separate from other neuronal cells and grow for long periods in culture. This makes it difficult to obtain the quantities required for biochemical studies.

Inducible cell models are therefore powerful tools to investigate the molecular events in the pathogenesis of ALS and have several advantages over using post-mortem material and animal models. Cell models can be grown for long periods in culture and therefore it is easy to produce the large quantities required for biochemical studies. The ability to turn on and off expression of a particular gene of interest allows the study of disease progression. This may be useful when trying to distinguish between causal and downstream events. It is also a useful tool if over-expression of the gene of interest is toxic or detrimental to the cell's growth and survival.

The Invitrogen Tetracycline-inducible Flp In system allows stable, isogenic integration at a single FRT site. This isogenic integration enables cell models to be built with genetically identical backgrounds, where the only difference will be the inserted transgene. The Tetracycline-inducible control of the gene of interest enables the construction of cell models with possible knock down/mutations of essential genes and/or production of toxic proteins at convenience.

The Invitrogen Flp In system has been successfully used to generate HEK Flp *GFP* reporter cell models and non-neuronal HEK Flp cellular models of TDP-43 ALS. The FRT cassette has also been successfully integrated into NSC-34 cells and has been used to generate cellular models of TDP-43 ALS which exhibit Q331K specific TDP-43 proteinopathy.

One of the limitations of using HEK Flp In and NSC-34 Flp In cell models is that these cells are dividing, whereas motor neurons do not divide. Cells that are continuously dividing in culture will be able to clear protein aggregation and accumulation. Converting Tetracycline inducible cells into iNPCs, which can be differentiated into motor neurons or astrocytes, may provide a more disease relevant inducible model of disease. These cells could be converted into neurons or astrocytes before the induction of cells with Tetraycycline. This would still allow for the study of early events following induction with Tetracycline, but the cells would be in a more disease relevant context.

6.2 Challenges and importance of identifying cellular translatomes

Transcriptomic studies have been used over the last twenty years to interrogate the transcriptomes of prokaryotes and eukaryotes in development, immunity, ageing and disease. The challenges of investigating the translatome are illustrated by the small number of publications investigating the translatome (<100 in *PubMed*) compared to those investigating the transcriptome (>25,000 in *PubMed*).

Due to cell compensatory mechanisms, several studies highlighted that proteomes and transcriptomes do not correlate well (Vogel et al., 2010; Schwanhäusser et al., 2011). An up regulation in mRNA level does not necessarily correlate with an increase in protein level, but rather a down

regulation as a cell tries to counteract the down regulation of protein by increasing synthesis/stability of the mRNA (Domínguez-Sánchez et al., 2011). The abundance of mammalian proteins is mostly regulated by their translation rate (Schwanhäusser et al., 2011). Less than one third of human protein abundance can be attributed to global mRNA concentration (Vogel et al., 2010).

Cellular translatomes should provide a more accurate and complete measure of gene expression than by analysing total mRNA levels alone. This is because the identification of mRNA molecules actively being translated at the ribosome (the translatome) directly reflects protein expression changes and the directionality of altered biological processes in health and disease.

The biggest challenge in the identification of the translatome is ensuring that isolated mRNA is in fact undergoing active translation and is not merely associated with the ribosome. There is an oversight in the literature where methods such as TRAP (translating ribosome affinity purification) were developed without checking that at least some of the identified translated mRNAs were actually actively being translated by ribosomes.

In order to test this, inducible GFP reporter cell models were built in order to test whether current methodologies were identifying translating mRNA. HEK Flp *GFP* cells produced GFP protein, whereas HEK Flp *M1V* cells did not produce GFP protein due to lack of an AUG start codon (Figure 3.5). Both models produced comparable levels of GFP transcripts (Figure 3.8).

TRAP methodology was carried out using these cells models and it was surprising to see that there was no significant difference between the number of

GFP reporter mRNA identified using TRAP (Figure 3.10). This showed that TRAP recognised translating and non-AUG translating mRNAs equally.

This was interesting as it indicates that mRNAs lacking AUG start codons and failing to be translated into their cognate protein can still remain stably associated with ribosomes. This signal cannot be due to scanning ribosomes, as only the 43S small subunit (40S subunit and initiation factors) is involved in scanning before the recruitment of the 60S subunit and identification of the start codon (Kozak, 1978; Hinnebusch, 2011; Paek et al., 2015) and we tagged the large RPL10a ribosome subunits as in other TRAP studies. This indicates that the 60S subunit could still be recruited in absence of start codon. It is noteworthy that the Kozak sequence is still present in both GFP cell lines. Mutating the Kozak sequence could determine if this were the case, as if there was a decrease in the GFP detected in the HEK Flp M1V cells with a mutation in the Kozak sequence, this would very interestingly imply that the Kozak sequence could also recruit the 60S subunit in absence of AUG start codon.

This signal is not due to a later AUG start codon, as the transcripts lack all potential AUG start codons in all frames within the first two thirds of the GFP sequence. The qPCR primers were designed against a region in those two thirds, so any transcripts detected correspond to non-AUG driven initiation of translation. This suggests initiation from non-conventional codons might occur (S. Lee, Liu, Lee, Huang, Shen, and Qian, 2012a).

The GFP reporter cell models were used to develop the GRASPS method, to try and reduce contamination of non-AUG translating transcripts. A direct comparison of TRAP to GRASPS showed that the optimised GRASPS method

identified approximately 75 % less non-AUG translated GFP reporter mRNA compared to TRAP (Figure 3.19).

6.3 Comparison of GRASPS to other methods that identify ribosome-associated RNA

Polysome profiling, the first method developed to study translatome, allows the qualitative separation of polysomes using ultracentrifugation of sucrose gradients before the identification of polysomal mRNA by qRT-PCR, microarray or next generation RNA sequencing. These experiments are useful for measuring ribosome density and occupancy on mRNA, for screening translational changes and tracking the translational status of known mRNAs (Chassé et al., 2017).

However, the use of polysome profiling to study the translatome is hindered by the delicate manipulation needed for accurate fractionation using sucrose gradients. High molecular weight ribonucleoprotein complexes named pseudo-polysomes have been shown to contaminate polysomal fractions, further highlighting the need to investigate multiple sucrose gradient fractions (Thermann et al., 2007). These affect the reproducibility of polysome profiling and seriously limit its use in high throughput studies. Sucrose gradient fractionation involves typically two-three days of work and involves the collection and further analysis of many samples.

Ribosome profiling allows the mapping of ribosomes onto RNA molecules at a near nucleotide resolution (Ingolia et al., 2009). This method involves RNase digestion prior to sucrose gradient separation of ribosome-protected fragments (ribosome/monosome footprints) before identification of RNA by next generation sequencing. This method can identify the ribosomal positional information in

order to dissect initiation and elongation events, as well as identifying mechanisms of translational control (Ingolia, 2014).

The accuracy of ribosome profiling can be improved by PAGE size separation of ribosome protected fragments (D. W. Reid et al., 2015). However, the biggest limitation in analysing ribosome-protected fragments comes with the analysis and limited bioinformatics tools currently available. Ribosome profiling data differs from traditional RNA deep sequencing, the template is the position of one ribosome and therefore it measures the number of ribosomes translating a transcript, rather than the abundance of a transcript. Ribosome-protected fragments are short (approximately 25-30 nucleotides) preventing paired-end sequencing and making it challenging to accurately align repetitive sequences and alternative transcripts in these data (Ingolia, 2014). Ribosome profiling also degrades the 3' and 5' UTRs of transcripts, which may contain regulatory information.

Ribosome affinity purification does not rely on the use of sucrose gradients to purify ribosome-associated RNA (Gerber et al., 2006; Halbeisen et al., 2009). This method allows cell-specific purification of tagged ribosomal subunits. TRAP has been further developed using restricted neural expression of GFP tagged RPL10a in the molecular characterisation of CNS cell types (Heiman et al., 2008) and YFP-tagged RPL10a in subcellular domains in Purkinje neurons (Kratz et al., 2014). More recently it has also been used to investigate axonal translation in visual circuits (Shigeoka et al., 2016) and neuronal differentiation in *Drosophila* (Kelvin Xi Zhang et al., 2016). The use of tagging allows cell-type specific translatomes to be obtained. However, the tagging and over-expression of ribosomal subunits in TRAP may affect the assembly and/or structure of ribosomes, potentially disrupting regulatory mechanisms. The generation of

transgenic cell lines or animal models and the tagging of ribosomal subunits is also a lengthy and costly process. The TRAP method itself is also complicated by the need to tag ribosomal proteins. This tagging may cause steric hindrance to the ribosomes, which could affect the assembly of the ribosome and binding to the ribosome. Transfection of tagged ribosomal proteins also causes issues with reproducibility with differences in transfection efficiencies between experiments. Transfection also causes cellular stress, which can have an effect upon global translation within a cell, so also needs to be controlled for.

Unlike polysome profiling, GRASPS does not require the delicate manipulation of sucrose gradients. It also does not require the tagging of ribosomal subunits or generation of transgenic cell lines or animal models. The GRASPS protocol from the point of cell lysis to purified mRNA takes less than 48 h, making the length of time required for performing the experiment much shorter.

The GRASPS protocol allows for the purification of full length RNAs that are subsequently fragmented into approximately 100 nucleotide fragments for next generation RNA sequencing. However, these cover full length mRNA and the larger length considerably improves mapping to the genome up to 80-90% in this study (Table 4.4).

Moreover, GRASPS does not require the use of any translational inhibitors. Recently, ribosome profiling experiments using cycloheximide have revealed that translation inhibitors can induce artefactual ribosome occupancy in response to stress (Gerashchenko et al., 2014).

However, similarly to the analysis of ribosome-protected footprints, the bioinformatics tools to analyse translatome data from GRASPS are currently

being developed and the application of GRASPS will become more accessible once these methods have been established.

In summary, the simple biochemical precipitation of ribosomes in GRASPS provides a quick and simple, biochemical precipitation of ribosomes, without the need for transfection or the generation of transgenic cell or animal models, as it can be apply to any cells or tissue. Sufficient mRNA has been obtained when applying the GRASPS methodology to mouse brain (Data not shown).

6.4 Discovery of novel disease pathways and potential therapeutic strategies in TDP-43 ALS

Analysis of the TDP-43 Q331K translatome revealed novel pathways implicated in disease. The most down regulated transcript in the DEG list involved in neurodegeneration was CRABP1 (Figure 4.14). Since completion of this work, conversion of the HEK Flp Sham and TDP-43 Q331K cell models into iNPCs before retinoic differentiation has shown a Q331K specific delay in neuronal differentiation. Validation using qRT-PCR has shown approximately 100 times less CRABP1 transcripts in the HEK Flp TDP-43 Q331K cell line compared to the Sham. These data show that induction of TDP-43 Q331K has a negative effect on neuronal differentiation due to the down regulation of CRABP1. This is consistent with iPSC motor neurons derived from patients with M337V producing neurites which were half the length of healthy iPSC motor neurons (Egawa et al., 2012). Altered splicing and expression of profilin 1 (PFN1), granulin (GRN) and (ELP3), all of which are involved in neurite outgrowth, have been identified in motor neurons of patients with sporadic ALS with TDP-43 pathology (Highley et al., 2014).

Another enriched dysregulated pathway highlighted in the TDP-43 Q331K ALS translatome was mRNA transport/export (Figure 4.13). DEG involved with this pathway included several components of the TREX complex including DDX39B, THO subunits and NXF1, several nucleoporins and the cytoplasmic RNA helicase DDX19b/DBP5. There are also four nuclear export adaptors included in this list: SRSF1, SRSF3, SRSF7 and CHTOP.

Since, other works carried out in the RNA Biology group indicate that TDP-43 effectively binds the nuclear export receptor NXF1 and is part of the TREX complex, both *in vitro* using recombinant proteins and *in vivo* by co-immunoprecipitation in cell extracts. Since alterations in the TREX complex are likely to affect the co-transcriptional processing and the nuclear export of transcripts, Actinomycin D was used to artificially slow the nuclear biogenesis and processing of transcripts. When NSC-34 Flp TDP-43 Q331K were induced with Tetracycline for 7 days and treated with Actinomycin D, nuclear loss and cytoplasmic accumulation of TDP-43 was significantly exacerbated. This occurred in approximately 30% of TDP-43 Q331K cells and was significantly higher than the number of cells displaying proteinopathy in NSC-34 Flp Sham and TDP-43 WT cells (only 2-3%) (Figure 5.15).

Interestingly, this also indicates that the WT protein has a propensity to form TDP-43 proteinopathy under compromised co-transcriptional processing. Both cell models expressing TDP-43 and TDP-43 Q331K show impaired growth. TDP-43 protein levels are tightly regulated and over-expression of wild type protein can cause cytotoxic effects (Wils et al., 2010; Tsai et al., 2010; Xu et al., 2010; Shan et al., 2010) in full agreement with the data presented here. Interestingly, the Q331K mutation found in ALS is less toxic to the cells. This may be because in disease neurodegeneration takes several decades and it

might be that the mechanisms of toxicity are different affecting or promoting binding to other cellular proteins for example. Also the Flp In cell models express less TDP-43 Q331K protein compared to TDP-43 WT in both HEK and NSC-34 cells (Figure 4.1, Figure 5.7).

The HEK Flp TDP-43 cell models did not display TDP-43 proteinopathy, as is also reported by other groups (Ling et al., 2010). TDP-43 proteinopathy was only observed in a small proportion of NSC-34 Flp TDP-43 Q331K cells under normal culture conditions. Actinomycin D exacerbated this phenotype. When the HEK Flp cell models were treated with Actinomycin D, no TDP-43 proteinopathy was observed (Figure 5.20). This highlights the need for a neuronal context for proteinopathy to occur.

Ongoing, preliminary data from the RNA Biology lab has also shown differences in the levels of TDP-43 targets in whole cell, cytoplasmic and nuclear fractions from the HEK Flp In cell models. The nuclear reduction and accumulation of TDP-43 targets in the HEK Flp TDP-43 Q331K cell model is consistent with the disruption of nuclear export reported in the GRASPS translatome and cytoplasmic accumulation of poly(A)+ RNA with TDP-43 proteinopathy. The nuclear accumulation and cytoplasmic decrease of TDP-43 targets in the HEK Flp TDP-43 WT cell model is consistent with the effects of over-expression of a nuclear export adaptor. This would be consistent with the new role of TDP-43 as a nuclear export adaptor and would only occur in the TDP-43 WT model, as this model shows greater expression of TDP-43 than the TDP-43 Q331K cell model.

Transcription inhibition has been used to decrease nuclear import and visualise the presence of proteins in the cytoplasm. The distribution of other shuttling proteins, such as hnRNPs, has been shown to be affected by Actinomycin D,

but the mechanism by which transcription is required in the nuclear localisation of shuttling proteins is unknown (Piñol-Roma et al., 1992; 1993). Transcription is highly dependent on the tightly regulated TDP-43 protein levels and it has been shown that Actinomycin D can affect the distribution of TDP-43 WT by a yet unknown mechanism (Ayala et al., 2008). Interestingly, we observed that TDP-43 proteinopathy is associated with the nuclear loss and cytoplasmic accumulation of poly-A+ mRNA. We suggest that altered TREX functioning and impairments in the co-transcriptional processing and nuclear export of transcripts is associated with the TDP-43 proteinopathy. This is the first time that the loss of bulk mRNA has been linked to TDP-43 proteinopathy. TDP-43 has over 4000 identified binding targets (Sephton et al., 2011) and if TDP-43 was mislocalised from the nucleus it is perhaps not surprising that the transcripts it binds may also be mislocalised. Further investigations into the mechanisms that cause this to occur are required.

Nucleocytoplasmic transport has recently been implicated in C9ORF72-ALS using loss of function screen in *Drosophila* and yeast (Jovičić et al., 2015; Freibaum et al., 2015; Ke Zhang et al., 2015). Polypeptides produced from the C9ORF72 repeat expansion can aggregate and sequester proteins involved in nucleocytoplasmic transport (Yong-Jie Zhang et al., 2016). Cytoplasmic aggregation of disease proteins such as fragments of huntingtin and TDP-43 have been shown to interfere with the nucleocytoplasmic transport of RNA and proteins (Woerner et al., 2016). These studies using different methodologies have shown the involvement of this pathway in disease, supporting its identification from the GRASPS translatome.

Impaired recruitment of TDP-43 ALS mutant to mRNP complexes and altered TREX functioning is likely to slow down the rate of co-transcriptional processing,

which may result in the formation of R-loops and DNA damage. Several omics data, including ours, indicate that the DNA damage/repair pathways are altered in ALS and this might be linked to potential contribution from the TDP-43 proteinopathy. Recently another group has shown that TDP-43 (and FUS) are involved in the repair of R-loop formation due to DNA damage and that DNA damage due to loss of TDP-43 function leads to R-loop formation (Hill et al., 2016). The work described in this thesis supports this model. Furthermore, ongoing work in the RNA Biology group and by Dr Robin Highley has also shown that motor neurons exhibiting TDP-43 proteinopathy always show R-loops in post mortem CNS tissue from patients with ALS, while approximately one third of motor neurons show R-loops in the CNS of control individuals (unpublished data).

6.5 Transcriptome and translatome in health and disease: are all processed mRNAs translated into proteins?

Our study clearly highlights that the transcriptome and translatome are very poorly correlated (Figure 4.18). The Pheatmaps of DEG involved in neurodegeneration and nuclear export (Figure 4.13, Figure 4.14) again highlights this lack of correlation, with the majority of the DEGs in the WCT and CyT showing no difference compared to large up/down regulation in the GRASPS translatome. Also, there are striking differences between the CyT and WCT. There is no overall correlation between the three omics data sets. There is a small amount of correlation between the CyT and GRASPS translatome, which is perhaps not surprising given that translation occurs in the cytoplasm. These data really highlight the need to investigate the translatome rather than the transcriptome in order to investigate the functional consequences of RNA metabolism dysregulation/ modification, as not all mRNAs are translated into

proteins and each WCT or CyT transcriptomes and translatomes give rise to various DEG profiles. The CyT contains mRNAs that are held in stress granules for example, as well as non-coding RNAs. The WCT contains numerous non-coding RNA species held in various nuclear and cytoplasmic bodies or under co-transcriptional processing/quality control mechanisms. Investigating mRNA molecules which are preferentially translated into proteins is likely to allow reduction in the signal to noise ratio in GRASPS experiments and statistically improve the detection of DEGs – hence greater fold changes and the discovery of novel cellular processes.

Correctly spliced transcripts are marked with the EJC as part of an mRNA surveillance mechanism (Dreyfuss et al., 2002). mRNA decay pathways operate during translation to remove abnormally spliced mRNA with premature termination codons, no termination codon and those bound to ribosomes stuck in elongation (Kervestin et al., 2012). Splicing analysis showed that were approximately 28 times more abnormal splicing events in the transcriptome (WCT or CyT) than observed in the GRASPS translatome, approximately 900 in the transcriptome compared to 33 in the translatome (Table 4.5). This suggests, for the first time, that very few abnormally spliced mRNA isoforms are associated with selectively enriched translating ribosomes and are translated into protein. This is consistent with previous transcriptomic studies reporting large numbers of abnormal splicing events (Table 1.4). However, again, these results highlight the weaknesses of transcriptomic analysis, as not all abnormally spliced transcripts are translated into protein. These results suggest that abnormally spliced mRNA can be exported into the cytoplasm, but a yet unknown control mechanism seems to prevent the translation of the majority of abnormally spliced mRNA. This is very interesting but needs much more investigation.

There were 33 abnormally spliced events in 14 different transcripts, which did reach the ribosome for translation (Table 4.6). Of these, seven had gene products, which had already been implicated in neurodegeneration (*TBK1*, *PADI2*, *IFRD1*). Notably, *TBK1* mutations have already been implicated in sALS, fALS and FTD (Freischmidt et al., 2015; 2016). TBK1 is a kinase involved in the autophagosome-mediated degradation of ubiquitinated cargoes (Weidberg et al., 2011). ALS-associated proteins optineurin and p62 are both substrates of TBK1 (Domínguez-Escobar et al., 2011; Pilli et al., 2012; Komatsu et al., 2012). Mutations in TBK1 have been shown to cause ALS and it is of great interest to find that in TDP-43 Q331K ALS, a mutant TBK1 is potentially produced.

The majority of gene products reported as not involved in neurodegeneration are still applicable for a neurodegenerative disease or associated with pathways that have already been implicated in ALS. For example, *GMFB* encodes the glia maturation factor beta protein responsible for glia maturation (Kato et al., 1987; Lim et al., 1989).

Kaiso (encoded by *ZBTB33*) is a transcriptional repressor which interacts with importins (Kelly et al., 2004; Klose et al., 2006). Importins are responsible for the transport of proteins into the nucleus and in ALS TDP-43 is depleted from the nucleus. Altered import of TDP-43 into the nucleus is a potential reason for the nuclear depletion of TDP-43 in TDP-43 proteinopathy.

Other abnormally spliced products in the translatome are involved in the regulation of transcription: *IFRD1*, *ZNF558*, *NUPR1*, *ZNF614*. Altered regulation of transcription could also affect the rate of transcription/processing resulting in altered mRNA transport/export in TDP-43 ALS.

6.6 Transcriptome and translatome in TDP-43 ALS

The transcriptomic analysis in this study, identified processes which have also been reported in other TDP-43 ALS using transcriptomic studies such as protein transport/localisation, RNA splicing/processing, protein degradation/ubiquitin and cell death apoptosis (Table 1.4). This shows that the processes identified from the transcriptomes using the HEK Flp cell models in this study are reliable and consistent with already published transcriptomic studies.

The ALS translatome also identifies biological processes that have already been implicated in ALS, but not using genome wide -omics studies. For example, aerobic respiration and metabolism have already been implicated in ALS from measuring mitochondrial respiration and glycolytic flux in living cells (Allen et al., 2014).

Some pathways have already been implicated in disease, but in addition to these pathways the translatome study gives a list of the genes that are altered and this list contains genes that have not yet been implicated in disease before.

Comparisons of the transcriptomes (WCT and CyT) and translatome have shown that GRASPS is more sensitive and more robust. However, it is also interesting to compare the GRASPS analysis in this study with a recent paper which performed TRAP on TDP-43 Q331K mice (MacNair et al., 2015). This paper reports the identification of 28 DEGs with FC up to 16, compared to the 4264 DEGs with FC up to approximately 250 fold in the GRASPS translatome. The huge increase in the amount of data obtained, further highlights the increased sensitivity of the GRASPS methodology.

Interestingly, tRNA aminoacylation alterations have not previously been implicated in ALS. tRNA aminoacylation is carried out by tRNA aminoacylation synthetases and matches the correct amino acid to the anticodon of the tRNA during protein synthesis. Protein synthesis is vital to maintain cell homeostasis and function. tRNA aminoacylation has been implicated in disease with mutations in mitochondrial tRNA causing diseases related to cellular energy metabolism (Abbott et al., 2014). All known disease mutations in tRNA synthetases are associated with Charcot-Marie-Tooth and related neuropathies, which are characterised by the progressive degeneration of distal motor and sensory neuron function (Yao et al., 2013).

6.7 Future work

Future experiments from this work, will be to perform pulse SILAC (pSILCAC) proteomics analysis to measure the protein synthesis rate of candidate proteins to validate the finding from GRASPS and determine whether GRASPS translatomes would better correlate with pSILAC proteomes. It has already been reported that WCT transcriptomes poorly correlate with proteomes in agreement to the GRASP translatome findings presented in this thesis.

Further validation of these findings is to be performed using polysome profiling (Lui et al., 2014; Costello et al., 2015). Although not a high throughput technique, qRT-PCR on polysome fractions could validate the translatome analysis, such as the down regulation of CRABP1 for example.

The limitations of the GRASPS methodology are the lack of cell type specificity. GRASPS experiments have only so far been performed on cultured cells. In future work the combination of cell-type specific purificiation prior to GRASPS purifications could be explored. The limit of detection for the GRASPS method is

not currently known. Future experiments exploring the type and amount of material upon which to perform GRASPS would enable more specific targeting of GRASPS and identify its limit of detection. These are important considerations in the long-term application of GRASPS.

Although successful purification of ribosomes has been shown by the GRASPS method, approximately 25% of GFP M1V mRNA is still co-purified with the ribosomes in induced HEK Flp M1V, even though no GFP protein is produced (Figure 3.19). In the future, some improvements could be brought in by trying to add inhibitory peptides that prevent the recruitment of the 60S large subunit to the 43S pre-initiation complex.

No significant differences in cell proliferation or growth of NSC-34 Flp TDP-43 Q331K cell lines compared to NSC-34 Flp Sham or TDP-43 WT under normal culture conditions were observed. Now a method to cause the nuclear loss and cytoplasmic accumulation of TDP-43 has been established, it would be interesting to repeat the MTT assays at later time points following Actinomycin D treatment. The loss of TDP-43 from the nucleus may have an effect upon the cell growth and proliferation, which is not visible immediately after treatment. After Actinomycin D treatment cells were fixed and stained immediately to identify mRNA and TDP-43 localisation. Changing the media and continuing to grow cells following the nuclear loss would determine whether or not the nuclear loss of TDP-43 was reversible if TDP-43 was relocalised and whether the nuclear loss had a toxic affect upon the cell's growth and survival.

Nuclear loss of TDP-43 is present in approximately 30 % of NSC-34 Flp TDP-43 Q331K following treatment with Actinomycin D. This quantifiable measure of

nuclear loss could make the cell models suitable for use in drug screening of therapeutic compounds, which may relocalise TDP-43.

NSC-34 Flp In TDP-43 A315T and M337V models were also constructed. Identifying whether or not Actinomycin D can induce the nuclear loss and cytoplasmic accumulation of TDP-43 in these models and if so, the quantification of the nuclear loss in these models would be interesting to see whether a particular mutation has a more severe phenotype and whether they all exhibit proteinopathy. For example, TDP-43 M337V ALS displays a more aggressive disease progression in patients and it would be interesting to see whether the TDP-43 proteinopathy was also more pronounced in these cells.

Greater understanding of the nuclear loss could be used to identify therapeutic strategies. What is preventing the relocalisation of TDP-43 into the nucleus? Is there an import/export problem preventing the shuttling of TDP-43 protein between the nucleus and cytoplasm. Several importins were included in the DEG involved in dysregulated mRNA export/transport. If importins were knocked down it would be interesting to see whether TDP-43 could get back into the nucleus in NSC-34 Flp TDP-43 Q331K cells and whether it would have the same effect upon the other cell lines.

The splicing analysis suggests that abnormally spliced mRNA is exported into the cytoplasm but is prevented from reaching the ribosome for translation. These abnormally spliced variants identified will need to be validated by RT-PCR analysis. Future bioinformatics experiments would also aim at determining whether specific sequence elements exist between the two different groups of abnormally spliced transcripts. Also, synthetic abnormally spliced mRNA could be produced and used *in vitro* pull down assays with cellular assays for

identification of potentially different mRNA binding proteins by mass spectrometry.

6.8 Conclusions

The use of GRASPS to study the translatome is a powerful method to study at genome wide level mRNAs that are undergoing active protein synthesis. It can be applied to any cell type for the investigation of any disease.

This methodology has been used on HEK Flp TDP-43 cell lines to identify new disease relevant pathways and the genes involved in these pathways. In combination with the characterisation of motor neuron-like TDP-43 ALS cell models, which replicate TDP-43 proteinopathy, a new role for TDP-43 in mRNA nuclear export has been identified, together with an association between the concomitant nuclear loss and cytoplasmic mislocalisation of TDP-43 and bulk mRNA in proteinopathy.

The comparison of the GRASPS translatome with whole cell transcriptome and cytoplasmic transcriptomes, has demonstrated the importance of using a different approach and translatome technologies in the analysis of gene expression. It has also identified potential new mechanisms in gene expression with the discovery that only a small fraction of abnormally spliced transcripts that are exported into the cytoplasm, reach the ribosome to be translated.

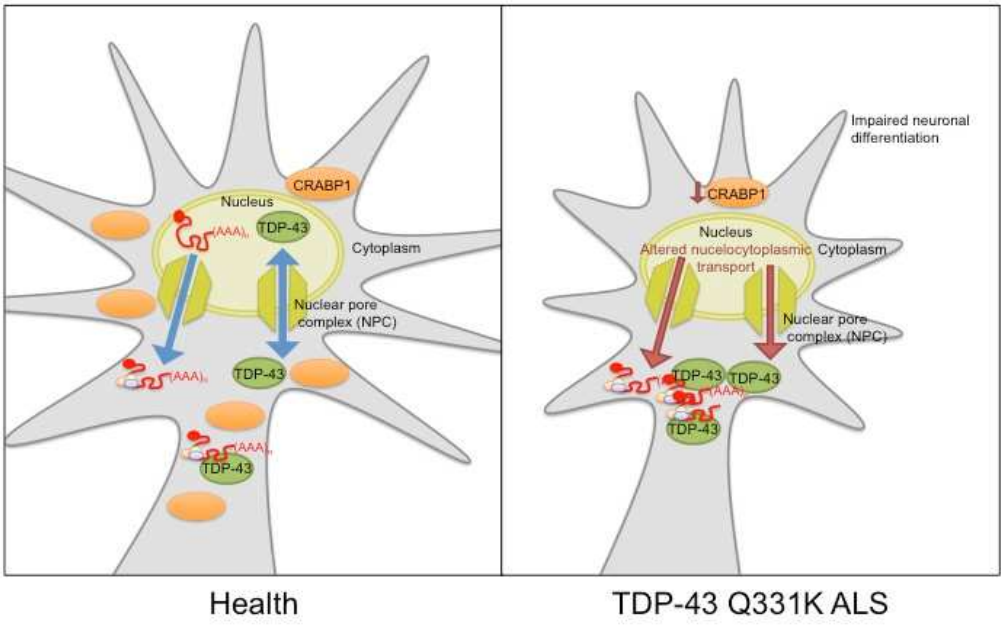


Figure 6.1 The functional consequences of widespread RNA dysregulation in TDP-43 Q331K ALS cell models

In healthy cells TDP-43 is located predominantly in the nucleus, but continuously shuttles between the nucleus and cytoplasm. Alterations to nucleocytoplasmic transport in TDP-43 Q331K ALS cause the nuclear loss and cytoplasmic mislocation of TDP-43 and bulk poly(A)$^+$ mRNA. Down regulation of CRABP1 contributes to impaired neuronal function in TDP-43 Q331K ALS.

Outputs from the body of work in my PhD

Dodd JE et al. Molecular mechanisms of nuclear loss in TDP-43 related proteinopathy and neurodegeneration. (*in preparation*)

Dodd JE et al. GRASPSing translatomes reveal hidden disease-associated pathways in neurodegeneration and a novel function for TDP-43 (*in preparation*)

Mihaylov SR, **Dodd JE** et al. Development of a glial and neuronal reprogramming protocol for stable inducible cell models. (*in preparation*)

Gene Editing Symposium, Research Staff Association for Medicine, Dentistry and Health April 2016 – invited oral presentation
RNA UK Jan 2016 – poster presentation
Sheffield Molecular Biology Symposium Jan 2015 – poster presentation
25th International Symposium on ALS/MND, Dec 2014 – poster presentation, Research Care and Community blog post 'Protein Misfolding and Toxicity'
Neuroscience Research Day Oct 2014 – oral presentation
Neuroscience Research Day Oct 2013 – prize for poster presentation

References

Abbott, Jamie A, Francklyn, Christopher S and Robey-Bond, Susan M (2014), Transfer RNA and Human Disease. *Frontiers in genetics*, 5(122).

Abhyankar, Mayuresh M, Urekar, Craig and Reddi, Prabhakara P (2007), A Novel CpG-Free Vertebrate Insulator Silences the Testis-Specific SP-10 Gene in Somatic Tissues: Role for TDP-43 in Insulator Function. *Journal of Biological Chemistry*, 282(50): 36143–36154.

Acharya, Kshitish K, Govind, Chhabi K, Shore, Amy N, Stoler, Mark H and Reddi, Prabhakara P (2006), Cis-Requirement for the Maintenance of Round Spermatid-Specific Transcription. *Developmental Biology*, 295(2): 781–790.

Al-Chalabi, Ammar, Jones, Ashley, Troakes, Claire, King, Andrew, Al-Sarraj, Safa and van den Berg, Leonard H (2012), The Genetics and Neuropathology of Amyotrophic Lateral Sclerosis. *Acta neuropathologica*, 124(3): 339–352.

Al-Saif, Amr, Al-Mohanna, Futwan and Bohlega, Saeed (2011), A Mutation in Sigma-1 Receptor Causes Juvenile Amyotrophic Lateral Sclerosis. *Annals of neurology*, 70(6): 913–919.

Al-Sarraj, Safa, King, Andrew, Troakes, Claire, Smith, Bradley, Maekawa, Satomi, Bodi, Istvan, Rogelj, Boris, Al-Chalabi, Ammar, Hortobagyi, Tibor and Shaw, Christopher E (2011), P62 Positive, TDP-43 Negative, Neuronal Cytoplasmic and Intranuclear Inclusions in the Cerebellum and Hippocampus Define the Pathology of C9orf72-Linked FTLD and MND/ALS. *Acta neuropathologica*, 122(6): 691–702.

Alami, Nael H, Smith, Rebecca B, Carrasco, Monica A, Williams, Luis A, Winborn, Christina S, Han, Steve S W, Kiskinis, Evangelos, Winborn, Brett, Freibaum, Brian D, Kanagaraj, Anderson, Clare, Alison J, Badders, Nisha M, Bilican, Bilada, Chaum, Edward, Chandran, Siddharthan, Shaw, Christopher E, Eggan, Kevin C, Maniatis, Tom and Taylor, J Paul (2014), Axonal Transport of TDP-43 mRNA Granules Is Impaired by ALS-Causing Mutations. *Neuron*, 81(3): 536–543.

Alexander, Adanna G, Marfil, Vanessa and Li, Chris (2014), Use of Caenorhabditis Elegans as a Model to Study Alzheimer's Disease and Other Neurodegenerative Diseases. *Frontiers in genetics*, 5(JUL): 279.

Alexander, Michael D, Traynor, Bryan J, Miller, Nicole, Corr, Bernie, Frost, Eithne, McQuaid, Shirley, Brett, Francesca M, Green, Andrew and Hardiman, Orla (2002), 'True' Sporadic ALS Associated with a Novel SOD-1 Mutation. *Annals of neurology*, 52(5): 680–683.

Alfieri, Julio A, Pino, Natalia S and Igaz, Lionel M (2014), Reversible Behavioral Phenotypes in a Conditional Mouse Model of TDP-43 Proteinopathies. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 34(46): 15244–15259.

Allen, Scott P, Rajan, Sandeep, Duffy, Lynn, Mortiboys, Heather, Higginbottom, Adrian, Grierson, Andrew J and Shaw, Pamela J (2014), Superoxide Dismutase 1 Mutation in a Cellular Model of Amyotrophic Lateral Sclerosis Shifts Energy Generation From Oxidative Phosphorylation to Glycolysis. *Neurobiology of Aging*, 35(6): 1499–1509.

Anderson, Paul and Kedersha, Nancy (2008), Stress Granules: the Tao of RNA Triage. *Trends in biochemical sciences*, 33(3): 141–150.

Anger, Andreas M, Armache, Jean-Paul, Berninghausen, Otto, Habeck, Michael, Subklewe, Marion, Wilson, Daniel N and Beckmann, Roland (2013), Structures of the Human and Drosophila 80S Ribosome. *Nature*, 497(7447): 80–85.

Antonellis, Anthony, Ellsworth, Rachel E, Sambuughin, Nyamkhishig, Puls, Imke, Abel, Annette, Lee-Lin, Shih-Queen, Jordanova, Albena, Kremensky, Ivo, Christodoulou, Kyproula, Middleton, Lefkos T, Sivakumar, Kumaraswamy, Ionasescu, Victor, Funalot, Benoit, Vance, Jeffery M, Goldfarb, Lev G, Fischbeck, Kenneth H and Green, Eric D (2003), Glycyl tRNA Synthetase Mutations in Charcot-Marie-Tooth Disease Type 2D and Distal Spinal Muscular Atrophy Type v. *American journal of human genetics*, 72(5): 1293–1299.

Arai, Tetsuaki, Hasegawa, Masato, Akiyama, Haruhiko, Ikeda, Kenji, Nonaka, Takashi, Mori, Hiroshi, Mann, David, Tsuchiya, Kuniaki, Yoshida, Mari, Hashizume, Yoshio and Oda, Tatsuro (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): 602–611.

Arava, Yoav, Wang, Yulei, Storey, John D, Liu, Chih Long, Brown, Patrick O and Herschlag, Daniel (2003), Genome-Wide Analysis of mRNA Translation Profiles in Saccharomyces Cerevisiae. *PNAS*, 100(7): 3889–3894.

Argüelles, Sandro, Camandola, Simonetta, Hutchison, Emmette R, Cutler, Roy G, Ayala, Antonio and Mattson, Mark P (2013), Molecular Control of the Amount, Subcellular Location, and Activity State of Translation Elongation Factor 2 in Neurons Experiencing Stress. *Free radical biology & medicine*, 61C: 61–71.

Armakola, Maria, Higgins, Matthew J, Figley, Matthew D, Barmada, Sami J, Scarborough, Emily A, Diaz, Zamia, Fang, Xiaodong, Shorter, James, Krogan, Nevan J, Finkbeiner, Steven, Farese, Robert V and Gitler, Aaron D (2012), Inhibition of RNA Lariat Debranching Enzyme Suppresses TDP-43 Toxicity in ALS Disease Models. *Nature genetics*, 44(12): 1302–1309.

Armstrong, Gary A B and Drapeau, Pierre (2013a), Calcium Channel Agonists Protect Against Neuromuscular Dysfunction in a Genetic Model of TDP-43 Mutation in ALS. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(4): 1741–1752.

Armstrong, Gary A B and Drapeau, Pierre (2013b), Loss and Gain of FUS Function Impair Neuromuscular Synaptic Transmission in a Genetic Model of ALS. *Human molecular genetics*, 22(21): 4282–4292.

Arnold, Eveline S, Ling, Shuo-Chien, Huelga, Stephanie C, Lagier-Tourenne, Clotilde, Polymenidou, Magdalini, Ditsworth, Dara, Kordasiewicz, Holly B, McAlonis-Downes, Melissa, Platoshyn, Oleksandr, Parone, Philippe a, Da Cruz, Sandrine, Clutario, Kevin M, Swing, Debbie, Tessarollo, Lino, Marsala, Martin, Shaw, Christopher E, Yeo, Gene W and Cleveland, Don W (2013), PNAS Plus: ALS-Linked TDP-43 Mutations Produce Aberrant RNA Splicing and Adult-Onset Motor Neuron Disease Without Aggregation or Loss of Nuclear TDP-43. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8): E736–E745.

Ash, Peter E a, Bieniek, Kevin F, Gendron, Tania F, Caulfield, Thomas, Lin, Wen-Lang, DeJesus-Hernandez, Mariely, van Blitterswijk, Marka M, Jansen-West, Karen, Paul, Joseph W, Rademakers, Rosa, Boylan, Kevin B, Dickson, Dennis W and Petrucelli, Leonard (2013), Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron*, 77(4): 639–646.

Ash, Peter E a, Zhang, Yong-Jie, Roberts, Christine M, Saldi, Tassa, Hutter, Harald, Buratti, Emanuele, Petrucelli, Leonard and Link, Christopher D (2010), Neurotoxic Effects of TDP-43 Overexpression in C. Elegans. *Human molecular genetics*, 19(16): 3206–3218.

Ashley, C T and Warren, S T (1995), Trinucleotide Repeat Expansion and Human Disease. *Annual Review of Genetics*, 29: 703–728.

Avendaño-Vázquez, S Eréndira, Dhir, Ashish, Bembich, Sara, Buratti, Emanuele, Proudfoot, Nicholas and Baralle, Francisco E (2012), Autoregulation of TDP-43 mRNA Levels Involves Interplay Between Transcription, Splicing, and Alternative polyA Site Selection. *Genes \& development*, 26(15): 1679–1684.

Ayala, Youhna M, De Conti, Laura, Avendaño-Vázquez, S Eréndira, Dhir, Ashish, Romano, Maurizio, D'Ambrogio, Andrea, Tollervey, James, Ule, Jernej, Baralle, Marco, Buratti, Emanuele and Baralle, Francisco E (2011), TDP-43 Regulates Its mRNA Levels Through a Negative Feedback Loop. *The EMBO journal*, 30(2): 277–288.

Ayala, Youhna M, Pantano, Sergio, D'Ambrogio, Andrea, Buratti, Emanuele, Brindisi, Antonia, Marchetti, Caterina, Romano, Maurizio and Baralle, Francisco E (2005), Human, Drosophila, and C.Elegans TDP43: Nucleic Acid Binding Properties and Splicing Regulatory Function. *Journal of molecular biology*, 348(3): 575–588.

Ayala, Youhna M, Zago, Paola, D'Ambrogio, Andrea, Xu, Ya-Fei, Petrucelli, Leonard, Buratti, Emanuele and Baralle, Francisco E (2008), Structural Determinants of the Cellular Localization and Shuttling of TDP-43. *Journal of cell science*, 121(Pt 22): 3778–3785.

Baechtold, H, Kuroda, M, Sok, J, Ron, D, Lopez, B S and Akhmedov, a T (1999), Human 75-kDa DNA-Pairing Protein Is Identical to the Pro-Oncoprotein TLS/FUS and Is Able to Promote D-Loop Formation. *The Journal of biological chemistry*, 274(48): 34337–34342.

Bandmann, Oliver and Burton, Edward A (2010), Genetic Zebrafish Models of Neurodegenerative Diseases. *Neurobiology of disease*, 40(1): 58–65.

Bannwarth, Sylvie, Ait-El-Mkadem, Samira, Chaussenot, Annabelle, Genin, Emmanuelle C, Lacas-Gervais, Sandra, Fragaki, Konstantina, Berg-Alonso, Laetitia, Kageyama, Yusuke, Serre, Valérie, Moore, David G, Verschueren, Annie, Rouzier, Cécile, Le Ber, Isabelle, Augé, Gaëlle, Cochaud, Charlotte, Lespinasse, Françoise, N'Guyen, Karine, de Septenville, Anne, Brice, Alexis, Yu-Wai-Man, Patrick, Sesaki, Hiromi, Pouget, Jean and Paquis-Flucklinger, Véronique (2014), A Mitochondrial Origin for Frontotemporal Dementia and Amyotrophic Lateral Sclerosis Through CHCHD10 Involvement. *Brain : a journal of neurology*, 137(Pt 8): 2329–2345.

Barber, Sian C and Shaw, Pamela J (2010), Oxidative Stress in ALS: Key Role in Motor Neuron Injury and Therapeutic Target. *Free radical biology & medicine*, 48(5): 629–641.

Barmada, Sami J, Skibinski, Gaia, Korb, Erica, Rao, Elizabeth J, Wu, Jane Y and Finkbeiner, Steven (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 : the official journal of the Society for Neuroscience*, 30(2): 639–649.

Belin, Stéphane, Hacot, Sabine, Daudignon, Lionel, Therizols, Gabriel, Pourpe, Stéphane, Mertani, Hichem C, Rosa-Calatrava, Manuel and Diaz, Jean-Jacques (2010), Purification of Ribosomes From Human Cell Lines. *Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.]*, Chapter 3: Unit 3.40.

Belzil, Veronique V, Gendron, Tania F and Petrucelli, Leonard (2013), RNA-Mediated Toxicity in Neurodegenerative Disease. *Molecular and Cellular Neuroscience*, 56: 406–419.

Benajiba, Lina, Le Ber, Isabelle, Camuzat, Agnès, Lacoste, Mathieu, Thomas-Anterion, Catherine, Couratier, Philippe, Legallic, Solenn, Salachas, Francois, Hannequin, Didier, Decousus, Marielle, Lacomblez, Lucette, Guedj, Eric, Golfier, Véronique, Camu, William, Dubois, Bruno, Campion, Dominique, Meininger, Vincent, Brice, AlexisFrench Clinical and Genetic Research Network on Frontotemporal Lobar Degeneration/Frontotemporal Lobar Degeneration with Motoneuron Disease (2009), TARDBP Mutations in Motoneuron Disease with Frontotemporal Lobar Degeneration. *Annals of neurology*, 65(4): 470–473.

Bertin, Benjamin, Renaud, Yoan, Aradhya, Rajaguru, Jagla, Krzysztof and Junion, Guillaume (2015), TRAP-Rc, Translating Ribosome Affinity Purification From Rare Cell Populations of Drosophila Embryos. *Journal of visualized experiments : JoVE*, (103).

Bilican, Bilada, Serio, Andrea, Barmada, Sami J, Nishimura, Agnes Lumi, Sullivan, Gareth J, Carrasco, Monica, Phatnani, Hemali P, Puddifoot, Clare a, Story, David, Fletcher, Judy, Park, In-Hyun, Friedman, Brad a, Daley, George Q, Wyllie, David J a, Hardingham, Giles E, Wilmut, Ian, Finkbeiner, Steven, Maniatis, Tom, Shaw, Christopher E and Chandran, Siddharthan (2012), Mutant Induced Pluripotent Stem Cell Lines Recapitulate Aspects of TDP-43 Proteinopathies and Reveal Cell-Specific Vulnerability. *Proceedings of the National Academy of Sciences of the United States of America*, 109(15): 5803–5808.

Black, Douglas L (2003), Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annual Review of Biochemistry*, 72: 291–336.

Borchelt, D R, Davis, J, Fischer, M, Lee, M K, Slunt, H H, Ratovitsky, T, Regard, J, Copeland, N G, Jenkins, N a, Sisodia, S S and Price, D L (1996), A Vector for Expressing Foreign Genes in the Brains and Hearts of Transgenic Mice. *Genetic Analysis-Biomolecular Engineering*, 13(6): 159–163.

Borroni, B, Bonvicini, C, Alberici, a, Buratti, E, Agosti, C, Archetti, S, Papetti, a, Stuani, C, Di Luca, M, Gennarelli, M and Padovani, a (2009), Mutation Within TARDBP Leads to Frontotemporal Dementia Without Motor Neuron Disease. *Human mutation*, 30(11): E974–83.

Boyd, Justin D, Lee-Armandt, J Peter, Feiler, Marisa S, Zaarur, Nava, Liu, Min, Kraemer, Brian, Concannon, John B, Ebata, Atsushi, Wolozin, Benjamin and Glicksman, Marcie A (2014), A High-Content Screen Identifies Novel Compounds That Inhibit Stress-Induced TDP-43 Cellular Aggregation and Associated Cytotoxicity. *Journal of biomolecular screening*, 19(1): 44–56.

Brady, Owen A, Meng, Peter, Zheng, Yanqiu, Mao, Yuxin and Hu, Fenghua (2011), Regulation of TDP-43 Aggregation by Phosphorylation Andp62/SQSTM1. *Journal of neurochemistry*, 116(2): 248–259.

Brettschneider, Johannes, Del Tredici, Kelly, Toledo, Jon B, Robinson, John L, Irwin, David J, Grossman, Murray, Suh, EunRan, Van Deerlin, Vivianna M, Wood, Elisabeth M, Baek, Young, Kwong, Linda, Lee, Edward B, Elman, Lauren, McCluskey, Leo, Fang, Lubin, Feldengut, Simone, Ludolph, Albert C, Lee, Virginia M-Y, Braak, Heiko and Trojanowski, John Q (2013), Stages of pTDP-43 Pathology in Amyotrophic Lateral Sclerosis. *Annals of neurology*, 74(1): 20–38.

Brkanac, Zoran, Spencer, David, Shendure, Jay, Robertson, Peggy D, Matsushita, Mark, Vu, Tiffany, Bird, Thomas D, Olson, Maynard V and Raskind, Wendy H (2009), IFRD1 Is a Candidate Gene for SMNA on Chromosome 7q22-Q23. *American journal of human genetics*, 84(5): 692–697.

Brockington, Alice, Ning, Ke, Heath, Paul R, Wood, Elizabeth, Kirby, Janine, Fusi, Nicolò, Lawrence, Neil, Wharton, Stephen B, Ince, Paul G and Shaw, Pamela J (2013), Unravelling the Enigma of Selective Vulnerability in Neurodegeneration: Motor Neurons Resistant to Degeneration in ALS Show Distinct Gene Expression Characteristics and Decreased Susceptibility to Excitotoxicity. *Acta neuropathologica*, 125(1): 95–109.

Brook, J D, McCurrach, M E, Harley, H G, Buckler, a J, Church, D, Aburatani, H, Hunter, K, Stanton, V P, Thirion, J P and Hudson, T (1992), Molecular Basis of Myotonic Dystrophy: Expansion of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member. *Cell*, 69(2): 385.

Budini, Mauricio, Buratti, Emanuele, Stuani, Cristiana, Guarnaccia, Corrado, Romano, Valentina, De Conti, Laura and Baralle, Francisco E (2012), Cellular Model of TAR DNA-Binding Protein 43 (TDP-43) Aggregation Based on Its C-Terminal Gln/Asn-Rich Region. *The Journal of biological chemistry*, 287(10): 7512–7525.

Budini, Mauricio, Romano, Valentina, Quadri, Zainuddin, Buratti, Emanuele and Baralle, Francisco E (2014), TDP-43 Loss of Cellular Function Through Aggregation Requires Additional Structural Determinants Beyond Its C Terminal Q/N Prion-Like Domain. *Human molecular genetics*.

Bunina, T L (1962), On Intracellular Inclusions in Familial Amyotrophic Lateral Sclerosis. *Korsakov J Neuropathol Psychiatry*, 62: 1293–1299.

Buratti, E, Brindisi, A, Giombi, M, Tisminetzky, S, Ayala, Y M and Baralle, F E (2005), TDP-43 Binds Heterogeneous Nuclear Ribonucleoprotein a/B Through Its C-Terminal Tail - an Important Region for the Inhibition of Cystic Fibrosis Transmembrane Conductance Regulator Exon 9 Splicing. *The Journal of biological chemistry*, 280(45): 37572–37584.

Buratti, E, Dork, T, Zuccato, E, Pagani, F, Romano, M and Baralle, F E (2001), Nuclear Factor TDP-43 and SR Proteins Promote in Vitro and in Vivo CFTR Exon 9 Skipping. *EMBO Journal*, 20(7): 1774–1784.

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

Buratti, Emanuele and Baralle, Francisco E (2011), TDP-43: New Aspects of Autoregulation Mechanisms in RNA Binding Proteins and Their Connection with Human Disease. *The FEBS journal*, 278(19): 3530–3538.

Buratti, Emanuele and Baralle, Francisco Ernesto (2010), The Multiple Roles of TDP-43 in Pre-mRNA Processing and Gene Expression Regulation. *RNA biology*, 7(4): 420–429.

Burkhardt, Matthew F, Martinez, Fernando J, Wright, Sarah, Ramos, Carla, Volfson, Dmitri, Mason, Michael, Garnes, Jeff, Dang, Vu, Lievers, Jeffery, Shoukat-Mumtaz, Uzma, Martinez, Rita, Gai, Hui, Blake, Robert, Vaisberg, Eugeni, Grskovic, Marica, Johnson, Charles, Irion, Stefan, Bright, Jessica, Cooper, Bonnie, Nguyen, Leane, Griswold-Prenner, Irene and Javaherian, Ashkan (2013), A Cellular Model for Sporadic ALS Using Patient-Derived Induced Pluripotent Stem Cells. *Molecular and cellular neurosciences*, 56: 355–364.

Burrell, James R, Halliday, Glenda M, Kril, Jillian J, Ittner, Lars M, Götz, Jürgen, Kiernan, Matthew C and Hodges, John R (2016), The Frontotemporal Dementia-Motor Neuron Disease Continuum. *Lancet*, 388(10047): 919–931.

Byrne, Susan, Walsh, Cathal, Lynch, Catherine, Bede, Peter, Elamin, Marwa, Kenna, Kevin, McLaughlin, Russell and Hardiman, Orla (2011), Rate of Familial Amyotrophic Lateral Sclerosis: a Systematic Review and Meta-Analysis. *Journal of neurology, neurosurgery, and psychiatry*, 82(6): 623–627.

Caccamo, Antonella, Shaw, Darren M, Guarino, Francesca, Messina, Angela, Walker, Aaron W and Oddo, Salvatore (2015), Reduced Protein Turnover Mediates Functional Deficits in Transgenic Mice Expressing the 25 kDa C-Terminal Fragment of TDP-43. *Human molecular genetics*, 24(16): 4625–4635.

Cannon, Ashley, Yang, Baoli, Knight, Joshua, Farnham, Ian M, Zhang, Yongjie, Wuertzer, Charles A, D'Alton, Simon, Lin, Wen-Lang, Castanedes-Casey, Monica, Rousseau, Linda, Scott, Brittany, Jurasic, Michael, Howard, John, Yu, Xin, Bailey, Rachel, Sarkisian, Matthew R, Dickson, Dennis W, Petrucelli, Leonard and Lewis, Jada (2012), Neuronal Sensitivity to TDP-43 Overexpression Is Dependent on Timing of Induction. *Acta neuropathologica*, 123(6): 807–823.

Carmody, Sean R and Wente, Susan R (2009), mRNA Nuclear Export at a Glance. *Journal of cell science*, 122(Pt 12): 1933–1937.

Casari, G, De Fusco, M, Ciarmatori, S, Zeviani, M, Mora, M, Fernandez, P, De Michele, G, Filla, a, Cocozza, S, Marconi, R, Dürr, A, Fontaine, B and Ballabio, a (1998), Spastic Paraplegia and OXPHOS Impairment Caused by Mutations in Paraplegin, a Nuclear-Encoded Mitochondrial Metalloprotease. *Cell*, 93(6): 973–983.

Casci, Ian and Pandey, Udai Bhan (2015), A Fruitful Endeavor: Modeling ALS in the Fruit Fly. *Brain research*, 1607: 47–74.

Cashman, N R, Durham, H D, Blusztajan, J K, Oda, K, Tabira, T, Shaw, I T, Dahrouge, S and Antel, J P (1992), Neuroblastoma X Spinal-Cord (NSC) Hybrid Cell-Lines Resemble Developing Motor Neurons. *Developmental Dynamics*, 194(3): 209–221.

Chang, Yueming, Kong, Qiongman, Shan, Xiu, Tian, Guilian, Ilieva, Hristelina, Cleveland, Don W, Rothstein, Jeffrey D, Borchelt, David R, Wong, Philip C and Lin, Chien-Liang Glenn (2008), Messenger RNA Oxidation Occurs Early in Disease Pathogenesis and Promotes Motor Neuron Degeneration in ALS. *PloS one*, 3(8): e2849.

Chassé, Héloïse, Boulben, Sandrine, Costache, Vlad, Cormier, Patrick and Morales, Julia (2017), Analysis of Translation Using Polysome Profiling. *Nucleic acids research*, 45(3): e15.

Che, Mei-Xia, Jiang, Lei-Lei, Li, Hai-Yin, Jiang, Ya-Jun and Hu, Hong-Yu (2015), TDP-35 Sequesters TDP-43 Into Cytoplasmic Inclusions Through Binding with RNA. *FEBS letters*.

Che, Mei-Xia, Jiang, Ya-Jun, Xie, Yuan-Yuan, Jiang, Lei-Lei and Hu, Hong-Yu (2011), Aggregation of the 35-kDa Fragment of TDP-43 Causes Formation of Cytoplasmic Inclusions and Alteration of RNA Processing. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 25(7): 2344–2353.

Chen, Ying-Zhang, Bennett, Craig L, Huynh, Huy M, Blair, Ian P, Puls, Imke, Irobi, Joy, Dierick, Ines, Abel, Annette, Kennerson, Marina L, Rabin, Bruce a, Nicholson, Garth a, Auer-Grumbach, Michaela, Wagner, Klaus, De Jonghe, Peter, Griffin, John W, Fischbeck, Kenneth H, Timmerman, Vincent, Cornblath, David R and Chance, Phillip F (2004), DNA/RNA Helicase Gene Mutations in a Form of Juvenile Amyotrophic Lateral Sclerosis (ALS4). *American journal of human genetics*, 74(6): 1128–1135.

Chen, Ying-Zhang, Hashemi, Sayed H, Anderson, Susan K, Huang, Yongzhao, Moreira, Maria-Ceu, Lynch, David R, Glass, Ian a, Chance, Phillip F and

Bennett, Craig L (2006), Senataxin, the Yeast Sen1p Orthologue: Characterization of a Unique Protein in Which Recessive Mutations Cause Ataxia and Dominant Mutations Cause Motor Neuron Disease. *Neurobiology of disease*, 23(1): 97–108.

Chiang, Po-Min, Ling, Jonathan, Jeong, Yun Ha, Price, Donald L, Aja, Susan M and Wong, Philip C (2010), Deletion of TDP-43 Down-Regulates Tbc1d1, a Gene Linked to Obesity, and Alters Body Fat Metabolism. *PNAS*, 107(37): 16320–16324.

Chiò, Adriano, Calvo, Andrea, Moglia, Cristina, Ossola, Irene, Brunetti, Maura, Sbaiz, Luca, Lai, Shiao-lin, Abramzon, Yevgeniya, Traynor, Bryan J and Restagno, Gabriella (2011), A De Novo Missense Mutation of the FUS Gene in a 'True' Sporadic ALS Case. *Neurobiology of Aging*, 32(3).

Chow, Clement Y, Landers, John E, Bergren, Sarah K, Sapp, Peter C, Grant, Adrienne E, Jones, Julie M, Everett, Lesley, Lenk, Guy M, McKenna-Yasek, Diane M, Weisman, Lois S, Figlewicz, Denise, Brown, Robert H and Meisler, Miriam H (2009), Deleterious Variants of FIG4, a Phosphoinositide Phosphatase, in Patients with ALS. *American journal of human genetics*, 84(1): 85–88.

Cifra, Alessandra, Nani, Francesca and Nistri, Andrea (2011), Riluzole Is a Potent Drug to Protect Neonatal Rat Hypoglossal Motoneurons in Vitro From Excitotoxicity Due to Glutamate Uptake Block. *The European journal of neuroscience*, 33(5): 899–913.

Cirulli, Elizabeth T, Lasseigne, Brittany N, Petrovski, Slavé, Sapp, Peter C, Dion, Patrick a, Leblond, Claire S, Couthouis, Julien, Lu, Yi-Fan, Wang, Quanli, Krueger, Brian J, Ren, Zhong, Keebler, Jonathan, Han, Yujun, Levy, Shawn E, Boone, Braden E, Wimbish, Jack R, Waite, Lindsay L, Jones, Angela L, Carulli, John P, Day-Williams, Aaron G, Staropoli, John F, Xin, Winnie W, Chesi, Alessandra, Raphael, Alya R, Mckenna-yasek, Diane, Cady, Janet, Vianney de Jong, J M B, Kenna, Kevin P, Smith, Bradley N, Topp, Simon, Miller, Jack, Gkazi, Athina, FALS Sequencing Consortium, Al-Chalabi, Ammar, van den Berg, Leonard H, Veldink, Jan, Silani, Vincenzo, Ticozzi, Nicola, Shaw, Christopher E, Baloh, Robert H, Appel, Stanley, Simpson, Ericka, Lagier-Tourenne, Clotilde, Pulst, Stefan M, Gibson, Summer, Trojanowski, John Q, Elman, Lauren, McCluskey, Leo, Grossman, Murray, Shneider, Neil A, Chung, Wendy K, Ravits, John M, Glass, Jonathan D, Sims, Katherine B, Van Deerlin, Vivianna M, Maniatis, Tom, Hayes, Sebastian D, Ordureau, Alban, Swarup, Sharan, Landers, John, Baas, Frank, Allen, Andrew S, Bedlack, Richard S, Harper, J Wade, Gitler, Aaron D, Rouleau, Guy a, Brown, Robert, Harms, Matthew B, Cooper, Gregory M, Harris, Tim, Myers, Richard M and Goldstein, David B (2015), Exome Sequencing in Amyotrophic Lateral Sclerosis Identifies Risk Genes and Pathways. *Science (New York, N.Y.)*, 347(6229): 1436–1441.

Clement, A M, Nguyen, M D, Roberts, E A, Garcia, M L, Boillee, S, Rule, M, McMahon, A P, Doucette, W, Siwek, D, Ferrante, R J, Brown, R H, Julien, J P, Goldstein, LSB and Cleveland, D W (2003), Wild-Type Nonneuronal Cells Extend Survival of SOD1 Mutant Motor Neurons in ALS Mice. *Science*, 302(5642): 113–117.

Cohen, Todd J, Hwang, Andrew W, Restrepo, Clark R, Yuan, Chao-Xing,

Trojanowski, John Q and Lee, Virginia M-Y (2015), An Acetylation Switch Controls TDP-43 Function and Aggregation Propensity. *Nature communications*, 6: 5845.

Colombrita, Claudia, Onesto, Elisa, Buratti, Emanuele, la Grange, de, Pierre, Gumina, Valentina, Baralle, Francisco E, Silani, Vincenzo and Ratti, Antonia (2015), From Transcriptomic to Protein Level Changes in TDP-43 and FUS Loss-of-Function Cell Models. *Biochimica et biophysica acta*.

Colombrita, Claudia, Zennaro, Eleonora, Fallini, Claudia, Weber, Markus, Sommacal, Andreas, Buratti, Emanuele, Silani, Vincenzo and Ratti, Antonia (2009), TDP-43 Is Recruited to Stress Granules in Conditions of Oxidative Insult. *Journal of neurochemistry*, 111(4): 1051–1061.

Cooper, Thomas a, Wan, Lili and Dreyfuss, Gideon (2009), RNA and Disease. *Cell*, 136(4): 777–793.

Cooper-Knock, Johnathan, Hewitt, Christopher, Highley, J Robin, Brockington, Alice, Milano, Antonio, Man, Somai, Martindale, Joanne, Hartley, Judith, Walsh, Theresa, Gelsthorpe, Catherine, Baxter, Lynne, Forster, Gillian, Fox, Melanie, Bury, Joanna, Mok, Kin, McDermott, Christopher J, Traynor, Bryan J, Kirby, Janine, Wharton, Stephen B, Ince, Paul G, Hardy, John and Shaw, Pamela J (2012), Clinico-Pathological Features in Amyotrophic Lateral Sclerosis with Expansions in C9ORF72. *Brain : a journal of neurology*, 135(Pt 3): 751–764.

Cooper-Knock, Johnathan, Higginbottom, Adrian, Stopford, Matthew J, Highley, J Robin, Ince, Paul G, Wharton, Stephen B, Pickering-Brown, Stuart, Kirby, Janine, Hautbergue, Guillaume M and Shaw, Pamela J (2015), Antisense RNA Foci in the Motor Neurons of C9ORF72-ALS Patients Are Associated with TDP-43 Proteinopathy. *Acta neuropathologica*: 1–13.

Cooper-Knock, Johnathan, Kirby, Janine, Ferraiuolo, Laura, Heath, Paul R, Rattray, Magnus and Shaw, Pamela J (2012), Gene Expression Profiling in Human Neurodegenerative Disease. *Nature Reviews Neurology*, 8(9): 518–530.

Cooper-Knock, Johnathan, Walsh, Matthew J, Higginbottom, Adrian, Robin Highley, J, Dickman, Mark J, Edbauer, Dieter, Ince, Paul G, Wharton, Stephen B, Wilson, Stuart A, Kirby, Janine, Hautbergue, Guillaume M and Shaw, Pamela J (2014), Sequestration of Multiple RNA Recognition Motif-Containing Proteins by C9orf72 Repeat Expansions. *Brain : a journal of neurology*, 137(Pt 7): 2040–2051.

Corrado, Lucia, Del Bo, Roberto, Castellotti, Barbara, Ratti, Antonia, Cereda, Cristina, Penco, Silvana, Sorarù, Gianni, Carlomagno, Yari, Ghezzi, Serena, Pensato, Viviana, Colombrita, Claudia, Gagliardi, Stella, Cozzi, Lorena, Orsetti, Valeria, Mancuso, Michelangelo, Siciliano, Gabriele, Mazzini, Letizia, Comi, Giacomo Pietro, Gellera, Cinzia, Ceroni, Mauro, D'Alfonso, Sandra and Silani, Vincenzo (2010), Mutations of FUS Gene in Sporadic Amyotrophic Lateral Sclerosis. *Journal of medical genetics*, 47(3): 190–194.

Costello, Joseph, Castelli, Lydia M, Rowe, William, Kershaw, Christopher J, Talavera, David, Mohammad-Qureshi, Sarah S, Sims, Paul F G, Grant, Christopher M, Pavitt, Graham D, Hubbard, Simon J and Ashe, Mark P

(2015), Global mRNA Selection Mechanisms for Translation Initiation. *Genome biology*, 16(1): 10.

Coyne, Alyssa N, Yamada, Shizuka B, Siddegowda, Bhavani Bagevalu, Estes, Patricia S, Zaepfel, Benjamin L, Johannesmeyer, Jeffrey S, Lockwood, Donovan B, Pham, Linh T, Hart, Michael P, Cassel, Joel A, Freibaum, Brian, Boehringer, Ashley V, Taylor, J Paul, Reitz, Allen B, Gitler, Aaron D and Zarnescu, Daniela C (2015), Fragile X Protein Mitigates TDP-43 Toxicity by Remodeling RNA Granules and Restoring Translation. *Human molecular genetics*: ddv389.

Cozzolino, Mauro, Ferri, Alberto, Valle, Cristiana and Carrì, Maria Teresa (2013), Mitochondria and ALS: Implications From Novel Genes and Pathways. *Molecular and cellular neurosciences*, 55: 44–49.

Crabtree, Benedict, Thiyagarajan, Nethaji, Prior, Stephen H, Wilson, Peter, Iyer, Shalini, Ferns, Tyrone, Shapiro, Robert, Brew, Keith, Subramanian, Vasanta and Acharya, K Ravi (2007), Characterization of Human Angiogenin Variants Implicated in Amyotrophic Lateral Sclerosis. *Biochemistry*, 46(42): 11810–11818.

Cragnaz, L, Klima, R, De Conti, L, Romano, G, Feiguin, F, Buratti, E, Baralle, M and Baralle, F E (2015), An Age-Related Reduction of Brain TBPH/TDP-43 Levels Precedes the Onset of Locomotion Defects in a Drosophila ALS Model. *Neuroscience*, 311: 415–421.

Cragnaz, Lucia, Klima, Raffaela, Skoko, Natasa, Budini, Mauricio, Feiguin, Fabian and Baralle, Francisco E (2014), Aggregate Formation Prevents dTDP-43 Neurotoxicity in the Drosophila Melanogaster Eye. *Neurobiology of disease*, 71: 74–80.

D'Alessandro, Giuseppina, Calcagno, Eleonora, Tartari, Silvia, Rizzardini, Milena, Invernizzi, Roberto William and Cantoni, Lavinia (2011), Glutamate and Glutathione Interplay in a Motor Neuronal Model of Amyotrophic Lateral Sclerosis Reveals Altered Energy Metabolism. *Neurobiology of disease*, 43(2): 346–355.

D'Alton, Simon, Altshuler, Marcelle and Lewis, Jada (2015), Studies of Alternative Isoforms Provide Insight Into TDP-43 Autoregulation and Pathogenesis. *RNA (New York, N.Y.)*, 21(8): 1419–1432.

De Conti, Laura, Akinyi, Maureen V, Mendoza-Maldonado, Ramiro, Romano, Maurizio, Baralle, Marco and Buratti, Emanuele (2015), TDP-43 Affects Splicing Profiles and Isoform Production of Genes Involved in the Apoptotic and Mitotic Cellular Pathways. *Nucleic acids research*.

De Vos, Kurt J, Chapman, Anna L, Tennant, Maria E, Manser, Catherine, Tudor, Elizabeth L, Lau, Kwok-Fai, Brownlees, Janet, Ackerley, Steven, Shaw, Pamela J, McLoughlin, Declan M, Shaw, Christopher E, Leigh, P Nigel, Miller, Christopher C J and Grierson, Andrew J (2007), Familial Amyotrophic Lateral Sclerosis-Linked SOD1 Mutants Perturb Fast Axonal Transport to Reduce Axonal Mitochondria Content. *Human molecular genetics*, 16(22): 2720–2728.

De Vos, Kurt J, Grierson, Andrew J, Ackerley, Steven and Miller, Christopher C

J (2008), Role of Axonal Transport in Neurodegenerative Diseases. *Annual review of neuroscience*, 31: 151–173.

DeJesus-Hernandez, Mariely, Mackenzie, Ian R, Boeve, Bradley F, Boxer, Adam L, Baker, Matt, Rutherford, Nicola J, Nicholson, Alexandra M, Finch, NiCole a, Flynn, Heather, Adamson, Jennifer, Kouri, Naomi, Wojtas, Aleksandra, Sengdy, Pheth, Hsiung, Ging-Yuek R, Karydas, Anna, Seeley, William W, Josephs, Keith a, Coppola, Giovanni, Geschwind, Daniel H, Wszolek, Zbigniew K, Feldman, Howard, Knopman, David S, Petersen, Ronald C, Miller, Bruce L, Dickson, Dennis W, Boylan, Kevin B, Graff-Radford, Neill R and Rademakers, Rosa (2011), Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2): 245–256.

Deng, Han-Xiang, Chen, Wenjie, Hong, Seong-Tshool, Boycott, Kym M, Gorrie, George H, Siddique, Nailah, Yang, Yi, Fecto, Faisal, Shi, Yong, Zhai, Hong, Jiang, Hujun, Hirano, Makito, Rampersaud, Evadnie, Jansen, Gerard H, Donkervoort, Sandra, Bigio, Eileen H, Brooks, Benjamin R, Ajroud, Kaouther, Sufit, Robert L, Haines, Jonathan L, Mugnaini, Enrico, Pericak-Vance, Margaret a and Siddique, Teepu (2011), Mutations in UBQLN2 Cause Dominant X-Linked Juvenile and Adult-Onset ALS and ALS/Dementia. *Nature*, 477(7363): 211–215.

Dewey, Colleen M, Cenik, Basar, Sephton, Chantelle F, Dries, Daniel R, Mayer, Paul, Good, Shannon K, Johnson, Brett A, Herz, Joachim and Yu, Gang (2011), TDP-43 Is Directed to Stress Granules by Sorbitol, a Novel Physiological Osmotic and Oxidative Stressor. *Molecular and cellular biology*, 31(5): 1098–1108.

Diaper, Danielle C, Adachi, Yoshitsugu, Sutcliffe, Ben, Humphrey, Dickon M, Elliott, Christopher J H, Stepto, Alan, Ludlow, Zoe N, Vanden Broeck, Lies, Callaerts, Patrick, Dermaut, Bart, Al-Chalabi, Ammar, Shaw, Christopher E, Robinson, Iain M and Hirth, Frank (2013), Loss and Gain of Drosophila TDP-43 Impair Synaptic Efficacy and Motor Control Leading to Age-Related Neurodegeneration by Loss-of-Function Phenotypes. *Human molecular genetics*, 22(8): 1539–1557.

Domínguez-Escobar, Julia, Chastanet, Arnaud, Crevenna, Alvaro H, Fromion, Vincent, Wedlich-Söldner, Roland and Carballido-López, Rut (2011), Processive Movement of MreB-Associated Cell Wall Biosynthetic Complexes in Bacteria. *Science (New York, N.Y.)*, 333(6039): 225–228.

Domínguez-Sánchez, María S, Sáez, Carmen, Japón, Miguel A, Aguilera, Andrés and Luna, Rosa (2011), Differential Expression of THOC1 and ALY mRNP Biogenesis/Export Factors in Human Cancers. *BMC cancer*, 11(1): 77.

Dormann, Dorothee, Capell, Anja, Carlson, Aaron M, Shankaran, Sunita S, Rodde, Ramona, Neumann, Manuela, Kremmer, Elisabeth, Matsuwaki, Takashi, Yamanouchi, Keitaro, Nishihara, Masugi and Haass, Christian (2009), Proteolytic Processing of TAR DNA Binding Protein-43 by Caspases Produces C-Terminal Fragments with Disease Defining Properties Independent of Progranulin. *Journal of neurochemistry*, 110(3): 1082–1094.

Dormann, Dorothee, Rodde, Ramona, Edbauer, Dieter, Bentmann, Eva,

Fischer, Ingeborg, Hruscha, Alexander, Than, Manuel E, Mackenzie, Ian R a, Capell, Anja, Schmid, Bettina, Neumann, Manuela and Haass, Christian (2010), ALS-Associated Fused in Sarcoma (FUS) Mutations Disrupt Transportin-Mediated Nuclear Import. *The EMBO journal*, 29(16): 2841–2857.

Dreyfuss, G, Matunis, M J, Piñol-Roma, S and Burd, C G (1993), hnRNP Proteins and the Biogenesis of mRNA. *Annual Review of Biochemistry*, 62: 289–321.

Dreyfuss, Gideon, Kim, V Narry and Kataoka, Naoyuki (2002), Messenger-RNA-Binding Proteins and the Messages They Carry. *Nature reviews. Molecular cell biology*, 3(3): 195–205.

Egawa, Naohiro, Kitaoka, Shiho, Tsukita, Kayoko, Naitoh, Motoko, Takahashi, Kazutoshi, Yamamoto, Takuya, Adachi, Fumihiko, Kondo, Takayuki, Okita, Keisuke, Asaka, Isao, Aoi, Takashi, Watanabe, Akira, Yamada, Yasuhiro, Morizane, Asuka, Takahashi, Jun, Ayaki, Takashi, Ito, Hidefumi, Yoshikawa, Katsuhiro, Yamawaki, Satoko, Suzuki, Shigehiko, Watanabe, Dai, Hioki, Hiroyuki, Kaneko, Takeshi, Makioka, Kouki, Okamoto, Koichi, Takuma, Hiroshi, Tamaoka, Akira, Hasegawa, Kazuko, Nonaka, Takashi, Hasegawa, Masato, Kawata, Akihiro, Yoshida, Minoru, Nakahata, Tatsutoshi, Takahashi, Ryosuke, Marchetto, Maria C N, Gage, Fred H, Yamanaka, Shinya and Inoue, Haruhisa (2012), Drug Screening for ALS Using Patient-Specific Induced Pluripotent Stem Cells. *Science translational medicine*, 4(145): –145ra104.

Eggett, C J, Crosier, S, Manning, P, Cookson, M R, Menzies, F M, McNeil, C J and Shaw, P J (2000), Development and Characterisation of a Glutamate-Sensitive Motor Neurone Cell Line. *Journal of neurochemistry*, 74(5): 1895–1902.

Elden, Andrew C, Kim, Hyung-Jun, Hart, Michael P, Chen-Plotkin, Alice S, Johnson, Brian S, Fang, Xiaodong, Armakola, Maria, Geser, Felix, Greene, Robert, Lu, Min Min, Padmanabhan, Arun, Clay-Falcone, Dana, McCluskey, Leo, Elman, Lauren, Juhr, Denise, Gruber, Peter J, Rüb, Udo, Auburger, Georg, Trojanowski, John Q, Lee, Virginia M-Y, Van Deerlin, Vivianna M, Bonini, Nancy M and Gitler, Aaron D (2010), Ataxin-2 Intermediate-Length Polyglutamine Expansions Are Associated with Increased Risk for ALS. *Nature*, 466(7310): 1069–1075.

Elkon, Ran, Ugalde, Alejandro P and Agami, Reuven (2013), Alternative Cleavage and Polyadenylation: Extent, Regulation and Function. *Nature Publishing Group*, 14(7): 496–506.

Elvira, George, Wasiak, Sylwia, Blandford, Vanessa, Tong, Xin-Kang, Serrano, Alexandre, Fan, Xiaotang, del Rayo Sánchez-Carbente, Maria, Servant, Florence, Bell, Alexander W, Boismenu, Daniel, Lacaille, Jean-Claude, McPherson, Peter S, DesGroseillers, Luc and Sossin, Wayne S (2006), Characterization of an RNA Granule From Developing Brain. *Molecular & cellular proteomics : MCP*, 5(4): 635–651.

Estes, Patricia S, Boehringer, Ashley, Zwick, Rebecca, Tang, Jonathan E, Grigsby, Brianna and Zarnescu, Daniela C (2011), Wild-Type and A315T Mutant TDP-43 Exert Differential Neurotoxicity in a Drosophila Model of

ALS. *Human molecular genetics*, 20(12): 2308–2321.

Estes, Patricia S, Daniel, Scott G, McCallum, Abigail P, Boehringer, Ashley V, Sukhina, Alona S, Zwick, Rebecca A and Zarnescu, Daniela C (2013), Motor Neurons and Glia Exhibit Specific Individualized Responses to TDP-43 Expression in a Drosophila Model of Amyotrophic Lateral Sclerosis. *Disease models & mechanisms*, 6(3): 721–733.

Evans, M C, Couch, Y, Sibson, N and Turner, M R (2013), Inflammation and Neurovascular Changes in Amyotrophic Lateral Sclerosis. *Molecular and cellular neurosciences*, 53: 34–41.

Fallini, Claudia, Bassell, Gary J and Rossoll, Wilfried (2012), The ALS Disease Protein TDP-43 Is Actively Transported in Motor Neuron Axons and Regulates Axon Outgrowth. *Human molecular genetics*, 21(16): 3703–3718.

Fecto, Faisal, Yan, Jianhua, Vemula, S Pavan, Liu, Erdong, Yang, Yi, Chen, Wenjie, Zheng, Jian Guo, Shi, Yong, Siddique, Nailah, Arrat, Hasan, Donkervoort, Sandra, Ajroud-Driss, Senda, Sufit, Robert L, Heller, Scott L, Deng, Han-Xiang and Siddique, Teepu (2011), SQSTM1 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Archives of Neurology*, 68(11): 1440–1446.

Feiguin, Fabian, Godena, Vinay K, Romano, Giulia, D'Ambrogio, Andrea, Klima, Raffaella and Baralle, Francisco E (2009), Depletion of TDP-43 Affects Drosophila Motoneurons Terminal Synapsis and Locomotive Behavior. *FEBS letters*, 583(10): 1586–1592.

Ferraiuolo, Laura, Kirby, Janine, Grierson, Andrew J, Sendtner, Michael and Shaw, Pamela J (2011), Molecular Pathways of Motor Neuron Injury in Amyotrophic Lateral Sclerosis. *Nature Reviews Neurology*, 7(11): 616–630.

Fiesel, Fabienne C, Voigt, Aaron, Weber, Stephanie S, Van den Haute, Chris, Waldenmaier, Andrea, Görner, Karin, Walter, Michael, Anderson, Marlene L, Kern, Jeannine V, Rasse, Tobias M, Schmidt, Thorsten, Springer, Wolfdieter, Kirchner, Roland, Bonin, Michael, Neumann, Manuela, Baekelandt, Veerle, Alunni-Fabbroni, Marianna, Schulz, Jörg B and Kahle, Philipp J (2010), Knockdown of Transactive Response DNA-Binding Protein (TDP-43) Downregulates Histone Deacetylase 6. *The EMBO journal*, 29(1): 209–221.

Fiesel, Fabienne C, Weber, Stephanie S, Supper, Jochen, Zell, Andreas and Kahle, Philipp J (2012), TDP-43 Regulates Global Translational Yield by Splicing of Exon Junction Complex Component SKAR. *Nucleic acids research*, 40(6): 2668–2682.

Forman, Mark S, Trojanowski, John Q and Lee, Virginia M-Y (2004), Neurodegenerative Diseases: a Decade of Discoveries Paves the Way for Therapeutic Breakthroughs. *Nature medicine*, 10(10): 1055–1063.

Freeberg, Lindsay, Kuersten, Scott and Syed, Fraz (2013), Isolate and Sequence Ribosome-Protected mRNA Fragments Using Size-Exclusion Chromatography. *Nature Publishing Group*, 10(5): i–ii.

Freibaum, Brian D, Chitta, Raghu K, High, Anthony A and Taylor, J Paul (2010),

Global Analysis of TDP-43 Interacting Proteins Reveals Strong Association with RNA Splicing and Translation Machinery. *Journal of Proteome Research*, 9(2): 1104–1120.

Freibaum, Brian D, Lu, Yubing, Lopez-Gonzalez, Rodrigo, Kim, Nam Chul, Almeida, Sandra, Lee, Kyung-Ha, Badders, Nisha, Valentine, Marc, Miller, Bruce L, Wong, Philip C, Petrucelli, Leonard, Kim, Hong Joo, Gao, Fen-Biao and Taylor, J Paul (2015), GGGGCC Repeat Expansion in C9orf72 Compromises Nucleocytoplasmic Transport. *Nature*, 525(7567): 129–133.

Freischmidt, Axel, Müller, Kathrin, Ludolph, Albert C, Weishaupt, Jochen H and Andersen, Peter M (2016), Association of Mutations in TBK1 with Sporadic and Familial Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *JAMA neurology*.

Freischmidt, Axel, Wieland, Thomas, Richter, Benjamin, Ruf, Wolfgang, Schaeffer, Veronique, Müller, Kathrin, Marroquin, Nicolai, Nordin, Frida, Hübers, Annemarie, Weydt, Patrick, Pinto, Susana, Press, Rayomond, Millecamps, Stéphanie, Molko, Nicolas, Bernard, Emilien, Desnuelle, Claude, Soriani, Marie-Hélène, Dorst, Johannes, Graf, Elisabeth, Nordström, Ulrika, Feiler, Marisa S, Putz, Stefan, Boeckers, Tobias M, Meyer, Thomas, Winkler, Andrea S, Winkelman, Juliane, de Carvalho, Mamede, Thal, Dietmar R, Otto, Markus, Brännström, Thomas, Volk, Alexander E, Kursula, Petri, Danzer, Karin M, Lichtner, Peter, Dikic, Ivan, Meitinger, Thomas, Ludolph, Albert C, Strom, Tim M, Andersen, Peter M and Weishaupt, Jochen H (2015), Haploinsufficiency of TBK1 Causes Familial ALS and Fronto-Temporal Dementia. *Nature neuroscience*, 18(5): 631–636.

Fuentealba, Rodrigo A, Udan, Maria, Bell, Shaughn, Wegorzewska, Iga, Shao, Jieya, Diamond, Marc I, Weihl, Conrad C and Baloh, Robert H (2010), Interaction with Polyglutamine Aggregates Reveals a Q/N-Rich Domain in TDP-43. *Journal of Biological Chemistry*, 285(34): 26304–26314.

Fugier, Charlotte, Klein, Arnaud F, Hammer, Caroline, Vassilopoulos, Stéphane, Ivarsson, Ylva, Toussaint, Anne, Tosch, Valérie, Vignaud, Alban, Ferry, Arnaud, Messaddeq, Nadia, Kokunai, Yosuke, Tsuburaya, Rie, la Grange, de, Pierre, Dembele, Doulaye, Francois, Virginie, Precigout, Guillaume, Boulade-Ladame, Charlotte, Hummel, Marie-Christine, Lopez de Munain, Adolfo, Sergeant, Nicolas, Laquerrière, Annie, Thibault, Christelle, Deryckere, François, Auboeuf, Didier, Garcia, Luis, Zimmermann, Pascale, Udd, Bjarne, Schoser, Benedikt, Takahashi, Masanori P, Nishino, Ichizo, Bassez, Guillaume, Laporte, Jocelyn, Furling, Denis and Charlet-Berguerand, Nicolas (2011), Misregulated Alternative Splicing of BIN1 Is Associated with T Tubule Alterations and Muscle Weakness in Myotonic Dystrophy. *Nature medicine*, 17(6): 720–725.

Fukuda, Toru, Yamagata, Kaoru, Fujiyama, Sally, Matsumoto, Takahiro, Koshida, Iori, Yoshimura, Kimihiro, Mihara, Masatomo, Naitou, Masanori, Endoh, Hideki, Nakamura, Takashi, Akimoto, Chihiro, Yamamoto, Yoko, Katagiri, Takenobu, Foulds, Charles, Takezawa, Shinichiro, Kitagawa, Hirochika, Takeyama, Ken-ichi, O'Malley, Bert W and Kato, Shigeaki (2007), DEAD-Box RNA Helicase Subunits of the Drosha Complex Are Required for Processing of rRNA and a Subset of microRNAs. *Nature cell biology*, 9(5): 604–611.

Gendron, Tania F, Bieniek, Kevin F, Zhang, Yong-Jie, Jansen-West, Karen, Ash, Peter E a, Caulfield, Thomas, Daughrity, Lillian, Dunmore, Judith H, Castanedes-Casey, Monica, Chew, Jeannie, Cosio, Danielle M, van Blitterswijk, Marka, Lee, Wing C, Rademakers, Rosa, Boylan, Kevin B, Dickson, Dennis W and Petrucelli, Leonard (2013), Antisense Transcripts of the Expanded C9ORF72 Hexanucleotide Repeat Form Nuclear RNA Foci and Undergo Repeat-Associated Non-ATG Translation in c9FTD/ALS. *Acta neuropathologica*.

Gerashchenko, Maxim V and Gladyshev, Vadim N (2014), Translation Inhibitors Cause Abnormalities in Ribosome Profiling Experiments. *Nucleic acids research*, 42(17): e134.

Gerber, André P, Luschnig, Stefan, Krasnow, Mark A, Brown, Patrick O and Herschlag, Daniel (2006), Genome-Wide Identification of mRNAs Associated with the Translational Regulator PUMILIO in Drosophila Melanogaster. *PNAS*, 103(12): 4487–4492.

Giordana, Maria Teresa, Piccinini, Marco, Grifoni, Silvia, De Marco, Giovanni, Vercellino, Marco, Magistrello, Michela, Pellerino, Alessia, Buccinnà, Barbara, Lupino, Elisa and Rinaudo, Maria Teresa (2010), TDP□43 Redistribution Is an Early Event in Sporadic Amyotrophic Lateral Sclerosis. *Brain Pathology*, 20(2): 351–360.

Gitcho, Michael A, Bigio, Eileen H, Mishra, Manjari, Johnson, Nancy, Weintraub, Sandra, Mesulam, Marsel, Rademakers, Rosa, Chakraverty, Sumi, Cruchaga, Carlos and Morris, John C (2009), TARDBP 3′-UTR Variant in Autopsy-Confirmed Frontotemporal Lobar Degeneration with TDP-43 Proteinopathy. *Acta neuropathologica*, 118(5): 633–645.

Glaus, Peter, Honkela, Antti and Rattray, Magnus (2012), Identifying Differentially Expressed Transcripts From RNA-Seq Data with Biological Variation. *Bioinformatics (Oxford, England)*, 28(13): 1721–1728.

Glisovic, Tina, Bachorik, Jennifer L, Yong, Jeongsik and Dreyfuss, Gideon (2008), RNA-Binding Proteins and Post-Transcriptional Gene Regulation. *FEBS letters*, 582(14): 1977–1986.

Greenway, Matthew J, Andersen, Peter M, Russ, Carsten, Ennis, Sean, Cashman, Susan, Donaghy, Colette, Patterson, Victor, Swingler, Robert, Kieran, Dairin, Prehn, Jochen, Morrison, Karen E, Green, Andrew, Acharya, K Ravi, Brown, Robert H and Hardiman, Orla (2006), ANG Mutations Segregate with Familial and 'Sporadic' Amyotrophic Lateral Sclerosis. *Nature genetics*, 38(4): 411–413.

Gregory, Richard I, Chendrimada, Thimmaiah P, Cooch, Neil and Shiekhattar, Ramin (2005), Human RISC Couples microRNA Biogenesis and Posttranscriptional Gene Silencing. *Cell*, 123(4): 631–640.

Gregory, Richard I, Yan, Kai-ping, Amuthan, Govindasamy, Chendrimada, Thimmaiah, Doratotaj, Behzad, Cooch, Neil and Shiekhattar, Ramin (2004), The Microprocessor Complex Mediates the Genesis of microRNAs. *Nature*, 432(7014): 235–240.

Guo, Huili, Ingolia, Nicholas T, Weissman, Jonathan S and Bartel, David P

(2010), Mammalian microRNAs Predominantly Act to Decrease Target mRNA Levels. *Nature*, 466(7308): 835–840.

Guo, Yansu, Wang, Qian, Zhang, Kunxi, An, Ting, Shi, Pengxiao, Li, Zhongyao, Duan, Weisong and Li, Chunyan (2012), HO-1 Induction in Motor Cortex and Intestinal Dysfunction in TDP-43 A315T Transgenic Mice. *Brain research*, 1460: 88–95.

Hadano, S, Hand, C K, Osuga, H, Yanagisawa, Y, Otomo, a, Devon, R S, Miyamoto, N, Showguchi-Miyata, J, Okada, Y, Singaraja, R, Figlewicz, D A, Kwiatkowski, T, Hosler, B A, Sagie, T, Skaug, J, Nasir, J, Brown, R H, Scherer, S W, Rouleau, G a, Hayden, M R and Ikeda, J E (2001), A Gene Encoding a Putative GTPase Regulator Is Mutated in Familial Amyotrophic Lateral Sclerosis 2. *Nature genetics*, 29(2): 166–173.

Haeusler, Aaron R, Donnelly, Christopher J, Periz, Goran, Simko, Eric A J, Shaw, Patrick G, Kim, Min-Sik, Maragakis, Nicholas J, Troncoso, Juan C, Pandey, Akhilesh, Sattler, Rita, Rothstein, Jeffrey D and Wang, Jiou (2014), C9orf72 Nucleotide Repeat Structures Initiate Molecular Cascades of Disease. *Nature*, 507(7491): 195–200.

Halbeisen, Regula E, Scherrer, Tanja and Gerber, André P (2009), Affinity Purification of Ribosomes to Access the Translatome. *Methods (San Diego, Calif.)*, 48(3): 306–310.

Hand, Collette K, Khoris, Jawad, Salachas, Francois, Gros-Louis, François, Lopes, Ana Amélia Simões, Mayeux-Portas, Veronique, Brewer, Carl G, Brown, Robert H, Meininger, Vincent, Camu, William and Rouleau, Guy a (2002), A Novel Locus for Familial Amyotrophic Lateral Sclerosis, on Chromosome 18q. *American journal of human genetics*, 70(1): 251–256.

Hasegawa, Masato, Arai, Tetsuaki, Nonaka, Takashi, Kametani, Fuyuki, Yoshida, Mari, Hashizume, Yoshio, Beach, Thomas G, Buratti, Emanuele, Baralle, Francisco, Morita, Mitsuya, Nakano, Imaharu, Oda, Tatsuro, Tsuchiya, Kuniaki and Akiyama, Haruhiko (2008), Phosphorylated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Annals of neurology*, 64(1): 60–70.

Hatzipetros, Theo, Bogdanik, Laurent P, Tassinari, Valerie R, Kidd, Joshua D, Moreno, Andy J, Davis, Crystal, Osborne, Melissa, Austin, Andrew, Vieira, Fernando G, Lutz, Cathleen and Perrin, Steve (2014), C57BL/6J Congenic Prp-TDP43A315T Mice Develop Progressive Neurodegeneration in the Myenteric Plexus of the Colon Without Exhibiting Key Features of ALS. *Brain research*, 1584: 59–72.

Hazelett, Dennis J, Chang, Jer-Cherng, Lakeland, Daniel L and Morton, David B (2012), Comparison of Parallel High-Throughput RNA Sequencing Between Knockout of TDP-43 and Its Overexpression Reveals Primarily Nonreciprocal and Nonoverlapping Gene Expression Changes in the Central Nervous System of Drosophila. *G3 (Bethesda, Md.)*, 2(7): 789–802.

He, Fang, Krans, Amy, Freibaum, Brian D, Paul Taylor, J and Todd, Peter K (2014), TDP-43 Suppresses CGG Repeat-Induced Neurotoxicity Through Interactions with HnRNP A2/B1. *Human molecular genetics*, 23(19): 5036–5051.

Heath, Paul R, Kirby, Janine and Shaw, Pamela J (2013), Investigating Cell Death Mechanisms in Amyotrophic Lateral Sclerosis Using Transcriptomics. *Frontiers in Cellular Neuroscience*, 7.

Heiman, Myriam, Schaefer, Anne, Gong, Shiaoching, Peterson, Jayms D, Day, Michelle, Ramsey, Keri E, Suárez-Fariñas, Mayte, Schwarz, Cordelia, Stephan, Dietrich A, Surmeier, D James, Greengard, Paul and Heintz, Nathaniel (2008), A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell*, 135(4): 738–748.

Henkel, Jenny S, Engelhardt, Joseph I, Siklós, László, Simpson, Ericka P, Kim, Seung H, Pan, Tianhong, Goodman, J Clay, Siddique, Teepu, Beers, David R and Appel, Stanley H (2004), Presence of Dendritic Cells, MCP-1, and Activated Microglia/Macrophages in Amyotrophic Lateral Sclerosis Spinal Cord Tissue. *Annals of neurology*, 55(2): 221–235.

Hentati, a, Bejaoui, K, Pericak-Vance, M A, Hentati, F, Speer, M C, Hung, W Y, Figlewicz, D A, Haines, J, Rimmler, J and Ben Hamida, C (1994), Linkage of Recessive Familial Amyotrophic Lateral Sclerosis to Chromosome 2q33-Q35. *Nature genetics*, 7(3): 425–428.

Herdewyn, Sarah, Cirillo, Carla, Van Den Bosch, Ludo, Robberecht, Wim, Vanden Berghe, Pieter and Van Damme, Philip (2014), Prevention of Intestinal Obstruction Reveals Progressive Neurodegeneration in Mutant TDP-43 (A315T) Mice. *Molecular Neurodegeneration*, 9(1): 24.

Hewamadduma, Channa A A, Grierson, Andrew J, Ma, Taylur P, Pan, Luyuan, Moens, Cecilia B, Ingham, Philip W, Ramesh, Tennore and Shaw, Pamela J (2013), Tardbpl Splicing Rescues Motor Neuron and Axonal Development in a Mutant Tardbp Zebrafish. *Human molecular genetics*, 22(12): 2376–2386.

Higashi, Shinji, Kabuta, Tomohiro, Nagai, Yoshitaka, Tsuchiya, Yukihiro, Akiyama, Haruhiko and Wada, Keiji (2013), TDP-43 Associates with Stalled Ribosomes and Contributes to Cell Survival During Cellular Stress. *Journal of neurochemistry*, 126(2): 288–300.

Higashi, Shinji, Tsuchiya, Yukihiro, Araki, Toshiyuki, Wada, Keiji and Kabuta, Tomohiro (2010), TDP-43 Physically Interacts with Amyotrophic Lateral Sclerosis-Linked Mutant CuZn Superoxide Dismutase. *Neurochemistry international*, 57(8): 906–913.

Highley, J Robin, Kirby, Janine, Jansweijer, Joeri A, Webb, Philip S, Hewamadduma, Channa A, Heath, Paul R, Higginbottom, Adrian, Raman, Rohini, Ferraiuolo, Laura, Cooper-Knock, Johnathan, McDermott, Christopher J, Wharton, Stephen B, Shaw, Pamela J and Ince, Paul G (2014), Loss of Nuclear TDP-43 in Amyotrophic Lateral Sclerosis (ALS) Causes Altered Expression of Splicing Machinery and Widespread Dysregulation of RNA Splicing in Motor Neurones. *Neuropathology and applied neurobiology*, 40(6): 670–685.

Hill, Sarah J, Mordes, Daniel A, Cameron, Lisa A, Neuberg, Donna S, Landini, Serena, Eggan, Kevin and Livingston, David M (2016), Two Familial ALS Proteins Function in Prevention/Repair of Transcription-Associated DNA Damage. *Proceedings of the National Academy of Sciences of the United States of America*.

Hinds, H L, Ashley, C T, Sutcliffe, J S, Nelson, D L, Warren, S T, Housman, D and Schalling, M (1993), Tissue Specific Expression of FMR-1 Provides Evidence for a Functional Role in Fragile X Syndrome. *Nature genetics*, 3(1): 36–43.

Hinnebusch, Alan G (2011), Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. *Microbiology and molecular biology reviews : MMBR*, 75(3): 434–67– first page of table of contents.

Hirano, a, Kurland, L T and Sayre, G P (1967), Familial Amyotrophic Lateral Sclerosis - a Subgroup Characterized by Posterior and Spinocerebellar Tract Involvement and Hyaline Inclusions in Anterior Horn Cells. *Archives of Neurology*, 16(3): 232–&.

Hirokawa, Nobutaka (2006), mRNA Transport in Dendrites: RNA Granules, Motors, and Tracks. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(27): 7139–7142.

Holmes, S E, O'Hearn, E, Rosenblatt, a, Callahan, C, Hwang, H S, Ingersoll-Ashworth, R G, Fleisher, a, Stevanin, G, Brice, a, Potter, N T, Ross, C a and Margolis, R L (2001), A Repeat Expansion in the Gene Encoding Junctophilin-3 Is Associated with Huntington Disease-Like 2. *Nature genetics*, 29(4): 377–378.

Hong, Kun, Li, Yi, Duan, Weisong, Guo, Yansu, Jiang, Hong, Li, Wenju and Li, Chunyan (2012), Full-Length TDP-43 and Its C-Terminal Fragments Activate Mitophagy in NSC34 Cell Line. *Neuroscience letters*, 530(2): 144–149.

Hsin, Jing-Ping and Manley, James L (2012), The RNA Polymerase II CTD Coordinates Transcription and RNA Processing. *Genes \& development*, 26(19): 2119–2137.

Huang, Cao, Huang, Bo, Bi, Fangfang, Yan, Linda H, Tong, Jianbin, Huang, Jufang, Xia, Xu Gang and Zhou, Hongxia (2014), Profiling the Genes Affected by Pathogenic TDP-43 in Astrocytes. *Journal of neurochemistry*, 129(6): 932–939.

Huang, Cao, Tong, Jianbin, Bi, Fangfang, Zhou, Hongxia and Xia, Xu Gang (2012), Mutant TDP-43 in Motor Neurons Promotes the Onset and Progression of ALS in Rats. *The Journal of clinical investigation*, 122(1): 107–118.

Huang, Chi-Chen, Bose, Jayarama Krishnan, Majumder, Pritha, Lee, Kuen-Haur, Huang, Jen-Tse Joseph, Huang, Jeffrey K and Shen, Che-Kun James (2014), Metabolism and Mis-Metabolism of the Neuropathological Signature Protein TDP-43. *Journal of cell science*, 127(14): 3024–3038.

Huang, Da Wei, Sherman, Brad T and Lempicki, Richard A (2008), Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. 4(1): 44–57.

Huang, Da Wei, Sherman, Brad T and Lempicki, Richard A (2009), Bioinformatics Enrichment Tools: Paths Toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic acids research*, 37(1): 1–

13.

Igaz, Lionel M, Kwong, Linda K, Chen-Plotkin, Alice, Winton, Matthew J, Unger, Travis L, Xu, Yan, Neumann, Manuela, Trojanowski, John Q and Lee, Virginia M-Y (2009), Expression of TDP-43 C-Terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies. *Journal of Biological Chemistry*, 284(13): 8516–8524.

Iguchi, Yohei, Katsuno, Masahisa, Niwa, Jun-ichi, Yamada, Shin-ichi, Sone, Jun, Waza, Masahiro, Adachi, Hiroaki, Tanaka, Fumiaki, Nagata, Koh-ichi, Arimura, Nariko, Watanabe, Takashi, Kaibuchi, Kozo and Sobue, Gen (2009), TDP-43 Depletion Induces Neuronal Cell Damage Through Dysregulation of Rho Family GTPases. *Journal of Biological Chemistry*, 284(33): 22059–22066.

Ince, P G and Codd, G a (2005), Return of the Cycad Hypothesis - Does the Amyotrophic Lateral Sclerosis/Parkinsonism Dementia Complex (ALS/PDC) of Guam Have New Implications for Global Health? *Neuropathology and applied neurobiology*, 31(4): 345–353.

Ince, Paul G, Highley, J Robin, Kirby, Janine, Wharton, Stephen B, Takahashi, Hitoshi, Strong, Michael J and Shaw, Pamela J (2011), Molecular Pathology and Genetic Advances in Amyotrophic Lateral Sclerosis: an Emerging Molecular Pathway and the Significance of Glial Pathology. *Acta neuropathologica*, 122(6): 657–671.

Ingolia, Nicholas T (2014), Ribosome Profiling: New Views of Translation, From Single Codons to Genome Scale. *Nature reviews. Genetics*, 15(3): 205–213.

Ingolia, Nicholas T, Brar, Gloria A, Stern-Ginossar, Noam, Harris, Michael S, Talhouarne, Gaeele J S, Jackson, Sarah E, Wills, Mark R and Weissman, Jonathan S (2014), Ribosome Profiling Reveals Pervasive Translation Outside of Annotated Protein-Coding Genes. *Cell reports*, 8(5): 1365–1379.

Ingolia, Nicholas T, Ghaemmaghami, Sina, Newman, John R S and Weissman, Jonathan S (2009), Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science (New York, N.Y.)*, 324(5924): 218–223.

Ingolia, Nicholas T, Lareau, Liana F and Weissman, Jonathan S (2011), Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell*, 147(4): 789–802.

Ishigami, Akihito, Ohsawa, Takako, Hiratsuka, Masaharu, Taguchi, Hiromi, Kobayashi, Saori, Saito, Yuko, Murayama, Shigeo, Asaga, Hiroaki, Toda, Tosifusa, Kimura, Narimichi and Maruyama, Naoki (2005), Abnormal Accumulation of Citrullinated Proteins Catalyzed by Peptidylarginine Deiminase in Hippocampal Extracts From Patients with Alzheimer's Disease. *Journal of neuroscience research*, 80(1): 120–128.

Jackson, Kasey L, Dayton, Robert D, Fisher-Perkins, Jeanne M, Didier, Peter J, Baker, Kate C, Weimer, Maria, Gutierrez, Amparo, Cain, Cooper D, Mathis, J Michael, Gitcho, Michael A, Bunnell, Bruce A and Klein, Ronald L (2015), Initial Gene Vector Dosing for Studying Symptomatology of Amyotrophic

Lateral Sclerosis in Non-Human Primates. *Journal of medical primatology*, 44(2): 66–75.

Jang, Byungki, Jin, Jae-Kwang, Jeon, Yong-Chul, Cho, Han Jeong, Ishigami, Akihito, Choi, Kyung-Chan, Carp, Richard I, Maruyama, Naoki, Kim, Yong-Sun and Choi, Eun-Kyoung (2010), Involvement of Peptidylarginine Deiminase-Mediated Post-Translational Citrullination in Pathogenesis of Sporadic Creutzfeldt-Jakob Disease. *Acta neuropathologica*, 119(2): 199–210.

Jang, Byungki, Kim, Eunah, Choi, Jin-Kyu, Jin, Jae-Kwang, Kim, Jae-Il, Ishigami, Akihito, Maruyama, Naoki, Carp, Richard I, Kim, Yong-Sun and Choi, Eun-Kyoung (2008), Accumulation of Citrullinated Proteins by Up-Regulated Peptidylarginine Deiminase 2 in Brains of Scrapie-Infected Mice: a Possible Role in Pathogenesis. *American Journal of Pathology*, 173(4): 1129–1142.

Johannes, G, Carter, M S, Eisen, M B, Brown, P O and Sarnow, P (1999), Identification of Eukaryotic mRNAs That Are Translated at Reduced Cap Binding Complex eIF4F Concentrations Using a cDNA Microarray. *PNAS*, 96(23): 13118–13123.

Johnson, Janel O, Mandrioli, Jessica, Benatar, Michael, Abramzon, Yevgeniya, Van Deerlin, Vivianna M, Trojanowski, John Q, Gibbs, J Raphael, Brunetti, Maura, Gronka, Susan, Wuu, Joanne, Ding, Jinhui, McCluskey, Leo, Martinez-Lage, Maria, Falcone, Dana, Hernandez, Dena G, Arepalli, Sampath, Chong, Sean, Schymick, Jennifer C, Rothstein, Jeffrey, Landi, Francesco, Wang, Yong-Dong, Calvo, Andrea, Mora, Gabriele, Sabatelli, Mario, Monsurrò, Maria Rosaria, Battistini, Stefania, Salvi, Fabrizio, Spataro, Rossella, Sola, Patrizia, Borghero, Giuseppe, ITALSGEN Consortium, Galassi, Giuliana, Scholz, Sonja W, Taylor, J Paul, Restagno, Gabriella, Chiò, Adriano and Traynor, Bryan J (2010), Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron*, 68(5): 857–864.

Johnson, Janel O, Pioro, Erik P, Boehringer, Ashley, Chia, Ruth, Feit, Howard, Renton, Alan E, Pliner, Hannah A, Abramzon, Yevgeniya, Marangi, Giuseppe, Winborn, Brett J, Gibbs, J Raphael, Nalls, Michael a, Morgan, Sarah, Shoai, Maryam, Hardy, John, Pittman, Alan, Orrell, Richard W, Malaspina, Andrea, Sidle, Katie C, Fratta, Pietro, Harms, Matthew B, Baloh, Robert H, Pestronk, Alan, Weihl, Conrad C, Rogaeva, Ekaterina, Zinman, Lorne, Drory, Vivian E, Borghero, Giuseppe, Mora, Gabriele, Calvo, Andrea, Rothstein, Jeffrey D, Drepper, Carsten, Sendtner, Michael, Singleton, Andrew B, Taylor, J Paul, Cookson, Mark R, Restagno, Gabriella, Sabatelli, Mario, Bowser, Robert, Chiò, Adriano and Traynor, Bryan J (2014), Mutations in the Matrin 3 Gene Cause Familial Amyotrophic Lateral Sclerosis. *Nature neuroscience*.

Jordanova, Albena, Irobi, Joy, Thomas, Florian P, Van Dijck, Patrick, Meerschaert, Kris, Dewil, Maarten, Dierick, Ines, Jacobs, An, De Vriendt, Els, Guergueltcheva, Velina, Rao, Chitharanjan V, Tournev, Ivailo, Gondim, Francisco a a, D'Hooghe, Marc, Van Gerwen, Veerle, Callaerts, Patrick, Van Den Bosch, Ludo, Timmermans, Jean-Pièrre, Robberecht, Wim, Gettemans, Jan, Thevelein, Johan M, De Jonghe, Peter, Kremensky, Ivo and Timmerman, Vincent (2006), Disrupted Function and Axonal Distribution of Mutant Tyrosyl-tRNA Synthetase in Dominant Intermediate Charcot-Marie-

Tooth Neuropathy. *Nature genetics*, 38(2): 197–202.

Jovičić, Ana, Mertens, Jerome, Boeynaems, Steven, Bogaert, Elke, Chai, Noori, Yamada, Shizuka B, Paul, Joseph W, Sun, Shuying, Herdy, Joseph R, Bieri, Gregor, Kramer, Nicholas J, Gage, Fred H, Van Den Bosch, Ludo, Robberecht, Wim and Gitler, Aaron D (2015), Modifiers of C9orf72 Dipeptide Repeat Toxicity Connect Nucleocytoplasmic Transport Defects to FTD/ALS. *Nature neuroscience*, 18(9): 1226–1229.

Joyce, Peter I, Fratta, Pietro, Fisher, Elizabeth M C and Acevedo-Arozena, Abraham (2011), SOD1 and TDP-43 Animal Models of Amyotrophic Lateral Sclerosis: Recent Advances in Understanding Disease Toward the Development of Clinical Treatments. *Mammalian genome : official journal of the International Mammalian Genome Society*, 22(7-8): 420–448.

Kabashi, Edor, Bercier, Valérie, Lissouba, Alexandra, Liao, Meijiang, Brustein, Edna, Rouleau, Guy a and Drapeau, Pierre (2011), FUS and TARDBP but Not SOD1 Interact in Genetic Models of Amyotrophic Lateral Sclerosis. *PLoS genetics*, 7(8): e1002214.

Kabashi, Edor, Brustein, Edna, Champagne, Nathalie and Drapeau, Pierre (2011), Zebrafish Models for the Functional Genomics of Neurogenetic Disorders. *Biochimica et biophysica acta*, 1812(3): 335–345.

Kabashi, Edor, Lin, Li, Tradewell, Miranda L, Dion, Patrick a, Bercier, Valérie, Bourgouin, Patrick, Rochefort, Daniel, Bel Hadj, Samar, Durham, Heather D, Vande Velde, Christine, Rouleau, Guy a and Drapeau, Pierre (2010), Gain and Loss of Function of ALS-Related Mutations of TARDBP (TDP-43) Cause Motor Deficits in Vivo. *Human molecular genetics*, 19(4): 671–683.

Kabashi, Edor, Valdmanis, Paul N, Dion, Patrick, Spiegelman, Dan, McConkey, Brendan J, Vande Velde, Christine, Bouchard, Jean-Pierre, Lacomblez, Lucette, Pochigaeva, Ksenia, Salachas, Francois, Pradat, Pierre-Francois, Camu, William, Meininger, Vincent, Dupre, Nicolas and Rouleau, Guy a (2008), TARDBP Mutations in Individuals with Sporadic and Familial Amyotrophic Lateral Sclerosis. *Nature genetics*, 40(5): 572–574.

Kanai, Yoshimitsu, Dohmae, Naoshi and Hirokawa, Nobutaka (2004), Kinesin Transports RNA: Isolation and Characterization of an RNA-Transporting Granule. *Neuron*, 43(4): 513–525.

Kanjilal, Baishali, Keyser, Brian M, Andres, Devon K, Nealley, Eric, Benton, Betty, Melber, Ashley A, Andres, Jaclynn F, Letukas, Valerie A, Clark, Offie and Ray, Radharaman (2014), Differentiated NSC-34 Cells as an in Vitro Cell Model for VX. *Toxicology mechanisms and methods*, 24(7): 488–494.

Kapeli, Katannya and Yeo, Gene W (2012), Genome-Wide Approaches to Dissect the Roles of RNA Binding Proteins in Translational Control: Implications for Neurological Diseases. *Frontiers in neuroscience*, 6: 144.

Karni, Rotem, de Stanchina, Elisa, Lowe, Scott W, Sinha, Rahul, Mu, David and Krainer, Adrian R (2007), The Gene Encoding the Splicing Factor SF2/ASF Is a Proto-Oncogene. *Nature structural \& molecular biology*, 14(3): 185–193.

Kasprzyk, Arek (2011), BioMart: Driving a Paradigm Change in Biological Data Management. *Database : the journal of biological databases and curation*, 2011: bar049.

Kato, T, Ito, J and Tanaka, R (1987), Functional Dissociation of Dual Activities of Glia Maturation Factor: Inhibition of Glial Proliferation and Preservation of Differentiation by Glial Growth Inhibitory Factor. *Brain research*, 430(1): 153–156.

Kawahara, Yukio and Mieda-Sato, Ai (2012), TDP-43 Promotes microRNA Biogenesis as a Component of the Drosha and Dicer Complexes. *PNAS*, 109(9): 3347–3352.

Keller, Brian a, Volkening, Kathryn, Droppelmann, Cristian a, Ang, Lee Cyn, Rademakers, Rosa and Strong, Michael J (2012), Co-Aggregation of RNA Binding Proteins in ALS Spinal Motor Neurons: Evidence of a Common Pathogenic Mechanism. *Acta neuropathologica*, 124(5): 733–747.

Kelly, Kevin F, Otchere, Abena A, Graham, Monica and Daniel, Juliet M (2004), Nuclear Import of the BTB/POZ Transcriptional Regulator Kaiso. *Journal of cell science*, 117(Pt 25): 6143–6152.

Kervestin, Stephanie and Jacobson, Allan (2012), NMD: a Multifaceted Response to Premature Translational Termination. *Nature reviews. Molecular cell biology*, 13(11): 700–712.

Kiebler, Michael A and Bassell, Gary J (2006), Neuronal RNA Granules: Movers and Makers. *Neuron*, 51(6): 685–690.

Kiernan, Matthew C, Vucic, Steve, Cheah, Benjamin C, Turner, Martin R, Eisen, Andrew, Hardiman, Orla, Burrell, James R and Zoing, Margaret C (2011), Amyotrophic Lateral Sclerosis. *Lancet*, 377(9769): 942–955.

Kim, Hong Joo, Kim, Nam Chul, Wang, Yong-Dong, Scarborough, Emily A, Moore, Jennifer, Diaz, Zamia, MacLea, Kyle S, Freibaum, Brian, Li, Songqing, Molliex, Amandine, Kanagaraj, Anderson P, Carter, Robert, Boylan, Kevin B, Wojtas, Aleksandra M, Rademakers, Rosa, Pinkus, Jack L, Greenberg, Steven A, Trojanowski, John Q, Traynor, Bryan J, Smith, Bradley N, Topp, Simon, Gkazi, Athina-Soragia, Miller, Jack, Shaw, Christopher E, Kottlors, Michael, Kirschner, Janbernd, Pestronk, Alan, Li, Yun R, Ford, Alice Flynn, Gitler, Aaron D, Benatar, Michael, King, Oliver D, Kimonis, Virginia E, Ross, Eric D, Weihl, Conrad C, Shorter, James and Taylor, J Paul (2013), Mutations in Prion-Like Domains in hnRNPA2B1 and hnRNPA1 Cause Multisystem Proteinopathy and ALS. *Nature*, 495(7442): 467–473.

Kim, Hyung-Jun, Raphael, Alya R, LaDow, Eva S, McGurk, Leeanne, Weber, Ross A, Trojanowski, John Q, Lee, Virginia M-Y, Finkbeiner, Steven, Gitler, Aaron D and Bonini, Nancy M (2014), Therapeutic Modulation of eIF2α Phosphorylation Rescues TDP-43 Toxicity in Amyotrophic Lateral Sclerosis Disease Models. *Nature genetics*, 46(2): 152–160.

Kim, Janghwan, Efe, Jem A, Zhu, Saiyong, Talantova, Maria, Yuan, Xu, Wang, Shufen, Lipton, Stuart A, Zhang, Kang and Ding, Sheng (2011), Direct Reprogramming of Mouse Fibroblasts to Neural Progenitors. *PNAS*,

108(19): 7838–7843.

King, Andrew, Maekawa, Satomi, Bodi, Istvan, Troakes, Claire and Al-Sarraj, Safa (2011), Ubiquitinated, P62 Immunopositive Cerebellar Cortical Neuronal Inclusions Are Evident Across the Spectrum of TDP-43 Proteinopathies but Are Only Rarely Additionally Immunopositive for Phosphorylation-Dependent TDP-43. *Neuropathology : official journal of the Japanese Society of Neuropathology*, 31(3): 239–249.

King, Isabelle N, Yartseva, Valeria, Salas, Donaldo, Kumar, Abhishek, Heidersbach, Amy, Ando, D Michael, Stallings, Nancy R, Elliott, Jeffrey L, Srivastava, Deepak and Ivey, Kathryn N (2014), The RNA-Binding Protein TDP-43 Selectively Disrupts MicroRNA-1/206 Incorporation Into the RNA-Induced Silencing Complex. *Journal of Biological Chemistry*, 289(20): 14263–14271.

Kirby, Janine, Goodall, Emily F, Smith, William, Highley, J Robin, Masanzu, Rudo, Hartley, Judith a, Hibberd, Rachel, Hollinger, Hannah C, Wharton, Stephen B, Morrison, Karen E, Ince, Paul G, McDermott, Christopher J and Shaw, Pamela J (2010), Broad Clinical Phenotypes Associated with TAR-DNA Binding Protein (TARDBP) Mutations in Amyotrophic Lateral Sclerosis. *Neurogenetics*, 11(2): 217–225.

Klose, Robert J and Bird, Adrian P (2006), Genomic DNA Methylation: the Mark and Its Mediators. *Trends in biochemical sciences*, 31(2): 89–97.

Komatsu, Masaaki, Kageyama, Shun and Ichimura, Yoshinobu (2012), P62/SQSTM1/A170: Physiology and Pathology. *Pharmacological research*, 66(6): 457–462.

Kong, Qiongman and Lin, Chien-Liang Glenn (2010), Oxidative Damage to RNA: Mechanisms, Consequences, and Diseases. *Cellular and molecular life sciences : CMLS*, 67(11): 1817–1829.

Koob, M D, Moseley, M L, Schut, L J, Benzow, K a, Bird, T D, Day, J W and Ranum, L P (1999), An Untranslated CTG Expansion Causes a Novel Form of Spinocerebellar Ataxia (SCA8). *Nature genetics*, 21(4): 379–384.

Kozak, M (1978), How Do Eucaryotic Ribosomes Select Initiation Regions in Messenger RNA? *Cell*, 15(4): 1109–1123.

Kozak, M (1987), An Analysis of 5'-Noncoding Sequences From 699 Vertebrate Messenger-RNAs. *Nucleic acids research*, 15(20): 8125–8148.

Köhler, Alwin and Hurt, Ed (2007), Exporting RNA From the Nucleus to the Cytoplasm. *Nature reviews. Molecular cell biology*, 8(10): 761–773.

Kraemer, Brian C, Schuck, Theresa, Wheeler, Jeanna M, Robinson, Linda C, Trojanowski, John Q, Lee, Virginia M-Y and Schellenberg, Gerard D (2010), Loss of Murine TDP-43 Disrupts Motor Function and Plays an Essential Role in Embryogenesis. *Acta neuropathologica*, 119(4): 409–419.

Kratz, Anton, Beguin, Pascal, Kaneko, Megumi, Chimura, Takahiko, Suzuki, Ana Maria, Matsunaga, Atsuko, Kato, Sachi, Bertin, Nicolas, Lassmann, Timo, Vigot, Réjan, Carninci, Piero, Plessy, Charles and Launey, Thomas

(2014), Digital Expression Profiling of the Compartmentalized Translatome of Purkinje Neurons. *Genome research*, 24(8): 1396–1410.

Kulicke, Ruth, Fenster, Robert J, Greengard, Paul, Heintz, Nathaniel and Heiman, Myriam (2014), Cell Type-Specific mRNA Purification by Translating Ribosome Affinity Purification (TRAP). *Nature Protocols*, 9(6): 1282–1291.

Kuusisto, E, Suuronen, T and Salminen, A (2001), Ubiquitin-Binding Protein P62 Expression Is Induced During Apoptosis and Proteasomal Inhibition in Neuronal Cells. *Biochemical and Biophysical Research Communications*, 280(1): 223–228.

Kwiatkowski, T J, Bosco, D A, Leclerc, A L, Tamrazian, E, Vanderburg, C R, Russ, C, Davis, A, Gilchrist, J, Kasarskis, E J, Munsat, T, Valdmanis, P, Rouleau, G a, Hosler, B A, Cortelli, P, de Jong, P J, Yoshinaga, Y, Haines, J L, Pericak-Vance, M A, Yan, J, Ticozzi, N, Siddique, T, McKenna-Yasek, D, Sapp, P C, Horvitz, H R, Landers, J E and Brown, R H (2009), Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science (New York, N.Y.)*, 323(5918): 1205–1208.

Kwong, Linda K, Neumann, Manuela, Sampathu, Deepak M, Lee, Virginia M-Y and Trojanowski, John Q (2007), TDP-43 Proteinopathy: the Neuropathology Underlying Major Forms of Sporadic and Familial Frontotemporal Lobar Degeneration and Motor Neuron Disease. *Acta neuropathologica*, 114(1): 63–70.

Lagier-Tourenne, C, Baughn, M and Rigo, F (2013), Targeted Degradation of Sense and Antisense C9orf72 RNA Foci as Therapy for ALS and Frontotemporal Degeneration,

Lagier-Tourenne, Clotilde, Polymenidou, Magdalini, Hutt, Kasey R, Vu, Anthony Q, Baughn, Michael, Huelga, Stephanie C, Clutario, Kevin M, Ling, Shuo-Chien, Liang, Tiffany Y, Mazur, Curt, Wancewicz, Edward, Kim, Aneeza S, Watt, Andy, Freier, Sue, Hicks, Geoffrey G, Donohue, John Paul, Shiue, Lily, Bennett, C Frank, Ravits, John, Cleveland, Don W and Yeo, Gene W (2012), Divergent Roles of ALS-Linked Proteins FUS/TLS and TDP-43 Intersect in Processing Long Pre-mRNAs. *Nature neuroscience*, 15(11): 1488–1497.

Lalmansingh, Avin S, Urekar, Craig J and Reddi, Prabhakara P (2011), TDP-43 Is a Transcriptional Repressor: the Testis-Specific Mouse Acrv1 Gene Is a TDP-43 Target in Vivo. *The Journal of biological chemistry*, 286(13): 10970–10982.

Lambrechts, Diether, Storkebaum, Erik, Morimoto, Masafumi, Del-Favero, Jurgen, Desmet, Frederik, Marklund, Stefan L, Wyns, Sabine, Thijs, Vincent, Andersson, Jörgen, van Marion, Ingrid, Al-Chalabi, Ammar, Bornes, Stephanie, Musson, Rhiannon, Hansen, Valerie, Beckman, Lars, Adolfsson, Rolf, Pall, Hardev Singh, Prats, Hervé, Vermeire, Severine, Rutgeerts, Paul, Katayama, Shigehiro, Awata, Takuya, Leigh, Nigel, Lang-Lazdunski, Loïc, Dewerchin, Mieke, Shaw, Christopher, Moons, Lieve, Vlietinck, Robert, Morrison, Karen E, Robberecht, Wim, Van Broeckhoven, Christine, Collen, Désiré, Andersen, Peter M and Carmeliet, Peter (2003), VEGF Is a Modifier of Amyotrophic Lateral Sclerosis in Mice and Humans and Protects

Motoneurons Against Ischemic Death. *Nature genetics*, 34(4): 383–394.

Langmead, Ben, Trapnell, Cole, Pop, Mihai and Salzberg, Steven L (2009), Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome biology*, 10(3): R25.

Lauranzano, Eliana, Pozzi, Silvia, Pasetto, Laura, Stucchi, Riccardo, Massignan, Tania, Paolella, Katia, Mombrini, Melissa, Nardo, Giovanni, Lunetta, Christian, Corbo, Massimo, Mora, Gabriele, Bendotti, Caterina and Bonetto, Valentina (2015), Peptidylprolyl Isomerase a Governs TARDBP Function and Assembly in Heterogeneous Nuclear Ribonucleoprotein Complexes. *Brain : a journal of neurology*, 138(Pt 4): 974–991.

Le Quesne, John P C, Spriggs, Keith a, Bushell, Martin and Willis, Anne E (2010), Dysregulation of Protein Synthesis and Disease. *The Journal of pathology*, 220(2): 140–151.

Lebreton, Alice and Séraphin, Bertrand (2008), Exosome-Mediated Quality Control: Substrate Recruitment and Molecular Activity. *Biochimica et biophysica acta*, 1779(9): 558–565.

Lee, Sooncheol, Liu, Botao, Lee, Soohyun, Huang, Sheng-Xiong, Shen, Ben and Qian, Shu-Bing (2012a), Global Mapping of Translation Initiation Sites in Mammalian Cells at Single-Nucleotide Resolution. *PNAS*, 109(37): E2424–E2432.

Lee, Sooncheol, Liu, Botao, Lee, Soohyun, Huang, Sheng-Xiong, Shen, Ben and Qian, Shu-Bing (2012b), Global Mapping of Translation Initiation Sites in Mammalian Cells at Single-Nucleotide Resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37): E2424–32.

Lee, Youn-Bok, Chen, Han-Jou, Peres, João N, Gomez-Deza, Jorge, Attig, Jan, Štalekar, Maja, Troakes, Claire, Nishimura, Agnes L, Scotter, Emma L, Vance, Caroline, Adachi, Yoshitsugu, Sardone, Valentina, Miller, Jack W, Smith, Bradley N, Gallo, Jean-Marc, Ule, Jernej, Hirth, Frank, Rogelj, Boris, Houart, Corinne and Shaw, Christopher E (2013), Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *Cell reports*, 5(5): 1178–1186.

Lefebvre, S, Bürglen, L, Reboullet, S, Clermont, O, Burlet, P, Viollet, L, Benichou, B, Cruaud, C, Millasseau, P and Zeviani, M (1995), Identification and Characterization of a Spinal Muscular Atrophy-Determining Gene. *Cell*, 80(1): 155–165.

Lewis, Coral-Ann, Manning, John, Rossi, Fabio and Krieger, Charles (2012), The Neuroinflammatory Response in ALS: the Roles of Microglia and T Cells. *Neurology research international*, 2012: 803701.

Li, Zhongwei, Wu, Jinhua and Deleo, Christopher J (2006), RNA Damage and Surveillance Under Oxidative Stress. *IUBMB life*, 58(10): 581–588.

Liachko, Nicole F, Guthrie, Chris R and Kraemer, Brian C (2010), Phosphorylation Promotes Neurotoxicity in a Caenorhabditis Elegans Model of TDP-43 Proteinopathy. *The Journal of neuroscience : the official journal*

of the Society for Neuroscience, 30(48): 16208–16219.

Lim, R, Miller, J F and Zaheer, A (1989), Purification and Characterization of Glia Maturation Factor Beta: a Growth Regulator for Neurons and Glia. *PNAS*, 86(10): 3901–3905.

Ling, Shuo-Chien, Albuquerque, Claudio P, Han, Joo Seok, Lagier-Tourenne, Clotilde, Tokunaga, Seiya, Zhou, Huilin and Cleveland, Don W (2010), ALS-Associated Mutations in TDP-43 Increase Its Stability and Promote TDP-43 Complexes with FUS/TLS. *PNAS*, 107(30): 13318–13323.

Ling, Shuo-Chien, Polymenidou, Magdalini and Cleveland, Don W (2013), Converging Mechanisms in ALS and FTD: Disrupted RNA and Protein Homeostasis. *Neuron*, 79(3): 416–438.

Liquori, C L, Ricker, K, Moseley, M L, Jacobsen, J F, Kress, W, Naylor, S L, Day, J W and Ranum, L P (2001), Myotonic Dystrophy Type 2 Caused by a CCTG Expansion in Intron 1 of ZNF9. *Science*, 293(5531): 864–867.

Liu, Botao and Qian, Shu-Bing (2014), Translational Reprogramming in Cellular Stress Response. *Wiley interdisciplinary reviews. RNA*, 5(3): 301–305.

Liu-Yesucevitz, Liqun, Lin, Amy Y, Ebata, Atsushi, Boon, Joon Y, Reid, Whitney, Xu, Ya-Fei, Kobrin, Kendra, Murphy, George J, Petrucelli, Leonard and Wolozin, Benjamin (2014), ALS-Linked Mutations Enlarge TDP-43-Enriched Neuronal RNA Granules in the Dendritic Arbor. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 34(12): 4167–4174.

Lopez, a J (1998), Alternative Splicing of Pre-mRNA: Developmental Consequences and Mechanisms of Regulation. *Annual Review of Genetics*, 32: 279–305.

Lu, Liang, Zheng, Lei, Si, Ying, Luo, Wenyi, Dujardin, Gwendal, Kwan, Thaddaeus, Potochick, Nicholas R, Thompson, Sunnie R, Schneider, David A and King, Peter H (2014), Hu Antigen R (HuR) Is a Positive Regulator of the RNA-Binding Proteins TDP-43 and FUS/TLS: Implications for Amyotrophic Lateral Sclerosis. *Journal of Biological Chemistry*, 289(46): 31792–31804.

Lui, Jennifer, Castelli, Lydia M, Pizzinga, Mariavittoria, Simpson, Clare E, Hoyle, Nathaniel P, Bailey, Kathryn L, Campbell, Susan G and Ashe, Mark P (2014), Granules Harboring Translationally Active mRNAs Provide a Platform for P-Body Formation Following Stress. *Cell reports*, 9(3): 944–954.

Luna, Rosa, Gaillard, Hélène, González-Aguilera, Cristina and Aguilera, Andrés (2008), Biogenesis of mRNPs: Integrating Different Processes in the Eukaryotic Nucleus. *Chromosoma*, 117(4): 319–331.

Luty, Agnes A, Kwok, John B J, Dobson-Stone, Carol, Loy, Clement T, Coupland, Kirsten G, Karlström, Helena, Sobow, Tomasz, Tchorzewska, Joanna, Maruszak, Aleksandra, Barcikowska, Maria, Panegyres, Peter K, Zekanowski, Cezary, Brooks, William S, Williams, Kelly L, Blair, Ian P, Mather, Karen A, Sachdev, Perminder S, Halliday, Glenda M and Schofield, Peter R (2010), Sigma Nonopioid Intracellular Receptor 1 Mutations Cause

Frontotemporal Lobar Degeneration-Motor Neuron Disease. *Annals of neurology*, 68(5): 639–649.

Mackenzie, Ian R a, Bigio, Eileen H, Ince, Paul G, Geser, Felix, Neumann, Manuela, Cairns, Nigel J, Kwong, Linda K, Forman, Mark S, Ravits, John, Stewart, Heather, Eisen, Andrew, McClusky, Leo, Kretzschmar, Hans a, Monoranu, Camelia M, Highley, J Robin, Kirby, Janine, Siddique, Teepu, Shaw, Pamela J, Lee, Virginia M-Y and Trojanowski, John Q (2007), Pathological TDP-43 Distinguishes Sporadic Amyotrophic Lateral Sclerosis From Amyotrophic Lateral Sclerosis with SOD1 Mutations. *Annals of neurology*, 61(5): 427–434.

MacNair, Laura, Xiao, Shangxi, Miletic, Denise, Ghani, Mahdi, Julien, Jean-Pierre, Keith, Julia, Zinman, Lorne, Rogaeva, Ekaterina and Robertson, Janice (2015), MTHFSD and DDX58 Are Novel RNA-Binding Proteins Abnormally Regulated in Amyotrophic Lateral Sclerosis. *Brain : a journal of neurology*.

Maekawa, Satomi, Leigh, P Nigel, King, Andrew, Jones, Edith, Steele, John C, Bodi, Istvan, Shaw, Christopher E, Hortobagyi, Tibor and Al-Sarraj, Safa (2009), TDP-43 Is Consistently Co-Localized with Ubiquitinated Inclusions in Sporadic and Guam Amyotrophic Lateral Sclerosis but Not in Familial Amyotrophic Lateral Sclerosis with and Without SOD1 Mutations. *Neuropathology : official journal of the Japanese Society of Neuropathology*, 29(6): 672–683.

Magrané, Jordi, Cortez, Czrina, Gan, Wen-Biao and Manfredi, Giovanni (2014), Abnormal Mitochondrial Transport and Morphology Are Common Pathological Denominators in SOD1 and TDP43 ALS Mouse Models. *Human molecular genetics*, 23(6): 1413–1424.

Maier, Oliver, Böhm, Julia, Dahm, Michael, Brück, Stefan, Beyer, Cordian and Johann, Sonja (2013), Differentiated NSC-34 Motoneuron-Like Cells as Experimental Model for Cholinergic Neurodegeneration. *Neurochemistry international*, 62(8): 1029–1038.

Majounie, Elisa, Renton, Alan E, Mok, Kin, Dopper, Elise G P, Waite, Adrian, Rollinson, Sara, Chiò, Adriano, Restagno, Gabriella, Nicolaou, Nayia, Simon-Sanchez, Javier, van Swieten, John C, Abramzon, Yevgeniya, Johnson, Janel O, Sendtner, Michael, Pamphlett, Roger, Orrell, Richard W, Mead, Simon, Sidle, Katie C, Houlden, Henry, Rohrer, Jonathan D, Morrison, Karen E, Pall, Hardev, Talbot, Kevin, Ansorge, Olaf, Chromosome 9-ALS/FTD Consortium, French research network on FTLD/FTLD/ALS, ITALSGEN Consortium, Hernandez, Dena G, Arepalli, Sampath, Sabatelli, Mario, Mora, Gabriele, Corbo, Massimo, Giannini, Fabio, Calvo, Andrea, Englund, Elisabet, Borghero, Giuseppe, Floris, Gian Luca, Remes, Anne M, Laaksovirta, Hannu, McCluskey, Leo, Trojanowski, John Q, Van Deerlin, Vivianna M, Schellenberg, Gerard D, Nalls, Michael a, Drory, Vivian E, Lu, Chin-Song, Yeh, Tu-Hsueh, Ishiura, Hiroyuki, Takahashi, Yuji, Tsuji, Shoji, Le Ber, Isabelle, Brice, Alexis, Drepper, Carsten, Williams, Nigel, Kirby, Janine, Shaw, Pamela, Hardy, John, Tienari, Pentti J, Heutink, Peter, Morris, Huw R, Pickering-Brown, Stuart and Traynor, Bryan J (2012), Frequency of the C9orf72 Hexanucleotide Repeat Expansion in Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia: a Cross-Sectional Study. *Lancet neurology*, 11(4): 323–330.

Maniatis, Tom and Reed, Robin (2002), An Extensive Network of Coupling Among Gene Expression Machines. *Nature*, 416(6880): 499–506.

Mann, David M A and Snowden, Julie S (2017), Frontotemporal Lobar Degeneration: Pathogenesis, Pathology and Pathways to Phenotype. *Brain Pathology*.

Mann, David M A, Rollinson, Sara, Robinson, Andrew, Bennion Callister, Janis, Thompson, Jennifer C, Snowden, Julie S, Gendron, Tania, Petrucelli, Leonard, Masuda-Suzukake, Masami, Hasegawa, Masato, Davidson, Yvonne and Pickering-Brown, Stuart (2013), Dipeptide Repeat Proteins Are Present in the P62 Positive Inclusions in Patients with Frontotemporal Lobar Degeneration and Motor Neurone Disease Associated with Expansions in C9ORF72. *Acta neuropathologica communications*, 1: 68.

Mannen, T, Iwata, M, Toyokura, Y and Nagashima, K (1977), Preservation of a Certain Motoneurone Group of the Sacral Cord in Amyotrophic Lateral Sclerosis: Its Clinical Significance. *Journal of neurology, neurosurgery, and psychiatry*, 40(5): 464–469.

Martin, Marcel (2011), Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnet.journal*, 17(1): pp. 10–12–3.

Maruyama, Hirofumi, Morino, Hiroyuki, Ito, Hidefumi, Izumi, Yuishin, Kato, Hidemasa, Watanabe, Yasuhito, Kinoshita, Yoshimi, Kamada, Masaki, Nodera, Hiroyuki, Suzuki, Hidenori, Komure, Osamu, Matsuura, Shinya, Kobatake, Keitaro, Morimoto, Nobutoshi, Abe, Koji, Suzuki, Naoki, Aoki, Masashi, Kawata, Akihiro, Hirai, Takeshi, Kato, Takeo, Ogasawara, Kazumasa, Hirano, Asao, Takumi, Toru, Kusaka, Hirofumi, Hagiwara, Koichi, Kaji, Ryuji and Kawakami, Hideshi (2010), Mutations of Optineurin in Amyotrophic Lateral Sclerosis. *Nature*, 465(7295): 223–226.

Massman, P J PJ, Sims, J J, Cooke, N N, Haverkamp, L J LJ, Appel, V V and Appel, S H SH (1996), Prevalence and Correlates of Neuropsychological Deficits in Amyotrophic Lateral Sclerosis. *Journal of neurology, neurosurgery, and psychiatry*, 61(5): 450–455.

McGoldrick, Philip, Joyce, Peter I, Fisher, Elizabeth M C and Greensmith, Linda (2013), Rodent Models of Amyotrophic Lateral Sclerosis. *Biochimica et biophysica acta*, 1832(9): 1421–1436.

Mercado, Pablo Arrisi, Ayala, Youhna M, Romano, Maurizio, Buratti, Emanuele and Baralle, Francisco E (2005), Depletion of TDP 43 Overrides the Need for Exonic and Intronic Splicing Enhancers in the Human apoA-II Gene. *Nucleic acids research*, 33(18): 6000–6010.

Meyer, Kathrin, Ferraiuolo, Laura, Miranda, Carlos J, Likhite, Shibi, McElroy, Sohyun, Renusch, Samantha, Ditsworth, Dara, Lagier-Tourenne, Clotilde, Smith, Richard A, Ravits, John, Burghes, Arthur H, Shaw, Pamela J, Cleveland, Don W, Kolb, Stephen J and Kaspar, Brian K (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): 829–832.

Mi, Huaiyu, Muruganujan, Anushya, Casagrande, John T and Thomas, Paul D (2013), Large-Scale Gene Function Analysis with the PANTHER Classification System. *Nature Protocols*, 8(8): 1551–1566.

Miller, J W, Urbinati, C R, Teng-Umnuay, P, Stenberg, M G, Byrne, B J, Thornton, C a and Swanson, M S (2000), Recruitment of Human Muscleblind Proteins to (CUG)(N) Expansions Associated with Myotonic Dystrophy. *The EMBO journal*, 19(17): 4439–4448.

Mitchell, Jacqueline C, McGoldrick, Philip, Vance, Caroline, Hortobagyi, Tibor, Sreedharan, Jemeen, Rogelj, Boris, Tudor, Elizabeth L, Smith, Bradley N, Klasen, Christian, Miller, Christopher C J, Cooper, Jonathan D, Greensmith, Linda and Shaw, Christopher E (2013), Overexpression of Human Wild-Type FUS Causes Progressive Motor Neuron Degeneration in an Age- and Dose-Dependent Fashion. *Acta neuropathologica*, 125(2): 273–288.

Mizielinska, Sarah, Lashley, Tammaryn, Norona, Frances E, Clayton, Emma L, Ridler, Charlotte E, Fratta, Pietro and Isaacs, Adrian M (2013), C9orf72 Frontotemporal Lobar Degeneration Is Characterised by Frequent Neuronal Sense and Antisense RNA Foci. *Acta neuropathologica*.

Mizuno, Yuji, Amari, Masakuni, Takatama, Masamitsu, Aizawa, Hitoshi, Mihara, Ban and Okamoto, Koichi (2006a), Transferrin Localizes in Bunina Bodies in Amyotrophic Lateral Sclerosis. *Acta neuropathologica*, 112(5): 597–603.

Mizuno, Yuji, Amari, Masakuni, Takatama, Masamitsu, Aizawa, Hitoshi, Mihara, Ban and Okamoto, Koichi (2006b), Immunoreactivities of P62, an Ubiqutin-Binding Protein, in the Spinal Anterior Horn Cells of Patients with Amyotrophic Lateral Sclerosis. *Journal of the neurological sciences*, 249(1): 13–18.

Mori, Kohji, Lammich, Sven, Mackenzie, Ian R a, Forné, Ignasi, Zilow, Sonja, Kretzschmar, Hans, Edbauer, Dieter, Janssens, Jonathan, Kleinberger, Gernot, Cruts, Marc, Herms, Jochen, Neumann, Manuela, Van Broeckhoven, Christine, Arzberger, Thomas and Haass, Christian (2013), hnRNP A3 Binds to GGGGCC Repeats and Is a Constituent of P62-Positive/TDP43-Negative Inclusions in the Hippocampus of Patients with C9orf72 Mutations. *Acta neuropathologica*, 125(3): 413–423.

Mori, Kohji, Weng, Shih-Ming, Arzberger, Thomas, May, Stephanie, Rentzsch, Kristin, Kremmer, Elisabeth, Schmid, Bettina, Kretzschmar, Hans a, Cruts, Marc, Van Broeckhoven, Christine, Haass, Christian and Edbauer, Dieter (2013), The C9orf72 GGGGCC Repeat Is Translated Into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science (New York, N.Y.)*, 339(6125): 1335–1338.

Morohoshi, F, Ootsuka, Y, Arai, K, Ichikawa, H, Mitani, S, Munakata, N and Ohki, M (1998), Genomic Structure of the Human RBP56/hTAFII68 and FUS/TLS Genes. *Gene*, 221(2): 191–198.

Moujalled, Diane, James, Janine L, Yang, Shu, Zhang, Katharine, Duncan, Clare, Moujalled, Donia M, Parker, Sarah J, Caragounis, Aphrodite, Lidgerwood, Grace, Turner, Bradley J, Atkin, Julie D, Grubman, Alexandra, Liddell, Jeffrey R, Proepper, Christian, Boeckers, Tobias M, Kanninen, Katja M, Blair, Ian, Crouch, Peter J and White, Anthony R (2015), Phosphorylation

of hnRNP K by Cyclin-Dependent Kinase 2 Controls Cytosolic Accumulation of TDP-43. *Human molecular genetics*, 24(6): 1655–1669.

Mórotz, Gábor M, De Vos, Kurt J, Vagnoni, Alessio, Ackerley, Steven, Shaw, Christopher E and Miller, Christopher C J (2012), Amyotrophic Lateral Sclerosis-Associated Mutant VAPBP56S Perturbs Calcium Homeostasis to Disrupt Axonal Transport of Mitochondria. *Human molecular genetics*, 21(9): 1979–1988.

Muhle, Rebecca, Trentacoste, Stephanie V and Rapin, Isabelle (2004), The Genetics of Autism. *Pediatrics*, 113(5): e472–86.

Mutihac, R, Alegre-Abarrategui, J, Gordon, D, Farrimond, L, Yamasaki-Mann, M, Talbot, K and Wade-Martins, R (2015), TARDBP Pathogenic Mutations Increase Cytoplasmic Translocation of TDP-43 and Cause Reduction of Endoplasmic Reticulum Ca2+ Signaling in Motor Neurons. *Neurobiology of disease*, 75(C): 64–77.

Nagai, Makiko, Re, Diane B, Nagata, Tetsuya, Chalazonitis, Alcmène, Jessell, Thomas M, Wichterle, Hynek and Przedborski, Serge (2007), Astrocytes Expressing ALS-Linked Mutated SOD1 Release Factors Selectively Toxic to Motor Neurons. *Nature neuroscience*, 10(5): 615–622.

Neumann, Manuela, Sampathu, Deepak M, Kwong, Linda K, Truax, Adam C, Micsenyi, Matthew C, Chou, Thomas T, Bruce, Jennifer, Schuck, Theresa, Grossman, Murray, Clark, Christopher M, McCluskey, Leo F, Miller, Bruce L, Masliah, Eliezer, Mackenzie, Ian R, Feldman, Howard, Feiden, Wolfgang, Kretzschmar, Hans a, Trojanowski, John Q and Lee, Virginia M-Y (2006), Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science (New York, N.Y.)*, 314(5796): 130–133.

Nicholas, Anthony P, Sambandam, Thiagarajan, Echols, Joshua D and Tourtellotte, Wallace W (2004), Increased Citrullinated Glial Fibrillary Acidic Protein in Secondary Progressive Multiple Sclerosis. *The Journal of comparative neurology*, 473(1): 128–136.

Nishimoto, Yoshinori, Ito, Daisuke, Yagi, Takuya, Nihei, Yoshihiro, Tsunoda, Yoshiko and Suzuki, Norihiro (2010), Characterization of Alternative Isoforms and Inclusion Body of the TAR DNA-Binding Protein-43. *Journal of Biological Chemistry*, 285(1): 608–619.

Nishimura, Agnes L, Mitne-Neto, Miguel, Silva, Helga C a, Richieri-Costa, Antônio, Middleton, Susan, Cascio, Duilio, Kok, Fernando, Oliveira, João R M, Gillingwater, Tom, Webb, Jeanette, Skehel, Paul and Zatz, Mayana (2004), A Mutation in the Vesicle-Trafficking Protein VAPB Causes Late-Onset Spinal Muscular Atrophy and Amyotrophic Lateral Sclerosis. *American journal of human genetics*, 75(5): 822–831.

Nonaka, Takashi, Arai, Tetsuaki, Buratti, Emanuele, Baralle, Francisco E, Akiyama, Haruhiko and Hasegawa, Masato (2009), Phosphorylated and Ubiquitinated TDP-43 Pathological Inclusions in ALS and FTLD-U Are Recapitulated in SH-SY5Y Cells. *FEBS letters*, 583(2): 394–400.

Nonaka, Takashi, Kametani, Fuyuki, Arai, Tetsuaki, Akiyama, Haruhiko and

Hasegawa, Masato (2009), Truncation and Pathogenic Mutations Facilitate the Formation of Intracellular Aggregates of TDP-43. *Human molecular genetics*, 18(18): 3353–3364.

Okamoto, K, Hirai, S, Amari, M, Watanabe, M and Sakurai, A (1993), Bunina Bodies in Amyotrophic-Lateral-Sclerosis Immunostained with Rabbit Anti-Cystatin-C Serum. *Neuroscience letters*, 162(1-2): 125–128.

Okamoto, Koichi, Mizuno, Yuji and Fujita, Yukio (2008), Bunina Bodies in Amyotrophic Lateral Sclerosis. *Neuropathology*, 28(2): 109–115.

Oosthuyse, B, Moons, L, Storkebaum, E, Beck, H, Nuyens, D, Brusselmans, K, Van Dorpe, J, Hellings, P, Gorselink, M, Heymans, S, Theilmeier, G, Dewerchin, M, Laudenbach, V, Vermylen, P, Raat, H, Acker, T, Vleminckx, V, Van Den Bosch, L, Cashman, N, Fujisawa, H, Drost, M R, Sciot, R, Bruyninckx, F, Hicklin, D J, Ince, C, Gressens, P, Lupu, F, Plate, K H, Robberecht, W, Herbert, J M, Collen, D and Carmeliet, P (2001), Deletion of the Hypoxia-Response Element in the Vascular Endothelial Growth Factor Promoter Causes Motor Neuron Degeneration. *Nature genetics*, 28(2): 131–138.

Ou, S H, Wu, F, Harrich, D, García-Martínez, L F and Gaynor, R B (1995), Cloning and Characterization of a Novel Cellular Protein, TDP-43, That Binds to Human Immunodeficiency Virus Type 1 TAR DNA Sequence Motifs. *Journal of virology*, 69(6): 3584–3596.

Paek, Ki Young, Hong, Ka Young, Ryu, Incheol, Park, Sung Mi, Keum, Sun Ju, Kwon, Oh Sung and Jang, Sung Key (2015), Translation Initiation Mediated by RNA Looping. *Proceedings of the National Academy of Sciences of the United States of America*, 112(4): 1041–1046.

Pandey, Udai Bhan, Nie, Zhiping, Batlevi, Yakup, McCray, Brett A, Ritson, Gillian P, Nedelsky, Natalia B, Schwartz, Stephanie L, DiProspero, Nicholas A, Knight, Melanie A, Schuldiner, Oren, Padmanabhan, Ranjani, Hild, Marc, Berry, Deborah L, Garza, Dan, Hubbert, Charlotte C, Yao, Tso-Pang, Baehrecke, Eric H and Taylor, J Paul (2007), HDAC6 Rescues Neurodegeneration and Provides an Essential Link Between Autophagy and the UPS. *Nature*, 447(7146): 859–863.

Parker, Roy and Sheth, Ujwal (2007), P Bodies and the Control of mRNA Translation and Degradation. *Molecular cell*, 25(5): 635–646.

Parker, Sarah J, Meyerowitz, Jodi, James, Janine L, Liddell, Jeffrey R, Crouch, Peter J, Kanninen, Katja M and White, Anthony R (2012), Endogenous TDP-43 Localized to Stress Granules Can Subsequently Form Protein Aggregates. *Neurochemistry international*, 60(4): 415–424.

Parkinson, N, Ince, P G, Smith, M O, Highley, R, Skibinski, G, Andersen, P M, Morrison, K E, Pall, H S, Hardiman, O, Collinge, J, Shaw, P J, Fisher, E M C, MRC Proteomics in ALS StudyFReJA Consortium (2006), ALS Phenotypes with Mutations in CHMP2B (Charged Multivesicular Body Protein 2B). *Neurology*, 67(6): 1074–1077.

Passoni, Monica, De Conti, Laura, Baralle, Marco and Buratti, Emanuele (2012), UG Repeats/TDP-43 Interactions Near 5′ Splice Sites Exert Unpredictable

Effects on Splicing Modulation. *Journal of molecular biology*, 415(1): 46–60.

Paulson, H L and Fischbeck, K H (1996), Trinucleotide Repeats in Neurogenetic Disorders. *Annual review of neuroscience*, 19: 79–107.

Pesiridis, G Scott, Tripathy, Kalyan, Tanik, Selçuk, Trojanowski, John Q and Lee, Virginia M-Y (2011), A 'Two-Hit' Hypothesis for Inclusion Formation by Carboxyl-Terminal Fragments of TDP-43 Protein Linked to RNA Depletion and Impaired Microtubule-Dependent Transport. *Journal of Biological Chemistry*, 286(21): 18845–18855.

Pilli, Manohar, Arko-Mensah, John, Ponpuak, Marisa, Roberts, Esteban, Master, Sharon, Mandell, Michael A, Dupont, Nicolas, Ornatowski, Wojciech, Jiang, Shanya, Bradfute, Steven B, Bruun, Jack-Ansgar, Hansen, Tom Egil, Johansen, Terje and Deretic, Vojo (2012), TBK-1 Promotes Autophagy-Mediated Antimicrobial Defense by Controlling Autophagosome Maturation. *Immunity*, 37(2): 223–234.

Piñol-Roma, S and Dreyfuss, G (1992), Shuttling of Pre-mRNA Binding Proteins Between Nucleus and Cytoplasm. *Nature*, 355(6362): 730–732.

Piñol-Roma, S and Dreyfuss, G (1993), hnRNP Proteins: Localization and Transport Between the Nucleus and the Cytoplasm. *Trends in cell biology*, 3(5): 151–155.

Polymenidou, Magdalini, Lagier-Tourenne, Clotilde, Hutt, Kasey R, Huelga, Stephanie C, Moran, Jacqueline, Liang, Tiffany Y, Ling, Shuo-Chien, Sun, Eveline, Wancewicz, Edward, Mazur, Curt, Kordasiewicz, Holly, Sedaghat, Yalda, Donohue, John Paul, Shiue, Lily, Bennett, C Frank, Yeo, Gene W and Cleveland, Don W (2011), Long Pre-mRNA Depletion and RNA Missplicing Contribute to Neuronal Vulnerability From Loss of TDP-43. *Nature neuroscience*, 14(4): 459–468.

Puls, Imke, Jonnakuty, Catherine, LaMonte, Bernadette H, Holzbaur, Erika L F, Tokito, Mariko, Mann, Eric, Floeter, Mary Kay, Bidus, Kimberly, Drayna, Dennis, Oh, Shin J, Brown, Robert H, Ludlow, Christy L and Fischbeck, Kenneth H (2003), Mutant Dynactin in Motor Neuron Disease. *Nature genetics*, 33(4): 455–456.

Reaume, A G, Elliott, J L, Hoffman, E K, Kowall, N W, Ferrante, R J, Siwek, D F, Wilcox, H M, Flood, D G, Beal, M F, Brown, R H, Scott, R W and Snider, W D (1996), Motor Neurons in Cu/Zn Superoxide Dismutase-Deficient Mice Develop Normally but Exhibit Enhanced Cell Death After Axonal Injury. *Nature genetics*, 13(1): 43–47.

Reid, David W, Shenolikar, Shirish and Nicchitta, Christopher V (2015), Simple and Inexpensive Ribosome Profiling Analysis of mRNA Translation. *Methods (San Diego, Calif.)*.

Reid, Evan, Kloos, Mark, Ashley-Koch, Allison, Hughes, Lori, Bevan, Simon, Svenson, Ingrid K, Graham, Felicia Lennon, Gaskell, Perry C, Dearlove, Andrew, Pericak-Vance, Margaret a, Rubinsztein, David C and Marchuk, Douglas a (2002), A Kinesin Heavy Chain (KIF5A) Mutation in Hereditary Spastic Paraplegia (SPG10). *American journal of human genetics*, 71(5): 1189–1194.

Renoux, Abigail J and Todd, Peter K (2012), Neurodegeneration the RNA Way. *Progress in neurobiology*, 97(2): 173–189.

Renton, Alan E, Majounie, Elisa, Waite, Adrian, Simon-Sanchez, Javier, Rollinson, Sara, Gibbs, J Raphael, Schymick, Jennifer C, Laaksovirta, Hannu, van Swieten, John C, Myllykangas, Liisa, Kalimo, Hannu, Paetau, Anders, Abramzon, Yevgeniya, Remes, Anne M, Kaganovich, Alice, Scholz, Sonja W, Duckworth, Jamie, Ding, Jinhui, Harmer, Daniel W, Hernandez, Dena G, Johnson, Janel O, Mok, Kin, Ryten, Mina, Trabzuni, Danyah, Guerreiro, Rita J, Orrell, Richard W, Neal, James, Murray, Alex, Pearson, Justin, Jansen, Iris E, Sondervan, David, Seelaar, Harro, Blake, Derek, Young, Kate, Halliwell, Nicola, Callister, Janis Bennion, Toulson, Greg, Richardson, Anna, Gerhard, Alex, Snowden, Julie, Mann, David, Neary, David, Nalls, Michael a, Peuralinna, Terhi, Jansson, Lilja, Isoviita, Veli-Matti, Kaivorinne, Anna-Lotta, Hölttä-Vuori, Maarit, Ikonen, Elina, Sulkava, Raimo, Benatar, Michael, Wuu, Joanne, Chiò, Adriano, Restagno, Gabriella, Borghero, Giuseppe, Sabatelli, Mario, ITALSGEN Consortium, Heckerman, David, Rogaeva, Ekaterina, Zinman, Lorne, Rothstein, Jeffrey D, Sendtner, Michael, Drepper, Carsten, Eichler, Evan E, Alkan, Can, Abdullaev, Ziedulla, Pack, Svetlana D, Dutra, Amalia, Pak, Evgenia, Hardy, John, Singleton, Andrew, Williams, Nigel M, Heutink, Peter, Pickering-Brown, Stuart, Morris, Huw R, Tienari, Pentti J and Traynor, Bryan J (2011), A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*, 72(2): 257–268.

Richardson, Katie, Allen, Scott P, Mortiboys, Heather, Grierson, Andrew J, Wharton, Stephen B, Ince, Paul G, Shaw, Pamela J and Heath, Paul R (2013), The Effect of SOD1 Mutation on Cellular Bioenergetic Profile and Viability in Response to Oxidative Stress and Influence of Mutation-Type (W. Le, Ed.). *PloS one*, 8(6): e68256.

Ringholz, G M, Appel, S H, Bradshaw, M, Cooke, N A, Mosnik, D M and Schulz, P E (2005), Prevalence and Patterns of Cognitive Impairment in Sporadic ALS. *Neurology*, 65(4): 586–590.

Ritchie, Matthew E, Phipson, Belinda, Wu, Di, Hu, Yifang, Law, Charity W, Shi, Wei and Smyth, Gordon K (2015), Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic acids research*, 43(7): e47.

Rodríguez-Navarro, Susana and Hurt, Ed (2011), Linking Gene Regulation to mRNA Production and Export. *Current opinion in cell biology*, 23(3): 302–309.

Rogelj, Boris, Easton, Laura E, Bogu, Gireesh K, Stanton, Lawrence W, Rot, Gregor, Curk, Tomaž, Zupan, Blaž, Sugimoto, Yoichiro, Modic, Miha, Haberman, Nejc, Tollervey, James, Fujii, Ritsuko, Takumi, Toru, Shaw, Christopher E and Ule, Jernej (2012), Widespread Binding of FUS Along Nascent RNA Regulates Alternative Splicing in the Brain. *Scientific reports*, 2: 603.

Romano, Giulia, Appocher, Chiara, Scorzeto, Michele, Klima, Raffaella, Baralle, Francisco E, Megighian, Aram and Feiguin, Fabian (2015), Glial TDP-43 Regulates Axon Wrapping, GluRIIA Clustering and Fly Motility by Autonomous and Non-Autonomous Mechanisms. *Human molecular*

genetics, 24(21): 6134–6145.

Romano, Maurizio, Buratti, Emanuele, Romano, Giulia, Klima, Raffaella, Del Bel Belluz, Lisa, Stuani, Cristiana, Baralle, Francisco and Feiguin, Fabian (2014), Evolutionarily Conserved Heterogeneous Nuclear Ribonucleoprotein (hnRNP) a/B Proteins Functionally Interact with Human and Drosophila TAR DNA-Binding Protein 43 (TDP-43). *The Journal of biological chemistry*, 289(10): 7121–7130.

Romano, Maurizio, Feiguin, Fabian and Buratti, Emanuele (2012), Drosophila Answers to TDP-43 Proteinopathies. *Journal of amino acids*, 2012: 356081.

Rosen, D R, Siddique, T, Patterson, D, Figlewicz, D A, Sapp, P, Hentati, a, Donaldson, D, Goto, J, O'Regan, J P and Deng, H X (1993), Mutations in Cu/Zn Superoxide Dismutase Gene Are Associated with Familial Amyotrophic Lateral Sclerosis. *Nature*, 362(6415): 59–62.

Rutherford, Nicola J, Zhang, Yong-Jie, Baker, Matt, Gass, Jennifer M, Finch, NiCole a, Xu, Ya-Fei, Stewart, Heather, Kelley, Brendan J, Kuntz, Karen, Crook, Richard J P, Sreedharan, Jemeen, Vance, Caroline, Sorenson, Eric, Lippa, Carol, Bigio, Eileen H, Geschwind, Daniel H, Knopman, David S, Mitsumoto, Hiroshi, Petersen, Ronald C, Cashman, Neil R, Hutton, Mike, Shaw, Christopher E, Boylan, Kevin B, Boeve, Bradley, Graff-Radford, Neill R, Wszolek, Zbigniew K, Caselli, Richard J, Dickson, Dennis W, Mackenzie, Ian R, Petrucelli, Leonard and Rademakers, Rosa (2008), Novel Mutations in TARDBP (TDP-43) in Patients with Familial Amyotrophic Lateral Sclerosis. *PLoS genetics*, 4(9): e1000193.

Sabatelli, Mario, Zollino, Marcella, Conte, Amelia, Del Grande, Alessandra, Marangi, Giuseppe, Lucchini, Matteo, Mirabella, Massimiliano, Romano, Angela, Piacentini, Roberto, Bisogni, Giulia, Lattante, Serena, Luigetti, Marco, Rossini, Paolo Maria and Moncada, Alice (2015), Primary Fibroblasts Cultures Reveal TDP-43 Abnormalities in Amyotrophic Lateral Sclerosis Patients with and Without SOD1 Mutations. *Neurobiology of Aging*, 36(5): 2005.e5–2005.e13.

Sanz, Elisenda, Yang, Linghai, Su, Thomas, Morris, David R, McKnight, G Stanley and Amieux, Paul S (2009), Cell-Type-Specific Isolation of Ribosome-Associated mRNA From Complex Tissues. *PNAS*, 106(33): 13939–13944.

Sareen, Dhruv, O'Rourke, Jacqueline G, Meera, Pratap, Muhammad, A K M G, Grant, Sharday, Simpkinson, Megan, Bell, Shaughn, Carmona, Sharon, Ornelas, Loren, Sahabian, Anais, Gendron, Tania, Petrucelli, Leonard, Baughn, Michael, Ravits, John, Harms, Matthew B, Rigo, Frank, Bennett, C Frank, Otis, Thomas S, Svendsen, Clive N and Baloh, Robert H (2013), Targeting RNA Foci in iPSC-Derived Motor Neurons From ALS Patients with a C9ORF72 Repeat Expansion. *Science translational medicine*, 5(208): 208ra149.

Schmid, Bettina, Hruscha, Alexander, Hogl, Sebastian, Banzhaf-Strathmann, Julia, Strecker, Katrin, van der Zee, Julie, Teucke, Mathias, Eimer, Stefan, Hegermann, Jan, Kittelmann, Maike, Kremmer, Elisabeth, Cruts, Marc, Solchenberger, Barbara, Hasenkamp, Laura, van Bebber, Frauke, Van Broeckhoven, Christine, Edbauer, Dieter, Lichtenthaler, Stefan F and

Haass, Christian (2013), Loss of ALS-Associated TDP-43 in Zebrafish Causes Muscle Degeneration, Vascular Dysfunction, and Reduced Motor Neuron Axon Outgrowth. *PNAS*, 110(13): 4986–4991.

Schmid, Manfred and Jensen, Torben Heick (2008), Quality Control of mRNP in the Nucleus. *Chromosoma*, 117(5): 419–429.

Schwanhäusser, Björn, Busse, Dorothea, Li, Na, Dittmar, Gunnar, Schuchhardt, Johannes, Wolf, Jana, Chen, Wei and Selbach, Matthias (2011), Global Quantification of Mammalian Gene Expression Control. *Nature*, 473(7347): 337–342.

Scotter, Emma L, Vance, Caroline, Nishimura, Agnes L, Lee, Youn-Bok, Chen, Han-Jou, Urwin, Hazel, Sardone, Valentina, Mitchell, Jacqueline C, Rogelj, Boris, Rubinsztein, David C and Shaw, Christopher E (2014), Differential Roles of the Ubiquitin Proteasome System and Autophagy in the Clearance of Soluble and Aggregated TDP-43 Species. *Journal of cell science*, 127(Pt 6): 1263–1278.

Seibenhener, M Lamar, Babu, Jeganathan Ramesh, Geetha, Thangiah, Wong, Hing C, Krishna, N Rama and Wooten, Marie W (2004), Sequestosome 1/P62 Is a Polyubiquitin Chain Binding Protein Involved in Ubiquitin Proteasome Degradation. *Molecular and cellular biology*, 24(18): 8055–8068.

Sephton, Chantelle F, Cenik, Can, Kucukural, Alper, Dammer, Eric B, Cenik, Basar, Han, Yuhong, Dewey, Colleen M, Roth, Frederick P, Herz, Joachim, Peng, Junmin, Moore, Melissa J and Yu, Gang (2011), Identification of Neuronal RNA Targets of TDP-43-Containing Ribonucleoprotein Complexes. *The Journal of biological chemistry*, 286(2): 1204–1215.

Sephton, Chantelle F, Good, Shannon K, Atkin, Stan, Dewey, Colleen M, Mayer, Paul, Herz, Joachim and Yu, Gang (2010), TDP-43 Is a Developmentally Regulated Protein Essential for Early Embryonic Development. *The Journal of biological chemistry*, 285(9): 6826–6834.

Shan, Xiu, Chiang, Po-Min, Price, Donald L and Wong, Philip C (2010), Altered Distributions of Gemini of Coiled Bodies and Mitochondria in Motor Neurons of TDP-43 Transgenic Mice. *Proceedings of the National Academy of Sciences of the United States of America*, 107(37): 16325–16330.

Shaw, P J (2005), Molecular and Cellular Pathways of Neurodegeneration in Motor Neurone Disease. *Journal of neurology, neurosurgery, and psychiatry*, 76(8): 1046–1057.

Shi, Ping, Wei, Yanming, Zhang, Jiayu, Gal, Jozsef and Zhu, Haining (2010), Mitochondrial Dysfunction Is a Converging Point of Multiple Pathological Pathways in Amyotrophic Lateral Sclerosis. *Journal of Alzheimer's disease : JAD*, 20 Suppl 2: S311–24.

Shibata, N, Nagai, R, Uchida, K, Horiuchi, S, Yamada, S, Hirano, a, Kawaguchi, M, Yamamoto, T, Sasaki, S and Kobayashi, M (2001), Morphological Evidence for Lipid Peroxidation and Protein Glycoxidation in Spinal Cords From Sporadic Amyotrophic Lateral Sclerosis Patients. *Brain research*, 917(1): 97–104.

Shigeoka, Toshiaki, Jung, Hosung, Jung, Jane, Turner-Bridger, Benita, Ohk, Jiyeon, Lin, Julie Qiaojin, Amieux, Paul S and Holt, Christine E (2016), Dynamic Axonal Translation in Developing and Mature Visual Circuits. *Cell*, 166(1): 181–192.

Siddique, T, Figlewicz, D A, PERICAKVANCE, M A, Haines, J L, ROULEAU, G, JEFFERS, A J, Sapp, P, Hung, W Y, BEBOUT, J, MCKENNAYASEK, D, DENG, G, Horvitz, H R, GUSELLA, J F, Brown, R H and ROSES, A D (1991), Linkage of a Gene Causing Familial Amyotrophic-Lateral-Sclerosis to Chromosome-21 and Evidence of Genetic-Locus Heterogeneity. *New England Journal of Medicine*, 324(20): 1381–1384.

Singh, Minati (2012), Dysregulated a to I RNA Editing and Non-Coding RNAs in Neurodegeneration. *Frontiers in genetics*, 3: 326–326.

Skehel, P a, Fabian-Fine, R and Kandel, E R (2000), Mouse VAP33 Is Associated with the Endoplasmic Reticulum and Microtubules. *Proceedings of the National Academy of Sciences of the United States of America*, 97(3): 1101–1106.

Smith, Bradley N, Ticozzi, Nicola, Fallini, Claudia, Gkazi, Athina-Soragia, Topp, Simon, Kenna, Kevin P, Scotter, Emma L, Kost, Jason, Keagle, Pamela, Miller, Jack W, Calini, Daniela, Vance, Caroline, Danielson, Eric W, Troakes, Claire, Tiloca, Cinzia, Al-Sarraj, Safa, Lewis, Elizabeth A, King, Andrew, Colombrita, Claudia, Pensato, Viviana, Castellotti, Barbara, de Belleroche, Jacqueline, Baas, Frank, Asbroek, ten, Anneloor L M A, Sapp, Peter C, Mckenna-yasek, Diane, McLaughlin, Russell L, Polak, Meraida, Asress, Seneshaw, Esteban-Pérez, Jesús, Muñoz-Blanco, JoséLuis L, Simpson, Michael, van Rheenen, Wouter, Diekstra, Frank P, Lauria, Giuseppe, Duga, Stefano, Corti, Stefania, Cereda, Cristina, Corrado, Lucia, Sorarù, Gianni, Morrison, Karen E, Williams, Kelly L, Nicholson, Garth a, Blair, Ian P, Dion, Patrick a, Leblond, Claire S, Rouleau, Guy a, Hardiman, Orla, Veldink, Jan H, van den Berg, Leonard H, Al-Chalabi, Ammar, Pall, Hardev, Shaw, Pamela J, Turner, Martin R, Talbot, Kevin, Taroni, Franco, García-Redondo, Alberto, Wu, Zheyang, Glass, Jonathan D, Gellera, Cinzia, Ratti, Antonia, Brown, Robert H, Silani, Vincenzo, Shaw, Christopher E and Landers, John E (2014), Exome-Wide Rare Variant Analysis Identifies TUBA4A Mutations Associated with Familial ALS. *Neuron*, 84(2): 324–331.

Sobue, G, Hashizume, Y, Yasuda, T, Mukai, E, Kumagai, T, Mitsuma, T and Trojanowski, J Q (1990), Phosphorylated High-Molecular-Weight Neurofilament Protein in Lower Motor Neurons in Amyotrophic Lateral Sclerosis and Other Neurodegenerative Diseases Involving Ventral Horn Cells. *Acta neuropathologica*, 79(4): 402–408.

Sreedharan, Jemeen, Blair, Ian P, Tripathi, Vineeta B, Hu, Xun, Vance, Caroline, Rogelj, Boris, Ackerley, Steven, Durnall, Jennifer C, Williams, Kelly L, Buratti, Emanuele, Baralle, Francisco, de Belleroche, Jacqueline, Mitchell, J Douglas, Leigh, P Nigel, Al-Chalabi, Ammar, Miller, Christopher C, Nicholson, Garth and Shaw, Christopher E (2008), TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science*, 319(5870): 1668–1672.

Stallings, Nancy R, Puttaparthi, Krishna, Luther, Christina M, Burns, Dennis K and Elliott, Jeffrey L (2010), Progressive Motor Weakness in Transgenic

Mice Expressing Human TDP-43. *Neurobiology of disease*, 40(2): 404–414.

Steitz, Joan Argetsinger (1969), Polypeptide Chain Initiation: Nucleotide Sequences of the Three Ribosomal Binding Sites in Bacteriophage R17 RNA. *Nature*, 224(5223): 957–964.

Stribl, Carola, Samara, Aladin, Trümbach, Dietrich, Peis, Regina, Neumann, Manuela, Fuchs, Helmut, Gailus-Durner, Valerie, De Angelis, Martin Hrabě, Rathkolb, Birgit, Wolf, Eckhard, Beckers, Johannes, Horsch, Marion, Neff, Frauke, Kremmer, Elisabeth, Koob, Sebastian, Reichert, Andreas S, Hans, Wolfgang, Rozman, Jan, Klingenspor, Martin, Aichler, Michaela, Karl Walch, Axel, Becker, Lore, Klopstock, Thomas, Glasl, Lisa, Hölter, Sabine M, Wurst, Wolfgang and Floss, Thomas (2014), Mitochondrial Dysfunction and Decrease in Body Weight of a Transgenic Knock-in Mouse Model for TDP-43. *Journal of Biological Chemistry*, 289(15): 10769–10784.

Subramanian, Vasanta, Crabtree, Benedict and Acharya, K Ravi (2008), Human Angiogenin Is a Neuroprotective Factor and Amyotrophic Lateral Sclerosis Associated Angiogenin Variants Affect Neurite Extension/Pathfinding and Survival of Motor Neurons. *Human molecular genetics*, 17(1): 130–149.

Suzuki, Hiroaki, Lee, Kikyo and Matsuoka, Masaaki (2011), TDP-43-Induced Death Is Associated with Altered Regulation of BIM and Bcl-xL and Attenuated by Caspase-Mediated TDP-43 Cleavage. *Journal of Biological Chemistry*, 286(15): 13171–13183.

Swarup, Vivek, Phaneuf, Daniel, Bareil, Christine, Robertson, Janice, Rouleau, Guy a, Kriz, Jasna and Julien, Jean-Pierre (2011), Pathological Hallmarks of Amyotrophic Lateral Sclerosis/Frontotemporal Lobar Degeneration in Transgenic Mice Produced with TDP-43 Genomic Fragments. *Brain*, 134(9): 2610–2626.

Takahashi, Yuji, Fukuda, Yoko, Yoshimura, Jun, Toyoda, Atsushi, Kurppa, Kari, Moritoyo, Hiroyoko, Belzil, Veronique V, Dion, Patrick a, Higasa, Koichiro, Doi, Koichiro, Ishiura, Hiroyuki, Mitsui, Jun, Date, Hidetoshi, Ahsan, Budrul, Matsukawa, Takashi, Ichikawa, Yaeko, Moritoyo, Takashi, Ikoma, Mayumi, Hashimoto, Tsukasa, Kimura, Fumiharu, Murayama, Shigeo, Onodera, Osamu, Nishizawa, Masatoyo, Yoshida, Mari, Atsuta, Naoki, Sobue, Gen, JaCALS, Fifita, Jennifer A, Williams, Kelly L, Blair, Ian P, Nicholson, Garth a, Gonzalez-Perez, Paloma, Brown, Robert H, Nomoto, Masahiro, Elenius, Klaus, Rouleau, Guy a, Fujiyama, Asao, Morishita, Shinichi, Goto, Jun and Tsuji, Shoji (2013), ERBB4 Mutations That Disrupt the Neuregulin-ErbB4 Pathway Cause Amyotrophic Lateral Sclerosis Type 19. *American journal of human genetics*, 93(5): 900–905.

Tan, Adelene Y and Manley, James L (2009), The TET Family of Proteins: Functions and Roles in Disease. *Journal of molecular cell biology*, 1(2): 82–92.

Tan, Chun-Feng, Eguchi, Hiroto, Tagawa, Asako, Onodera, Osamu, Iwasaki, Takuya, Tsujino, Akira, Nishizawa, Masatoyo, Kakita, Akiyoshi and Takahashi, Hitoshi (2007), TDP-43 Immunoreactivity in Neuronal Inclusions in Familial Amyotrophic Lateral Sclerosis with or Without SOD1 Gene Mutation. *Acta neuropathologica*, 113(5): 535–542.

Tanaka, Fumiaki, Ikenaka, Kensuke, Yamamoto, Masahiko and Sobue, Gen (2011), Neuropathology and Omics in Motor Neuron Diseases. *Neuropathology*, 32(4): 458–462.

Tanaka, Mikiei, Chock, P Boon and Stadtman, Earl R (2007), Oxidized Messenger RNA Induces Translation Errors. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1): 66–71.

Thermann, Rolf and Hentze, Matthias W (2007), Drosophila miR2 Induces Pseudo-Polysomes and Inhibits Translation Initiation. *Nature*, 447(7146): 875–878.

Therrien, Martine and Parker, J Alex (2014), Worming Forward: Amyotrophic Lateral Sclerosis Toxicity Mechanisms and Genetic Interactions in Caenorhabditis Elegans. *Frontiers in genetics*, 5(APR): 85.

Tian, Bin and Manley, James L (2013), Alternative Cleavage and Polyadenylation: the Long and Short of It. *Trends in biochemical sciences*, 38(6): 312–320.

Timchenko, L T, Miller, J W, Timchenko, N a, DeVore, D R, Datar, K V, Lin, L, Roberts, R, Caskey, C T and Swanson, M S (1996), Identification of a (CUG)N Triplet Repeat RNA-Binding Protein and Its Expression in Myotonic Dystrophy. *Nucleic acids research*, 24(22): 4407–4414.

Tollervey, James R, Curk, Tomaž, Rogelj, Boris, Briese, Michael, Cereda, Matteo, Kayikci, Melis, König, Julian, Hortobagyi, Tibor, Nishimura, Agnes L, Župunski, Vera, Patani, Rickie, Chandran, Siddharthan, Rot, Gregor, Zupan, Blaž, Shaw, Christopher E and Ule, Jernej (2011), Characterizing the RNA Targets and Position-Dependent Splicing Regulation by TDP-43. *Nature neuroscience*, 14(4): 452–458.

Tripathi, Vineeta Bhasker, Baskaran, Pranetha, Shaw, Christopher E and Guthrie, Sarah (2014), Tar DNA-Binding Protein-43 (TDP-43) Regulates Axon Growth in Vitro and in Vivo. *Neurobiology of disease*, 65: 25–34.

Tsai, Kuen-Jer, Yang, Chun-Hung, Fang, Yen-Hsin, Cho, Kuan-Hung, Chien, Wei-Lin, Wang, Wei-Ting, Wu, Tzu-Wei, Lin, Ching-Po, Fu, Wen-Mei and Shen, Che-Kun James (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): 1661–1673.

Tsao, William, Jeong, Yun Ha, Lin, Sophie, Ling, Jonathan, Price, Donald L, Chiang, Po-Min and Wong, Philip C (2012), Rodent Models of TDP-43: Recent Advances. *Brain research*, 1462: 26–39.

Uchida, Azusa, Sasaguri, Hiroki, Kimura, Nobuyuki, Tajiri, Mio, Ohkubo, Takuya, Ono, Fumiko, Sakaue, Fumika, Kanai, Kazuaki, Hirai, Takashi, Sano, Tatsuhiko, Shibuya, Kazumoto, Kobayashi, Masaki, Yamamoto, Mariko, Yokota, Shigefumi, Kubodera, Takayuki, Tomori, Masaki, Sakaki, Kyohei, Enomoto, Mitsuhiro, Hirai, Yukihiko, Kumagai, Jiro, Yasutomi, Yasuhiro, Mochizuki, Hideki, Kuwabara, Satoshi, Uchihara, Toshiki, Mizusawa, Hidehiro and Yokota, Takanori (2012), Non-Human Primate Model of Amyotrophic Lateral Sclerosis with Cytoplasmic Mislocalization of TDP-43.

Brain : a journal of neurology, 135(Pt 3): 833–846.

Uranishi, H, Tetsuka, T, Yamashita, M, Asamitsu, K, Shimizu, M, Itoh, M and Okamoto, T (2001), Involvement of the Pro-Oncoprotein TLS (Translocated in Liposarcoma) in Nuclear Factor-Kappa B P65-Mediated Transcription as a Coactivator. *The Journal of biological chemistry*, 276(16): 13395–13401.

Urushitani, Makoto, Sato, Takashi, Bamba, Hitoshi, Hisa, Yasuo and Tooyama, Ikuo (2010), Synergistic Effect Between Proteasome and Autophagosome in the Clearance of Polyubiquitinated TDP-43. *Journal of neuroscience research*, 88(4): 784–797.

Valencia-Sanchez, Marco Antonio, Liu, Jidong, Hannon, Gregory J and Parker, Roy (2006), Control of Translation and mRNA Degradation by miRNAs and siRNAs. *Genes \& development*, 20(5): 515–524.

Vance, Caroline, Rogelj, Boris, Hortobagyi, Tibor, De Vos, Kurt J, Nishimura, Agnes Lumi, Sreedharan, Jemeen, Hu, Xun, Smith, Bradley, Ruddy, Deborah, Wright, Paul, Ganesalingam, Jeban, Williams, Kelly L, Tripathi, Vineeta, Al-Saraj, Safa, Al-Chalabi, Ammar, Leigh, P Nigel, Blair, Ian P, Nicholson, Garth, de Belleroche, Jackie, Gallo, Jean-Marc, Miller, Christopher C and Shaw, Christopher E (2009), Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science (New York, N.Y.)*, 323(5918): 1208–1211.

Veyrat-Durebex, C, Corcia, P, Dangoumau, A, Laumonnier, F, Piver, E, Gordon, P H, Andres, C R, Vourc'h, P and Blasco, H (2013), Advances in Cellular Models to Explore the Pathophysiology of Amyotrophic Lateral Sclerosis. *Molecular neurobiology*.

Viphakone, Nicolas, Hautbergue, Guillaume M, Walsh, Matthew, Chang, Chung-te, Holland, Arthur, Folco, Eric G, Reed, Robin and Wilson, Stuart A (2012), TREX Exposes the RNA-Binding Domain of Nxf1 to Enable mRNA Export. *Nature communications*, 3: 1006.

Vogel, Christine, Abreu, Raquel de Sousa, Ko, Daijin, Le, Shu-Yun, Shapiro, Bruce A, Burns, Suzanne C, Sandhu, Devraj, Boutz, Daniel R, Marcotte, Edward M and Penalva, Luiz O (2010), Sequence Signatures and mRNA Concentration Can Explain Two-Thirds of Protein Abundance Variation in a Human Cell Line. *Molecular systems biology*, 6: 400.

Voigt, Aaron, Herholz, David, Fiesel, Fabienne C, Kaur, Kavita, Müller, Daniel, Karsten, Peter, Weber, Stephanie S, Kahle, Philipp J, Marquardt, Till and Schulz, Jörg B (2010), TDP-43-Mediated Neuron Loss in Vivo Requires RNA-Binding Activity. *PloS one*, 5(8): e12247.

Voorhoeve, P Mathijs, le Sage, Carlos, Schrier, Mariette, Gillis, Ad J M, Stoop, Hans, Nagel, Remco, Liu, Ying-Poi, van Duijse, Josyanne, Drost, Jarno, Griekspoor, Alexander, Zlotorynski, Eitan, Yabuta, Norikazu, De Vita, Gabriella, Nojima, Hiroshi, Looijenga, Leendert H J and Agami, Reuven (2006), A Genetic Screen Implicates miRNA-372 and miRNA-373 as Oncogenes in Testicular Germ Cell Tumors. *Cell*, 124(6): 1169–1181.

Wahl, Markus C, Will, Cindy L and Lührmann, Reinhard (2009), The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*, 136(4):

701–718.

Walker, Adam K, Spiller, Krista J, Ge, Guanghui, Zheng, Allen, Xu, Yan, Zhou, Melissa, Tripathy, Kalyan, Kwong, Linda K, Trojanowski, John Q and Lee, Virginia M-Y (2015), Functional Recovery in New Mouse Models of ALS/FTLD After Clearance of Pathological Cytoplasmic TDP-43. *Acta neuropathologica*: 1–18.

Walsh, Matthew J, Cooper-Knock, Johnathan, Dodd, Jennifer E, Stopford, Matthew J, Mihaylov, Simeon R, Kirby, Janine, Shaw, Pamela J and Hautbergue, Guillaume M (2015), Invited Review: Decoding the Pathophysiological Mechanisms That Underlie RNA Dysregulation in Neurodegenerative Disorders: a Review of the Current State of the Art. *Neuropathology and applied neurobiology*, 41(2): 109–134.

Wang, I-Fan, Wu, Lien-Szn, Chang, Hsiang-Yu and Shen, C K James (2008), TDP-43, the Signature Protein of FTLD-U, Is a Neuronal Activity-Responsive Factor. *Journal of neurochemistry*, 105(3): 797–806.

Wang, Li-Chun, Chen, Kuan-Yu, Pan, Huichin, Wu, Chia-Chieh, Chen, Po-Hsuan, Liao, Yuan-Ting, Li, Chin, Huang, Min-Lang and Hsiao, Kuang-Ming (2011), Muscleblind Participates in RNA Toxicity of Expanded CAG and CUG Repeats in Caenorhabditis Elegans. *Cellular and molecular life sciences : CMLS*, 68(7): 1255–1267.

Wang, Wenzhang, Li, Li, Lin, Wen-Lang, Dickson, Dennis W, Petrucelli, Leonard, Zhang, Teng and Wang, Xinglong (2013), The ALS Disease-Associated Mutant TDP-43 Impairs Mitochondrial Dynamics and Function in Motor Neurons. *Human molecular genetics*, 22(23): 4706–4719.

Wang, Wenzhang, Wang, Luwen, Lu, Junjie, Siedlak, Sandra L, Fujioka, Hisashi, Liang, Jingjing, Jiang, Sirui, Ma, Xiaopin, Jiang, Zhen, da Rocha, Edroaldo Lummertz, Sheng, Max, Choi, Heewon, Lerou, Paul H, Li, Hu and Wang, Xinglong (2016), The Inhibition of TDP-43 Mitochondrial Localization Blocks Its Neuronal Toxicity. *Nature medicine*, 22(8): 869–878.

Wächter, Nicole, Storch, Alexander and Hermann, Andreas (2015), Human TDP-43 and FUS Selectively Affect Motor Neuron Maturation and Survival in a Murine Cell Model of ALS by Non-Cell-Autonomous Mechanisms. *Amyotrophic lateral sclerosis & frontotemporal degeneration*, 16(7-8): 431–441.

Webster, Christopher P, Smith, Emma F, Bauer, Claudia S, Moller, Annekathrin, Hautbergue, Guillaume M, Ferraiuolo, Laura, Myszczynska, Monika A, Higginbottom, Adrian, Walsh, Matthew J, Whitworth, Alexander J, Kaspar, Brian K, Meyer, Kathrin, Shaw, Pamela J, Grierson, Andrew J and De Vos, Kurt J (2016), The C9orf72 Protein Interacts with Rab1a and the ULK1 Complex to Regulate Initiation of Autophagy. *The EMBO journal*: e201694401–21.

Wegorzewska, Iga, Bell, Shaughn, Cairns, Nigel J, Miller, Timothy M and Baloh, Robert H (2009), TDP-43 Mutant Transgenic Mice Develop Features of ALS and Frontotemporal Lobar Degeneration. *PNAS*, 106(44): 18809–18814.

Weidberg, Hilla and Elazar, Zvulun (2011), TBK1 Mediates Crosstalk Between

the Innate Immune Response and Autophagy. *Science Signaling*, 4(187): – pe39.

Wils, Hans, Kleinberger, Gernot, Janssens, Jonathan, Pereson, Sandra, Joris, Geert, Cuijt, Ivy, Smits, Veerle, Ceuterick-de Groote, Chantal, Van Broeckhoven, Christine and Kumar-Singh, Samir (2010), TDP-43 Transgenic Mice Develop Spastic Paralysis and Neuronal Inclusions Characteristic of ALS and Frontotemporal Lobar Degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 107(8): 3858–3863.

Winton, Matthew J, Igaz, Lionel M, Wong, Margaret M, Kwong, Linda K, Trojanowski, John Q and Lee, Virginia M-Y (2008), Disturbance of Nuclear and Cytoplasmic TAR DNA-Binding Protein (TDP-43) Induces Disease-Like Redistribution, Sequestration, and Aggregate Formation. *The Journal of biological chemistry*, 283(19): 13302–13309.

Wirth, Brunhilde, Brichta, Lars and Hahnen, Eric (2006), Spinal Muscular Atrophy: From Gene to Therapy. *Seminars in pediatric neurology*, 13(2): 121–131.

Woerner, Andreas C, Frottin, Frédéric, Hornburg, Daniel, Feng, Li R, Meissner, Felix, Patra, Maria, Tatzelt, Joerg, Mann, Matthias, Winklhofer, Konstanze F, Hartl, F Ulrich and Hipp, Mark S (2016), Cytoplasmic Protein Aggregates Interfere with Nucleocytoplasmic Transport of Protein and RNA. *Science*, 351(6269): 173–176.

Wojciechowska, Marzena and Krzyzosiak, Wlodzimierz J (2011), Cellular Toxicity of Expanded RNA Repeats: Focus on RNA Foci. *Human molecular genetics*, 20(19): 3811–3821.

Wolin, S L and Walter, P (1988), Ribosome Pausing and Stacking During Translation of a Eukaryotic mRNA. *EMBO Journal*, 7(11): 3559–3569.

Wong, P C, Pardo, C a, Borchelt, D R, Lee, M K, Copeland, N G, Jenkins, N a, Sisodia, S S, Cleveland, D W and Price, D L (1995), An Adverse Property of a Familial ALS-Linked SOD1 Mutation Causes Motor Neuron Disease Characterized by Vacuolar Degeneration of Mitochondria. *Neuron*, 14(6): 1105–1116.

Wood, J D, Beaujeux, T P and Shaw, P J (2003), Protein Aggregation in Motor Neurone Disorders. *Neuropathology and applied neurobiology*, 29(6): 529–545.

Wu, Chi-Hong, Fallini, Claudia, Ticozzi, Nicola, Keagle, Pamela J, Sapp, Peter C, Piotrowska, Katarzyna, Lowe, Patrick, Koppers, Max, Mckenna-yasek, Diane, Baron, Desiree M, Kost, Jason E, Gonzalez-Perez, Paloma, Fox, Andrew D, Adams, Jenni, Taroni, Franco, Tiloca, Cinzia, Leclerc, Ashley Lyn, Chafe, Shawn C, Mangroo, Dev, Moore, Melissa J, Zitzewitz, Jill A, Xu, Zuo-Shang, van den Berg, Leonard H, Glass, Jonathan D, Siciliano, Gabriele, Cirulli, Elizabeth T, Goldstein, David B, Salachas, Francois, Meininger, Vincent, Rossoll, Wilfried, Ratti, Antonia, Gellera, Cinzia, Bosco, Daryl A, Bassell, Gary J, Silani, Vincenzo, Drory, Vivian E, Brown, Robert H and Landers, John E (2012), Mutations in the Profilin 1 Gene Cause Familial Amyotrophic Lateral Sclerosis. *Nature*, 488(7412): 499–503.

Wu, David, Yu, Wenhao, Kishikawa, Hiroko, Folkerth, Rebecca D, Iafrate, a John, Shen, Yiping, Xin, Winnie, Sims, Katherine and Hu, Guo-Fu (2007), Angiogenin Loss-of-Function Mutations in Amyotrophic Lateral Sclerosis. *Annals of neurology*, 62(6): 609–617.

Wu, Lien-Szu, Cheng, Wei-Cheng and Shen, C K James (2012), Targeted Depletion of TDP-43 Expression in the Spinal Cord Motor Neurons Leads to the Development of Amyotrophic Lateral Sclerosis-Like Phenotypes in Mice. *Journal of Biological Chemistry*, 287(33): 27335–27344.

Wu, Lien-Szu, Cheng, Wei-Cheng, Hou, Shin-Chen, Yan, Yu-Ting, Jiang, Si-Tse and Shen, C K James (2010), TDP-43, a Neuro-Pathosignature Factor, Is Essential for Early Mouse Embryogenesis. *Genesis (New York, N.Y. : 2000)*, 48(1): 56–62.

Xiao, Shangxi, Sanelli, Teresa, Chiang, Helen, Sun, Yulong, Chakrabartty, Avijit, Keith, Julia, Rogaeva, Ekaterina, Zinman, Lorne and Robertson, Janice (2015), Low Molecular Weight Species of TDP-43 Generated by Abnormal Splicing Form Inclusions in Amyotrophic Lateral Sclerosis and Result in Motor Neuron Death. *Acta neuropathologica*.

Xiao, Shangxi, Sanelli, Teresa, Dib, Samar, Sheps, David, Findlater, Joseph, Bilbao, Juan, Keith, Julia, Zinman, Lorne, Rogaeva, Ekaterina and Robertson, Janice (2011), RNA Targets of TDP-43 Identified by UV-CLIP Are Deregulated in ALS. *Molecular and cellular neurosciences*, 47(3): 167–180.

Xu, Ya-Fei, Gendron, Tania F, Zhang, Yong-Jie, Lin, Wen-Lang, D'Alton, Simon, Sheng, Hong, Casey, Monica Castanedes, Tong, Jimei, Knight, Joshua, Yu, Xin, Rademakers, Rosa, Boylan, Kevin, Hutton, Mike, McGowan, Eileen, Dickson, Dennis W, Lewis, Jada and Petrucelli, Leonard (2010), Wild-Type Human TDP-43 Expression Causes TDP-43 Phosphorylation, Mitochondrial Aggregation, Motor Deficits, and Early Mortality in Transgenic Mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(32): 10851–10859.

Xu, Ya-Fei, Zhang, Yong-Jie, Lin, Wen-Lang, Cao, Xiangkun, Stetler, Caroline, Dickson, Dennis W, Lewis, Jada and Petrucelli, Leonard (2011), Expression of Mutant TDP-43 Induces Neuronal Dysfunction in Transgenic Mice. *Molecular Neurodegeneration*, 6(1): 73.

Yang, Chunxing, Tan, Weijia, Whittle, Catheryne, Qiu, Linghua, Cao, Lucheng, Akbarian, Schahram and Xu, Zuoshang (2010), The C-Terminal TDP-43 Fragments Have a High Aggregation Propensity and Harm Neurons by a Dominant-Negative Mechanism. *PloS one*, 5(12): e15878.

Yang, L, Embree, L J, Tsai, S and Hickstein, D D (1998), Oncoprotein TLS Interacts with Serine-Arginine Proteins Involved in RNA Splicing. *The Journal of biological chemistry*, 273(43): 27761–27764.

Yang, Y, Hentati, a, Deng, H X, Dabbagh, O, Sasaki, T, Hirano, M, Hung, W Y, Ouahchi, K, Yan, J, Azim, a C, Cole, N, Gascon, G, Yagmour, a, Ben-Hamida, M, Pericak-Vance, M, Hentati, F and Siddique, T (2001), The Gene Encoding Alsin, a Protein with Three Guanine-Nucleotide Exchange Factor Domains, Is Mutated in a Form of Recessive Amyotrophic Lateral Sclerosis.

Nature genetics, 29(2): 160–165.

Yao, Peng and Fox, Paul L (2013), Aminoacyl-tRNA Synthetases in Medicine and Disease. *EMBO molecular medicine*, 5(3): 332–343.

Yonath, Ada (2009), Large Facilities and the Evolving Ribosome, the Cellular Machine for Genetic-Code Translation. *Journal of the Royal Society, Interface / the Royal Society*, 6 Suppl 5: S575–85.

Zhang, Dapeng, Iyer, Lakshminarayan M, He, Fang and Aravind, L (2012), Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Frontiers in genetics*, 3: 283.

Zhang, Ke, Donnelly, Christopher J, Haeusler, Aaron R, Grima, Jonathan C, Machamer, James B, Steinwald, Peter, Daley, Elizabeth L, Miller, Sean J, Cunningham, Kathleen M, Vidensky, Svetlana, Gupta, Saksham, Thomas, Michael A, Hong, Ingie, Chiu, Shu-Ling, Huganir, Richard L, Ostrow, Lyle W, Matunis, Michael J, Wang, Jiou, Sattler, Rita, Lloyd, Thomas E and Rothstein, Jeffrey D (2015), The C9orf72 Repeat Expansion Disrupts Nucleocytoplasmic Transport. *Nature*, 525(7567): 56–61.

Zhang, Kelvin Xi, Tan, Liming, Pellegrini, Matteo, Zipursky, S Lawrence and McEwen, Jason M (2016), Rapid Changes in the Translatome During the Conversion of Growth Cones to Synaptic Terminals. *Cell reports*, 14(5): 1258–1271.

Zhang, Tao, Hwang, Ho-Yon, Hao, Haiping, Talbot, Conover Jr and Wang, Jiou (2012), Caenorhabditis Elegans RNA-Processing Protein TDP-1 Regulates Protein Homeostasis and Life Span. *Journal of Biological Chemistry*, 287(11): 8371–8382.

Zhang, Yong-Jie, Gendron, Tania F, Grima, Jonathan C, Sasaguri, Hiroki, Jansen-West, Karen, Xu, Ya-Fei, Katzman, Rebecca B, Gass, Jennifer, Murray, Melissa E, Shinohara, Mitsuru, Lin, Wen-Lang, Garrett, Aliesha, Stankowski, Jeannette N, Daughrity, Lillian, Tong, Jimei, Perkerson, Emilie A, Yue, Mei, Chew, Jeannie, Castanedes-Casey, Monica, Kurti, Aishe, Wang, Zizhao S, Liesinger, Amanda M, Baker, Jeremy D, Jiang, Jie, Lagier-Tourenne, Clotilde, Edbauer, Dieter, Cleveland, Don W, Rademakers, Rosa, Boylan, Kevin B, Bu, Guojun, Link, Christopher D, Dickey, Chad A, Rothstein, Jeffrey D, Dickson, Dennis W, Fryer, John D and Petrucelli, Leonard (2016), C9ORF72 Poly(GA) Aggregates Sequester and Impair HR23 and Nucleocytoplasmic Transport Proteins. *Nature neuroscience*.

Zhang, Yong-Jie, Xu, Ya-Fei, Cook, Casey, Gendron, Tania F, Roettges, Paul, Link, Christopher D, Lin, Wen-Lang, Tong, Jimei, Castanedes-Casey, Monica, Ash, Peter, Gass, Jennifer, Rangachari, Vijayaraghavan, Buratti, Emanuele, Baralle, Francisco, Golde, Todd E, Dickson, Dennis W and Petrucelli, Leonard (2009), Aberrant Cleavage of TDP-43 Enhances Aggregation and Cellular Toxicity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18): 7607–7612.

Zhang, Yong-Jie, Xu, Ya-Fei, Dickey, Chad A, Buratti, Emanuele, Baralle, Francisco, Bailey, Rachel, Pickering-Brown, Stuart, Dickson, Dennis and Petrucelli, Leonard (2007), Progranulin Mediates Caspase-Dependent

Cleavage of TAR DNA Binding Protein-43. *Journal of Neuroscience*, 27(39): 10530–10534.

Zhang, Zhenxi, Lotti, Francesco, Dittmar, Kimberly, Younis, Ihab, Wan, Lili, Kasim, Mumtaz and Dreyfuss, Gideon (2008), SMN Deficiency Causes Tissue-Specific Perturbations in the Repertoire of snRNAs and Widespread Defects in Splicing. *Cell*, 133(4): 585–600.

Zhang, Zhijun, Almeida, Sandra, Lu, Yubing, Nishimura, Agnes L, Peng, Lingtao, Sun, Danqiong, Wu, Bei, Karydas, Anna M, Tartaglia, Maria C, Fong, Jamie C, Miller, Bruce L, Farese, Robert V, Moore, Melissa J, Shaw, Christopher E and Gao, Fen-Biao (2013), Downregulation of microRNA-9 in iPSC-Derived Neurons of FTD/ALS Patients with TDP-43 Mutations. *PloS one*, 8(10): e76055.

Zhou, Hongxia, Huang, Cao, Chen, Han, Wang, Dian, Landel, Carlisle P, Xia, Pedro Yuxing, Bowser, Robert, Liu, Yong-Jian and Xia, Xu Gang (2010), Transgenic Rat Model of Neurodegeneration Caused by Mutation in the TDP Gene. *PLoS genetics*, 6(3): e1000887.

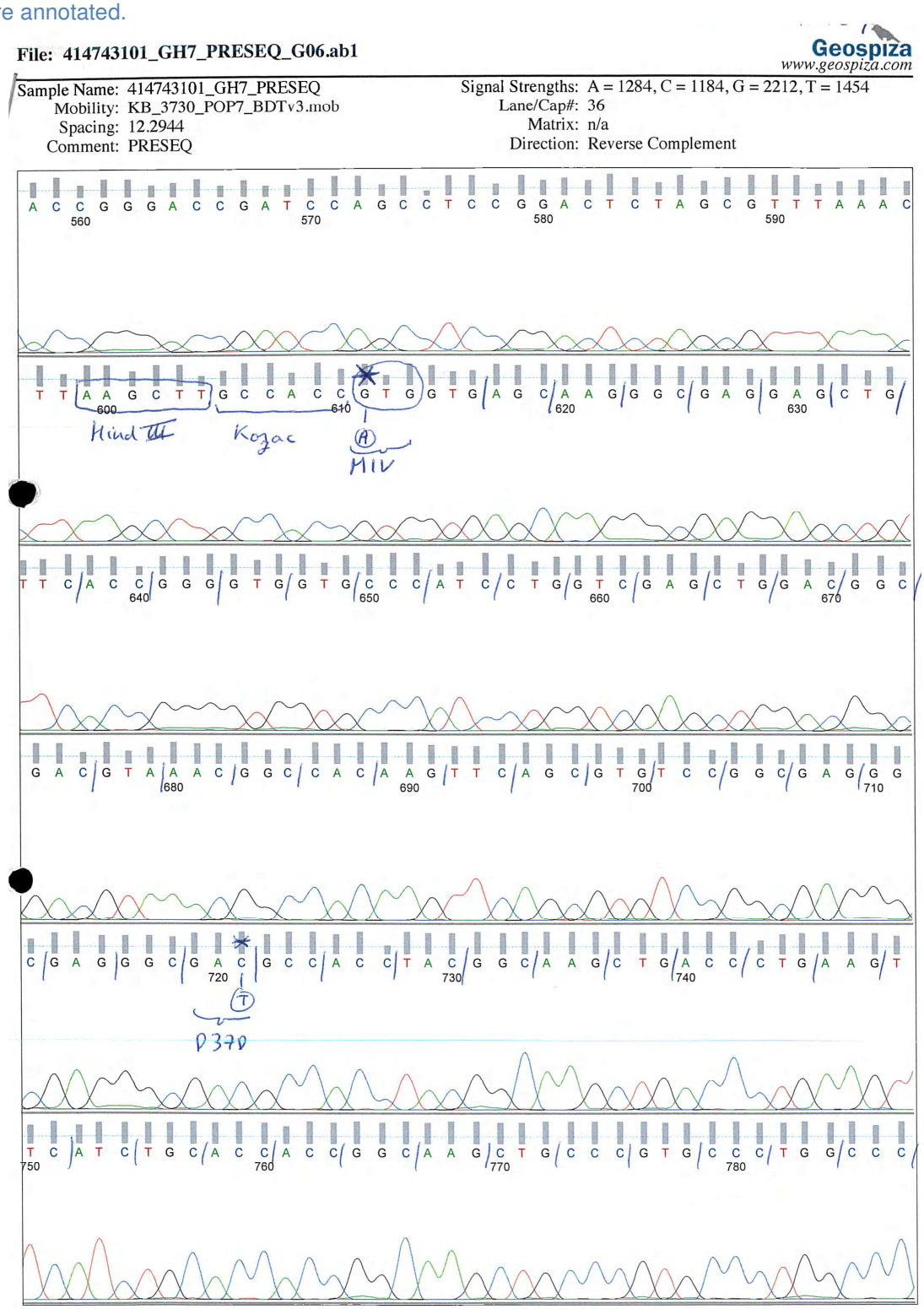
Zhou, Yueqin, Liu, Songyan, Liu, Guodong, Oztürk, Arzu and Hicks, Geoffrey G (2013), ALS-Associated FUS Mutations Result in Compromised FUS Alternative Splicing and Autoregulation. *PLoS genetics*, 9(10): e1003895.

Zinszner, H, Sok, J, Immanuel, D, Yin, Y and Ron, D (1997), TLS (FUS) Binds RNA in Vivo and Engages in Nucleo-Cytoplasmic Shuttling. *Journal of cell science*, 110 (Pt 15): 1741–1750.

Appendix 1

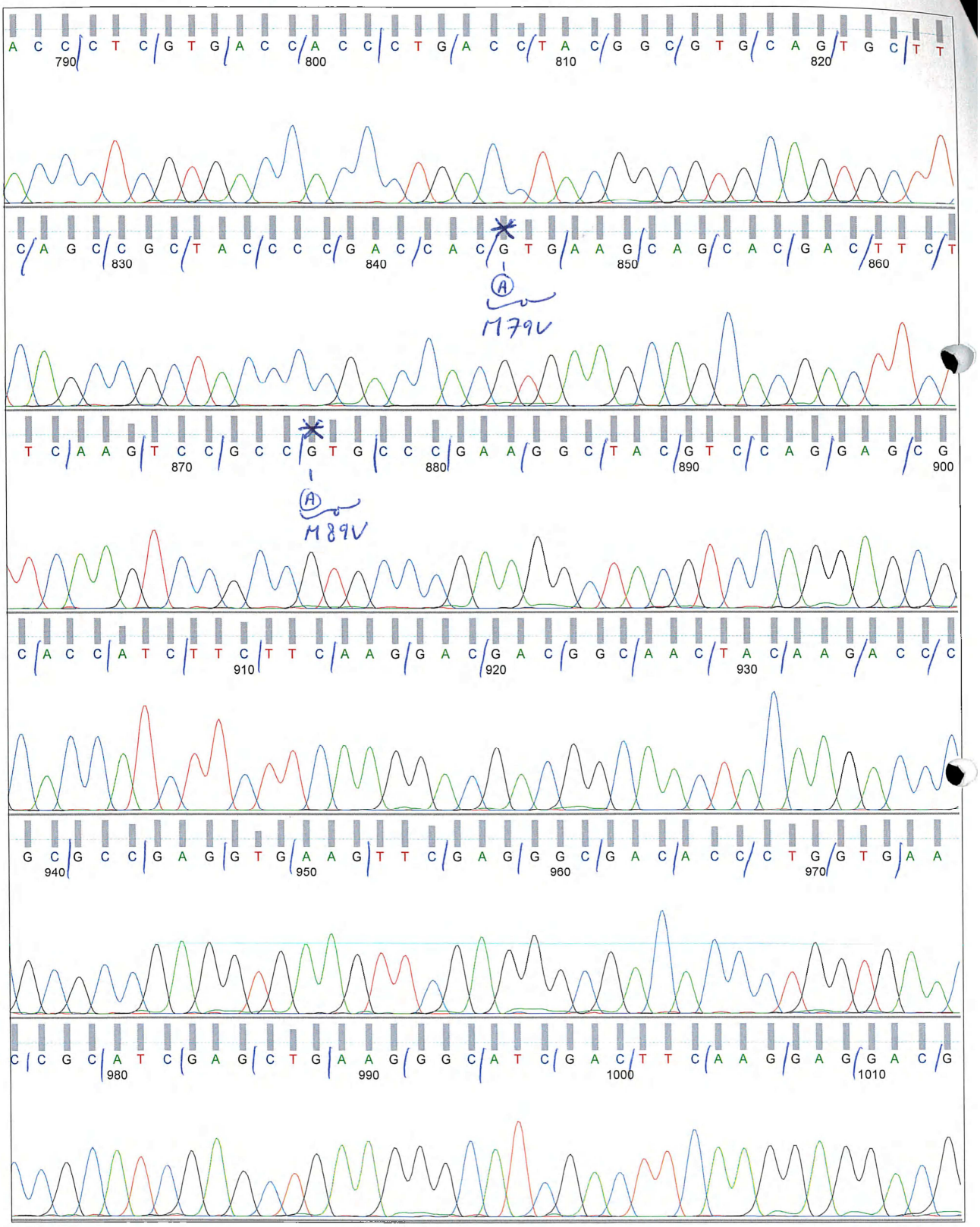
Example sequencing trace from sequencing of pcDNA5 FRT/TO/*M1V* plasmid

The plasmid was generated by site directed mutagenesis of start codon (*M1V*) of pcDNA5 FRT/TO/*GFP* plasmid. The pcDNA5 FRT/TO/*GFP* plasmid was generated by mutation of ATG triplet sequences in all frames in the first two thirds of the GFP sequence (*D37D M79V M89V*) of pcDNA5 FRT/TO/GFP AH. Restriction sites, the Kozak sequence and mutations are annotated.



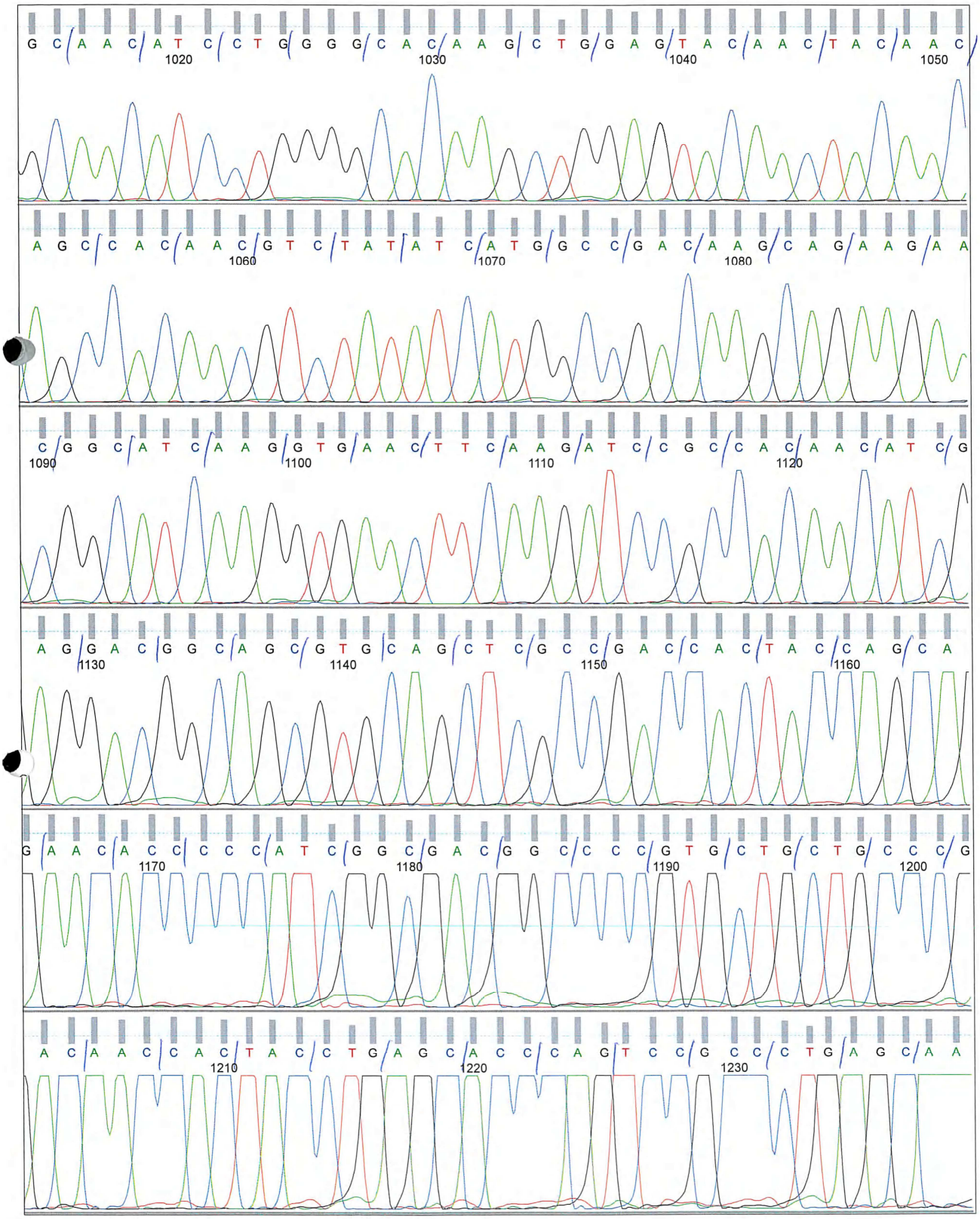
Sample Name: 414743101_GH7_PRESEQ
Mobility: KB_3730_POP7_BDTv3.mob
Spacing: 12.2944
Comment: PRESEQ

Signal Strengths: A = 1284, C = 1184, G = 2212, T = 1454
Lane/Cap#: 36
Matrix: n/a
Direction: Reverse Complement



ample Name: 414743101_GH7_PRESEQ
Mobility: KB_3730_POP7_BDTv3.mob
Spacing: 12.2944
Comment: PRESEQ

Signal Strengths: A = 1284, C = 1184, G = 2212, T = 1454
Lane/Cap#: 36
Matrix: n/a
Direction: Reverse Complement



Appendix 2

Appendix Table A1. Gene ontology analysis of the differentially expressed TDP-43 Q331K whole-cell transcriptome at gene level

The functional annotation chart module of DAVID was used to statistically assess the significance of biological process enrichment scores. GO: gene ontology. N: total number of human genes associated with the given functional category. WCT: number of human genes differentially expressed in the Q331K whole-cell transcriptome (WCT) for the given functional category. FE: fold enrichment. GO terms (52) are ranked by ascending P-Values ($p < 0.005$) for the differentially expressed gene list with $\log_2FC > 1$.

Biological process	GO term	N	WCT	FE	P-Value
RNA processing	GO:0006396	547	67	1.9	3.6E-07
Cell cycle	GO:0007049	776	86	1.7	5.1E-07
Golgi vesicle transport	GO:0048193	131	25	3.0	2.4E-06
Cell cycle process	GO:0022402	565	65	1.8	4.8E-06
mRNA metabolic process	GO:0016071	370	46	2.0	2.2E-05
Intracellular transport	GO:0046907	657	70	1.7	2.6E-05
Establishment of protein localization	GO:0045184	769	79	1.6	2.7E-05
Ubiquitin-dependent protein catabolic process	GO:0006511	242	34	2.2	2.8E-05
Response to DNA damage stimulus	GO:0006974	373	45	1.9	5.6E-05
Cellular macromolecular complex assembly	GO:0034622	318	40	2.0	6.4E-05
Macromolecular complex assembly	GO:0065003	665	69	1.6	6.8E-05
Ribonucleoprotein complex biogenesis	GO:0022613	180	27	2.4	7.5E-05
mRNA processing	GO:0006397	321	40	2.0	7.9E-05
Cellular macromolecular complex subunit organization	GO:0034621	357	43	1.9	8.8E-05
Ribonucleoprotein complex assembly	GO:0022618	69	15	3.4	9.5E-05
Mitotic cell cycle	GO:0000278	370	44	1.9	9.7E-05
Macromolecular complex subunit organization	GO:0043933	710	72	1.6	9.8E-05
Protein transport	GO:0015031	762	76	1.6	1.0E-04
Vesicle-mediated transport	GO:0016192	576	61	1.7	1.1E-04
Macromolecule catabolic process	GO:0009057	781	77	1.54	1.3E-04
DNA metabolic process	GO:0006259	506	55	1.70	1.3E-04
Transcription	GO:0006350	2101	174	1.30	1.5E-04
Modification-dependent macromolecule catabolic process	GO:0043632	574	60	1.64	1.8E-04
Modification-dependent protein catabolic process	GO:0019941	574	60	1.64	1.8E-04
Protein localization	GO:0008104	882	84	1.49	2.0E-04
Intracellular protein transport	GO:0006886	374	43	1.80	2.5E-04
DNA repair	GO:0006281	284	35	1.93	2.9E-04
Cellular macromolecule catabolic process	GO:0044265	725	71	1.53	3.0E-04
Proteolysis involved in cellular protein catabolic process	GO:0051603	600	61	1.59	3.3E-04
Cellular protein catabolic process	GO:0044257	603	61	1.58	3.8E-04
Cellular response to stress	GO:0033554	566	58	1.60	3.9E-04
DNA damage response, signal transduction	GO:0042770	80	15	2.94	4.8E-04
Protein catabolic process	GO:0030163	622	62	1.56	4.9E-04
Cellular protein localization	GO:0034613	411	45	1.71	5.0E-04
Cellular macromolecule localization	GO:0070727	414	45	1.70	5.8E-04
RNA splicing	GO:0008380	284	34	1.87	6.1E-04
Spliceosome assembly	GO:0000245	32	9	4.40	7.1E-04
Nuclear mRNA splicing, via spliceosome	GO:0000398	153	22	2.25	7.1E-04
RNA splicing, via transesterification reactions	GO:0000375	153	22	2.25	7.1E-04
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	GO:0000377	153	22	2.25	7.1E-04
Post-Golgi vesicle-mediated transport	GO:0006892	58	12	3.24	9.6E-04
Cell cycle phase	GO:0022403	414	44	1.66	1.1E-03
Cell cycle arrest	GO:0007050	103	16	2.43	2.2E-03
Translation	GO:0006412	331	36	1.70	2.2E-03
Cellular amide metabolic process	GO:0043603	56	11	3.08	2.6E-03
DNA recombination	GO:0006310	105	16	2.39	2.7E-03

Chromosome organization	GO:0051276	485	48	1.55	2.7E-03
DNA damage checkpoint	GO:0000077	48	10	3.26	3.0E-03
ncRNA metabolic process	GO:0034660	230	27	1.84	3.2E-03
Regulation of transcription	GO:0045449	2601	198	1.19	3.8E-03
Posttranscriptional regulation of gene expression	GO:0010608	211	25	1.86	4.2E-03
Cellular amino acid biosynthetic process	GO:0008652	51	10	3.07	4.5E-03

Appendix Table A2. Gene ontology analysis of the TDP-43 Q331K differentially expressed cytoplasmic transcriptome at gene level

The functional annotation chart module of DAVID () was used to statistically assess the significance of biological process enrichment scores. GO: gene ontology. N: total number of human genes associated with the given functional category. CyT: number of human genes differentially expressed in the Q331K cytoplasmic transcriptome (CyT) for the given functional category. FE: fold enrichment. GO terms (95) are ranked by ascending P-Values ($p < 0.005$) for the differentially expressed gene list with $log_2FC > 1$.

Biological process	GO term	N	CyT	FE	P-Value
Cell cycle	GO:0007049	776	160.0	2.0	1.0E-17
Cell cycle process	GO:0022402	565	124.0	2.1	1.0E-15
Mitotic cell cycle	GO:0000278	370	87.0	2.2	5.8E-13
Cell cycle phase	GO:0022403	414	91.0	2.1	9.0E-12
Cell division	GO:0051301	295	65.0	2.1	1.0E-08
Intracellular transport	GO:0046907	657	116.0	1.7	1.6E-08
Cellular response to stress	GO:0033554	566	103.0	1.7	2.4E-08
M phase	GO:0000279	329	69.0	2.0	2.8E-08
Transcription	GO:0006350	2101	293.0	1.3	5.2E-08
M phase of mitotic cell cycle	GO:0000087	224	52.0	2.2	6.5E-08
Protein localization	GO:0008104	882	141.0	1.5	2.4E-07
Mitosis	GO:0007067	220	50.0	2.2	2.4E-07
Nuclear division	GO:0000280	220	50.0	2.2	2.4E-07
Microtubule-based process	GO:0007017	253	55.0	2.1	2.6E-07
Golgi vesicle transport	GO:0048193	131	35.0	2.5	4.1E-07
Organelle fission	GO:0048285	229	50.0	2.1	8.5E-07
Regulation of transcription	GO:0045449	2601	343.0	1.3	9.8E-07
Establishment of protein localization	GO:0045184	769	123.0	1.5	1.5E-06
Interphase of mitotic cell cycle	GO:0051329	103	29.0	2.7	1.6E-06
Protein transport	GO:0015031	762	121.0	1.5	2.7E-06
Interphase	GO:0051325	106	29.0	2.6	2.9E-06
Response to DNA damage stimulus	GO:0006974	373	69.0	1.8	3.6E-06
DNA metabolic process	GO:0006259	506	87.0	1.6	3.9E-06
Vesicle-mediated transport	GO:0016192	576	96.0	1.6	4.5E-06
Cytoskeleton organization	GO:0007010	436	77.0	1.7	5.5E-06
Microtubule cytoskeleton organization	GO:0000226	147	34.0	2.2	1.8E-05
Modification-dependent macromolecule catabolic process	GO:0043632	574	93.0	1.5	2.0E-05
Modification-dependent protein catabolic process	GO:0019941	574	93.0	1.5	2.0E-05
DNA repair	GO:0006281	284	54.0	1.8	2.2E-05
Protein catabolic process	GO:0030163	622	99.0	1.5	2.2E-05
Spindle organization	GO:0007051	45	16.0	3.4	3.1E-05
Macromolecule catabolic process	GO:0009057	781	118.0	1.4	3.8E-05
Protein modification by small protein conjugation or removal	GO:0070647	160	35.0	2.1	4.5E-05
Post-Golgi vesicle-mediated transport	GO:0006892	58	18.0	3.0	6.2E-05
Macromolecular complex subunit organization	GO:0043933	710	108.0	1.4	6.4E-05
Proteolysis involved in cellular protein catabolic process	GO:0051603	600	94.0	1.5	6.8E-05
Chromosome organization	GO:0051276	485	79.0	1.6	7.7E-05
Cellular protein catabolic process	GO:0044257	603	94.0	1.5	8.2E-05
Cellular macromolecular complex subunit organization	GO:0034621	357	62.0	1.7	8.3E-05
Protein modification by small protein conjugation	GO:0032446	132	30.0	2.2	8.6E-05
Translational elongation	GO:0006414	101	25.0	2.4	9.4E-05
Cellular macromolecule catabolic process	GO:0044265	725	109.0	1.4	9.5E-05
Regulation of microtubule-based process	GO:0032886	49	16.0	3.1	9.6E-05
Regulation of RNA metabolic process	GO:0051252	1813	238.0	1.2	1.1E-04
Translation	GO:0006412	331	58.0	1.7	1.1E-04
Chromatin modification	GO:0016568	274	50.0	1.7	1.3E-04
Translational initiation	GO:0006413	45	15.0	3.2	1.3E-04
Protein complex biogenesis	GO:0070271	505	80.0	1.5	1.8E-04
Protein complex assembly	GO:0006461	505	80.0	1.5	1.8E-04
Macromolecular complex assembly	GO:0065003	665	100.0	1.4	1.9E-04
Regulation of transcription, DNA-dependent	GO:0006355	1773	231.0	1.2	2.2E-04
Negative regulation of macromolecule metabolic process	GO:0010605	734	108.0	1.4	2.3E-04

Regulation of microtubule cytoskeleton organization	GO:0070507	42	14.0	3.2	2.4E-04
Regulation of cell cycle	GO:0051726	331	56.0	1.6	3.8E-04
Apoptosis	GO:0006915	602	90.0	1.4	5.0E-04
Programmed cell death	GO:0012501	611	91.0	1.4	5.2E-04
Protein ubiquitination	GO:0016567	119	26.0	2.1	5.3E-04
Regulation of cell cycle process	GO:0010564	114	25.0	2.1	6.5E-04
ER to Golgi vesicle-mediated transport	GO:0006888	42	13.0	2.9	9.4E-04
Cellular macromolecule localization	GO:0070727	414	65.0	1.5	9.6E-04
Negative regulation of protein metabolic process	GO:0051248	187	35.0	1.8	9.9E-04
Positive regulation of apoptosis	GO:0043065	430	67.0	1.5	9.9E-04
Spindle assembly	GO:0051225	17	8.0	4.5	1.0E-03
Protein amino acid phosphorylation	GO:0006468	667	96.0	1.4	1.1E-03
Positive regulation of programmed cell death	GO:0043068	433	67.0	1.5	1.2E-03
Chromatin organization	GO:0006325	378	60.0	1.5	1.2E-03
Posttranscriptional regulation of gene expression	GO:0010608	211	38.0	1.7	1.2E-03
Cell death	GO:0008219	719	102.0	1.4	1.3E-03
Cellular protein localization	GO:0034613	411	64.0	1.5	1.3E-03
Positive regulation of cell death	GO:0010942	435	67.0	1.5	1.3E-03
Cellular macromolecular complex assembly	GO:0034622	318	52.0	1.6	1.4E-03
Death	GO:0016265	724	102.0	1.3	1.6E-03
Negative regulation of macromolecule biosynthetic process	GO:0010558	547	80.0	1.4	2.0E-03
Negative regulation of cellular protein metabolic process	GO:0032269	180	33.0	1.7	2.0E-03
Regulation of organelle organization	GO:0033043	217	38.0	1.7	2.0E-03
Regulation of mitotic cell cycle	GO:0007346	152	29.0	1.8	2.2E-03
Regulation of protein catabolic process	GO:0042176	52	14.0	2.6	2.3E-03
Induction of apoptosis	GO:0006917	320	51.0	1.5	2.6E-03
Cytokinesis	GO:0000910	41	12.0	2.8	2.6E-03
Maintenance of protein location in cell	GO:0032507	41	12.0	2.8	2.6E-03
Covalent chromatin modification	GO:0016569	126	25.0	1.9	2.8E-03
Anti-apoptosis	GO:0006916	206	36.0	1.7	2.8E-03
Induction of programmed cell death	GO:0012502	321	51.0	1.5	2.8E-03
Organelle localization	GO:0051640	92	20.0	2.1	2.9E-03
Regulation of cellular protein metabolic process	GO:0032268	474	70.0	1.4	3.0E-03
Maintenance of protein location	GO:0045185	48	13.0	2.6	3.3E-03
Phosphorylation	GO:0016310	800	109.0	1.3	3.4E-03
Regulation of microtubule polymerization or depolymerization	GO:0031110	31	10.0	3.1	3.6E-03
Negative regulation of molecular function	GO:0044092	334	52.0	1.5	3.9E-03
Regulation of apoptosis	GO:0042981	804	109.0	1.3	4.0E-03
Intracellular protein transport	GO:0006886	374	57.0	1.5	4.0E-03
Regulation of translation	GO:0006417	137	26.0	1.8	4.1E-03
Ubiquitin-dependent protein catabolic process	GO:0006511	242	40.0	1.6	4.3E-03
Centrosome organization	GO:0051297	32	10.0	3.0	4.5E-03
Chromatin remodeling	GO:0006338	56	14.0	2.4	4.5E-03

Appendix Table A3. Gene ontology analysis of the TDP-43 Q331K differentially expressed GRASPS-translatome at gene level

The functional annotation chart module of DAVID was used to statistically assess the significance of biological process enrichment scores. GO: gene ontology. N: total number of human genes associated with the given functional category. GRASPS: number of human genes differentially expressed in the Q331K GRASPS-translatome for the given functional category. FE: fold enrichment. GO terms (172) are ranked by ascending P-Values ($p < 0.005$) for the differentially expressed gene list with $log_2FC > 1$.

Biological process	GO term	N	GRASPS	FE	P-Value
Cell cycle	GO:0007049	776	227	1.9	7.8E-23
Cell cycle process	GO:0022402	565	176	2.0	6.9E-21
Mitotic cell cycle	GO:0000278	370	129	2.2	6.2E-20
Cellular macromolecule catabolic process	GO:0044265	725	204	1.8	1.8E-18
Macromolecule catabolic process	GO:0009057	781	214	1.7	6.4E-18
Modification-dependent macromolecule catabolic process	GO:0043632	574	167	1.9	2.2E-16
Modification-dependent protein catabolic process	GO:0019941	574	167	1.9	2.2E-16
Protein catabolic process	GO:0030163	622	177	1.8	2.4E-16
Ubiquitin-dependent protein catabolic process	GO:0006511	242	91	2.4	2.4E-16
Proteolysis involved in cellular protein catabolic process	GO:0051603	600	172	1.8	2.8E-16
Cellular protein catabolic process	GO:0044257	603	172	1.8	5.0E-16
Positive regulation of ligase activity	GO:0051351	73	41	3.6	1.4E-14
Positive regulation of ubiquitin-protein ligase activity	GO:0051443	70	40	3.6	1.5E-14
Regulation of ubiquitin-protein ligase activity during mitotic cell cycle	GO:0051439	71	40	3.6	2.8E-14
Positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	GO:0051437	68	39	3.7	2.8E-14
Positive regulation of protein ubiquitination	GO:0031398	84	43	3.3	2.1E-13
Regulation of ligase activity	GO:0051340	81	42	3.3	2.4E-13
Regulation of ubiquitin-protein ligase activity	GO:0051438	78	41	3.3	2.7E-13
Cellular response to stress	GO:0033554	566	156	1.8	3.1E-13
Proteasomal protein catabolic process	GO:0010498	102	48	3.0	3.8E-13
Proteasomal ubiquitin-dependent protein catabolic process	GO:0043161	102	48	3.0	3.8E-13
Intracellular transport	GO:0046907	657	174	1.7	5.2E-13
Regulation of protein ubiquitination	GO:0031396	100	47	3.0	7.3E-13
Anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	GO:0031145	65	36	3.5	1.3E-12
Negative regulation of protein ubiquitination	GO:0031397	74	38	3.3	5.7E-12
Negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle	GO:0051436	65	35	3.4	8.0E-12
Translation	GO:0006412	331	101	1.9	1.4E-11
Negative regulation of protein modification process	GO:0031400	119	50	2.7	2.0E-11
Negative regulation of ligase activity	GO:0051352	67	35	3.3	2.4E-11
Negative regulation of ubiquitin-protein ligase activity	GO:0051444	67	35	3.3	2.4E-11
DNA metabolic process	GO:0006259	506	137	1.7	4.0E-11
Chromatin modification	GO:0016568	274	87	2.0	4.4E-11
Cell cycle phase	GO:0022403	414	117	1.8	7.2E-11
M phase of mitotic cell cycle	GO:0000087	224	74	2.1	2.0E-10
Mitosis	GO:0007067	220	73	2.1	2.1E-10
Nuclear division	GO:0000280	220	73	2.1	2.1E-10
Response to DNA damage stimulus	GO:0006974	373	106	1.8	4.3E-10
Cell division	GO:0051301	295	89	1.9	5.0E-10
Organelle fission	GO:0048285	229	74	2.1	6.0E-10
Negative regulation of protein metabolic process	GO:0051248	187	64	2.2	7.6E-10

Negative regulation of cellular protein metabolic process	GO:0032269	180	61	2.2	3.0E-09
Regulation of cellular protein metabolic process	GO:0032268	474	123	1.7	7.1E-09
M phase	GO:0000279	329	93	1.8	7.6E-09
Negative regulation of catalytic activity	GO:0043086	277	80	1.8	3.4E-08
Negative regulation of macromolecule metabolic process	GO:0010605	734	171	1.5	3.6E-08
Chromatin organization	GO:0006325	378	101	1.7	3.8E-08
DNA repair	GO:0006281	284	81	1.8	5.1E-08
Regulation of protein modification process	GO:0031399	295	82	1.8	1.4E-07
Chromosome organization	GO:0051276	485	120	1.6	1.9E-07
Protein modification by small protein conjugation or removal	GO:0070647	160	52	2.1	2.3E-07
Covalent chromatin modification	GO:0016569	126	44	2.2	2.4E-07
Positive regulation of protein modification process	GO:0031401	187	58	2.0	2.5E-07
Regulation of cell cycle	GO:0051726	331	88	1.7	4.0E-07
Negative regulation of molecular function	GO:0044092	334	88	1.7	6.0E-07
Proteolysis	GO:0006508	1054	224	1.4	6.3E-07
Positive regulation of macromolecule metabolic process	GO:0010604	857	188	1.4	6.5E-07
Positive regulation of protein metabolic process	GO:0051247	243	69	1.8	6.6E-07
Histone modification	GO:0016570	122	42	2.2	7.3E-07
Regulation of programmed cell death	GO:0043067	812	179	1.4	9.1E-07
Protein ubiquitination	GO:0016567	119	41	2.2	9.9E-07
Protein modification by small protein conjugation	GO:0032446	132	44	2.1	1.0E-06
Regulation of apoptosis	GO:0042981	804	177	1.4	1.1E-06
Regulation of cell death	GO:0010941	815	179	1.4	1.2E-06
Golgi vesicle transport	GO:0048193	131	43	2.1	2.2E-06
Positive regulation of cellular protein metabolic process	GO:0032270	233	65	1.8	2.8E-06
Nucleotide-excision repair	GO:0006289	55	24	2.8	3.2E-06
Cell death	GO:0008219	719	159	1.4	3.3E-06
Vesicle-mediated transport	GO:0016192	576	132	1.5	3.8E-06
Death	GO:0016265	724	159	1.4	5.0E-06
Apoptosis	GO:0006915	602	136	1.4	5.9E-06
Negative regulation of programmed cell death	GO:0043069	359	89	1.6	7.7E-06
Negative regulation of apoptosis	GO:0043066	354	88	1.6	7.8E-06
Programmed cell death	GO:0012501	611	137	1.4	8.1E-06
Negative regulation of cell death	GO:0060548	360	89	1.6	8.6E-06
Regulation of organelle organization	GO:0033043	217	60	1.8	9.4E-06
Protein localization	GO:0008104	882	186	1.3	1.1E-05
Anti-apoptosis	GO:0006916	206	57	1.8	1.6E-05
Protein transport	GO:0015031	762	163	1.4	1.8E-05
Macromolecular complex assembly	GO:0065003	665	145	1.4	2.0E-05
Cellular macromolecule localization	GO:0070727	414	98	1.5	2.0E-05
Establishment of protein localization	GO:0045184	769	164	1.4	2.0E-05
Cytoskeleton organization	GO:0007010	436	102	1.5	2.3E-05
Cell proliferation	GO:0008283	436	102	1.5	2.3E-05
Cellular protein localization	GO:0034613	411	97	1.5	2.5E-05
mRNA metabolic process	GO:0016071	370	89	1.5	2.6E-05
Macromolecular complex subunit organization	GO:0043933	710	152	1.4	3.4E-05
Protein complex assembly	GO:0006461	505	114	1.4	3.6E-05
Protein complex biogenesis	GO:0070271	505	114	1.4	3.6E-05

Positive regulation of molecular function	GO:0044093	586	129	1.4	3.8E-05
Regulation of protein catabolic process	GO:0042176	52	21	2.6	5.9E-05
Posttranscriptional regulation of gene expression	GO:0010608	211	56	1.7	6.7E-05
Intracellular protein transport	GO:0006886	374	88	1.5	6.9E-05
Phosphorylation	GO:0016310	800	166	1.3	8.1E-05
Membrane organization	GO:0016044	381	89	1.5	8.2E-05
Positive regulation of catalytic activity	GO:0043085	520	115	1.4	8.6E-05
RNA splicing	GO:0008380	284	70	1.6	9.6E-05
mRNA processing	GO:0006397	321	77	1.5	1.1E-04
RNA processing	GO:0006396	547	119	1.4	1.3E-04
Nuclear mRNA splicing, via spliceosome	GO:0000398	153	43	1.8	1.4E-04
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	GO:0000377	153	43	1.8	1.4E-04
RNA splicing, via transesterification reactions	GO:0000375	153	43	1.8	1.4E-04
Response to UV	GO:0009411	59	22	2.4	1.4E-04
DNA catabolic process	GO:0006308	60	22	2.3	1.9E-04
Protein amino acid N-linked glycosylation	GO:0006487	44	18	2.6	1.9E-04
Microtubule cytoskeleton organization	GO:0000226	147	41	1.8	2.4E-04
Regulation of catabolic process	GO:0009894	96	30	2.0	2.5E-04
Translational initiation	GO:0006413	45	18	2.5	2.7E-04
Protein amino acid phosphorylation	GO:0006468	667	139	1.3	2.7E-04
Phosphate metabolic process	GO:0006796	973	193	1.3	2.8E-04
Phosphorus metabolic process	GO:0006793	973	193	1.3	2.8E-04
Positive regulation of I-kappaB kinase/NF-kappaB cascade	GO:0043123	97	30	2.0	3.1E-04
Regulation of I-kappaB kinase/NF-kappaB cascade	GO:0043122	107	32	1.9	3.6E-04
Interphase of mitotic cell cycle	GO:0051329	103	31	1.9	4.0E-04
Regulation of cytoskeleton organization	GO:0051493	136	38	1.8	4.0E-04
Protein import into nucleus, docking	GO:0000059	17	10	3.7	4.2E-04
Acetyl-CoA metabolic process	GO:0006084	31	14	2.9	4.3E-04
Nucleocytoplasmic transport	GO:0006913	156	42	1.7	4.5E-04
Regulation of mitotic cell cycle	GO:0007346	152	41	1.7	5.1E-04
Nuclear transport	GO:0051169	158	42	1.7	5.9E-04
Histone acetylation	GO:0016573	48	18	2.4	6.5E-04
Regulation of cellular component size	GO:0032535	271	64	1.5	6.7E-04
Interphase	GO:0051325	106	31	1.9	6.8E-04
Regulation of cell morphogenesis	GO:0022604	131	36	1.7	8.2E-04
Microtubule-based process	GO:0007017	253	60	1.5	9.0E-04
Nucleotide-excision repair, DNA damage removal	GO:0000718	22	11	3.2	9.3E-04
Hexose metabolic process	GO:0019318	192	48	1.6	9.8E-04
Positive regulation of RNA metabolic process	GO:0051254	481	102	1.3	1.0E-03
Positive regulation of transcription, DNA-dependent	GO:0045893	477	101	1.3	1.2E-03
Nitrogen compound biosynthetic process	GO:0044271	325	73	1.4	1.2E-03
DNA replication	GO:0006260	190	47	1.6	1.4E-03
Positive regulation of transcription	GO:0045941	564	116	1.3	1.5E-03
Biopolymer methylation	GO:0043414	69	22	2.0	1.5E-03
Protein amino acid acetylation	GO:0006473	52	18	2.2	1.8E-03
Cellular macromolecular complex subunit organization	GO:0034621	357	78	1.4	1.8E-03
Regulation of transcription from RNA polymerase II promoter	GO:0006357	727	144	1.3	1.9E-03
Negative regulation of cellular component organization	GO:0051129	142	37	1.7	1.9E-03

Regulation of cellular response to stress	GO:0080135	103	29	1.8	2.0E-03
Positive regulation of gene expression	GO:0010628	581	118	1.3	2.1E-03
Androgen receptor signaling pathway	GO:0030521	36	14	2.5	2.2E-03
Regulation of Ras GTPase activity	GO:0032318	104	29	1.8	2.3E-03
Regulation of gene-specific transcription	GO:0032583	134	35	1.7	2.5E-03
Steroid hormone receptor signaling pathway	GO:0030518	58	19	2.1	2.6E-03
Positive regulation of programmed cell death	GO:0043068	433	91	1.3	2.6E-03
Regulation of cell size	GO:0008361	206	49	1.5	2.6E-03
Cellular macromolecular complex assembly	GO:0034622	318	70	1.4	2.6E-03
Release of cytochrome c from mitochondria	GO:0001836	21	10	3.0	2.8E-03
Golgi organization	GO:0007030	21	10	3.0	2.8E-03
Regulation of translational initiation	GO:0006446	41	15	2.3	2.8E-03
Positive regulation of cell death	GO:0010942	435	91	1.3	3.0E-03
Translational elongation	GO:0006414	101	28	1.8	3.0E-03
Microtubule-based transport	GO:0010970	29	12	2.6	3.1E-03
Positive regulation of apoptosis	GO:0043065	430	90	1.3	3.1E-03
Transcription from RNA polymerase II promoter	GO:0006366	234	54	1.5	3.1E-03
Negative regulation of organelle organization	GO:0010639	82	24	1.9	3.1E-03
Positive regulation of gene-specific transcription	GO:0043193	87	25	1.8	3.2E-03
Ribonucleotide metabolic process	GO:0009259	147	37	1.6	3.6E-03
Regulation of translation	GO:0006417	137	35	1.6	3.6E-03
Endocytosis	GO:0006897	220	51	1.5	3.7E-03
Membrane invagination	GO:0010324	220	51	1.5	3.7E-03
Positive regulation of cell development	GO:0010720	69	21	1.9	3.7E-03
Actin cytoskeleton organization	GO:0030036	226	52	1.5	4.0E-03
Glucose metabolic process	GO:0006006	153	38	1.6	4.0E-03
Cell cycle arrest	GO:0007050	103	28	1.7	4.1E-03
Purine ribonucleotide metabolic process	GO:0009150	138	35	1.6	4.1E-03
Positive regulation of nitrogen compound metabolic process	GO:0051173	644	127	1.3	4.1E-03
Nucleobase, nucleoside, nucleotide and nucleic acid transport	GO:0015931	113	30	1.7	4.2E-03
Regulation of GTPase activity	GO:0043087	123	32	1.7	4.2E-03
Generation of precursor metabolites and energy	GO:0006091	313	68	1.4	4.2E-03
G1/S transition of mitotic cell cycle	GO:0000082	56	18	2.0	4.3E-03
Chromatin remodeling	GO:0006338	56	18	2.0	4.3E-03
Transcription, DNA-dependent	GO:0006351	292	64	1.4	4.5E-03
Cellular response to oxidative stress	GO:0034599	43	15	2.2	4.6E-03

Appendix Table A4. Altered biological processes and RNA-binding targets of TDP-43 in the differentially expressed TDP-43 Q331K whole cell transcriptome

Major biological processes are ranked by ascending P-Values (p <0.005) for the differentially expressed gene list with log_2FC >1 and FDR <0.2. A total of 384 DEG transcripts (31 %) bound by TDP-43 were identified in the differentially expressed list (1233 gene transcripts are differentially expressed in the whole cell transcriptome). Gene ontology-clustered RNA-binding targets of TDP-43 are highlighted in red.

Biological process	Count	%	P-Value	Official gene symbols
RNA splicing / processing:				ADAT1, APLP1, APP, C1D, CELF1, CPSF4, DEDD2, DICER1, DUS1L, DUS4L, ELAVL1, FBL, FIP1L1, FTSJ2, GEMIN4, GEMIN8, GSPT1, HNRNPF, HNRNPUL1, HSPA1A, INTS10, KHSRP, METTL1, MLH1, MOV10, MTO1, NHP2, NOLC1, PA2G4, PABPN1, PAIP1, PAN2, PCBP2, POLR2B, PPIG, PRMT5, PRMT7, PRPF39, PUF60, RALY, RBM10, RBM28, RBM3, RBM4, RBM5, RBM6, RBMX, RNPC3, RPP14, RPS15, RPS19, RSL1D1, SART3, SF3A1, SF3A2, SF3A3, SNRNP200, SNRNP25, SNRPD1, SNRPG, SPOP, TBL3, TFIP11, TRMT11, TXNL4A, UPF3A, UTP14A, WDR36, WDR83, YBX1, YTHDC1, ZCCHC11, ZCCHC8
GO:0006396	67	6.0	3.6E-07	
GO:0016071	46	4.1		
GO:0006397	40	3.6	2.2E-05	
GO:0008380	34	3.1		
GO:0000375	22	2.0	7.9E-05	
GO:0000377	22	2.0		
GO:0000398	22	2.0	6.0E-04	
			7.1E-04	
			7.1E-04	
			7.1E-04	
Cell cycle related:				ANAPC1, ANAPC7, APC, APP, AURKB, BCAT1, BTRC, CALM2, CCNK, CDCA5, CDK20, CENPE, CEP250, CETN3, CHMP1B, CUL3, DDIT3, DDX11, DMTF1, DTYMK, E2F6, EID1, EIF4G2, ESCO2, FBXO31, GADD45A, GSPT1, HAUS5, HINFP, ITGB1, JMY, KIF15, KIF20B, KRT18, LIG1, MACF1, MAP2K6, MAP9, MCM2, MCM8, MDM2, MLH1, MPHOSPH9, MSH2, MSH6, NASP, NBN, NCAPD3, NOLC1, PA2G4, PARD6B, PARD6G, PHGDH, PKD1, POLD1, POLE, PPP5C, PRC1, PRUNE2, PSMA2, PSMA3, PSMC3, PSMD1, PSRC1, RACGAP1, RAD1, RAD17, RAD21, RAD52, RASSF1, SMC3, SPAG5, SPAST, STAG2, STRADA, SUN2, SUPT5H, TAF1, TP53BP2, TP73, TRIP13, TUBE1, TXNL4A, UBE2D1, WDR6, ZWILCH
GO:0007049	86	7.7	5.1E-07	
GO:0022402	65	5.9		
GO:0000278	44	4.0	4.8E-06	
GO:0022403	44	4.0		
GO:0007050	16	1.4	9.7E-05	
			1.1E-03	
			2.2E-03	
Golgi / vesicle transport:				AP1B1, AP1S3, AP2B1, APLP1, APP, ARCN1, ARRB2, ATL2, ATP5B, ATP6V1H, BET1, CADPS, CADPS2, CLTC, COG2, COPG2, COPZ1, CXCL16, DOPEY1, ELMOD2, EPS15L1, EXOC7, FLNA, GULP1, HGS, IGF2R, KRT18, LIN7A, LMBR1L, LRP4, NME1-NME2, OPTN, PACSIN2, PICALM, PREB, PUM1, RAB18, RAB2A, RAB5A, RHOBTB3, SCFD1, SCFD2, SEC16A, SEC23B, SEC24C, SEC31A, SH3GL3, SNAP91, SORL1, SPAST, SPTBN4, STEAP2, STX10, STX3, STX4, SYNRG, TRIM36, VPS13A, VTI1B, WASF2, WDR19
GO:0048193	25	2.3	2.4E-06	
GO:0016192	61	5.5		
GO:0006892	12	1.1	1.1E-04	
			9.6E-04	
Protein transport/ localization:				AP1B1, AP1S3, AP2B1, APP, ARCN1, ATG16L1, ATG4B, ATL2, ATP2C1, AURKB, BET1, CADPS, CADPS2, CENPE, CEP250, CHMP1B, CLIC5, CLTC, COG2, COPG2, COPZ1, DOPEY1, DSCR3, EIF4ENIF1, EXOC7, EXPH5, FGF2, FLNA, FRAS1, GRIP1, HDAC6, HGS, HINFP, IPO11, IPO4, KIF13A, KPNA4, KPNA5, KRT18, LIN7A, MACF1, MGEA5, MYL6, MYO5A, NASP, NECAB3, NFKBIA, NPC2, NUP160, NUP98, OPTN, PABPN1, POM121, PREB, PRICKLE1, PUM1, RAB11FIP3, RAB18, RAB2A, RAB5A, RANBP3, RERE, RHOBTB3, RIMS2, RPS15, SCAMP4, SCFD1, SCFD2, SDAD1, SEC16A, SEC23B, SEC24C, SEC31A, SLC25A27, SNF8, SNX14, SPAST, SPTBN4, SRP14, SRP19, SSR2, STEAP2, STRADA, STX10, STX3, STX4, SYNRG, TIMM17B, TIMM23, TIMM23B, TNPO2, TPM1, VPS13A, VPS26B, VPS53, VTI1B, WASF2, WDR19, XPO6, YWHAE
GO:0046907	70	6.3	2.6E-05	
GO:0045184	79	7.1		
GO:0015031	76	6.8	2.7E-05	
GO:0008104	84	7.6		
GO:0006886	43	3.9	1.0E-04	
GO:0034613	45	4.1		
GO:0070727	45	4.1	2.0E-04	
			2.5E-04	
			5.0E-04	
			5.8E-	

			04	
Protein catabolic process and ubiquitin-dependent protein degradation: GO:0006511 GO:0009057 GO:0019941 GO:0043632 GO:0044265 GO:0051603 GO:0044257 GO:0030163	 34 77 60 60 71 61 61 62	 3.1 6.9 5.4 5.4 6.4 5.5 5.5 5.6	 2.8E-05 1.3E-04 1.8E-04 1.8E-04 3.0E-04 3.2E-04 3.8E-04 4.9E-04	ANAPC1, ANAPC7, ASB13, ASB3, ATE1, ATG4B, BTRC, CBLL1, CUL3, CYCS, DCUN1D2, DEDD2, ENDOG, ERLEC1, ERLIN2, FBXL12, FBXO16, FBXO31, FBXO4, FBXW2, FOXRED2, GNS, GSPT1, GUSB, HACE1, HDAC6, HECTD1, HLTF, HSPA1A, KDM2B, LRRC41, MARCH6, MDM2, MGEA5, MLH1, NGLY1, NSMCE2, OVGP1, PAN2, PSMA2, PSMA3, PSMC3, PSMD1, RNF14, RNF216, RNF40, SENP3, SMURF2, SPOP, SUMO1, TXNDC12, UBA5, UBE2A, UBE2D1, UBE2G1, UBE2S, UBE2V2, UBXN6, UCHL3, UPF3A, USP3, USP5, USP8, USP19, USP20, USP28, USP36, USP37, USP44, USP45, USP47, WDR48, WSB1, XPC, YME1L1, ZCCHC11, ZRANB1
DNA damage/ repair response: GO:0006974 GO:0006281 GO:0042770 GO:0000077	 45 35 15 10	 4.1 3.2 1.4 0.9	 5.6E-05 2.9E-04 4.8E-04 3.0E-03	ANKRD17, APC, BRSK1, DDIT3, ESCO2, FAM175A, FBXO31, FOXO3, GADD45A, HINFP, HMGB1, IGHMBP2, JMY, LIG1, MAP2K6, MICA, MLH1, MORF4L1, MSH2, MSH6, NBN, NSMCE2, NUPR1, POLD1, POLE, POLH, PTPN11, RAD1, RAD17, RAD21, RAD52, RBM4, RFC2, SMC3, SMC6, SUMO1, TDP1, TP53BP1, TP73, TRIP13, TXNDC12, UBE2A, UBE2V2, XPC, ZSWIM7
Macromolecular / ribonucleoprotein complex assembly: GO:0034622 GO:0065003 GO:0022613 GO:0034621 GO:0022618 GO:0043933 GO:0000245	 40 69 27 43 15 72 9	 3.6 6.2 2.4 3.9 1.4 6.5 0.8	 6.3E-05 6.8E-05 7.5E-05 8.8E-05 9.5E-05 9.8E-05 7.1E-04	ABI2, ADSL, APC, ATG16L1, ATL2, ATPAF1, ATPIF1, BCS1L, C1D, CELF1, CENPE, DAG1, DGKH, DIAPH1, DICER1, EBNA1BP2, EIF2AK3, FBL, FLNA, FTSJ2, GEMIN4, GEMIN8, GRB2, H2AFY, HIST1H2BJ, HIST1H4B, HIST2H2AB, HIST2H4A, HP1BP3, IGHMBP2, IPO11, IPO4, LIN7A, MAGI1, MCM2, MDM2, MED1, MED16, MIF, MLH1, MRRF, NAP1L1, NASP, NCK2, NHP2, NOLC1, NUP98, PA2G4, PARD6B, PDSS1, PICALM, POLR2B, PPP5C, PRKAB1, PRKAG1, PRMT5, PRMT7, RBM5, RPS15, RPS19, RSF1, SBF2, SDAD1, SF3A1, SF3A2, SF3A3, SKIL, SNAP91, SNRNP200, SNRPD1, SNRPG, SPAST, STX4, TAF1, TAF13, TAF6, TBL3, TFAM, TNPO2, TUBE1, TXNL4A, UTP14A, WASF1, WDR36, XPO6
DNA metabolic process: GO:0006259	 55	 5.0	 1.3E-04	ANKRD17, CYCS, ENDOG, ESCO2, FAM175A, FBXO4, GADD45A, GIN1, HINFP, HMGB1, IGHMBP2, JMY, LIG1, MCM2, MCM8, MED1, MLH1, MORF4L1, MSH2, MSH6, NAP1L1, NASP, NBN, NSMCE2, NUP98, POLD1, POLE, POLH, PPIA, PRIM2, PRMT7, RAD1, RAD17, RAD21, RAD52, RBM4, RFC2, RMI1, SMARCAL1, SMC3, SMC6, SUMO1, TDP1, TFAM, TK2, TP53BP1, TP73, TRIP13, TSN, TXNDC12, UBE2A, UBE2V2, WRAP53, XPC, ZSWIM7
Transcription, transcriptional regulation:				ABCA2, ACVR2B, AES, AFF1, AFF3, APP, ARID2, ARID4B, ASCC3, ASXL1, ATF6B, ATF7, BACH1, BACH2, BCOR, BMP3, C1D, CAMTA1, CBFA2T2, CCNK, CD3EAP, CHD1, CHD2, CHD9, CNOT6, COPS2, CRABP2, CREG1, DDIT3, DEDD, DEDD2,

GO:0006350 GO:0045449	174 198	15.7 17.8	1.5E-04 3.8E-03	DEK, DENND4A, DMTF1, DNAJB6, DPF2, DYRK1B, E2F6, EDA2R, EED, EID1, ELF4, ELP2, ENO1, ERF, FAM120B, FGF2, FLNA, FOXC1, FOXO3, GRAMD4, GRIP1, GTF3C3, HDAC6, HINFP, HKR1, HLTF, HMG20A, HMGB1, HMGN5, HNRNPUL1, HTATSF1, IGHMBP2, JARID2, JMJD1C, JMY, KDM2B, KDM5C, KHSRP, KLF4, LRCH4, LRRFIP1, MAFB, MBTD1, MCM2, MCM8, MDM2, MECP2, MED1, MED16, MED20, MED25, MED29, MEIS1, MKL1, MORF4L1, MOV10, NAA15, NAB1, NCOA3, NCOA4, NCOA5, NCOR2, NFAT5, NFATC3, NFE2L1, NFKBIA, NFYB, NFYC, NME1-NME2, NR2C1, PA2G4, PBX1, PDLIM1, PHB2, PHTF1, PITX2, POLR1C, POLR1D, POLR2B, PPP5C, PRDM10, PREB, PRIM2, PRKAR1A, PRMT5, PRMT7, PSRC1, PTMA, PTOV1, PUF60, RAB18, RASSF7, RBM4, RERE, RNF14, RSF1, SAP18, SARNP, SCAI, SETD1B, SKIL, SLTM, SMAD7, SMARCAL1, SMURF2, SNF8, SUMO1, SUPT4H1, SUPT5H, TAF1, TAF13, TAF6, TCERG1, TCF20, TCF25, TCOF1, TEAD4, TFAM, TFCP2, TGIF2, TIGD1, TLE1, TLE4, TP53BP1, TP73, TRIP13, TROVE2, TSC22D1, TSC22D4, VGLL4, YBX1, ZBTB25, ZBTB44, ZFAT, ZFP30, ZFP62, ZFP82, ZFX, ZHX1, ZMYM2, ZNF117, ZNF124, ZNF131, ZNF142, ZNF146, ZNF181, ZNF182, ZNF225, ZNF236, ZNF28, ZNF3, ZNF302, ZNF362, ZNF394, ZNF48, ZNF500, ZNF506, ZNF525, ZNF529, ZNF530, ZNF532, ZNF544, ZNF562, ZNF565, ZNF568, ZNF583, ZNF585A, ZNF607, ZNF682, ZNF701, ZNF717, ZNF77, ZNF780A, ZNF787, ZNF79, ZNF808, ZNF91, ZZZ3
Cellular response to stress: GO:0033554	58	5.2	3.9E-04	ADM, ANKRD17, APC, ATG16L1, BRSK1, DDIT3, EIF2AK2, EIF2AK3, ESCO2, FAM129A, FAM175A, FBXO31, FOXO3, GADD45A, HDAC6, HINFP, HMGB1, IGHMBP2, JMY, LIG1, MAP2K4, MAP2K6, MAP3K4, MAPK9, MBIP, MICA, MLH1, MORF4L1, MSH2, MSH6, NBN, NSMCE2, NUPR1, POLD1, POLE, POLH, PTPN11, RAD1, RAD17, RAD21, RAD52, RBM4, RFC2, SMC3, SMC6, STAC, SUMO1, TDP1, TP53BP1, TP73, TPM1, TRIP13, TXNDC12, UBE2A, UBE2V2, VAPB, XPC, ZSWIM7
Translation: GO:0006412	36	3.2	2.2E-03	EIF2AK2, EIF2AK3, EIF3G, EIF4A1, EIF4E, EIF4G2, FARSA, GSPT1, IGHMBP2, LARS2, MARS, MRPL13, MRPL51, MRPS18A, MRPS33, MRPS7, MRRF, NCOA5, PAIP1, QARS, RARS, RBM3, RPL17, RPL19, RPL21, RPL26L1, RPL27, RPL27A, RPL31, RPL34, RPL9, RPS15, RPS19, RPS5, RSL1D1, SECISBP2
Cellular amide/ amino acid metabolic process: GO:0043603 GO:0008652	 11 10	 1.0 0.9	 2.6E-03 4.5E-03	ALDH4A1, ASL, ASS1, BCAT1, DCXR, DHFRL1, DLD, ENOPH1, GLYR1, IDH1, IDH3B, LDHB, MCCC1, NADSYN1, PHGDH, PLOD1, PSPH, SEPHS2, TPI1
DNA recombination: GO:0006310	16	1.4	2.6E-03	HMGB1, IGHMBP2, LIG1, MLH1, MSH2, MSH6, NBN, NSMCE2, PPIA, RAD21, RAD52, RBM4, SMC6, TRIP13, TSN, ZSWIM7
Chromosome organization: GO:0051276	48	4.3	2.7E-03	ACIN1, APC, ARID2, ARID4B, BCOR, CDCA5, CENPE, CHD1, CHD2, CHD9, DAPK3, DDX11, EED, FAM175A, FBXO4, H2AFY, HDAC6, HIST1H2BJ, HIST1H4B, HIST2H2AB, HIST2H4A, HLTF, HMG20A, HMGB1, HP1BP3, JMJD1C, KDM2B, KDM5C, MCM2, MORF4L1, MSH2, MSH6, NAP1L1, NASP, NBN, NCAPD3, PRMT5, PRMT7, RBM4, RERE, RNF40, RSF1, SETD1B, SMARCAL1, SMC3, SUPT4H1, SUPT5H, UBE2A, WRAP53
ncRNA metabolic process: GO:0034660	27	2.4	3.2E-03	ADAT1, C1D, DEDD2, DUS1L, DUS4L, FARSA, FBL, FTSJ2, GEMIN4, INTS10, LARS2, MARS, METTL1, MTO1, NHP2, NOLC1, PA2G4, QARS, RARS, RPP14, RPS15, RPS19, TBL3, TRMT11, UTP14A, WDR36, ZCCHC11
Posttranscriptional regulation of gene expression: GO:0010608	25	2.3	4.2E-03	ANKHD1, APLP1, APP, CELF1, DICER1, EIF2AK3, EIF4E, EIF4G2, ELAVL1, FAM129A, FLNA, MOV10, NANOS1, NCK2, PA2G4, PAIP1, PUM1, RBM3, RPS5, SARNP, SMAD7, SRP14, TNRC6C, YBX1, ZCCHC11

Appendix Table A5. Altered biological processes and RNA-binding targets of TDP-43 in the differentially expressed TDP-43 Q331K cytoplasmic transcriptome

Major biological processes are ranked by ascending P-Values (p <0.005) for the differentially expressed gene list with log_2FC >1 and FDR <0.2. A total of 662 DEG transcripts (33 %) bound by TDP-43 were identified in the differentially expressed list (2001 gene transcripts are differentially expressed in the cytoplasmic transcriptome). Gene ontology-clustered RNA-binding targets of TDP-43 are highlighted in red.

Biological process	Count	%	P-Value	Official gene symbols
Cell cycle/ nuclear division related:				AATF, ABL1, AHCTF1, AK1, AKT1, ANAPC7, ANXA1, ANXA11, APBB1, APC, ARL2, AURKA, BBS4, BIRC5, BOP1, BUB1, CAMK2D, CCNB1, CCNE2, CCNG2, CDC20, CDC25B, CDC25C, CDC34, CDC45, CDC5L, CDC7, CDK20, CDK5RAP3, CDKN2C,
GO:0007049	160	8.7	1.0E-17	CENPF, CENPJ, CEP120, CEP164, CEP250, CEP63, CEP76, CETN3, CHAF1A,
GO:0022402	124	6.8		CHMP1A, CKAP2, CKS1B, CNTROB, CSRP2BP, DCLRE1A, DDX11, DLGAP5,
GO:0000278	87	4.8	1.0E-15	DMTF1, DNM2, DST, DTYMK, DYNC1LI1, E2F3, E2F6, E2F8, ERBB2IP, ERCC2,
GO:0022403	91	5.0		ESCO1, EXO1, FAM83D, FANCD2, FBXO5, FMN2, FNTA, FOXN3, GAK, GAS1,
GO:0051301	65	3.6	5.8E-13	GPS1, GSK3B, H2AFX, HAUS2, HAUS5, HAUS6, HELLS, HSPA8, ILF3, ITGB1, JMY,
GO:0000279	69	3.8		KATNA1, KATNB1, KCTD11, KIF20B, KNTC1, LIG1, LIG3, LLGL2, MAP3K11,
GO:0000087	52	2.8	9.0E-12	MDM2, MDM4, MEN1, MLF1, MLH1, MLH3, MPHOSPH9, MTBP, NAE1, NASP,
GO:0007067	50	2.7		NCOR1, NDC80, NEDD1, NEK3, NEK9, NFKBIL1, NIPBL, NOTCH2, NPM1,
GO:0000280	50	2.7	1.0E-08	NUMA1, NUSAP1, OIP5, PAPD7, PARD3, PBK, PEBP1, PIM1, PIN1, PKD2, PKIA,
GO:0048285	50	2.7		POLD1, PPP1CB, PPP2R3B, PPP3CA, PPP6C, PRC1, PRMT5, PSMA4, PSMA6,
GO:0051329	29	1.6	2.8E-08	PSMC6, PSMD1, PSMD13, PSMD2, PSMD9, PSME3, RACGAP1, RAN, RASSF2,
GO:0051325	29	1.6		RBM7, RHOU, RPL24, RPRM, RPS15A, SEH1L, SEPT2, SF1, SH3BP4, SIAH2,
GO:0007051	16	0.9	6.5E-08	SIRT2, SKA2, SKA3, SMC2, SPAST, STMN1, STRADB, TACC1, TAF1, TCF3,
GO:0051726	56	3.1		TCF7L2, TERF1, TET2, TEX15, TFDP2, TLK2, TOP3A, TP53BP2, TP53INP1, TUBB,
GO:0010564	25	1.4	2.4E-07	TUBE1, TUBG1, UBA52, UBB, UIMC1, USP22, VPS4A, WDR6, WEE1, ZC3HC1,
GO:0051225	8	0.4		ZWINT
GO:0007346	29	1.6	2.4E-07	
GO:0000910	12	0.7		
GO:0051297	10	0.6	8.5E-07	
			1.6E-06	
			2.9E-06	
			3.1E-05	
			3.8E-04	
			6.5E-04	
			1.1E-04	
			2.2E-03	
			2.6E-03	
			4.5E-03	
Protein transport/ localization:				AGAP1, AKAP10, AKT1, ALS2, ANK3, AP1M1, AP1S2, AP2B1, AP2S1, AP3B1, ARF4, ARF5, ARFGAP2, ARFIP1, ARL17B, ARRB1, ATG16L1, ATG3, ATG4B, ATG7, ATL2, ATL3, ATP2C1, ATP7B, BBS4, BID, BLZF1, BNIP3, CCHCR1, CCS,
GO:0046907	116	6.3	1.6E-08	CENPF, CEP250, CHMP1A, CHMP2A, CHMP6, CLTA, COG1, COG6, COX18,
GO:0008104	141	7.7		CPT1B, CTSA, CUX1, DDX19A, DDX19B, DNAJC19, DNM2, DST, EPS15, ERBB2IP,
GO:0045184	123	6.7	2.4E-07	ERC1, ERGIC2, EXOC7, EXPH5, FGF9, FLNA, FRAS1, GGA1, GIPC1, GOLGA5,
GO:0015031	121	6.6		GOSR1, GSK3B, HDAC6, HNRNPA1, HPS4, HSPA8, IPO4, KATNA1, KATNB1,
GO:0034613	64	3.5	1.5E-06	KDELR2, KIF1A, KIF20A, KIF5B, KLC1, LRPPRC, MGEA5, MLH3, MON2, MYO5A,
GO:0032507	12	0.7		MYO9B, NASP, NDC80, NDUFA13, NECAP1, NFKBIA, NFKBIE, NFKBIL1, NPM1,
GO:0045185	13	0.7	2.7E-06	NUP50, NUP98, NXT2, OS9, OSBPL5, PABPN1, PAX6, PCSK5, PEX13, PPARG,
GO:0006886	57	3.1		PPP3CA, PUM1, RAB10, RAB11A, RAB11B, RAB11FIP4, RAB14, RAB15, RAB18,
			1.3E-	RAB31, RAB40B, RAB6A, RAB7A, RABEP1, RAN, RANBP3, RANGRF, RBM22,

			03 2.6E-03 3.3E-03 4.0E-03	REEP1, RFFL, RGPD8, RHOBTB3, RHOT1, RPS15, RSRC1, SAR1B, SCAMP3, SCFD1, SEC23B, SEC31A, SEC61G, SEH1L, SH3D19, SLC15A4, SLC25A1, SLC25A12, SMG1, SNAP23, SNAP29, SNX13, SNX14, SNX5, SNX8, SORT1, SPAST, SPG7, SPTBN4, SQSTM1, SRP54, SRP9, STARD3, STAU1, STON1, STRADB, STX10, STX3, STX7, STXBP1, SYNRG, TEX15, THOC5, TIMM10, TIMM23B, TMSB10, TOM1L1, TOMM40, TOMM5, TOMM7, TOPORS, TRAM1, TRAPPC2, TRAPPC4, UPF2, USE1, VAMP7, VCP, VPS13B, VPS26A, VPS37C, VPS41, VPS4A, VTI1B, YIF1A, YIPF5
Cellular response to stress: GO:0033554	103	5.6	2.4E-08	AATF, ABL1, ALKBH2, APBB1, APC, ATAD5, ATF4, ATG16L1, ATMIN, ATRX, ATXN3, BAZ1B, BECN1, CAT, CCDC47, CCM2, CEP164, CEP63, CHAF1A, CHD1L, CHST3, CSNK1D, DCLRE1A, DUSP10, DUSP9, EEPD1, EIF2B4, ERCC2, ERCC5, ESCO1, EXO1, FADS1, FANCD2, FANCF, FANCL, FOXN3, GPS1, GSK3B, GTF2H2C, GTF2H4, H2AFX, HDAC6, INSIG1, JMY, LIG1, LIG3, LONP1, MAP1B, MAP2K4, MAP2K7, MAP3K11, MAP3K9, MAPK8, MBIP, MDFIC, MEN1, MLH1, MLH3, MMS19, MRPS35, MUS81, MYOF, NAE1, NSMCE2, OS9, PAPD7, PARP2, PARP3, PKN1, PMS1, POLD1, POLD3, POLI, POLK, PPP1CB, PPP1R15B, PRDX3, PXDN, REV3L, RFC3, RFC4, RPA2, RPS3, RTEL1, RUVBL2, SLC11A2, SMC6, SMG1, STAC, STRADB, SUMO1, TLK2, TNRC6A, TOPORS, TYMS, UBE2A, UBE2V2, UBR5, UIMC1, VCP, XRCC3, XRCC6BP1, ZSWIM7
Transcription, transcriptional regulation: GO:0006350 GO:0045449 GO:0051252 GO:0006355	293.0 343 238 231	16.0 19.0 13.0 13.0	5.2E-08 9.8E-07 1.1E-04 2.2E-04	AATF, ABL1, ADNP2, AEBP2, AHCTF1, APBB1, ARNT2, ASCC2, ASXL1, ATF2, ATF4, ATF5, ATF7, ATF7IP, ATRX, ATXN3, ATXN7L3, BAZ1A, BAZ1B, BAZ2B, BBX, BCKDHA, BCLAF1, BHLHE40, BLZF1, BPTF, BRF1, BRWD1, CAMTA1, CAT, CBFA2T2, CBX3, CCNC, CD3EAP, CDC5L, CDCA7, CDX2, CELSR2, CENPF, CHAF1A, CHD3, CHD9, CHMP1A, CLOCK, CNBP, CNOT4, CNOT6, CNOT8, COPS2, CREB3L2, CSDE1, CTNNBIP1, CUX1, DACH1, DDX5, DEDD, DEK, DIDO1, DMTF1, DNAJB6, DNM2, DNMT1, DNMT3A, DNMT3B, E2F3, E2F6, E2F8, EDA2R, EED, ELF2, ELP3, ELP4, ENO1, ERC1, ERCC2, ERF, ESRRA, ETV4, EWSR1, FADS1, FLNA, FOXN3, GABPB1, GATAD1, GON4L, GTF2A2, GTF2F1, GTF2H2C, GTF2H4, GTF2IRD2B, GTF3C5, HAT1, HDAC4, HDAC6, HDAC9, HELLS, HES1, HIVEP1, HIVEP2, HKR1, HMBOX1, HMGA1, HNF4G, HNRNPD, HOMEZ, HOXA9, HOXC6, HOXC9, HOXD13, IGF2BP1, IKBKG, IKZF5, ILF3, IRAK1, JMJD6, JMY, KAT2A, KDM1A, KDM2A, KDM2B, KHSRP, KLF3, KRBA2, L3MBTL3, LBH, LEF1, LIMD1, LIN54, LRPPRC, LRRFIP1, MAML2, MBD1, MDFIC, MDM2, MDM4, MED17, MED24, MED25, MED29, MED31, MED6, MEF2A, MEF2C, MEIS1, MEIS2, MEN1, MIER3, MKL1, MKL2, MLF1, MMS19, MURC, MYBL1, MYBL2, MYNN, MYPOP, NAA15, NCOA1, NCOA7, NCOR1, NCOR2, NDUFA13, NFAT5, NFATC3, NFE2L1, NFIA, NFIB, NFKBIA, NFKBIZ, NKX3-1, NOTCH2, NPM1, PAIP1, PAPOLA, PAX3, PAX6, PBX1, PBXIP1, PCGF3, PDCD6, PFN1, PHF1, PHF12, PHF3, PHF6, PIAS3, PIM1, PKIA, POLR1A, POLR1D, POLR1E, POLR2H, POLR2I, POLR2J2, POLR2J3, POLR2L, POLR3D, PPARG, PPM1A, PQBP1, PRDM16, PRDX3, PRMT5, PSMD9, PTOV1, PUF60, RAB18, RAN, RBM5, RCOR2, RIPK2, RNASEK, RNF4, RNPS1, RPS13, RPS3, RREB1, RRN3, RUVBL2, S1PR1, SARNP, SERBP1, SETD2, SF1, SFMBT1, SIM2, SIN3A, SIN3B, SIRT2, SIRT5, SLC30A9, SMAD7, SMAD9, SMARCD1, SNAPC3, SNAPC5, SOX5, SP4, SQSTM1, SRFBP1, STAT3, STON1, SUMO1, SUPT3H, TAF1, TAF1A, TAF4B, TAF6L, TAF8, TARBP2, TBX1, TCEAL8, TCEB2, TCF3, TCF7L2, TCFL5, TEAD2, TERF1, TFDP2, TGFBR1, TGIF1, TIGD2, TIGD7, TMPO, TOPORS, TRAPPC2, TRIB3, TROVE2, TSC22D3, TTF1, TTLL5, UBA52, UBB, UIMC1, USF1, USF2, USP22, VDR, VGLL4, WHSC1, YY1, ZBTB44, ZBTB6, ZFAT, ZFP64, ZFP91, ZHX1, ZKSCAN2, ZNF107, ZNF138, ZNF148, ZNF160, ZNF189, ZNF195, ZNF202, ZNF207, ZNF211, ZNF221, ZNF224, ZNF226, ZNF227, ZNF229, ZNF233, ZNF248, ZNF263, ZNF267, ZNF268, ZNF280D, ZNF329, ZNF331, ZNF337, ZNF35, ZNF382, ZNF404, ZNF407, ZNF419, ZNF426, ZNF431, ZNF433, ZNF436, ZNF44, ZNF441, ZNF470, ZNF48, ZNF493, ZNF518A, ZNF519, ZNF529, ZNF548, ZNF549, ZNF555, ZNF562, ZNF57, ZNF571, ZNF573, ZNF574, ZNF582, ZNF586, ZNF595, ZNF605, ZNF615, ZNF621, ZNF625, ZNF641, ZNF644, ZNF649, ZNF670, ZNF700, ZNF701, ZNF709, ZNF71, ZNF714, ZNF720, ZNF721, ZNF747, ZNF76, ZNF766, ZNF780B, ZNF800, ZNF81, ZNF813, ZNF821, ZNHIT3, ZSCAN18, ZSCAN23, ZXDA

Microtubule/ cytoskeleton based process: GO:0007017 GO:0007010 GO:0000226 GO:0032886 GO:0070507 GO:0031110	 55 77 34 16 14 10	 3.0 4.2 1.9 0.9 0.8 0.6	 2.6E-07 5.5E-06 1.8E-05 9.6E-05 2.4E-04 3.6E-03	ABI2, ABL1, ACTR10, ADD1, ALDOA, AMOT, ANK3, APC, ARHGAP10, ATP2C1, AURKA, BBS4, CALD1, CAP2, CDC42BPB, CENPJ, CEP120, CEP250, CEP63, CEP76, CETN3, CNN3, CNTROB, CYTH2, DAAM2, DLC1, DNAJB6, DST, DYNC1I2, DYNC1LI1, DYNLL1, EPB41L1, EPB41L2, EPB41L3, ERBB2IP, FBXO5, FERMT2, FLNA, FMN2, HAUS2, HAUS5, HAUS6, HDAC6, INPP5K, ITGB1, KATNA1, KATNB1, KIF16B, KIF1A, KIF20A, KIF20B, KIF21A, KIF24, KIF5B, KIF7, KIFC3, KLC1, KTN1, LASP1, LRPPRC, MAP1B, MAP2, MAP3K11, MID1IP1, MLH1, MYO9B, NCK2, NCOR1, NDC80, NPHP4, NPM1, NUSAP1, OBSL1, OFD1, PACSIN2, PARVA, PEX13, PFN1, PLD2, PRC1, RACGAP1, RAN, RHOT1, RHOU, RICTOR, SEH1L, SEMA6A, SGCB, SKA2, SKA3, SPAST, SPG7, SPTBN4, SSH1, STMN1, TERF1, TMSB10, TRPM7, TUBA1B, TUBB, TUBD1, TUBE1, TUBG1, TUBG2, TUBGCP2, ZWINT
Golgi /vesicle transport: GO:0048193 GO:0016192 GO:0006892 GO:0006888	 35 96 18 13	 1.9 5.2 1.0 0.7	 4.1E-07 4.5E-06 6.2E-05 9.4E-04	ALS2, AP1M1, AP1S2, AP2B1, AP2S1, AP3B1, ARF4, ARF5, ARFGAP2, ARL17B, ARRB1, ATL2, ATL3, BLZF1, CHMP1A, CLTA, COG1, CPNE1, CUX1, CYTH1, CYTH2, DNM2, EHD4, ELMO2, EPN2, ERC1, ERGIC2, EXOC7, FLNA, GAPVD1, GGA1, GLRB, GOLGA5, GOSR1, GTF2H2C, GULP1, HSPA8, ITSN1, ITSN2, JMJD6, KALRN, KDELR2, KIF20A, LLGL2, LRP4, MON2, NECAP1, OSBPL1A, OSBPL5, PACSIN2, PLD2, PUM1, RAB14, RAB18, RAB6A, RAB7A, RABEP1, RABEPK, RBM12, RHOBTB3, SAR1B, SCAMP3, SCFD1, SDF4, SEC23B, SEC31A, SH3BP4, SH3D19, SNAP23, SNAP29, SNAP91, SORT1, SPAST, SPTBN4, SQSTM1, SRC, STON1, STX10, STX3, STX7, STXBP1, SYNRG, TRAPPC2, TRAPPC4, TRIM36, USE1, VAMP3, VAMP4, VAMP7, VCP, VLDLR, VPS26A, VPS41, VPS4A, VTI1B, YIF1A, YIPF5
DNA damage/ repair response: GO:0006974 GO:0006281	 69 54	 3.8 3.0	 3.6E-06 2.2E-05	AATF, ABL1, ALKBH2, APBB1, APC, ATAD5, ATMIN, ATRX, ATXN3, BAZ1B, CEP164, CEP63, CHAF1A, CHD1L, CSNK1D, DCLRE1A, EEPD1, ERCC2, ERCC5, ESCO1, EXO1, FANCD2, FANCF, FANCL, FOXN3, GTF2H2C, GTF2H4, H2AFX, JMY, LIG1, LIG3, MEN1, MLH1, MLH3, MMS19, MRPS35, MUS81, NAE1, NSMCE2, PAPD7, PARP2, PARP3, PMS1, POLD1, POLD3, POLI, POLK, PPP1CB, REV3L, RFC3, RFC4, RPA2, RPS3, RTEL1, RUVBL2, SMC6, SMG1, SUMO1, TLK2, TOPORS, TYMS, UBE2A, UBE2V2, UBR5, UIMC1, VCP, XRCC3, XRCC6BP1, ZSWIM7
DNA metabolic process: GO:0006259	 87	 4.8	 3.9E-06	ABL1, ALKBH2, ATF7IP, ATRX, ATXN3, BNIP3, CCNE2, CDC25C, CDC34, CDC45, CDC7, CENPF, CEP164, CHAF1A, CHD1L, CSNK1D, DCLRE1A, RIT1, DNMT1, DNMT3A, DNMT3B, EEPD1, ERCC2, ERCC5, ESCO1, EXO1, FANCD2, FANCF, FANCL, GIN1, GINS1, GTF2H2C, GTF2H4, H2AFX, HELLS, HMGA1, HSPD1, JMY, KCTD13, KRBA2, LIG1, LIG3, LONP1, MCM9, MEN1, MLH1, MLH3, MMS19, MUS81, MYO18A, NAE1, NASP, NFIA, NFIB, NOL8, NSMCE2, NUP98, PAPD7, PARP2, PARP3, PMS1, POLD1, POLD3, POLI, POLK, PPIA, RAN, REV3L, RFC3, RFC4, RPA2, RRM1, RTEL1, RUVBL2, SMC6, SMG1, SUMO1, TERF1, TOP3A, TYMS, UBE2A, UBE2V2, UIMC1, VCP, XRCC3, XRCC6BP1, ZSWIM7
Protein catabolic process and ubiquitin-dependent protein degradation: GO:0043632 GO:0019941 GO:0030163 GO:0009057 GO:0051603 GO:0044257 GO:0044265 GO:0016567 GO:0051248 GO:0032269	 93 93 99 118 94 94 109 26 35 33	 5.1 5.1 5.4 6.4 5.1 5.1 6.0 1.4 1.9 1.8	 2.0E-05 2.0E-05 2.2E-05 3.8E-05 6.8E-05	ADAMTS9, AKT1, ANAPC7, APC, ARIH2, ASB13, ATG3, ATG4B, ATG7, AXIN1, BNIP3, CACYBP, CCNB1, CCNB1IP1, CDC20, CDC34, CNOT4, CPS1, DCAF15, RIT1, DLC1, DNMT1, DNMT3B, EDEM1, EIF2A, EIF2B4, EIF3K, EIF4E3, EIF4G3, ERCC2, ERCC5, ETF1, FANCL, FBXO25, FBXO33, FBXO41, FBXO45, FBXO5, FBXO7, FBXO9, FBXW2, FEM1B, FLNA, FNTA, FURIN, GIPC1, GTF2H2C, GTF2H4, HDAC4, HDAC6, HNRNPD, IGF2BP1, IGF2BP2, IGF2BP3, KDM2A, KDM2B, KLHL12, LONP1, MAP3K11, MAP3K9, MARCH5, MARCH7, MARCH8, MDFIC, MDM2, MDM4, MEN1, MGEA5, MGRN1, MIB2, MKNK1, MLH1, MTBP, MTIF3, MTRF1, NAE1, NCK2, NDUFA13, NEDD4L, NFKBIA, NGDN, NGLY1, NOSIP, NSMCE2, OS9, PAIP1, PEBP1, PIAS3, PIN1, PKN1, POP1, PPIL2, PPP1R15B, PPP2R4, PRKCA, PSMA4, PSMA6, PSMC6, PSMD1, PSMD13, PSMD2, PSMD9, PSME3, PUM1, RAB40B, RFFL, RICTOR, RNF128, RNF34, RNF40, RNF7, RNH1, RNPS1, RPA2, RPS5, SAE1, SARNP, SH3D19, SIAH2, SMAD7, SMG1, SOCS2, SOCS4, SOCS7, SPG7, SPSB1, SQSTM1, SRP9, SUMO1, TAF1, TARBP2, TCEB2, TGFBR1, TIA1, TIMP1, TNRC6A, TOM1L1, TOPORS, TRIM23, UBA52, UBB,

GO:0042176 GO:0032268 GO:0006511	14 70 40	0.8 3.8 2.2	8.2E-05 9.5E-05 5.3E-04 9.9E-04 2.0E-03 2.3E-03 3.0E-03 4.3E-03	UBE2A, UBE2D2, UBE2D3, UBE2J2, UBE2K, UBE2O, UBE2V2, UBE3B, UBL5, UBL7, UBR3, UBR4, UBR5, UPF2, UPF3A, USE1, USP10, USP22, USP33, USP38, USP42, USP43, USP44, USP47, USP48, USP54, USP8, VCP, WDSUB1, YME1L1, ZC3HC1, ZFP91
Protein modification by small protein conjugation or removal GO:0070647 GO:0032446	35 30	1.9 1.6	4.5E-05 8.6E-05	AKT1, ATG3, ATXN7L3, CDC34, FBXO25, FBXO9, HDAC6, KAT2A, MDM2, MIB2, NAE1, NEDD4L, NOSIP, OS9, PIAS3, PPIL2, RNF40, RNF7, SAE1, SIAH2, SOCS7, SUMO1, SUPT3H, TRIM23, UBA52, UBB, UBE2A, UBE2D2, UBE2D3, UBE2V2, UIMC1, USP22, USP33, VCP, WDSUB1
Macromolecular / ribonucleoprotein complex assembly: GO:0043933 GO:0034621 GO:0070271 GO:0006461 GO:0065003 GO:0010605 GO:0034622	108 62 80 80 100 108 52	5.9 3.4 4.4 4.4 5.5 5.9 2.8	6.4E-05 8.3E-05 1.8E-04 1.8E-04 1.9E-04 2.3E-04 1.4E-03	AATF, ABI2, ADSL, AHCTF1, ANAPC7, AP2S1, APBB1, APC, ARL2, ATF7IP, ATG16L1, ATL2, ATL3, ATPIF1, BCLAF1, BHLHE40, BNIP3, BNIP3L, BPTF, BRF1, CAT, CBFA2T2, CBX3, CDC20, CDC45, CDX2, CELF1, CENPF, CENPJ, CHAF1A, CHMP1A, COPS2, COX11, COX18, CTNNBIP1, CUX1, CYBA, DEDD, DGKH, DNAJB6, DNMT1, DNMT3A, DNMT3B, E2F3, E2F6, EED, EIF2A, EIF2B4, ENO1, ERCC2, ETF1, FBXO5, FLNA, FNTA, FOXN3, FURIN, GIPC1, GLRB, GTF2A2, GTF2F1, GTF2H2C, GTF2H4, H1F0, H2AFV, H2AFX, H3F3B, HAT1, HDAC4, HDAC6, HDAC9, HELLS, HES1, HIST1H3B, HIST1H4H, HIVEP1, HMGA1, HP1BP3, HSD17B10, HSPD1, IGF2BP1, IGF2BP2, IGF2BP3, ILF3, IPO4, IRAK1, KDM1A, KNTC1, LEF1, LONP1, LRRFIP1, MAP3K11, MBD1, MDM2, MDM4, MED17, MED24, MEF2C, MEIS2, MEN1, MGEA5, MLH1, MTIF3, MTRF1, MTRF1L, MYPOP, NASP, NCK2, NCOR1, NCOR2, NDUFA13, NPM1, NUP98, PARD3, PARVA, PAX3, PBXIP1, PEBP1, PEX13, PFDN6, PFKL, PHF12, PKIA, POLR2H, POLR2I, POLR2L, PPARG, PPP2R4, PRDM16, PRKAG1, PRKCA, PRMT5, PRPF31, PSMA4, PSMA6, PSMC6, PSMD1, PSMD13, PSMD2, PSMD9, PSME3, RBM5, RCOR2, RNF128, RPL24, RPS13, RPS15, RPS3, RRM1, SBF2, SCO2, SEH1L, SF1, SHMT1, SIM2, SIN3A, SIN3B, SIRT2, SIRT5, SMAD7, SNAP91, SNAPC5, SNRPC, SNRPD2, SPAST, SRC, SRP54, SRP9, STAT3, STMN1, STON1, SUMO1, TAF1, TAF4B, TAF6L, TARBP2, TAZ, TCEB2, TERF1, TGIF1, TIA1, TIMP1, TNRC6A, TSR1, TTF1, TUBA1B, TUBB, TUBD1, TUBE1, TUBG1, TUBG2, TUBGCP2, UBA52, UBB, UIMC1, VAMP3, VCP, VDR, ZFP91, ZHX1, ZNF148, ZNF189, ZNF202, ZNF382
Chromatin organization/ modification: GO:0051276 GO:0016568 GO:0006325 GO:0016569 GO:0006338	79 50 60 25 14	4.3 2.7 3.3 1.4 0.8	7.7E-05 1.3E-04 1.2E-03 2.8E-03 4.5E-03	ACIN1, AEBP2, APBB1, APC, ATXN7L3, BAZ1B, BNIP3, BPTF, CABIN1, CBX3, CENPF, CHAF1A, CHD1L, CHD3, CHD9, CHMP1A, CHRAC1, CSRP2BP, DAPK3, DDX11, RIT1, DLGAP5, DNMT1, DNMT3A, DNMT3B, EED, EHMT2, FANCD2, H1F0, H2AFV, H2AFX, H3F3B, HAT1, HDAC4, HDAC6, HDAC9, HELLS, HIST1H3B, HIST1H4H, HMGA1, HP1BP3, JMJD6, KAT2A, KDM1A, KDM2A, KDM2B, KDM6A, MEN1, MLH3, MSL1, NASP, NCAPH2, NCOR1, NDC80, NIPBL, NPM1, NUSAP1, PAPD7, PHF1, PRMT5, RNF40, RTEL1, RUVBL2, SEH1L, SETD2, SIRT2, SMARCD1, SMC2, SUPT3H, TAF6L, TERF1, TEX15, TLK2, TTF1, UBE2A, UIMC1, USP22, WHSC1, ZWINT

Translation/ translational regulation: GO:0006414 GO:0006412 GO:0006413 GO:0006417	 25 58 15 26	 1.4 3.2 0.8 1.4	 9.4E-05 1.1E-04 1.3E-04 4.1E-03	AKT1, BRF1, DHPS, DMXL2, EEF1D, EFTUD1, EIF1AD, EIF2A, EIF2B4, EIF3C, EIF3F, EIF3G, EIF3J, EIF3K, EIF4E3, EIF4G3, ETF1, HBS1L, IGF2BP1, IGF2BP2, IGF2BP3, LARS, MKNK1, MRPL11, MRPL2, MRPL22, MRPL55, MRPS16, MRPS18B, MRPS36, MTIF3, MTRF1, MTRF1L, NCK2, NDUFA13, NGDN, PAIP1, PPP1R15B, PRKCA, PUM1, RPL10, RPL14, RPL17, RPL18, RPL19, RPL21, RPL24, RPL26, RPL3, RPL31, RPL38, RPLP0, RPS13, RPS15, RPS15A, RPS18, RPS19, RPS24, RPS3, RPS3A, RPS5, SARNP, SARS2, SRP9, TARBP2, TARS, TARS2, TIA1, TNRC6A, TSFM, UBA52, UBB, VARS2
Cell death/ apoptosis: GO:0006915 GO:0012501 GO:0043065 GO:0043068 GO:0008219 GO:0010942 GO:0016265 GO:0006917 GO:0012502 GO:0042981	 90 91 67 67 102 67 102 51 51 109	 4.9 5.0 3.7 3.7 5.6 3.7 5.6 2.8 2.8 6.0	 5.0E-04 5.2E-04 9.9E-04 1.2E-03 1.3E-03 1.3E-03 1.6E-03 2.6E-03 2.8E-03 4.0E-03	AATF, ABL1, ACIN1, ACVR1C, AKT1, ALS2, ANXA1, APBB1, APC, ARNT2, ATF5, ATXN3, AXIN1, AXIN2, BAG4, BAG5, BCL2L11, BCL2L13, BCLAF1, BECN1, BID, BIRC5, BNIP2, BNIP3, BNIP3L, CADM1, CAMK1D, CASP7, CAT, CDKN2C, CFLAR, CKAP2, COL4A3, CTSB, DAP3, DAPK1, DAPK3, DDX19B, DEDD, RIT1, DIDO1, DLC1, DNAJB6, DNM2, DYNLL1, EI24, ELMO2, EPHA7, ERCC2, ERCC5, FASTK, FASTKD1, FASTKD2, FEM1B, FNTA, FURIN, FXR1, GAS1, GSK3B, GULP1, HDAC6, HELLS, HSPD1, IKBKG, IP6K2, IRAK1, ITSN1, JMJD6, JMY, KALRN, MAP3K11, MAP3K7, MAPK8, MCF2L, MCL1, MDM4, MEF2A, MEF2C, MEN1, MGEA5, MKL1, MLH1, MTCH1, MYO18A, NAE1, NDUFA13, NDUFS1, NFKBIA, NFKBIL1, NOTCH2, NPM1, NQO1, PAX3, PDCD5, PDCD6, PEA15, PEG10, PIM1, PRDX3, PRKCA, PSME3, PUF60, RABEP1, RASGRF2, RASSF5, RBM5, REEP1, RFFL, RHOT1, RIPK2, RNF34, RNF7, RPS3, RPS3A, RTEL1, RTKN, SEMA6A, SGK1, SIAH2, SLC11A2, SOCS2, SON, SORT1, SPAST, SPG20, SPG7, SQSTM1, SRC, STRADB, TCF7L2, TERF1, TGFBR1, TIA1, TNFRSF10B, TOPORS, TP53BP2, TP53INP1, TPT1, TRAF5, TRIB3, TSC22D3, TTBK2, TUBB, UBA52, UBB, UNC5C, VCP, VDAC1, VDR, ZC3HC1, ZDHHC16, ZFP91, ZFYVE27
Phosphorylation: GO:0006468 GO:0016310	 96 109	 5.2 6.0	 1.1E-03 3.4E-03	ABI2, ABL1, ACVR1C, ADRBK1, AGK, AKT1, AKT3, ARAF, ATP5E, ATP6V1D, AURKA, BAZ1B, BCKDK, BMP2K, BUB1, CAMK1D, CAMK2D, CDC42BPB, CDC7, CDK16, CDK20, CHRFAM7A, CHUK, CSNK1A1, CSNK1D, CSNK1G1, CSNK1G2, CSNK2A1, DAPK1, DAPK3, DDR1, DGUOK, DSTYK, EIF2A, EPHA7, ERC1, FASTK, FRS2, GAK, GOLGA5, GSK3B, IP6K2, IRAK1, KALRN, KSR1, LIMK2, MAK, MAP2, MAP2K4, MAP2K5, MAP2K7, MAP3K11, MAP3K7, MAP3K9, MAP4K4, MAPK8, MATK, MDFIC, MEX3B, MKNK1, MON2, MVD, NDUFA10, NDUFA5, NDUFB10, NDUFS1, NEK3, NEK9, NIN, NRBP2, PBK, PIM1, PINK1, PKN1, PRKACB, PRKAG1, PRKCA, PTPRA, PXK, RIOK3, RIPK2, RPS6KB1, RSRC1, SCYL2, SCYL3, SGK1, SMAD7, SMG1, SRC, STK35, STRADB, TAF1, TAZ, TBCK, TBK1, TESK1, TFG, TGFBR1, TLK2, TRIB3, TRPM7, TTBK2, TYRO3, UQCR10, UQCRC1, VRK1, VRK3, WEE1, YES1
Posttranscriptional regulation of gene expression: GO:0010608	 38.0	 2.1	 1.2E-03	AKT1, AURKA, BRF1, CELF1, EIF2A, EIF2B4, EIF3K, EIF4E3, EIF4G3, ETF1, FBXO7, FLNA, GIPC1, HNRNPD, HPS4, HSPD1, IGF2BP1, IGF2BP2, IGF2BP3, MDM4, MKNK1, MTIF3, MTRF1, NCK2, NDUFA13, NGDN, PAIP1, PPP1R15B, PRKCA, PUM1, RPS5, SARNP, SERBP1, SMAD7, SRP9 TARBP2, TIA1, TNRC6A
Organelle organization/ localization: GO:0033043 GO:0051640	 38 20	 2.1 1.1	 2.0E-03 2.9E-	ADD1, AMOT, APC, ARPC3, BBS4, BIRC5, BUB1, CDC25C, CENPF, CEP120, CEP250, CEP76, DLC1, DLGAP5, DNMT1, DNMT3B, DST, DYNC1LI1, FMN2, HDAC6, KATNB1, KIF20B, LRPPRC, MAP1B, MAP2, MEN1, MID1IP1, MLH1, NCK2, NDC80, NEXN, NFKBIL1, NIN, NPM1, NUSAP1, PEBP1, PEX13, PIN1, PRMT5, RHOT1, RICTOR, RPS15, S1PR1, SEH1L, SEMA6A, SKA2, SKA3, SNAP23, SNAP29, SPTBN4, TERF1, TLK2, TMSB10

			03	
Anti-apoptosis: GO:0006916	36.	2.0	2.8E-03	AATF, AKT1, ANXA1, ATF5, BAG4, BECN1, BIRC5, BNIP2, BNIP3, BNIP3L, CFLAR, DAPK1, GSK3B, HELLS, IRAK1, MCL1, MEF2C, MKL1, MYO18A, NFKBIA, NOTCH2, NPM1, PEA15, RIPK2, RNF7, RTEL1, SOCS2, SON, SQSTM1, STRADB, TCF7L2, TGFBR1, TPT1, UBA52, UBB, ZC3HC1

Appendix Table A6. Altered biological processes and RNA-binding targets of TDP-43 in the differentially expressed TDP-43 Q331K GRASPS translatome

Major biological processes are ranked by ascending P-Values (p <0.005) for the differentially expressed gene list with log_2FC >1 and FDR <0.2. A total of 1377 DEG transcripts (32 %) bound by TDP-43 were identified in the differentially expressed list (4264 gene transcripts are differentially expressed in the GRASPS translatome). Gene ontology-clustered RNA-binding targets of TDP-43 are highlighted in red.

Biological process	Count	%	P-Value	Official gene symbols
Cell cycle/ nuclear division related:				AATF, ADCY3, AKT1, ANAPC1, ANAPC11, ANAPC5, ANXA11, APBB1, ARAP1, ARHGAP8, ARHGEF2, ASNS, ASPM, AURKB, BARD1, BBS4, BCAT1, BIN3, BIRC5, BMP2, BOD1, BRE, BRSK1, BUB1B, CABLES1, CALM3, CASP3, CCNA1, CCNB1,
GO:0007049	227	8.3	7.8E-23	CCNE1, CCNE2, CCNG1, CCNH, CDC123, CDC16, CDC20, CDC25A, CDC25B,
GO:0022402	176	6.4	6.9E-21	CDC27, CDC37, CDC42, CDC45, CDCA3, CDCA7L, CDCA8, CDK1, CDK10, CDK13,
GO:0000278	129	4.7	6.2E-20	CDK4, CDK5RAP1, CDK5RAP3, CDK7, CDKN1A, CDKN1C, CDKN3, CENPV,
GO:0022403	117	4.3	7.2E-11	CEP250, CETN3, CFL1, CGREF1, CHEK2, CHFR, CHTF18, CHTF8, CKAP5, CLIP1,
GO:0000087	74	2.7	2.0E-10	CSRP2BP, CUL1, CUL7, CUL9, CYLD, DAXX, DBF4B, DCTN1, DCTN2, DDB1,
GO:0007067	73	2.7	2.1E-10	DDX11, DGKZ, DLG1, DMTF1, DNM2, DSN1, E2F6, E4F1, EGFR, EIF4G2, ERCC2,
GO:0000280	73	2.7	2.1E-10	ERCC3, EXO1, FANCA, FANCI, FOXM1, FZR1, GADD45B, GAK, GMNN, GPS1,
GO:0051301	89	3.3	5.0E-10	GSPT1, HAUS4, HAUS5, HAUS6, HAUS8, HBP1, HECTD3, HERC5, HEXIM2, HHEX,
GO:0048285	74	2.7	6.0E-10	HJURP, HMG20B, HSPA8, ILF3, ILK, ILKAP, INCENP, ING4, INTS3, IRF6, ITGB1,
GO:0000279	93	3.4	7.6E-09	JAG2, KATNB1, KLHDC3, LIG1, LIG3, LZTS1, MACF1, MAD1L1, MAD2L2, MADD,
GO:0051726	88	3.2	4.0E-07	MAEA, MAP2K1, MAP3K11, MAP9, MAPK14, MAPK3, MAPK7, MAPRE2,
GO:0051329	31	1.1	4.0E-04	MARK4, MCM2, MCM3, MCM5, MCM7, MDM4, MEN1, MIS12, MKI67, MLF1,
GO:0007346	41	1.5	5.1E-04	MLH1, MPHOSPH6, MSH5, MSH6, MYC, MYCBP2, MYH10, NASP, NCAPD2,
GO:0051325	31	1.1	6.8E-04	NCAPD3, NCAPH, NCOR1, NDE1, NEK1, NEK4, NEK6, NOLC1, NPM1, NUDC,
GO:0007050	28	1.0	4.1E-03	NUP37, OIP5, PA2G4, PBK, PBRM1, PCNP, PCNT, PDPN, PEX11B, PIM3, PKIA,
GO:0000082	18	0.7	4.3E-03	PKMYT1, PLK1, PML, POLE, PPP1CA, PPP3CA, PPP5C, PRC1, PRMT5, PRR5, PSMA1, PSMA2, PSMA3, PSMA4, PSMA6, PSMA7, PSMB7, PSMC2, PSMC3, PSMC3IP, PSMC4, PSMC5, PSMC6, PSMD1, PSMD10, PSMD12, PSMD13, PSMD3, PSMD4, PSMD6, PSMD8, PSME1, PSME3, PSMF1, PSMG2, RAD21, RAN, RASSF1, RB1CC1, RBBP4, RBBP8, RBM7, RCC1, RIF1, RNF167, RNF2, ROCK2, RPS27A, RUVBL1, SBDS, SCRIB, SEPT1, SEPT2, SEPT5, SEPT9, SESN3, SF1, SGSM3, SKP1, SKP2, SMARCA4, SMARCB1, SPC24, SPIN2B, STEAP3, STK11, STRADB, SUPT5H, SUV39H2, SYCE1L, TACC3, TAF6, TARDBP, TBRG4, TCF3, TCF7L2, TERF1, TFDP2, TMEM8B, TP53, TP73, TPD52L1, TPX2, TSPYL2, TUBE1, UBA3, UBC, UBE2C, UBE2E1, UBE2I, UIMC1, WDR6, WEE1, WTAP, XPO1, ZBTB16, ZBTB17, ZC3HC1, ZWINT
Cellular/ protein catabolic process and ubiquitin-dependent protein degradation:				ACR, ADAM15, ADAM23, ADAMTS10, AIFM1, AKT1, AKT2, AMZ2, ANAPC1, ANAPC11, ANAPC5, APOE, ARNT, ARNTL, ASB6, ATG4A, ATG4B, AUH, BACE2, BAP1, BARD1, BECN1, BIRC6, BMP1, BNIP3, BRAP, BUB1B, C2, CAND1, CAND2, CAPN1, CAPN2, CAPNS1, CASP3, CASP6, CASP7, CBLL1, CCNB1, CCNB1IP1, CCNH, CD46, CD55, CDC16, CDC20, CDC27, CDCA3, CDK1, CDK7, CHFR, CLN3, CLN5, CLN6, CLU, CNDP2, CNOT4, CR2, CST3, CTSB, CTSD, CTSH, CUL1, CUL7,
GO:0044265	204	7.5	1.8E-18	CUL9, CYCS, CYLD, DCUN1D2, DDB1, DDB2, DERL1, DERL2, DLD, DNAJA3,
GO:0009057	214	7.8	6.4E-18	DNASE1L1, DNMT1, DNPEP, DPP9, E4F1, ECD, EIF2B1, EIF2B4, EIF3E, EML2,
GO:0043632	167	6.1	2.2E-16	ERCC1, ERCC2, ERCC3, ERCC4, ERLEC1, FAF1, FAM129A, FANCL, FBXL16,
GO:0019941	167	6.1	2.2E-16	FBXL20, FBXL4, FBXO22, FBXO36, FBXO8, FBXO9, FBXW2, FBXW7, FKBP1A,
GO:0030163	177	6.5	2.4E-16	FLNA, FZR1, G6PD, GAA, GAN, GCLC, GIPC1, GNS, GSPT1, GTF2H2, GTF2H2C,
GO:0006511	91	3.3	2.4E-16	GTF2H3, GUSB, HDAC6, HECTD2, HECTD3, HERC5, HERC6, HIF1A, HLTF,
GO:0051603	172	6.3	2.8E-16	HNRNPD, HSP90B1, HTRA1, ITCH, ITGAV, KCMF1, KDM2B, LGMN, LMO7,
GO:0044257	172	6.3	5.0E-16	LONP1, LONP2, LRRC29, LTA4H, MALT1, MARCH6, MARCH7, MDM4, MEN1,
GO:0051351	41	1.5	1.4E-14	METAP2, MIB2, MLH1, MMP2, MPG, MUTYH, MYC, MYCBP2, NEDD4, NEDD4L,
GO:0051443	40	1.5	1.5E-14	NF2, NGLY1, NHEJ1, NSMCE2, OS9, OSGEP, OTUB1, OTUD5, PABPC4, PACSIN3,
GO:0051439	40	1.5	2.8E-14	PAF1, PCNP, PCSK1N, PCSK6, PEG10, PIAS1, PIAS2, PIGK, PITRM1, PJA1, PLAT,
GO:0051437	39	1.4	2.8E-14	PLK1, PML, PMPCA, PMPCB, PNRC2, PPARA, PPP2R1A, PPP2R4, PRKAG1,
GO:0031398	43	1.6	2.1E-13	PRKCZ, PRPF19, PSEN2, PSMA1, PSMA2, PSMA3, PSMA4, PSMA6, PSMA7,
GO:0051340	42	1.5	2.4E-13	PSMB7, PSMC2, PSMC3, PSMC4, PSMC5, PSMC6, PSMD1, PSMD10, PSMD12,
GO:0051438	41	1.5	2.7E-13	PSMD13, PSMD3, PSMD4, PSMD6, PSMD8, PSME1, PSME3, PSMF1, PYGL,
GO:0010498	48	1.8	3.8E-13	RAD23A, RBCK1, RCE1, RELA, RNASEH1, RNF123, RNF14, RNF167, RNF2,
GO:0043161	48	1.8	3.8E-13	RNF216, RNF25, RNF34, RNF40, RNF41, RNH1, RNPEP, RNPS1, RPA2, RPS27A,
GO:0031396	47	1.7	7.3E-13	RPS6KA5, SAE1, SCPEP1, SEC11A, SENP3, SET, SGSM3, SHPRH, SKP1, SKP2,

GO:0031145 GO:0031397 GO:0051436 GO:0051352 GO:0051444 GO:0051248 GO:0032269 GO:0006508 GO:0016567 GO:0042176 GO:0009894	36 38 35 35 35 64 61 224 41 21 30	1.3 1.4 1.3 1.3 1.3 2.3 2.3 8.2 1.5 0.8 1.1	1.3E-12 5.7E-12 8.0E-12 2.4E-11 2.4E-11 7.6E-10 3.0E-09 6.3E-07 9.9E-07 5.9E-05 2.5E-04	SMG7, SOCS2, SOD1, SPCS1, SPCS2, SPG7, STAMBP, STAMBPL1, STUB1, SUMO1, SUMO2, SUMO3, SYVN1, TAF9, TBL1XR1, TCEB1, TGFB1I1, THOP1, TIA1, TIMP1, TMEM189, TOM1L1, TSC1, TSPAN17, UBA2, UBA3, UBC, UBE2A, UBE2C, UBE2D2, UBE2D3, UBE2E1, UBE2E2, UBE2G1, UBE2H, UBE2I, UBE2J2, UBE2N, UBE2V1, UBE2V2, UBE3B, UBL7, UBR1, UBR3, UBR4, UBR5, UCHL1, UQCRC1, UQCRC2, USP11, USP14, USP15, USP19, USP21, USP28, USP3, USP33, USP38, USP39, USP44, USP5, USP54, USP7, USP8, VPS28, WDR48, WFDC2, WWP2, XPNPEP1, XPO1, YWHAE, ZBTB16, ZC3HC1, ZFP36L1, ZNRF1
Cellular response to stress including oxidative stress: GO:0033554 GO:0080135 GO:0034599	 156 29 15	 5.7 11 0.5	 3.1E-13 2.0E-03 4.6E-03	AATF, AIFM1, AKT1, AKT2, ANKRD17, APBB1, APEX1, APTX, ARNT, ASNS, ATG16L1, ATG9A, ATMIN, ATRX, BARD1, BECN1, BRE, BRSK1, CASP3, CBS, CCNH, CDK1, CDK7, CDKN1A, CHD1L, CHEK2, CIDEB, CLN3, CLPB, CRKL, CTSD, DAXX, DBNL, DDB1, DDB2, DERL1, DERL2, DGKZ, DHRS2, DHX9, DYRK2, EIF2AK2, EIF2B1, EIF2B4, EIF2S1, ERCC1, ERCC2, ERCC3, ERCC4, ETV5, EXO1, EYA2, FAM129A, FANCA, FANCC, FANCI, FANCL, FGD4, FOS, FOXO3, G6PD, GNL1, GPS1, GPX3, GTF2H2, GTF2H2C, GTF2H3, HDAC6, HERPUD1, HIF1A, HIPK1, ING4, INTS3, KAT5, LIG1, LIG3, LONP1, MAP1B, MAP2K4, MAP2K7, MAP3K10, MAP3K11, MAP3K6, MAP3K7, MAPK14, MAPK9, MCM7, MEN1, MGMT, MLH1, MMS19, MORF4L1, MORF4L2, MPG, MSH5, MSH6, MUS81, MUTYH, NCOR1, NDEL1, NEDD4, NEFL, NEIL2, NEK1, NHEJ1, NONO, NPM1, NSMCE2, NUPR1, OS9, PCNA, PGAP2, PML, POLB, POLD2, POLE, POLE2, POLL, PRDX1, PRKDC, PRMT6, PRPF19, PXDN, PYCR1, RAD21, RAD23A, RB1CC1, RBBP8, RBM4, RECQL4, RECQL5, RFC2, RFC5, RIF1, RIPK2, RPA2, RPS3, RTEL1, RUVBL2, SCAMP5, SCAP, SHPRH, SIRT1, SOD1, SREBF1, STRADB, SUMO1, TAOK2, TAOK3, TERF2IP, TMEM189, TP53, TP53BP1, TP73, TYMS, UBE2A, UBE2N, UBE2V1, UBE2V2, UBR5, UIMC1, UVRAG, VAPB, WRNIP1, XRCC1, XRCC3, XRCC6, ZSWIM7
Intracellular/ protein transport: GO:0046907 GO:0015031 GO:0006886 GO:0000059	 174 163 88 10	 6.4 6.0 3.2 0.4	 5.2E-13 1.8E-05 6.9E-05 4.2E-04	AAAS, AAGAB, ACTN4, AFTPH, AGAP3, AKT1, ANP32A, AP1B1, AP1M1, AP1S1, AP2A2, AP2B1, AP2M1, AP3M2, AP4B1, APOE, ARF3, ARF5, ARFGAP1, ARFGAP2, ARHGEF2, ARL1, ARNTL, ATG16L1, ATG4A, ATG4B, ATG9A, ATL2, ATP2C1, ATP5O, ATP7B, BACE2, BBS4, BCAP31, BCL2L1, BID, BNIP3, CAMK1, CDC37, CEP290, CEP57, CHM, CHMP2A, CHMP4A, CHMP6, CLINT1, CLTA, CLTB, CLTCL1, COG2, COG4, COG5, COPA, COPZ1, CSE1L, CUX1, DCLK1, DDX19A, DDX19B, DERL1, DERL2, DNM2 DTNBP1, EIF5A, ENAH, ERC1, EXOC7, FLNA, FTH1, GABARAPL2, GDI2, GGA1, GGA3, GIPC1, GNAS, GORASP1, GOSR2, HDAC6, HHEX, HNRNPA1, HOOK3, HSP90B1, HSPA8, HTATIP2, ICMT, IPO11, IPO13, IPO4, IPO5, KATNB1, KIF17, KLC1, KLC2, KPNA4, KPNB1, LMAN2, LMF1, LONP2, M6PR, MACF1, MALT1, MAP1S, MAP2K1, MLX, MYH10, MYL6, MYO1C, MYO9B, NACA, NASP, NDE1, NDEL1, NECAP2, NEDD4, NEFL, NPC2, NPM1, NUP37, NUP54, NUP62CL, NUP85, NUP88, NUP93, NUP98, NUTF2, NXF1, OSBPL5, PCNA, PEX10, PEX19, PEX5, PITPNM1, PML, PPIF, PPP3CA, PREB, PRKCZ, PSEN2, PUM1, RAB11FIP3, RAB11FIP4, RAB18, RAB1B, RAB32, RAB38, RAB3C, RAB3IP, RAB5B, RAB7A, RAE1, RAMP2, RAN, RANBP3, RHOT2, RPH3AL, RRBP1, RUFY1, RYR2, SAR1A, SAR1B, SCAMP5, SCFD1, SCYL1, SEC16A, SEC23B, SEC24D, SEC31A, SEC61A1, SET, SFT2D3, SLC25A1, SLC25A10, SLC25A12, SLC25A13, SLC25A20, SMG7, SNUPN, SNX17, SNX27, SNX3, SNX5, SNX6, SNX9, SPG7, SRP54, SSR2, STARD3, STRADB, STX5, STX8, STXBP2, SYNGR1, SYTL1, SYTL5, TAOK2, THOC3, THOC5, THOC6, TIMM50, TIMM9, TMED10, TMX1, TNPO2, TOM1, TOM1L1, TOM1L2, TOMM40, TP53, TRAPPC1, TRAPPC3, TRAPPC4, TSC1, UCHL1, VAMP7, VAMP8, VPS11, VPS13A, VPS28, VPS33A, VPS37D, VPS41, VPS45, VTA1, VTI1B, WASF2, XPO1, XPO6, YIPF5, YKT6, YWHAE, YWHAZ, ZNF384
Translation/ translational regulation:				AIMP1, AKT1, BRF1, CDK4, CYLD, DDX1, DHPS, DPH1, EEF1A1, EEF1B2, EEF1D, EEF2K, EGFR, EIF2A, EIF2AK2, EIF2B1, EIF2B4, EIF2S1, EIF3C, EIF3CL, EIF3D, EIF3E, EIF3H, EIF3J, EIF3K, EIF3L, EIF3M, EIF4A1, EIF4A2, EIF4B, EIF4E, EIF4E2, EIF4G1, EIF4G2, EIF4H, EIF5, EIF5A, EIF5B, EIF6, EPRS, FAM129A, GFM1, GFM2,

GO:0006412 GO:0006413 GO:0006446 GO:0006414 GO:0006417	101 18 15 28 35	3.7 0.7 0.5 1.0 1.3	1.4E-11 2.7E-04 2.8E-03 3.0E-03 3.6E-03	GSPT1, HARS, HBS1L, HSPB1, KARS, LARS, LRRC47, MARS, MIF4GD, MKNK2, MRPL10, MRPL27, MRPL3, MRPL30, MRPL32, MRPL33, MRPL37, MRPL42, MRPL48, MRPL49, MRPL55, MRPL9, MRPS7, MRRF, NACA, NARS2, PA2G4, PABPC4, PAIP1, PML, PTRH2, PUM1, PUM2, QARS, RBM3, RPL13, RPL15, RPL17, RPL18, RPL18A, RPL26L1, RPL3, RPL31, RPL36A, RPL4, RPL6, RPL7, RPL7L1, RPL9, RPLP0, RPS11, RPS2, RPS27A, RPS3, RPS3A, RPS6KB2, RPS7, RRBP1, RSL1D1, TARBP2, TARS, TARS2, TIA1, TNIP1, TNRC6B, TNRC6C, TSC1, TSFM, TUFM, UBC, VARS2, WARS, WARS2, YARS, ZFP36L1
DNA metabolic/ catabolic process: GO:0006259 GO:0006308 GO:0006260	 137 22 47	 5.0 0.8 1.7	 4.0E-11 1.9E-04 1.4E-03	AIFM1, ANKRD17, APEX1, APTX, ATRX, BARD1, BNIP3, BRE, CASP3, CCDC88A, CCNE2, CCNH, CDC25A, CDC45, CDK7, CHD1L, CHTF18, CHTF8, CYCS, DDB1, DDB2, DMAP1, DNAJA3, DNASE1L1, DNMT1, DNMT3A, E4F1, ERCC1, ERCC2, ERCC3, ERCC4, EXO1, EYA2, FANCA, FANCC, FANCI, FANCL, FOS, GTF2H2, GTF2H2C, GTF2H3, HEMK1, HMGA1, HSPD1, IGFBP4, ING4, INTS3, KAT5, KCTD13, KLHDC3, LIG1, LIG3, LONP1, MCM2, MCM3, MCM4, MCM5, MCM7, MEN1, MGMT, MLH1, MMS19, MORF4L1, MORF4L2, MPG, MSH5, MSH6, MUS81, MUTYH, MYC, NAP1L1, NASP, NEIL2, NFIA, NFIC, NHEJ1, NONO, NSMCE2, NUP98, PCNA, PHB, PML, POLB, POLD2, POLE, POLE2, POLL, POLM, PPFIBP1, PRKDC, PRMT6, PRMT7, PRPF19, PSMC3IP, RAD21, RAD23A, RAN, RBBP4, RBBP8, RBM4, RECQL4, RECQL5, REPIN1, RFC2, RFC5, RNASEH1, RPA2, RRM1, RTEL1, RUVBL1, RUVBL2, SET, SHPRH, SIRT1, SMARCAL1, SMARCB1, SOD1, SUMO1, TERF1, TERF2IP, TFAM, TMX1, TP53, TP53BP1, TP73, TSN, TYMS, UBE2A, UBE2N, UBE2V2, UIMC1, UVRAG, WRAP53, WRNIP1, XRCC1, XRCC3, XRCC6, ZSWIM7
Chromatin / histone modification and organization: GO:0016568 GO:0006325 GO:0051276 GO:0016569 GO:0016570 GO:0016573 GO:0006338 GO:0006473	 87 101 120 44 42 18 18 18	 3.2 3.7 4.4 1.6 1.5 0.7 0.7 0.7	 4.4E-11 3.8E-08 1.9E-07 2.4E-07 7.3E-07 6.5E-04 4.3E-03 1.8E-03	ACIN1, ACTL6A, AIFM1, APBB1, ASF1B, ATXN7L3, BCORL1, BNIP3, BRD1, BRD8, BRE, CABIN1, CARM1, CBX6, CDCA8, CENPV, CHD1L, CHD4, CHRAC1, CSRP2BP, DAPK3, DDX11, DMAP1, DNMT1, DNMT3A, EED, ERCC1, ERCC4, EYA2, EZH2, H2AFY, H3F3B, HDAC10, HDAC6, HDAC7, HDAC8, HIRA, HIST1H2AC, HIST2H4A, HJURP, HLTF, HMG20B, HMGA1, HMGN2, ING4, KAT2A, KAT5, KDM2B, KDM4B, KDM6A, MCM2, MEN1, MIS12, MORF4L1, MORF4L2, MSH5, MSH6, MSL3, MTA2, NAP1L1, NAP1L4, NASP, NCAPD2, NCAPD3, NCAPH, NCOR1, NPM1, PAF1, PBRM1, PCGF2, PHB, PHF1, PRKDC, PRMT5, PRMT6, PRMT7, RBBP4, RBM4, RNF2, RNF40, RPS6KA5, RTEL1, RUVBL1, RUVBL2, SAFB, SAP130, SET, SETDB1, SHPRH, SIRT1, SMARCA4, SMARCAL1, SMARCB1, SMARCD1, SMARCE1, SMYD3, SOX6, SRCAP, SUPT5H, SUPT6H, SUPT7L, SUV39H2, TAF5L, TAF9, TBL1XR1, TERF1, TERF2IP, TP53, TRIM16, TSPYL2, UBE2A, UBE2E1, UBE2N, UBN1, UIMC1, USP21, WHSC1, WRAP53, XRCC6, ZWINT
DNA damage/ repair response: GO:0006974 GO:0006281 GO:0006289 GO:0009411 GO:0000718	 106 81 24 22 11	 3.9 3.0 0.9 0.8 0.4	 4.3E-10 5.1E-08 3.2E-06 1.4E-04 9.3E-04	AATF, AIFM1, AKT1, ANKRD17, APBB1, APEX1, APTX, ATMIN, ATRX, BARD1, BRE, BRSK1, CASP3, CASP7, CCNH, CDK1, CDK7, CDKN1A, CHD1L, CHEK2, CIDEB, DDB1, DDB2, DGKZ, DYRK2, EGFR, ERCC1, ERCC2, ERCC3, ERCC4, EXO1, EYA2, FANCA, FANCC, FANCI, FANCL, FECH, FOXO3, GNL1, GTF2H2, GTF2H2C, GTF2H3, HIPK1, ING4, INTS3, KAT5, LIG1, LIG3, MAPK14, MCM7, MEN1, MGMT, MLH1, MMS19, MORF4L1, MORF4L2, MPG, MSH5, MSH6, MUS81, MUTYH, NEDD4, NEIL2, NEK1, NHEJ1, NONO, NSMCE2, NUPR1, PCNA, PGAP2, PML, POLB, POLD2, POLE, POLE2, POLL, PRKDC, PRMT6, PRPF19, RAD21, RAD23A, RBBP8, RBM4, RECQL4, RECQL5, RELA, RFC2, RFC5, RIF1, RPA2, RPS3, RTEL1, RUVBL2, SDF4, SHPRH, SIRT1, SOD1, SUMO1, TDP1, TERF2IP, TP53, TP53BP1, TP73, TYMS, UBE2A, UBE2N, UBE2V2, UBR5, UIMC1, USF1, UVRAG, WRNIP1, XRCC1, XRCC3, XRCC6, ZSWIM7
Macromolecule/ protein metabolic process regulation; protein modification: GO:0032268 GO:0010605 GO:0043086	 123 171 80 82	 4.5 6.3 2.9 3.0	 7.1E-09 3.6E-08 3.4E-08 1.4E-07	AATF, ADCY3, AES, AKT1, AKT1S1, AKT2, ANAPC1, ANAPC11, ANAPC5, APBB1, APOE, ARNT, ARNTL, ATXN7L3, BACE2, BARD1, BCL2L1, BCL7A, BIRC5, BMP2, BNIP3, BRE, BRF1, BUB1B, CALCOCO1, CALM3, CAMKK2, CAND1, CASK, CASP3, CAST, CBLL1, CBS, CCDC88A, CCNB1, CCNE1, CCNH, CD276, CD81, CDC16, CDC20, CDC27, CDC45, CDK1, CDK4, CDK5RAP1, CDK7, CDKN1A, CDKN1C, CENPK, CHFR, CHURC1, CLN3, CLN6, CNBP, CNOT7, CREB1, CRTC1, CST3, CTBP1, CTBP2, CUL1, CUX1, DAXX, DBNL, DDB1, DDB2, DDX1, DDX5, DEDD, DGCR8, DICER1, DLC1, DMAP1, DNAJA3, DNAJB6, DNMT1, DNMT3A, DOCK7, DR1, DYRK2, E2F6, EED, EGFR, EIF2A, EIF2B1, EIF2B4, EIF2S1, EIF3E, EIF3H,

GO:0031399 GO:0070647 GO:0031401 GO:0044092 GO:0010604 GO:0051247 GO:0032446 GO:0032270	52 58 88 188 69 44 65	1.9 2.1 3.2 6.9 2.5 1.6 2.4	2.3E-07 2.5E-07 6.0E-07 6.5E-07 6.6E-07 1.0E-06 2.8E-06	EIF3K, EIF4A2, EIF4B, EIF4E, EIF4E2, EIF4G1, EIF4G2, EIF4H, EIF5, EIF5A, EIF5B, ELF1, ELK1, ENO1, ERCC1, ERCC2, ERCC3, ERCC4, ETV4, FAM129A, FBXO9, FBXW7, FGD4, FKBP1A, FLNA, FOS, FOXA2, FOXH1, FOXO3, FOXP4, FZR1, G6PD, GADD45B, GATA4, GCLC, GDF7, GDNF, GIPC1, GMNN, GNAI2, GNAZ, GPS1, GSPT1, GTF2F1, HAND2, HDAC10, HDAC6, HDAC8, HERPUD1, HES1, HEXIM2, HHEX, HIF1A, HMGA1, HNRNPAB, HOPX, HOXA1, HSPB1, IFI6, ILF3, ILK, ING4, IRAK1, IRF6, ITCH, ITGAV, JAZF1, KAT2A, KAT5, LDB1, LMCD1, LMO7, MAF1, MAML2, MAP2K1, MAP3K10, MAP3K11, MAP3K6, MAPK14, MAPK9, MBD1, MDM4, MED6, MEF2C, MEN1, MIB2, MIF4GD, MKL1, MKNK2, MLH1, MLST8, MLX, MMS19, MORF4L2, MOV10, MSH6, MTA2, MYC, NCOA1, NCOR1, NEDD4, NEDD4L, NF2, NFAT5, NFIC, NFYC, NIF3L1, NPAS2, NPM1, NPRL2, NQO1, NR2F2, OS9, PA2G4, PACSIN3, PAF1, PAIP1, PALM, PAWR, PBX1, PCGF2, PCNP, PDGFA, PHB, PIAS1, PIAS2, PKIA, PKNOX1, PLK1, PML, POU2F1, PPARA, PPARD, PPM1A, PPP2R1A, PPP2R4, PRKAR1B, PRKCZ, PRKDC, PRKRIP1, PRMT6, PRPF19, PSEN2, PSMA1, PSMA2, PSMA3, PSMA4, PSMA6, PSMA7, PSMB7, PSMC2, PSMC3, PSMC3IP, PSMC4, PSMC5, PSMC6, PSMD1, PSMD10, PSMD12, PSMD13, PSMD3, PSMD4, PSMD6, PSMD8, PSME1, PSME3, PSMF1, PTPRU, PUM1, PUM2, RAN, RAPGEF3, RB1CC1, RBM3, RBM4, RELA, REST, RNF10, RNF14, RNF167, RNF2, RNF4, RNF40, RNF41, RPS27A, RPS3, SAE1, SCAP, SCMH1, SCRIB, SET, SGSM3, SH3BP5, SIRT1, SKP1, SKP2, SMAD4, SMAD5, SMARCA4, SMARCB1, SMARCD1, SMARCE1, SND1, SNX6, SOX4, SOX6, SPRY1, SREBF1, SREBF2, SRRT, STUB1, SUMO1, SUMO2, SUMO3, SUPT5H, TAF9, TARBP2, TBL1XR1, TBP, TBX1, TCF25, TCF3, TCF7L2, TERF1, TERF2IP, TFE3, TGFB1I1, TIA1, TIMP1, TMEM189, TMX1, TNFRSF1A, TNRC6B, TNRC6C, TP53, TP53BP1, TP73, TRIM16, TRIM28, TSC1, TSPAN17, TSPYL2, UBA3, UBC, UBE2A, UBE2C, UBE2D2, UBE2D3, UBE2E1, UBE2H, UBE2I, UBE2N, UBE2V1, UBE2V2, UCHL1, UIMC1, USF1, USF2, USP21, USP33, USP7, UTF1, VEGFA, VEGFC, VPS28, WDR48, WDTC1, WWP2, XRCC6, YWHAE, ZBTB16, ZEB1, ZFHX3, ZFP36L1, ZFYVE28, ZNF423, ZNF462, ZNF496, ZNF639, ZSCAN21
Cell death/ apoptosis: GO:0043067 GO:0042981 GO:0010941 GO:0008219 GO:0016265 GO:0006915 GO:0043069 GO:0043066 GO:0012501 GO:0060548 GO:0043068 GO:0010942 GO:0043065	 179 177 179 159 159 136 89 88 137 89 91 91 90	 6.5 6.5 6.5 5.8 5.8 5.0 3.3 3.2 5.0 3.3 3.3 3.3 3.3	 9.1E-07 1.1E-06 1.2E-06 3.3E-06 5.0E-06 5.9E-06 7.7E-06 7.8E-06 8.1E-06 8.6E-06 2.6E-03 3.0E-03 3.1E-03	AATF, ACIN1, ACTN4, ADNP, AIFM1, AIMP1, AKT1, AKT1S1, ANXA4, ANXA5, APBB1, APOE, APTX, ARHGDIA, ARHGEF16, ARHGEF2, ARHGEF4, ASNS, ATXN10, BAG1, BARD1, BCAP31, BCL2L1, BCL2L13, BECN1, BID, BIRC5, BIRC6, BLCAP, BMF, BNIP3, BRE, BUB1B, CADM1, CARD6, CASP3, CASP6, CASP7, CD44, CDK1, CDKN1A, CFL1, CHEK2, CIAPIN1, CIDEB, CLN3, CLN5, CLN6, CLPTM1L, CLU, COL18A1, CREB1, CSE1L, CTSB, CTSD, CUL1, CYCS, CYFIP2, DAP3, DAPK3, DAXX, DCTN1, DDX19B, DEDD, DHRS2, DIDO1, DLC1, DNAJA3, DNAJB6, DNAJC5, DNM2, DPF1, DPF2, DYRK2, EGFR, EGLN3, EIF2AK2, EIF4G2, EIF5A, ELMO2, ERBB2, ERCC1, ERCC2, ERCC3, EYA2, FADD, FAF1, FAIM, FASTK, FGD4, FKBP8, FOXO3, FUS, FXR1, GADD45B, GAN, GCH1, GCLC, GDNF, GGCT, GSN, GSPT1, GSTP1, HAND2, HDAC6, HERPUD1, HIPK1, HSP90B1, HSPB1, HSPD1, HTATIP2, IFI6, IFT57, IKBKB, IKBKG, ILK, ING4, IP6K2, IRAK1, ITSN1, JAG2, KCNMA1, KRT8, LITAF, LUC7L3, MADD, MAEA, MALT1, MAP1S, MAP3K10, MAP3K11, MAP3K7, MAPK7, MAPK9, MARK4, MCL1, MDM4, MEF2A, MEF2C, MEF2D, MEN1, MGMT, MKL1, MLH1, MOAP1, MSH6, MYC, MYD88, NDUFS1, NEFL, NEK6, NGEF, NME3, NPM1, NQO1, NUDT2, NUPR1, P2RX4, PACS2, PAWR, PCGF2, PCSK6, PDCD10, PDCD2, PDCD6, PEA15, PEG10, PGAP2, PGP, PHB, PIGT, PIM3, PML, PNPLA6, POLB, POLR2G, PPARD, PPIF, PPP2R1A, PRAME, PRDX1, PRKCZ, PRKDC, PSEN2, PSMC5, PSME3, PSMG2, PTRH2, PUF60, RAD21, RAF1, RB1CC1, RELA, RHOA, RHOT2, RIPK2, RNF216, RNF34, RPS27A, RPS3, RPS3A, RTEL1, RTN3, RTN4, RYR2, SCRIB, SEMA4D, SFRP1, SHARPIN, SHISA5, SIRT1, SKP2, SMNDC1, SOCS2, SOD1, SON, SOX4, SPG7, SPIN2B, SPTBN2, STAMBP, STEAP3, STRADB, SYNE1, SYVN1, TAF9, TAOK2, TARDBP, TBP, TBRG4, TCF7L2, TERF1, TIA1, TIAM2, TMEM173, TMX1, TNFRSF1A, TP53, TP73, TPT1, TRAF3, TSC22D3, TSTA3, UBC, UNC5B, VAPB, VDAC1, VEGFA, WWOX, YARS, YWHAE, YWHAZ, ZBTB16, ZC3HC1, ZDHHC16, ZFYVE27
Golgi /vesicle transport: GO:0048193 GO:0016192	 43 132	 1.6 4.8	 2.2E-06 3.8E-06	ABCA7, ACR, AGAP3, AP1B1, AP1M1, AP1S1, AP2A2, AP2B1, AP2M1, AP3M2, AP4B1, APOE, ARAP3, ARF3, ARF5, ARFGAP1, ARFGAP2, ARL1, ARRB2, ATL2, ATP5B, ATP6V1H, BCAP31, CD302, CLINT1, CLN3, CLTA, CLTB, CLTCL1, COG2, COG4, COG5, COPA, COPZ1, CORO1C, CPNE1, CPNE3, CTBP1, CUX1, CYTH1, DAB2, DBNL, DCLK1, DENND1A, DNM2, DTNBP1, DYNC2H1, EHD4, ELMO2,

GO:0007030	10	0.4	2.8E-03	EPN1, EPN2, EPS15L1, ERC1, ERGIC3, EXOC7, FLNA, FNBP1, FTH1, GABARAPL2, GGA1, GGA3, GOLGB1, GORASP1, GOSR2, GSN, GTF2H2, GTF2H2C, HOOK3, HSPA8, ITGAV, ITSN1, ITSN2, KIF17, LMBR1L, LRP1, M6PR, MAP2K1, MFGE8, MRC2, MYH10, NECAP2, NEDD4, NKD2, OSBPL5, PACSIN3, PI4KB, PICALM, PREB, PUM1, RAB18, RAB3IP, RAB7A, RABEPK, RAMP2, RBM12, RPH3AL, RTN3, RUFY1, SAR1A, SAR1B, SCAMP5, SCARB1, SCFD1, SCYL1, SDF4, SEC16A, SEC23B, SEC24D, SEC31A, SEPT5, SH3GL1, SNX17, SNX3, SORL1, SPTBN2, STX5, STX8, STXBP2, STXBP6, SYNE1, SYTL1, SYTL5, TFRC, TMED10, TMX1, TOM1, TRAPPC1, TRAPPC3, TRAPPC4, VAMP7, VAMP8, VPS13A, VPS33A, VPS41, VPS45, VTI1B, WASF2, YIPF5, YKT6, YWHAZ
Organelle organization/ localization: GO:0033043 GO:0051129 GO:0010639	 60 37 24	 2.2 1.4 0.9	 9.4E-06 1.9E-03 3.1E-03	ACTR3, ACTR3B, ARAP1, ARHGDIA, ARHGEF1, ARHGEF2, ARPC1A, ARPC3, ARPC5, CAPG, CAPZA2, CAPZB, CCDC88A, CDC16, CDC42, CDK13, CENPV, CEP250, CFL1, CLU, CST3, CUL7, CUL9, DLC1, DNMT1, ERCC1, ERCC4, GSN, HDAC6, HECTD3, ILK, KATNB1, LINGO1, MACF1, MAD2L2, MAP1B, MAP4, MEN1, MID1IP1, MLST8, MYC, MYCBP2, NEK6, NF2, NPM1, PACSIN3, PDGFA, PFN2, PKMYT1, PML, PRKCZ, PRMT5, PSMG2, RCC1, RHOA, ROCK2, RTN4, SBDS, SCAMP5, SEMA4F, SET, SPTBN2, STMN3, SYNPO, TERF1, TERF2IP, TGFBR3, TMSB10, TSC1, UBE2N, XPO1
Protein localization: GO:0008104 GO:0070727 GO:0045184 GO:0034613	 186 98 164 97	 6.8 3.6 6.0 3.5	 1.1E-05 2.0E-05 2.0E-05 2.5E-05	AAGAB, ACTN4, AFTPH, AGAP3, AKT1, ALG2, AP1B1, AP1M1, AP1S1, AP2A2, AP2B1, AP2M1, AP3M2, AP4B1, ARF3, ARF5, ARFGAP1, ARFGAP2, ARHGEF2, ARNTL, ATG16L1, ATG4A, ATG4B, ATG9A, AURKB, BACE2, BBS4, BCAP31, BID, BIN3, CD81, CDC37, CDC42, CEP250, CEP290, CEP57, CHM, CHMP2A, CHMP4A, CHMP6, CLTA, CLTB, CLTCL1, COG2, COG4, COG5, COPA, COPZ1, CSE1L, DDX19A, DDX19B, DERL1, DERL2, DYNC2H1, EGFR, EIF5A, ERC1, ERCC3, EXOC7, FAF1, FLNA, FLNB, GABARAPL2, GDI2, GGA1, GGA3, GIPC1, GNAS, GORASP1, GOSR2, GPAA1, HDAC6, HOOK3, HSP90B1, ICMT, IPO11, IPO13, IPO4, IPO5, KATNB1, KIF17, KPNA4, KPNB1, LMAN2, LMF1, LONP2, MACF1, MYO1C, NACA, NASP, NECAP2, NEDD4, NFKBIB, NPM1, NUP37, NUP54, NUP62CL, NUP85, NUP88, NUP93, NUP98, NUTF2, OS9, PALM, PCNA, PEX10, PEX19, PEX5, PITPNM1, PML, PPP3CA, PREB, PSEN2, PTPRU, RAB11FIP3, RAB11FIP4, RAB18, RAB1B, RAB32, RAB38, RAB3C, RAB3IP, RAB5B, RAB7A, RAMP2, RAN, RANBP3, RPH3AL, RRBP1, RUFY1, SAR1A, SAR1B, SCAMP5, SCFD1, SEC16A, SEC23B, SEC24D, SEC31A, SEC61A1, SFT2D3, SNUPN, SNX17, SNX27, SNX3, SNX5, SNX6, SNX9, SRP54, SSR2, STAU1, STRADB, STX5, STXBP2, SUPT7L, SYNGR1, SYTL1, SYTL5, TACC3, TAOK2, TERF2IP, TIMM50, TIMM9, TLN2, TMED10, TMSB10, TNPO2, TOM1, TOM1L1, TOM1L2, TOMM40, TP53, VAMP7, VPS11, VPS13A, VPS28, VPS33A, VPS37D, VPS41, VPS45, VTA1, VTI1B, XPO1, XPO6, YIPF5, YKT6, YWHAE, YWHAZ
Anti-apoptosis: GO:0006916	 57	 2.1	 1.6E-05	AATF, AKT1, AKT1S1, ANXA4, ANXA5, APOE, ARHGDIA, BAG1, BCL2L1, BECN1, BIRC5, BIRC6, BNIP3, CDK1, CFL1, CIAPIN1, CLU, GCLC, GDNF, GSTP1, HSP90B1, HSPB1, HTATIP2, IFI6, IKBKB, IRAK1, MALT1, MCL1, MEF2C, MKL1, MYC, MYD88, NPM1, PEA15, PGAP2, POLB, PRKCZ, RELA, RIPK2, RPS27A, RTEL1, SEMA4D, SFRP1, SKP2, SOCS2, SOD1, SON, STAMBP, STRADB, SYVN1, TCF7L2, TMX1, TPT1, UBC, VEGFA, YWHAZ, ZC3HC1
Macromolecular / ribonucleoprotein complex assembly: GO:0065003 GO:0043933 GO:0006461 GO:0070271 GO:0034621 GO:0034622	 145 152 114 114 78 47	 5.3 5.6 4.2 4.4 2.9 1.7	 2.0E-05 3.4E-05 3.6E-05 3.6E-05 1.8E-03 1.4E-03	AASS, ABI2, ADSL, ANXA5, APOE, ARPC4, ASF1B, ATG16L1, ATL2, ATPAF1, ATPAF2, BCS1L, BRF1, CAPG, CAPZA2, CCNH, CDK7, CENPV, CLNS1A, COG4, COX11, CSE1L, DCTPP1, DDX1, DECR1, DGKD, DICER1, EIF2A, EIF6, EPRS, ERCC2, ERCC3, EVL, FADD, FANCA, FANCC, FKBP1A, FLNA, GCH1, GEMIN6, GFM2, GPAA1, GPX3, GRB2, GSN, GTF2A1, GTF2F1, GTF2H2, GTF2H2C, GTF2H3, H2AFY, H3F3B, HIST1H2AC, HIST2H4A, HJURP, HMGA1, HSPA4, HSPD1, IKBKAP, ILK, IPO11, IPO13, IPO4, IPO5, IRAK1, KPNB1, LONP1, MALT1, MAP2K1, MAP3K11, MAZ, MCM2, MDM4, MED16, MED24, MIS12, MLH1, MPP6, MRRF, MYC, NAP1L1, NAP1L4, NASP, NDUFS7, NDUFS8, NEFL, NPM1, NUP98, P2RX4, PDSS1, PEX5, PFKP, PICALM, PIH1D1, PILRB, PML, POLR2D, POLR2G, POLR2H, POLR2I, PPP2R1A, PPP5C, PRKAB1, PRKAG1, PRKCZ, PRMT5, PRMT7, PRPF31, PSMG1, PSMG2, RRM1, SCARB1, SEPT9, SET, SF1, SF3A1, SHPRH, SMAD4, SMARCE1, SMNDC1, SNRPD3, SNUPN, SPAG9, SRP54, SRR,

				STK16, TAF6, TAF9, TARBP2, TBP, TBPL1, TCP1, TERF1, TFAM, TNFRSF1A, TNPO2, TP53, TRIM21, TRIM4, TSC1, TSPYL2, TUBA1B, TUBA1C, TUBA4A, TUBB6, TUBB8, TUBE1, TUBGCP2, TUBGCP4, USP39, VAMP8, XPO1, XPO6
Microtubule/ cytoskeleton based process: GO:0007010 GO:0000226 GO:0051493 GO:0007017 GO:0010970 GO:0030036	 102 41 38 60 12 52	 3.7 1.5 1.4 2.2 0.4 1.9	 2.3E-05 2.4E-04 4.0E-04 9.0E-04 3.1E-03 4.0E-03	ABI2, ABLIM1, ACTG1, ACTN4, ACTR10, ACTR3, ACTR3B, ALDOA, APOE, ARAP1, ARAP3, ARHGAP10, ARHGAP8, ARHGEF2, ARPC1A, ARPC3, ARPC4, ARPC5, ATP2C1, BBS4, BIN3, BUB1B, CAPG, CAPZA2, CAPZB, CCDC88A, CDC42, CEP250, CETN3, CFL1, CKAP5, CNN2, CNN3, DBN1, DCTN2, DIAPH3, DLC1, DLG1, DNAJB6, DOCK7, DYNC1I2, DYNC2H1, EPB41L1, EPB41L2, EVL, FERMT2, FGD4, FHOD1, FLNA, FLNB, GSN, HAUS4, HAUS5, HAUS6, HAUS8, HDAC6, HOOK3, ILK, INF2, INPP5K, ITGB1, KATNB1, KIF17, KIF21A, KIF2A, KIF3C, KIF5C, KLC1, KLC2, KRT19, KRT8, LMO7, MACF1, MAEA, MAP1B, MAP1S, MAP2K1, MAP3K11, MAP4, MARK1, MARK4, MID1IP1, MLH1, MLST8, MTSS1, MYCBP2, MYH10, MYO9B, NCOR1, NDE1, NDEL1, NEFL, NF2, NPHP4, NPM1, OBSL1, OFD1, PALM, PCM1, PCNT, PDGFA, PFN2, PLS3, PPP1R9A, PPP4C, PRC1, PRKCZ, PRR5, RAF1, RAN, RCC1, RHOA, RHOT2, RND1, ROCK2, SBDS, SOD1, SPG7, SPTBN2, STMN3, SYNPO, TACC3, TAOK2, TERF1, TLN2, TMSB10, TSC1, TUBA1B, TUBA1C, TUBA4A, TUBB6, TUBB8, TUBE1, TUBGCP2, TUBGCP4, TUBGCP6, UBE2C, UCHL1, WASF2, XPO1, ZWINT
Cell proliferation: GO:0008283	 102	 3.7	 2.3E-05	AMBN, ANXA7, ARHGEF1, ASPM, BCAT1, BST2, BUB1B, CBFA2T3, CD276, CD81, CDC16, CDC25A, CDC27, CDK7, CSE1L, CSRP2, CTF1, CUL1, DAB2, DCTN2, DLG1, DPH1, E4F1, EGFR, EMP2, EPHB4, ERBB2, ERBB4, ERCC1, ERCC2, FKBP1A, FTH1, GLUL, GNAI2, GNB1, HDGF, HHEX, HHIP, HSPD1, IFNAR2, IGFBP4, ILK, IMPDH1, IMPDH2, IRF6, LIPA, LRP1, MALT1, MAP2K1, MAP3K11, MAPRE2, MCM7, MDM4, MKI67, MORF4L1, MTCP1, MYC, MYH10, NASP, NDE1, NDEL1, NF2, NUDC, NUMB, PA2G4, PCNA, PDGFA, PDPN, PDXK, PLK1, PPARD, PRDX1, PRMT5, PRPF19, PSPH, RAF1, RAP1B, RAPGEF3, RIPK2, SBDS, SCARB1, SCRIB, SKP2, SRRT, STIL, TACC3, TBX1, TCF19, TCF7L2, TGFBR3, TMX1, TP53, TPX2, TXNRD1, UBE2V2, UBR5, UCHL1, USP8, VEGFA, VEGFC, VTI1B, ZEB1
RNA processing/ splicing: GO:0016071 GO:0008380 GO:0006397 GO:0006396 GO:0000398 GO:0000377 GO:0000375	 89 70 77 119 43 43 43	 3.3 2.6 2.8 4.4 1.6 1.6 1.6	 2.6E-05 9.6E-05 1.1E-04 1.3E-04 1.4E-04 1.4E-04 1.4E-04	AUH, BRF1, CDK5RAP1, CLNS1A, CPSF1, CPSF3L, CPSF4, CPSF7, CSTF1, CSTF2T, DDX1, DDX5, DDX56, DGCR8, DHX9, DICER1, DUS1L, DUS3L, EIF3E, ELAC1, ELAC2, FTSJ1, FTSJ2, FUS, GDNF, GEMIN6, GRSF1, GSPT1, GTF2F1, GTPBP3, HNRNPA1, HNRNPC, HNRNPD, HNRNPF, HNRNPH1, HNRNPH3, HNRNPK, HNRNPL, HNRNPM, HNRNPUL1, IMP4, INTS10, INTS3, INTS7, INTS8, KHDRBS3, KHSRP, KIAA0391, LSM4, LUC7L3, MLH1, MOV10, NOLC1, NONO, NOP2, NSUN2, PA2G4, PABPC1, PABPC4, PAIP1, PCBP2, PCBP3, PCF11, PLRG1, PNRC2, POLR2D, POLR2G, POLR2H, POLR2I, PPIE, PPIL3, PPP2R1A, PPP4R2, PRMT5, PRMT7, PRPF19, PRPF31, PRPF40A, PRPF8, PTBP1, PTBP2, PUF60, PUS1, PUS3, RBM23, RBM3, RBM39, RBM4, RBM4B, RNH1, RNPS1, RPL36A, RPL7, RPS7, RSL1D1, SART3, SBDS, SF1, SF3A1, SF3B1, SF3B2, SLBP, SMG7, SMNDC1, SNRNP27, SNRPD3, SNUPN, SRPK1, SRRM2, SRRT, SSU72, TARBP2, TARDBP, TFB2M, TFIP11, THOC3, TRA2B, TRMT1, TRMT2B, TRMU, TSEN2, U2AF2, USP39, UTP18, UTP6, VEGFA, WDR55, WTAP, YBX1, ZFP36L1, ZNF638
Positive regulation of molecular function/ catalytic activity: GO:0044093 GO:0043085	 129 115	 4.7 4.2	 3.8E-05 8.6E-05	ACR, ACTN4, ADCY3, ADCYAP1R1, AGK, AKT1, AMH, ANAPC1, ANAPC11, ANAPC5, APOE, ARAP1, BCL2L13, BUD31, CALM3, CCDC88A, CCNB1, CD81, CDC16, CDC20, CDC25B, CDC27, CDC42, CDK1, CUL1, CYCS, DAXX, DBNL, DGKA, DGKD, DGKZ, DLC1, DNAJA3, DOCK7, EFNA1, EGFR, ERBB2, ERC1, ERCC2, FGD4, FKBP1A, FZR1, GABARAPL2, GADD45B, GCH1, GDNF, GNAI2, GNAS, GNB1, GSPT1, HIF1A, HSPD1, IFT57, IKBKB, IKBKG, ILK, IRAK1, LRP1, MADD, MALT1, MAP2K1, MAP3K10, MAP3K11, MAP3K6, MAP3K7, MEN1, MOAP1, MSH6, MYC, MYD88, NDEL1, NHEJ1, NPM1, NTRK2, PILRB, PLK1, PML, PPP2R4, PRKACB, PRKAG1, PRKAR1B, PRKCZ, PSEN2, PSMA1, PSMA2, PSMA3, PSMA4, PSMA6, PSMA7, PSMB7, PSMC2, PSMC3, PSMC4, PSMC5, PSMC6, PSMD1, PSMD10, PSMD12, PSMD13, PSMD3, PSMD4, PSMD6, PSMD8, PSME1, PSME3, PSMF1, RELA, RIPK2, RPS27A, RPS3, SCARB1, SCRIB, SKP1, SMARCA4, SMARCB1, SOD1, SPAG9, STRADB, TAOK2, TERF1, TGFBR3, TMEM189, TP53, TSC1, UBC, UBE2C, UBE2E1, UBE2N, UBE2V1, WRAP53

Posttranscriptional regulation of gene expression: GO:0010608	 56	 2.0	 6.7E-05	AKT1, APTX, BRF1, CDK4, DDX1, DGCR8, DHX9, DICER1, DNAJA3, EIF2A, EIF2B1, EIF2B4, EIF2S1, EIF3E, EIF3H, EIF3K, EIF4A2, EIF4B, EIF4E, EIF4E2 EIF4G1, EIF4G2, EIF4H, EIF5, EIF5A, EIF5B, FAM129A, FLNA, GDNF, GIPC1, HNRNPD, HSPB1, HSPD1, MDM4, MIF4GD, MKNK2, MOV10, PA2G4, PABPC1, PAIP1, PML, PRKDC, PUM1, PUM2, RBM3, SND1, SOX4, SRRT, TARBP2, TIA1, TNRC6B, TNRC6C, TSC1, VEGFA, YBX1, ZFP36L1
Phosphorylation, regulation of phosphorylation: GO:0016310 GO:0006468 GO:0006796 GO:0006793	 166 139 193 193	 6.1 5.1 7.1 7.1	 8.1E-05 2.7E-04 2.8E-04 2.8E-04	ABI2, AGK, AKAP9, AKT1, AKT2, ATP5A1, ATP5B, ATP5G2, ATP5O, ATP6V0C, ATP6V1D, ATP6V1E1, ATP6V1H, AURKB, BMP2, BRSK1, BRSK2, CAMK1, CAMKK2, CAMKV, CASK, CD81, CDC25A, CDC25B, CDK1, CDK10, CDK13, CDK16, CDK4, CDK7, CDKN3, CFL1, CHEK2, CLK3, CREB1, CSK, CSNK1A1, CSNK1G2, CSNK1G3, CSNK2A1, CTBP1, DAPK3, DAXX, DBNL, DCLK1, DDR1, DDR2, DGUOK, DLD, DMPK, DUSP3, DYRK2, EEF2K, EFNA1, EGFR, EIF2A, EIF2AK2, EIF2S1, EPHA8, EPHB1, EPHB4, ERBB2, ERBB4, ERC1, ERCC3, EYA2, FASTK, FGD4, FGFR1, FGFR2, FGFR3, GADD45B, GAK, GALK2, GDF7, GK, GMFB, GNAI2, GRK6, HIPK1, IKBKAP, IKBKB, ILK, ILKAP, INPPL1, IP6K2, IPPK, IRAK1, MADD, MAP2K1, MAP2K4, MAP2K5, MAP2K7, MAP3K10, MAP3K11, MAP3K6, MAP3K7, MAP4K4, MAPK14, MAPK3, MAPK7, MAPK9, MAPKAPK5, MARK1, MARK2, MARK3, MARK4, MELK, MINPP1, MKNK2, MTMR14, MYLK, NADK, NDUFA10, NDUFA7, NDUFS1, NDUFS7, NDUFS8, NDUFV1, NEK1, NEK4, NEK6, NTRK2, PAK4, PBK, PDK2, PHKG2, PI4KB, PIK3C2A, PIK3CD, PIM3, PIP5K1A, PKMYT1, PKN2, PLK1, PML, PNCK, PPA2, PPEF1, PPM1A, PPP1CA, PPP2R1A, PPP3CA, PPP5C, PRKACB, PRKAG1, PRKAR1B, PRKCZ, PRKD2, PRKDC, PTP4A2, PTPN12, PTPN2, PTPRA, PTPRS, PTPRU, RAF1, RIPK2, ROCK2, ROR1, RPS6KA2, RPS6KA5, RPS6KB1, RPS6KB2, SBF1, SCYL1, SDHAF2, SIK3, SOD1, SPAG9, SRPK1, SSH3, STK11, STK16, STK24, STK25, STK32C, STK40, STRADB, STYXL1, TAOK2, TAOK3, TGFBR3, THTPA, TIMM50, TRIM28, TYK2, ULK3, UQCR10, UQCRC1, UQCRC2, WEE1, WNK2
Membrane organization, endocytosis and invagination: GO:0016044 GO:0006897 GO:0010324	 89 51 51	 3.3 1.9 1.9	 8.2E-05 3.7E-03 3.7E-03	A4GALT, ABCA7, AP1B1, AP1M1, AP1S1, AP2A2, APOE, ARFGAP1, ARRB2, ATP5B, ATP6V1H, BCL2L1, BID, BNIP3, CCDC88A, CD302, CLINT1, CLN3, CLTA, CLTCL1, COL5A1, COPA, COPZ1, CORO1C, DAB2, DBNL, DENND1A, DNM2, DTNBP1, EGFR, EHD4, ELMO2, EMD, EPN1, EPN2, EPS15L1, FNBP1, FTH1, GDNF, GNPAT, GOSR2, GTF2H2, GTF2H2C, HSPA4, HSPA8, ITGAV, ITSN1, ITSN2, LMBR1L, LMNA, LRP1, M6PR, MFGE8, MRC2, MYH10, NDEL1, NECAP2, NEDD4, PACSIN3, PEX5, PI4KB, PICALM, PPIF, PREB, PUM1, RAB18, RAB7A, RABEPK, RAMP2, RUFY1, SAR1B, SCARB1, SEC24D, SEC31A, SH3GL1, SNX17, SNX3, SOD1, SORL1, STX8, TFRC, TIMM50, TIMM9, TOM1, TP53, VAMP7, VAMP8, VAPA, VTI1B, WASF2
Protein amino acid N-linked glycosylation: GO:0006487	 18	 0.7	 1.9E-04	ALG2, ALG6, ALG8, DDOST, DOLPP1, FUT8, LIPA, MAGT1, MAN1B1, MGAT1, MGAT2, MGAT4A, MGAT4B, MOGS, MPDU1, RPN1, STT3A, TUSC3
I-kappaB kinase/NF-kappaB cascade regulation: GO:0043123 GO:0043122	 30 32	 1.1 1.2	 3.1E-04 3.6E-04	ATP2C1, BST2, CANT1, CC2D1A, DNAJA3, EEF1D, FADD, FKBP1A, FLNA, IKBKB, LITAF, MALT1, MAP3K7, MAVS, MIB2, MIER1, MYD88, NEK6, PPM1A, PPP5C, RELA, RHOA, RIPK2, SHISA5, SLC35B2, TMEM189, TMEM9B, TNFRSF1A, TRIM13, UBE2N, UBE2V1, VAPA, WLS
Acetyl-CoA metabolic process: GO:0006084	 14	 0.5	 4.3E-04	SDHA, SDHB, ACSS1, ACO2, SUCLG2, SDHC, DLD, ACACA, IDH2, IDH3B, DLAT, MDH2, IDH3A, MDH1
Nucleocytoplasmic transport of proteins and RNA: GO:0006913 GO:0051169	 42 42	 1.5 1.5	 4.5E-04 5.9E-04	AAAS, AKT1, ANP32A, ARNTL, CAMK1, CEP57, CHTOP, CKAP5, CSE1L, DDX19A, DDX19B, DDX39B, EIF5A, FMR1, HHEX, HNRNPA1, HTATIP2, IPO11, IPO13, IPO4, IPO5, KHSRP, KPNA4, KPNB1, MALT1, MLX, MYO1C, NPM1, NUP37, NUP54, NUP85, NUP88, NUP93, NUP98, NUTF2, NXF1, PML, PPP3CA, RAE1, RAN, SET, SLC23A2, SLC25A19, SLC29A1, SLC35B2, SMG7, SNUPN, SRSF1,

GO:0015931	30	1.1	4.2E-03	SRSF3, SRSF7, STRADB, THOC3, THOC5, THOC6, TNPO2, TP53, TSC1, VDAC3, XPO1, XPO6, ZNF384
Regulation of cellular component size: GO:0032535	64	2.3	6.7E-04	ACTR3, ACTR3B, AKT1, AKT1S1, APBB1, ARHGAP5, ARPC1A, ARPC3, ARPC5, BIN3, CAPG, CAPZA2, CAPZB, CCNG1, CDH4, CDK4, CDKN1A, CFL1, CREB1, CSRP2, CYFIP1, DAB2, DCLK1, DDX5, DERL2, DGKD, ENO1, FGFR1, FGFR2, FGFR3, FHL1, GNG4, GSN, ILK, IP6K2, MAP1B, MAP2K5, MLST8, NDRG3, NEFL, NPM1, NUBP1, NUPR1, PFN2, PML, PPP2R1A, RB1CC1, RERG, RPTOR, RTN4, SEMA4F, SLC3A2, SMAD4, SPTBN2, TAF9, TGFBR3, TMSB10, TP53, TP73, TRO, TSC1, TSPYL2, VAT1, WDTC1
Regulation of cell morphogenesis: GO:0022604	36	1.3	8.2E-04	ACTR3, ALDOA, ANXA7, APOE, ARAP1, ARAP3, ARHGDIA, ARHGEF1, CDH4, CYFIP1, DLC1, EFNA1, FERMT2, FGD4, FN1, ILK, LINGO1, LZTS1, MAP1B, MYH10, NDEL1, NEFL, NUMB, PALM, PALM2-AKAP2, PDPN, PLXNB2, RHOA, RTN4, SEMA4D, SEMA4F, SMAD4, TAOK2, TGFB1I1, TGFBR3, VEGFA
Hexose metabolic process: GO:0019318 GO:0006006	48 38	1.8 1.4	9.8E-04 4.0E-03	ADPGK, AIMP1, AKR1A1, AKT1, ALDOA, ATF3, CREM, DLAT, ENO1, ENO2, ENO3, FUT8, G6PD, GAA, GALE, GALK2, GBE1, GMDS, GYG1, H6PD, KHK, LDHA, LDHB, MAN2A2, MAN2C1, MAPK14, MDH1, MDH2, MYC, PC, PCK2, PDK2, PFKFB3, PFKP, PGD, PGM1, PGM3, PHKB, PHKG2, PPARD, PPP1CA, PYGL, SLC25A10, SLC35A2, SLC37A4, TSTA3, USF1, WDTC1
Transcriptional regulation, RNA polymerase II-dependent-transcription regulation: GO:0051254 GO:0045893 GO:0045941 GO:0006357 GO:0010628 GO:0032583 GO:0006366 GO:0043193 GO:0006351	102 101 116 144 118 35 54 25 64	3.7 3.7 4.2 5.3 4.3 1.3 2.0 0.9 2.3	1.0E-03 1.2E-03 1.5E-03 1.9E-03 2.1E-03 2.5E-03 3.1E-03 3.2E-03 4.5E-03	AATF, AES, AFF4, ANKRD1, APBB1, APEX1, ARNT, ARNTL, ASH2L, ATXN7L3, BMP2, BRD8, BRF1, BUD31, CALCOCO1, CAMKK2, CAND1, CASK, CCNE1, CCNH, CDK7, CDKN1C, CENPK, CHD4, CHURC1, CNBP, CNOT7, CREB1, CRTC1, CTBP1, CUX1, DBP, DDX5, DEAF1, DNMT1, DNMT3A, DR1, E2F6, ECD, ECSIT, ELF1, ELK1, ELP2, ELP3, ELP4, ENO1, ERCC2, ERCC3, ETV4, FOS, FOXA2, FOXH1, FOXM1, FOXO3, FOXP4, GABPB1, GATA3, GATA4, GDF7, GDNF, GMEB2, GTF2A1, GTF2F1, GTF2H2, GTF2H2C, GTF2H3, GTF3C5, HAND2, HDAC10, HDAC8, HES1, HEXIM2, HHEX, HIF1A, HIRA, HLTF, HMGA1, HNRNPAB, HOPX, HOXA1, HTATIP2, IFNAR2, IKBKAP, ILF3, IRAK1, IRF3, IRF6, JAZF1, KAT2A, KAT5, LITAF, LMCD1, MAML2, MAP2K1, MAPK14, MAPK9, MAX, MAZ, MBD1, MDM4, MED15, MED16, MED22, MED24, MED6, MEF2C, MEN1, MKL1, MMS19, MORF4L2, MSL3, MTA2, MYC, NCOA1, NCOR1, NEDD4, NFAT5, NFATC4, NFE2L1, NFIC, NFYC, NIF3L1, NPAS2, NR2F2, PAWR, PBX1, PCGF2, PIAS1, PIAS2, PIR, PKIA, PKNOX1, POLR1C, POLR2D, POLR2G, POLR2H, POLR2I, POLR3D, POU2F1, PPARA, PPARD, PPM1A, PRKDC, PSMC3IP, PSMC5, RAN, RBBP8, RBM4, RELA, RHOA, RNF10, RNF14, RNF2, RNF4, RPS27A, RUVBL1, SCAP, SIRT1, SMAD4, SMAD5, SMARCA4, SMARCAL1, SMARCB1, SMARCD1, SMARCE1, SNAPC3, SOX4, SOX6, SRCAP, SREBF1, SREBF2, SUPT5H, SUPT6H, TAF5L, TAF6, TAF9, TARBP2, TARDBP, TBL1XR1, TBP, TBPL1, TBX1, TCEB1, TCF19, TCF25, TCF3, TCF7L2, TCFL5, TCOF1, TFAM, TFB2M, TFCP2, TFE3, TGFB1I1, TMX1, TNFRSF1A, TP53, TP53BP1, TP73, TRIM16, TRIM28, TSC22D1, UBC, UBN1, USF1, USF2, USP21, UTF1, VEGFA, WDTC1, XRCC6, YBX1, ZEB1, ZFHX3, ZNF143, ZNF423, ZNF462, ZNF496, ZSCAN21
Nitrogen compound biosynthetic process: GO:0044271	73	2.7	1.2E-03	ADAL, ADCY3, ADSL, AKT1, ALAS1, ALDH18A1, ALDOA, AMD1, APRT, ASMTL, ASNS, ASNSD1, ASS1, ATIC, ATP13A1, ATP13A2, ATP1A3, ATP1B3, ATP2B1, ATP2B3, ATP2C1, ATP5A1, ATP5B, ATP5G2, ATP5O, ATP6V0C, ATP6V1D, ATP6V1E1, ATP6V1H, ATP7B, AZIN1, BCAT1, BCAT2, CBS, CTPS2, ENOPH1, FECH, GART, GCH1, GLUL, HAAO, IMPDH1, IMPDH2, MAT2A, MOCS2, NADSYN1, NFE2L1, NME3, NME4, NOS1, NQO1, ODC1, P2RX4, PADI2, PAICS, PNP, PPOX, PRPS1, PRPSAP2, PSAT1, PSPH, PTS, PYCR1, PYCR2, QDPR, RRM1, SEPHS2, SLC25A13, SRR, TBPL1, THTPA, TYMS, UMPS
Biopolymer methylation: GO:0043414	22	0.8	1.5E-03	ATRX, CARM1, DMAP1, DNMT1, DNMT3A, EZH2, FOS, FTSJ1, FTSJ2, GSPT1, HEMK1, ICMT, MEN1, METTL5, NSUN2, PRMT1, PRMT2, PRMT5, PRMT6, PRMT7, SUV39H2, TFB2M
Ras and GTPase activity regulation:				ACAP3, AGAP3, ARAP1, ARAP3, ARFGAP1, ARFGAP2, ASAP3, CHM, DOCK7, EVI5L, FGD4, GDI2, MAP2K1, MLST8, NDEL1, RAB3GAP1, SCRIB, SGSM2,

GO:0032318 GO:0043087	29 32	1.1 1.2	2.3E-03 4.2E-03	SGSM3, SMAP1, STMN3, TBC1D1, TBC1D10A, TBC1D14, TBC1D16, TBC1D22A, TBC1D23, TBC1D2B, TBC1D30, TBC1D4, TBC1D9B, TSC1,
Steroid hormone receptor signaling pathway: GO:0030518	19	0.7	2.6E-03	CALCOCO1, CARM1, CCNE1, CDK7, DAXX, KAT5, MED16, MED24, NCOA1, NEDD4, PIAS1, PIAS2, PMEPA1, RAN, RBM4, RNF14, RNF4, TGFB1I1, UBR5
Regulation of cell size: GO:0008361	49	1.8	2.6E-03	RTN4, AKT1, DAB2, ARHGAP5, NUBP1, ILK, GNG4, MAP2K5, IP6K2, PPP2R1A, SLC3A2, TP53, CDK4, RPTOR, VAT1, SEMA4F, BIN3, FGFR2, FGFR1, DERL2, FGFR3, FHL1, PML, CCNG1, CDH4, TSPYL2, AKT1S1, NDRG3, DGKD, RB1CC1, TRO, NPM1, TAF9, DCLK1, ENO1, WDTC1, CREB1, MAP1B, SMAD4, DDX5, CSRP2, TP73, RERG, CDKN1A, NUPR1, TSC1, CYFIP1, TGFBR3, APBB1
Release of cytochrome c from mitochondria: GO:0001836	10	0.4	2.8E-03	BID, CASP3, GGCT, CASP7, CLU, TP53, BCL2L1, MYC, IFI6, TP73
Ribonucleotide/ purine metabolic process: GO:0009259 GO:0009150	37 35	1.4 1.3	3.6E-03 4.1E-03	ACLY, ADAL, ADSL, ALDOA, ATIC, ATP13A1, ATP13A2, ATP1A3, ATP1B3, ATP2B1, ATP2B3, ATP2C1, ATP5A1, ATP5B, ATP5G2, ATP5O, ATP6V0C, ATP6V1D, ATP6V1E1, ATP6V1H, ATP7B, ERCC3, GART, GCH1, IMPDH1, IMPDH2, LONP1, MYO9B, NADK, NDUFS1, NME3, NME4, NUDT5, PAICS, PRPS1, SLC25A13, UMPS
Positive regulation of cell development GO:0010720	21	0.8	3.7E-03	BMP2, ARHGEF1, HOXA11, PLXNB2, RELA, MAP1B, XRCC6, SMAD4, CDH4, HOXD11, PRPF19, ACTR3, VEGFC, NDEL1, ILK, NUMB, RHOA, SEMA4D, TGFB1I1, NEFL, ARHGDIA
Positive regulation of nitrogen compound metabolic process: GO:0051173	127	4.6	4.1E-03	AATF, AKT1, APBB1, APOE, ARNT, ARNTL, ATXN7L3, BMP2, BRE, BRF1, CALCOCO1, CAMKK2, CAND1, CASK, CCNE1, CCNH, CDK7, CENPK, CHURC1, CNBP, CNOT7, COMT, CREB1, CRTC1, DDAH1, DDX5, EGFR, ELF1, ELK1, ERCC2, ERCC3, ETV4, FOS, FOXA2, FOXH1, FOXO3, GATA4, GDF7, GDNF, GTF2F1, HAND2, HES1, HHEX, HIF1A, HMGA1, HNRNPAB, HOXA1, ILF3, IRAK1, IRF6, KAT5, MAML2, MAP2K1, MAPK14, MED6, MEF2C, MEN1, MKL1, MMS19, MORF4L2, MTA2, MYC, NCOA1, NFAT5, NFIC, NFYC, NIF3L1, NPAS2, NR2F2, P2RX4, PBX1, PDGFA, PIAS1, PIAS2, PKNOX1, POU2F1, PPARA, PPM1A, PRKDC, PSMC3IP, PSMC5, RAN, RBM4, RELA, RNF10, RNF14, RNF4, RPS27A, SCAP, SMAD4, SMAD5, SMARCA4, SMARCB1, SMARCD1, SOX4, SOX6, SREBF1, SREBF2, SUPT5H, TAF9, TBL1XR1, TBP, TBX1, TCF3, TCF7L2, TFE3, TGFB1I1, TMX1, TNFRSF1A, TP53, TP53BP1, TP73, TRIM16, TRIM28, UBC, UBE2N, UIMC1, USF1, USF2, USP21, UTF1, VEGFA, XRCC6, ZEB1, ZNF423, ZNF462, ZSCAN21,
Generation of precursor metabolites and energy: GO:0006091	68	2.5	4.2E-03	ACAA1, ACADVL, ACO2, ADPGK, AKT1, ALDOA, ATP5A1, ATP5B, ATP5G2, ATP5O, ATP6V0C, ATP6V1D, ATP6V1E1, ATP6V1H, CYB5B, CYCS, DLAT, DLD, ECH1, ENO1, ENO2, ENO3, ETFA, FADS2, FDX1L, FDXR, FECH, GAA, GBE1, GNAS, GYG1, IDH2, IDH3A, IDH3B, LDHA, LDHB, MDH1, MDH2, NDUFA10, NDUFA7, NDUFS1, NDUFS7, NDUFS8, NDUFV1, NFATC4, PFKP, PGM1, PHKB, PHKG2, PPARD, PPP1CA, PYGL, RUNX1T1, SDHA, SDHAF2, SDHB, SDHC, SLC25A12, SLC25A13, SLC25A3, SUCLG2, THTPA, TMX1, TXNL1, TXNRD1, UQCR10, UQCRC1, UQCRC2
Neuronal differentiation: GO:0030182	41	1.5	4.3E-03	ACTR3, APOE, ARHGDIA, ARHGEF1, ASPM, BMP2, CDH4, CDK5RAP1, CDK5RAP2, CDK5RAP3, CRABP1, DBN1, EFNA1, FOXA2, HES1, ILK, IRX3, LINGO1, LZTS1, MAP1B, NDEL1, NEFL, NOTCH3, NTRK2, NUMB, PBX1, PLXNB2, PRPF19, RELA, REST, RHOA, RPS27A, RTN4, SEMA4D, SEMA4F, TGIF2, TIMP2, TP53, UBC, VEGFC, XRCC6
tRNA aminoacylation: GO:0043039	14	0.5	4.5E-03	AIMP1, EDRS, HARS, KARS, LARS, MARS, NARS2, QARS, TARS, TARS2, VARS2, WARS, WARS2, YARS

Appendix Table A7. Gene ontology analysis of the TDP-43 Q331K differentially expressed whole-cell transcriptome at gene level

The PANTHER overrepresentation test (version 10.0 released 2015-05-15) from the Gene Ontology (GO) database (http://www.pantherdb.org/) was also used to statistically assess significant enrichment of biological processes (PANTHER GO-Slim Biological Process) and validate the DAVID analysis. N: total number of human genes associated with the given functional category. WCT: number of genes differentially expressed in the Q331K whole-cell transcriptome (WCT) for the given functional category. Exp.: number of genes expected for the given functional category without statistically significant enrichment. R: representation (+: overrepresented; -: underrepresented). FE: fold enrichment. GO terms (34) are ranked by ascending P-Values ($p < 0.005$) for the differentially expressed gene list with $\log_2$FC >1.

Biological process	GO term	N	WCT	Exp.	R	FE	P-Value
Nucleobase-containing compound metabolic process	GO:0006139	3467	275	196.1	+	1.4	2.22E-09
Metabolic process	GO:0008152	8247	565	466.4	+	1.21	3.41E-09
Primary metabolic process	GO:0044238	6825	477	385.9	+	1.24	1.60E-08
Immune system process	GO:0002376	1391	39	78.7	-	0.5	2.74E-07
Immune response	GO:0006955	518	8	29.3	-	0.27	2.83E-06
RNA metabolic process	GO:0016070	2360	184	133.5	+	1.38	5.53E-06
Translation	GO:0006412	435	47	24.6	+	1.91	3.06E-05
System development	GO:0048731	1271	41	71.9	-	0.57	3.46E-05
Response to stimulus	GO:0050896	2170	83	122.7	-	0.68	4.21E-05
Protein transport	GO:0015031	1082	92	61.2	+	1.5	9.23E-05
Protein metabolic process	GO:0019538	2692	197	152.2	+	1.29	1.02E-04
Nitrogen compound metabolic process	GO:0006807	1099	92	62.2	+	1.48	1.56E-04
tRNA metabolic process	GO:0006399	82	14	4.6	+	3.02	3.26E-04
Nervous system development	GO:0007399	823	25	46.5	-	0.54	3.28E-04
Intracellular protein transport	GO:0006886	1052	87	59.5	+	1.46	3.43E-04
Cell communication	GO:0007154	3006	130	170.0	-	0.76	3.60E-04
mRNA splicing, via spliceosome	GO:0000398	183	23	10.4	+	2.22	4.39E-04
Protein localization	GO:0008104	116	17	6.6	+	2.59	4.55E-04
DNA repair	GO:0006281	172	22	9.7	+	2.26	4.60E-04
Cell cycle	GO:0007049	1107	90	62.6	+	1.44	4.65E-04
mRNA processing	GO:0006397	274	30	15.5	+	1.94	6.35E-04
DNA metabolic process	GO:0006259	379	37	21.4	+	1.73	1.26E-03
Multicellular organismal process	GO:0032501	1640	66	92.7	-	0.71	1.52E-03
Single-multicellular organism process	GO:0044707	1636	66	92.5	-	0.71	1.64E-03
Ectoderm development	GO:0007398	663	21	37.5	-	0.56	2.17E-03
Protein targeting	GO:0006605	112	15	6.3	+	2.37	2.26E-03
Transcription, DNA-dependent	GO:0006351	1941	139	109.8	+	1.27	2.62E-03
Vesicle-mediated transport	GO:0016192	895	71	50.6	+	1.4	3.22E-03
Regulation of cell cycle	GO:0051726	26	6	1.5	+	4.08	4.02E-03
Catabolic process	GO:0009056	407	37	23.0	+	1.61	4.04E-03
RNA splicing, via transesterification reactions	GO:0000375	132	16	7.5	+	2.14	4.28E-03

Appendix Table A8. Gene ontology analysis of the TDP-43 Q331K differentially expressed cytoplasmic transcriptome at gene level

The PANTHER overrepresentation test (version 10.0 released 2015-05-15) from the Gene Ontology (GO) database (http://www.pantherdb.org/) was also used to statistically assess significant enrichment of biological processes (PANTHER GO-Slim Biological Process) and validate the DAVID analysis. N: total number of human genes associated with the given functional category. WCT: number of genes differentially expressed in the Q331K whole-cell transcriptome (WCT) for the given functional category. Exp.: number of genes expected for the given functional category without statistically significant enrichment. R: representation (+: overrepresented; -: underrepresented). FE: fold enrichment. GO terms (49) are ranked by ascending P-Values (p <0.005) for the differentially expressed gene list with log_2FC >1.

Biological process	GO term	N	WCT	Exp.	R	FE	P-Value
Metabolic process	GO:0008152	8247	1002	768.3	+	1.3	5.12E-27
Primary metabolic process	GO:0044238	6825	830	635.8	+	1.31	2.28E-20
Nucleobase-containing compound metabolic process	GO:0006139	3467	452	323.0	+	1.4	3.54E-14
Protein metabolic process	GO:0019538	2692	359	250.8	+	1.43	2.23E-12
RNA metabolic process	GO:0016070	2360	317	219.9	+	1.44	2.74E-11
Transcription, DNA-dependent	GO:0006351	1941	254	180.8	+	1.4	3.46E-08
Cellular component organization or biogenesis	GO:0071840	1316	183	122.6	+	1.49	7.70E-08
Cellular protein modification process	GO:0006464	1317	180	122.7	+	1.47	3.01E-07
Translation	GO:0006412	435	76	40.5	+	1.88	3.05E-07
Immune system process	GO:0002376	1391	78	129.6	-	0.6	3.25E-07
Organelle organization	GO:0006996	571	90	53.2	+	1.69	1.86E-06
Cellular component organization	GO:0016043	1206	162	112.4	+	1.44	3.20E-06
Nitrogen compound metabolic process	GO:0006807	1099	149	102.4	+	1.46	5.02E-06
Ion transport	GO:0006811	714	34	66.5	-	0.51	6.24E-06
Blood circulation	GO:0008015	154	1	14.4	-	< 0.2	8.62E-06
Transcription from RNA polymerase II promoter	GO:0006366	1723	214	160.5	+	1.33	1.42E-05
Cell cycle	GO:0007049	1107	147	103.1	+	1.43	1.62E-05
Sensory perception	GO:0007600	455	18	42.4	-	0.42	1.72E-05
Multicellular organismal process	GO:0032501	1640	106	152.8	-	0.69	2.10E-05
Single-multicellular organism process	GO:0044707	1636	106	152.4	-	0.7	2.38E-05
System process	GO:0003008	1296	81	120.7	-	0.67	5.00E-05
Response to stimulus	GO:0050896	2170	152	202.2	-	0.75	6.42E-05
Cation transport	GO:0006812	571	28	53.2	-	0.53	9.35E-05
Purine nucleobase metabolic process	GO:0006144	88	21	8.2	+	2.56	1.25E-04
Immune response	GO:0006955	518	25	48.3	-	0.52	1.50E-04
Cellular component biogenesis	GO:0044085	310	50	28.9	+	1.73	2.01E-04
Regulation of translation	GO:0006417	148	29	13.8	+	2.1	2.21E-04
Protein transport	GO:0015031	1082	137	100.8	+	1.36	2.43E-04
DNA repair	GO:0006281	172	32	16.0	+	2	2.65E-04
Regulation of nucleobase-containing compound metabolic process	GO:0019219	1700	202	158.4	+	1.28	2.76E-04
Vesicle-mediated transport	GO:0016192	895	116	83.4	+	1.39	3.12E-04
Intracellular protein transport	GO:0006886	1052	133	98.0	+	1.36	3.16E-04
Biosynthetic process	GO:0009058	764	101	71.2	+	1.42	3.92E-04
Neurological system process	GO:0050877	1064	68	99.1	-	0.69	4.52E-04
Mitosis	GO:0007067	367	55	34.2	+	1.61	5.69E-04
Protein phosphorylation	GO:0006468	603	82	56.2	+	1.46	6.04E-04
rRNA metabolic process	GO:0016072	115	23	10.7	+	2.15	7.14E-04

Protein targeting	GO:0006605	112	22	10.4	+	2.11	1.15E-03
Cell-cell signaling	GO:0007267	633	37	59.0	-	0.63	1.28E-03
Catabolic process	GO:0009056	407	58	37.9	+	1.53	1.30E-03
Extracellular transport	GO:0006858	113	2	10.5	-	< 0.2	1.76E-03
Biological regulation	GO:0065007	3918	416	365.0	+	1.14	1.92E-03
mRNA processing	GO:0006397	274	41	25.5	+	1.61	2.72E-03
Mitochondrial transport	GO:0006839	25	8	2.3	+	3.44	2.77E-03
System development	GO:0048731	1271	90	118.4	-	0.76	3.07E-03
Regulation of cell cycle	GO:0051726	26	8	2.4	+	3.3	3.50E-03
Regulation of transcription from RNA polymerase II promoter	GO:0006357	1319	153	122.9	+	1.25	3.68E-03
DNA metabolic process	GO:0006259	379	52	35.3	+	1.47	4.64E-03
mRNA splicing, via spliceosome	GO:0000398	183	29	17.1	+	1.7	4.98E-03

Appendix Table A9. Gene ontology analysis of the TDP-43 Q331K differentially expressed GRASPS-translatome at gene level

The PANTHER overrepresentation test (version 10.0 released 2015-05-15) from the Gene Ontology (GO) database (http://www.pantherdb.org/) was also used to statistically assess significant enrichment of biological processes (PANTHER GO-Slim Biological Process) and validate the DAVID analysis. N: total number of human genes associated with the given functional category. WCT: number of genes differentially expressed in the Q331K whole-cell transcriptome (WCT) for the given functional category. Exp.: number of genes expected for the given functional category without statistically significant enrichment. R: representation (+: overrepresented; -: underrepresented). FE: fold enrichment. GO terms (69) are ranked by ascending P-Values ($p < 0.005$) for the differentially expressed gene list with $\log_2$FC >1.

Biological process	GO term	N	WCT	Exp.	R	FE	P-Value
Metabolic process	GO:0008152	8247	2048	1624	+	1.26	4.35E-41
Primary metabolic process	GO:0044238	6825	1710	1344	+	1.27	4.38E-33
Protein metabolic process	GO:0019538	2692	733	530.2	+	1.38	1.29E-19
Nucleobase-containing compound metabolic process	GO:0006139	3467	873	682.8	+	1.28	7.03E-15
Translation	GO:0006412	435	155	85.7	+	1.81	6.04E-12
DNA repair	GO:0006281	172	80	33.9	+	2.36	8.75E-12
Cellular protein modification process	GO:0006464	1317	370	259.4	+	1.43	1.22E-11
Sensory perception	GO:0007600	455	36	89.6	-	0.4	7.24E-11
DNA metabolic process	GO:0006259	379	134	74.6	+	1.8	2.60E-10
Immune response	GO:0006955	518	46	102.0	-	0.45	2.73E-10
Multicellular organismal process	GO:0032501	1640	221	323.0	-	0.68	2.79E-10
Cellular component organization or biogenesis	GO:0071840	1316	361	259.2	+	1.39	3.39E-10
Single-multicellular organism process	GO:0044707	1636	221	322.2	-	0.69	3.67E-10
Organelle organization	GO:0006996	571	180	112.5	+	1.6	1.57E-09
Immune system process	GO:0002376	1391	185	273.9	-	0.68	2.54E-09
Cellular component organization	GO:0016043	1206	329	237.5	+	1.39	3.80E-09
System process	GO:0003008	1296	175	255.2	-	0.69	2.75E-08
Regulation of translation	GO:0006417	148	63	29.2	+	2.16	3.25E-08
Natural killer cell activation	GO:0030101	99	1	19.5	-	< 0.2	6.70E-08
Neurological system process	GO:0050877	1064	142	209.5	-	0.68	2.55E-07
Cellular amino acid biosynthetic process	GO:0008652	81	40	16.0	+	2.51	2.84E-07
Cell-cell signaling	GO:0007267	633	74	124.7	-	0.59	4.59E-07
RNA metabolic process	GO:0016070	2360	566	464.8	+	1.22	6.99E-07
Cell cycle	GO:0007049	1107	291	218.0	+	1.33	7.07E-07
mRNA processing	GO:0006397	274	93	54.0	+	1.72	7.34E-07
Catabolic process	GO:0009056	407	123	80.2	+	1.53	4.24E-06
mRNA splicing, via spliceosome	GO:0000398	183	66	36.0	+	1.83	4.35E-06
Intracellular protein transport	GO:0006886	1052	272	207.2	+	1.31	5.49E-06
Protein transport	GO:0015031	1082	277	213.1	+	1.3	9.12E-06
Response to stimulus	GO:0050896	2170	347	427.4	-	0.81	1.35E-05
Blood circulation	GO:0008015	154	10	30.3	-	0.33	1.70E-05
Protein targeting	GO:0006605	112	43	22.1	+	1.95	4.81E-05
Cation transport	GO:0006812	571	74	112.5	-	0.66	6.01E-05
Nitrogen compound metabolic process	GO:0006807	1099	272	216.4	+	1.26	1.01E-04
Cellular component biogenesis	GO:0044085	310	91	61.1	+	1.49	1.82E-04
Protein phosphorylation	GO:0006468	603	159	118.8	+	1.34	2.01E-04
Ion transport	GO:0006811	714	101	140.6	-	0.72	2.21E-04
Protein complex biogenesis	GO:0070271	108	39	21.3	+	1.83	3.44E-04
RNA splicing	GO:0008380	135	46	26.6	+	1.73	3.77E-04
Regulation of vasoconstriction	GO:0019229	52	1	10.2	-	< 0.2	3.97E-04
Neuron-neuron synaptic transmission	GO:0007270	52	1	10.2	-	< 0.2	3.97E-04
RNA splicing, via transesterification reactions	GO:0000375	132	45	26.0	+	1.73	4.28E-04
Cell communication	GO:0007154	3006	519	592.0	-	0.88	5.32E-04
Chromatin organization	GO:0006325	250	74	49.2	+	1.5	5.48E-04
Protein complex assembly	GO:0006461	107	38	21.1	+	1.8	5.50E-04
Transcription from RNA polymerase II promoter	GO:0006366	1723	397	339.3	+	1.17	7.58E-04
Vesicle-mediated transport	GO:0016192	895	219	176.3	+	1.24	8.14E-04
System development	GO:0048731	1271	203	250.3	-	0.81	8.45E-04

Purine nucleobase metabolic process	GO:0006144	88	32	17.3	+	1.85	9.86E-04
Visual perception	GO:0007601	219	24	43.1	-	0.56	1.05E-03
Skeletal system development	GO:0001501	232	26	45.7	-	0.57	1.06E-03
Transcription, DNA-dependent	GO:0006351	1941	440	382.3	+	1.15	1.27E-03
Cellular amino acid metabolic process	GO:0006520	264	75	52.0	+	1.44	1.48E-03
Mitosis	GO:0007067	367	99	72.3	+	1.37	1.49E-03
DNA replication	GO:0006260	161	50	31.7	+	1.58	1.55E-03
Mesoderm development	GO:0007498	671	100	132.1	-	0.76	1.84E-03
Protein methylation	GO:0006479	14	9	2.8	+	3.26	2.19E-03
Carbohydrate metabolic process	GO:0005975	573	144	112.8	+	1.28	2.36E-03
Response to external stimulus	GO:0009605	378	51	74.4	-	0.69	2.40E-03
Regulation of cell cycle	GO:0051726	26	13	5.1	+	2.54	2.45E-03
Tricarboxylic acid cycle	GO:0006099	20	11	3.9	+	2.79	2.52E-03
Response to biotic stimulus	GO:0009607	41	1	8.1	-	< 0.2	2.81E-03
DNA recombination	GO:0006310	50	20	9.9	+	2.03	2.88E-03
B cell mediated immunity	GO:0019724	141	14	27.8	-	0.5	3.01E-03
Biosynthetic process	GO:0009058	764	185	150.5	+	1.23	3.02E-03
Regulation of molecular function	GO:0065009	1096	256	215.8	+	1.19	3.37E-03
Protein localization	GO:0008104	116	37	22.8	+	1.62	3.83E-03
Cytoskeleton organization	GO:0007010	78	27	15.4	+	1.76	4.42E-03
Synaptic transmission	GO:0007268	331	45	65.2	-	0.69	4.99E-03

Appendix Table A10. Commonly differentially expressed TDP-43 targets

Differentially expressed transcripts common to all the three datasets with log2 FC >1.

Gene symbol	Gene description	WCT (log_2 FC)	CyT (log_2 FC)	GRASPS (log_2 FC)
PUM1	pumilio RNA binding family member 1 [HGNC:14957]	-1.1	1.6	1.7
EED	embryonic ectoderm development [HGNC:3188]	1.7	1.1	2.8
UBE2A	ubiquitin conjugating enzyme E2A [HGNC:12472]	-1.9	-2.6	2.4
HNRNPDL	heterogeneous nuclear ribonucleoprotein D like [HGNC:5037]	-1.4	-1.0	1.5
ATF7	activating transcription factor 7 [HGNC:792]	1.7	-2.4	2.3
ABI2	abl-interactor 2 [HGNC:24011]	-2.4	1.0	2.1
PABPN1	poly(A) binding protein, nuclear 1 [HGNC:8565]	3.0	1.5	1.5
ATP2C1	ATPase, Ca++ transporting, type 2C, member 1 [HGNC:13211]	-1.2	-1.6	2.2
MYO5A	myosin VA [HGNC:7602]	1.2	1.3	-1.3
PDE4DIP	phosphodiesterase 4D interacting protein [HGNC:15580]	1.1	-1.1	1.7
ACIN1	apoptotic chromatin condensation inducer 1 [HGNC:17066]	1.1	-1.3	1.9
NOMO2	NODAL modulator 2 [HGNC:22652]	-1.5	1.0	2.9
KIAA0101	KIAA0101 [HGNC:28961]	3.1	-1.4	1.7
ATL2	atlastin GTPase 2 [HGNC:24047]	1.7	1.1	2.6
DNAJB6	DnaJ heat shock protein family (Hsp40) member B6 [HGNC:14888]	-2.0	1.0	1.0
CCT5	chaperonin containing TCP1, subunit 5 (epsilon) [HGNC:1618]	-1.9	1.3	-1.5
NUP98	nucleoporin 98kDa [HGNC:8068]	1.9	1.1	2.8
VDAC1	voltage-dependent anion channel 1 [HGNC:12669]	1.8	1.2	1.5
CELF1	CUGBP, Elav-like family member 1 [HGNC:2549]	1.0	1.2	-1.3
VGLL4	vestigial like family member 4 [HGNC:28966]	-1.5	3.7	1.3
ATG16L1	autophagy related 16 like 1 [HGNC:21498]	3.1	1.1	-2.1
DYM	dymeclin [HGNC:21317]	-2.1	-1.1	4.3
ZDHHC20	zinc finger, DHHC-type containing 20 [HGNC:20749]	-1.1	-1.7	1.2
GLYR1	glyoxylate reductase 1 homolog (Arabidopsis) [HGNC:24434]	1.1	-1.8	1.4
NCOR2	nuclear receptor corepressor 2 [HGNC:7673]	-1.0	-1.7	-1.0
TUSC3	tumor suppressor candidate 3 [HGNC:30242]	2.3	1.3	1.9
ENOPH1	enolase-phosphatase 1 [HGNC:24599]	-1.5	1.0	1.7
RAB18	RAB18, member RAS oncogene family [HGNC:14244]	1.1	-1.8	1.9
MED25	mediator complex subunit 25 [HGNC:28845]	-1.1	-1.1	-2.1
NSMCE2	NSE2/MMS21 homolog, SMC5-SMC6 complex SUMO ligase [HGNC:26513]	1.1	-1.1	-2.0
DCUN1D4	DCN1, defective in cullin neddylation 1, domain containing 4 [HGNC:28998]	1.1	1.2	1.7
C1orf198	chromosome 1 open reading frame 198 [HGNC:25900]	-1.9	1.0	-3.9
PIGK	phosphatidylinositol glycan anchor biosynthesis class K [HGNC:8965]	-1.6	1.1	-1.7
DDR1	discoidin domain receptor tyrosine kinase 1 [HGNC:2730]	-1.3	-1.2	-1.4
SRSF3	serine/arginine-rich splicing factor 3 [HGNC:10785]	1.2	1.1	1.9
TPT1	tumor protein, translationally-controlled 1 [HGNC:12022]	1.4	-1.0	1.3
KHSRP	KH-type splicing regulatory protein [HGNC:6316]	1.8	-1.8	1.1
TPM1	tropomyosin 1 (alpha) [HGNC:12010]	-2.7	1.1	-1.1
DAPK1	death-associated protein kinase 1 [HGNC:2674]	1.6	1.0	1.3
UBE2V2	ubiquitin conjugating enzyme E2 variant 2 [HGNC:12495]	-1.1	-1.3	2.3
NCAM1	neural cell adhesion molecule 1 [HGNC:7656]	1.0	1.0	1.6
NASP	nuclear autoantigenic sperm protein (histone-binding) [HGNC:7644]	1.1	-2.2	-1.1
MATK	megakaryocyte-associated tyrosine kinase [HGNC:6906]	2.9	1.0	1.4
HN1	hematological and neurological expressed 1 [HGNC:14569]	-1.3	1.0	1.1
RHOBTB3	Rho-related BTB domain containing 3 [HGNC:18757]	-1.1	1.9	-1.8
SMC6	structural maintenance of chromosomes 6 [HGNC:20466]	1.6	1.0	-2.1
PAIP1	poly(A) binding protein interacting protein 1 [HGNC:16945]	1.1	-1.0	-1.3

PTOV1	prostate tumor overexpressed 1 [HGNC:9632]	-1.4	-1.2	-1.8
ADSL	adenylosuccinate lyase [HGNC:291]	1.1	-1.2	1.6
AP2B1	adaptor related protein complex 2, beta 1 subunit [HGNC:563]	-1.3	-1.1	-4.5
FLNA	filamin A [HGNC:3754]	1.6	1.3	-2.1
ADAM23	ADAM metallopeptidase domain 23 [HGNC:202]	1.5	1.1	1.8
SUMO1	small ubiquitin-like modifier 1 [HGNC:12502]	-1.8	-1.1	1.6
ITGB1	integrin subunit beta 1 [HGNC:6153]	-1.5	1.2	-4.6