MRC | Mammalian Genetics Unit

Investigating TDP43 biological dysfunction through the characterisation of *Tardbp* ENU mouse mutants: Implications for Neurodegeneration

A thesis presented for the degree of Doctor of Philosophy (D.Phil)

Wolfson College, Oxford

Hugo De São José Martinho De Oliveira, BSc. MBBS

Mammalian Genetics Unit, MRC Harwell

Supervisors: Dr Abraham Acevedo-Arozena (MGU, MRC Harwell)

Dr Peter Oliver (DPAG, Oxford)

Trinity term 2014

Investigating TDP43 biological dysfunction through the characterisation of ENU *Tardbp* mutants: implications for Neurodegeneration

Hugo Martinho De Oliveira

Wolfson College, Oxford

Mammalian Genetics Unit, MRC Harwell

Department of Physiology, Anatomy and Genetics, University of Oxford

A thesis presented for D.Phil, Trinity term, 2014.

Abstract

TDP43 is a highly conserved ubiquitously expressed protein which performs multiple roles in RNA metabolism, including regulation of gene expression, pre-mRNA splicing and the regulation of micro-RNA biogenesis. It belongs to the hnRNP family and interacts extensively with other proteins and RNA to perform its known biological functions and, analogously to other proteins from the same family, is proposed to autoregulate its own protein levels.

TDP43 biological dysfunction plays a pivotal role in the aetiology and pathophysiology of ALS and FTLD-TDP given that abnormally cleaved mislocalised fragments of TDP43, which are polyubiquitinated and hyperphosphorylated, constitute histopathological hallmarks of these diseases. Moreover, mutations in *TARDBP* (the gene encoding TDP43), most of which located in exon 6 that encodes the Glycine-rich C-terminal of the protein, are causative of ALS and FTLD-TDP. Additionally, widespread dysfunction in RNA metabolism involving TDP43 target genes is found in tissue samples from patients who suffered from these two neurodegenerative diseases.

Mechanistic dissection between TDP43 dysfunction and downstream neurodegeneration has remained elusive given that both over-expressor cellular and mouse models and conditional deletion or knockdown models have been shown to develop behavioural dysfunction and, in some cases, neurodegeneration with associated dysfunction in RNA metabolism.

In this thesis a reverse genetics approach was taken and two *Tardbp* ENU mutant mouse models were characterised, $Tardbp^{F210I}$ (mutation in the RRM2 of TDP43) and $Tardbp^{M323K}$ (mutation in the C-terminal domain of TDP43).

Strikingly, both mutants showed significant phenotypes in TDP43-dependent splicing, with the F210I mutation leading to a "loss of normal splicing function" and the M323K mutation an "augmented alternative exon selection" leading to a splicing phenotype opposite to the observed as a consequence of the "loss of normal splicing function" caused by the F210I protein at the genome-wide level.

Full behavioural characterisation of heterozygous mice for these two ENU mutations did not identify any phenotypes associated with neurodegeneration, but preliminary data suggests that homozygous mice for the M323K mutation may suffer from synaptic dysfunction. Therefore, despite developing genome-wide splicing changes as a consequence of TDP43 biological dysfunction, neither of the mutant mouse lines developed neurodegeneration.

In conclusion, the characterisation of *Tardbp* ENU mouse mutants has revealed that TDP43 dysfunction leading to "loss of normal splicing function" or "augmented alternative exon selection" is not sufficient to cause, on its own, neurodegeneration in the laboratory mouse within its normal lifespan.

Dedication

I dedicate this thesis to my mother, Raquel Martinho, and to the memory of my grandmother, Teresa Gomes (26/07/1926-15/04/2014).

Their unwavering support and unconditional love has been paramount not only to the completion of this thesis but to everything else in my life.

Acknowledgments

A large number of people have contributed to the work I performed during my DPhil.

Firstly I would like to thank Abraham Acevedo-Arozena for welcoming me into his group and for the many interesting discussions about TDP43 and neurodegeneration and the supervision throughout the last four years.

I am grateful to all members of my lab, past and present. In particular, I would like to thank Thomas Ricketts who did the initial characterisation the *Tardbp*^{F210I} mutant and Allison Landman for proof reading and comments on the thesis. A number of students spent some time in our lab and have brought good humour with them, especially Javi and Valentina. I thank Gareth Banks and Ines Heise from the Neurobehavioural group for the suggestions and exchange of ideas regarding mouse behavioural phenotyping and their collaborations in the circadian phenotyping and the Home cage automated-wheel running test, respectively.

The core services of the MRC Harwell have been very helpful and I would like to acknowledge Chris Esapa and Helen Hilton of proteomics, everyone in the genotyping service, especially Deen Quwailid and Anne Southwell, and Jeremy Sanderson from microscopy. I thank George Nicholson for the advice on statistics.

All the mice were housed at the Mary Lyon Centre and I am grateful to the entire staff in ward 6 and in particular Michelle Stewart and Liz Joynson for looking after the animals and being always available to accommodate my requests.

The scientific projects on TDP43 involved close collaboration with Pietro Fratta in the research group of Prof Elizabeth Fisher (UCL) and I am grateful to both for their contributions, support and enthusiasm. Additionally, I thank Vincent Plagnol (UCL) and his group for the collaboration in analysing the RNA-seq data. I am grateful to Prof Linda

Greensmith and Bernadett Kalmar (UCL) for the time dedicated in accommodating the axotomies in the F210I mice.

I was lucky to have been able to spend two months at the ICGEB in Trieste where I performed some of the experiments presented here. I am grateful to the group leaders, specifically Prof Baralle and Dr Buratti, for giving me the opportunity to go there and to Mauricio Budini who helped me with experiments. Jeremias Herzog collaborated with me in the Northern blots and showed me that Trieste can also be fun.

At a more personal level I am grateful to Vishal Sahni, a good friend and skilled neuroscientist, for the many discussions about science and other topics and the encouragement and advice he gave me throughout all this process. I am grateful to Luis Santos not only for his friendship but also for dedicating his time to help me making the format of the thesis presentable.

I thank the Medical Research Council, the Company of Biologists and the Genetics Society for generous funding.

I cannot thank my mother, father and sister enough for all their support throughout the last 33 years and my grandmother for all her dedication to me and my sister when we were growing up, whom unfortunately is no longer with us and thus cannot share the happiness it comes with the completion of this thesis. I dedicate this thesis to my mother and the memory of my grandmother.

Abbreviations

3' UTR	3' Untranslated Region
AD	Alzheimer's Disease
ALS	Amyotrophic Lateral Sclerosis
ASPA	Animal Scientific Procedure's Act
CNS	Central Nervous System
CTFs	TDP43 C-terminal Fragments
dpc	days post conception
DSG	Disuccinimidyl Glutarate
EMSA	Electrophoretic Mobility Assay
ENU	*N*-Ethyl-*N*-Nitrosurea
F210I	mutation in TDP43 constituting of a change from Phelynalanine to Isoleucine at residue 210 of the amino acid chain
FTLD	Frontotemporal Lobe Degeneration
FTLD-TDP/FTLD-U	Frontotemporal Lobe Degeneration with TDP43 inclusions
hnRNPs	Heterogenous Nuclear Ribonuclear Proteins
IVF	In vitro Fertilisation
M323K	mutation in TDP43 constituting of a change from Methionine to Lysine at residue 323 of the amino acid chain
MEFs	Mouse Embryonic Fibroblasts
NES	Nuclear Export Signal (in TDP43)
NLS	Nuclear Localisation signal (in TDP43)
PAGE	Polyacrylamide Gel Electrophoresis
Poly(A)1	Canonical polyadenylation site in the TDP43 transcript
Poly(A)2	Alternative polyadenylation site 2 in the TDP43 transcript
Poly(A)4	Alternative polyadenylation site 4 in the TDP43 transcript
PPI	Pre-pulse inhibition
Q101X	mutation in TDP43 constituting of an early stop mutation at Glutamine 101
RRM	RNA Recognition Motif domain in TDP43 (1 or 2)
Tardbp	The mouse gene encoding TDP43
TARDBP	The human gene encoding TDP43
TDP43	Transactive DNA Binding protein of 43 kDa
TDP43-F210I	TDP43 protein with the F210I mutation
TDP43-M323K	TDP43 protein with the M323K mutation
TDPBR	TDP Binding Region (in the 3' UTR of the TDP43 transcript)
TDP-F210I-B6	Experimental cohort of *TardbpF210I* mutants in the C57BL/6J background
TDP-M323K-B6	Experimental cohort of $Tardbp^{M323K}$ mutants in the C57BL/6J background
TDP-M323K-B6-DBA	Cohort of $Tarbp^{M323K}$ mouse mutants in the hybrid C57BL/6J-DBA background
(UG)	Uracyl-Guanidine

TABLE OF CONTENTS

CHAPTER 1. **INTRODUCTION 1**

1. TDP43 AND ITS BIOLOGICAL FUNCTION 4

1.1.1 TDP43 binds a wide range of RNA species and has genome wide effects in pre-mRNA splicing and gene expression 8

1.2 ALS AND FTLD-U: OVERVIEW OF THE AETIOLOGY, PATHOPHYSIOLOGY AND CLINICAL DISEASE 13

1.2.1 ALS: Clinical Overview 14

1.2.2 FTD: Clinical Overview 18

1.3 AETIOLOGY AND PATHOPHYSIOLOGY OF ALS AND FTLD 20

1.3.1 Neurodegenerative pathways in TDP43 proteinopathies 27

1.4 *IN VIVO* MODELS OF TDP43 PROTEINOPATHIES: GAIN AND LOSS OF FUNCTION 31

1.4.1 Mouse models of TDP43 proteinopathies 32

1.4.1.1 "Loss of function" mouse models of TDP43 proteinopathies 32

1.4.1.2 "Gain of function" mouse models of TDP43 proteinopathies 35

1.5 TDP43 AUTOREGULATION: IMPLICATIONS FOR MOUSE MODELS OF TDP43 PROTEINOPATHIES 41

1.6 OVERALL LIMITATIONS OF THE MOUSE MODELS OF TDP43 PROTEINOPATHIES 44

1.7 ENU MUTAGENESIS: CREATING MODELS FOR A BETTER UNDERSTANDING OF TDP43 BIOLOGY AND DYSFUNCTION 45

1.7.1 Archive Screening and ENU Tardbp mouse mutants: A reverse genetics approach 48

1.7.2 ENU Tardbp mutants: advancing our understanding of TDP43 biology 50

CHAPTER 2. **MATERIALS AND METHODS 53**

2.1 MOLECULAR LABORATORY TECHNIQUES 56

2.1.1 Genomic DNA extraction from embryonic and adult mouse tissue 56

2.1.2 Genotyping for Tardbp ENU mutations 56

2.1.3 Genotyping TARDBP^{G348C} transgenic animals 60

2.1.4 Harvesting, culturing and storing Mouse Embryonic Fibroblasts (MEFs) and embryonic heads ... 61

2.1.5 CFTR minigene transfection in MEFs 62

2.1.6 RNA extraction from embryonic heads and MEFs (including MEFs transfected with the CFTR minigene) 63

2.1.7 cDNA synthesis from template RNA extracted from MEFs and embryo heads 64

2.1.8 PCR splicing assay of endogenous genes: Sortilin-1 Poldip-3, Dnajc5, Pdp-1 ..... *65*

2.1.9 PCR splicing assay of the CFTR minigene *65*

2.1.10 Agarose gel electrophoresis of PCR products and quantification *66*

2.1.11 Northern Blot for TDP43 RNA *67*

2.1.12 RNA-Seq *68*

2.1.13 Whole cell protein extraction from MEFs and embryo heads *68*

2.1.14 Protein concentration measurement in samples extracted with RIPA buffer *69*

2.1.15 Immunobloting for TDP43 protein levels *69*

2.1.16 Cross-linking and homodimerisation in Tardbp^{M323K} MEFs *71*

2.1.17 Extraction of Nuclear/Cytoplasmic Protein fractions *73*

2.1.18 Antibodies used in TDP43 Immunoblotting experiments *74*

2.1.19 Synthesis and purification of recombinant GST-TDP43 proteins *74*

2.1.20 Quantification of recombinant TDP43 proteins *76*

2.1.21 GST overlay/Far Western *77*

2.1.22 Electrophoretic Mobility Assay (EMSA) *78*

2.1.23 Custom made gels for the experiments performed in the ICGEB in Trieste (Purification of recombinant protein, GST overlay/far western and EMSAs) *80*

2.1.24 Immunofluorescence *81*

2.2 IN VIVO WHOLE ANIMAL PHENOTYPING TESTS 82

2.2.1 Body Weights and Brain Mass *82*

2.2.2 Body composition *82*

2.2.3 Modified SHIRPA procedure *82*

2.2.4 Open Field *83*

2.2.5 Light/Dark Box *84*

2.2.6 Grip Strength *85*

2.2.7 Rotarod *85*

2.2.8 Home cage automated-wheel running *86*

2.2.9 Locotronic/Horizontal Ladder *87*

2.2.10 Contextual and Cued Fear Conditioning *87*

2.2.11 Acoustic Startle Response and Pre-pulse inhibition (PPI) *89*

2.2.12 Circadian phenotyping *90*

2.3 STATISTICAL ANALYSIS AND GRAPHICAL REPRESENTATIONS 90

CHAPTER 3. **RESULTS: INHERITANCE AND LETHALITY OF THE *TARDBP*^{M323K} AND *TARDBP*^{F210I} ALLELES.....93**

3.1 HOMOZYGOSITY FOR *TARDBP*^{F210I} OR *TARDBP*^{M323K} LEADS TO EARLY LETHALITY IN LABORATORY MICE OF C57BL/6J BACKGROUND 96

3.2 A HYBRID C57BL/6J–DBA BACKGROUND INFLUENCES SURVIVAL OF *TARDBP*$^{F210I/F210I}$ AND *TARDBP*$^{M323K/M323K}$ LABORATORY MICE 99

3.3 BREEDING COMPOUND *TARDBP* ENU: ADDITIONAL TOOLS TO STUDY TDP43 BIOLOGY *IN VIVO* 103

3.3.1 Breeding Tardbp$^{M323K/Q101}$ *mice* *105*

3.3.2 Breeding Tardbp$^{F210I/M323K}$ *mice* *105*

3.4 BREEDING *TARDBP* ENU MUTANTS WITH BAC *TARDBP* TRANSGENIC MICE: DEMONSTRATING THAT SINGLE POINT MUTATIONS IN TDP43 CAN LEAD TO MULTIPLE FUNCTIONAL CONSEQUENCES 106

3.5 SUMMARY OF THE INHERITANCE AND LETHALITY OF THE *TARDBP*^{M323K} AND *TARDBP*^{F210I} ALLELES 109

CHAPTER 4. **RESULTS: MOLECULAR CHARACTERISATION OF *TARDBP*^{M323K} MUTANTS AND FURTHER CHARACTERISATION OF *TARDBP*^{F210I} MUTANTS 111**

4.1 *TARDBP*^{M323K} LEADS TO DOSE DEPENDENT ALTERNATIVE EXON SPLICING CHANGES IN TARGET GENES AND THE CFTR MINIGENE IN THE OPPOSITE DIRECTION TO THOSE OBSERVED IN TDP43 KNOCKDOWN AND *TARDBP*^{F210I} 113

4.2 THE M323K AND F210I ENU MUTATIONS LEAD TO GENOME-WIDE TRANSCRIPTOME CHANGES IN VIVO 120

4.3 INVESTIGATING THE BIOLOGICAL MECHANISM LEADING TO "AUGMENTED ALTERNATIVE EXON SELECTION" IN THE M323K MUTATION 125

4.3.1 The M323K mutation does not lead to significant changes in TDP43 protein levels or its subcellular localisation *127*

4.3.2 The TPD43-M323K transcript is differentially polyadenylated when compared to the wild type transcript *131*

4.3.3 TDP-M323K mutant protein interactions with RNA and other proteins: investigations into the biological mechanism leading to the "augmented alternative exon selection" *135*

4.3.4 TDP43-M323K mutant protein and RNA interactions *136*

4.3.5 TDP43-M323K protein-protein interactions *139*

4.3.6 GST overlay/Far Western: qualitative assessment of the interactions between TDP43 and other cellular proteins 139

4.3.7 TDP43 homodimerisation in MEFs 141

4.4 SUMMARY OF THE MOLECULAR CHARACTERISATION OF *TARDBP*^{M323K} MUTANTS AND FURTHER CHARACTERISATION OF *TARDBP*^{M323K} MUTANTS 144

CHAPTER 5. ***IN VIVO* CHARACTERISATION OF *TARDBP*^{F210I} AND *TARDBP*^{M323K} MUTANTS 147**

5.1 MULTISYSTEM PHENOTYPING OF *TARDBP*^{F210I} AND *TARDBP*^{M323K} 152

5.1.1 Body weight and body composition 152

5.1.1.1 TDP-F210I-B6 152

5.1.1.2 TDP-M323K-B6 158

5.1.1.3 TDP-M323K-B6-DBA 163

5.1.2 Brain mass: TDP-F210I-B6 and TDP-M323K-B6 165

5.1.2.1 TDP-F210I-B6 165

5.1.2.2 TDP-M323K-B6 166

5.1.3 Lifespan/longevity 168

5.1.3.1 TFP-F210I-B6 168

5.1.3.2 TDP-M323K-B6 169

5.1.4 Modified SHIRPA procedure 169

5.2 MOTOR FUNCTION PHENOTYPING OF *TARDBP*^{F210I} AND *TARDBP*^{M323K} ANIMALS 170

5.2.1 Open field: total distance travelled 170

5.2.1.1 TDP-F210I-B6 171

5.2.1.2 TDP-M323K-B6 171

5.2.1.3 TDP-M323K-B6-DBA 172

5.2.2 Grip Strength 172

5.2.2.1 TDP-F210I-B6 173

5.2.2.2 TDP-M323K-B6 174

5.2.3 Rotarod 176

5.2.3.1 TDP-F210I-B6 177

5.2.3.2 TDP-M323K-B6 177

5.2.4 Automated home cage wheel-running system 178

5.2.4.1 TDP-F210I-B6 179

5.2.4.2 TDP-M323K-B6 180

5.2.5 Locotronic/Horizontal ladder 182

5.2.5.1 TDP-F210I-B6 182

5.2.5.2 TDP-M323K-B6 183

5.3 ENDOPHENOTYPES OF NEUROPSYCHIATRIC DISEASES AND LEARNING AND MEMORY 185

5.3.1 Open Field and Light/Dark box 186

5.3.1.1 TDP-F210I-B6 186

5.3.1.2 TDP-M323K-B6 186

5.3.1.3 TDP-M323K-B6-DBA 187

5.3.1.4 TDP-F210I-B6 188

5.3.1.5 TDP-M323K-B6 189

5.3.2 Contextual and cued fear conditioning 190

5.3.2.1 TDP-F210I-B6 192

5.3.2.2 TDP-M323K-B6 194

5.3.2.3 TDP-M323K-B6-DBA 198

5.3.3 Pre-pulse inhibition and acoustic startle response 201

5.3.3.1 TDP-F210I-B6 203

5.3.3.2 TDP-M323K-B6 204

5.3.3.3 TDP-M323K-B6-DBA 205

5.3.4 Circadian phenotyping: TDP-F210I-B6, TDP-M323K-B6 208

5.3.4.1 TDP-F210I-B6 209

5.3.4.2 TDP-M323K-B6 210

5.4 SUMMARY OF THE *IN VIVO* CHARACTERISATION OF *TARDBP*^{F210I} AND *TARDBP*^{M323K} MUTANTS 211

CHAPTER 6. **DISCUSSION AND CONCLUSIONS 213**

6.1 THE *TARDBP*^{F210I} AND *TARDBP*^{M323K} MUTATIONS ARE FUNCTIONAL DURING EMBRYONIC DEVELOPMENT AND INTERACT WITH GENETIC BACKGROUND OF DIFFERENT LABORATORY MOUSE STRAINS 217

6.2 THE SPLICING PHENOTYPES OF TDP43-F210I AND TDP43-M323K AND THE UNDERLYING BIOLOGICAL MECHANISM 219

6.2.1 Mutant M323K protein and RNA interactions 220

6.2.2 Mutant M323K protein and interactions with other proteins and itself 222

6.2.3 Summary of the investigations into the biological mechanism leading to "augmented alternative exon selection" in the M323K protein and future work 224

6.3 EFFECTS OF THE M323K MUTATION ON TDP43 PROTEIN LEVELS AND RNA LEVELS; IMPLICATIONS FOR THE AUTOREGULATION PROCESS 224

6.3.1 Summary of implications of the experimental results of mutant M323K protein levels and RNA studies for the TDP43 autoregulation process and future work 227

6.4 *TARDBP^{M323K}* MOLECULAR CHARACTERISATION 228

6.4.1 Summary of Tardbp^{M323K} molecular characterisation and future work 229

6.5 *TARDBP^{F210I}* LEADS TO A UNIQUE EFFECT IN THE EXPRESSION OF LONG INTRON GENES: IMPLICATIONS FOR TDP43 BIOLOGY AND NEURODEGENERATION 230

6.5.1 Summary of the effect of the F210I mutation in the expression of genes with long introns and future work 232

6.6 *IN VIVO* CHARACTERISATION OF *TARDBP^{F210I}* AND *TARDBP^{M323K}* 233

6.6.1 In vivo Characterisation of Tardbp$^{F210I/+}$ 233

6.6.1.1 Summary of in vivo phenotyping of *Tardbp^{F210I}* mutants and future work 235

6.6.2 In vivo Characterisation of Tardbp^{M323K} 235

6.6.2.1 Summary of *in vivo* phenotyping of Tardbp^{M323K} *mutants and future work* 238

6.7 FROM MOUSE TO MAN: IMPLICATIONS OF THE CHARACTERISATION OF *TARDBP^{F210I}* AND *TARDBP^{M323K}* ON THE ROLE OF TDP43 IN NEURODEGENERATION 239

6.8 CONCLUDING REMARKS 246

CHAPTER 7. **BIBLIOGRAPHY 249**

APPENDIX 1 265

APPENDIX 2 293

Chapter 1.

Introduction

The Transactive DNA Binding Protein Binding of 43kDa (TDP43) was initially identified as a novel, ubiquitously expressed protein which binds the TAR DNA region of the Human Immunodeficiency Virus type I, repressing its transcription (Ou et al. 1995). Subsequent research has identified TDP43 as a major component of cytoplasmic inclusion bodies in the neurones of patients suffering from Amyotrophic Lateral Sclerosis (ALS) and Fronto-Temporal Lobar Degeneration with Ubiquitinated inclusions (FTLD-U/FTLD-TDP) (Neumann et al. 2006; Davidson et al. 2007; Cairns et al. 2007; Dickson et al. 2007; Fujita et al. 2008; Mori et al. 2008; Brandmeir et al. 2008; Pamphlett et al. 2009; Tan et al. 2013). Moreover, in a subset of patients diagnosed with ALS and FTLD-U, mutations in TDP43 are causative of these diseases (Sreedharan et al. 2008; Rutherford et al. 2008; Van Deerlin et al. 2008; Corrado et al. 2009; Chiang et al. 2012; Chiò et al. 2010; Borghero et al. 2011; Borroni et al. 2009; Synofzik et al. 2014).

The experimental work presented in this thesis consists of the characterisation of *N*-Ethyl-*N*-Nitrosurea (ENU) derived *Tardbp* mouse mutants, with the aim of contributing towards a better understanding of TDP43 biology and the functional consequences of mutations in its coding sequence, thus increasing our understanding of the pathophysiology of ALS and FTLD-U. Before the results are described and discussed, the current knowledge of TDP43's biological function together with the evidence for its role in the aetiology and pathophysiology of ALS and FTLD-U will be reviewed. In addition, the existing models of TDP43 proteinopathies will be discussed, including their limitations, in order to contextualise the advantages and limitations of characterising mouse models which are generated employing an ENU mutagenesis methodology.

1. TDP43 and its biological function

TDP43 constitutes the major protein isoform translated from the *TAR DNA Binding Protein* gene (*TARDBP*/*Tardbp)* (figure 1.1), is highly conserved across species (figure 1.2) and ubiquitously expressed (Ou et al. 1995; Wang et al. 2004; Buratti et al. 2001). It belongs to the family of Heterogenous Nuclear Ribonuclear Proteins (hnRNPs) and contains two RNA Recognitions Motifs, designated RNA Recognition Motif 1 and 2 (RRM1 and RRM2, respectively), a Nuclear Localisation Signal (NLS), a Nuclear Export Signal (NES) and, additionally, a Glycine-rich C-terminal domain (Buratti & Baralle 2001; Wang et al. 2004) (figure 1.3).

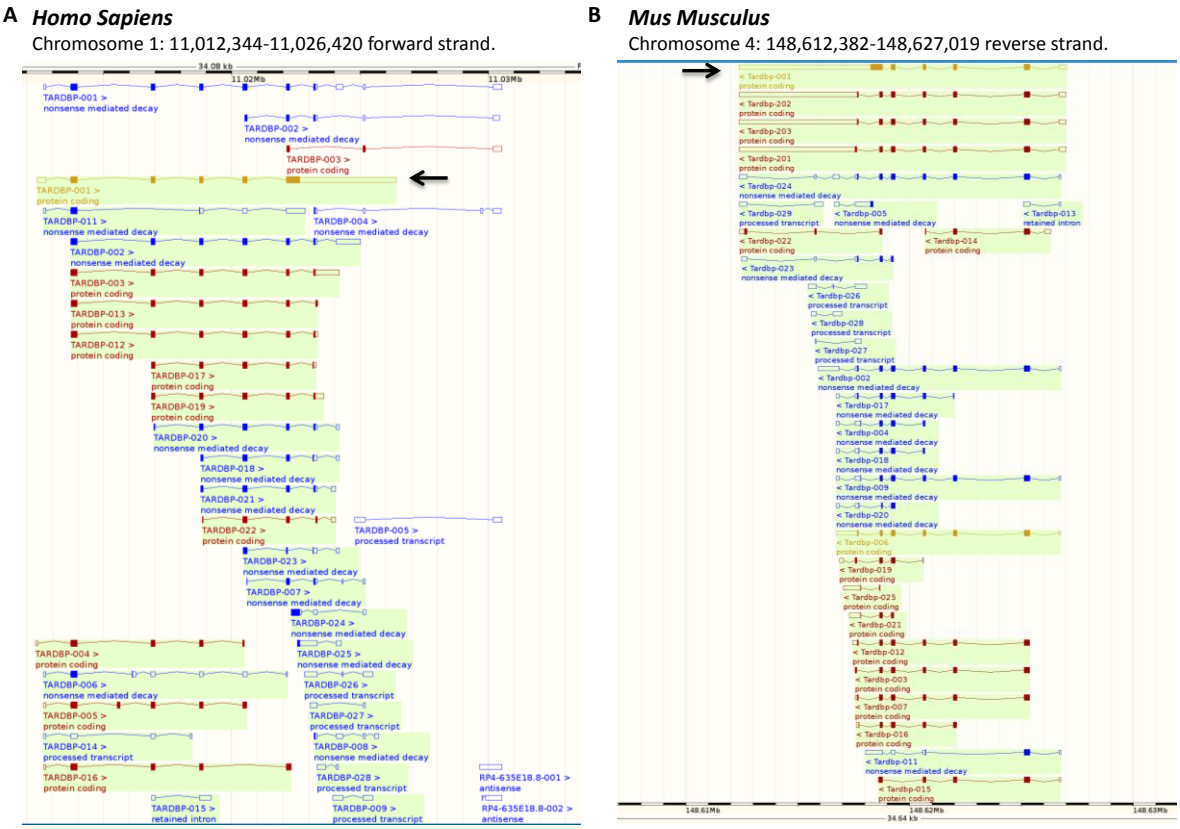

Figure 1.1: *TARDBP*/*Tardbp* transcripts and TDP43
TDP43 is the major protein isoform translated from the (A) human *TARDBP* and (B) mouse *Tardbp* genes and its corresponding transcript are indicated with an arrow at its 3' end. It is immediately obvious that both in Mouse and Man, *TARDBP*/*Tardbp* gives rise to a multitude of transcripts, some of which are translated and others that never reach translation, which may have implications in TDP43 biology (e.g. playing a role in TDP43 autoregulation as described in section 1.5). Images adapted from the genomic database Ensembl (http://www.ensembl.org).

The structure of all TDP43 domains has remained undetermined until relatively recently, with the structure of the Glycine-rich region remaining so to the present day. Nevertheless, *in vitro* and *ex vivo* studies, mostly using TDP43 constructs with deletions and point mutations to either produced recombinant protein or transfect mammalian cells, have shed light on TDP43's biological functions and interactions, mapping it to specific domains and amino acid residues. These studies have suggested that TDP43 binds to both single stranded DNA (ssDNA) and RNA, a binding interaction which was initially proposed to require a minimal sequence of six Uracyl-Guanine repeats [$(UG)_6$] in the case of RNA and which is enhanced by Thymine-Guanine (TG) repeat sequences in ssDNA oligonucleotides (Buratti & Baralle 2001). Furthermore, it has been suggested that the RRM2 is not strictly necessary for RNA binding, but acts as a modulator of the TDP43-RNA interaction, conferring it different characteristics to those observed when only the RRM1 is present in recombinant protein constructs (Buratti & Baralle 2001).

Subsequent studies have, however, identified several regions without the canonical $(UG)_6$ sequence, which are nevertheless UG/GU rich, with the UG/GU tandem repeats being interspersed by other nucleotides, to which TDP43 binds to, including the TDP Binding Region (TDPBR) on its own pre-mRNA (Tollervey et al. 2011; Colombrita et al. 2012; Xiao et al. 2011). These sequences are distributed genome-wide and are present mainly in deep intronic sites and at the 3' Untranslated Region (3' UTR) of many genes, in addition to non-coding RNAs and pre-mRNAS (Polymenidou et al. 2011; Tollervey et al. 2011), therefore suggesting a wide range of interactions between TDP43 and various RNA species and consequently, its involvement in a wide range of processes in RNA metabolism**.**

It has thus been established that TDP43 binds a wide range of RNA sequences with various affinities which are largely dependent on the number of UG repeats and lead to a

conformational change, at least *in vitro*, of the binding RNA but not of the TDP43 protein (Bhardwaj et al. 2013).

Of the constitutive TDP43 domains, a crystallography structure has been identified for the RRM2, which has an atypical RRM-fold with an additional β-strand when compared to the RNA binding domains of other hnRNPs. The RRM2 is involved not only in nucleic acid binding but also thermal-stable interactions, including TDP43 homodimerisation (Kuo et al. 2009). Additionally, the crystal structure of the RRM1 has also been reported, consisting overall of an αβ sandwich structure which interacts extensively with nucleic acids (Kuo et al. 2014). The determination of the crystal structures from the RRM domains of TDP43 and characterisation of its properties, *in vitro*, have corroborated their role in nucleic acid binding and continued to suggest that the RRM1 plays a more direct role in the protein-RNA binding than the RRM2 (Kuo et al. 2009; Kuo et al. 2014). The Glycine rich C-terminal has been identified as a region which is critical for the interaction of TDP43 with other proteins, including other members hnRNP family, specifically hnRNPA1 and hnRNPA2/B1 (Buratti et al. 2005). Moreover, the interaction between TDP43 and hnRNPA2/B1 has been finely mapped, in a cellular system, to the C-terminal region between the amino acid residues 321 and 366 (D'Ambrogio et al. 2009).

The experimental evidence therefore suggested different interactions and functions for the different domains of TDP43, with the RRM1 binding very strongly to RNA and the RRM2 not being strictly necessary for RNA binding but modulating the nature of the protein-RNA interactions and being able to promote homodimerisation (Buratti & Baralle 2001; Kuo et al. 2009), whilst the Glycine-rich C-terminal interacts with other proteins, including those of the same family (Buratti et al. 2005; D'Ambrogio et al. 2009) (figure 1.3).

Figure 1.2: Alignments of Human, Chimpanzee and Mouse TDP43 illustrating the high conservation of the protein's amino acid sequence
Alignment of the amino acid sequence of TDP43 from Man (*Homo Sapiens*), Chimpanzee (*Pan Troglodytes*) and Mouse (*Mus Musculus*). Single amino acid residues are represented according to single-letter code and the degree of conservation can be seen by the colour of each column. The homogeneity in the colour of the alignment figure immediately makes obvious the high degree of conservation in TDP43's amino acid sequence across these species.

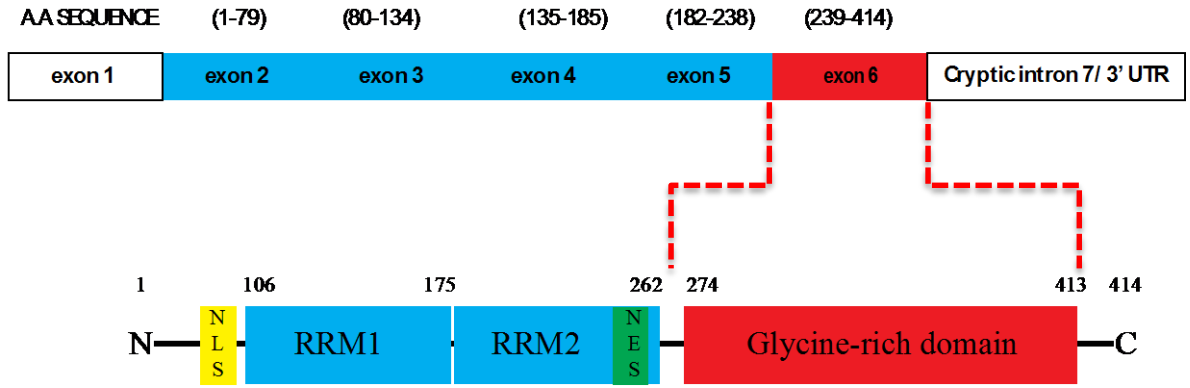

Figure 1.3: Structural domains of TDP43
TDP43 is the main protein isoform translated from the *TARDBP*/*Tardbp* gene. It is translated from 5 coding exons (exons 2-6) and has 414 amino acid residues. Its constitutive domains include a Nuclear Localisation Signal (NLS), two RNA Recognition Motifs (RRM1 and RRM2) and a Nuclear Export Signal (NES) within the RRM2. The Glycine-rich domain is encoded by exon 6 and is located at the C-terminal of TDP43. Note also in the cryptic intron 7 at the 3' Untranslated Region (3' UTR) of the transcript, which is proposed to be involved in autoregulation (section 1.5).

The first insight into a specific biological function of TDP43 in RNA metabolism came from studies investigating the splicing the *Cystic Fibrosis Transmembrane Conductance Regulator* gene (*CFTR*), which identified TDP43 as a splicing factor that acts in *trans* to promote exclusion of exon 9 from *CFTR*, in a dose dependent manner, requiring [UG(m)] repeats to be present at the 5' splicing donor site (Buratti et al. 2001). Alternative splicing of exon 9 from *CFTR* requires not only the RRM1 domain (Buratti & Baralle 2001) but also the presence of the Glycine-rich C-terminal domain in recombinant protein constructs (Wang et al. 2004).

Following from its identification as an RNA-binding protein and splicing factor, a wide range of studies have increased our understanding of TDP43-dependent splicing, a function which is conserved across species, including humans and mice.

1.1.1 TDP43 binds a wide range of RNA species and has genome wide effects in pre-mRNA splicing and gene expression

Alternative exon splicing is an extremely common event in the mammalian transcriptome and it is estimated that around 95% of all multiexon genes in the human transcriptome undergo alternative splicing (Pan et al. 2008). The current understanding of this complex process has been comprehensively reviewed (Kornblihtt et al. 2013; Gamazon & Stranger 2014). The splicing reaction consists of a two-step reaction, specifically two transesterification reactions, which result in the excised RNA sequence being released in the lariat conformation (figure 1.4) and, eventually, degraded (Kornblihtt et al. 2013). The difference between constitutive splicing and alternative splicing resides primarily in the strength of the splice site, with the former generally occurring when strong splice sites are present, which are fully used.

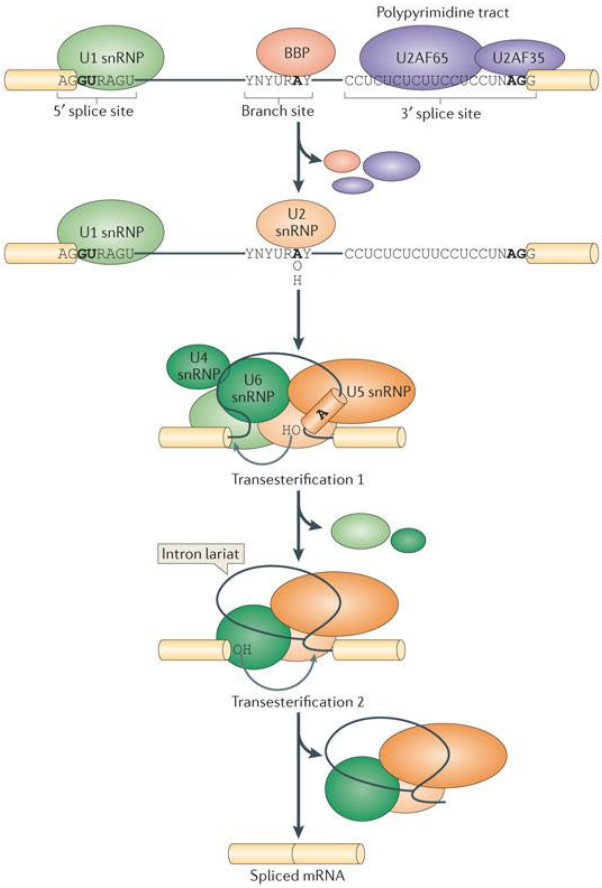

Figure 1.4: The RNA splicing reaction
The pre-mRNA contains consensus sequences at the 5' splice site, branch site and 3' splice site which are recognised by small nuclear ribonuclearproteins (snRNPs) and auxiliary factors (e.g. U2AF65 and U2AF35), which together form the spliceosome complex. Once assembled the spliceosome performs two transesterification reactions leading to release of the intron in the lariat conformation and joining of the exons. Adapted from (Kornblihtt et al. 2013).

At the regulatory level, alternative splicing involves a variety of *cis* acting factors which are present in the pre-mRNA sequence, including intronic and exonic splicing enhancers and silencers and *trans* acting factors, which include two main families of proteins, the Serine/Arginine-rich proteins and the hnRNPs, which target the components of the spliceosome (Kornblihtt et al. 2013) (figure 1.5).

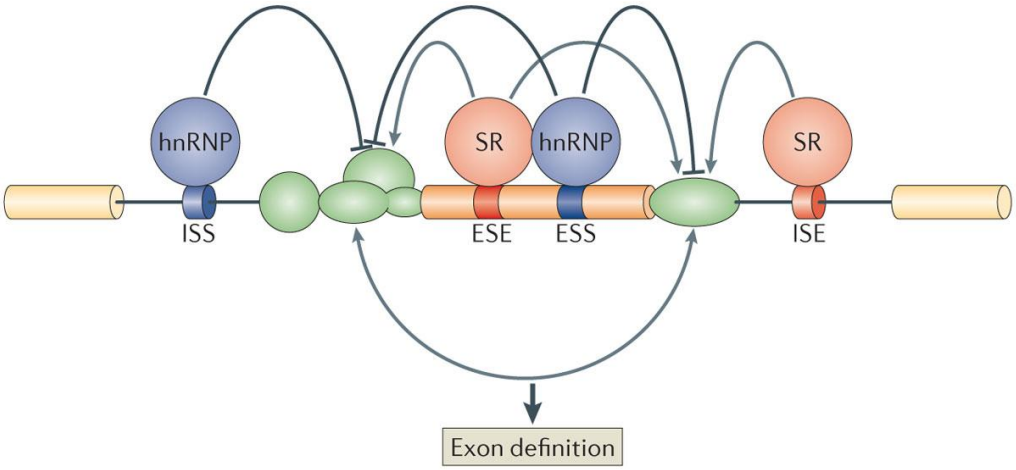

Figure 1.5: Factor determining alternative exon splicing
Alternative exon splicing is dependent on *cis*-regulatory factors and *trans*-regulatory factors. *Cis* regulatory factors are sequences present in the pre-mRNA and include Intronic Splicing Silencers (ISS), Exonic Splicing Enhancers (ESE), Exonic Splicing Silencers (ESS) and Intronic Splicing Enhancers (ISE). Cis acting factors influence binding of the spliceosome complex to the pre-mRNA. *Trans* regulatory factors include Serine/Arginine-rich (SR) proteins and the hnRNPs (represented as green shapes in the figure) and target the spliceosome, assembling at the splice sites flanking alternative exons and can either inhibit or activate the recognition of the splice site. (Kornblihtt et al. 2013).

Multiple classes of alternative splicing have been described, with the most prevalent type of alternative splicing being exon skipping/cassette exon, which amounts to approximately 40% of alternative splicing events, followed by the use of alternative 5' and 3' splice sites (Gamazon & Stranger 2014). The different classes of alternative splicing have been diagrammatically represented and further described in figure 1.6.

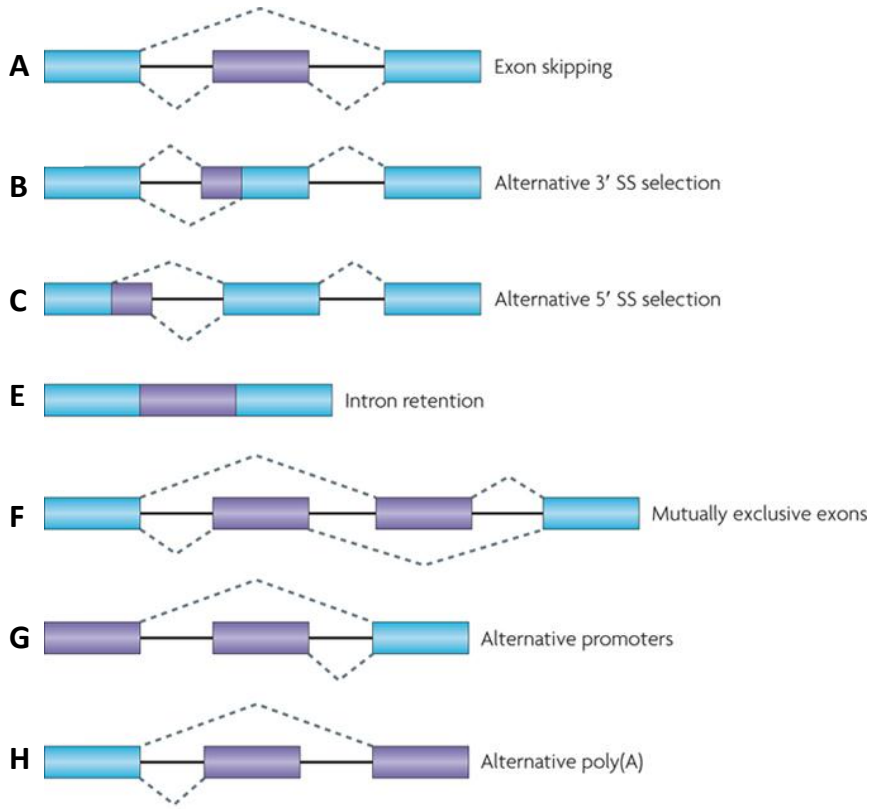

Figure 1.6: Types of alternative exon splicing
Multiple types of alternative splicing have been described. This include (A) cassette exon, in which specific exons are sometimes included and sometimes excluded of the mRNA, (B) alternative 3' splice site selection and (C) alternative 5' splice site selection as well as (E) retained intron and (G) the use of alternative splice site promoters. H) Alternative polyadenylation at the 3' Untranslated Region is also a form of alternative splicing. Adapted from (Keren et al. 2010).

Following from the early evidence that TDP43 acts in *trans* to modulate alternative splicing of exon 9 from *CFTR* (Buratti et al. 2001), additional studies demonstrated that TDP43 also acts as a splicing factor for other selected genes (Mercado et al. 2005; Bose et al. 2008; Dreumont et al. 2010), hence consolidating the evidence for its role in this process of RNA metabolism. In 2011, two seminal studies were published demonstrating the genome-wide effects of TDP43 in alternative exon splicing and gene expression.

Polymenidou et al. 2011 used the mouse as its experimental model and performed Cross-Linking and Immunoprecipitation (CLIP) of TDP43-RNA complexes from the mouse brain, using and antibody against TDP-43, followed by sequencing. Additionally, RNA-seq from the striatum of mice in which TDP43 had been knockdown with antisense oligonucleotides was also performed, thus investigating the genome wide effects of TDP43 knockdown in a specific CNS tissue (Polymenidou et al. 2011). The results were striking and demonstrated

that TDP43 binds approximately 30% of the murine transcriptome, particularly UG-rich repeats located at distal intronic sites, and that knockdown of the protein resulted in significant upregulation of 362 genes and downregulation of 239 genes. Moreover, clustering of TDP43 binding sites was observed in the downregulated transcripts suggesting that TDP43 binds directly to specific transcripts in order to maintain its levels. Interestingly, most of the genes which were downregulated by TDP43 knockdown had conspicuously long introns and are involved in synaptic function (Polymenidou et al. 2011). Additionally, TDP43 knockdown also had a remarkable effect in pre-mRNA splicing, with 203 cassette exons being differentially included or excluded, with the former (i.e. significantly more frequently included exons on TDP43 knockdown), and to a lesser extent the latter, being significantly enriched for TDP43 binding sites (Polymenidou et al. 2011).

Using a similar methodology, Tollervey et al. 2011 performed individual resolution nucleotide CLIP, followed by reverse transcription, in post-mortem cortical brain tissue from both healthy individuals and from individuals suffering from FTLD-U. In addition, using the same methodology, the authors also performed similar experiments in human Embryonic Stem Cells (ESC) and a neuroblastoma cell line (SH-SY5Y) (Tollervey et al. 2011). The results showed that TDP43 binds to a large number of nucleic acid sequences, mostly UG-rich sequences which are located in deep intronic regions. Furthermore, subcellular fractionation revealed that, in the cytoplasmic fraction of neuroblastoma cells, TDP43 was mainly bound to the 3' UTR of transcripts (Tollervey et al. 2011). Analysis of splicing changes from TDP43 knockdown in SH-SY5Y cells, assessed by splice-junction microarrays, detected 158 splicing changes in alternative cassette exons and an additional 71 splice changes of other types (Tollervey et al. 2011).

The two seminal studies succinctly described above are supported by additional studies and in conjunction, the evidence demonstrates that TDP43 binds a wide range of RNA species,

including mRNAs (Strong et al. 2007; Fiesel et al. 2010; Tollervey et al. 2011) long-non coding RNAs (Liu et al. 2012) and micro-RNAs (Buratti et al. 2010). Additionally, TDP43 is further involved in micro-RNA biogenesis by associating itself with the nuclear Drosha complex and the cytoplasmic Dicer complex (Kawahara & Mieda-Sato 2012), which cleave the precursor micro-RNAs. Mature micro-RNAs can bind other RNAs and trigger gene silencing or RNA degradation through RNA interference and are not only involved in multiple biological processes but have also been implicated in a wide range of diseases (Sayed & Abdellatif 2011). TDP43 is thus able to regulate not only mRNA levels and micro-RNA levels but also alternative exon splicing at the genome-wide level (Strong et al. 2007; Fiesel et al. 2010; Buratti et al. 2010; Tollervey et al. 2011; Polymenidou et al. 2011).

The genome wide effects observed when TDP43 expression is altered suggest that TDP43 has a pivotal role as a regulatory protein, modulating and enriching the diversity of the transcriptome and, consequently, the proteome. Notwithstanding, a deeper understanding of TDP43 biology remains elusive and, as described above, has been mainly investigated using *in vitro* and *ex vivo* experimental methodologies. Moreover, TDP43 dysfunction plays a pivotal role in the aetiology and pathophysiology of ALS and FTLD-U, two devastating neurodegenerative diseases for which very limited therapeutic interventions are currently available.

1.2 ALS and FTLD-U: Overview of the Aetiology, pathophysiology and clinical disease

ALS was firstly described by the eminent French neurologist Jean-Martin Charcot in the late nineteenth century and since then, progress has been made in further classifying the disease. Frontotemporal Dementia (FTD) was firstly described by the Czech neurologist Arnold Pick, towards the end of the eighteenth century and beginning of the nineteenth century, who described patients who presented with loss of speech but retained knowledge of action and

usage and were found to have temporal lobe degeneration in post-mortem examination (Hornberger & Piguet 2012). Over a century later, and analogously to ALS, progress has been made in further describing and classifying FTD.

Notwithstanding, a definite understanding of ALS and FTD is yet to be achieved. It has, however, become evident that these two diseases overlap in its aetiology, pathophysiology and clinical presentation, which has led to the suggestion that they may represent a clinical spectrum of the same heterogeneous disease (Strong 2008). ALS and FTD will be described in more detail, starting with the clinical aspects of the diseases and epidemiology, followed by the role of TDP43 in its aetiology and pathophysiology.

1.2.1 ALS: Clinical Overview

The incidence of ALS is approximately 2.6 per 100,000 population/year throughout most European countries, with a bias towards the male sex, with a male to female ratio of approximately 1.25:1, and the peak incidence is observed in individuals between the ages of 50 and 75 (Logroscino et al. 2010).

It is a heterogeneous disease, both in its clinical presentation and progression (Haverkamp et al. 1995), which has as its hallmark the progressive degeneration of the upper motor neurones (UMNs) in the primary motor cortex, and thus the corticospinal tracts, the brainstem and the lower motor neurones (LMNs) located in the anterior (or ventral) horn of the spinal cord (Mitchell & Borasio 2007). In the majority of cases, the disease presents with limb symptoms (in around 65% of patients), with additional clinical presentations including bulbar dysfunction (30% of patients present with dysarthria and dysphagia) and a in small percentage of cases (5%) present with respiratory onset disease (Kiernan et al. 2011; Hardiman et al. 2011).

However, the classical clinical signs of combined UMN and LMN degeneration may be absent in the initial stages of disease and the variability in clinical presentation often mimics symptoms and signs that are common to other neurodegenerative diseases. These factors, together with the fact that ALS is a relatively rare disease, contribute to a significant delay in patients being referred to a neurologist and, consequently, a significant delay in diagnosing ALS has been reported. It takes on average 14 months from symptom onset until a primary diagnosis of ALS is achieved (Chiò 1999); ALS remains ultimately a clinical diagnosis given the inexistence, at present, of specific diagnostic tests and biomarkers (Hardiman et al. 2011; Bowser et al. 2011).

Disease progression, albeit variable, tends to be fast and the median survival has been reported to be between 3 and 5 years from symptom onset (Haverkamp et al. 1995) and therefore, considering that the average time necessary to achieve a firm diagnosis of ALS is over one year, approximately half of the patients with ALS die within 2 to 4 years of diagnosis; the majority as a result of bronchopneumonia as determined by post-mortem examination (Corcia et al. 2008). In general, patients with respiratory onset disease and those presenting with bulbar dysfunction have worse prognosis than those presenting with limb-onset disease (Chiò et al. 2009; Wijesekera et al. 2009; Burrell, Vucic, et al. 2011).

The clinical management of ALS remains mostly supportive and best delivered by multidisciplinary team and tailored to each individual patient. Its aims are to minimise the symptoms and promote general wellbeing of the patients and thus improve their quality of life. Consideration must be given to the patient's nutritional status and feeding, particularly when dysphagia is a prominent symptom, and the psychosocial aspects affecting the patient and his family (Hardiman et al. 2011). Non-invasive ventilation has been shown to increase quality of life and has a moderate effect on survival, particularly in patients with better preserved bulbar function (mean survival benefit of 205 days) (Bourke et al. 2006) and

riluzole remains the only licensed pharmacotherapy shown to have a beneficial effect on survival (Bensimon et al. 1994; Lacomblez et al. 1996; Miller et al. 2012). A recent clinical trial failed to show efficacy of dexpramipexole (Cudkowicz et al. 2013), a mitochondrial agent which had revealed promising results in *in vitro* and animal studies, thus reflecting the complexity of the underlying pathophysiology and the challenge from translating results from *in vitro* and animal studies into the clinic.

Despite having been firstly described over a century ago ALS remains a diagnostic and therapeutic challenge. Its progressive disability and the scarcity of disease modifying therapeutic options have, in addition to the physical suffering, devastating psychosocial effects on the patients and their families. The degree of despair and hopelessness caused by ALS challenges the core ethical values of our societies to the extreme, which is reflected in the fact that, in a recent study performed in the Netherlands (where euthanasia is legal), almost one third of patients with ALS requested euthanasia/physician assisted suicide in spite of the majority of those patients rating the clinical care they received as very good or excellent (Maessen et al. 2014).

At a time when a private member's bill on assisted dying has been tabled in the House of Lords and is likely to be read in the commons[1], taking decisive steps forward, even if small, to gain a better understanding of the pathophysiology of ALS with the ultimate goal of achieving better diagnostic and therapeutic success, and in so far as possible carry out primary prevention, has never been more urgent.

In addition to the already described motor degeneration, in the last decade, increasing amounts of evidence have demonstrated that cognitive deficits are a common feature in ALS and therefore the underlying degenerative process does not affect the motor functions

[1] The referred bill constitutes the *Assisted Dying Bill* sponsored by Lord Falconer which had its second reading on the house of Lords on 18 July 2014, for more information visit the UK parliament website: http://services.parliament.uk/bills/2013-14/assisteddying.html

selectively, as it was proposed in the initial description of the disease. In fact, neuropsychiatric symptoms in ALS are not only associated with cortical atrophy (Mioshi et al. 2013) but often predate motor symptoms (Mioshi et al. 2014). Moreover, the cognitive deficits observed are commonly related to frontal executive functions (Phukan et al. 2007; Strong 2008; Elamin et al. 2011) and approximately 15% of patients with ALS also fulfil the diagnostic criteria for FTD (Lomen-Hoerth et al. 2002; Elamin et al. 2011).

At present, whether ALS and FTD constitute different clinical presentations from the same underlying disease, constituting part of the clinico-pathological heterogeneity, or, in contrast, constitute different entities, with ALS-FTD constituting yet a different disease itself, remains a matter of academic debate. Notwithstanding, the overlap in clinical signs and symptoms and in the aetiology and pathobiology of these diseases are highly suggestive that, regardless of whether ALS, FTD and ALS-FTD are different diseases or heterogeneous clinical manifestation of the same disease, there is a degree of convergence in the pathophysiology which leads to neuronal dysfunction and neuronal death. Consequently, the aetiology and pathophysiology of ALS and FTD are better discussed together after a clinical overview of FTD.

1.2.2 FTD: Clinical Overview

Following from its initial description by Arnold Pick, FTD has become recognised to be part of a group of heterogeneous clinical presentations that involve degeneration of the prefrontal and, or, temporal cortices, which are collectively referred to as Frontotemporal Lobar Degeneration (FTLD) (Neary et al. 2005). The classification of FTLD can be made according to the clinical symptomatology and includes three clinical syndromes, namely:

1. FTD, which presents with changes in behaviour and social conduct (sometimes referred to as behaviour variant FTLD).
2. Semantic Dementia (SD), presenting with loss of semantic knowledge of faces, objects and languages.
3. Progressive Non-Fluent Aphasia (PNFA) which presents with progressive loss of fluency of speech with preservation of intact word comprehension.

The clinical presentation is correlated with the degree of degeneration in the brain regions involved. Atrophy of the prefrontal and anterotemporal cortices is predominant in FTD; atrophy of the middle and inferotemporal cortices in SD and asymmetric atrophy of the left temporal and frontal cortices is predominant in PNFA (Neary et al. 2005; Sleegers et al. 2010). However, as the disease progresses, these different clinical phenotypes coalesce and overlapping symptomatology is observed (Sleegers et al. 2010).

In addition to the three variants of FTLD described above, symptoms characteristic of other neurodegenerative diseases, including Progressive Supranuclear Palsy, Corticobasal Syndrome and symptoms and signs of Parkinsonism can be present in patients who meet the diagnostic criteria for FTLD (Boeve 2007). Hence, the coexistence of signs and symptoms from other neurological diseases in FTLD contribute towards the diagnostic challenge of this heterogeneous disease as the additional symptoms and signs are often considered pathognomonic of other neurodegenerative diseases.

Moreover, as already discussed in the preceding section on the clinical overview of ALS, there is significant comorbidity between FTLD and ALS (commonly designated by FTLD-ALS), with approximately 15% of patients meeting the diagnostic criteria for both diseases (Lomen-Hoerth et al. 2002; Elamin et al. 2011; Burrell, Kiernan, et al. 2011). Additionally, subtle motor dysfunction can be present in patients with FTLD who do not meet the diagnostic criteria for ALS (Burrell, Kiernan, et al. 2011), mirroring the presence of cognitive dysfunction in patients with ALS which do not meet the criteria for the diagnosis of FTLD, thus suggesting a clinico-pathological relationship between the two diseases.

Achieving a primary diagnosis of FTLD represents an even a greater challenge than achieving a diagnosis of ALS, with a longer time hiatus between symptom onset and diagnosis being reported. Overall, the average time from symptom onset to diagnosis is 3 years, reduced to approximately 1 year in patients with FTLD-ALS (Hodges et al. 2003), in concordance to the reported diagnostic delay from symptom onset to diagnosis in patients with ALS (Chiò 1999). Symptom onset in FTLD occurs, on average, between the ages of 45 and 65 (Hodges et al. 2003), with the incidence being difficult to determine but estimated to be between 2.7-4.1 per 100,000 population/year, affecting both female and male individuals without gender bias (Onyike & Diehl-Schmid 2013). Disease progression, reflects the heterogeneity of the disease itself, and therefore is highly variable and correlates with the clinical phenotype. The mean survival for all patients with FTLD is reported to be 6 years from symptom onset, with far worse prognosis in patients with FTLD-ALS who have a mean survival of 2 years from symptom onset (Hodges et al. 2003).

Analogously to ALS, no disease modifying therapies exist, and despite the common use of anticholinergic medication and memantine in FTLD (Bei Hu et al. 2010), the only class of medications for which some evidence exists for amelioration of the behavioural symptoms, but not for an effect on cognitive function, is Serotonin Selective Reuptake inhibitors (Huey

et al. 2006). Notwithstanding, the evidence is suboptimal and has not been supported, to date, by any large randomised clinical trials.

In the last two sections, clinical overviews of ALS and FTLD were given. The disability caused by the symptoms and the natural progression of these diseases, allied to the absence of effective therapeutic tools, leave no doubt regarding the suffering that patients and their relatives undergo. At present, little comfort can be offered to patients from the onset of symptoms, through the lengthy diagnostic process and a death that occurs usually with a decline in respiratory function and bulbar dysfunction, which are difficult to manage clinically and cause high levels of distress in patients and their families. ALS and FTLD do not only share clinical symptomatology but also have a heterogeneous and yet overlapping aetiology and pathophysiology, with TDP43 constituting a pivotal factor common to both diseases, as discussed in the subsequent sections.

1.3 Aetiology and pathophysiology of ALS and FTLD

The clinical heterogeneity of ALS and FTLD is accompanied by remarkable heterogeneity in its aetiology.

Epidemiological data published in the second half of the XX century unveiled that in a small number of cases of ALS a clear autosomal dominant Mendelian inheritance pattern could be observed (Kurland and Mulder 1955), which demonstrated that inheritable genetic factors are implicated in its aetiology. Nevertheless, given that the cases of familial ALS constitute a minority, estimated to be around 6% (Byrne et al. 2011), suggests that ALS is a multifactorial disease, with environmental risk factors interacting with individual genetic susceptibility to precipitate the development of the disease[2]. Notwithstanding, evidence has failed, to date, to

[2] The study cited predated the identification of ALS causative mutations in *C9ORF*. Therefore, further studies are likely to show a significant increase in percentage of patients in whom ALS is caused by genetic mutations. However, the vast majority of sporadic ALS cases continue to have idiopathic aetiology.

support the involvement of specific environmental factors in the development of ALS (Al-Chalabi & Hardiman 2013).

Advances in DNA sequencing technology and in the knowledge of our genetic code generated by the Human Genome Project has enabled scientist to build on the epidemiological data and pursue the investigation into the genetic risk factors involved in the aetiology of ALS. Thus, in 1993 the first set of dominant missense mutations associated with familial ALS were reported in the gene Superoxide Dismutase 1 (*SOD1*) (Rosen et al. 1993). Following the discovery that mutations in *SOD1* were causative of familial ALS, genetic studies identified an ever-growing number of genes in which mutations were found to be causative of ALS. A review of all genetic factors causing ALS or acting as risk factors for the development of the disease merits a standalone publication, falls outside the scope of this thesis, and is available in several peer-reviewed article (Andersen & Al-Chalabi 2011; Al-Chalabi et al. 2012; Finsterer & Burgunder 2014). A list of the genetic factors involved in ALS is given in table 1.1 and has been collated using the publications referenced and the ALSoD[3] database (Andersen & Al-Chalabi 2011; Al-Chalabi et al. 2012; Finsterer & Burgunder 2014; Abel et al. 2013).

Of particular relevance for the experimental work in this thesis was the identification of a large number of mutations in *TARDBP* which are causative of both familial and sporadic ALS (Sreedharan et al. 2008; Rutherford et al. 2008; Van Deerlin et al. 2008; Corrado et al. 2009; Chiang et al. 2012). The mutations identified constitute mainly, albeit not exclusively, dominant missense mutations leading to single amino acid changes in TDP43 and are predominately clustered around the Glycine-rich C-terminal of the protein (table 1.2, figure 1.7).

[3] The ALSoD database is maintained by King's College London and available in the following web address: http://alsod.iop.kcl.ac.uk/home.aspx

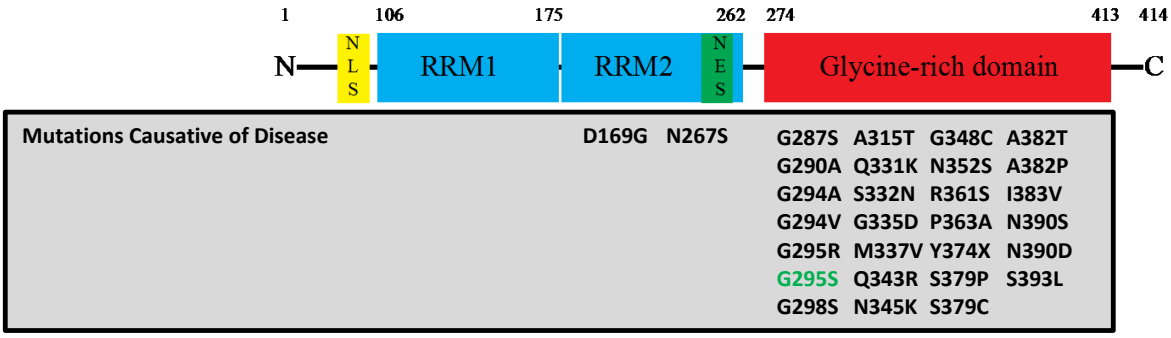

Figure 1.7: Representative diagram of TDP43 and the location of mutations which cause ALS and FTLD-TDP

Mutations in TDP43 are causative of ALS and FTLD-TDP. Most disease causative mutations constitute missense mutations and are present in the Glycine-rich C-terminal. In the figure mutations are represented as single letter of the amino acid code and the residue of TDP43 it changes (e.g. D169G consist of a mutation from Aspartic Acid to Glycine at residue 169). The mutation in green is the example of a mutation has been found to cause both FTLD-TDP and ALS.

Gene symbol	Gene name	G.I	Disease	Protein Function/pathways
ALS2	Amyotrophic lateral sclerosis 2 (juvenile) homolog (human). Alsin	AR	fALS & jALS	endosomal sorting complex/protein trafficking
FIG4	FIG4 homolog, SAC1 lipid phosphatase domain containing (S. cerevisiae)	AD	fALS & sALS	Vesicle Traficking/protein trafficking
DCTN1	Dynactin	AD	fALS;sALS & FTLD	Vesicle Traficking/protein traficking inc. axonal transport
OPTN	Optineurin	AD & AR	fALS &sALS	Protein traficking/inflammatory response
NEFH	Neurofilament, heavy polypeptide 200kDa, heavy chain	RF	sALS	Protein traficking
PRPH	Peripherin	RF	sALS	Cytoskeletal organisation/protein traficking
PFN1	Profilin 1	AD	fALS; sALS & FTLD	Actin Binding Protein
SIGMAR1	Sigma non-opioid intracellular receptor 1	AR	fALS; jALS & FTLD	Lipid transport from ER/plasma membrane biogenesis
SOD1	Cu/Zn superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	AD & AR	fALS and sALS	Antioxidant Enzyme
DAO	D-amino-acid oxidase	AD	fALS	Aminoacid oxidation
ATXN2	Ataxin 2	AD	sALS	Unknown
C9orf72	Chromosome 9 open reading frame 72	AD	fALS; sALS & FTLD	Unknown
SPG11	Spastic paraplegia 11 (AR)	AR	fALS & jALS	Unknown
CHMP2B	Chromatin modifying protein 2B	AD	fALS, sALS & FTLD	Endocytosis/Lysosomal recycling and degradation
SQSTM1	Sequestosome 1 (p62 protein)	AD	fALS & sALS	Protein degradation
UBQLN2	Ubiquilin 2	XLD	fALS; jALS; sALS & FTLD	Active in the Ubiquiting-Proteosome system
VAPB	Vesicle-associated membrane protein-associated protein B	AD	fALS	ER-associated Protein/protein unfolded response
VCP	Valosin-containing protein	AD	fALS & FTLD	Cell Division/DNA repair
ERBB4	V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 4	AD	fALS	Tyrosine Kinase family/mitogenesis & cell differentiation
GRN	Granulin	RF	sALS only; FTLD	Cellular growth
SETX	Senataxin	AR	fALS & jALS	Helicase involved in DNA repair
UNC13A	Unc-13 homolog A (C. elegans)	RF	sALS only	vesicle-mediated neurotransmission
TARDBP	TAR DNA binding protein	AD	fALS; sALS & FTLD	RNA Binding Protein/RNA metabolism
FUS	Fusion (involved in t(12;16) in malignant liposarcoma)	AD & AR	fALS; jALS; sALS & FTLD	RNA Binding Protein/RNA metabolism
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	AD	fALS & sALS	RNA Binding Protein/RNA metabolism
ANG	Angiogenin	AD	fALS; sALS & FTLD	RNA Binding Protein/Angiogenesis
HNRNPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	AD	fALS	RNA Binding Protein/RNA metabolism

Table 1.1: Genetic risk factors for ALS

Mutations in a wide number of genes involve in many cellular pathways have been found to cause ALS as illustrated in this table. Notice particularly the significant number of genes coding for proteins involved in RNA metabolism (shaded in grey). Many of the genes are also associated with FTLD as shown in the table. G.I: Genetic Inheritance Mode; AD: Autosomal dominant, AR: Autosomal Recessive, XLD: X linked Dominant, fALS: familial ALS, sALS: sporadic ALS, jALS: juvenile ALS.

Mutation name	Type	Seq. Original	Seq. Mutated	AA. Original	AA. Mutated	Codon Number	Exon/Intron Number
5' UTR	Polymorphism	C	T	N/A	N/A	0	exon 1
5' UTR	Polymorphism	G	T	N/A	N/A	0	exon 1
p.Leu27	Polymorphism	G	A	N/A	N/A	0	intron 2
c.-12-54G>A	Polymorphism	G	A	N/A	N/A	0	intron 1
p.Ala66	Polymorphism	T	C	N/A	N/A	0	intron 2
p.Ser104	Polymorphism	G	A	N/A	N/A	0	intron 3
p.Lys137	Polymorphism	A	G	N/A	N/A	0	intron 4
c.403-80G>A	Polymorphism	G	A	N/A	N/A	0	intron 3
c.543+112C>A	Polymorphism	C	A	N/A	N/A	0	intron 4
Gly40Gly	Polymorphism	GGG	GGA	Gly	Gly	40	exon 2
Ala66Ala	Substitution	GCT	GCC	Ala	Ala	66	exon 2
Ala90Val	Substitution	GCT	GTT	Ala	Val	90	exon 3
Asp169Gly	Substitution	GAT	GGT	Asp	Gly	169	exon 4
Asn267Ser	Substitution	AAT	AGT	Asn	Ser	267	exon 6
Gly287Ser	Substitution	GGT	AGT	Gly	Ser	287	exon 6
Gly290Ala	Substitution	GGT	GCT	Gly	Ala	290	exon 6
Ser292Asn	Substitution	AGC	AAC	Ser	Asn	292	exon 6
Gly294Ala	Substitution	GGG	GCG	Gly	Ala	294	exon 6
Gly294Val	Substitution	GGG	GTG	Gly	Val	294	exon 6
Gly295Arg	Substitution	GGT	CGT	Gly	Arg	295	exon 6
Gly295Ser	Substitution	GGT	AGT	Gly	Ser	295	exon 6
Gly298Ser	Substitution	GGT	AGT	Gly	Ser	298	exon 6
Met311Val	Substitution	ATG	GTG	Met	Val	311	exon 6
Ala315Thr	Substitution	GCG	ACG	Ala	Thr	315	exon 6
Ala321Gly	Substitution	GCC	GGC	Ala	Gly	321	exon 6
Ala321Val	Substitution	GCC	GTC	Ala	Val	321	exon 6
Gln331Lys	Substitution	CAG	AAG	Gln	Lys	331	exon 6
Ser332Asn	Substitution	AGC	AAC	Ser	Asn	332	exon 6
Gly335Asp	Substitution	GGT	GAT	Gly	Asp	335	exon 6
Met337Val	Substitution	ATG	GTG	Met	Val	337	exon 6
Gln343Arg	Substitution	CAG	CGG	Gln	Arg	343	exon 6
Asn345Lys	Substitution	AAC	AAA	Asn	Lys	345	exon 6
Gly348Cys	Substitution	GGC	TGC	Gly	Cys	348	exon 6
Gly348Val	Substitution	GGC	GTC	Gly	Val	348	exon 6
Asn352Ser	Substitution	AAT	AGT	Asn	Ser	352	exon 6
Arg361Ser	Substitution	AGG	AGT	Arg	Ser	361	exon 6
Pro363Ala	Substitution	CCA	GCA	Pro	Ala	363	exon 6
Ala366Ala	Polymorphism	GCC	GCG	Ala	Ala	366	exon 6
Tyr374STOP	Substitution	TAT	TAA	Tyr	STOP	374	exon 6
Asn378Asp	Substitution	AAT	GAT	Asn	Asp	378	exon 6
Ser379Pro	Substitution	TCT	CCT	Ser	Pro	379	exon 6
Ser379Cys	Substitution	TCT	TGT	Ser	Cys	379	exon 6
Ala382Thr	Substitution	GCA	ACA	Ala	Thr	382	exon 6
Ala382Pro	Substitution	GCA	CCA	Ala	Pro	382	exon 6
Ile383Val	Substitution	ATT	GTT	Ile	Val	383	exon 6
Gly384Arg	Substitution	GGT	CGT	Gly	Arg	384	exon 6
Trp385Gly	Substitution	TGG	GGG	Trp	Gly	385	exon 6
Asn390Asp	Substitution	AAT	GAT	Asn	Asp	390	exon 6
Asn390Ser	Substitution	AAT	AGT	Asn	Ser	390	exon 6
Ser393Leu	Substitution	TCG	TTG	Ser	Leu	393	exon 6

Table 1.2: *TARDBP* mutations which are causative of ALS

A large number of mutations in *TARDBP*, mostly missense mutations, are causative of ALS. Exon 6, which encodes TDP43's Glycine-rich C-terminal is a hotspot for ALS causative mutations (area of table shaded in grey).

In contrast to ALS, FTLD diagnoses carry a positive family history in 30-50% of cases, with an Autosomal Dominant inheritance pattern constituting the most common mode of inheritance (Bird et al. 2003; Goldman et al. 2005). Analogously to ALS, FTLD-causative mutations have been identified in several genes (Sleegers et al. 2010; Halliday et al. 2012), including *GRN*, *VCP*, *CHMP2B*, *MAPT*, thus showing a partial overlap with the genetic risk factors of ALS (table 1.1). Moreover, mutations in *TARDBP* are not only causative of ALS but also FTLD, both in co-morbidity with ALS (Chiò et al. 2010; Borghero et al. 2011) and FTLD alone (Borroni et al. 2009; Synofzik et al. 2014) (figure 1.7). Additionally, ALS and FTLD causing mutations in other genes which encode RNA-binding proteins have been described and include hnRNPA1 (Kim et al. 2013) and FUS (Vance et al. 2009; Kwiatkowski et al. 2009) thus reinforcing the suggestion that cellular dysfunction in pathways related to RNA metabolism is a pivotal process in disease pathophysiology.

The large number of genes involved in the aetiology of ALS and FTLD is perhaps not surprising given the heterogeneity of these diseases. Despite evidence that the type of mutation, even within the same gene, can influence the clinical presentation and to a certain extent survival (Corcia et al. 2012), studies demonstrating a clear genotype to phenotype relationship in ALS and FTLD have not been published.

Whilst the identification of the genetic risk factors of ALS and FTLD has not stopped since the discovery of TDP43 mutations in the aetiology of these diseases, the identification of a disease-causing intronic hexanucleotide repeat in the *C9ORF72* gene (Renton et al. 2011; DeJesus-Hernandez et al. 2011; van der Zee et al. 2013) deserves a special mention. The discovery of disease-causing expansion mutations in *C9ORF72* finally unravelled the genetic mutations underlying the linkage of ALS and FTLD to chromosome 9 (Morita et al. 2006; Vance et al. 2006; Le Ber et al. 2009) and proved to constitute the most common genetic

lesion found in sporadic and familial ALS and FTLD (Majounie et al. 2012; Smith et al. 2013).

Histopathological analysis has revealed that in the majority of cases of ALS, regardless of the underlying genetic mutations, with the exception of patients with *SOD1* mutations (Maekawa et al. 2009) and possibly *FUS* mutations (Bäumer et al. 2010), but including patients with the *C9ORF72* hexanucleotide expansion, the affected neurones have abnormally aggregated, cleaved, polyubiquitinated and hyperphosphorilated TDP43 cytoplasmic inclusions (Neumann et al. 2006; Dickson et al. 2007; Fujita et al. 2008; Pamphlett et al. 2009). Moreover, TDP43 is also a major component of cytoplasmic inclusion bodies in neuronal tissue from FTLD patients, specifically Ubiquitin-rich/Tau-negative inclusions that had been described in FTLD with Ubiquitinated inclusions (FTLD-U) previously to the identification of TDP43 as its major component (Neumann et al. 2006; Davidson et al. 2007; Cairns et al. 2007; Brandmeir et al. 2008; Tan et al. 2013). FTLD-U is now commonly referred to as FTLD-TDP.

Recently, an increasing number of publications have also found that TDP43 histopathological features are correlated with mild cognitive impairment and Alzheimer's disease (AD), the most common form of Dementia in the elderly. A positive dose-dependent correlation has been described between TDP43 cytoplasmic inclusions, mild cognitive impairment and AD (Tremblay et al. 2011). Additionally, patients with AD and TDP43 histopathology have been found to be more severely affected than the group of patients with AD and TDP43-negative disease, with the former being ten times more likely to be markedly cognitively impaired at death (Josephs et al. 2014). In spite of TDP43 immunoreactivity also being found in healthy older adults (Geser et al. 2010), the positive correlation between the severity of cognitive impairment in both pre-clinical stages of dementia and in AD, with up to 25% of patients with AD displaying TDP43 histopathology, suggests that TDP43 dysfunction may also contribute to additional neurodegenerative disease processes and the increase in cognitive symptom

burden. Interest is therefore surfacing about the possible roles of TDP43 in AD and further information on this subject can be found reviewed in (Wilson et al. 2011).

TDP43 is thus at the crux of the aetiology of ALS and FTLD-TDP and of the pathological hallmarks of these diseases. Therefore, TDP43 dysfunction is postulated to play a major role in the pathophysiology of ALS and FTLD-TDP and ultimately lead to neuronal degeneration. Additionally, evidence is emerging that TDP43 dysfunction may contribute to cognitive impairment more generally as well as to the onset and severity of AD. Given the shared histopathological features between ALS and FTLD-TDP, these two diseases are often referred to, collectively, as TDP43 proteinopathies. The degenerative process itself is incompletely understood; nevertheless the pathways that are proposed to lead to degeneration in TDP43 proteinopathies are discussed in the next section.

1.3.1 Neurodegenerative pathways in TDP43 proteinopathies

In the healthy physiological state, TDP43 shuttles between the cytoplasm and the nucleus but is a predominately localised to the nucleus (Ayala et al. 2008), performing a wide range of functions in RNA metabolism and gene expression, hence contributing to the diversity of the transcriptome and proteome (refer to section 1).

Post-mortem histological examination has shown that in ALS and FTLD-TDP, TDP43 is misslocalised to the cytoplasm, where it accumulates and aggregates with prominent post-translational modifications, being cleaved in C-terminal fragments (CTFs), polyubiquitinated and hyperphosphorylated (figure 1.8) (Neumann et al. 2006; Davidson et al. 2007; Cairns et al. 2007; Dickson et al. 2007; Fujita et al. 2008; Mori et al. 2008; Brandmeir et al. 2008; Pamphlett et al. 2009; Tan et al. 2013). The C-terminal fragments which are found in the neuronal inclusion bodies have consistently been reported to have a molecular weight of 25kDa and 35kDa. The former has been observed in CNS tissue from patients with ALS and FTLD-TDP and the latter exclusively in brain tissue of patients with FTLD (Neumann et al.

2006; Igaz et al. 2009). The consistent observation of the abnormal 25kDa and 35kDa CTFs in inclusion bodies, as opposed to full length TDP43 which has a molecular weight of 43kDa, has led to suggestions that these fragments may themselves represent aberrant TDP43 species which are toxic to neurones.

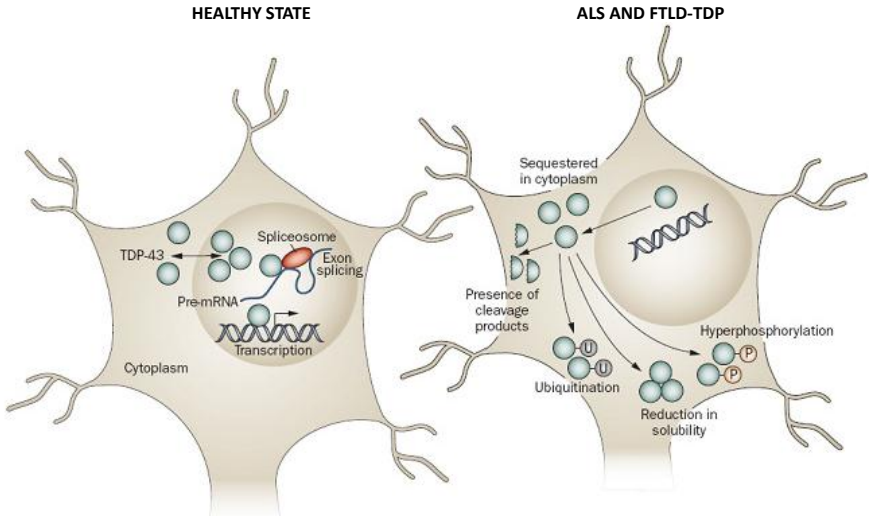

Figure 1.8: TDP43 histopathological changes observed in ALS and FTLD-TDP
In the healthy state TDP43 shuttles between the cytoplasm and the nucleus, but is predominately nuclear. In post-mortem CNS tissue from patients who had suffered from ALS and FTLD-TDP, TDP43 is sequestered in the cytoplasm, cleaved into C-terminal fragments (CTFs), hyperphosphorylated and polyubiquitinated, being aggregated into inclusion bodies. Adapted from (Chen-Plotkin et al. 2010).

The aggregation of TDP43 CTFs, which are subjected to posttranslational polyubiquitination and hyperphosphorytion, has been recapitulated in cells lines (Nonaka et al. 2009; Igaz et al. 2009; Pesiridis et al. 2011), and the aggregation potential has been shown to be enhanced by a mutation associated with ALS (Guo et al. 2011). The cellular experiments have produced evidence that abnormal TDP43 CTF aggregation can be induced by direct overexpression of TDP43 exogenous constructs and proteolytic cleavage of full length TDP43, both endogenous or exogenous in origin, in immortalised cell lines (Nonaka et al. 2009; Igaz et al. 2009; Pesiridis et al. 2011; Che et al. 2011). Furthermore, in some of the experiments, TDP43-CTFs misslocalising to the cytoplasm were able to recruit full length TDP43 (Che et al. 2011) and CTFs, generated *de novo* by proteolytic cleavage (Pesiridis et al. 2011), and translocate it to the cytoplasm.

In addition, the C-terminal of TDP43, under specific experimental conditions, has been shown to self-interact and interact with prion-like domains of other proteins, promoting aggregation (Wang et al. 2012). In some instances, mislocalisation and aggregation of TDP43 led to inhibition of TDP43-dependent splicing as assessed by CFTR minigene splicing assay (Nonaka et al. 2009; Igaz et al. 2009). However these results should be interpreted with warranted scepticism, given that the constructs transfected often included deletions in regions shown to be required for the splicing activity and therefore the "loss of splicing function" observed cannot be unambiguously attributed to CTF aggregation and mislocalisation.

The disruption of additional cellular processes, including microtubule transport and proteosomal function, was necessary in order to achieve the mislocalisation and aggregation of the abnormal TDP43 species in some studies (Nonaka et al. 2009; Pesiridis et al. 2011), which has led to the postulation that TDP43 dysfunction acts as a risk factor in neurodegeneration which is typically triggered by an another cellular stressor event or "second hit", a theory which is commonly referred to as the "two hit hypothesis" (Pesiridis et al. 2011).

However, the abnormal CTF aggregation recapitulated in cell models has not always been associated with ubiquitination and phosphorylation (Che et al. 2011), hence creating ambiguity on whether the posttranslational modifications precede or follow the cleavage and aggregation of TDP43-CTFs, given that at least one other study found that full length phosphorylated TDP43 was found in the insoluble cytoplasmic fraction with CTFs when cells were transfected with TDP43-CTF constructs (Igaz et al. 2009).

The histopathological TDP43 changes observed in FTLD-TDP and ALS have therefore been recapitulated in cellular models through over-expression of C-terminal fragments or induced cleavage of TDP43, leading to mislocalisation and aggregation and, in some cases, partial loss of function of TDP-43 dependent splicing (Nonaka et al. 2009; Igaz et al. 2009; Pesiridis et al.

2011; Che et al. 2011; Guo et al. 2011). Based on the histopathological features of ALS and FTLD-TDP and the results from cellular experiments, it has been hypothesised that TDP43 dysfunction leads to progressive and abnormal accumulation of TDP43 species in cytoplasmic inclusion bodies which are polyubiquitinated and hyperphosphorylated leading to biological dysfunction such as abnormal RNA metabolism and eventually neuronal apoptosis. Notwithstanding, despite the success in recapitulating the mislocalisation of abnormal TDP43 CTFs, the full replication of the posttranslational modifications has not been uniformly achieved and the "loss of splicing function" observed in some studies has to be interpreted cautiously given the caveat that the referred studies used TDP43 constructs with significant deletions in areas shown to be required for TDP43 splicing function. The discrepancies between the experimental findings of (i.e. TDP43 mislocalisation and aggregation, as well as splicing effects, achieved by over-expression in some studies but not others), could have technical explanations ranging from the type of immortal cell line used through to the sequences of the specific constructs and antibodies used, which in itself has biological significance given that in ALS and FTLD-TDP the disease process affects most severely primary neuronal populations in the CNS. Moreover, the key events in ALS and FTLD-TDP, as in any other neurodegenerative disease, are neuronal dysfunction and death, not protein aggregation.

Whilst informative, the experimental results in cellular system fall short of replicating the complexity of a whole organism and are clearly insufficient to shed light on the underlying neurodegenerative process occurring in TDP43 proteinopathies, particularly as, to date, there is no evidence of a causal relationship between protein aggregation and neurodegeneration in most neurodegenerative diseases. Therefore, TDP43 inclusions, despite histopathological hallmarks of ALS and FTLD-TDP may represent epiphenomena rather than pathophysiological events.

Taking into account whole animal complexity, a wide variety of animal models of TDP43 proteinopathies have been characterised in the quest for a better understanding of pathophysiology of neurological diseases associated with TDP43 dysfunction. Advances have been made using *in vivo* murine models that have shaped our understanding and working hypotheses of the underlying neurodegenerative mechanisms in ALS and FTLD-TDP. The existing experimental mouse models will be succinctly described and the mechanisms proposed to underlie neurodegeneration in TDP43 proteinopathies summarised. A critical appraisal of the limitations of the existing models will also be done, revealing the opportunities gained by studying ENU-induced mutant mouse models.

1.4 *In vivo* models of TDP43 proteinopathies: Gain and Loss of function

Cellular models have suggested that TDP43 dysfunction is associated with protein mislocalisation and aggregation. However, cell models fall short of answering the question whether aggregation is necessary and/or sufficient to trigger neuronal death. Additionally, and of particular relevance in the disease pathophysiology and constituting pivotal information necessary for the development of successful therapies, the dissection of the mechanism associated with TDP43 dysfunction that leads directly to neurodegeneration requires supporting data obtained from *in vivo* models.

With the aim of determining whether TDP43 dysfunction leads to neuronal toxicity through a "loss of function", "alterations in alternative exon selection", a "gain of a novel function" or, alternatively, a combination of any of those modalities, a variety of TDP43 proteinopathy *in vivo* models have been produced and characterised. Different organisms have been used to produce disease models, including Drosophila, C. elegans, yeast systems and even a primate model (Uchida et al. 2012), in addition to murine models. The variety of existing models available have been subjected of reviews, most commonly together with other "ALS and

FTLD animal models" with manipulation of other genes involved in these diseases (Joyce et al. 2011; McGoldrick et al. 2013; Liu et al. 2013).

In this thesis, a focus is given to the mouse models already described and its contribution to our understanding of TDP43 toxicity. Moreover, the limitations of the existing mouse models will be discussed thus revealing the need for additional models and providing evidence for the advantages and disadvantages of an ENU mutagenesis approach.

1.4.1 Mouse models of TDP43 proteinopathies

The mouse models of TDP43 proteinopathies described can be divided into "loss of function" and "gain of function" models, with the former including knockout models, conditionally deletion models and knockdown models, and the latter transgenic over-expressors.

Given the aim of determining the mechanism through which TDP43 biological dysfunction leads to neurodegeneration, the existing models will be described under the categories "loss of function" and "gain of function", outlining their relevant contributions to our understanding of TDP43 biology and integrating the knowledge acquired in a working hypothesis, emphasising the particular aspects which require support and answers from further scientific evidence.

1.4.1.1 "Loss of function" mouse models of TDP43 proteinopathies

Early attempts to produce a global TDP43 mouse knockout (KO) proved unsuccessful and embryos were shown to die early during development. The non-viability of *Tardbp* null embryos, allied to the observations that *Tardbp* is highly expressed during development and that *Tardbp*$^{-/-}$ blastocysts fail to expand the inner cell mass efficiently in cell culture has led to the conclusion that TDP43 fulfils essential roles in early embryonic development, including pluripotent stem cell proliferation (Sephton et al. 2010; Wu et al. 2010; Chiang et al. 2010). Interestingly, heterozygous animals, *Tardbp*$^{+/-}$ have been reported to be phenotypically

normal, fertile and without motor deficits up to the age of 14 months (Wu et al. 2010). Furthermore, TDP43 protein levels, determined by western blot (Wu et al. 2010), were equivalent in heterozygous animals and wild type controls, despite the former only having one copy of the *Tardbp* allele, as were *Tardbp* mRNA levels assessed by real time PCR (Sephton et al. 2010).

Conditional deletion of TDP43 has proved informative and has provided confirmation for roles of TDP43 in gene expression, metabolism and trophic support making it an essential factor for the wellbeing of the whole organism. Global deletion of *Tardbp* performed in animals between 4 and 6 weeks of age, thus bypassing the embryonic lethality, leads to dramatic weight loss after conditional deletion, specifically a significant reduction in fat mass, not attributable to reduced food intake but associated to transcriptional repression of the gene *Tbc1d1* in skeletal muscle (Chiang et al. 2010). Strikingly, the metabolic consequences of conditional *Tardbp*-KO were so severe that the animals died within 9 days when the Cre-recombinase was driven by the ubiquitous Rosa26 promoter (which is very potent in most cell populations but not in the CNS) under tamoxifen induction and within 18 days if the recombination event was driven by the weaker hybrid CAG promoter (Chiang et al. 2010); hence demonstrating that global loss of TDP43 function (i.e. at the whole organism level) is incompatible with life both in the prenatal and postnatal periods.

Deletion of *Tardbp* in spinal cord motor neurones has been reported to lead to the development of hindlimb clasp, exaggerated thoracic kyphosis (which can be associated with weakening of the spinal and paraspinal muscles), progressive rotarod deficits and weight loss; ALS-related phenotypes which were male dominant and were associated with premature death with an average survival of 10 months (Wu et al. 2012). Moreover, the ALS-like phenotypes were accompanied by histological features of motor neurone death, specifically reactive astrogliosis and an increase in polyubiquitinated cytoplasmic proteins present in the

insoluble urea fraction of extracts from the spinal cords of the experimental animals (Wu et al. 2012).

TDP43 depletion from motor neurones has thus been demonstrated to lead to neuronal toxicity. It is relevant however to interpret the results with caution as two caveats remain. Firstly, as acknowledged by Wu and colleagues, expression of the promoter used to drive the recombination event (HB9) could not be excluded in additional cell types to spinal motor neurones (Wu et al. 2012). Secondly, the observation of increased levels of polyubiquitinated proteins (not immunoreactive to TDP43) in the cytoplasmic fraction of spinal cord extracts from experimental animals, whilst not necessarily incompatible with the pathophysiology of ALS, does not have an apparent correlation with its histopathology. The histopathological hallmark, as described in previous sections, is the aggregation of TDP43 CTFs which are hyperphosphorylated and polyubiquitinated and not vast amounts of other proteins and therefore the experimental results described by Wu and colleagues cannot be said to recapitulate the histopathology of ALS or FTLD-TDP.

In vivo TDP43 knockdown has been achieved through the generation of transgenic lines with inducible artificial microRNA complementary to TDP43 mRNA, which binds the latter and knocks down its levels though RNA interference (Yang et al. 2014). Technical difficulties precluded the authors from achieving CNS-selective TDP43 knockdown, which was instead achieved at the whole organism level using a germ line Cytomegalovirus promoter. Mice that were subjected to TDP43 knockdown developed, from 4 weeks onwards, progressive neurological phenotypes, transitioning from early hyperactivity, to muscle weakness and finally paralysis, with confirmed positive sharp waves on electromyography of the limbs (Yang et al. 2014). Weight also declined from day 50 onwards until premature death occurred, for all animals including males and females, between 43.9 and 100.2 days.

TDP43 “loss of function” mouse models have clearly shown that TDP43 depletion is highly detrimental to life. Furthermore, TDP43 is essential in early embryogenesis and complete knockout prohibits viable embryonic development and leads to early lethality. Even in the postnatal period, global deletion of TDP43 results in profound metabolic changes and loss of fat mass and progression to rapid death in mice (Sephton et al. 2010; Wu et al. 2010; Chiang et al. 2010). More restricted TDP43 depletion, specifically aimed at targeting spinal cord motor neurones, and a knockdown approach which modestly reduces TDP43 protein levels (by approximately 10%), leads to motor neurone degeneration and recapitulates, at least partly, pathological features of ALS (Wu et al. 2012; Yang et al. 2014).

The existing “gain of function” mouse models will now be reviewed.

1.4.1.2 “Gain of function” mouse models of TDP43 proteinopathies

In this section, mouse models generated using transgenic technology will be discussed, including animals engineered to express the wild type *TARDBP* gene or the *TARDBP* gene with point mutations associated with ALS. Transgenic animals have been engineered using a variety of different promoters which lead to overexpression of the exogenous gene and thus, regardless of any additional effects of the transgene, *TARDBP* “gain of function” is necessarily a consequence of the promoter technology used and is present, at different levels, in these mouse models.

Human TDP43 transgenic animals have been described and shown to develop motor dysfunction and learning and memory deficits. Human wild type TDP43 (hTDP43-WT) has been expressed under regulation of the mouse prion promoter, leading to increased TDP43 protein levels, axonal degeneration and reactive gliosis, effects which were critically dose dependent (Xu et al. 2010). Additionally, the histopathological hallmarks of ALS and FTLD-TDP were partially recapitulated with cytoplasmic phosphorylated TDP43 species being detected in the experimental animals with higher protein levels, specifically being present in

homozygous transgenics animals but not hemizygous animals (Xu et al. 2010). Corroboration that over-expression of human wild type TDP43 induces neurotoxicity has been provided by additional studies including the report from Wils and colleagues that hTDP43-WT transgenic mice develop dose-dependent loss of cortical neurones and spinal motor neurones, and that the accumulation of the abnormal 25kDa CTF in the nuclear fraction of brain tissue was correlated with disease progression (Wils et al. 2010).

Mice with mutant human TDP43 transgenes have also been studied. Expression of transgenes with mutations associated with ALS, specifically the A315T mutation, develop motor neurone loss with increased ubiquitin immunoreactivity in the spinal cord but no TDP43 inclusions (Wegorzewska et al. 2009).

Transgenic animals with mutations in human TDP43, specifically the NLS, have also been studied using the hTDP43-WT as controls. Remarkably, but perhaps not unexpectedly, given the reported dose-dependent neurotoxicity of hTDP43-WT in transgenic mice, both the animals with the mutant NLS transgene and wild type transgene developed degeneration of forebrain neuronal populations and corticospinal tracts, which was more severe in the mice with the mutant NLS transgene (Igaz et al. 2011).

Expression of both human wild type and ALS-associated mutant TDP43 has been shown to lead to neurotoxicity. Furthermore, overexpression of mouse wild type TDP43 in the forebrain of mice also leads to enhanced apoptosis and impaired learning and memory identified on behavioural testing, at least in some models (Tsai et al. 2010).

Transgenic *TARDBP* animals engineered using the Bacterial Artificial Chromosome (BAC transgenics) and with similar levels of transgene expression have also been characterised. Mice carrying the wild type *TARDBP* transgene, $TARDBP^{A315T}$ or $TARDBP^{G348C}$ have all been shown to develop motor deficits and deficits in learning and memory which are accompanied

by astrogliosis and histopathological changes, with the mutant BAC transgenics being more severely affected (Swarup et al. 2011).

The mechanistic dissection of how TDP43 dysfunction leads to neurodegeneration has become particularly challenging given that conditional deletion or knockdown of TDP43 leads to neurological dysfunction and in some instances neuronal death (Wu et al. 2012; Yang et al. 2014), as does over-expression of wild type and disease-associated mutant *TARDBP* alleles in transgenic mouse models (Wegorzewska et al. 2009; Wils et al. 2010; Tsai et al. 2010; Igaz et al. 2011). Additionally, the *in vivo* data is inconsistent in recapitulating the histopathology of ALS and FTLD-TDP; abnormal CTFs have been reported in the nucleus and correlate with phenotype severity (Wils et al. 2010) and an increase ubiquitinated proteins, not recognised by anti-TDP43 antibodies, have been reported in the cytoplasm upon conditional deletion of TDP43 (Wu et al. 2012). Furthermore, cytoplasmic TDP43 inclusions with cleaved and ubiquitinated CTFs have been reported upon over-expression of a human TDP43 transgene with a defective NLS but not on expression of the wild type human transgene (Igaz et al. 2011).

A recent publication reported the effects of expression of transgenes with wild type human TDP43 alleles and alleles with pathogenic mutations, pulling together animals with similar levels of transgenic protein, and reported that only mice carrying the transgene with pathogenic mutations, specifically the Q331K and M337V, developed neuromuscular deficits, severe in the former and mild on the latter (Arnold et al. 2013). Furthermore, neurodegeneration was not associated with TDP43 mislocalisation, aggregation or abnormal insoluble CTFs, and endogenous gene splicing assays suggested that, overall, the pathogenic mutations and degeneration were associated with a gain of function, at least at the splicing level (Arnold et al. 2013).Whilst informative, the results from Arnold and colleagues need to be interpreted together with the additional data provided by other mouse models of TDP43

proteinopathies, whose characteristics are summarised in table 1.3 (Wu et al. 2012; Wegorzewska et al. 2009; Guo et al. 2012; Xu et al. 2010; Xu et al. 2011; Stallings et al. 2010; Wils et al. 2010; Caccamo et al. 2012; Tsai et al. 2010; Igaz et al. 2011; Cannon et al. 2012; Swarup et al. 2011; Yang et al. 2014; Arnold et al. 2013; McGoldrick et al. 2013).

The inescapable conclusions from the existing TDP43 proteinopathy animal models are necessarily that both "loss of function" and "gain of function" are associated with neurodegeneration in the mouse, remaining undetermined whether the gain of function that leads to neurodegeneration constitutes "augmented alternative exon selection" or the "gain of a novel toxic function". Moreover, the data is suggestive that TDP43 mislocalisation and aggregation is not necessary for "TDP43 dysfunction"-induced neurodegeneration, at least in some over-expression models (Arnold et al. 2013).

The models described, including "gain of function" or "loss of function" models, share in common the alteration of TDP43 protein levels as a consequence of the technology used. TDP43 levels are, however, tightly autoregulated by TDP43 itself when wild type and mutant alleles are expressed isogenically (Ayala et al. 2011; Avendaño-Vázquez et al. 2012; Bembich et al. 2013; Ricketts et al. 2014**).** Heterozygous animals for *Tardbp* deletions have protein levels equivalent to wild type controls and are phenotypically normal (Sephton et al. 2010; Chiang et al. 2010). In addition, downregulation of mouse endogenous TDP43 has been reported in transgenic over-expressors (Xu et al. 2010; Igaz et al. 2011). Analysing the previous two observations together with the observation that changes in TDP43 protein levels lead to degeneration in mouse models, it could be concluded that the autoregulation process is highly relevant for health and disease. Thus disruption of TDP43 autoregulation, allowing an increase or decrease in TDP43 levels to be uncorrected, could potentially lead to neurodegeneration.

TDP43 autoregulation will be succinctly described, contextualising the limitations of relying on over-expression or conditional deletion models and underlining the advantages of using ENU-derived *Tardbp* mouse models.

Promoter	Protein	Protein (fold change)	Symptom onset (weeks)	Survival (weeks)	Abnormal motor behaviour	LMN	UMN	Cortex	Axonal	Gliosis	Cytopalmic TDP43	CFTs	References
Hb9:Cre	MN specific depletion*	0	13	40	Y	60% loss	nd	nd	nd	Y	nd	nd	Wu et al. 2012
mPrp	hTDP43 A315T	3	13	22	Y	20% loss	Y	nd	Y	Y	Y	Y	Wegorzewska et al. 2009; Guo et al. 2012
mPrp	hTDP43 WT	1.9	Not affected										Xu et al. 2010
mPrp	hTDP43 WT	2.5	2	4–8	Y	N	nd	nd	Y	Y	Y	Y	Xu et al. 2010
mPrp	TDP43 M337V	1.9	Not affected										Xu et al. 2011
mPrp	TDP43 M337V	2.5	2	4	Y	nd	nd	nd	Y	Y	Y	Y	Xu et al. 2011
mPrp	TDP43 WT	3–4	nd	nd	N	nd	nd	nd	nd	Y	nd	nd	Stallings et al. 2010
mPrp	TDP43 A315T	4	4	37.5	Y	nd	nd	nd	nd	Y	Y	Y	Stallings et al. 2010
Thy1.2	TDP43 WT	Males: 3.6	2-2.5	nd	Y	N	nd	nd	Y	Y	N	N	Shan et al. 2010
Thy1.2	TDP43 WT	Females: 1.3	13	nd	Y	N	nd	nd	nd	nd	N	N	Shan et al. 2010
Thy1.2	TDP43 WT	1.9	56	nd	Y	nd	nd	nd	nd	nd	nd	nd	Wils et al. 2010
Thy1.2	hTDP43 WT	3.8	8	Approx. 27	Y	10% loss	15% loss	nd	nd	Y	Y	nd	Wils et al. 2010
Thy1.2	hTDP43 WT	5.1	2	4	Y	25% loss	30% loss	nd	nd	Y	Y	Y	Wils et al. 2010
Thy1.2	TDP25	4.7	nd	nd	nd	nd	nd	N	nd	nd	Y	Y	Caccamo et al. 2012
CaMKII	TDP43 WT	2	8	71	Y	nd	nd	Y	nd	Y	Y	Y	Tsai et al. 2010
CaMKII (TRE)	hTDP43 WT	0.8 (induced at 4 weeks)	8–49	nd	Y	nd	Y	Y	nd	Y	nd	nd	Igaz et al. 2011
CaMKII (TRE)	hTDP43 ΔNLS	7.9 (induced at 4 weeks)	5	26	Y	nd	Y	Y	Y	Y	Y	nd	Igaz et al. 2011
CaMKIIα (TRE)	TDP43 WT	3 (induced from P21)	nd	nd	nd	nd	nd	Y	nd	Y	Y	nd	Cannon et al. 2012
TDP43 (BAC)	TDP43 WT	3	28–32	nd	Y	nd	nd	nd	N	Y	nd	nd	Swarup et al. 2011
TDP43 (BAC)	TDP43 A315T	3	28–32	nd	Y	nd	nd	nd	N	Y	Y	Y	Swarup et al. 2011
TDP43 (BAC)	TDP43 G348C	3	28–32	nd	nd	nd	nd	nd	N	Y	Y	Y	Swarup et al. 2011
CMV (germline)	TDP43 amiRNA (knd)	0.9	4	11	Y	60% loss	nd	Y	nd	Y	nd	nd	Yang et al. 2014
mPrp	hTDP43 WT	1.5-2	Not affected										Arnold et al. 2014
mPrp	hTDP43 M337V	2.5-3	70	nd	Y	N	N	nd	Y		N	N	Arnold et al. 2014
mPrp	hTDP43 Q331K	2-3	21	nd	Y	35% loss	N	nd	Y	Y	N	N	Arnold et al. 2014

Table 1.3: Mouse models of TDP43 proteinopathies

Table summarising the phenotypes of mouse models of TDP43 proteinopathies described in the literature. Parameters given in the table include neuronal degeneration (LMN, UMN, cortex) Axonal loss and histopathological changes (gliosis, TDP43 mislocalisation and CTF formation). The phenotypes are variable, however the key point of this table is that wild type or mutant TDP43 over-expression as well as conditional deletion can lead to neurodegeneration and associated phenotypes in mouse models. Y: yes, N: no, nd: not done. *motor neurone specific depletion not strictly proven.

1.5 TDP43 autoregulation: implications for mouse models of TDP43 proteinopathies

The *in vivo* evidence that transgenic *Tardbp*$^{+/-}$ mice have TDP43 levels comparable to wild type controls and the fact that other RNA binding proteins autoregulate its own expression (Sureau et al. 2001; Wollerton et al. 2004) led to further investigation into hypothesis that TDP43 also autoregulates its own protein levels. Additionally, as described in the previous section, ectopic expression of TDP43 in transgenic mice can lead to suppression of mouse endogenous TDP43 (Xu et al. 2010; Igaz et al. 2011).

Exogenous TDP43 constructs have been inserted under stable conditions into HEK293 cells and induced by tetracycline to express a single copy, hence leading to stable expression of the exogenous TDP43, which downregulated endogenous TDP43 at the transcript and protein level (Ayala et al. 2011). Moreover, downregulation of endogenous TDP43 was abolished in construct with a mutagenized RRM1 or partial deletion of the C-terminal from TDP43, specifically the region encoding for amino acid residues 321 to 366, but not in constructs with a mutagenized RRM2, thus demonstrating the requirement of the RRM1 and C-terminal regions, but not the RRM2, for the autoregulation process (Ayala et al. 2011).

The evidence from two separate research groups demonstrate that TDP43 binds to the 3' UTR of its own transcript, confirming it is an interaction which is pivotal for the autoregulation process (Ayala et al. 2011; Polymenidou et al. 2011). However, the exact mechanism downstream from TDP43 binding to its 3'UTR in repressing its own expression are not consensual and two models have been proposed.

The group from Emanuelle Buratti and Francisco Baralle, based on their experimental evidence, suggest that when present in high levels, TDP43 binds to its transcript's 3' UTR, at the TDPBR, leading to polymerase II stalling and enhanced *TARDBP* intron 7 splicing (a

cryptic intron located in the 3'URT, downstream from the last coding exon of TDP43), thus physically removing its canonical polyadenylation site Poly(A)1. Consequently, polyadenylation is enhanced at the cryptic Poly(A)2 site, disrupting cytoplasmic transport and thus promoting nuclear retention of the TDP43 pre-mRNA, which cannot therefore be translated and is eventually degraded. Additionally, when TDP43 is bound to the TDPBR, some transcripts are terminated prematurely given that RNA polymerase II stalls, and hence it is suggested that abnormally short transcript may undergoes nonsense mediated decay, further contributing to the inhibition of TDP43 protein levels (Ayala et al. 2011; Bembich et al. 2013) (Figure 1.9).

The Cleveland group has suggested that the key event in TDP43 autoregulation is alternative splicing of the intron 7 in the 3'UTR by TDP43 when it binds the TDPBR of its own transcript, which deposits exon-junction complexes 3' of the stop codon thus marking it for nonsense medicated decay (Polymenidou et al. 2011).

Despite the underlying biological mechanisms involved in TDP43 autoregulation lacking consensus at present, the evidence both *in vivo* and *ex vivo* suggests that this process keeps TDP43 proteins levels within a narrow physiological range at whole organism level. Moreover, the C-terminal domain of TDP43, where disease causing mutations cluster, is essential for the autoregulation of the protein's levels (Ayala et al. 2011).

The mouse models of TDP43 proteinopathies (table 1.3) have as a common characteristic changes in TDP43 levels which result from ectopic expression of a transgene or conditional deletion. Consequently, even when constructs which carry pathogenic mutations are used, bias is being introduced by artificially bypassing TDP43 autoregulation as a direct result of the technology used rather than an effect which could be attributed to the mutations.

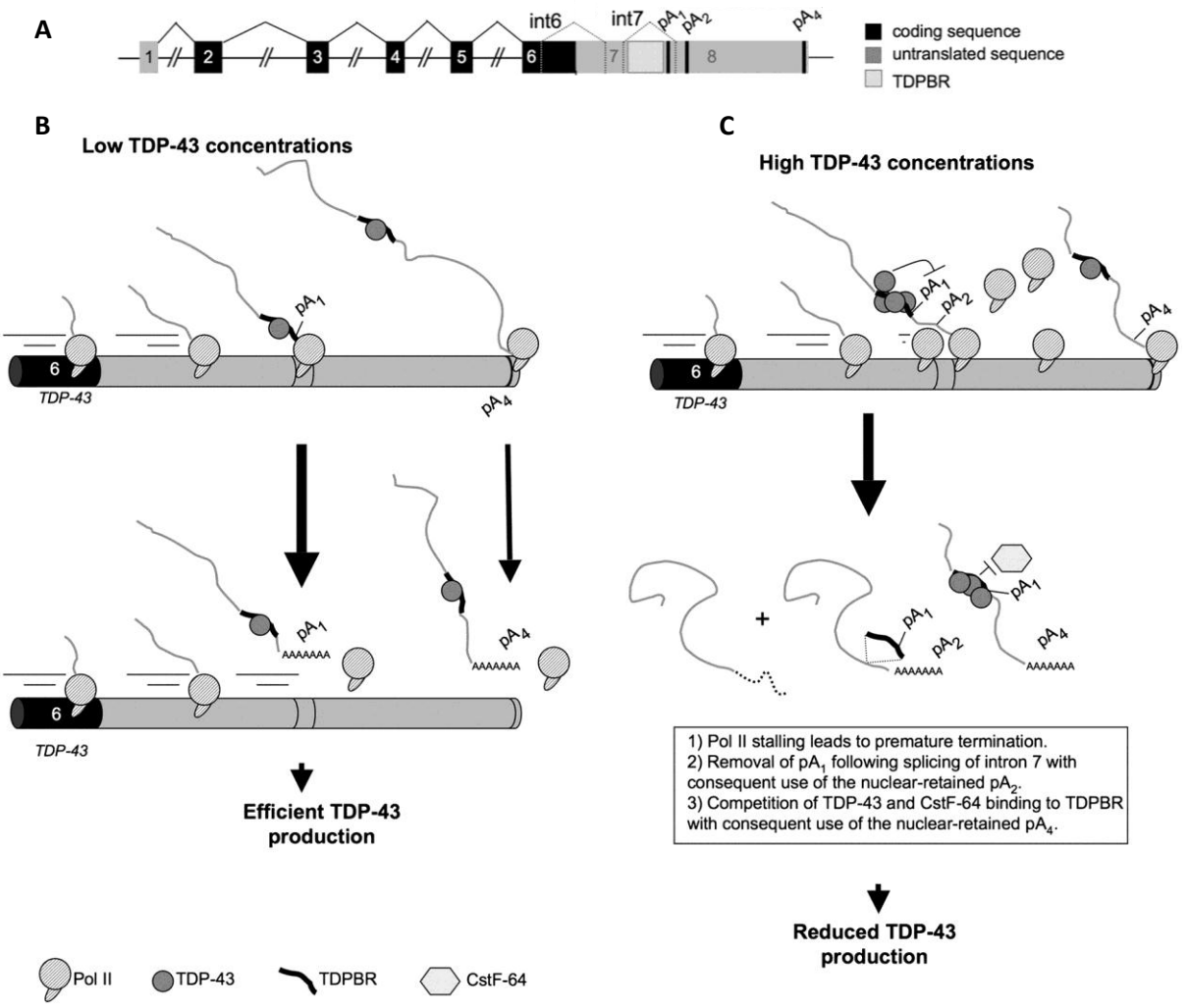

Figure 1.9: Models of TDP43 autoregulation

Two models have been proposed for the autoregulation of TDP43 protein levels. Both involve biding of TDP43 to its own (A) transcript at the 3' UTR in the TDPBR (TDP Binding Region); an interaction that has been shown to occur *in vivo* (Polymenidou et al. 2011).

After binding of TDP43 to the TDPBR, the model form the Cleveland research group proposes that the cryptic intron 7 at the 3'UTR of the *Tardbp* transcript is spliced, depositing exon-junction complexes 3' of the stop codon which marks the mRNA for nonsense medicated decay (note that this process is not diagrammatically represented in this figure). The model from the research group from Buratti and Baralle proposes that when (B) TDP43 concentrations are low the TDPBR is free and polyadenylation occurs mainly at the canonical Poly(A)1 site and the alternative Poly(A)4 site, with transcripts polyadenylated at either of those sites being efficiently translated. C) When TDP43 concentrations are high, TDP43 binds the TDPBR which leads to both stalling of RNA polymerase II (poll II) and splicing of the cryptic intron 7 and hence removal of the canonical Poly(A)1 site from the *TARDBP*/*Tardbp* transcript. It is suggested that polyadenylation then occurs at the alternative Poly(A)2 site which disables transport of the alternatively polyadenylated transcripts to the cytoplasm and therefore a reduction in TDP43 translation. Additionally, the model also suggests that some transcripts terminate early due to pol II stalling and are degraded via nonsense mediated decay. Adapted from (Avendaño-Vázquez et al. 2012).

1.6 Overall limitations of the mouse models of TDP43 proteinopathies

The implications for TDP43 autoregulation for the existing mouse models of TDP43 proteinopathies have been described in the preceding section. Additional limitations are apparent, specifically the onset of expression of the transgene, or conditional deletion, may influence the phenotype in such way where it becomes too difficult to translate the results into meaningful processes inherent to the TDP43 dysfunction associated with ALS and FTLD-TDP. In the cases in which TDP43 mutations are causative of these diseases, the mutant allele is present from conception, through early embryonic and late foetal development, postnatal growth and until the onset of disease occurs, generally after the fifth decade of life. The effects of mutations are therefore occurring in the majority of cell types of the organisms, given TDP43's ubiquitous expression, throughout its life, eventually leading to disease.

The limitations of the existing models are thus not restricted to the bias introduced in the autoregulation process or the temporal expression of the transgene but also include the differential expression of the transgene in the different cells of the mice, including cells of the CNS. Tissue-specific effects are particularly relevant not only for the obvious reason that neuronal degeneration is associated with the dysfunction of ubiquitously expressed TDP43, but also because the promoters used in the transgenic models lead to variable levels of expression and often lack the specificity anticipated.

To illustrate the last point, Wu and colleagues could not confirm that their HB9 promoter driving the Cre-recombinase enzyme specifically deleted TDP43 in motor neurones admitting it could also have done so in additional cells in the spinal cord (Wu et al. 2012), which has particular relevance given that glial cell dysfunction is associated with the pathophysiology of ALS (Philips & Rothstein 2014).

Further understating of TDP43's biological function and dysfunction caused by point mutations in different protein domains can be gained by characterising ENU *Tardbp* mouse mutants. The ENU mutagenesis process will be succinctly described, highlighting the advantages of ENU mouse models of TDP43 proteinopathies over the existing models and its limitations.

1.7 ENU mutagenesis: Creating models for a better understanding of TDP43 biology and dysfunction

ENU is an alkylating agent which acts as a mutagen by transferring an ethyl group to nucleophilic nitrogen or oxygen species present in DNA purine and pyrimidine bases (Justice et al. 1999; Kondo et al. 2010), causing it to be misread during DNA replication and resulting in inheritable mutations when DNA repair mechanism fail to correct the error (Noveroske et al. 2000; Bielas & Heddle 2000; O'Neill 2000; Shrivastav et al. 2010).

Despite being a random process, the nature of the alkylation reaction and DNA replication leads to point mutations being the most frequently induced mutations in ENU mutagenesis, with a minority of cases resulting in small deletions (Justice et al. 1999). A-T base pairs are the most frequent site where mutations occur, with A-T to T-A transversions or A-T to G-C transitions representing up to 85% for all mutagenic events. Additionally, of all induced mutations it is estimated that approximately 70% result in non-synonymous amino acid changes in the translated protein, two thirds of which are missense mutations and the remaining one third nonsense or splicing mutations (Justice et al. 1999; Takahasi et al. 2007).

Protocols have been designed and different dosing regimens of ENU indicated for optimal mutagenesis in different laboratory mouse strains calculated and available in the literature (Justice et al. 2000). Typically, ENU is injected in male mice intraperitoneally, leading to high frequency mutations in the premeiotic spermatogonia cells (in addition to somatic cells).

After a period of infertility, the injected males can be mated with wild type females, generating offspring with a variety of mutations in the parentally inherited alleles (Acevedo-Arozena et al. 2008) (figure 1.10).

The first generation offspring (G1) from a wild type female and an ENU-injected male will carry, on average, 30-50 potentially functional mutations and segregation can be achieved by crossing the G1 animals to wild type mice, reducing the mutations burden by half in the offspring of every backcross, provided there is no linkage disequilibrium (Keays et al. 2006). ENU mutagenesis thus constitute a versatile methodology that can be employed in a variety of genetic screens, including dominant, recessive and modifier screens, reviewed in detail in (Acevedo-Arozena et al. 2008). This mutagenesis methodology can therefore shed further light into gene function and genetic pathways, potentially unravelling new targets for pharmacotherapeutic agents in different diseases.

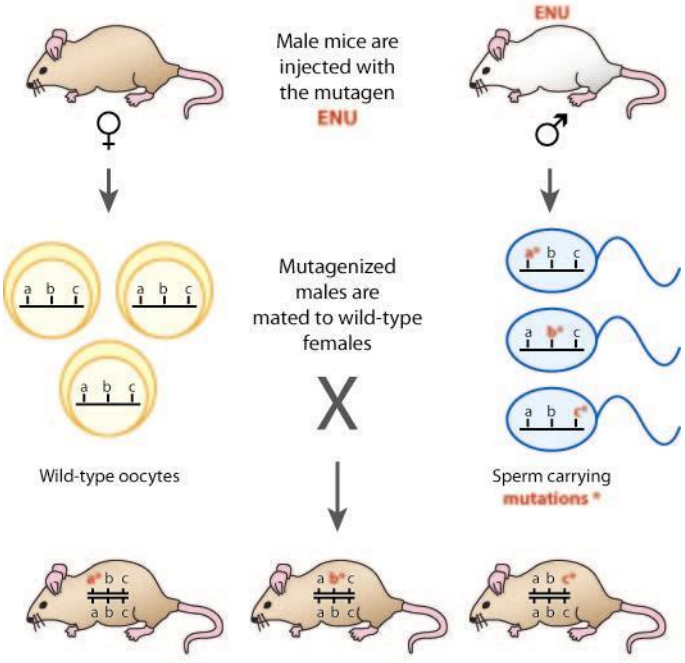

Figure 1.10: Generation of mouse mutants through ENU mutagenesis
Male mice are typically injected with ENU intraperitoneally and, following a period of infertility, when mated with wild type females will transmit mutations to the offspring. Each G1 animal has on average 30-50 potentially functional mutations. Adapted from (Acevedo-Arozena et al. 2008).

In the context of the aetiology and pathophysiology of ALS and FTLD-TDP, ENU mutagenesis offers the unique opportunity of generating TDP43 mice with missense mutations in TDP43, with the respective *Tardbp* gene in its native locus and therefore physiological levels of expression can be expected even in mutant alleles. Provided that the necessary steps are taken to ensure the segregation of any additional mutations, changes in protein levels and TDP43 dysfunction can then be wholly attributable to the mutations.

ENU mutagenesis thus provides a valuable tool to further characterise TDP43's biological function and study the effects of missense mutations in the protein's known functions, including the autoregulation process and RNA metabolism. Moreover, the missense mutation will be present throughout the organism's life, hence enabling the longitudinal study of the animals. Therefore, ENU-induced mutants can be extremely informative regarding the intrinsic relationship between TDP43 mutations, biological dysfunction, neuronal dysfunction and, ultimately neurodegeneration; a dynamic process that can be postulated to ultimately lead to the onset of neurodegeneration and clinical disease. Limitations also exist and these include the randomness of the mutagenic process (and therefore specific mutations cannot be engineered), the extensive breeding of G1 animals required to reduce the mutation burden of the mice and the risk of linkage disequilibrium which can confound the genotype to phenotype relationship.

A reverse genetics approach was taken by the neurodegeneration group at the MRC Mammalian Genetics Unit to identify mutations in the mouse *Tardbp* gene and pursue investigations on TDP43's biological function and dysfunction.

1.7.1 ***Archive Screening and ENU Tardbp mouse mutants: A reverse genetics approach***

The DNA archives of G1 males resulting from crosses between ENU injected males and wild type female mice from the MRC Mammalian Genetics Unit and the Riken Institute in Japan were screen for mutations in all *Tardbp* coding exons. Methodologically, the screening process consisted of PCR amplification of exonic regions followed by high resolution DNA melting analysis. High resolution DNA melting analysis has been shown to be an efficient way to screen for mutations (Ishikawa et al. 2010) and identifies changes in the melting temperature of the complementary DNA strands, thus detecting changes in the melting temperature of the complementary PCR-amplified strands when mismatches exist, even if they consist of a single base pair. The technique is thus suited to screen G1 animals who will be heterozygous given that they can inherit a mutant *Tardbp* allele from the mutagenized male parent but will inherit a wild type allele form the female parent.

The archive screening was performed before I joined the research group by members of the MRC Mammalian Genetics Unit and collaborators at the Riken Institute. Overall, screening identified ten coding mutations in *Tardbp*, nine missense mutations and one nonsense mutation, with selected mouse lines being re-derived by *in vitro* fertilisation using sperm matched to the archived DNA. The reverse genetics approach taken is diagrammatically displayed in figure 1.11 and the position of the mutations identified in the TDP43 protein displayed in figure 1.12.

The first ENU *Tardbp* mutants were thus identified and their characterisation started to contribute towards our understanding of the biological consequences of mutations in TDP43.

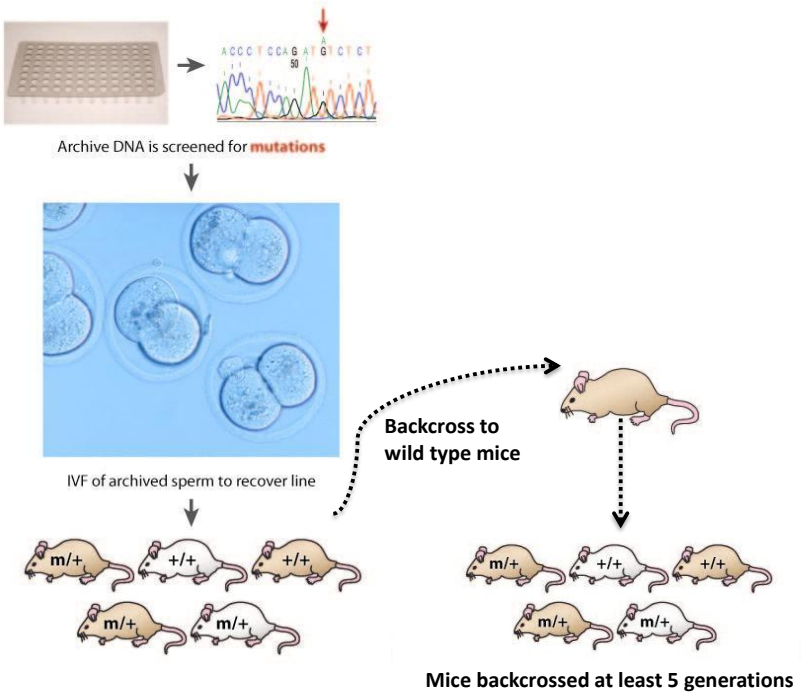

Figure 1.11: Reverse genetics approached used to identify *Tardbp* ENU mutants
The DNA archive from G1 animals from the MRC Mammalian Genetics Unit and the Riken Research institute in Japan was screened for mutations in all TDP43 coding exons. Once mutations were identified, matching sperm archive was used and the mice re-derived by in-vitro fertilisation. Additional mutations were segregated from the mutations of interest in *Tardbp* by backcrossing the mice to wild type animals. Testing cohorts and mice and experimental samples were only taken after the mice had been backcrossed at least 5 generations to ensure appropriate segregation of any associated mutations. Adapted from (Acevedo-Arozena et al. 2008).

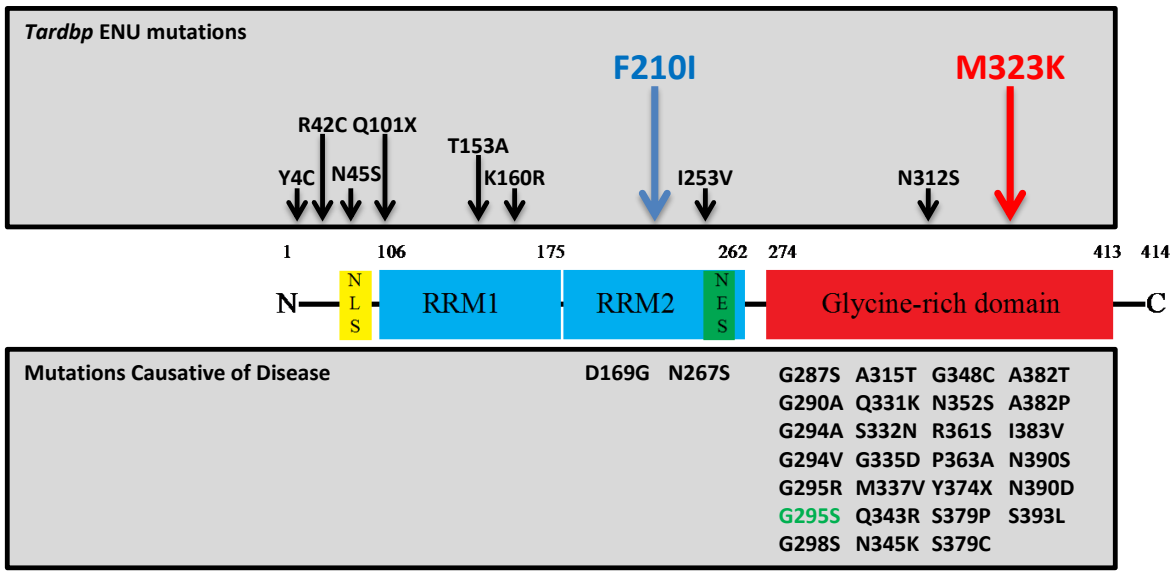

Figure 1.12: *Tardbp* ENU mutations in the context of TDP43 and disease causative mutations
Tardbp ENU mutations were identified throughout the TDP43 protein. The Q101X mutation constitutes an early stop codon which replaces the Glutamine amino acid at residue 101 of the protein chain. *Tardbp*^{Q101X} mutants have been characterised and the results published (Ricketts et al. 2014). The F210I mutation constitutes a single amino acid change at residue 210 in the RRM2 from Phenylalanine to Isoleucine. The M323K mutation constitutes a single amino acid change form Methionine to Lysine at residue 323 of the amino acid change, located in the C terminal region where most disease causative mutations cluster.

1.7.2 ENU Tardbp mutants: advancing our understanding of TDP43 biology

Thomas Ricketts characterised the initially ENU mutants that were re-derived after the archive screening as part of his PhD thesis. Molecular and behavioural characterisation of the *Tardbp*^{Q101X} mutant has been completed and published. This mutation causes an early termination of TDP43, before the RRM1, and the resulting mutant transcript is not translated (Ricketts et al. 2014). In the homozygous state, the Q101X mutation therefore represents a full loss of TDP43 function.

Tardbp$^{Q101X/Q101X}$ animals are not viable, dying at an early stage during embryonic development with embryos never having been detected even as early as 6.5 days post conception (dpc) (Ricketts et al. 2014), in line with previous studies suggesting that complete TDP43 loss of function precluded early embryonic development and therefore is not compatible with life (Sephton et al. 2010; Wu et al. 2010; Chiang et al. 2010). Interestingly, the protein levels in *Tardbp*$^{Q101X/+}$ were found to be identical to wild type animals, whilst the RNA from the mutant allele was markedly reduced, strengthening the concept that TDP43 autoregulation is a sophisticated and highly effective process which is incompletely understood (Ricketts et al. 2014). However, despite normal protein levels *Tardbp*$^{Q101X/+}$ showed changes in alternative splicing of two target genes in the direction of loss of function (*Sort1* and *Pdp1*). *In vivo* phenotyping revealed that the heterozygous animals had softer abdominal tone and females exhibited abnormal limb clasping; neurodegeneration or neuromuscular dysfunction (tested by electromyography) were not found and neurite outgrowth assessed in primary neuronal cultures was normal (Ricketts et al. 2014).

The first published ENU *Tardbp* mutant has thus proved to be informative revealing intricacies in the TDP43 biology which had not been identified in transgenic animals. Remarkably, specific splicing deficits were identified in *Tardbp*$^{Q101X/+}$ with normal levels of wild type TDP43 protein, raising the possibility that the mutant RNA is in some way

interfering with the protein's function or that the commonly used western blots are not sensitive enough to detect physiologically relevant changes in TDP43.

In his thesis "Assessment of an allelic series of mouse TDP43 mutations", deposited at University College London, Thomas Ricketts described the characterisation of additional ENU *Tardbp* mutants[4]. Of particular relevance to the experimental work described on this thesis was the initial molecular characterisation of the F210I mutation (figure 1.12) which is summarised below.

Affecting the RRM2, the F210I amino acid substitution decreases the affinity of mutant TDP43-F210I for RNA considerably as assessed by Electrophoretic Mobility Assay (EMSA). Whilst the homozygous animals die before birth on the C57BL/6J background, foetuses were found to be alive up to 15.5dpc, enabling the assessment of protein levels in embryonic tissue and the culture of primary cells. Across all genotypes, protein levels remained identical, as did TDP43 subcellular localisation. However, a clear dose-dependent effect was found between the mutant allele and loss of function in TDP43-dependent splicing of target genes. Strikingly, despite the decreased RNA binding affinity of the F210I mutant allele, overall no significant difference in the proteins levels were observed in embryonic heads or cultured mouse embryonic fibroblasts across all there genotypes ($Tardbp^{+/+}$, $Tardbp^{F210I/+}$, $Tardbp^{F210I/F210I}$), suggesting either autoregulation is preserved or is not required during embryonic development.

$Tardbp^{F210I}$ ENU mutant mice thus represent the first loss of function model generated from isogenically expressed mutant TDP43 with a well-defined mechanistic consequence, specifically decreased RNA binding affinity. Additionally, it partially replicates, at the splicing level, the effects of TDP43 knockdown in the adult CNS (Polymenidou et al. 2011).

[4] The original work included in Thomas Ricketts' in his PhD thesis which is relevant for the original work being presented here will be reference as (Ricketts 2012), and the thesis itself referenced in the bibliography section.

The identification of a loss of function ENU mutant provided an opportunity to investigate further the molecular consequences of mutation-induced TDP43 loss of function and, critically, assess the *in vivo* effects at the whole animal level of the TDP43 "loss of normal function" through phenotypical behavioural testing for motor function, learning and memory and endophenotypes of neuropsychiatric diseases. I therefore continued working on the *Tardbp*^{F210I} mutant mouse line pursuing further molecular characterisation and performing the complete phenotypic characterisation and the results are presented in this thesis.

In addition, the M323K mutation (figure 1.12) presented the opportunity to study the consequences of mutations in the C-terminal of TDP43. Despite not being analogous to a human mutation, the M323K mutation affects a conserved residue in TDP43 and is located in the region where disease causing mutations cluster. Furthermore, as exposed in the earlier sections of this introduction, the C-terminal is essential in a variety of key processes in TDP43 metabolism including protein-protein interactions, alternative exon splicing and the autoregulation process. Characterising the M323K mouse line both molecularly and behaviourally thus offered a significant opportunity to further understand TDP43 biology and how dysregulation can lead to neurodegeneration, particularly when integrated with the results obtain with the F210I loss of function mutant.

The characterisation of the mouse mutants was undertaken with the aim of understanding TDP43 biology, endeavouring to unravel the functional consequences of TDP43 mutations and its mechanistic links with neurodegeneration, and, in particularly, ALS and FTLD-TDP.

Chapter 2.

Materials and Methods

The methodologies and materials used in the experimental work presented in this thesis will be described, divided into sections of breeding and colony maintenance, molecular characterisation and *in vivo* phenotyping; the latter section includes all behavioural tests. At the end for the chapter a table will be given with the manufacturer and catalogue numbers of the reagents used.

All procedure using animals were performed in accordance to the Animal Scientific Procedures Act 1998 and experiments performed under project license 30/2995 and personal license L314.

The mice used in the work presented in this thesis were housed at the Mary Lyon Centre, in individually ventilated cages under certified "Pathogen-free environment" and husbandry procedures were performed by the staff from the Mary Lyon Centre. The mice were observed daily ensuring its welfare and the humane end points were determined by the project license and included weight loss of 20% body weight and general signs of unwelness (e.g piloerect coat, limping which conditions access to food, wounds). Additionally, the Named Care and Welfare Officer (NACWO) had the discretion of determining when the health status of a mouse required it to be sacrificed.

The *Tardbp* mutants, $Tardbp^{F210I}$, $Tardbp^{M323K}$, $Tardbp^{Q101X}$ were re-derived by using sperm matched to the archived DNA to perform *in vitro* fertilisation. $Tardbp^{Q101X}$ was identified in the archive screening from the MRC Mammalian Genetics Unit; $Tardbp^{F210I}$ and $Tardbp^{M323K}$ were identified in the archive of the Riken institute in Japan, where the archive was screened in collaboration with the research group of Professor Yoichi Gondo.

At weaning age (around postnatal day 21) animals were ear marked with ear punch biopsy, which provided a means of identifying individual animals in a cage and, additionally, provided tissue for DNA extraction and genotyping. The offspring of all crosses was selected

by genotyping for the specific *Tardbp* mutations and mice kept being backcrossed with C57BL/6J wild type animals. Breeding was achieved by mating, as a rule, one male (*Tardbp* mutant) with two wild type females. Mice were only included in the testing cohorts or its tissue included in molecular experiments after being backcrossed for at least five generations onto the C57BL/6J background.

The *TARDBP* BAC transgenic mice were a generous gift from Professor Jean-Pierre Julien and its generation and characterisation has been previously described (Swarup et al. 2011). These mice were kept in the same husbandry conditions as the ENU mutants.

2.1 Molecular laboratory techniques

2.1.1 Genomic DNA extraction from embryonic and adult mouse tissue

A small fragment of embryonic or adult mouse tissue (typically ear clip tissue or fragment of tail) was incubated in 200µl of TENS buffer (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 0.1N NaOH, 0.5% SDS) and 10µl of Proteinase K in a 55°C water bath for a time period between 2 hours and 48 hours. Subsequently 200µl of isopropanol was added to each sample and immediately mixed by inverting the tube several times. The DNA was precipitated by centrifugation at 14,000rpm, 4°C for 10 minutes, the supernatant discarded and the DNA pellet air dried. Lastly, the DNA was eluted by adding 200µl of TE buffer (10 mM Tris-Cl pH 8.0 and 1 mM EDTA) to each sample.

2.1.2 Genotyping for Tardbp ENU mutations

The concentration of the extracted genomic DNA was measured by spectrophotometry (Nanodrop 8000, *Thermo Scientific*) and diluted in water to a final concentration of 5ng/µl and genotyping performed via pyrosequencing as described below.

PCR was performed with specific biotinylated primers for each mutant line (table 2.1) in order to amplify the region of interest, where the mutation was located for each mouse line, using the reaction mix displayed in table 2.2 and thermal cycle displayed in table 2.3.

The PCR products were subjected to pyrosequencing using the specific forward sequencing primer for each mutation (table 2.1).

***Tardbp* genotyping primers**		
Mutation	**Primer Sequence**	
M323K	Forward Primer	5'-GAATCAGGGTGGGTTTGGTAACAG-3' (*MWG Operon*)
	Reverse primer (biotinylated)	5'-GGTTCTGCTGGCTGGCTAACA-3' (*MWG Operon*)
	Sequencing Primer (Pyrosequencing)	5'-GCATTAACCCAGCGA-3' (*MWG Operon*)
Q101X	Forward Primer	5'-CAAAAGGAAAATGGATGAGAC-3' (*MWG Operon*)
	Reverse primer (biotinylated)	5'-AGTTGTTTTCCAGGGGAGAC-3' (*MWG Operon*)
	Sequencing Primer (Pyrosequencing)	5'-GAAAGTGAAAAGAGCAGTC-3' (*MWG Operon*)
F210I	Forward Primer	5'-GGACGTTGTACAGAGGACATGACT-3' (*MWG Operon*)
	Reverse primer (biotinylated)	5'-TTGGGAATGAAGACATCTACCAC -3' (*MWG Operon*)
	Sequencing Primer (Pyrosequencing)	5'-TTCTCCATACTGACAGAAA-3' (*MWG Operon*)

Table 2.1: Sequence of primers used to genotype *Tardbp*^{M323K}, *Tardbp*^{Q101X} and *Tardbp*^{F210I} ENU mutants

PCR for pyrosequencing genotyping reaction mix	
Reagent	**Volume**
Master Mix (*Qiagen* taq pcr mix 201445)	5µl
Forward Pimer (10µM)	0.2µl
Reverse Primer (10µM)	0.2µl
Distilled H_20	2.6µl
Genomic DNA (5ng/µl)	2µl

Table 2.2: Reaction mix used in in the PCR amplification of *Tardbp*^{M323K} and *Tardbp*^{Q101X} genomic DNA for pyrosequencing genotyping

Temperature	**Time**	
95°C	5 minutes	
95°C	15 seconds	44 Cycles
55°C	30 seconds	
72°C	15 seconds	
72°C	5 minutes	

Table 2.3: Thermal cycle used in PCR amplification of the target region from *Tardbp*^{M323K}, *Tardbp*^{Q101X} and *Tardbp*^{F210I} genomic DNA for pyrosequencing genotyping

Pyrosequencing consists of a single base sequencing methodology which is achieved by mixing the PCR products (where the base of interest is located) with the sequencing primer,

DNA polymerase, ATP sulfurylase, luciferase, apyrase, adenosine 5'-phosphosulfurate and luciferin, the latter two acting as substrates for the enzymatic reaction.

The reaction mixture is heated in the Pyrosequencer (PSQ 96, *Qiagen*), allowing the sequencing primer to hybridise with the biotinylated strand. Deoxynuribonucleotide triphosphates (DNTPs) are added step by step (automatically by the pyrosequencer) and when the DNTP containing the complementary base to the one in the template is added, its incorporation in the replicating strand leads to the release of pyrophosphate. The latter is converted to Adenosine triphosphate (ATP) which drives the luciferase-catalysed reaction of luciferin to oxyluciferin to occur, leading to the emission of a light pulse. The light pulse emitted is detected by the machine enabling determination of the DNTP incorporated and therefore the base pair present at that genomic site of interest.

Further detailed information regarding the pyrosequencing technology can be found on the Qiagen website (http://www.qiagen.com/gb/resources/technologies/pyrosequencing-resource-center/).

The output images from the pyrosequencing enable rapid and efficient identification of the genotypes. Representative genotyping images for *Tardbp*^{M323K}, *Tardbp*^{Q101X} and *Tardbp*^{F210I} can be seen in figures 2.1, 2.2 and 2.3.

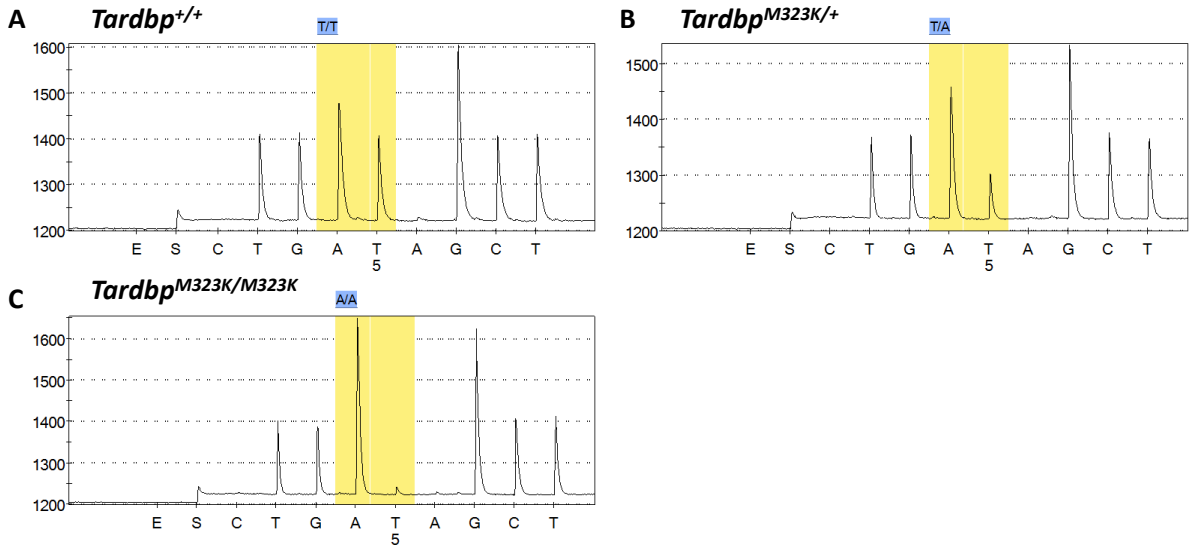

Figure 2.1: Representative *Tardbp*^{M323K} genotyping images from the graphical output produced by the pyrosequencer.

The M323K mutation results from a single base pair change in the coding sequence of *Tardbp*, specifically ATG>AAG.

The amplitude of the light signal emitted during pyrosequencing is proportional to the number of bases incorporated and is represented graphically by the amplitude of the peak seen in the graph.

A) In *Tardbp*$^{+/+}$ animals two Thymine DNTPs are incorporated at the position of interest (yellow shading, represented as T) leading to a peak with an amplitude of 1400 (arbitrary units).

B) In *Tardbp*$^{M323K/+}$ animals one Thymine DNTP is incorporated and one Adenine DNTP is incorporated leading to an increase in the amplitude of the light signal peak of the Adenine base and decrease in the amplitude of the peak in the T base.

C) In *Tardbp*$^{M323K/M323K}$ animals only Adenine DNTPs are incorporated which is reflected in the amplitude of the light peaks.

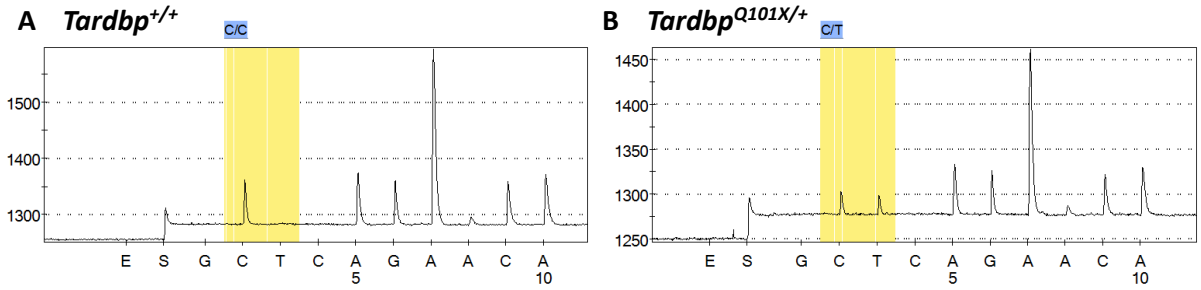

Figure 2.2: Representative Tardbp^{Q101X} genotyping images from the graphical output produced by the pyrosequencer.

The Q101 stop mutation results from a single base pair change in the coding sequence of *Tardbp*, specifically CAG>TAG.

A) In *Tardbp*$^{+/+}$ two Cytosine deoxyribonucleotides are incorporated leading to a single peak.

B) In *Tardbp*$^{Q101X/+}$ one Cytosine and one Thymine deoxyribonuleotides are incorporated leading to two peaks of similar amplitudes. Homozygous animals or embryos were never identified.

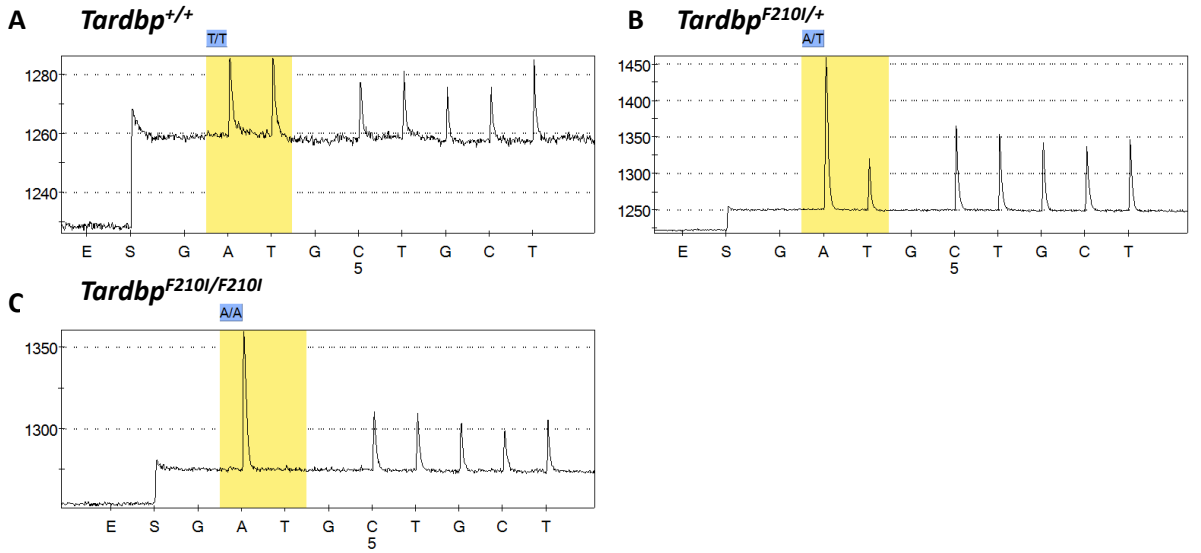

Figure 2.3: Representative Tardbp^{F210IX} genotyping images from the graphical output produced by the pyrosequencer.

The F210I mutation results from a single base pair change in the coding sequence of *Tardbp*, specifically TTT>ATT

A) In *Tardbp*$^{+/+}$ two Thymine deoxyribonucleotides are incorporated leading to a tall peak in the region of interest. The Adenine (A) peak is related to the incorporation of an Adenine deoxyribonucleaotude in the immediate 5' upstream codon.

(B) In *Tardbp*$^{F210I/+}$ one Thymine and one Adenine deoxyribonuleotides are incorporated leading to the change in the relative amplitude of the A and T peaks.

C) In *Tardbp*$^{F210I/F210I}$ only Adenine deoxyribonucleotides are incorporated leading to a single A peak in the region of interest.

2.1.3 Genotyping TARDBP^{G348C} transgenic animals

Transgenic mice were genotyped by PCR with primers aligning on the transgene sequence. After DNA was extracted, a 1:10 dilution was made and the PCR performed using the primers and conditions detailed below. The PCR products were subjected to agarose gel electrophoresis (1% agarose) containing gel red and visualised under Ultraviolet light. Given that the primers aligned with the transgene, a band was only seen in reactions using template DNA of animals carrying the transgene.

- Forward Primer: 5'-CTCTTTGTGGAGAGGAC-3'
- Reverse Primer: 5'-GGATTAATGCTGAACGT-3'

Temperature	Time	
95°C	3 minutes	
95°C	45 seconds	34 Cycles
55°C	45 seconds	
72°C	45 seconds	
72°C	10 minutes	

Table 2.4: Thermal cycle used in PCR reaction for genotyping TARDBP BAC transgenics

2.1.4 Harvesting, culturing and storing Mouse Embryonic Fibroblasts (MEFs) and embryonic heads

Tardbp$^{M323K/+}$ animals were mated and vaginal plugging monitored (day that plug is observed counts as ½ day post conception). Females were sacrificed at 14.5dpc and embryos incubated in a 10cm cell culture dish with 8mls of Phosphate Buffered Saline (PBS) on ice.

Embryos were excised from its sac and dissociated from the placenta, the head was severed and immediately frozen in dry ice in a 1.5ml microcentrifuge tube and a small segment of tail was used for DNA extraction and genotyping as described above. Subsequently, embryos were macroscopically dissected, being dismembered and eviscerated, washed in PBS and the resulting carcass transferred into its individual 5cm cell culture dish and incubated in PBS, on ice, until all dissections were completed. When all dissections were completed, the carcasses were subjected to enzymatic digestions by incubation with 3mls of Trypsin-EDTA in a 5cm cell culture dish, further severed and dissociated using 2 scalpel blades, homogenised with a 5ml pipette, transferred into individual 50ml tubes and incubated in a 37°C water bath for 5 minutes with gentle swirling. After the trypsin digestion, the tubes containing the suspension of tissue from individual embryos was taken into the tissue culture room and all procedure from that point on performed in a class II laminar flow cabinet.

The tubes were opened and the tissue suspension further homogenised by pipetting up and down with a 5 ml pipette and added directly into 10cm cell culture dishes containing 7 ml of pre-warmed “MEFs” culture media (Dulbecco’s Complete Media containing 4.5g/l glucose,

Invitrogen, supplemented with 10% Foetal Bovine Serum, Ultraglutamine II, 1xpenicillin/streptomycin).

When harvesting MEFs from embryos, tissue culture dishes were equilibrated with PBS-0.1% gelatine (PBS-0.1%G) for 1 hour at 37°C (in the same incubator the MEFs are grown) after which the PBS-0.1%G was removed and MEFs media added.

Cells were observed daily and passaged when they were 70-80% confluent. Passaging was achieved by washing MEFs twice with 10mls of PBS, digesting with 3mls of trypsin, adding MEFs media to the dishes following digestion and transferring the cell suspension into new tissue culture dishes with a final volume of 10mls (typically MEFs were passaged 1:3). Tissue culture dishes did not require incubation with PBS-0.1%G to received MEFs that had already been in culture.

Freezing MEFs Stocks: Cells harvested by centrifugation were re-suspended in 1ml of media containing 90% Foetal Bovine Serum and 10% Dimethyl Sulfoxide (DMSO) and transferred into cryotubes. The cryotubes were placed in a recipient containing isopropanol at the bottom and frozen at -80°C, to allow for a more gradual decrease in the temperature. Once frozen, the cell suspension was transferred into liquid nitrogen for longer term storage.

Harvesting and storing MEF pellets: Cell pellets were harvested by after trypsin digestion centrifugation and immediately frozen in dry ice and kept at -80°C.

2.1.5 CFTR minigene transfection in MEFs

The CFTR minigene was a generous gift from the Buratti and Baralle lab and the construct's design and execution has been previously described, refer to the "human CFTR minigene" in (Niksic et al. 1999); a diagram of the minigene can be found in chapter 4, figure 4.1.

Frozen MEF aliquots were thawed on the 37° water bath and pelleted by centrifugation at (5,000rpm, 1 minute). The freezing media was removed by aspiration, the cells re-suspended into 1ml of MEFs media and added to a 10cm dish containing 9mls of MEFs media. Cells were passaged according to their growth rate and harvested by centrifugation when all individual dishes were 60-80% confluent (by digestion with trypsin and centrifugation).

After centrifugation, the supernatant was discarded and the MEFs subsequently re-suspended in 100µl of supplemented mouse/rat Nucleofactor solution previously warmed to room temperature, 2.5µg of CFTR minigene plasmid DNA (concentration of 500ng/µl) was added to each sample, the suspension was transferred into the amaxa cuvette and each sample electroporated individually in a Nucleofector II device (*Amaxa Biosystems*). After electroporation, 500µl of MEFs media was added to each cuvette and the cell suspension transferred to a 10cm tissue culture dish containing 9.4mls of MEFs media pre-warmed to 37°C and incubated at 37°C and 5%CO_2 in a humidified incubator. Approximately 36 hours post-transfection, the cells were pelleted by centrifugation and the RNA extraction performed, followed by cDNA synthesis and splicing PCRs described in the next sections.

2.1.6 RNA extraction from embryonic heads and MEFs (including MEFs transfected with the CFTR minigene)

Frozen heads were thawed on ice and transferred to tubes containing ceramic spheres. 1ml of Trizol was added to each head and homogenisation performed in a FastPrep homogeniser in cycles of 30 seconds until no solid tissue remained (Fastprep FP 120, *Thermo Scientific*) , followed by incubation at room temperature for 5 minutes. 200µl of chloroform was added to each sample and the tubes briefly shaken, incubated at room temperature for 25 minutes and submitted to centrifugation at 12,000g, 4°C for 15 minutes. The aqueous phase was transferred into a clean 1.5ml microcentrifuge tube, 500µl of isopropanol was added, and the samples were mixed by light vortex action and incubated at room temperature for 20 minutes.

The RNA was precipitated by centrifugation (12,000g, 4°C, for 5 minutes) and the supernatant discarded. The pellet was washed with 1ml of 75% ethanol and a last centrifugation step (7,500g, 4°C, for 5 minutes) performed. The supernatant was discarded, the pellet briefly air dried (1-3 minutes) and the RNA eluted in 100µl of nuclease-free water.

For RNA extraction from MEFs the same sequence of experimental procedures were taken, with the only difference being the volumes of reagents used at each step, 100µl of Trizol, 20 µl of chloroform and 30 µl of nuclease free water was used. Homogenisation was achieved by pipetting the cell pellet up and down instead for using the Fastprep homogeniser.

2.1.7 cDNA synthesis from template RNA extracted from MEFs and embryo heads

The concentration of RNA extracted from MEFs and embryo heads was measured by spectrophotometry (Nanodrop 8000, *Thermo Scientific*) and cDNA synthesis reactions performed using the Taqman™ Reverse transcription Kit according to the reaction mix and thermal cycling specifications in tables 2.5 and 2.6.

cDNA synthesis reaction mix	
Reagent	**Volume**
10xRT reaction buffer	2µl
MgCL2	4.4µl
DNTP mix	3µl
oligo DT	0.5µl
RNAi	0.3µl
RT	0.5µl
RNA	Volume equivalent to1 microgram
H_2O	To achieve final volume of 20µl

Table 2.5: Reaction mix used for cDNA synthesis from RNA extracted from *Tardbp* mutant embryo heads and MEFs

Temperature	**Time**
25°C	10 minutes
48°C	35 minutes
95°C	5 minutes

Table 2.6: Thermal cycled used in the cDNA synthesis from RNA extracted from *Tardbp* mutant embryos and MEFs

2.1.8 PCR splicing assay of endogenous genes: Sortilin-1 Poldip-3, Dnajc5, Pdp-1

After cDNA synthesis, assessment of alternative splicing of endogenous genes were performed using the specific primers for each gene (table 2.7) and following the conditions described in the tables below.

Gene	Primer Sequence	
Sort1	Forward Primer	5'-CAGGAGACAAATGCCAAGGT-3' (*Sigma*)
	Reverse Pimer	5'-ACAAGCATCAGTCCCACGAT-3' (*Sigma*)
Poldip-3	Forward Primer	5'-CATTGGGACTGTAACCCCAG-3' (*Sigma*)
	Reverse Pimer	5'-TGCAAACTTCATTCTGCTTGG-3' (*Sigma*)
Dnajc5	Forward Primer	5'-CTCTATGTGGCGGAGCAGTT-3' (*Sigma*)
	Reverse Pimer	5'-GCTGTATGACGATCGGTGTG-3' (*Sigma*)
Pdp-1	Forward Primer	5'-GTGCTGAGTGAGGGAAGGAC-3' (*Sigma*)
	Reverse Pimer	5'-TGCAGTGCCATAGATTCTGC-3' (*Sigma*)

Table 2.7: Sequences of the primers used to perform the RT-PCR splicing assay of endogenous target genes in *Tardbp* ENU mutants

Endogenous genes splicing PCR reaction mix	
Reagent	**Volume**
Master Mix (Qiagen taq pcr mix 201445)	5µl
Forward Primer (20µM)	0.3µl
Reverse Primer (20µM)	0.3µl
H_2O	2.4µl
cDNA	2µl

Table 2.8: Reaction mix used in the RT-PCR splicing assay of endogenous target genes in *Tardbp* ENU mutants

Temperature	Time	
94°C	3 minutes	
94°C	45 seconds	34 cycles
55°C	45 seconds	
72°C	45 seconds	
72°C	10 minutes	

Table 2.9: Thermal cycle used for the RT-PCR splicing assay of endogenous target genes in *Tardbp* mutants

2.1.9 PCR splicing assay of the CFTR minigene

cDNA from the MEFs transfected with the CFTR minigene was used as the template for the RT-PCR splicing assay. The primers used align flanking regions of the cassette exon (refer to

figure 4.1 in chapter 4 for a diagrammatic representation of the minigene and primer alignment region) and its sequences are;

Forward primer: 5'-CAACTTCAAGCTCCTAAGCCACTGC-3' (*MWG Operon*)
Reverse primer: 5'-TAGGATCCGGTCACCAGGAAGTTGGTTAAATCA-3' (*MWG Operon*)

The reaction mix and thermal cycle used are detailed in tables 2.10 and 2.11.

CFTR minigene splicing assay reaction mix	
Reagent	**Volume**
Master Mix (Qiagen taq pcr mix 201445)	15.83µl
Forward primer (10µM)	2µl
Reverse Primer (10µM)	2µl
H_2O	1.17µl
cDNA	4µl

Table 2.10: Reaction mix used in the CFTR minigene splicing assay

Temperature	**Time**	
94°C	5 minutes	
94°C	30 seconds	25 cycles
55°C	30 seconds	
72°C	30 seconds	
72°C	7 minutes	

Table 2.11: Thermal cycle used in the CFTR minigene splicing RT-PCR

2.1.10Agarose gel electrophoresis of PCR products and quantification

The nucleic acid products amplified in the splicing RT-PCRs of target endogenous genes and CFTR minigene splicing assay were analysed by agarose gel electrophoresis. After 2µl of loading dye was added to the products of each reaction:

- The products from the endogenous genes' splicing assays were loaded in a 3% agarose (Nusieve 3:1) gel with Etidium Bromide and electrophoresis performed at 100mv; 100 base-pair ladder was used as a size marker.
- The products from the CFTR minigene splicing assay were loaded into a 1% agarose gel, with Ethidium Bromide. Electrophoresis was performed at 100 mv (for approximately 1 hour). 100 base-pair ladder was used as a size marker.

Quantification for the purposes of determining the ratio of exon exclusion was achieved by using the freely available *ImageJ* software to quantify band intensity. The percentage of exon exclusion was determined by using the formula:

% exon exclusion=[Lower band intensity/(Top band intensity+Lower band intensity)]X100

2.1.11 Northern Blot for TDP43 RNA

This experiment was planned and designed together with Jeremias Herzog, a PhD student at the ICGEB in Trieste. MEFs were grown and harvested at the MRC Genetics Mammalian Unit and all other experimental steps performed by Jeremias Herzog at the ICGEB in Trieste. The probe was designed to hybridise within exon 5, which is present in the vast majority of *Tardbp* transcripts identified in the genomic database Ensembl (figure 1.1). It was amplified from cDNA using the forward (5'-CAAAGCCCAGACGAGCCT-3') and reverse (5-CTTATCATCTGCAAAGGTGACG-3') primers and labelled radioactively with Phosphorus 32 (P^{32}) using the Rediprime II DNA Labeling System (GE Healthcare).

MEFs were cultured and harvested at the MRC Mammalian Genetics Unit and the pellets immediately suspended in 150µl of Trizol and sent to Jeremias Herzog frozen in dry ice. RNA was extracted with EuroGoldtriFast (Euroclone) following manufacturer's instructions, the samples loaded on 1.2% formal- dehydeagarose gels and after electrophoresis transferred to Hybond-N+ nylon membranes (Amersham Biosciences). Prehybridisation was performed with ULTRAhyb® ultrasensitive hybridisation buffer (Ambion) followed by hybridisation at 65ºC. Visualisation was achieved by exposing a phosphor imaging screen to the membrane (Cyclone Phosphoimager Screen) and revealed in the Cyclone system (Cyclone Phosphoimager hardware and software). Quantification was performed using *ImageJ* software.

2.1.12 RNA-Seq

Embryo heads and MEF pellets were harvested at the MRC Mammalian Genetics Unit as already described and sent to Piettro Fratta at the Institute of Neurology. RNA was extracted using the miRNeasy Mini Kit (*Qiagen*) according to the manufacturer's instructions and quality verified by Agilent Bioanalyzer NanoChip. Ribo-Zero columns (*Epicenter Biotechnologies*) were used to reduce ribosomal RNA abundance and RNA-seq libraries were performed (paired-end 2 × 36 cycles by Illumina GAIIx). Sequencing reads were aligned against the NCBI build 37.2 of the *Mus musculus* reference genome using the software *tophat* and alignment incorporated the reference genes annotations (GTF format) provided as part of the Illumina iGenomes project. Duplicate reads were filtered using the PICARD tools. Read counts per gene and per exon were computed using the python scripts provided with the *DESeq/DEXSeq* packages. Differential expression at the gene level was computed using the *R package DESeq* and at the exon level using the *R package DEXSeq*.

RNAseq was performed with 3 biological replicates for each genotype for both embryo heads and MEFs (n=3 $Tardbp^{+/+}$, 3 $Tardbp^{M323K/+}$ and 3 $Tardbp^{M323K/M323K}$ for the M323K mouse line on the C57BL/6J background and n= 3 $Tardbp^{+/+}$, 3 $Tardbp^{F210I/+}$ and 3 $Tardbp^{F210I/F210I}$ for the F210I mouse line on the C57BL/6J background).

2.1.13 Whole cell protein extraction from MEFs and embryo heads

In MEFs, protein was extracted by adding 100µl RIPA buffer (150 mM sodium chloride, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) with protease and phosphatase inhibitors to the MEF pellets, pipetting up and down and incubated on ice for 15 minutes. The samples were centrifuged at 10,000rpm, 4°C for 5 minutes and the supernatant transferred into a new microcentrifuge tubes.

Protein extraction from embryonic heads was performed by transferring the thawed embryo heads to tubes containing ceramic spheres adding 1ml of RIPA buffer and homogenising in a

FastPrep homogeniser for 30 seconds (Fastprep FP 120, *Thermo Scientific*) followed by incubation on ice for 15 minutes. Subsequently, the homogenate was transferred into a new microcentrifuge tube, the phases separated by centrifugation 10,000rpm, 4°C for 5 minutes and the supernatant, containing the protein, aliquoted into new microcentrifuge tubes. All eluted protein samples were stored at -80ºC.

2.1.14 Protein concentration measurement in samples extracted with RIPA buffer

In all protein samples extracted using RIPA buffer, protein concentrations were measured by the Bradford Reagent method. Calibration standards and a blank reading were prepared with known concentrations of BSA in RIPA buffer (0, 0.125, 0.25, 0.5, 1, 2 micrograms/ml).

Protein samples were diluted 1:5 so that the detergents in the buffer did not interfere with the Bradford reaction. 5µl of the diluted samples (and standards) were added to an optical reading plate, in three biological replicates followed by 200µl of room-temperature Bradford reagent. Subsequently the plate was read in the spectrophotometer at an absorbance of 595nm and the software *KC junior* estimated the protein concentrations accordingly.

2.1.15 Immunobloting for TDP43 protein levels

In assessing TDP43 levels, the same amounts of protein extracted from MEFs, embryo heads or brains for animals of all *Tardbp*^{M323K} genotypes (typically 15 to 30mcg) were prepared in a loading mix as per table 2.12 and denaturation completed by incubating the sample mixes at 95ºC for 5 minutes, after which they were briefly centrifuged (10,00rpm for 1 minute) and loaded in precast 10% Bis-Tris polyacrylamide gels, together with commercial molecular weight markers (See Blue Plus Two). Polyacrylamide Gel Electrophoresis (PAGE) performed with voltages between 130mv-200mv in MOPS running buffer until all sample die ran off the gel and good separation of the molecular marker was observed.

Reagent	Volume
4x SDS Sample Buffer	2.5µl
20x Reducing Agent	1µl
Protein	Equivalent to 15mcg to 30mcg
H_20	To a final volume of 10µl

Table 2.12: Sample mix for protein SDS-PAGE in commercially acquired pre-cast mini gels

After electrophoresis, the protein was transferred from the gel into a nitrocellulose membrane using the iBlot or a low fluorescence Polyvinyl Difluoride (PVDF) membrane using a semi-dry system.

In the semi-dry system, the PVDF membrane was activated by briefly immersing it in absolute methanol, and a "sandwich" consisting of filter paper, membrane, polyacrylamide gel and filter paper, all soaked in transfer buffer (0.05M Tris, 0.03M Glycine, 0.04% SDS, 20% Methanol), layered. Transfer was completed by applying an electrical current of 15 volts for 90 minutes.

Using the commercial iBlot system from Invitrogen (catalogue number IB1001) a "sandwich" consisting of the pre-packed anode, which contains the nitrocellulose membrane, polyacrylamide gel, filter paper and cathode was layered. The standard programme was selected in the iBlot machine which applied an electrical current for seven minutes, concluding the transfer.

After transfer, the membrane (regardless of the transfer method) was blocked in 5% skimmed milk in PBS-0.2% Tween (PBST0.2%) for 1 hour at room temperature. If interrupting the experiment was necessary at this point, blocking was performed for a maximum time window extending for up to 48 hours at 4ºC. Incubation with the primary antibodies followed blocking and was performed in 5% skimmed/PBST0.2% for 1 hour at room temperature. The antibodies used, including its catalogue numbers and concentrations can be found in table 2.12. Subsequently, the membrane was washed three times with PBS-0.2% Tween for 20

minutes and incubated with the fluorescent secondary antibodies (table 2.12) in blocking solution with 0.01% SDS added, for 1 hour at room temperature and dark light conditions.

Lastly, the membrane was washed three times with PBST0.2%, allowed to dry and visualised using the Odyssey system. Quantification was performed using the *Odyssey* software to calculate band intensity.

2.1.16 Cross-linking and homodimerisation in Tardbp^{M323K} MEFs

Frozen MEFs were placed back into culture as described in the experimental methodology for the CFTR minigene transfection. Disuccinimidyl Glutarate (DSG) is a homobifunctional cross-linker which has been used successfully to covalently cross-link proteins which interact natively in cell systems or *in vitro* (Bartels et al. 2011; Lo et al. 2006). Additionally, its use has been proven to be effective in covalently cross-linking TDP43 proteins expressed in mammalian cells (Zhang et al. 2013). The experimental methodology used was performed to replicate the Cross linking protocol used by Zhang and colleagues and I am grateful to Yongie Zhang for his advice and clarifications when planning the experiment.

MEFs were passaged upon reaching 70% confluence and allowed to proliferate until reaching 60-80% confluence, at which point the cells were harvested, transferred into individual microcentrifuge tubes, washed twice with ice cold PBS and resuspended in 100µl of room temperature PBS.

MEFs of each original vial were exposed to different concentrations of DSG, specifically 25µM, 50µM and 100µM suspended in DMSO. Working DSG solutions of 2.5mM, 5mM and 10mM were made and 1µl was added to each experimental sample to achieved the final concentrations specified above. As a negative control, 1µl of DMSO only was added to the cells. The cells were incubated at room temperature for 30 minutes with shaking. The cross-

linking reaction was quenched by the addition of 2µl of 1M Tris-Base (to a final concentration of 20mM) and incubation for 15 minutes, with shaking, at room temperature.

Subsequently, protein was extracted by adding Lysis Buffer (50mM Tris-HCL pH 7.4, 300mM Nacl, 1% Triton X-100, 1% SDS, phenylmethylsulfonyl fluoride, 1 X Phosphatase and Protease inhibitors) and incubating on ice for 15 minutes. The samples were centrifuged at 16,000g, 4 ºC for 20 minutes and the supernatant, containing the protein, aliquoted into new microcentrifuge tubes. The protein concentration was measured using the BCA assay from Biorad (http://www.bio-rad.com/en-uk/product/dc-protein-assay) according to the manufacturer's instructions. BSA standards were the same as used for the Bradford reagent protein measuring method. Briefly, 20µl of reagent S were added to 1ml of reagent A, creating a working reagent A'. Protein samples were diluted 1:5 and 5µl added to each well in an optical reading plate; in triplicate, followed by 25µl of reagent A' and 200µl of reagent B, allowing it to mix. After 15 minutes, the plate was read in the spectrophotometer at the absorbance of 750nm and the software provided an output of the concentrations.

Samples were prepared as described in table 2.11, heated to 95°C for 5 minutes and all steps including SDS-Page and immunoblot, as well as quantification, were performed as described in the preceding section, using the iBlot transfer system. HiMark protein standard was used as the molecular weight marker. The gel selected to show the result of the experiment in this thesis consists of the experiment that was performed with a final concentration of 100µM of DSG.

2.1.17 Extraction of Nuclear/Cytoplasmic Protein fractions

MEFs were harvested directly from culture by trypsin digestion when confluence reached 70% and transferred into a microcentrifuge tube.

Extraction of nuclear/cytoplasmic fractionated protein performed using the NE-PER kit from *Roche* (http://www.piercenet.com/product/ne-per-nuclear-protein-extraction-kit).

Once in the microcentrifuge tube, MEFs were washed with PBS, and re-suspended in 100µl of ice-cold CER I buffer, vortexed vigorously for 15 seconds and incubated on ice for 10 minutes. 5.5µl of ice-cold CER II buffer was added to the samples followed by vortexing for 5 seconds, incubation on ice for 1 minute and centrifugation at 12,000g, 4ºC for 5 minutes. The supernatant, containing the cytoplasmic protein fraction was transferred into new, pre-chilled, microcentrifuge tube.

The remaining pellet, containing the nuclei with the intact nuclear membranes, was re-suspended in 50µl of ice-cold NER buffer by vortexing vigorously for 15 seconds, and the suspension incubated on ice, being vigorously vortexed for 15 seconds every 10 minutes for a total of 40 minutes. Lastly, the two phases were separated by centrigation at 12,000g, 4ºC for 10 minutes and the supernatant containing the nuclear protein fraction transferred into new microcentrifuge tubes. Sample preparation and loading, SDS-PAGE, semi-dry transfer and immunoblotting were performed as described in section 2.1.1.5.

Detailed information concerning the antibodies used in TDP43 immunoblotting can be found in table 2.13.

2.1.18 Antibodies used in TDP43 Immunoblotting experiments

The table below gives detailed information regarding all antibodies used in TDP43 immunoblotting experiments.

Antigenic protein	Characteristics	Manufacturer and catalogue number	Experiments	Concentration
TDP43 (primary)	Rabbit Polycloncal	Protein-tech; 10782-2-AP	TDP43 immunoblotting in MEFs (including homodimerisation and nuclear/cytoplasmis fractionation), embryonic heads and adult brain	1: 1,500
Alpha-Tubulin (primary)	Mouse Monoclonal	Developmental Studies Hybridoma Bank (University of Iowa); 12G10	Loading control in TDP43 immunoblots	1:5,000
Nuclear Matrix protein P84 (primary)	Rabbit Polyclonal	Abcam; ab102684	TDP43 immunoblotting of Nuclear and Cytoplasmic protein fractions	1:5,000
Mouse Immunoglobulins (secondary)	Goat Polyclonal	LI-COR Biosciences; 926-32210	Secondary antibody against anti-Tubulin primary antibody	1:15,000
Rabbit Immunoglobulins (secondary)	Goat Polyclonal	LI-COR Biosciences; 926-32221	Secondary antibody against anti-TDP43 primary antibody	1:15,000

Table 2.13: Properties, manufacturers and concentrations of antibodies used in immunoblotting experiments

2.1.19 Synthesis and purification of recombinant GST-TDP43 proteins

The synthesis and purification of recombinant TDP43 proteins and the experiments in which the recombinant proteins were used for were performed at the Buratti and Barralle and lab in the ICGEB in Trieste. Mouse TDP43 cDNA, both with the wild type sequence and sequences of ENU mutants (e.g. M323K, F210I and N312S) had been previously cloned by Thomas Ricketts into the pGEX5X3 bacterial expression vector (GE Healthcare). The cDNA was inserted into the multiple cloning site of the vector in frame with the upstream Glutathione S-Transferase (GST) tag, which enabled the synthesis of N-terminal GST tagged TDP43 wild type or mutant proteins.

BL-21 competent *E. coli* were transformed with by adding 1µl of previously purified plasmid (at a concentration of 10ng/µl) to a 50µl aliquot of cells (provided by the Buratti and Baralle lab), incubating on ice for 20 minutes, followed by heat shock in the water bath for 2 minutes and a final incubation on ice for an additional 2 minutes and 30 seconds. Subsequently 60µl of Luria Broth (LB) media was added and the cells incubated ay 37ºC, with shaking, for 30 minutes and finally spread into LB-agar plates with 75mcg/L of ampicillin, which were incubated overnight at 37ºC.

The following day, single colonies were picked with an autoclaved pipette tip which was added to a 50ml conical tube containing 10mls of LB/Ampicillin (100mcg/L) and grown overnight at 37ºC with shaking.

On the following morning, 4mls of the bacterial suspension grown overnight were inoculated onto 1L conical glass flasks containing 400mls of LB/Ampicillin (100mcg/L) and grown at 37ºC, with shaking until the optical density reached 0.5, at which point the BL-21 cells were induced to express the recombinant protein by adding Isopropylthio-β-Galactoside (IPTG) to a final concentration of 0.3mM (12µl of 1M IPTG was added). Protein synthesis was allowed to progress by incubation at 37ºC, with shaking, for 3 hours, after which the bacterial cell suspension was centrifuged at 5000g for 10 minutes, supernatant discarded, the pellet washed once with PBS and resuspended in 40mls in PBS-1% Triton with 1x protease inhibitors. The bacterial suspension was sonicated twice for 5 minutes (10-15 cycles, 20 seconds on – 30 seconds off, 15 microns) and centrifuged at 4000g, 4ºC for 20 minutes.

The supernatant was transferred into new conical tubes containing 500µl of GST-resin, pre-washed with 10ml PBS-1% Triton and incubated for 2 hours, 4ºC in rotary wheel-shaking movement. The samples were centrifuged at 500g, 4ºC, 10 minutes, the supernatant decanted out and the pellet/resin washed twice with PBS-1% Triton for 10 minutes with agitation. After the second wash and a new centrifugation step (500g, 4ºC, 10 minutes), the supernatant was discarded and the resin was transferred into a 2ml microcentrifuge tube, which was again centrifuged removing the remaining supernatant and leaving the resin only.

Serial elutions were then performed (table 2.14) and the supernatant (containing the recombinant protein) stored after phase separation was performed by centrifugation (500g, 4ºC, 10 minutes) for each elution.

	Elution Buffer Composition	Conditions
Elution 1	5mM Gluthatione, 100mM Tris pH8.8, 1mM DTT and 2x protease inhibitor cocktail	1ml buffer, 10 minutes at 4ºC with agitation
Elution 2	10mM Gluthatione, 100mM Tris pH 8.8, 1mM DTT and 2x Protease inhibitor cocktail	1ml buffer, 20 minutes at 4ºC with agitation
Elution 3	10mM Gluthatione, 100mM Tris pH 8.8, 1mM DTT and 2x Protease inhibitor cocktail	1ml buffer, 20 minutes at 4ºC with agitation
Elution 4	10mM Gluthatione, 100mM Tris pH 8.8, 1mM DTT and 2x Protease inhibitor cocktail	1ml buffer, 20 minutes at 4ºC with agitation
Elution 5	10mM Gluthatione, 100mM Tris pH 8.8, 1mM DTT and 2x Protease inhibitor cocktail	1ml buffer, overnight at 4ºC with agitation

Table 2.14: Elution buffer and conditions for each elution of the purified recombinant GST-TDP43 fusion proteins

2.1.20 Quantification of recombinant TDP43 proteins

The quantification of recombinant TDP43 proteins was performed by PAGE (12% polycrylamide gel) and coomassie staining by comparison of the sample's band intensity to the curve of BSA standards. 15µl of eluted recombinant protein was mixed with 5µl of 4 x loading dye, as were the BSA samples with concentrations of 1mcg, 2mcg, 4mcg and 8mcg. After electrophoresis, the gel was stained with coomassie blue solution (50% H_2O, 40% methanol, 10% acetic acid, 0.2% Coomassie powder) and scanned (figure 2.4), with quantification performed using *imageJ* software. The best protein, as seen by gel analysis, was from elution 3 and it was used for the GST Overlay and Electrophoretic Mobility Assay experiments. The polyacrylamide gels used in the experiment performed at the ICGEB in Trieste (EMSA and GST Overlay) were individually made and cast in the lab. After the experimental methodologies are described for the GST overlay/far western and EMSA, the details on how the gels were made will be provided on a dedicated section.

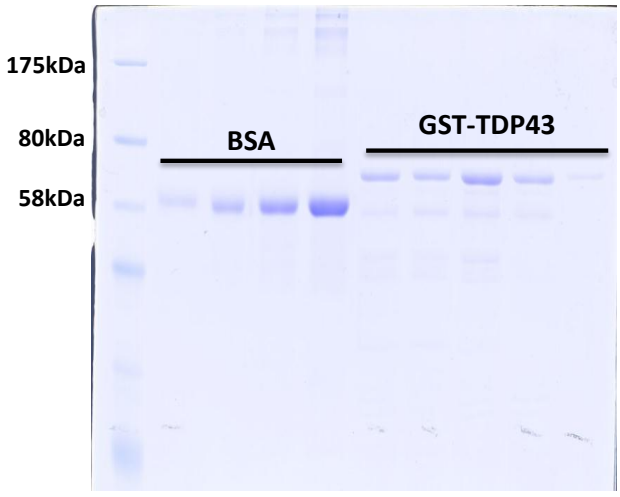

Figure 2.4: Recombinant GST-tagged TDP43 protein and BSA PAGE, post-stained with Coomassie blue

2.1.21 GST overlay/Far Western

15µl of commercially acquired S100 HeLa cell nuclear and cytoplasmic fractionated protein extracts were mixed with 4 µl of 5 x SDS running buffer (250mM Tris-HCL pH 6.8, 10% SDS, 30% Glycerol, 5% β-mercapthoethanol, 0.02% Bromophenol Blue) and 1 µl of water, incubated at 95°C for 3 minutes and loaded in a 10% polyacrylamide gel. Electrophoresis was performed with an initial current of 100mA through stacking and subsequently at 60mA until the loading die was no longer visible in the sample lanes and the molecular weight marker showed good separation.

After electrophoresis, transfer was performed using a fully wet methodology. A PVDF membrane was activated by immersion in absolute methanol and then immersed in transfer buffer (25mM Tris, 0.2M Glycine, 10% Methanol). The "filter paper, gel, membrane, filter paper" sandwich assembled in a plastic scaffold, the region of the membrane corresponding to every two wells of the gel marked with a pencil, and slotted into the transfer tank which was filled with transfer buffer. The transfer was performed by applying an electrical current of 200mA for 1 hour and 30 minutes. Once transfer was concluded, the membrane was cut into strips according to the previous marked areas, thus separating strips of membranes containing two groups of lanes (nuclear and cytoplasmic fractionated proteins).

The membrane strips were blocked with 5% skimmed milk in PBS-0.1% Tween (PBST0.1%) overnight at 4ºC and, the following day, individual strips were incubated with 2mcg of either wild type recombinant GST-TDP43 protein or mutant recombinant GST-TDP43 in 10% skimmed milk/ PBST0.1% for 2 hours at room temperature.

Subsequently, the membranes were washed four times with PBST0.1% (for 10 minutes) and incubated with anti-GST antibody (GE Healthcare; 27-457701) in a final concentration of 1:2,000 in 5% skimmed milk/PBS-0.1% Tween for 1 hour at room temperature, followed by four washes of 10 minutes in PBS-0.1% Tween and incubation with secondary anti-goat IgG

antibody with conjugated Horseradish Peroxidase (GE Healthcare, P0448) in the same solution used for blocking also at a final concentration of 1:2,000 for 1 hour at room temperature.

Finally, the unbound secondary antibody was removed by washing four times in PBST0.1% and Enhanced Cheminolumenescence reaction performed using commercial acquired reagents, which was allowed to progress less than 5 minutes. The membrane was then exposed to a radiographic film which was revealed using the adequate solutions.

2.1.22 Electrophoretic Mobility Assay (EMSA)

Two type of Electrophoretic mobility assays were performed; qualitative EMSA and semi-quantitative EMSA. The difference between the two is that in the former increasing amounts of recombinant TDP43 protein is added to the same concentration of UG_6 RNA radioactively labelled with Phosphorous 32 isotope (UG_6P^{32}) and, in the latter the same amount of recombinant protein is added to increasing amounts of UG_6P^{32}.

Synthetic UG_6 oligonucleotides were purchased from *Sigma* at a concentration of 100ng/µl and radioactive labelling was performed by mixing UG_6 oligonucleotides, T4 Polynucleotide Kinase (T4-PNK), PNK binding buffer and radioactive Gamma Adenosine Triphosphate (γ-ATP) in the proportions described in table 2.14 and incubating at 37ºC for 1 hour, with shaking. After labelling, the UG_6P^{32} RNA was precipitated by adding 2µl of $Mg(CH_3COO)_2$ and 80µl of absolute ethanol, incubating on dry ice for 20 minutes and centrifugation at 13,000rpm for 20 minutes. The RNA pellet was washed once with 70% ethanol and resuspended in 400µl of water, giving a final concentration of UG_6P^{32} of 0.5ng/µl.

Reagent	Volume
$UG_{6\ RNA}$ (100ng/µl)	2µl
10XPNK Buffer	2µl
T4-PNK	1µl
g-ATP	2µl
H_2O	13µl

Table 2.15: UG_6 oligonucleotide radioactive labelling mix

For the qualitative EMSA, increasing amounts of wild type or mutant recombinant TDP43 proteins, specifically 250ng, 500ng and 1mcg, were added to UG_6P^{32} (0.5ng/µl), 10xBinding Buffer (52mM Hepes Ph7.5, 10mM $MgCl_2$, 8mM $Mg(CH_3COO)_2$, 5.2mM DTT, 38% Glycerol) and H_2O, in quantities specified in table 2.15, and allowed to bind for 10 minutes at room temperature.

Reagent	Volume
Recombinant TDP43 protein	equivalent to 250ng, 500ng or 1mcg
10XBinding Buffer	2.5µl
UG_6P^{32} (0.5ng/µl)	1µl
H_2O	to final volume of 25µl

Table 2.16: Reagent mix for recombinant TDP43 and radioactively labelled UG_6 RNA binding reaction

After binding, 1µl of loading die was added to the GST-TDP43/UG_6P^{32} samples and the whole volumes loaded in a 5% polyacrylamide gel (refer to the next section) which had been subjected to pre-loading electrophoresis for 30 minutes in the cold room at 4ºC using 0.5xTBE as running buffer and electrophoresis performed at 60mA, 4ºC, which typically took two hours until completion (the gel was observed for the migration of the loading die down the gel). Following electrophoresis, the gel was transferred onto thick filter paper (*3MM*), covered with cling film and dried for 1 hour and 30 minutes in a gel dryer (*BioRad*). Subsequently, a phosphor membrane (Cyclone Phosphoimager Screen) was exposed to the gel enclosed in an autoradiograph film cassette and revealed in the membrane reader hardware and associated software (Cyclone Phosphoimager hardware and software).

In the semi-quantitative EMSA, 50ng of protein were incubated with exponentially increasing concentrations of UG_6P^{32} RNA (µM concentrations were typically the following: 2, 4, 8, 10,

20, 40, 60, 80, 100) and the same binding protocol, gel electrophoresis and phosphor membrane exposition and revelation methodologies followed. Quantification was performed by determining band intensities using *Image J* software, and dissociation plots for each semi-quantitative EMSA gel performed in *Prism/Graph-pad* software as "non-linear phase decay" in the same way as used in a previous publication (Bhardwaj et al. 2013). The dissociation constants (KD) were determined using the following equations:

$\boldsymbol{Xs} = -1/K \times \log [Ns/(100 \times (Ns - Y0))]$ and $\boldsymbol{KD} = (Y0 - Ns)e^{-(K \times Xs/2)} + Ns$

The values for K, Ns and Y0 were given for individual gels by the *Prism/Graph Pad* software upon plotting the dissociation curve.

2.1.23 Custom made gels for the experiments performed in the ICGEB in Trieste (Purification of recombinant protein, GST overlay/far western and EMSAs)

All gels for the protein purification and GST overlay/far western were custom made in the lab. Acrylamide/bis (40%) was made by mixing Acrylamide (38.7%); N,N'-Methylenebisacrylamide (1.33%) and water (60%). Polyacrylamide gels were then made at the desired densities by mixing the reagents as instructed in table 2.17 and stacking performed (table 2.18).

	Gel density (volumes)		
Reagent	**8%**	**10%**	**12%**
Acrylamide/bis (40%)	4ml	5ml	6ml
Tris HCl 1.5M pH 8.8	5ml	5ml	5ml
10% APS	200µl	200µl	200µl
TEMED	200µl	200µl	200µl
10% SDS	20µl	20µl	20µl
H2O (distilled)	11ml	10ml	9ml

Table 2.17: Reagent mix used to make acrylamide gels of different densities

Reagent	Volume
Acrylamide/bis (40%)	420µl
Tris HCl 0.5M pH 6.8	2.5ml
APS 10%	50µl
TEMED	10µl
10% SDS	50µl
H_2O (distilled)	2ml

Table 2.18: Reagent mix for stacking acrylamide gels

The EMSA gels were prevented to contact SDS and were 5% Acrylamide in 0.5xTBE, made as per the instructions in table 2.19.

Reagent	Volume
0.5xTBE	17.06ml
Acrylamide/bis (40%)	2.5ml
TEMED	200µl
10% APS	20µl

Table 2.19: Reagent Mix used to make EMSA gels

2.1.24 Immunofluorescence

MEFs were grown in 10cm dishes containing square (22mmx22mm) coverslips previously treated with 70% Ethanol. When confluent, the coverslips were transferred a 6-whell plates, washed three times with PBS and cells fixed by adding 4% paraformaldehyde (PFA) and incubating 10 minutes at room temperature. The 4% PFA was removed and the cells washed with PBS. The fixed MEFs were permeabilised by incubation with ice cold 100% methanol for 15 minutes at room temperature, followed by two PBS washes and blocking was performed with a solution of 5% Bovine Serum Albumine (BSA) in PBS for 1 hour. The MEFs were incubated with anti-TDP43 antibody (*Protein-tech*, 107-82-2-AP) at a concentration of 1:200 in the blocking solution at room temperature for 1 hour, followed by three washes with PBS and incubation with a fluorescent-tagged secondary antibody (Fluorochrome Alexa 488, anti-rabbit IgG; *Invitrogen*, A11008), also in blocking solution at a concentration of 1:500 in dark lighting conditions; from this point onwards, exposition of the coverslips to light was minimised as much as possible. After incubation with the secondary antibody, the cells were washed with PBS and the coverslips mounted onto slides using DAPI

mounting media and the edges of the coverslips sealed with nail varnish. After drying at room temperature the cells were stored at 4ºC and visualised by confocal microscopy (Zeiss LSM 710 NLO multi-photon microscope) within a space of two weeks before the immunofluorescence faded.

2.2 In vivo whole animal phenotyping tests

2.2.1 Body Weights and Brain Mass

Live animals were weighed in common scales inside a beaker without being anaesthetised and the scales took measurements for five seconds calculating an average body weight value to correct for the mouse moving inside the beaker. For brain mass measurements, mice were sacrificed according the Schedule 1 of the ASPA and the bodies weighed immediately after death, as were the brains, immediately after dissection.

2.2.2 Body composition

Body composition was measured by placing the non-anaesthetised animals in a cylindrical, transparent plastic tube and inside the EchoMRI™-100H machine (http://www.echomri.com/Body_Composition_Mice_2MHz.aspx) which using quantitative magnetic resonance technology calculated the body composition automatically. Three measurements were taken and the average values for each animal used for statistical analysis.

2.2.3 Modified SHIRPA procedure

The modified SHIRPA procedure constitutes of a battery of tests that assess mouse behaviour and is able to detect neuromuscular deficits. It starts by placing the mouse on a transparent Perspex viewing jar (14cm diameter, 18 cm high) on a grid floor and observe for tremors, body position and respiratory rate. Subsequently the mouse is transferred into the arena by placing a thin aluminium base to contain the mouse inside the viewing jar and dropping the animal into the Perspex arena (60cmx 37cm with walls 18cm high). Transfer arousal and gait

are observed, and the acoustic startle reflex tested with a click box that generates a 20KHz tone at 90 decibels (lasting approximately 2 seconds). The mouse is then suspended by its tail and observed for hind limb clasping. Grip strength is tested by allowing the animal to grab the grid mesh on top of the arena and body tone tested by palpation. The mouse is then "scruffed" and hind limb tone is tested, placed on a wire on the side of the cage thus testing the ability to perform the wire manoeuvre and negative geotaxis tested by placing the mouse in a vertical mesh grid and making it walk down. The self-righting reflex is tested by placing the mouse in a cylindrical tube and spinning it in the air. The main equipment used in the SHIRPA procedure can be found in figure 2.5 and further details of the modified SHIRPA procedure are available in the empress website, which is itself available through the MRC Harwell's website[5].

Figure 2.5: Main equipment used in the modified SHIRPA procedure
A mouse can be seen in the grid floor on the viewing jar and to the right the Perspex arena is visible.

2.2.4 Open Field

An open field automated set up was used. Mice are placed in square Perspex arenas (44cmx44cm) with opaque walls and flooring in a quiet dedicated room with controlled lighting conditions; light intensity is adjusted to be between 150-200 Lux. Central (8cmx8cm from the centre) and peripheral (11cm from the outer edge) areas are programmed into the

[5] http://har.mrc.ac.uk/services/mouse-husbandry-phenotyping/phenotyping/neurology

arenas in the *Ethovision Noldus* software (figure 2.6). Mice are allowed to move freely in the arenas for a test time of either 30 minutes (in the case of TDP-F210I-B6 and TDP-M323K-B6 cohorts) or 20 minutes (in the case of TDP-M323K-B6-DBA male cohort) and recorded by a ceiling camera, being tracked by the *Ethovision* software which calculates the time the mouse spends in the centre and periphery of the arena as well as the total distance moved during the test.

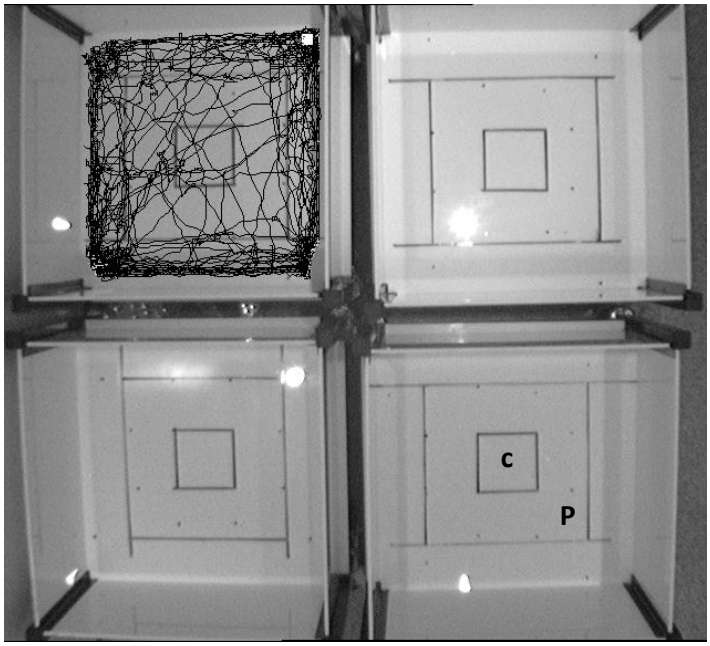

Figure 2.6: Automated open field set up
View of the open field arenas as captured by the overlying camera. The arenas are divided into two areas; centre (inner square, marked C) and periphery (outer area, marked P). The mice are tracked by the software and movement tracing can be seen in the arena at the top left hand side of the image, automatically calculating the total distance moved and the time each animal spent in the centre or periphery of the arena.

2.2.5 Light/Dark Box

The Light/Dark box test used the same automatic set up as the open field (camera and tracking software). The custom made box was divided into a roofless area exposed to light and a covered area with opaque walls and roof which remains in darkness; a small inlet that allows the mouse to travel between the two areas. The box is placed in a dedicated room with the lights turned on maximum intensity and individual mice tested. The mouse is placed in the area exposed to light, facing away from the access "door" to the dark area and left there for 5

minutes. The software tracts the mouse and outputs the length of time the animal spends in either the area of the box exposed to light and the dark area.

2.2.6 Grip Strength

Forelimb grip strength was measured by a grip strength meter (bioSeb). The mice allowed to grab the grid on the grip strength meter and traction is applied in the tail. The resistance the mouse applies on the grid determines the grip strength in grams (figure 2.7). Each mouse was submitted to two trials and an average of the two trials taken as the grip strength for each time point.

Figure 2.7: Grip strength test
A mouse can be seen gripping the grid while the operator applies traction on the tail.

2.2.7 Rotarod

Mice were tested in the rotarod (UgoBasilie, figure 2.8) on the accelerating mode, with the protocol used extending over 2 weeks. Initially, for 3 consecutive days, mice were subjected to a training regime consisting of 3 runs of five minutes each. The first practice run was at a fixed speed of 4 revolutions per minute (rpm) and the 2 subsequent runs had a starting speed of 4 rpm with constant acceleration to reach a maximum speed of 40 rpm. During the practice training, mice were placed back in the rolling drum each time it fell, thus conditioning the

animals to continue running until physical causes leads it to fall. The week following training, testing was performed during 3 consecutive days for the first time point and 2 consecutive days for the second time point. Each day of testing involved 3 runs in the accelerated mode and the time each individual animal fell recorded. The latency to fall for each animal was taken as the average time to fall resulting from all test runs between the 3 or 2 days of testing.

Figure 2.8: Rotarod used for testing *Tardbp* ENU mutant mice
The animals walk on the rolling drum and when falling, the time is automatically recorded.

2.2.8 Home cage automated-wheel running

Female mice were singly housed in cages with a running wheel (TSE Systems), being subject to a 12 hour light/dark cycle. For 7 days the wheel running activity is recorded by the *TSE Phenomaster* Software. Output measurements include total distance travelled, rotations, velocity, number of runs and time spent running. The validation of this system in identifying motor deficits can be found in (Mandillo et al. 2014). Each individual mouse had ad lib access to food and water. After 7 days the animals were removed and the data analysed. All equipment and software used is available commercially from TSE systems (http://www.tse-systems.com/).

2.2.9 Locotronic/Horizontal Ladder

Mice are placed at one end of the horizontal ladder (124 centimetres in length) where flashing lights are used to stimulate the animal to walk to the other end which is dark. As the mouse walks, each time its paws miss a ladder step, falling below its level, beams record it as errors and output it automatically as an excel file. Each animal was tested 3 times and the average number of errors from all three trials for each category (forelimb errors and hind limb errors) used for statistical analysis. The Locotronic system was commercially acquired from IntelliBio Innovation (http://www.intelli-bio.com/).

2.2.10 Contextual and Cued Fear Conditioning

The fear conditioning protocol used extended over 2 days and the equipment used was acquired from UgoBasile (http://www.ugobasile.com/). The test is recorded by a video camera and the mice tracked by *Any Maze* software, which gives an output of the time the mouse spends freezing automatically. On day 1, mice were conditioned to associate a mild foot shock with an acoustic sound. The animals are placed in the individual cage (figure 2.9A) and allowed to explore the new cage environment, during which baseline freezing is measured. Subsequently, three events of foot shocks coupled with an acoustic tone were performed. In the morning of day 2 the contextual fear conditioning was tested by placing the animals in the same cage (figure 2.9A) and the freezing times measured. In the afternoon, mice were placed in a different container (figure 2.9B) and baseline freezing measurements taken, following which freezing in response to the acoustic cue was measured by playing the tone three times at one minute intervals.

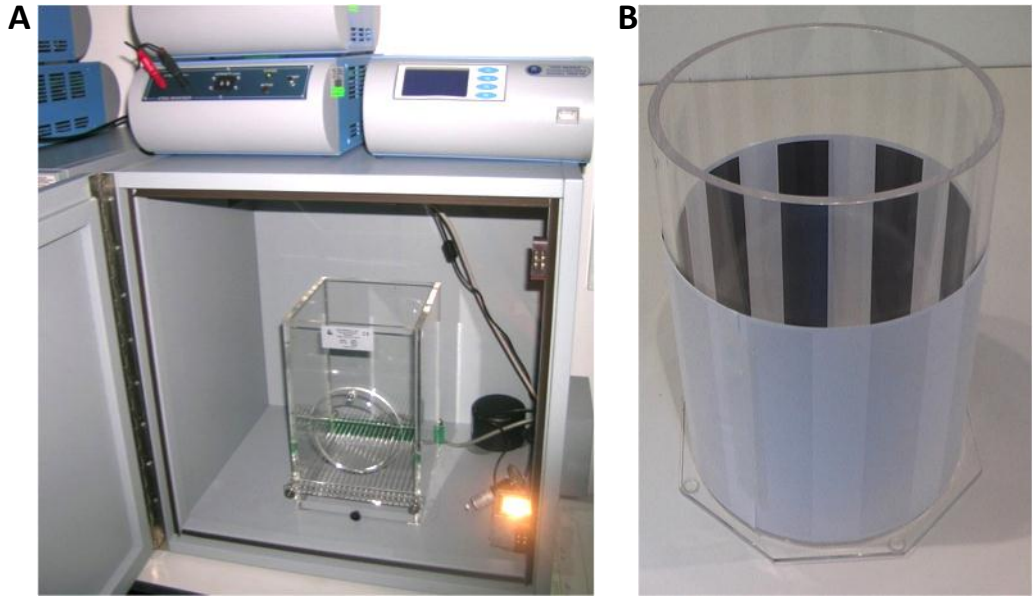

Figure 2.9: Equipment used in the contextual and cued fear conditioning
A) Fear conditioning cage placed in its isolation cupboard. A shock is delivered by the grid floor and the speaker (bottom right corner) delivers the acoustic cue. The following day contextual fear conditioning is tested by placing the mouse back on this cage.
B) Cued fear conditioning is performed by placing the mouse in the container shown, inside the same isolation cupboard seen in A and playing the acoustic cue.

The freezing times are calculated as percentage of test time and the following parameters are measured:

- "Tone-Noxius stimulus" (i.e. freezing time with each shock-cue conditioning event), which provides a measurement of whether mice from different genotypes have conditioned appropriately (day 1).
- "Context Baseline", which gives a measurement of the time the mice spend freezing in the new environment (cage) before being submitted to any foot shocks (day 1).
- "Context Freezing" (i.e. time freezing in context), which gives an output of the time the mice spend freezing in the same cage where conditioning took place on the second day, thus giving an output of the contextual memory (day 2).
- "Cue Baseline", which gives a measurement of the time the mice spend freezing in the new container before being submitted to the cue, thus enabling to correct the influence of the new environment in the measurement (percentage of time freezing, measured in day 2).

- "Cue freezing", which is the measurement of time freezing when mice are exposed to the cue only after conditioning (day 2).

After measurements are taken, the Context-(Context Baseline) and Cue-(Cue Baseline) were calculated.

2.2.11 Acoustic Startle Response and Pre-pulse inhibition (PPI)

Mice were placed in individual transparent perforated (to conduct sound better) plastic tubes and slotted into a platform in individual insulated cupboards with a camera and speaker (figure 2.10). The equipment and accompanying software was acquired commercially from Med Associates Inc (http://www.med-associates.com/).

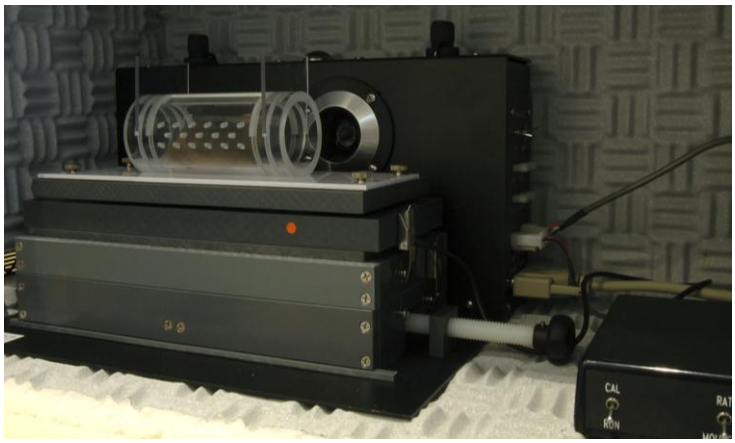

Figure 2.10: PPI and Startle equipment and set up
A mouse can be seen in its individual perforated plastic tube and slotted into the platform in the proximity of the speaker and camera. Notice the insulation material on the walls and floor of each individual cupboard.

Once the mice are in the platforms and start is selected in the software the cycle commences automatically and consists of an initial acclimatisation period without any acoustic stimuli followed by a pre-pulse inhibition session of 10 different trial types in which the reflex movement ("small jump") of the animal inside the chamber is measured. Further specifications can be found in the manufacturer's website (http://www.med-associates.com/product/acoustic-startle-reflex-starter-package-for-rat-or-mouse/). The data outputted included the startle reflex to an acoustic sound of 110 decibels and the startle reflex

when the acoustic startle stimulus is preceded by pulses of white noise with amplitudes of 65, 70 and 75 decibels.

2.2.12 Circadian phenotyping

Circadian phenotyping was performed in an automated system which is described in detail in (Banks & Nolan 2011).

Female animals were singly housed on a 12 hour light/dark cycle initially and wheel running activity automatically measured by the software. Over a period of 7 days circadian activity is monitored not only on the 12 hour light/dark cycle but also in conditions of constant light and constant darkness.

2.3 Statistical analysis and graphical representations

All statistical analysis for the Immunoblots, EMSA, and RT-PCR splicing assays as well as the behavioural characterisation was performed in proprietary *Spss* statistical software, except for the Chi-square test which was performed using an online freely available facility (http://graphpad.com/quickcalcs/chisquared1.cfm) and the Benjamini Hoechber correction for multiple testing also performed using an online resource (http://www.sdmproject.com/utilities/?show=FDR). All graphs were plotted on *Prism/Graph-pad* software.

The Chi-square test was performed to assess whether genetic inheritance of mutant alleles deviated from the expected Mendelian ratios and the Anova test performed for quantitative molecular experiments and behavioural characterisation. In the results section reference of the statistical test performed for each experiment is always included.

Statistical analysis of the RNA-seq data was performed by Pietro Fratta in collaboration with the group of Dr Vincent Plagnol as specified in the RNA-seq section of this chapter.

Product (including reagents and labware)	Manufacturer	catalogue number
Protease inhibitor cocktail tablets	Roche	11697498001
Phosphatase inhibitor cocktail tablets	Roche	4906845005
Trizol	Quiagen	15596018
Bradford Reagent	Sigma	B6916-500ML
Nupage LDS Sample Buffer (4X)	Life Technologies	NP0007
Sample Reducing agent (10X)	Life Technologies	NP0004
Nupage 10% Bis-Tris Gel (pre-cast gel)	Life Technologies	NP0316
Nupage MOPS SDS Running buffer	Life Technologies	NP0001
SeeBluePlus2 pre-stained protein standards	Life Technologies	LC5925
Nupage Antioxidant	Life Technologies	NP0005
iBlot gel Transfer Stacks (nitrocellulose membrane)	Life Technologies	IB301002
Skim milk powder	Sigma	70166-500G
Agarose	Sigma	A9539-500G
Tween 20	Fisher	10113103
Nusieve 3:1 Agarose	Lonza	50090
Triton X-100	Fisher	10102913
Broad range Protein standards	New England Biolabs	P7708S
NE-PER Nuclear/Cytoplasmics Protein extraction kit	Thermo Scientific	78833
Isopropanol	Fisher	11358461
Methanol	Fisher	10767665
Ethanol	Fisher	10610813
Proteinase K	Roche	3115844001
Trypsin-EDTA	Life Technologies	25300-054
Dulbecco's Complete Media	Life Technologies	61965-026
Foetal Bovine Serum	Life Technologies	10270-098
Ultra Glutamine II	Lonza	BE04-684E
Amaxa Nucleofector Kit for MEFs	Lonza	VPL-1004
BSA	New England Biolabs	B9000S
Penicillin/Strptomycin	Life Technologies	15140-122
100 base pair nucleic acid ladder	New England Biolabs	N3231S
Nuclease free water	Qiagen	129115
Ethidium Bromide	Sigma	E1510
DMSO	Sigma	D2650
Gelatine	Sigma	91393
Phosphate Buffered Saline (PBS)	Life Technologies	14190-094
T4 PNK (inclides 10x PNK buffer)	New England Biolabs	M0236S
HiMark pre-stained protein standards	Life Technologies	LC5699
See Blue Plus Two pre-stained protein standards	Life Technologies	LC5925
Gel Red	Biotium	41003-1
BCA "DC protein Assay" kit	Biorad	500-0111
Acrylamide	Gerbu	1138
N,N'-Methylenebisacrylamide Water	Gerbu	1138
TEMED	Fisher	BP150-20
APS	Fisher	AC40116-0250
Amersham ECL Western Blotting Detection Reagents	GE Healthcare	RPN2106
Low fluorescence PVDF immunoblotting membrane	Millipore	IPFL00010

Table 2.20: Manufacturers and catalogue numbers for reagents used in molecular experiments

Chapter 3.

Results: Inheritance and lethality of the *Tardbp*^{M323K} and *Tardbp*^{F210I} alleles

Tardbp mutants were produced by ENU mutagenesis (C57BL/6J were injected with ENU and crossed with DBA females) and after the mutations were identified from the DNA archive and the animals re-derived with matching sperm through IVF, the offspring were mated with C57BL/6J wild type mice. Mating the re-derived ENU mutants to C57BL/6J wild type is referred to as backcrossing and aims to segregate the mutation of interest from additional mutations carried by the G1 animals (G1 animals are the direct offspring of the mating between a ENU-injected male mouse and a wild type female). The C57BL/6J mouse strain was selected given that most mouse models of TDP43 proteinopathies published are in this particular strain.

In general, however, when backcrossing animals with mutations achieved by ENU mutagenesis, it is preferable to use a different strain for backcrossing in case doubts arise regarding linkage disequilibrium between the mutation of interest and additional mutations in neighbouring genetic regions. When the mutagenized animals and animals used for backcrossing are of different strains, assessment of single nucleotide polymorphisms carried by the offspring of the different matings can be used to assess whether linkage disequilibrium between two mutations, is in fact, present

Tardbp mutants in the offspring of backcross matings were selected by genotyping, given that a reverse genetics approach was followed and therefore the mutations of interest had been known from the outset (refer to relevant section in the introduction for a more detailed description of ENU mutagenesis).

Animals were backcrossed at least five generations before embryos, foetuses[6] or adult mice were included in the experimental sample, thus ensuring that other potential functional

[6] The transition between embryonic stage and foetal stage occurs between 14.5dpc and 15.5dpc. In this thesis, the convention of describing gestating mice up to and including 14.5dpc as embryos and from 15.5dpc until birth as foetuses has been adopted.

mutations were segregated appropriately from the experimental cohorts and embryonic or foetal samples.

At present, the *Tardbp*^{M323K} and *Tardbp*^{F210I} mouse colonies are on congenic C57BL/6J background (i.e. the mutant mice have been backcrossed over ten generations since the mice were re-derived).

3.1 Homozygosity for *Tardbp*^{F210I} or *Tardbp*^{M323K} leads to early lethality in laboratory mice of C57BL/6J background

The inheritance of the *Tardbp*^{F210I} allele has been previously described and leads to prenatal lethality in homozygosity in mice which have been backcrossed at least four generations onto C57BL/6J background, with foetuses dying shortly after 15.5dpc, whilst *Tardbp*$^{F210I/+}$ are born and develop up to weaning age at Mendelian ratios (Ricketts 2012).

Assessment of the inheritance pattern of the *Tardbp*^{M323K} allele was performed by analysing the offspring from matings between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ animals and matings between *Tardbp*$^{M323K/+}$ mice; all mice used in the matings had been backcrossed at least five generations onto the C57BL/6J background.

The offspring *Tardbp*$^{M323K/+}$ backcrosses to *Tardbp*$^{+/+}$ mice revealed that the *Tardbp*^{M323K} allele was inherited in a Mendelian pattern at weaning (table 3.1, refer to table for statistical analysis).

Genotype [*Tardbp*$^{+/+}$ (C57BL/6Jx *Tardbp*$^{M323K/+}$ (C57BL/6J)]	Weaning
Tardbp$^{+/+}$ (C57BL/6Jbackground)	67
Tardbp$^{M323K/+}$ (C57BL/6Jbackground)	54
Total	121
x^2 (Mendelian inheritance)	p=0.2373

Table 3.1: Inheritance pattern of *Tardbp*^{M323K} allele in backcross matings
When wild type animals, on the C57BL/6J genetic background are mated with *Tardbp*$^{M323K/+}$ animals, the wild type and mutant *Tardbp*^{M323K} alleles are inherited as expected according to a Mendelian pattern. Mendelian inheritance is statistically demonstrated by p=0.237 obtained in the Chi-square test (x^2) comparing the observed number of animals with the number of animals expected for each genotype according to Mendelian prediction.

However, when *Tardbp*$^{M323K/+}$ mice were mated with *Tardbp*$^{M323K/+}$ mice, no homozygous animals were found alive at weaning age (table 3.2, refer to table for statistical analysis).

Genotype [*Tardbp*$^{M323K/+}$ (C57BL/6J) x *Tardbp*$^{M323K/+}$ (C57BL/6J)]	14.5dpc	P1	P5	Weaning
Tardbp$^{+/+}$ (C57BL/6J background)	27	3	2	23
Tardbp$^{M323K/+}$ (C57BL/6J background)	53	6	9	25
Tardbp$^{M323K/M323K}$ (C57BL/6J background)	38	4	**0**	**0**
Subtotal	118	13	11	48
x^2 (Mendelian inheritance)	p=0.1948	p=0.8910	p=0.0750	**p<0.0001**

Table 3.2:Inheritance pattern of the *Tardbp*^{M323K} allele in heterozygous matings (intercrosses) of mice in the C57BL/6J inbred background
When *Tardbp*$^{M323K/+}$ mice are intercrossed, *Tardbp*$^{M323K/M323K}$ animals are not found alive at weaning age (p<0.0001, chi-square test). At 14.5 days post conception (14.5dpc), embryos are found alive and inherit the *Tardbp*^{M323K} allele according to the expected Mendelian pattern (p=0.1948, Chi-square test). At postnatal day 1 *Tardbp*$^{M323K/M323K}$ pups are found alive and, with the caveat of small numbers having been assessed, the inheritance pattern of the *Tardbp*^{M323K} allele is suggested to be Mendelian (p=0.8910, chi-square test). No animals have been found alive at postnatal day 5 and despite the inheritance at postnatal day 5 falling below statistical significance, which is most likely due to the small sample size (p=0.075, Chi-square test), the observation is suggestive that death of homozygous animals is likely to occur between postnatal days 1 and 5.

Nevertheless, *Tardbp*$^{M323K/M323K}$ embryos were found to be present in their expected Mendelian ratios at 14.5dpc and, at postnatal day 1, *Tardbp*$^{M323K/M323K}$ pups were also found alive (table 3.2).

Moreover, all *Tardbp*$^{M323K/M323K}$ animals observed at P1 (n=13) were macroscopically indistinguishable from its littermates and all were able to feed as milk was observed in the stomach of neonates, suggesting that, despite the small sample, the cause of death is not immediate respiratory failure (in which case animals would not be able to suckle) or hypoglycaemia secondary to inability to feed. At postnatal day 5 (n=11), no *Tardbp*$^{M323K/M323K}$

pups were found alive (table 3.2, refer to table for statistical analysis), suggesting that homozygosity for the *Tardbp*^{M323K} allele leads to mortality between postnatal days 1 and 5, which needs to be interpreted with the necessary caution given the suboptimal number of animals in the sample (n=13 for postnatal day 1 and n=11 for postnatal day 5).

It was therefore concluded that, on the C57BL/6J background, homozygosity for the *Tardbp*^{M323K} allele leads to perinatal mortality.

Aiming to determine the developmental consequences of homozygosity for the *Tardbp*^{M323K} allele, full body Computer Tomography (CT) scanning was performed in foetuses.

To describe the process briefly, foetuses were harvested from *Tardbp*$^{M323K/+}$ intercrosses at 16.5dpc and 18.5dpc, fixed in 4% paraformaldehyde, exposed to contrast solution by immersion and washed with Phosphate Buffered Saline (PBS). Subsequently the foetuses were embedded in 1% agarose in a transparent plastic cylindrical tubes and full body Computer Tomography (CT) scanning performed with the micro CT scan at the MRC Harwell. These particular ages were selected given neonatal pup size exceeds the size limits of the micro CT scanner, which therefore precluded CT scanning from being performed at postnatal time points.

The CT scans enabled good quality images to be obtained with all organs in situ and eliminated some artefacts associated with dissection and mounting histological preparations. However, it must be remarked that, as with any other imaging technique, CT scan images are also vulnerable to artefact, which has to be considered when analysing the images.

Preliminary analysis did not reveal any obvious abnormalities in organogenesis, but have raised interest in more detailed assessment of the angle of the heart in *Tardbp*$^{M323K/M323K}$ and brain ventricle sizes in *Tardbp*$^{F210I/F210I}$ (figure 3.1, please refer to figure legend for further information). The images have been sent to collaborators at the Institute of Child Health who

will perform expert phenotypic assessment of the foetuses at 16.5dpc and 18.5dpc; sample images, together with a succinct description of the future work can be found in figure 3.1.

In addition to the aim of determining the developmental consequences of mutant TDP43-M3232K biological dysfunction, it is hoped that a structural abnormality may be identified that will be informative regarding the cause of perinatal mortality in *Tardbp*$^{M323K/M323K}$ and prenatal mortality in *Tardbp*$^{F210I/F210I}$.

3.2 A hybrid C57BL/6J–DBA background influences survival of *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{M323K/M323K}$ laboratory mice

Interestingly, when *Tardbp*$^{M323K/+}$ mutants in congenic C57BL/6J background were mated with *Tardbp*$^{M323K/+}$ mutants in the DBA background (which had been backcrossed at least five generations onto DBA), hence leading to offspring of hybrid C57BL/6J-DBA genetic background, *Tardbp*$^{M323K/M32K}$ animals were found to be viable at weaning age and followed a Mendelian pattern of inheritance (table 3.3, refer to table for statistical analysis).

Genotype [(*Tardbp*$^{M323K/+}$ (C57BL/6J) X *Tardbp*$^{M323K/+}$ (DBA)]	**Weaning**
***Tardbp*$^{+/+}$ (C57BL/6J-DBA hybrid background)**	39
***Tardbp*$^{M323K/+}$ (C57BL/6J-DBA hybrid background)**	66
***Tardbp*$^{M323K/M323K}$ (C57BL/6J-DBA hybrid background)**	30
Total	135
x^2 (Mendelian inheritance)	p=0.5308

Table 3.3: Inheritance pattern of the *Tardbp*^{M323K} allele in matings of heterozygous animals of two genetics backgrounds, C57BL/6J and DBA.
Mating *Tardbp*$^{M323K/+}$ animals in the C57BL/6J background with *Tardbp*$^{M323K/+}$ in the DBA background leads to Mendelian inheritance of the *Tardbp*^{M323K} allele at weaning age (p=0.5308, Chi-square test). Thus, in contrast to what is observed in the C57BL/6J background, *Tardbp*$^{M323K/M323K}$ animals are viable at weaning age in the hybrid C57BL6/J-DBA background.

Consistent with the rescue of perinatal lethality observed in *Tardbp*$^{M323K/M323K}$ animals by a hybrid C57BL/6J–DBA genetic background, partial rescue is also observed in *Tardbp*$^{F210I/F210I}$ animals in a C57BL/6J–DBA genetic background.

As already described, in the C57BL/6J background, *Tardbp*$^{F210I/F210I}$ die *in utero* at around 15.5dpc, without any live animals being observed at birth (Ricketts 2012). However, when *Tardbp*$^{F210I/+}$ animals in congenic C57BL/6J background are mated with *Tardbp*$^{F210I/+}$ animals in the DBA background (which had been backcrossed at least five generations onto DBA), *Tardbp*$^{F210I/F210I}$ foetuses are still found alive at 18.5dpc, following a Mendelian pattern of inheritance (table 3.4, refer to table for statistical analysis). However, a trend is observed towards reduced numbers of *Tardbp*$^{F210I/F210I}$ foetuses at 18.5dpc and therefore, it is possible that further breeding will reveal that *Tardbp*^{F210I} inheritance falls below Mendelian ratios at 18.5dpc (table 3.4).

Genotype [*Tardbp*$^{F210I/+}$ (C57BL/6J) x *Tardbp*$^{F210I/+}$ (DBA)]	**16.5dpc**	**18.5dpc**	**Weaning**
***Tardbp*$^{+/+}$ (C57BL/6J-DBA hybrid background)**	13	23	15
***Tardbp*$^{F210I/+}$ (C57BL/6J-DBA hybrid background)**	40	30	29
***Tardbp*$^{F210I/F210I}$ (C57BL/6J-DBA hybrid background)**	20	11	**0**
Subtotal	73	64	44
x^2 (Mendelian inheritance)	p=0.3654	p=0.0930	**p=0.0006**

Table 3.4: Inheritance pattern of the *Tardbp*^{F210I} allele in matings of heterozygous animals of two genetics backgrounds, C57BL/6Jand DBA.
In contrast to *Tardbp*$^{F210I/F210I}$ animals in the C57BL/6J background which die at around 15.5dpc, in the C57BL/6J–DBA hybrid background foetuses are found alive at 18.5dpc with the *Tardbp*^{F210I} allele being inherited in a Mendelian pattern at both 16.5dpc (p=0.365, Chi-square test) and 18.5dpc (p=0.0930, Chi-square test). Despite the reasonable sample size at 18.5dpc (n=64), it is possible that further breeding will lead to a significant p value on Chi-square statistical test and that, in fact, the *Tardbp*^{F210I} allele is not inherited in a Mendelian pattern at 18.5dpc. Even in the hybrid background, no animals have been found alive at weaning age (p=0.0006, Chi-square test).
The partial rescue of *Tardbp*$^{F210I/F210I}$ lethality in the hybrid background reinforces the suggestion that genetic background of laboratory mouse strains interacts with *Tardbp* mutations, given that a rescue effect is also observed in *Tardbp*^{M323K} mutants.

Analogously to *Tardbp*^{M323K} foetuses in the C57BL/6J background, *Tardbp*^{F210I} foetuses in the hybrid C57BL/6J– DBA background were harvested, CT scan imaging performed (figure 3.1) and the images sent to collaborators at the Institute of Child Health so that we can analyse the data together with researchers with expert knowledge in embryonic and foetal development.

It can therefore be concluded that the genetic background of laboratory mouse inbred strains interacts with *Tardbp*^{F210I} and *Tardbp*^{M323K} and influences undetermined processes that enable

partial rescue of prenatal lethality in *Tardbp*$^{F210I/F210I}$ and complete rescue of perinatal lethality in *Tardbp*$^{M323K/M323K}$.

The influence of *Tardbp*^{M323K} allele homozygosity in the lifespan/longevity of laboratory mice in hybrid C57BL/6J– DBA genetic background remains, at present, undetermined. However, given that adult *Tardbp*$^{M323K/M323K}$ animals are viable we have proceeded with a breeding programme to establish testing cohorts of *Tardbp*^{M323K} mice in C57BL/6J–DBA hybrid background which will enable full assessment of the physiological consequences of M323K allele homozygosity at the whole animal level, including a potential role in neuronal dysfunction and neurodegeneration.

The inheritance patterns of *Tardbp*^{M323K} and *Tardbp*^{F210I} alleles have now been described both in the C57BL/6J and C57BL/6J-DBA hybrid background. Additionally, the ongoing embryo phenotyping work aimed at dissecting the developmental consequences for homozygozity of these alleles has also been detailed.

Further genetic breeding was performed, with the objective of assessing how different mutant alleles interact *in vivo*, including whether compound heterozygous mutants, without a wild type *Tardbp* allele, are viable.

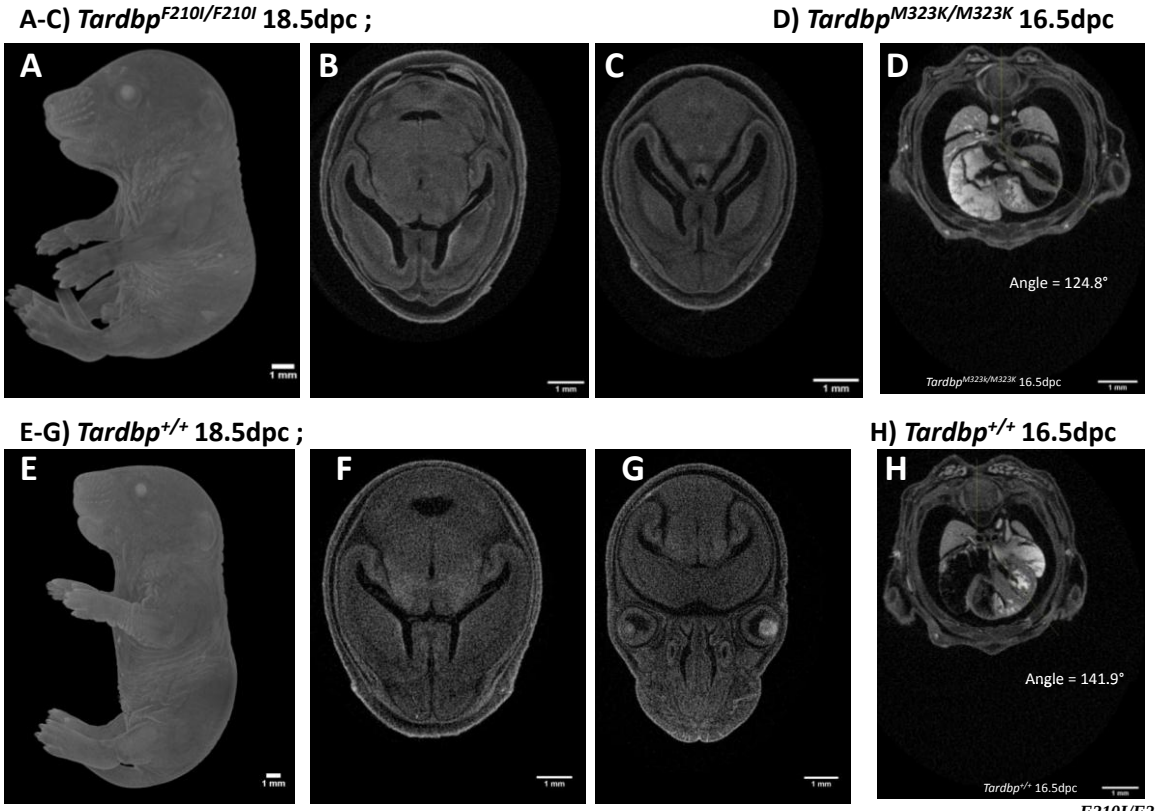

Figure 3.1: Representative Computer Tomography (CT) images of *Tardbp*$^{F210I/F210I}$, *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$ whole-foetuses.

Tardbp^{F210I} and *Tardbp*^{M323K} mutant foetuses, together with wild type littermate controls were harvested at 16.5 and 18.5dpc, fixed and immersed in contrast, after which CT scanning was performed.

Panels A and E show representative 3 dimensional body images from *Tardbp*$^{+/+}$ and *Tardbp*$^{F210I/F210I}$ foetuses.

Panels B, C and F show representative images of horizontal sections through the head for *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{+/+}$ foetuses and panel G a coronal plan through a *Tardbp*$^{+/+}$ foetal head.

Analysis of the images will include measuring brain ventricle sizes and cortical parenchyma thickness to look for hydrocephalus or microcephaly which has been reported in transgenic mice with TDP43 knockdown (Yang et al. 2014).

Panels D and H show representative images of a transverse sections through the thorax of *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/M323K}$ foetus harvested at 16.5dpc. An empirical observation is that the angle of the heart is different in *Tardbp*$^{M323K/M323K}$ foetuses and controls, which will be further analysed in a systematic manner.

The analysis will include measurement of the ventricles, cortical parenchyma and heart angle as well as exhaustive structural assessment of all other internal organs within the resolution sensitivity of the CT images.

3.3 Breeding compound *Tardbp* ENU: additional tools to study TDP43 biology *in vivo*

Viable compound heterozygous for *Tardbp* ENU mutations can constitute an informative tool for further study of the effects of mutations on TDP43's biological functions. Its use can be illustrated with matings between $Tardbp^{Q101X/+}$ and $Tardbp^{F210I/+}$ to generate $Tardbp^{Q101X/F210I}$, as described below and recapitulating some of the information given in the introduction.

The Q101X mutation in TDP43 functions as a null allele and in homozygosity ($Tardbp^{Q101X/Q101X}$) leads to very early embryonic lethality. Additionally, in $Tardbp^{Q101X/+}$, a partial "loss of normal splicing function" in the target genes *Sort1* and *Pdp1* is observed, in the context of normal TDP43 protein levels (when compared to $Tardbp^{+/+}$ controls) but reduction in $Tardbp^{Q101X}$ transcript levels and upregulation of the *Tardbp* wild type transcript (Ricketts et al. 2014).

The data on the study of the $Tardbp^{Q101X}$ mutant line suggests that wild type mouse TDP43 is able to autoregulate its own levels (Ricketts 2012) and is consistent with the proposed models, described in more detail in the introduction, which require TDP43 to directly bind the 3' UTR of its own transcript (Ayala et al. 2011; Polymenidou et al. 2011; Avendaño-Vázquez et al. 2012). However, given that normal protein levels are observed in $Tardbp^{Q101X/+}$ the underlying mechanism leading to the splicing phenotype remains obscure.

The F210I mutation in TDP43 leads to dose –dependent "loss of normal splicing function" in target genes, including *Sort1*, *Dnajc-5* and *Pdp1*, in the context of similar TDP43 levels across all genotypes ($Tardbp^{+/+}$, $Tardbp^{F210I/+}$ and $Tardbp^{F210I/F210I}$) and prenatal lethality in homozygosity (Ricketts 2012). Moreover, mutant TDP43-F210I protein has been shown to have reduced affinity for RNA when compared to wild type TDP43, which is proposed to consist of the underlying biological mechanism leading to the splicing phenotype (Ricketts

2012). Therefore, it is reasonable to expect that the reduced affinity of TDP43-F210I for RNA would inhibit it from binding the 3' UTR of *Tardbp* transcripts, thus disabling the autoregulation process, assuming that the two autoregulation models proposed to date are valid. Thus, *Tardbp*$^{Q101X/F210I}$ would provide a unique *in vivo* system in which to study the ability of the F210I mutant protein to autoregulate its own levels by upregulating *Tardbp*^{F210I} transcript in order to compensate for the degradation of null *Tardbp*^{Q101X} transcripts (most likely by non-sense mediated decay), mirroring what was observed in tissue of *Tardbp*$^{Q101X/+}$ animals, therefore testing the validity of the existing autoregulation models.

The same hypothesis regarding the opportunity to further study TDP43 dysfunction in the context of multiple endogenous mutations, including autoregulation, can be applied to breeding *Tardbp*^{M323K} with *Tardbp*^{Q101X} mice.

Furthermore, as it will be described in chapter 4, the M323K mutation leads to "augmented alternative exon selection" and thus a splicing phenotype opposed to the one observed in the F210I mutation and therefore, it would be valuable to assess whether the splicing phenotype is rescued in compound heterozygous *Tardbp*$^{F210I/M323K}$ mice.

Tardbp$^{F210I/+}$ animals were mated with *Tardbp*$^{Q101/+}$ mice (both mutant lines had been backcrossed at least five generations into the C57BL/9J background) by Thomas Ricketts, and the results included in his PhD thesis. *Tardbp*$^{Q101X/F210I}$ died prenatally and it was not possible to extract protein from the embryos to test the autoregulation properties of the mutant F210I protein (Ricketts 2012).

3.3.1 Breeding *Tardbp*$^{M323K/Q101}$ mice

Tardbp$^{M323K/+}$ males were bred with *Tardbp*$^{Q101X/+}$ females (both on congenic C57BL/6J background) and the offspring genotyped at weaning age. *Tardbp*$^{M323K/Q101X}$ animals are viable and found to survive to weaning age in Mendelian ratios (table 3.5, refer to table for statistical analysis)

Genotype [Mating: *Tardbp*$^{Q101X/+}$ (C57BL/6J) x *Tardbp*$^{M323K/+}$ (C57BL/6J)]	**Weaning**
***Tardbp*$^{+/+}$**	9
***Tardbp*$^{M323K/+}$**	9
***Tardbp*$^{Q101X/+}$**	10
***Tardbp*$^{M323K/Q101}$**	5
Total	33
x^2 (Mendelian)	p=0.6176

Table 3.5: inheritance of *Tardbp* mutant alleles upon mating *Tardbp*$^{Q101X/+}$ and *Tardbp*$^{M323K/+}$ animals
Tardbp alleles, both mutant (Q101X and M323K) and wild type, are inherited in Mendelian ratios when genotyping is performed in offspring mice at weaning age (p=0.6176).

Tardbp$^{M323K/Q101X}$ will therefore provide a unique tool for further investigating TDP43 biological function, including the ability of the M323K mutant protein to autoregulate TDP43 protein levels. Specific experiments planned with these compound heterozygous will be described in the discussion after the results of the molecular and *in vivo* characterisations are given.

3.3.2 Breeding *Tardbp*$^{F210I/M323K}$ mice

Breeding between *Tardbp*$^{F210I/+}$ females and *Tardbp*$^{M323K/+}$ males was also performed (both mutant lines were on congenic C57BL/6J background) and both embryos at 14.5dpc and mice at weaning age were genotyped. *Tardbp*$^{M323K/F210I}$ embryos were found in Mendelian ratios at 14.5dpc but at very reduced numbers at weaning age, falling significantly below the expected Mendelian ratios (table 3.6, refer to table for statistical analysis).

Genotype [Mating: *Tardbp*$^{F210I/+}$(C57BL/6J) x *Tardbp*$^{M323K/+}$(C57BL/6J)]	14.5dpc	Weaning
Tardbp$^{+/+}$	12	23
Tardbp$^{F210I/+}$	8	24
Tardbp$^{M323K/+}$	8	15
Tardbp$^{M323K/F210I}$	5	**6**
Total	33	68
x^2 (Mendelian)	p=0.3916	**p=0.0063**

Table 3.6: *Tardbp* allele inheritance in matings between *Tardbp*$^{F210I/+}$ and *Tardbp*$^{M323K/+}$ mice (congenic C57BL6/J background)
All *Tardbp* alleles are inherited in a Mendelian pattern at 14.5dpc thus leading to the expected Mendelian ratio of *Tardbp*$^{M323K/F210I}$ embryos (p=0.3916, Chi-square test). However, at weaning age only 6 *Tardbp*$^{M323K/F210I}$ animals were found alive, a significant reduction from the Mendelian expected 17 animals (p=0.0063, Chi-square test).

The data suggests that many *Tardbp*$^{M323K/F210I}$ animals die prematurely, between 14.5dpc and weaning age, and the low number of live mice at weaning age (6 observed to be alive out of 17 expected according to Mendelian ratios) may disable the establishment of adult testing cohorts. Nevertheless, the Mendelian ratios observed at 14.5dpc enable relevant molecular experiments to be performed and these will be timely discussed in chapter 4 and the discussion chapter.

3.4 Breeding *Tardbp* ENU mutants with BAC *TARDBP* transgenic mice: demonstrating that single point mutations in TDP43 can lead to multiple functional consequences

In addition to the ENU mutants, *TARDBP* BAC transgenic animals recently became available to our group, specifically animals over-expressing (in approximately three fold) human wild type full length *TARDBP* or mutant full length human *TARDBP*^{G348C} (Swarup et al. 2011). The transgenic mice were a generous gift from Professor Jean-Pierre Julien.

The wild type and G348C *TARDBP* BAC transgenic animals have been shown to develop histopathological hallmarks of TDP43 proteinopathies, cognitive and motor deficits from 7 months of age onwards (Swarup et al. 2011) and have impaired recovery from sciatic nerve crush injury (Swarup et al. 2012).

Breeding *TARDBP* BAC transgenics with *Tardbp* ENU mutants would enable further study on how the endogenous *Tardbp* alleles interact and possibly allow inferences regarding specific aspects of TDP43 dysfunction and neurodegeneration to be made. Moreover, the *TARDBP* transgene may be expected to rescue the lethality in $Tardbp^{Q101X/Q101X}$ and $Tardbp^{F210I/F210I}$ which is suggested to occur in these mutant animals due to TDP43 loss of function during gestation, and should therefore be overcome by *TARDBP* transgene expression.

Only genotyping of offspring from matings of $TARDBP^{G348C}$ BAC transgenic animals and $Tardbp^{F210I}$ ENU mutants has been achieved at present, due limitations in time and resources, with the initial goal of assessing whether the $TARDBP^{G348C}$ allele rescues lethality in $Tardbp^{F210I/F210I}$. Additionally, if $Tardbp^{F210I/F210I}$ mice with the $TARDBP^{G348C}$ transgene developed a phenotype different from the $TARDBP^{G348C}$ BAC transgenics published by (Swarup et al. 2011), further inferences and hypotheses could also be conceptualised regarding the toxicity associated with TDP43 dysfunction.

$Tardbp^{F210I/+}$ animals were mated with hemizygous (i.e. with a single copy of the transgene) $TARDBP^{G348C}$ BAC transgenic mice (both on C57BL/6J background) and the offspring genotyped for the presence of the transgene and the F210I mutation. Strikingly, presence of the $TARDBP^{G348C}$ transgene did not rescue lethality in $Tardbp^{F210I/F210I}$ (table 3.7, refer to table for statistical analysis). Failure of the $TARDBP^{G348C}$ transgene in rescuing the lethality of $Tardbp^{F210I}$ allele homozygosity suggests that the F210I mutation does not cause premature lethality as a simple consequence of loss of function, which would otherwise be expected to be compensated by the transgene expression, with the caveat that rescue has not yet been assessed with the wild type *TARDBP* transgene.

Genotype [Mating: (*TARDBP*^{G349C}*Tardbp*$^{F210I/+}$) x (*Tardbp*$^{F210I/+}$)]	Weaning	Mendelian Expected
Tardbp$^{+/+}$ (2 wild type mouse alleles and without human transgene)	15	12.50%
Tardbp$^{F210I/+}$(heterozygous for the mouse F210I mutation and without the human transgene)	17	25%
Tardbp$^{F210I/F210I}$(homozygous for the mouse F210I mutation and without the human transgene)	**0**	**12.50%**
TARDBP^{G349C}*Tardbp*$^{+/+}$ (2 wild type mouse Tardbp alleles and without human transgene)	9	12.50%
TARDBP^{G349C}*Tardbp*$^{F210I/+}$ (heterozygous for the mouse F210I mutation and with the human transgene)	16	25%
TARDBP^{G349C}*Tardbp*$^{F210I/F210I}$ (homozygous for the mouse F210I mutation and with the human transgene)	**0**	**12.50%**
Total	57	100.00%
x^2 (Mendelian)	**p=0.0002**	

Table 3.7: Inheritance of *Tardbp*^{F210I} allele in mice with *TARDBP*^{G348C} transgene
The presence of the mutant human transgene does not rescue homozygosity of the F210I mutation. The expected Mendelian ratios for each genotype are given. Remarkably, the presence of the human transgene does not rescue the already described lethality caused by homozygosity for the F210I mutation (p=0.0002, Chi-Square test), at least to the extent of allowing animals to survive until weaning age. Failure of the human transgene in rescuing the lethality in *Tardbp*$^{F210I/F210I}$ suggests that the F210I mutation leads to premature death through a mechanism which is not simply associated with loss of function, which would be compensated by *TARDBP*^{G348C} transgene expression.

Thus, the breeding data demonstrates that the functional consequences of single missense mutations in TDP43 can be multiple (e.g. “loss of splicing function” in the F210I protein which is not lethal and possibly gain of a toxic function leading to premature death in *Tardbp*$^{F210I/F210I}$) and therefore over-expression or conditional deletion models of TDP43 proteinopathies fall short of the complexity necessary in replicating the effects of TDP43 point mutations.

Further breeding between the *TARDBP* BAC transgenics (including the wild type *TARDBP* BAC transgenic) and *Tardbp* ENU mutants (F210I, Q101X and M323K) is therefore planned to assess the inheritance pattern in the offspring and, if rescue is observed, proceed to further molecular and possibly behavioural experiments.

3.5 Summary of the inheritance and lethality of the *Tardbp*^{M323K} and *Tardbp*^{F210I} alleles

Detailed description of the inheritance of the *Tardbp*^{M323K} and *Tardbp*^{F210I} alleles is given above.

In previous work presented in his thesis, Thomas Ricketts had shown that in the C57BL/6J genetic background *Tardbp*^{F210I} is inherited in a Mendelian pattern in heterozygosity and leads to prenatal mortality in homozygosity around 15.5dpc and, in addition, that *Tardbp*$^{Q101X/Q101X}$ die very early during gestation and that *Tardbp*$^{Q101X/F210I}$ also die prenatally (Ricketts 2012).

The observations regarding the inheritance of *Tardbp*^{M323K} and *Tardbp*^{F210I} alleles, constituting original work presented in this thesis are summarised as follows:

- In the C57BL/6J background *Tardbp*^{M323K} is inherited in a Mendelian pattern in heterozygosity (table 3.1) and leads to perinatal mortality in homozygosity, possibly between postnatal days1 and 5 (table 3.1).
- Hybrid C57BL/6J–DBA background rescues the prenatal mortality in *Tardbp*$^{M323K/M323K}$, which are found alive at weaning age in Mendelian ratios with *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ (table 3.3).
- Hybrid C57BL/6J–DBA background partly rescues the prenatal mortality in *Tardbp*$^{F210I/F210I}$ which are still found alive at 18.5dpc following a Mendelian pattern of inheritance (table 3.4).
- *Tardbp*$^{M323K/Q101X}$ animals (in congenic C57BL/6J background) are found in Mendelian ratios at weaning (table 3.5) and provide a valuable opportunity to investigate the autoregulation properties of the M323K mutant protein and how different TDP43 mutations interact *in vivo*.

- *Tardbp*$^{M323K/F210I}$ embryos are found in Mendelian ratios at 14.5dpc but in reduced numbers at weaning (table 3.6), suggesting reduced life expectancy and premature death. Whilst establishing adult phenotyping cohorts may be disabled by the premature death of the animals, it will be possible to use embryonic tissue to study how the mutations interact in modifying its individual molecular phenotypes (this point will be expanded in chapter 4 and the discussion).
- Hemizygosity for *TARDBP*^{G348C} does not rescue *Tardbp*$^{F210I/F210I}$ lethality (table 3.7), suggesting that homozygosity for the F210I mutation may cause premature death in mice through a mechanism which cannot be explained as a simple loss of function (which has been previously described as a functional consequence of the F210I mutation), illustrating how single point mutations can lead to more than one functional biological change in the TDP43's activity.

Having described the inheritance of the *Tardbp*^{M323K} allele, including in the context of additional *Tardbp* mutations, and further expanded on the inheritance of the *Tardbp*^{F210I} allele, the molecular characterisation of the *Tardbp*^{M323K} ENU mutants, and further characterisation of *Tardbp*^{F210I} mutants, expanding on the work from Thomas Ricketts on the latter (Ricketts 2012), will be described in the next chapter.

Chapter 4.

Results: Molecular Characterisation of *Tardbp*^{M323K} mutants and further characterisation of *Tardbp*^{F210I} mutants

As described in the introduction, TDP43 binds a large proportion of the murine and human transcriptomes and changes in its protein levels lead to widespread splicing and gene expression changes in both the mouse brain and human cell lines (Polymenidou et al. 2011; Tollervey et al. 2011).

Furthermore, in *Tardbp*^{F210I} ENU mutants, alternative exon splicing of target genes, including *Sort1*, *Dnajc5*, *Pdp1* and *Poldip3* is significantly altered in the same direction in which is altered upon TDP43 knockdown (Ricketts 2012), thus partly recapitulating, at least at the splicing level of specific transcripts, the effects of reduced protein levels as reported by (Polymenidou et al. 2011).

Hence, given the established function of TDP43 as an RNA binding protein involved in alternative exon splicing, the first step in the molecular characterisation of *Tardbp*^{M323K} mutants consisted in assessing whether the M323K mutation had an effect in alternative exon splicing of target genes.

4.1 *Tardbp*^{M323K} leads to dose dependent alternative exon splicing changes in target genes and the CFTR minigene in the opposite direction to those observed in TDP43 knockdown and *Tardbp*^{F210I}

Following the early identification of TDP43 as a critical factor in the splicing of exon 9 from *CFTR* (Buratti et al. 2001), a recombinant construct containing exon 9 from *CFTR* (referred to as the CFTR minigene) was frequently used to determine the consequences of mutations or deletions in TDP43 in its splicing activity.

Typically, splicing experiments were performed by co-transfecting cells with recombinant TDP43 constructs and the CFTR minigene (Buratti et al. 2001; Wang et al. 2004; Zhang et al. 2013).

The CFTR minigene was a generous gift from the Buratti and Baralle research group and the splicing assay has been modified from its usual protocol in the work presented in this thesis.

In our modified assay, only the CFTR minigene was transfected into MEFs harvested from different *Tardbp*^{M323K} embryos, with splicing of the exogenous construct (CFTR minigene) being performed by the endogenous TDP43 protein, either wild type or mutant (M323K), according to the genotype of the embryo from which the cells were harvested.

MEFs harvested from 14.5dpc *Tardbp*^{M323K} embryos originating from matings between *Tardbp*$^{M323K/+}$ mice (in the C57BL/6J background), were transfected with the CFTR minigene and the splicing activity of the endogenous TDP43 protein assessed by RT-PCR across all three genotypes (*Tardbp*$^{+/+}$, *Tardbp*$^{M323K/+}$and *Tardbp*$^{M323K/M323K}$).

Strikingly, an increase in the exclusion of exon 9 from the CFTR minigene was observed, in a dose-dependent fashion to *Tardbp*^{M323K} genotype (figure 4.1, refer to figure legend for statistical analysis).

In *Tardbp*$^{M323K/+}$ MEFs, exclusion of exon 9 was more frequent when compared to wild type controls falling just below statistical significance (p=0.055, figure 4.1) and, significantly more frequently excluded in *Tardbp*$^{M323K/M323K}$ MEFs when compared to the other two genotypes (figure 4.1, refer to figure legend for statistical analysis).

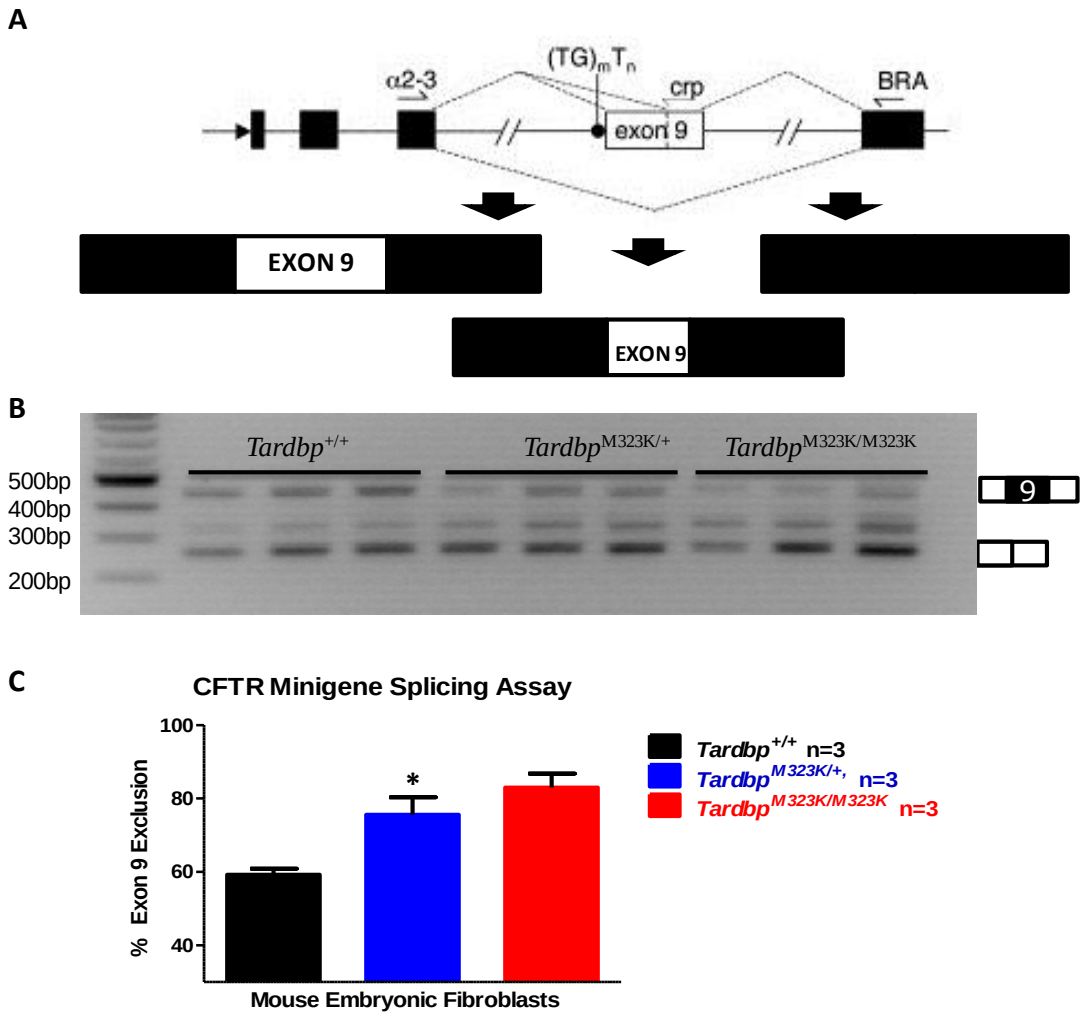

Figure 4.1: CFTR minigene splicing assay in MEFs harvested from *Tardbp*^{M323K} embryos.
The CFTR minigene (A) is transfected into MEFs. 36 hours post-transfection the RNA is extracted, cDNA synthesis performed, which is followed by RT-PCR, using the α23 and BRA primers, hence amplifying both transcript isoforms (with and without exon 9), and agarose gel electrophoresis (B). A dose dependent effect between the frequency of exon 9 exclusion from the CFTR minigene mRNA and presence of TDP43-M323K mutant protein is observed (B and C). In *Tardbp*$^{M323K/+}$ MEFs exon 9 is more frequently excluded than in wild type MEFs, a difference which falls just below statistical significance (mean percentage of exon 9 exclusion 59.24±1.63 for *Tardbp*$^{+/+}$ and 75.59±4.71 for *Tardbp*$^{M323K/+}$, n=3 for both genotypes; p=0.055, Anova and posthoc Bonferroni). In *Tardbp*$^{M323K/M323K}$ MEFs, exon 9 is significantly more frequently excluded when compared to *Tardbp*$^{+/+}$ MEFs (mean percentage of exon 9 exclusion 59.24±1.63 for *Tardbp*$^{+/+}$ and 83.02±3.76 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes; *****p=0.01, Anova and posthoc Bonferroni). Error bars on graph are ±SEM.
The intermediate band seen on the agarose gel (B) constitutes a splicing event occurring at a cryptic splice site within exon 9, identified as crp in the CFTR minigene diagram (A).
CFTR minigene diagram (A) adapted from (Ayala et al. 2006).

Splicing changes in the opposite direction (i.e. significantly less frequent exclusion of exon 9 from the CFTR minigene) is observed when the experiment is performed, using the same modified assay, in *Tardbp*$^{F210I/F210I}$ MEFs and in mammalian cell lines subjected to RNA interference leading to TDP43 knockdown using the original transfection protocol (Ayala et al. 2006).

In addition to the CFTR minigene splicing assay, and taking advantage of the knowledge of TDP43's endogenous murine splicing targets published in 2011 (Polymenidou et al. 2011), the investigation of the splicing activity of mutant TDP43-M323K protein was extended to endogenous targets.

After RNA was extracted from MEFs and embryonic heads harvested from 14.5dpc *Tardbp*^{M323K} embryos, cDNA synthesis was performed and assessment of alternative exon splicing of target genes, specifically *Sort1*, *Dnajc5*, *Pdp1* and *Poldip3* was perform by RT-PCR.

The primers used in the RT-PCR spanned the cassette exons of the named genes and had the same sequence as those used by (Polymenidou et al. 2011) and by Thomas Ricketts in the molecular characterisation of *Tardbp*^{F210I} mutants, thus enabling direct comparison of the splicing effects with these two previous studies (Polymenidou et al. 2011; Ricketts 2012).

In two of the four genes assessed, specifically *Sort1* (figure 4.2, refer to figure legend for statistical analysis) and *Dnajc5* (figure 4.3, refer to figure legend for statistical analysis), a significantly higher frequency of exclusion of the cassette exons was observed in a dose-dependent fashion to *Tardbp*^{M323K} on both MEFs and embryo heads. However, no significant differences in the exclusion of cassette exons from *Poldip3* and *Pdp1* was observed, despite a trend towards more frequent cassette exon exclusion being seen particularly in the latter (figure 4.4, refer to figure legend for statistical analysis).

The splicing pattern promoted by mutant TDP-M323K protein in *Sort1* and *Dnajc5* transcripts is thus diametrically opposed to the splicing pattern promoted by the mutant TDP43-F210I and by TDP43 knockdown, as illustrated in figure 4.5.

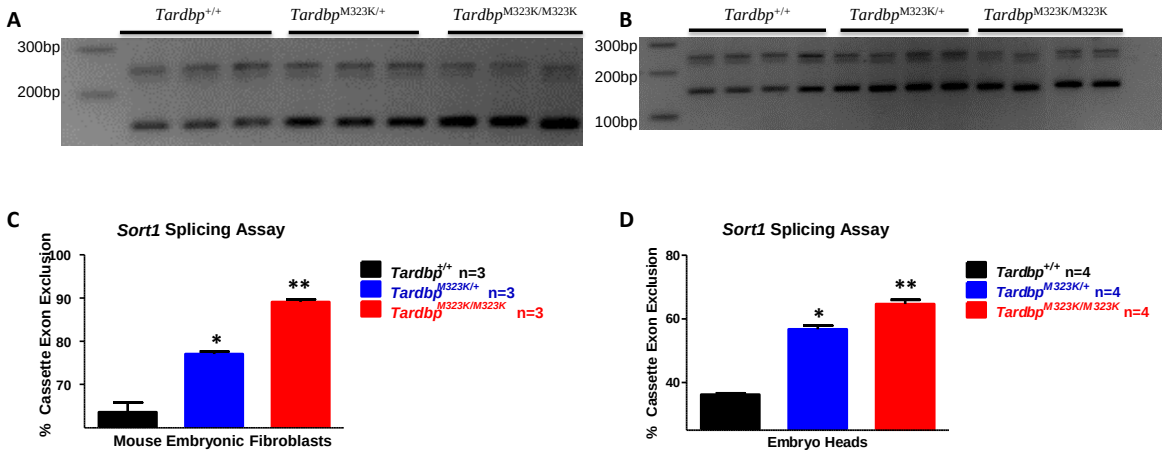

Figure 4.2: Splicing assay for the endogenous *Sort1* gene in MEFs and embryo heads harvested from *Tardbp*^{M323K} embryos at 14.5dpc.

Cassette exon exclusion from the endogenous *Sort1* gene was significantly more frequent in *Tardbp*$^{M323K/+}$ MEFs (mean percentage of cassette exon exclusion 77.8±0.58 for *Tardbp*$^{M323K/+}$ and 63.64±2.22 for *Tardbp*$^{+/+}$, n=3 for both genotypes; *****p=0.001, Anova and posthoc Bonferroni) and *Tardbp*$^{M323K/M323K}$ MEFs (mean percentage of cassette exon exclusion 89.11±0.57 for *Tardbp*$^{M323K/M323K}$ and 63.64±2.22 for *Tardbp*$^{+/+}$, n=3 for both genotypes ******p<0.001, Anova and posthoc Bonferroni) when compared to wild type controls (A and C). Error bars on graphs are ±SEM. Analogously, Cassette exon exclusion from the endogenous *Sort1* gene was significantly more frequent in *Tardbp*$^{M323K/+}$ embryo heads (mean percentage of cassette exon exclusion 56.76±1.14 for *Tardbp*$^{M323K/+}$ and 36.21±0.41 for *Tardbp*$^{+/+}$, n=4 for both genotypes; *****p=0.001, Anova and posthoc Bonferroni) and *Tardbp*$^{M323K/M323K}$ embryo heads (mean percentage of cassette exon exclusion 64.65±1.34 for *Tardbp*$^{M323K/M323K}$ and 36.21±0.41 for *Tardbp*$^{+/+}$, n=4 for both genotypes; ******p<0.001, Anova and posthoc Bonferroni) when compared to wild type controls (B and D). Error bars on graphs are ±SEM.

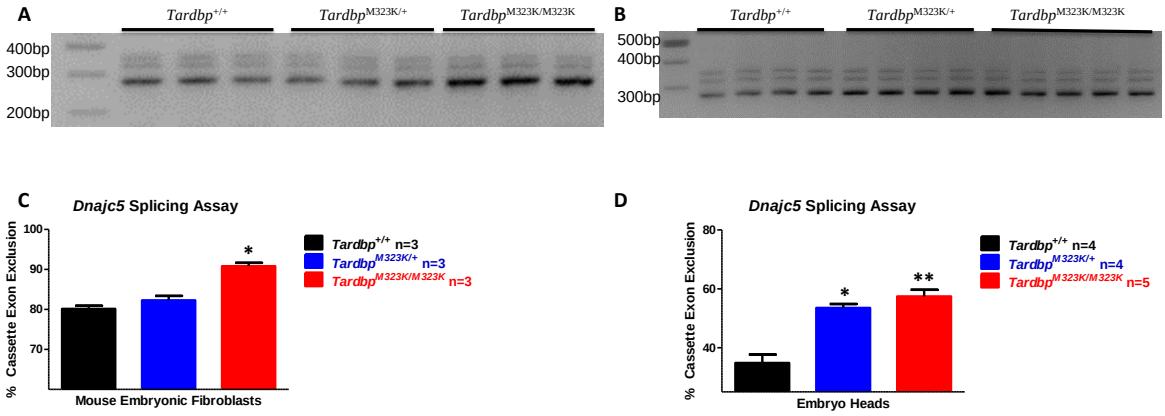

Figure 4.3: Splicing assay for the endogenous *Dnajc5* gene in MEFs and embryo heads harvested from *Tardbp*^{M323K} embryos at 14.5dpc.

Cassette exon exclusion from the endogenous *Dnajc5* gene was significantly more frequent in *Tardbp*$^{M323K/M323K}$. MEFs (mean percentage of cassette exon exclusion 90.82±0.84 for *Tardbp*$^{M323K/M323K}$ and 80.14±0.76 for *Tardbp*$^{+/+}$, n=3 for both genotypes; *****p=0.002, Anova and posthoc Bonferroni) (A and C) when compared to wild type controls. Error bars on graphs are ±SEM. Cassette exon exclusion from the endogenous *Dnajc5* gene was significantly more frequent in embryo heads from *Tardbp*$^{M323K/+}$ (mean percentage of cassette exon exclusion 53.51±1.33 for *Tardbp*$^{M323K/+}$ and 34.82±2.87 for *Tardbp*$^{+/+}$, n=4 for both genotypes; *****p=0.001, Anova and posthoc Bonferroni) and *Tardbp*$^{M323K/M323K}$ (mean percentage of cassette exon exclusion 57.43±2.22 for *Tardbp*$^{M323K/M323K}$, n=5 and 34.82±2.87 for *Tardbp*$^{+/+}$, n=4; ******p<0.001, Anova and posthoc Bonferroni) when compared to wild type controls (B and D). Error bars on graphs are ±SEM.

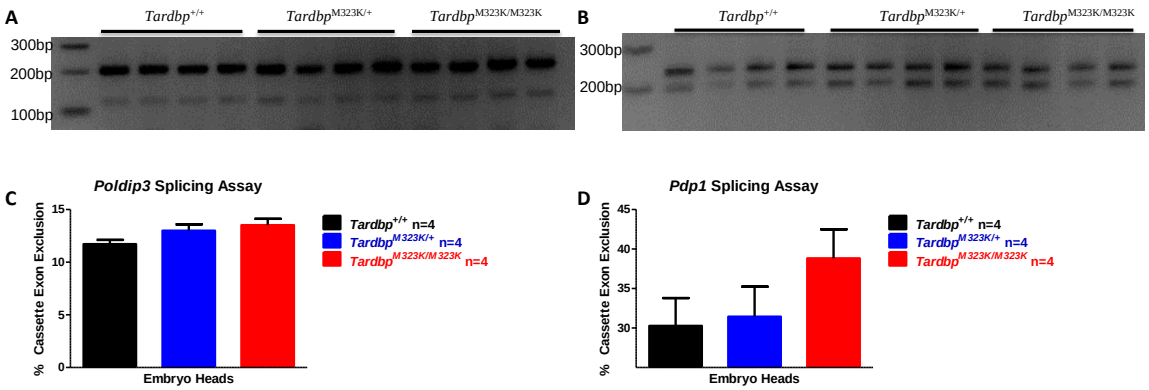

Figure 4.4: Splicing assay for the endogenous *Poldip3* and *Pdp1* genes in embryo heads harvested from *Tardbp*^{M323K} embryos at 14.5dpc

Cassette exon splicing for *Poldip3* (A and C) and *Pdp1* (B and D) genes did not reveal any significant changes in the frequency of exon exclusion associated with different M323K genotypes.

Mean percentage of *Poldip 3* cassette exon exclusion 11.71±0.41 for *Tardbp*$^{+/+}$, 12.00±0.60 for *Tardbp*$^{M323K/+}$ and 13.52±0.60 for *Tardbp*$^{M323K/M323K}$, n=4 per genotype; p=0.768 both the comparison between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ and the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$; Anova and posthoc Bonferroni. Error bars on graphs are ±SEM.

Mean percentage of *Pdp1* cassette exon exclusion 30.28±3.50 for *Tardbp*$^{+/+}$, 31.44±3.80 for *Tardbp*$^{M323K/+}$ and 38.82±3.67 for *Tardbp*$^{M323K/M323K}$, n=4 embryo per genotype; p=1 for comparison between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ and p=0.400 for comparison between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/M323K}$. Error bars on graphs are ±SEM.

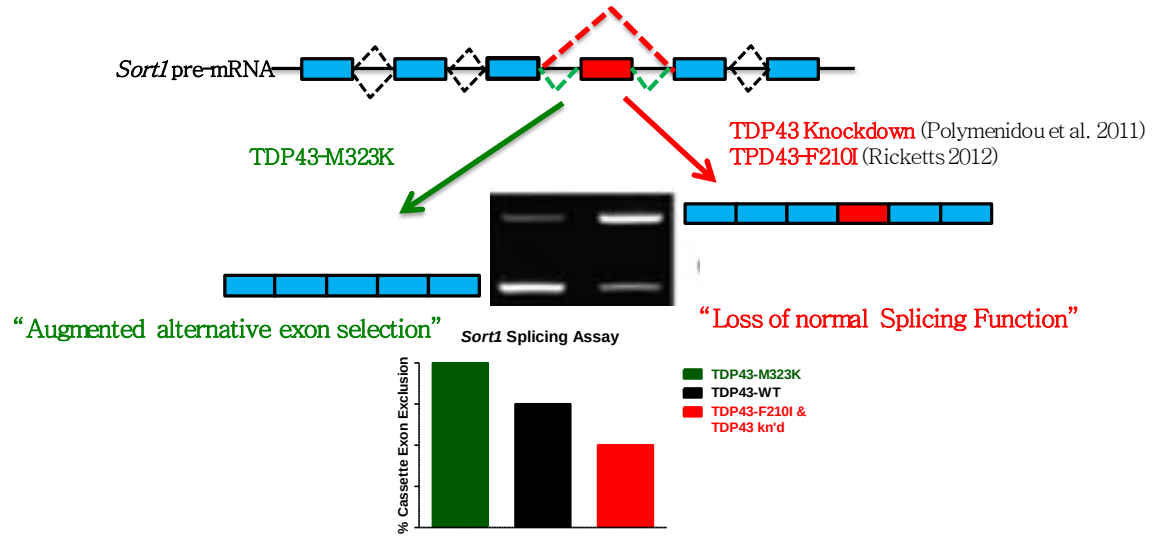

Figure 4.5: Diagrammatic representation of the opposite effect caused in TDP43 dependent splicing by the M323K mutation to the F210I mutation and TDP43 knockdown

The splicing of *Sort1* pre-mRNA is used to illustrate the opposite effects of the M323K mutation to the F210I ENU mutation and TDP43 knockdown in TDP43-dependent splicing.

The cassette exon in *Sort1* pre-mRNA (red box) is less frequently excluded from the mRNA when TDP43 protein levels are knocked down *in vivo* (Polymenidou et al. 2011) or in tissue from *Tardbp*$^{F210I/+}$ and *Tardbp*$^{F210I/F210I}$ ENU mouse mutants when compared to *Tardbp*$^{+/+}$.

In contrast, the same cassette exon is more frequently excluded in tissue from *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ ENU mouse mutants when compared to *Tardbp*$^{+/+}$.

The two isoforms, with and without the cassette exon, can be visualised on an agarose gel after RT-PCR from cDNA, with the brightness of the band enabling semi-quantitative determination of their relative abundance and graphical display the frequency of cassette exon exclusion in a bar chart (image of agarose gel adapted from Polymenidou et al. 2011).

Additionally, over-expression of wild type TDP43 in mammalian cell lines leads to more frequent exon 9 exclusion from the CFTR minigene (Buratti et al. 2001), whilst depletion of TDP43 through RNA interference leads to decreased exon 9 exclusion from the CFTR minigene (Ayala et al. 2006), thus validating the inference made from the observation of the splicing effects of the M323K and F210I mutations at the level of at least some endogenous genes:

- The M323K mutation leads to "augmented alternative exon selection".
- The F210I mutation leads to a "loss of normal splicing function".

The CFTR minigene splicing assay and assessment of alternative splicing in previously described TDP43 target genes have hence demonstrated that, at the splicing level, the M323K mutation leads to "augmented alternative exon selection", whilst the F210I mutation leads to a "loss of normal splicing function", both on the CFTR minigene splicing assay and, *in vivo*, at least on the splicing of selected target genes, including *Sort1* and *Dnajc5*.

Nevertheless, the opposite direction of the splicing changes is not extended to every individual cassette exon spliced by TDP43. Examples of specific splicing targets that are affected by only one of the mutations include the splicing of the cassette exons from *Poldip3* and *Pdp1* which are significantly changed in the same direction by the F210I mutation (Ricketts 2012) and TDP43 knockdown (Polymenidou et al. 2011) but unchanged by the M323K mutation (figure 4.4).

Having determined the two opposite effects of the M323K and F210I mutations in target genes, further research was performed with the objectives of continuing to characterising the effects of these mutations, and additionally, the underlying biological mechanism leading to the "augmented alternative exon selection" in the M323K mutant protein.

The mechanism leading to "loss of normal splicing function" in the F210I mutant protein was determined by Thomas Ricketts and is attributable to the decreased affinity of TDP43-F210I for RNA when compared to wild type TDP43 (Ricketts 2012).

In addition, the effects of both the M323K and F210I mutations were further investigated by performing RNA-Seq in embryo heads and MEFs, the results of which will be subsequently described, followed by the description of the results from the investigation into the mechanisms leading to the "augmented alternative exon selection" observed in the M323K protein.

4.2 The M323K and F210I ENU mutations lead to genome-wide transcriptome changes in vivo

RNA-seq was performed from MEFs and embryo heads harvested from *Tardbp*^{M323K} and *Tardbp*^{F210I} embryos at 14.5dpc.

The tissue was harvested at the MRC Harwell and the RNA-seq performed in collaboration with Pietro Fratta in the research group of Professor Elizabeth Fisher, at University College London's Institute of Neurology. Data analysis was performed in collaboration with the group of Dr Vincent Plagnol, at University College London.

Analysis from RNA-seq confirmed the splicing changes identified by RT-PCR in target endogenous genes caused by the M323K and F210I mutations.

Moreover, the effects of M323K in the transcriptome were shown to be vast, with 696 exons being differentially spliced in *Tardbp*$^{M323K/M323K}$ when compared to *Tardbp*$^{+/+}$ (table in Annex 1).

Changes in alternative exon splicing resulting from the M323K mutation were a consequence of several types of alternative splicing events, including cassette exon and retained intron

types of alternative splicing and, in addition, alternative finish of transcript's exons by differential splicing at the 3' UTR of the pre-mRNA, which constituted the most frequent event leading to splicing changes. The majority of differentially spliced exons in *Tardbp*^{M323K} mutant tissue were more frequently excluded when compared to wild type (*Tardbp*$^{+/+}$) controls (table in Annex 1).

In contrast to the RNA-seq data from *Tardbp*^{M323K} embryonic tissue, the F210I mutation, in homozygosity, caused significant changes in alternative splicing of 79 exons (table in Annex 2), the majority of which were in the opposite direction of the change observed in the same exons caused by the M323K mutation (figure 4.6).

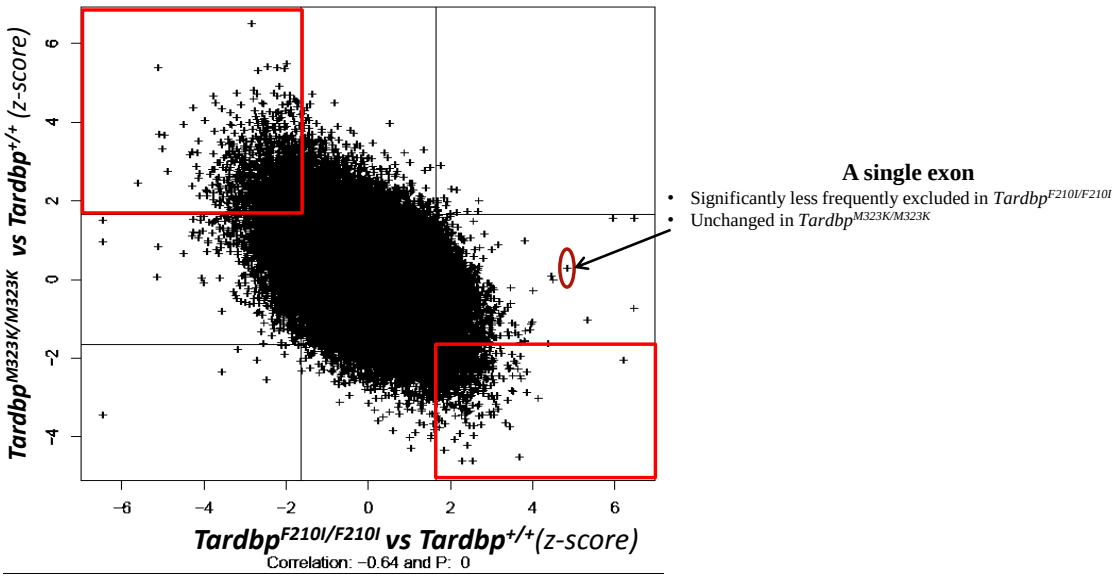

Figure 4.6: The M323K and F210I mutations cause genome wide opposite effects in alternative exon splicing.
Diagrammatic representation of genome-wide alternative splicing changes observed in *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{F210I/F210I}$ embryo heads harvested at 14.5dpc.
Individual exons are represented by black crosses and the changes in exon splicing are given for each mutant animal line in comparison with the wild type controls as a z-score.
A positive z-score identifies exons which are significantly less frequently excluded in the mutants; a z-score of 0 that the splicing is unchanged, and a negative z-score that the exon is significantly more frequently excluded in the mutants.
In the y-axis the z-scores between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$ are given and the in x-axis the z-scores between *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{+/+}$ are given for the same exons.
A large number of exons are located in areas of the graph in which the z-score is positive for the *Tardbp*$^{M323K/M323K}$ and negative for *Tardbp*$^{F210I/F210I}$ (top left corner, highlighted in red) and vice-versa (bottom right corner of graph, highlighted in red), demonstrating that splicing changes are observed in the opposite directions in the two mutants for a large number of exons.
The splicing of some exons is only changed as a consequence of one of the mutations and not others, an example of which has been highlighted (red circle) in the graph. 3 biological replicates were used for each mutant genotype and matched wild type controls.

The RNA-seq data has thus confirmed the findings regarding the alternative exon splicing changes caused by the M323K and F210I mutations. It has specifically shown that alternative exon splicing is significantly changed in opposite directions by TDP43-F210I and TDP43-M323K at the whole transcriptome level, therefore demonstrating that the "augmented alternative exon selection" of the M323K protein and the "loss of normal splicing function" of the F210I protein have effects at the genome-wide level.

However, as described earlier, it is necessary to register that despite the opposite changes seen in alternative splicing at the whole transcriptome level between *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{M323K/M323K}$ mutants, there are exons in which alternative splicing changes are observed in one of the mutants and remain unchanged in the other (figure 4.6), which is in concurrence with the previous splicing data generated by RT-PCR.

In addition to alternative exon splicing, RNA-seq also enabled the assessment of whether gene expression was influenced by the M323K and F210I mutations, which is of particular interest given that regulating gene expression is also one of TDP43's biological functions (Polymenidou et al. 2011).

Remarkably, whilst the F210I mutation recapitulates the splicing effects of TDP43 knockdown, it has the opposite effect in the regulation of the expression of long intron genes.

Genes with long introns, specifically with an average intron length of 28,707 base pairs, containing multiple TDP43 binding sites, were found to be downregulated on TDP43 knockdown. The average intron length in transcripts downregulated on TDP43 knockdown (28,707 base pairs) is six fold higher than the intron length of genes whose expression remained unaffected (Polymenidou et al. 2011), hence being described collectively as "long intron genes".

Moreover, many long intron genes which were downregulated on TDP43 knockdown have been demonstrated to be involved in synaptic transmission (e.g. *Grin2a*, *Glur6*) and implicated in the pathophysiology of neurological diseases (Polymenidou et al. 2011). Therefore, the effect of TDP43 dysfunction in the regulation of long intron gene expression is particularly relevant from a translational perspective.

Thus, as described above, TDP43 knockdown in the striatum leads to preferential depletion of transcripts from long intron genes, whilst in the embryo heads of *Tardbp*$^{F210I/+}$ and *Tardbp*$^{F210I/F210I}$, the opposite pattern is observed and the transcripts from the same long intron genes are upregulated (figure 4.7). The M323K mutation does not have an effect in the expression of long intron genes (figure 4.7).

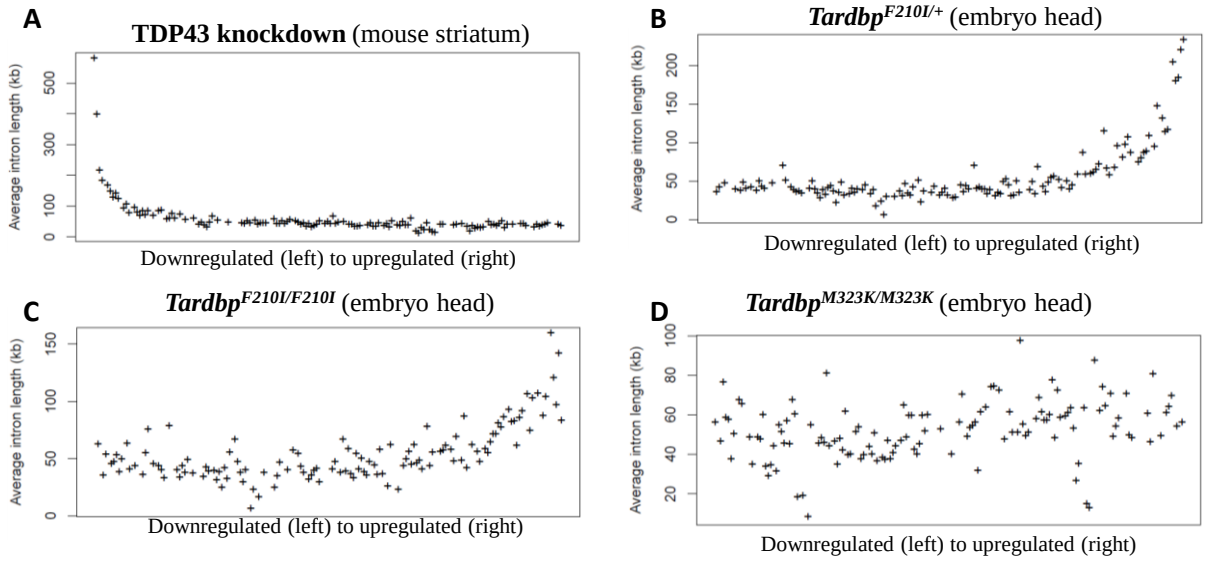

Figure 4.7: Effects of TDP43 knockdown, TDP43-F210I and TDP43-M323K in the expression of long intron genes *in vivo*.
The expression of long intron genes is affected by TDP43 knockdown and the F210I mutation but not the M323K mutation as demonstrated in this graphical representation of the RNA-seq data analysis.
Individual genes are represented as black crosses and plotted in the y-axis according to average intron length (in Kb). Expression, in relation to wild type controls is represented along the x-axis, with the left hand side of the graph representing downregulation and the right hand side of the graph upregulation.
In the graph for TDP43 knockdown, genes with long introns are clustered in the left side of the x-axis, as they are downregulated when compared to wild type (A). Data from (Polymenidou et al 2010).
In contrast, in the graphs for *Tardbp*$^{F210I/+}$ (B) and *Tardbp*$^{F210I/F210I}$ (C) genes with long introns are clustered on the right-side of the x-axis, given that they are upregulated.
In the graph for *Tardbp*$^{M323K/M323K}$ (D), genes with long introns are scattered along the x-axis, given that a specific effect in its expression levels was not observed. 3 biological replicates were used for each mutant genotype and matched wild type controls for *Tardbp* mutant embryo heads.

The observation of similar effects between TDP43 knockdown and the F210I mutation in alternative splicing and opposite effects in the levels of expression of genes with long introns demonstrates the relevance of investigating TDP43 biology using isogenic *Tardbp* ENU mutants, particularly in the context of mutation-associated TDP43 dysfunction causing downstream neurodegeneration.

The interpretation of the results of TDP43 knockdown and TDP43-F210I in the expression of genes with long introns needs to take into consideration that the experiments with the protein knockdown were performed in adult mouse brain tissue, specifically the striatum (Polymenidou et al. 2011) whilst for $Tardbp^{F210I}$ mutants embryo heads were used for the RNA-seq.

$Tardbp^{F210I/F210I}$ animals die prenatally (refer to chapter 3). Nevertheless, validation of mutant TDP-F210I protein in the expression of long intron genes, which is currently being planned, will include samples of adult brain tissue of $Tardbp^{F210I/+}$ mice.

Given that the effect on the expression of long intron genes is also seen in the heads of heterozygous embryos, successful validation of the RNA-seq data in adult brain tissue of $Tardbp^{F210I/+}$ will confirm that the opposite effects between $Tardbp^{F210I}$ and TDP43 knockdown in gene expression are not seen as a consequence of tissue specificity but constitute a *bona fide* global effect across developmental and adult neural tissue.

Moreover, the fact that the downregulation of long intron genes is also observed in $Tardbp^{F210I/+}$ embryo heads suggests that this is a dominant effect from the F210I protein, possibly a “dominant gain of function” in contrast with its dose dependent “loss of normal splicing function”.

The effects of a single point mutation in TDP43 can thus be multiple, suggesting that cellular dysfunction and neurodegeneration could be a consequence of one specific effect (e.g. gain of

normal function/loss of normal function/gain of novel function) or, alternatively, a toxic combination of multiple effects from a single mutation.

Now that the description of the molecular effects of the M323K and F210I mutations has been given, the subsequent sections will focus on the investigation into the underlying biological mechanism leading to the "augmented alternative exon selection" in TDP-M323K.

4.3 Investigating the biological mechanism leading to "augmented alternative exon selection" in the M323K mutation

The investigation into the biological mechanism leading to "augmented alternative exon selection" observed in the M323K mutation was pursued in a logical manner, making use of the published evidence for TDP43's splicing function and the domains and interactions of the protein associated with this biological function.

Given that TDP43 knockdown leads, *in vivo*, to suppression of its splicing function (Polymenidou et al. 2011), it was pivotal to determine whether protein levels were altered by the M323K mutation.

Moreover, given that, at least in some target genes (figures 4.2 and 4.3), there is a dose-dependent "augmented alternative exon selection" in *Tardbp*^{M323K} animals, if the M323K mutation led to higher protein levels, it could be concluded that the "augmented alternative exon selection" observed constituted a direct effect of the higher protein levels.

Higher TDP43 levels would itself constitute a molecular phenotype and dysfunction in one of TDP43's biological functions, as TDP43 has been shown to autoregulate its own levels (Ayala et al. 2011; Polymenidou et al. 2011; Avendaño-Vázquez et al. 2012; Bembich et al. 2013).

Furthermore, autoregulation has been shown to occur *in vivo* given that in brain tissue from *Tardbp*$^{Q101X/+}$ mice TDP43 protein levels are equivalent to wild type controls, but the levels of *Tardbp* wild type transcript are increased and the levels of the mutant *Tardbp*^{Q101X} transcript, which constitutes a null allele, reduced (Ricketts et al. 2014).

In addition to the biological importance of TDP43 levels, it has already been described that TDP43 mislocalisation, specifically the cytoplasmic aggregation of CTFs which are hyperphosphorylated and polyubiquitylated, constitutes a histopathological hallmark of ALS and FTLD-TDP (Neumann et al. 2006; Davidson et al. 2007; Cairns et al. 2007; Dickson et al. 2007; Fujita et al. 2008; Mori et al. 2008; Brandmeir et al. 2008; Pamphlett et al. 2009; Tan et al. 2013), including in patients in whom TDP43 C-terminal mutations are causative of these diseases.

Consequently, the subcellular distribution of TDP43 was also assessed in *Tardbp*^{M323K} mutants, aiming to determine whether the M323K mutation led to increased levels of cytoplasmic TDP43, either in the presence, or absence, of increased "whole cell" TDP43 levels.

Moreover, given the already described "augmented alternative exon selection" caused by the M323K mutation, if increased levels of TDP43 were found in the nucleus of cells from *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ when compared to wild type controls, it could explain, at least in part, the molecular phenotype. Nevertheless, since TDP43 is predominately nuclear, the sensitivity of this methology in identifying increased levels of nuclear TDP43 is reduced in comparison to its sensitivity in identifying increased levels of cytoplasmic protein.

4.3.1 The M323K mutation does not lead to significant changes in TDP43 protein levels or its subcellular localisation

Protein was extracted from MEFs, embryonic heads and adult brain and TDP43 levels assessed by immunoblotting and semi-quantitative analysis of the band intensity produced from the light emitted by the fluorochromes conjugated with secondary antibodies.

When the TDP43 levels were adjusted for individual sample loading by performing a ratio between TDP43 and the loading control marker (either Actin or Tubulin), TDP43 levels were not significantly different across genotypes in all tissues tested, including MEFs, embryonic heads and adult brains (figure 4.8, refer to figure legend for statistical analysis).

In the immunoblots shown in figure 4.8, an overall trend is seen towards higher protein levels in homozygous tissue, nevertheless it falls below statistical significance and the trend itself is not consistently seen in every immunoblot (see figure 4.15). The exception consists in the immunoblots from adult brain tissue, in which the trend towards higher protein levels could be more suggestive of a possible difference on the single immunoblot that was performed and therefore further immunoblotting from protein extracted from adult brain tissues is indicated.

In addition to assessing whole-cell protein levels as described above, the localisation and relative nuclear/cytoplasmic abundance of TDP43-M323K was also assessed using two methodologies, nuclear/cytoplasmic protein fractionation and immunoblotting and immunofluorescence.

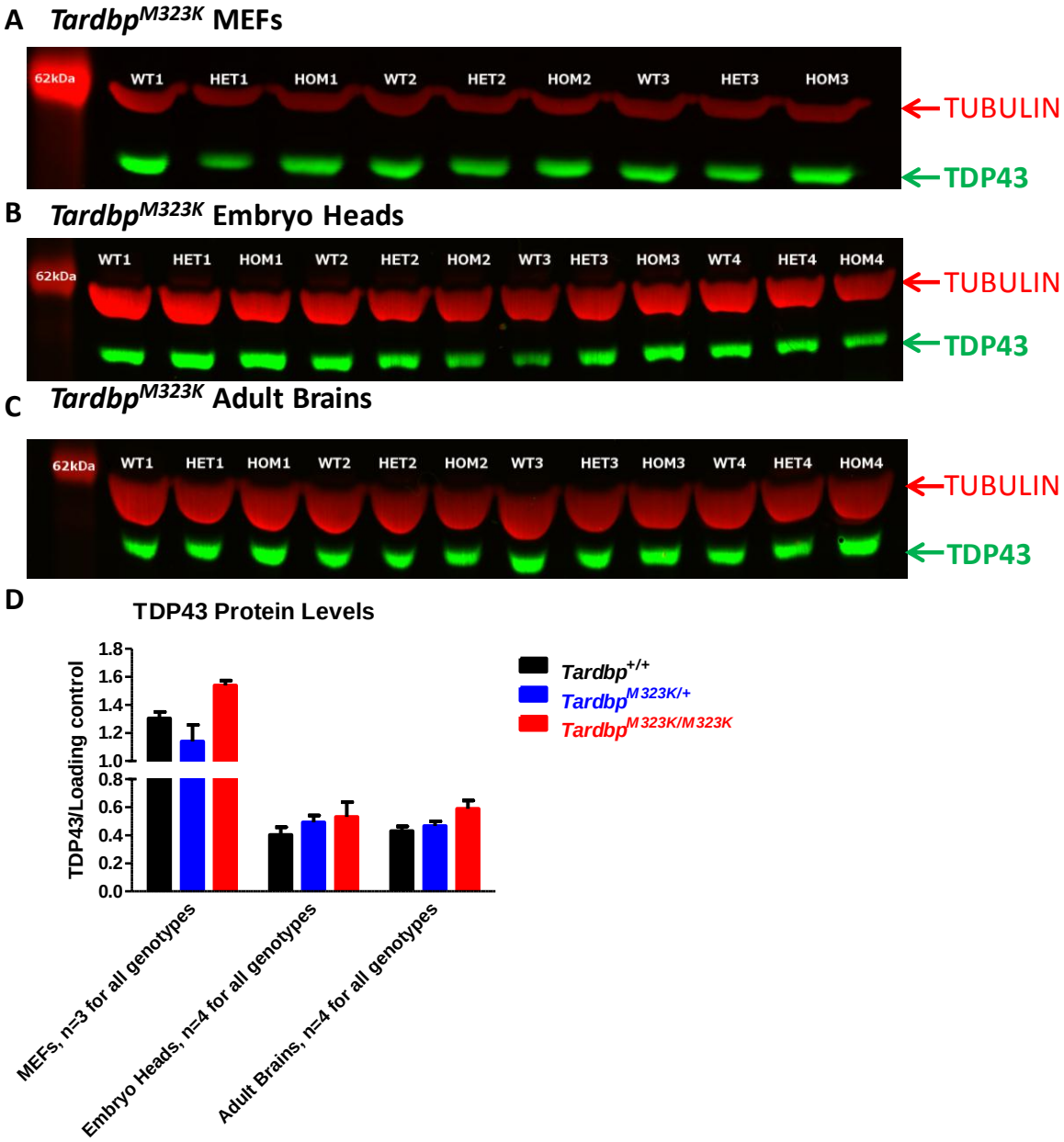

Figure 4.8: TDP43 protein levels in MEFs, embryo heads and adult brains harvested from *Tardbp*^{M323K} embryos at 14.5dpc.

No significant differences were seen in TDP43 protein levels between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ or *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/M323K}$ in (A) MEFs (mean ratio TDP43/Tubulin 1.30±0.046 for *Tardbp*$^{+/+}$, 1.14±0.12 for *Tardbp*$^{M323K/+}$ and 1.54±0.035 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes p=0.530 between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ and p=0.212 between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$, Anova and posthoc Bonferroni), (B) embryonic heads (mean ratio TDP43/Tubulin 0.40±0.056 for *Tardbp*$^{+/+}$, 0.49±0.049 for *Tardbp*$^{M323K/+}$ and 0.53±0.11 for *Tardbp*$^{M323K/M323K}$, n=4 for all genotypes p=1 between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ and p=0.750 between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$, Anova and posthoc Bonferroni) and (C) Adult brains (mean ratio TDP43/Tubulin 0.43±0.034 for *Tardbp*$^{+/+}$, 0.47±0.033 for *Tardbp*$^{M323K/+}$ and 0.59±0.060 for *Tardbp*$^{M323K/M323K}$, n=4 for all genotypes p=1 between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ and p=0.094 between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$, Anova and posthoc Bonferroni), as displayed graphically in (D). A more suggestive trend towards higher protein levels is however observed in adult brains and further immunoblotting will be useful in determining whether the trend is consistently reproducible in this tissue. Error bars in graph are ±SEM. Genotypes as labelled in immunoblots; WT: *Tardbp*$^{+/+}$; Het: *Tardbp*$^{M323K/+}$; HOM: *Tardbp*$^{M323K/M323K}$

Protein extraction from nuclear and cytoplasmic cellular compartments (referred to as nuclear/cytoplasmic protein fractionation) was performed in MEFs harvested from 14.5dpc embryos using a commercially available kit (NE-PER Nuclear cytoplasmic extraction reagents from Thermo Scientific).

Immunoblotting of nuclear/cytoplasmic fractionated protein from MEFs did not reveal any significant differences in the subcellular distribution of TDP43 (figure 4.9, refer to figure legend for statistical analysis).

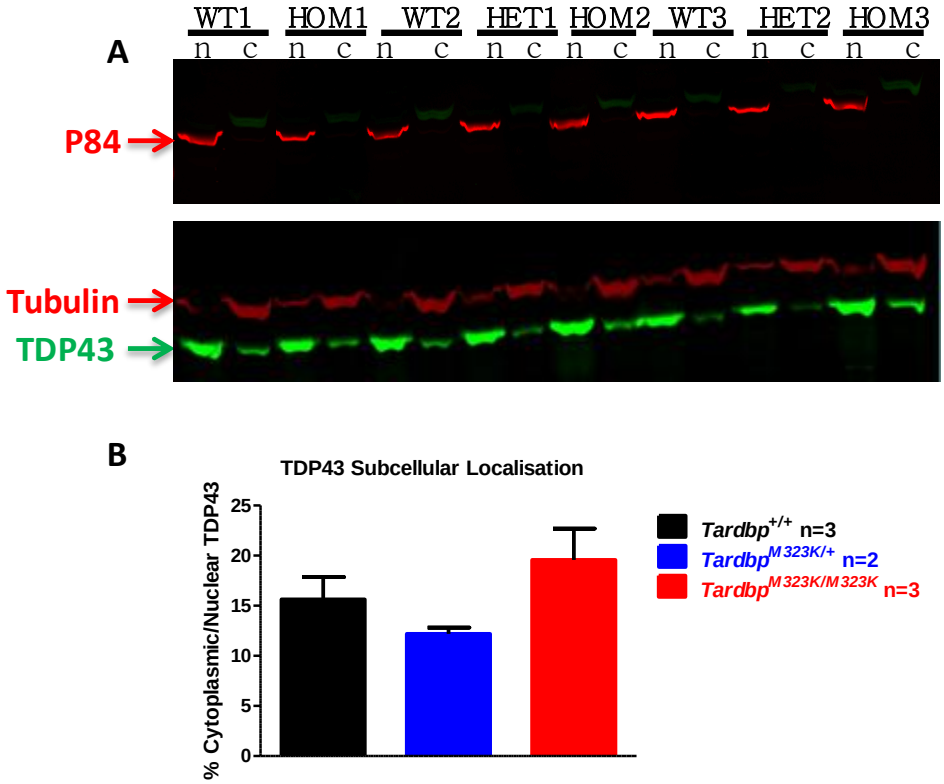

Figure 4.9: Assessment of TDP43-M323K subcellular distribution by immunoblotting
Immunoblotting of nuclear/cytoplasmic fractionated protein extracted from MEFs harvested from *Tardbp*^{M323K} embryos at 14.5dpc did not reveal any significant changes in the ratio of cytoplasmic/nuclear TDP43 across genotypes (mean percentage of cytoplasmic/nuclear TDP43 15.65±2.21 for *Tardbp*$^{+/+}$ n=3, 14.32±2.15 for *Tardbp*$^{M323K/+}$ n=2 and 20.09±5.31 for *Tardbp*$^{M323K/M323K}$ n=3; p=0.908 between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$; p=0.335 between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{M323K/+}$ and p=1 between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$, Anova and posthoc Bonferroni**).** Anti-p84 antibody was used as a nuclear specific marker and anti-Tubulin antibody as a cytoplasmic specific marker, thus confirming that the fractionation was successful.

Immunofluorescence was performed in MEFs harvested from *Tarbp*^{M323K} embryos at 14.5dpc, using DAPI staining as a nuclear marker and a primary antibody against TDP43 followed by a fluorescent secondary antibody and confocal microscopy.

Qualitatively, no differences were observed in the subcellular localisation of TDP43 in MEFs harvested from *Tardbp*$^{+/+}$, *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ embryos (figure 4.10).

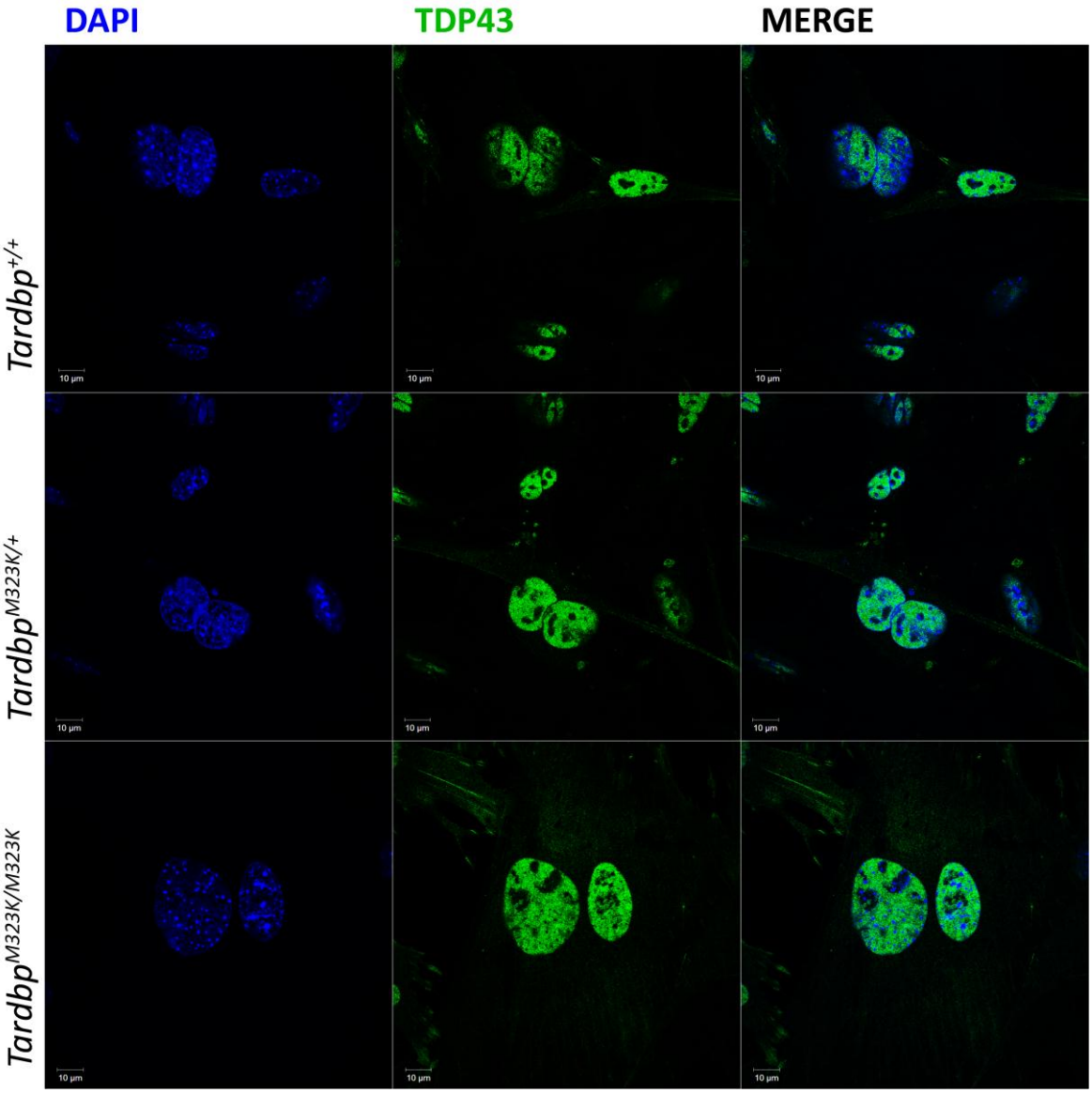

Figure: 4.10: Qualitative assessment of subcellular localisation of TDP43 by immunofluorescence in *Tardbp*^{M323K} MEFs
MEFs were incubated with a secondary fluorescent antibody and stained with DAPI after incubation with a primary anti-TDP43 antibody (*Protein-tech*). Confocal microscopy did not reveal any changes in the subcellular distribution of TDP43 protein in *Tardbp*$^{+/+}$, *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ cells.

Thus, in tissues and MEFs from *Tardbp*^{M323K} mutants, TDP43 protein levels and its subcellular distributions are similar across all genotypes (*Tardbp*$^{+/+}$ *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$).

Despite TDP43 protein levels in MEFs, embryonic heads and adult brains being comparable across all M323K genotypes (*Tardbp*$^{+/+}$, *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$), it remains unclear whether this observation is a consequence of autoregulation. Since autoregulation is proposed to occur at the transcript level (Ayala et al. 2011; Polymenidou et al. 2011), this process was further investigated by northern blot.

4.3.2 The TPD43-M323K transcript is differentially polyadenylated when compared to the wild type transcript

TDP43 is found in concentrations within a narrow physiological range and its levels are proposed to be regulated by the TDP43 protein itself. The autoregulation models, based on the evidence published to date, are described in the introduction in some detail.

The two existing models consensually suggest that TDP43 binds the 3' UTR of its own transcript at the TDPBR and enhances splicing of the cryptic intron 7, which is part of the transcripts 3' UTR. The downstream effects of intron 7 splicing leading to autoregulation diverge in the two models.

One model proposes that the exon-junction complexes deposited at the 3' end of the *Tardbp* transcript marks it for non-sense mediated decay (Polymenidou et al. 2011), whilst the other suggests that the canonical poly(A)1 site is removed upon splicing of intron 7, leading to polyadenylation of the *Tardbp* transcript at the non-canonical Poly(A)2 site, which disrupts cytoplasmic transport and disables translation of the transcripts polyadenylated at the alternative Poly(A)2 site (Ayala et al. 2011; Bembich et al. 2013).

Given that autoregulation is suggested to occur at the transcript level and that polyadenylation is proposed to be involved in this process, alterations to *Tardbp* RNA levels and its polyadenylation were investigated by northern blot and semi-quantitative determination of the band intensity.

This experiment was designed together with Jeremias Herzog, a PhD student at the ICGEB in Trieste, and the northern blot performed by him at that institution. Accounting for the multiple *Tardbp* transcript isoforms (figure 1.1), the probe was design to target exon 5, which is present in the main *Tardbp* transcripts identified in the genome database Ensembl.

Northern blot using RNA extracted from MEFs revealed that overall RNA levels were similar across genotypes (figure 4.11, refer to figure legend for mean RNA levels per genotype and statistical analysis). However, in $Tardbp^{M323K/M323K}$ MEFs, there was a significant increase in transcripts polyadenylated at the alternative Poly(A)4 site and consequent reduction in transcript levels polyadenylated at the canonical Poly(A)1 site (figure 4.11, refer to figure legend for statistical analysis).

It has already been described that the M323K mutation leads to "augmented alternative exon selection". Hence, it is possible that the mutant protein splices the cryptic intron 7 more frequently, thus removing the canonical poly(A)1 site, and promoting the usage of the alternative Poly(A)4 site.

Therefore, whilst the changes in polyadenylation site usage may be directly related to the TDP43 dysfunction caused by the M323K mutation, its physiological relevance for the autoregulation process remains unclear.

Despite not constituting the canonical polyadenylation site, *Tardbp* transcripts polyadenylated at the alternative Poly(A)4 site are observed at low frequencies and are translated efficiently into TDP43 (Avendaño-Vázquez et al. 2012). Moreover, the observations that total RNA

levels and TDP43 protein levels are similar across M323K genotypes suggest that the changes in polyadenylation site usage described are not influencing translation and thus could be unrelated to the autoregulation process.

The critical experiments in determining the autoregulation properties of the M323K mutant protein are the assessment of protein and RNA levels from *Tardbp*$^{Q101X/M323K}$ mouse tissue (refer to section 3.3.1 in the previous chapter) which will shed light on whether TDP43-M323K is in fact able to regulate its own protein levels *in vivo* and, additionally, whether differential polyadenylation site usage is seen to be correlated with the autoregulation process.

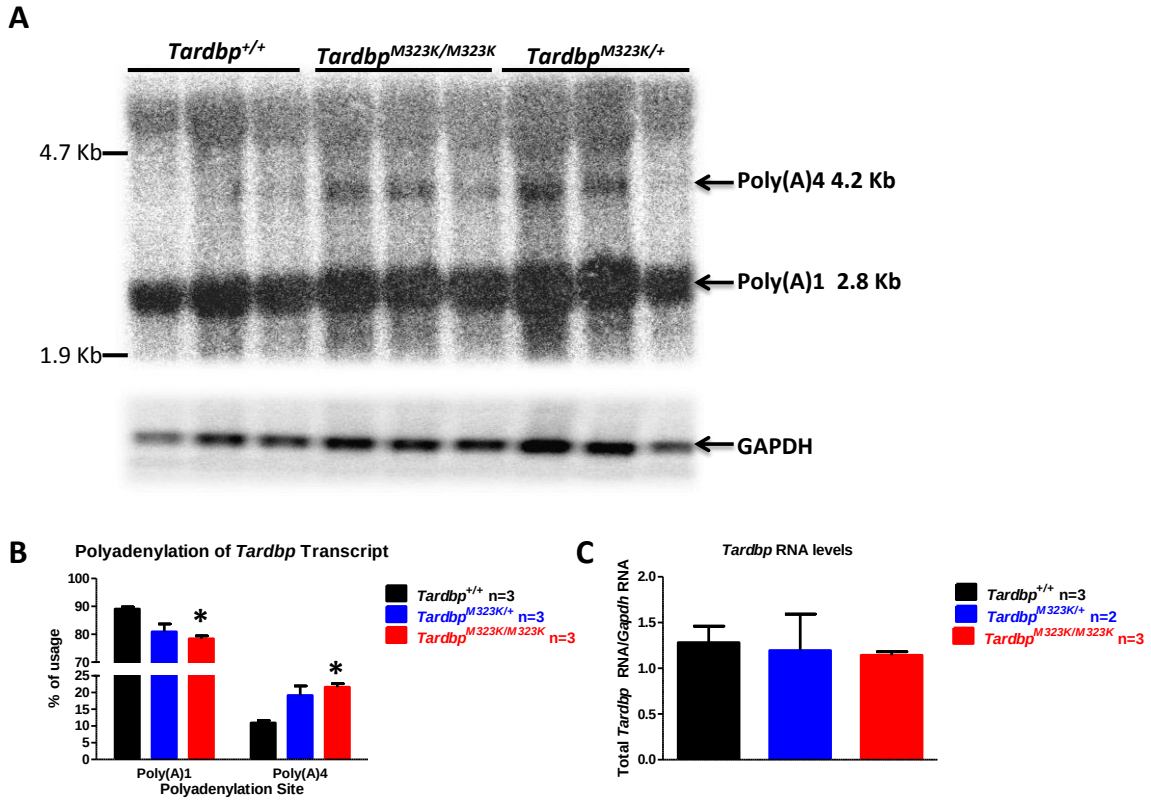

Figure 4.11: RNA levels and polyadenylation of wild type and *Tardbp*^{M323K} transcripts determined by northern blot.

Northern blot from MEFs harvested from *Tardbp*^{M323K} embryos at 14.5dpc (A) reveal (B) an increase in the usage of the alternative Poly(A)4 when the *Tardbp*^{M323K} transcript is present in homozygosity (mean percentage of Poly(A)4 usage 10.93±0.73 for *Tardbp*$^{+/+}$ and 21.6±1.03 for *Tardbp*$^{M323K/M323K}$, n=3 for both genotypes *p=0.017, Anova and posthoc Bonferroni) and a consequent decrease in the use of the canonical Poly(A)1 site (mean percentage of Poly(A)1 usage 89.03±0.75 for *Tardbp*$^{+/+}$ and 78.40±1.78 for *Tardbp*$^{M323K/M323K}$, n=3 for both genotypes *p=0.017, Anova and posthoc Bonferroni). A similar trend is observed between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$, but the difference falls just below statistical significance (p=0.054, n=3 for both genotypes, Anova and posthoc Bonferroni).

Total *Tardbp* RNA levels (C) are similar across all *Tardbp* genotypes (mean ratio of total *Tardbp* RNA/*Gapdh* RNA 1.2±018 for *Tardbp*$^{+/+}$, 1.19±0.40 for *Tardbp*$^{M323K/+}$ and 1.16±0.04 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes p>0.05 for comparisons across genotypes, Anova and posthoc Bonferroni).

4.3.3 TDP-M323K mutant protein interactions with RNA and other proteins: investigations into the biological mechanism leading to the "augmented alternative exon selection"

TDP43 is known to interact with RNA and other proteins.

TDP43 interactions with RNA involve both RNAs with the canonical $(UG)_6$ repeat sequences and RNAs with other (UG/GU)-rich sequences and this interactions are essential for TDP43 dependent splicing (Buratti & Baralle 2001; Tollervey et al. 2011).

Proteins which interact with TDP43 include proteins from the same hnRNP family (Buratti et al. 2005), for which the C-terminal region located between the amino acid residues 321 and 326 of the protein's 414 amino acid chain, and thus the region where the M323K mutation is located, is critically important (D'Ambrogio et al. 2009).

Furthermore, in addition to interacting with other proteins, TDP43 has also been shown to homodimerise (Kuo et al. 2009).

TDP43's interactions with other proteins and itself are suggested to be crucial processes in its splicing activity given that integrity of the C-terminal domain is required for TDP43 dependant splicing (Wang et al. 2004). Additionally, TDP43 homodimerisation has also been shown to influence TDP43 dependent splicing (Zhang et al. 2013).

In view of the known interactions of TDP43 with other proteins, itself and RNA, succinctly described above, the interactions between the mutant M323K protein and RNA, as well as its interactions with itself and other proteins were also investigated, in the quest to unravel the underlying biological mechanism responsible for the molecular phenotype observed (i.e "augmented alternative exon selection").

4.3.4 TDP43-M323K mutant protein and RNA interactions

It can be hypothesised that if the mutant TDP43-M323K protein has higher affinity for RNA, levels of mutant M323K protein similar to wild type protein would lead to enhanced TDP43-dependent splicing activity, as the former (M323K protein) would be more frequently bound to RNA than the latter (wild type protein), thus enhancing its splicing activity.

The affinity between mutant TDP43-M323K and RNA was assessed by me at the ICGEB in Trieste, using the Electrophoretic Mobility Assay (EMSA) as the methodology of choice, both in its qualitative variant and a more recently developed semi-quantitative variant, which was used at that research institute to characterise the interactions between TDP43 and its RNA targets (Bhardwaj et al. 2013).

In the strictly qualitative EMSA, a constant amount of radioactively labelled UG_6 RNA is incubated with increasing amounts of recombinant TDP43 protein, both wild type and mutant, before gel electrophoresis is performed, after which the gel is dried. The results revealed by exposing a phosphor membrane to the dried, radioactive, gel.

In the semi-quantitative variant of the EMSA, the amount of recombinant TDP protein is kept constant, and the amount of UG_6 increased exponentially, thus enabling the plotting of a dissociation curve in order to find a dissociation constant (KD) which is situated at the midpoint of the dissociation curve.

Except for the addition of increasing amounts of radioactively labelled RNA instead of protein, the semi-quantitative EMSA protocol is similar to that of the qualitative EMSA.

The qualitative EMSA revealed that increasing concentrations of TDP-M323K followed a similar pattern of binding to TDP-WT; increasing concentrations of wild type and mutant M323K protein bound increasing amounts of UG_6 RNA (figure 4.12).

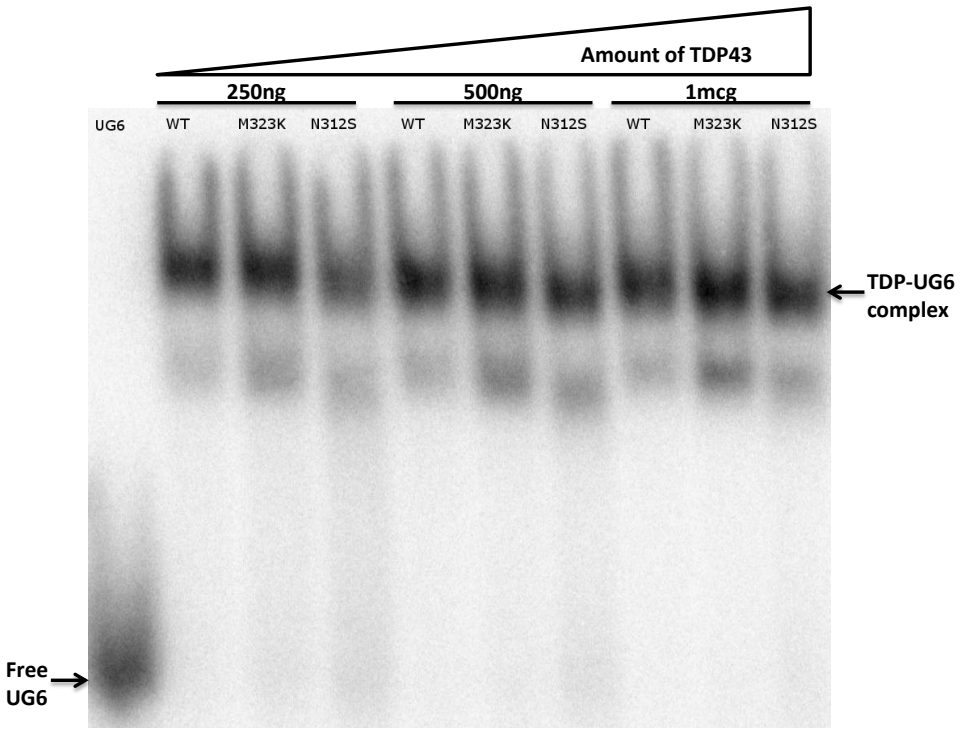

Figure 4.12: Qualitative EMSA for the assessment of TDP-WT and mutant TDP-M323K and its interactions with RNA (UG_6).
Increasing amounts of recombinant TDP-WT and mutant TDP-M323K lead to increased intensity of the shifted band, which represents the ability of increasing amounts of both proteins in binding increasing amounts of radioactively labelled UG_6 RNA. An additional C-terminal mutant protein (TDP43-N312S) was also tested, demonstrating a similar binding pattern to the other two TDP proteins and UG_6 RNA.

In the semi-quantitative EMSA, gels from repeated experiments were analysed individually and the results combined in order to perform statistical analysis of the dissociation constant between the recombinant TDP proteins and UG_6 RNA. The quantitative analysis of the shifted band was performed using *ImageJ* software and the dissociation curves were plotted in *Graphpad/Prism5* software (refer to chapter 2 for the equations used to find the KD values).

No significant differences were found between the dissociation constants (KD) of TDP-WT and RNA and mutant TDP-M323K and RNA (figure 4.13, refer to figure legend for mean KD values and statistical analysis).

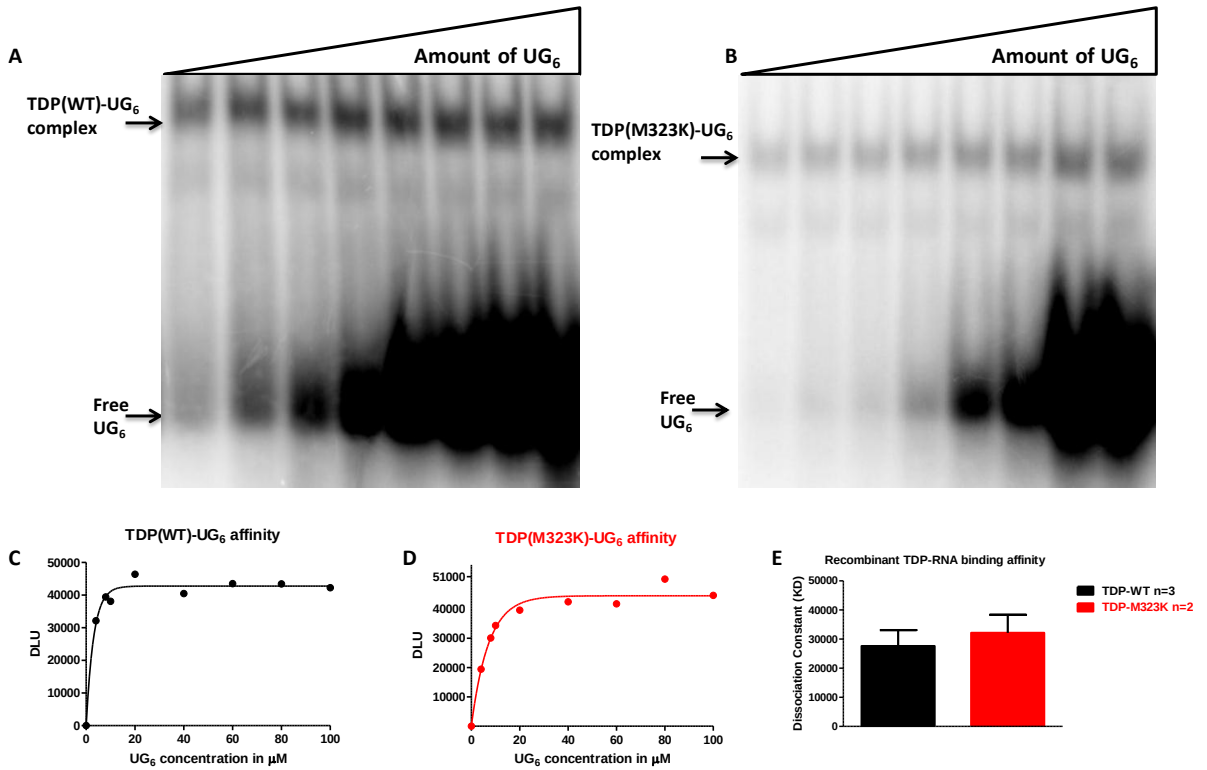

Figure 4.13: Semi-quantitative EMSA for the assessment of TDP-WT and TDP-M323K affinity for UG_6 RNA.
Gels from different semi-quantitative EMSA experiments were analysed together and statistical analysis performed on the dissociation constants (KD) between TDP-WT and UG_6 RNA and TDP-M323K and UG_6 RNA. Representative EMSA gels for (A) TDP-WT and (B) TDP-M323K are shown, as are the plotted dissociation curves (C and D). E) No statistical differences were found in the KDs (arbitrary units) of wild type and mutant TDP-M323K and UG_6 RNA (mean KD 27602.94±5487.31 for TDP-WT, n=3 and 32163.30±6191.93 for TDP-M323K, n=2 p=1, Anova).

Despite the fact that the "augmented alternative exon selection" observed in the TDP43-M323K mutant protein would be compatible with a higher affinity for RNA when compared to wild type, both the qualitative EMSA and the semi-quantitative EMSA failed to show an increase in the affinity between the mutant TDP-M323K protein and RNA.

It can be concluded that, at least using this methodology, there is no significant differences in the affinities between TDP-WT and TDP-M323K for RNA.

Nevertheless, given the high affinity of TDP43 for UG_6 RNA (Buratti & Baralle 2001), it should be remarked that the sensitivity of this technique for detecting decreasing affinity for RNA is far higher that its sensitivity for detecting an increase in affinity between the recombinant TDP43 proteins and UG_6 RNA (possible ways of overcoming this technical limitation will be expanded on in the discussion).

Being unable to attribute the "augmented alternative exon selection" to the interactions between TDP43-M323K and RNA, the possibility that the M323K mutation leads to changes in the protein-protein interactions and thus the phenotype observed also constitutes a reasonable hypothesis. The interactions between TDP43-M323K with other proteins and itself were investigated and the results are described in the next section.

4.3.5 TDP43-M323K protein-protein interactions

The interactions of TDP43-M323K with other cellular proteins was investigated using the GST overlay/far western methodology, which has previously been used to map the domains through which TDP interacts with other cellular proteins (Buratti et al. 2005; D'Ambrogio et al. 2009).

4.3.6 GST overlay/Far Western: qualitative assessment of the interactions between TDP43 and other cellular proteins

During the period I spent at the ICGEB in Trieste, in addition to performing the EMSA experiments described above, the interaction between TDP-M323K and other cellular proteins was assessed using the GST-overlay/far western.

Briefly, commercially acquired fractionated protein extracts from HeLA cells were subjected to polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane and the nuclear and cytoplasmic membrane bound proteins were subsequently incubated, individually, with GST-tagged recombinant TDP-WT or TDP-M323K protein.

After incubation with the recombinant TDP proteins, the membranes were washed and revealed with anti-GST antibody, thus revealing a band pattern corresponding to the nuclear and cytoplasmic proteins which were bound by the recombinant TDP proteins.

Hence, it was possible to assess, qualitatively, whether mutant TDP-M323K had different interactions with cellular proteins which would reveal itself as a different band pattern from that observed in the far western corresponding to the TDP-WT protein.

The GST overlay/far western, as succinctly described, involves the incubation of the membrane with recombinant protein and subsequent revelation with the use of an anti-GST antibody. The limitations in estimating protein concentrations, variability in the interaction of each incubation between the recombinant protein and the membrane bound proteins and the fact that immunoblotting constitutes, at best, a semi-quantitative technique, determine that inferences regarding the affinity of the binding interaction between recombinant TDP-protein and the membrane bound proteins cannot be made.

Qualitatively, the pattern of bands in the GST overlay/far-western obtained with TDP-WT and TDP-M323K proteins were analogous, suggesting that both proteins interact with the same subset of mammalian cellular proteins (figure 4.14).

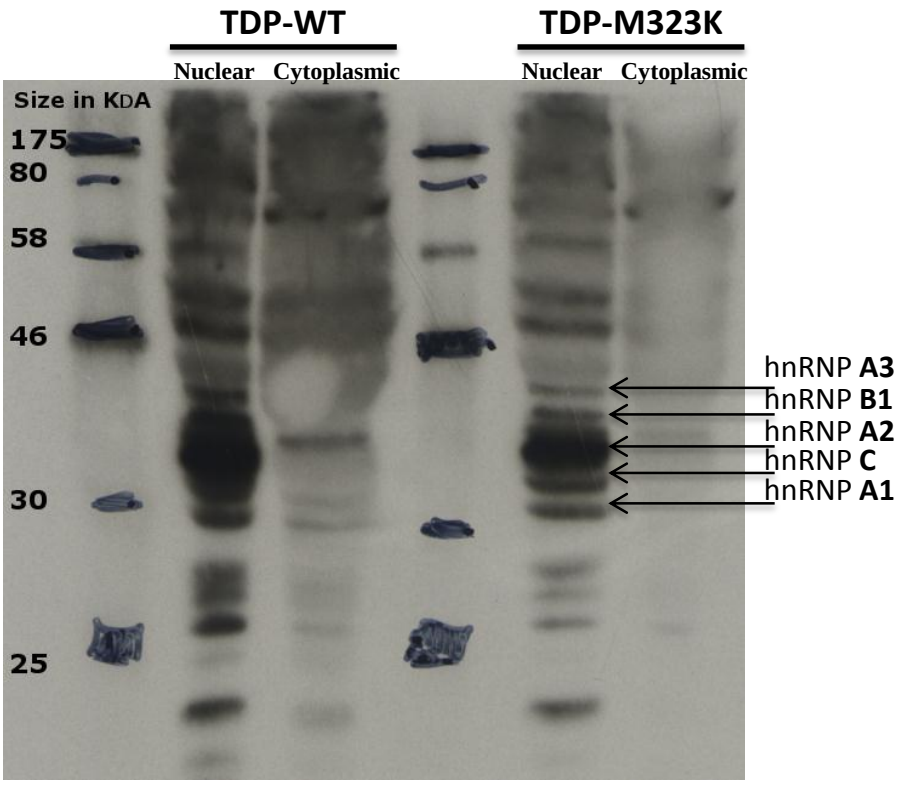

Figure 4.14: GST overlay/far western with TDP-WT and TDP43-M323K recombinant proteins
The pattern of bands resulting from the GST overlay/far western using recombinant TDP-WT or TDP-M323K and nuclear and cytoplasmic HeLa protein fractions is similar, suggesting that wild type TDP43 and mutant TDP43-M323K protein interact with the same subset of mammalian cellular proteins. The bands corresponding to known interactors of the same hnRNP family are indicated by arrows.

4.3.7 TDP43 homodimerisation in MEFs

The homodimerisation properties of the mutant TDP43 protein was determined by replicating the methodology employed by (Zhang et al. 2013) who used a covalent cross-linker to assess the homodimerisation properties of different exogenous TDP constructs which were over-expressed in HEK293T cells. In their study, the authors determined that homodimerisation of the exogenous TDP protein was positively correlated with its splicing activity (Zhang et al. 2013)

In contrast to the cited worked which used a cell line transfected with exogenous constructs, the homodimerisation experiment presented in this thesis was performed in MEFs harvested from *Tardbp*^{M323K} embryos at 14.5dpc. By using MEFs, the homodimerisation properties of TDP43-M323K was studied in the same tissue where splicing changes were identified and,

crucially, investigated the endogenous TDP43 proteins resulting from the isogenic expression of wild type *Tardbp* or *Tardbp*^{M323K} alleles.

Thus, the homodimerisation experiment was performed in a manner which minimised the possibility of the results being related to artefacts associated exogenous TDP43 overexpression.

MEFs cultured in 10cm dishes were harvested by trypsin digestion when confluence reached 70% and crosslinked with DSG, after which whole cell protein was extracted and polyacrylamide gel electrophoresis performed under denaturing conditions, given that the TDP43 homodimers had been covalently bound and were thus resistant to denaturation.

Immunoblotting with anti-TDP43 and anti-Tubulin antibodies was performed and band intensity quantified, with TDP43 homodimer being determined to be a band of the appropriate size which is only present in cross-linked samples and absent in negative controls (protein extracted from MEFs which were treated in the same manner but with cross-linking buffer without DSG were used as negative controls).

No significant differences were observed in the levels of TDP43 homodimer, free TDP43 or total TDP43 species across genotypes; the latter consisting of the summation of free TDP43 and TDP43 homodimer (figure 4.15, refer to figure for statistical analysis).

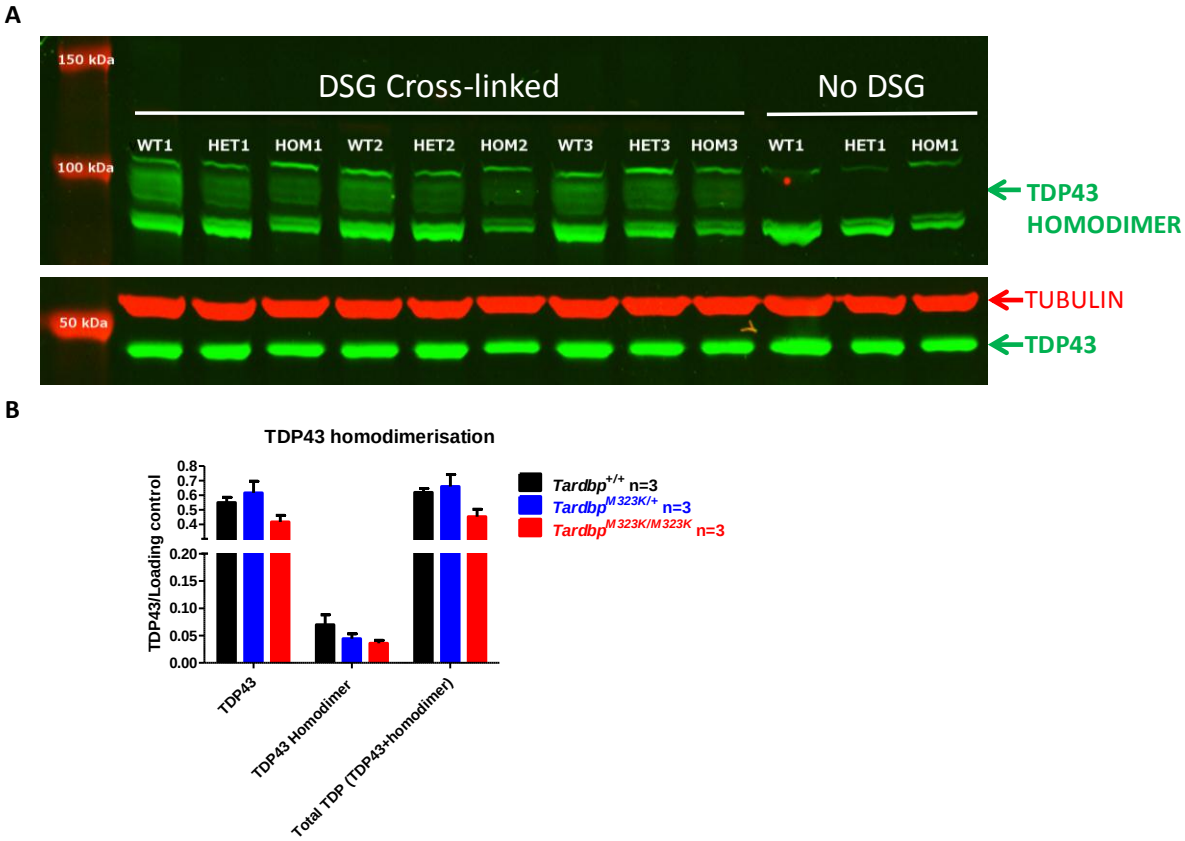

Figure 4.15: Homodimerisation properties of TDP43-WT and mutant TDP43-M323K
A) Immunoblot with protein extracted from MEFs (harvested from *Tardbp*^{M323K} embryos at 14.5dpc) treated with the covalent cross-linker DSG. B) Quantification revealed that no significant differences were seen in the levels of TDP43 homodimer (mean TDP43/Tubulin ratio 0.07±0.18 for *Tardbp*+/+, 0.04±0.09 for *Tardbp*$^{M323K/+}$ and 0.036±0.05 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes p>0.05 for every comparison, Anova and posthoc Bonferroni), free TDP43 (mean TDP43/Tubulin ratio 0.55±0.026 for *Tardbp*+/+, 0.62±0.80 for *Tardbp*$^{M323K/+}$ and 0.42±0.045 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes p>0.05 for every comparison, Anova and posthoc Bonferroni) and total TDP43 species (mean TDP43/Tubulin ratio 0.62$^{\pm}$0.027 for *Tardbp*+/+, 0.66±0.082 for *Tardbp*$^{M323K/+}$ and 0.45±0.050 for *Tardbp*$^{M323K/M323K}$, n=3 for all genotypes p>0.05 for every comparison, Anova and posthoc Bonferroni).

The studies investigating TDP43-M323K protein-protein interactions, as well as its homodimerisation properties have suggested that, within the limitations of the methodologies used (GST overlay/far western and DSG crosslinking and immunoblotting), the mutant M323K protein is able to interact with the same subset of mammalian cellular proteins when compared to wild type TDP43 and that its homodimerisation properties are similar to those of wild type TDP43, at isogenic levels of expression.

4.4 Summary of the Molecular characterisation of *Tardbp*^{M323K} mutants and further characterisation of *Tardbp*^{M323K} mutants

It has been previously shown that *Tardbp*^{F210I} leads to a "loss of normal splicing function" which recapitulates the changes observed on TDP43 knockdown due to a decrease affinity of the F210I mutant protein for RNA (Ricketts 2012).

A description of the molecular characterisation of *Tardbp*^{M323K} mutants, including a succinct description of the methodologies employed has been given in this chapter.

Additionally, work consisting of further characterisation of *Tardbp*^{F210I} mutants, specifically regarding the consequences of transcriptome changes at the whole genome level assessed by RNA-seq, has also been described.

In summary, the experimental results from original work presented in this thesis suggest that:

- *Tardbp*^{M323K} leads to "augmented alternative exon selection" in a large number of exons alternatively spliced by TDP43 (figures 4.1-4.5).
- Overall, at the genome-wide level, the splicing changes in the same alternative exons are seen in opposite directions in *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{F210I/F210I}$ (figure 4.6).
- *Tardbp*^{F210I} leads to upregulation in the expression of genes with long introns, at least in embryonic heads; the opposite effect observed to downregulation of genes with long introns upon TDP43 knockdown (figure 4.7).
- *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ have similar TDP43 protein levels to wild type controls (figure 4.8), which may be due to a functional autoregulation process which can be tested experimentally using tissue from *Tardbp*$^{Q101/M323K}$ mice.
- The M323K mutation does not affect the subcellular localisation of TDP43 (figures 4.9 and 4.10).

- In *Tardbp*$^{M323K/M323K}$ MEFs, there is significantly higher usage of the non-canonical Poly(A)4 site in the *Tardbp* transcript (figure 4.11) which may be a consequence of enhanced splicing of the crytic intron 7 by the mutant M323K protein; however its biological relevance remains obscure.
- Within the limitations of the methodologies employed, no significant differences were observed between the affinity of TDP43-M323K for RNA (figures 4.12 and 4.13).
- Using the GST overlay as a methodology, the results suggest wild type and mutant M323K TDP43 interact with the same subset of mammalian protein interactors (figure 4.14).
- The homodimerisation properties of TDP43-M323K are similar to those of the wild type protein when assessed by DSG-crosslinking and immunoblotting (figure 4.15).

In the next chapter, the results of the *in vivo* whole animal characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K} ENU mutants will be described.

Chapter 5.

In vivo **characterisation of** ***Tardbp*^{F210I}** **and** ***Tardbp*^{M323K}** **mutants**

In order to understand the consequences, at the whole animal level, of the F210I and M323K mutations in TDP43, comprehensive *in vivo* live animal phenotyping was performed.

As already described, TDP43 is ubiquitously expressed (Ou et al. 1995) and has a pivotal role in the aetiology and pathophysiology of ALS and FTLD-TDP (Neumann et al. 2006; Davidson et al. 2007; Cairns et al. 2007; Sreedharan et al. 2008; Van Deerlin et al. 2008; Chiang et al. 2012; Synofzik et al. 2014; Borroni et al. 2009). Therefore, phenotypic tests aimed at screening for behavioural deficits specifically associated with neurodegeneration, including deficits in motor function, learning and memory and endophenotypes of neuropsychiatric diseases. Additionally, general whole animal phenotypes which expand dysfunction in multiple body systems, including metabolism, were also performed given the ubiquitous nature of TDP43 expression.

A phenotyping pipeline (figure 5.1) was thus designed using establish paradigms and making use of tests which became validated in time to be performed in the determined experimental cohorts.

As described in the introduction and chapter 3, prenatal lethality of *Tardp*$^{F210I/F210I}$ and perinatal lethality of *Tardbp*$^{M323K/M323K}$, in the C57BL/6J background was observed, therefore disabling testing of homozygous animals for F210I and M323K mutations.

Within the limitation of not being able to test homozygous animals, testing cohorts were established early in the DPhil research using animals on the C57BL/6J background which had been backcrossed at least five generations onto the C57BL/6J inbred strain. Genetic modification in genes performing important functions in the CNS have been reported to display sexual dimorphism in laboratory mice (Fragkouli et al. 2006; O'Tuathaigh et al. 2006) and therefore it was decided to establish separate testing cohorts of male and female mice for *Tardbp*^{F210I} and *Tardbp*^{M323K} mutants.

Based on power calculations testing cohorts for each sex were established to have at least 10 mice per each genotype (mutant and wild type control littermates). Consequently, four experimental cohorts were established:

- Female *Tardbp*^{F210I} cohort (*Tardbp*$^{F210I/+}$ and *Tardbp*$^{+/+}$ animals)
- Male *Tardbp*^{F210I} cohort (*Tardbp*$^{F210I/+}$ and *Tardbp*$^{+/+}$ animals)
- Female *Tardbp*^{M323K} cohort (*Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$) animals
- Male *Tardbp*^{M323K} cohort (*Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$) animals

The *Tardbp*^{F210I} cohorts (both male and female), in the C57BL/6J background will be referred to as TDP-F210I-B6 and the *Tardbp*^{M323K} cohorts as TDP-M323K-B6.

In vivo phenotyping started with an excess of 80 animals and, in so far as possible, the phenotyping tests were performed in all animals of all cohorts at the same ages.

Nevertheless, during the lifetime of the animals in the experimental cohorts, periods in which the health status of individual mice prevented testing existed (e.g. due to fighting wounds) and, in addition, animals had to be euthanized according to the end humane points defined in the project licence or as determined by the discretionary judgement of named animal care or welfare officer. Humane endpoints in the project license included weight loss exceeding 20% of body weight and general signs that the animals were unwell (e.g. piloerect coat not improving in 24 hours, skin lesions not improving in 48 hours, abnormal gait interfering with ability to feed).

Consequently, the experimental numbers vary in the tests performed and will always be displayed in the graphs or data tables in this chapter; the error bars given in all graphs constitute positive and negative deflections of the Standard Error of the Mean (±SEM). When possible, and based on the power calculations, equivalent genetic animals (i.e. mutant and wild type *Tardbp*^{F210I} and *Tardbp*^{M323K} backcrossed at least five generations into the

C57BL/6J background) were added into the cohorts and tested to enable meaningful statistical analysis.

In addition, and also as described in chapter 3, it has been identified that, in a hybrid C57BL/6J-DBA genetic background, *Tardbp*$^{M323K/M323K}$ animals are viable. However, *Tardbp*$^{F210I/F210I}$, even on a C57BL/6J-DBA background, die before weaning age.

Regrettably, the viability of *Tardbp*$^{M323K/M323K}$ mice was detected late into the DPhil research timeline, and therefore it was not possible to proceed to full characterisation of these animals. Notwithstanding, despite the time constraints, it was possible to perform limited testing in a cohort of *Tardbp*^{M323K} animals on the hybrid background, designated TDP-M323K-B6-DBA.

The results from the phenotypic assessment, in conjunction with a brief description of the individual tests, will be given in this chapter, organised together into sections according to the body system and processes the tests assess.

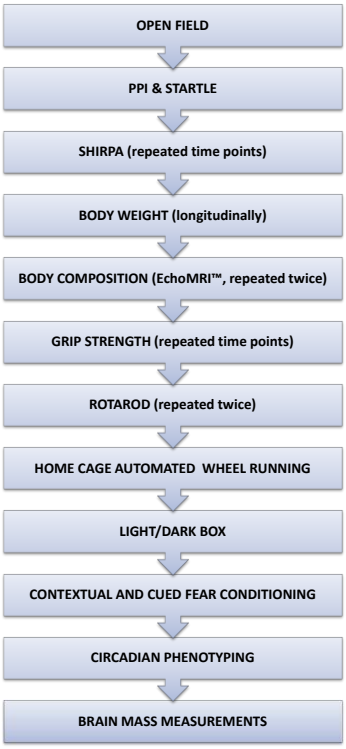

Figure 5.1: Phenotyping pipeline used in the characterisation for *Tardbp*^{F210I} and *Tardbp*^{M323K} mutants

5.1 Multisystem phenotyping of *Tardbp*^{F210I} and *Tardbp*^{M323K}

The *Tardbp* ENU mutants were assessed with a series of measurements and tests that can detect manifestations of biological dysfunction at the whole organism level. Whilst the data obtained in more general tests and measurements may not give an immediate indication of the underlying specific biological process affected by the mutation that leads to the observed phenotype, it is useful in determining whether mutations are functional *in vivo*.

Body weights of *Tardbp* mutants were collected longitudinally and their body composition assessed by Magnetic Resonance Imaging (EchoMRI™, which uses Quantitative Magnetic Resonance technology). Additionally, the modified SHIRPA procedure was also performed longitudinally and brain mass measured at one time point.

The influence of *Tardbp*^{F210I} and *Tardbp*^{M323K} in the lifespan/longevity of the animals was assessed (up to the age limit of 27 months), by performing Kaplan-Meier survival analysis in the TDP-F210I-B6 and TDP-M323K-B6 experimental cohorts.

5.1.1 Body weight and body composition

Weights were performed using common laboratory scales. Individual mice were placed inside a beaker in scales which took measures for five seconds, averaging the values to correct for the animal moving inside the beaker, thus determining the animal weight.

Body weight composition was measured in non-anaesthetised animals placed in a transparent plastic tube inside the Echo-MRI™ machine. Whole body MRI was performed three times and the average of the three tests used to determine the lean and fat body mass compositions.

5.1.1.1 TDP-F210I-B6

Tardbp$^{F210I/+}$ females weight less than wild type controls from 28 weeks onwards (mean weight 26.93±0.58 for *Tardbp*$^{+/+}$ and 24.95±0.30 for *Tardbp*$^{F210I/+}$, n=12 for both genotypes,

p=0.029, Anova with Benjamin Hochberg correction for multiple testing), a difference that continues to be observed throughout the majority of their lives (figure 5.2 and table 5.1, refer to table for mean weights and statistical analysis).

Moreover, when the difference in weights falls below significance, as observed from 76 weeks of age onwards (figure 5.2 and table 5.1), the lack of statistical significance is likely to be explained by the increased variability in weight amongst individual animals at such advanced age and, in addition, a reduction in the experimental numbers at the later time points.

The lower body weight observed in the *Tardbp*$^{F210I/+}$ females was attributable to a reduction in the percentage of fat body mass measured at 31 weeks (mean percentage of fat mass 19.62±1.28, n=11 for *Tardbp*$^{+/+}$ and 15.83±1.16, n=12 for *Tardbp*$^{F210I/+}$; p=0.040, Anova), an effect that was not reproduced at 100 weeks (mean percentage of fat mass 36.40±2.44, n=8 for *Tardbp*$^{+/+}$ and 31.0±3.51, n=7 for *Tardbp*$^{F210I/+}$; p=0.223, Anova) when the difference between the weights of the animals from the two genotypes also fails to reach statistical significance (figure 5.2 and table 5.1)

Tardbp$^{F210I/+}$ males have similar body weights and body composition to wild type controls throughout their lives (figure 5.3 and table 5.2, refer to table for mean weights and statistical analysis).

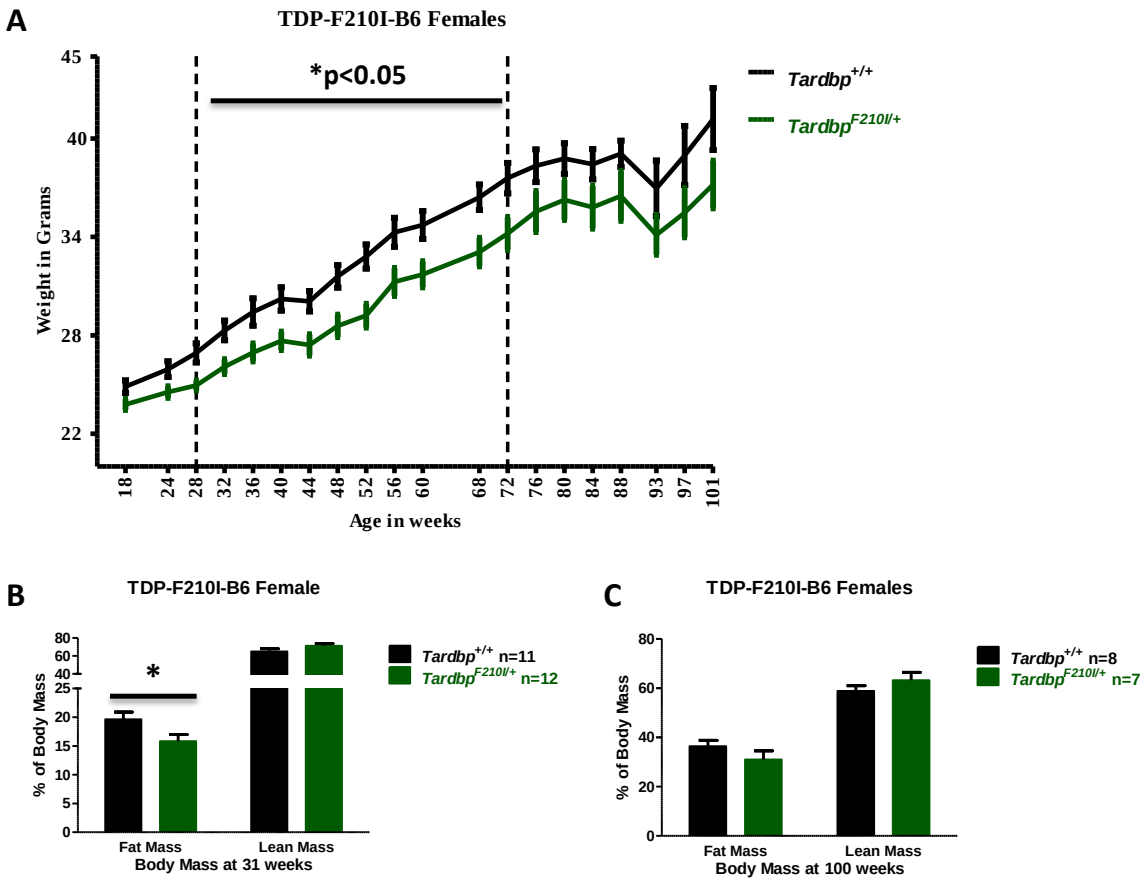

Figure 5.2: Weights and body composition of TDP-F210I-B6 females

A) *Tardbp*$^{F210I/+}$ females weight less from 28 weeks of age and throughout most of its lives when compared to *Tardbp*$^{+/+}$ (*p<0.05, refer to table 5.1 for statistical analysis).

B) At 31 weeks of age, the lower body weight from *Tardbp*$^{F210I/+}$ is attributable to reduced percentage of fat mass (*p=0.040; refer to main text for mean percentages of fat mass; mean percentage of lean mass 64.98±3.10 for *Tardbp*$^{+/+}$ and 71.31±2.45 for *Tardbp*$^{F210I/+}$; p=0.120, Anova), an effect which is not observed at (B) 100 weeks when weights are not significantly different between genotypes (refer to table 5.1 for mean weights and statistical analysis and to main text for mean percentages of fat mass and statistical analysis; mean percentage of lean mass 58.91±2.15 for *Tardbp*$^{+/+}$ and 63.21±3.25 for *Tardbp*$^{F210I/+}$; p=0.279, Anova). Error bars on all graphs are ±SEM.

Age (Weeks)	Genotype	Mean Weight (Grams)	SEM	n	p (Anova)	p (FDR corrected)
18	*Tardbp* $^{+/+}$	24.87	0.38	12	0.042	0.068
	Tardbp $^{F210I/+}$	23.79	0.32	12		
24	*Tardbp* $^{+/+}$	25.93	0.48	12	0.029	0.051
	Tardbp $^{F210I/+}$	24.54	0.35	12		
28	*Tardbp* $^{+/+}$	26.93	0.58	12	0.006	**0.029**
	Tardbp $^{F210I/+}$	24.95	0.30	12		
32	*Tardbp* $^{+/+}$	28.30	0.60	12	0.007	**0.029**
	Tardbp $^{F210I/+}$	26.09	0.45	12		
36	*Tardbp* $^{+/+}$	29.42	0.83	11	0.019	**0.044**
	Tardbp $^{F210I/+}$	26.96	0.54	12		
40	*Tardbp* $^{+/+}$	30.22	0.71	11	0.010	**0.030**
	Tardbp $^{F210I/+}$	27.66	0.56	12		
44	*Tardbp* $^{+/+}$	30.08	0.62	11	0.007	**0.029**
	Tardbp $^{F210I/+}$	27.41	0.64	12		
48	*Tardbp* $^{+/+}$	31.59	0.69	11	0.005	**0.029**
	Tardbp $^{F210I/+}$	28.58	0.68	12		
52	*Tardbp* $^{+/+}$	32.80	0.73	11	0.002	**0.029**
	Tardbp $^{F210I/+}$	29.20	0.72	12		
56	*Tardbp* $^{+/+}$	34.28	0.88	11	0.023	**0.044**
	Tardbp $^{F210I/+}$	31.26	0.85	12		
60	*Tardbp* $^{+/+}$	34.73	0.85	11	0.017	**0.044**
	Tardbp $^{F210I/+}$	31.72	0.79	12		
68	*Tardbp* $^{+/+}$	36.42	0.77	10	0.010	**0.030**
	Tardbp $^{F210I/+}$	33.08	0.86	11		
72	*Tardbp* $^{+/+}$	37.58	0.93	10	0.023	**0.044**
	Tardbp $^{F210I/+}$	34.20	0.99	11		
76	*Tardbp* $^{+/+}$	38.34	1.00	10	0.098	0.147
	Tardbp $^{F210I/+}$	35.55	1.25	10		
80	*Tardbp* $^{+/+}$	38.77	0.94	10	0.118	0.155
	Tardbp $^{F210I/+}$	36.26	1.21	10		
84	*Tardbp* $^{+/+}$	38.43	0.92	10	0.116	0.155
	Tardbp $^{F210I/+}$	35.81	1.29	10		
88	*Tardbp* $^{+/+}$	39.07	0.80	10	0.151	0.176
	Tardbp $^{F210I/+}$	36.50	1.51	10		
93	*Tardbp* $^{+/+}$	36.96	1.69	9	0.192	0.192
	Tardbp $^{F210I/+}$	34.12	1.21	9		
97	*Tardbp* $^{+/+}$	38.96	1.81	5	0.166	0.176
	Tardbp $^{F210I/+}$	35.47	1.49	6		
101	*Tardbp* $^{+/+}$	41.18	1.89	8	0.126	0.155
	Tardbp $^{F210I/+}$	37.19	1.45	7		
113	*Tardbp* $^{+/+}$	40.49	2.89	5	0.168	0.176
	Tardbp $^{F210I/+}$	35.86	0.99	5		

Table 5.1: Mean body weights of TDP-F210I-B6 females and statistical analysis
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the age and p-value numbers in red.

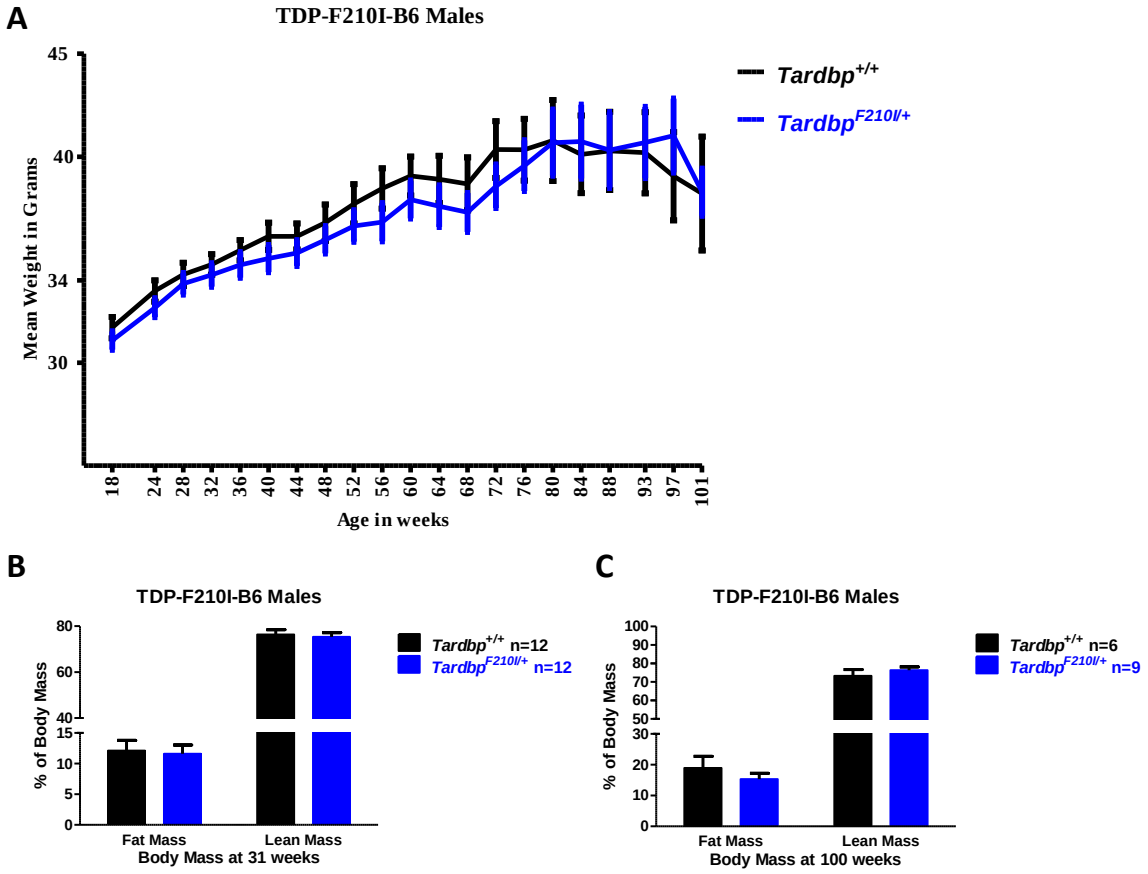

Figure 5.3: Weights and body composition for TDP-F210I-B6 Males
A) *Tardbp*$^{F210I/+}$ males have similar body weights to *Tardbp*$^{+/+}$ males throughout life (refer to table 5.2 for statistical analysis), as well as similar body composition at (B) 31 weeks (mean percentage of fat mass 12.13±1.66 for *Tardbp*+/+ and 11.62±1.42 for *Tardbp*$^{F210I/+}$; p=0.815, Anova; mean percentage lean mass 76.36±20.6 for *Tardbp*+/+ and 75.33±1.79 for *Tardbp*$^{F210I/+}$; p=0.711, Anova) and (C) 100 weeks. (mean percentage of fat mass 18.97±3.77 for *Tardbp*$^{+/+}$ and 15.33±1.85 for *Tardbp*$^{F210I/+}$; p=0.355, Anova; mean percentage lean mass 73.27±8.50 for *Tardbp*+/+ and 76.38±5.24 for *Tardbp*$^{F210I/+}$; p=0.393, Anova). Error bars on all graphs are ±SEM.

Age (Weeks)	Genotype	Mean Weight (Grams)	SEM	n	p (Anova)	p (FDR corrected)
18	*Tardbp* $^{+/+}$	31.71	0.52	12	0.369	0.759
	Tardbp $^{F210I/+}$	31.08	0.44	12		
24	*Tardbp* $^{+/+}$	33.48	0.52	12	0.250	0.759
	Tardbp $^{F210I/+}$	32.67	0.45	12		
28	*Tardbp* $^{+/+}$	34.29	0.56	12	0.561	0.785
	Tardbp $^{F210I/+}$	33.83	0.52	12		
32	*Tardbp* $^{+/+}$	34.77	0.50	12	0.506	0.759
	Tardbp $^{F210I/+}$	34.26	0.56	12		
36	*Tardbp* $^{+/+}$	35.46	0.49	11	0.379	0.759
	Tardbp $^{F210I/+}$	34.75	0.62	11		
40	*Tardbp* $^{+/+}$	36.14	0.66	11	0.261	0.759
	Tardbp $^{F210I/+}$	35.06	0.66	11		
44	*Tardbp* $^{+/+}$	36.13	0.63	11	0.373	0.759
	Tardbp $^{F210I/+}$	35.32	0.62	11		
48	*Tardbp* $^{+/+}$	36.80	0.88	8	0.450	0.759
	Tardbp $^{F210I/+}$	35.97	0.66	11		
52	*Tardbp* $^{+/+}$	37.72	0.95	8	0.384	0.759
	Tardbp $^{F210I/+}$	36.64	0.76	11		
56	*Tardbp* $^{+/+}$	38.46	0.98	8	0.248	0.759
	Tardbp $^{F210I/+}$	36.82	0.92	11		
60	*Tardbp* $^{+/+}$	39.07	0.94	8	0.406	0.759
	Tardbp $^{F210I/+}$	37.92	0.92	11		
64	*Tardbp* $^{+/+}$	38.91	1.14	7	0.407	0.759
	Tardbp $^{F210I/+}$	37.59	1.01	10		
68	*Tardbp* $^{+/+}$	38.68	1.29	8	0.387	0.759
	Tardbp $^{F210I/+}$	37.30	0.95	11		
72	*Tardbp* $^{+/+}$	40.35	1.38	8	0.309	0.759
	Tardbp $^{F210I/+}$	38.57	1.05	10		
76	*Tardbp* $^{+/+}$	40.34	1.50	9	0.692	0.908
	Tardbp $^{F210I/+}$	39.57	1.23	12		
80	*Tardbp* $^{+/+}$	40.79	1.97	9	0.968	0.990
	Tardbp $^{F210I/+}$	40.69	1.60	12		
84	*Tardbp* $^{+/+}$	40.12	1.89	8	0.818	0.989
	Tardbp $^{F210I/+}$	40.74	1.79	11		
88	*Tardbp* $^{+/+}$	40.29	1.90	9	0.990	0.989
	Tardbp $^{F210I/+}$	40.32	1.84	11		
93	*Tardbp* $^{+/+}$	40.20	1.97	9	0.853	0.990
	Tardbp $^{F210I/+}$	40.69	1.73	11		
97	*Tardbp* $^{+/+}$	39.06	2.14	9	0.485	0.759
	Tardbp $^{F210I/+}$	41.03	1.79	11		
101	*Tardbp* $^{+/+}$	38.22	2.76	6	0.981	0.990
	Tardbp $^{F210I/+}$	38.28	1.14	8		

Table 5.2: Mean body weights of TDP-F210I-B6 males and statistical analysis
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing.

5.1.1.2 TDP-M323K-B6

Tardbp$^{M323K/+}$ females had similar weights to *Tardbp*$^{+/+}$ controls throughout their lives (figure 5.4 and table 5.3, refer to table from mean body weights and statistical analysis), as well as similar body composition measured at 31 and 100 weeks (figure 5.4, refer to figure legend for statistical analysis).

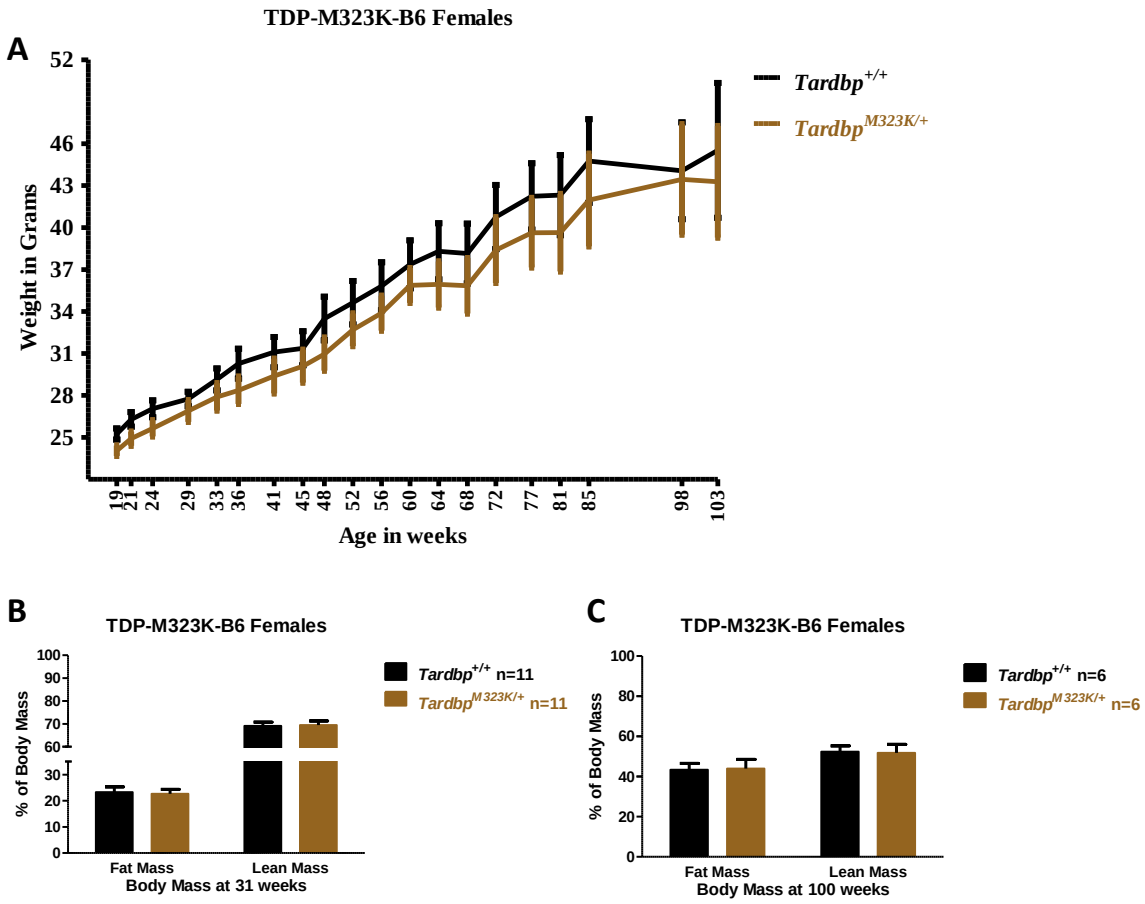

Figure 5.4: Weights and body composition of TDP-M323K-B6 females.
A) *Tardbp*$^{M323K/+}$ females have similar body weights to *Tardbp*$^{+/+}$ females throughout life (refer to table 5.3 for statistical analysis). Body composition was similar at (B) 31 weeks (mean percentage of lean mass 69.11±1.67 for *Tardbp*$^{+/+}$ and 69.49±1.79 for *Tardbp*$^{M323K/+}$; p=0.880, Anova; mean percentage of fat mass 23.30±1.99 for *Tardbp*$^{+/+}$ and 22.70±1.69 for *Tardbp*$^{M323K/+}$; p=0.820, Anova) and (B) 100 weeks (mean percentage of lean mass 52.37±2.91 for *Tardbp*$^{+/+}$ and 51.82±4.22 for *Tardbp*$^{M323K/+}$; p=0.919, Anova; mean percentage of fat mass 43.38±3.17 for *Tardbp*$^{+/+}$ and 44.00±4.53 for *Tardbp*$^{M323K/+}$; p=0.913, Anova). Error bars on all graphs are ±SEM.

Age (Weeks)	Genotype	Mean Weight (Grams)	SEM	n	p (Anova)	p (FDR corrected)
19	*Tardbp* $^{+/+}$	25.24	0.39	11	0.038	0.600
	Tardbp $^{M323K/+}$	24.04	0.37	11		
21	*Tardbp* $^{+/+}$	26.29	0.52	10	0.072	0.600
	Tardbp $^{M323K/+}$	24.91	0.51	10		
24	*Tardbp* $^{+/+}$	27.05	0.61	11	0.110	0.600
	Tardbp $^{M323K/+}$	25.65	0.58	11		
29	*Tardbp* $^{+/+}$	27.75	0.52	11	0.371	0.600
	Tardbp $^{M323K/+}$	26.89	0.77	11		
33	*Tardbp* $^{+/+}$	29.14	0.79	11	0.333	0.600
	Tardbp $^{M323K/+}$	27.90	0.98	11		
36	*Tardbp* $^{+/+}$	30.27	1.07	11	0.197	0.600
	Tardbp $^{M323K/+}$	28.37	0.95	11		
41	*Tardbp* $^{+/+}$	31.10	1.08	11	0.308	0.600
	Tardbp $^{M323K/+}$	29.38	1.24	10		
45	*Tardbp* $^{+/+}$	31.37	1.22	11	0.458	0.600
	Tardbp $^{M323K/+}$	30.09	1.18	11		
48	*Tardbp* $^{+/+}$	33.50	1.56	9	0.206	0.600
	Tardbp $^{M323K/+}$	30.95	1.19	10		
52	*Tardbp* $^{+/+}$	34.62	1.55	11	0.330	0.600
	Tardbp $^{M323K/+}$	32.69	1.16	11		
56	*Tardbp* $^{+/+}$	35.81	1.72	11	0.374	0.600
	Tardbp $^{M323K/+}$	33.88	1.24	11		
60	*Tardbp* $^{+/+}$	37.37	1.79	11	0.526	0.600
	Tardbp $^{M323K/+}$	35.86	1.49	11		
64	*Tardbp* $^{+/+}$	38.30	2.01	11	0.373	0.600
	Tardbp $^{M323K/+}$	35.93	1.65	11		
68	*Tardbp* $^{+/+}$	38.15	2.14	11	0.440	0.600
	Tardbp $^{M323K/+}$	35.84	1.98	10		
72	*Tardbp* $^{+/+}$	40.75	2.29	10	0.482	0.600
	Tardbp $^{M323K/+}$	38.39	2.35	9		
77	*Tardbp* $^{+/+}$	42.23	2.37	10	0.460	0.600
	Tardbp $^{M323K/+}$	39.63	2.48	9		
81	*Tardbp* $^{+/+}$	42.32	2.86	10	0.511	0.600
	Tardbp $^{M323K/+}$	39.64	2.77	9		
85	*Tardbp* $^{+/+}$	44.75	2.99	10	0.540	0.600
	Tardbp $^{M323K/+}$	41.97	3.31	9		
98	*Tardbp* $^{+/+}$	44.06	3.46	9	0.910	0.910
	Tardbp $^{M323K/+}$	43.45	3.94	6		
103	*Tardbp* $^{+/+}$	45.52	4.82	6	0.727	0.765
	Tardbp $^{M323K/+}$	43.28	3.99	6		

Table 5.3: Mean body weights of TDP-M323K-B6 females and statistical analysis
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing.

Interestingly, *Tardbp*$^{M323K/+}$ males were observed to tend to weigh generally more than *Tardbp*$^{+/+}$ mice (figure 5.5 and table 5.4, refer to table for statistical analysis). However, the difference in weight never reached statistical significance (table 5.4).

An additional observation about the weight data from *Tardbp*$^{M323K/+}$ males is that from 81 weeks of age onwards, *Tardbp*$^{+/+}$ males lost large amounts of weight (figure 5.5, table 5.4) and therefore, the difference in weight, despite continuing to remain below statistical significance after correction for multiple testing (table 5.4), is likely to be exacerbated by the weight loss in the wild type control cohort which may not be representative of the body weight of wild type animals at the population level.

In agreement with the weight data, the percentage of fat mass of *Tardbp*$^{M323K/+}$ males is higher than *Tardbp*$^{+/+}$ males at 31 weeks (figure 5.5), despite failing to reach statistical significance (mean percentage of fat mass 11.12±1.37 for *Tardbp*$^{+/+}$, n=12 and 15.40±1.95 for *Tardbp*$^{M323K/+}$, n=10; p=0.10, Anova).

At 100 weeks *Tardbp*$^{M323K/+}$ have a higher percentage of fat mass when compared to *Tardbp*$^{+/+}$(figure 5.5), which reaches statistical significance (15.81±1.67 for *Tardbp*$^{+/+}$, n=6, and 27.47±2.63 for *Tardbp*$^{M323K/+}$ n=8, p=0.005, Anova) and lower percentage of lean mass (75.29±1.74 for *Tardbp*$^{+/+}$, n=6, and 66.21±2.51 for *Tardbp*$^{M323K/+}$ n=8; p=0.017, Anova).

However, as discussed for the weights, the observation of different body compositions at 100 weeks is more likely to be related with the abnormal weight loss in the wild type control animals and decreasing sample numbers than an intrinsic genotype to phenotype effect.

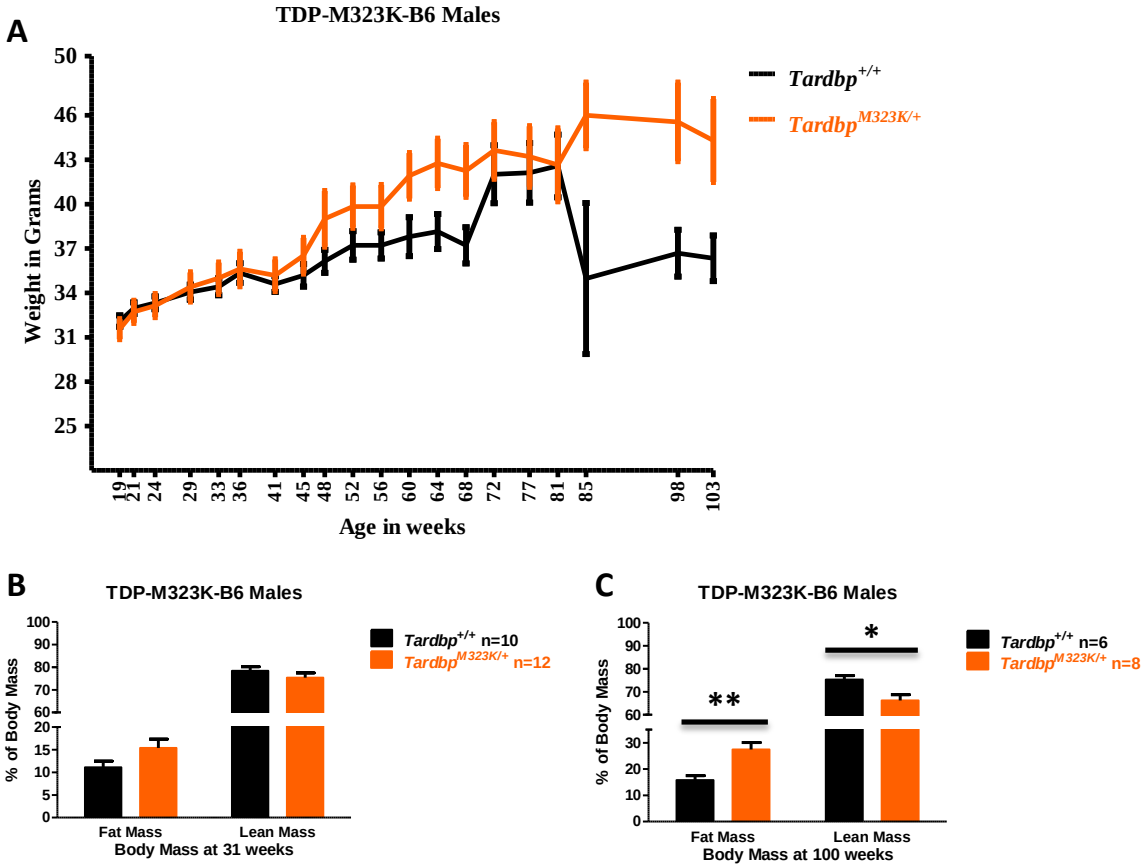

Figure 5.5: Weights and body composition of TDP-M323K-B6 males

A) *Tardbp*$^{M323K/+}$ males do not have significantly different weights from wild type controls throughout life (refer to table 5.4 for statistical analysis). From 81 weeks of age, *Tardbp*$^{+/+}$ in the experimental cohort lose large amounts of weight, which is unlikely to represent the body weight in the population of wild type C57BL/6J laboratory mice. At 85 weeks of age 6 *Tardbp*$^{+/+}$ animals had to be sacrificed due to losing over 20% of body weight, suggesting that the mice were not healthy and thus not representative of the population.

B) At 31 weeks of age the body composition is identical in heterozygous and wild type males (refer to main text for mean percentage of fat mass and for statistical analysis; mean percentage of lean mass 78.43±1.75 for *Tardbp*$^{+/+}$ and 75.39±2.10 for *Tardbp*$^{M323K/+}$; p=0.291, Anova).

C) At 100 weeks *Tardbp*$^{M323K/+}$ have significantly higher percentage of fat mass (**p=0.005) and significantly lower percentage of lean mass (*p=0.017), which is most likely to be a result from abnormal weight loss in the *Tardbp*$^{+/+}$ males (refer to main text for mean percentages of fat and lean mass and for statistical analysis). Error bars in all graphs are ±SEM.

Age (Weeks)	Genotype	Mean Weight (Grams)	SEM	n	p (Anova)	p (FDR corrected)
19	*Tardbp* $^{+/+}$	32.10	0.38	10	0.508	0.892
	Tardbp $^{M323K/+}$	31.55	0.67	12		
21	*Tardbp* $^{+/+}$	32.96	0.39	10	0.787	0.892
	Tardbp $^{M323K/+}$	32.72	0.74	12		
24	*Tardbp* $^{+/+}$	33.32	0.43	10	0.847	0.892
	Tardbp $^{M323K/+}$	33.14	0.78	12		
29	*Tardbp* $^{+/+}$	34.05	0.51	10	0.767	0.892
	Tardbp $^{M323K/+}$	34.40	0.98	12		
33	*Tardbp* $^{+/+}$	34.42	0.57	10	0.657	0.892
	Tardbp $^{M323K/+}$	34.99	1.06	12		
36	*Tardbp* $^{+/+}$	35.34	0.66	10	0.847	0.892
	Tardbp $^{M323K/+}$	35.62	1.16	12		
41	*Tardbp* $^{+/+}$	34.60	0.53	9	0.662	0.892
	Tardbp $^{M323K/+}$	35.19	1.07	12		
45	*Tardbp* $^{+/+}$	35.19	0.76	7	0.441	0.882
	Tardbp $^{M323K/+}$	36.52	1.20	12		
48	*Tardbp* $^{+/+}$	36.14	0.78	8	0.179	0.398
	Tardbp $^{M323K/+}$	39.01	1.87	8		
52	*Tardbp* $^{+/+}$	36.32	0.97	9	0.159	0.398
	Tardbp $^{M323K/+}$	39.03	1.42	12		
56	*Tardbp* $^{+/+}$	37.21	0.88	9	0.174	0.398
	Tardbp $^{M323K/+}$	39.83	1.46	12		
60	*Tardbp* $^{+/+}$	37.80	1.32	9	0.066	0.220
	Tardbp $^{M323K/+}$	41.92	1.53	12		
64	*Tardbp* $^{+/+}$	38.15	1.19	9	0.049	0.220
	Tardbp $^{M323K/+}$	42.76	1.67	12		
68	*Tardbp* $^{+/+}$	37.21	1.22	9	0.040	0.220
	Tardbp $^{M323K/+}$	42.25	1.74	12		
72	*Tardbp* $^{+/+}$	42.02	1.96	12	0.561	0.892
	Tardbp $^{M323K/+}$	43.64	1.91	12		
77	*Tardbp* $^{+/+}$	42.12	2.01	12	0.704	0.892
	Tardbp $^{M323K/+}$	43.22	2.04	13		
81	*Tardbp* $^{+/+}$	42.58	2.13	12	0.984	0.984
	Tardbp $^{M323K/+}$	42.65	2.46	13		
85	*Tardbp* $^{+/+}$	34.97	5.11	12	0.060	0.220
	Tardbp $^{M323K/+}$	46.00	2.21	12		
98	*Tardbp* $^{+/+}$	36.69	1.59	6	0.022	0.220
	Tardbp $^{M323K/+}$	45.56	2.65	8		
103	*Tardbp* $^{+/+}$	36.34	1.54	6	0.043	0.220
	Tardbp $^{M323K/+}$	44.31	2.80	8		

Table 5.4: Mean body weights of TDP-M323K-B6 males and statistical analysis
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the age and p-value numbers in red.

5.1.1.3 TDP-M323K-B6-DBA

Weights were also collected for TDP-M323K-B6-DBA male animals. Consistently with the results for heterozygous animals in the C57BL/6J background, a trend for *Tardbp*$^{M323K/+}$ males to weight more than wild types has also been observed in the hybrid C57BL/6J-DBA background at the early five time points measured. In contrast, however, *Tardbp*$^{M323K/M323K}$ weigh less from 19 weeks onwards, a difference which is significantly different from *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ at 29 and 32 weeks (p=0.005 for both time points) when the strict Bonferroni correction is applied to the Anova test and the significance is corrected for multiple testing (figure 5.6 and table 5.5, refer to table for mean weights and statistical analysis).

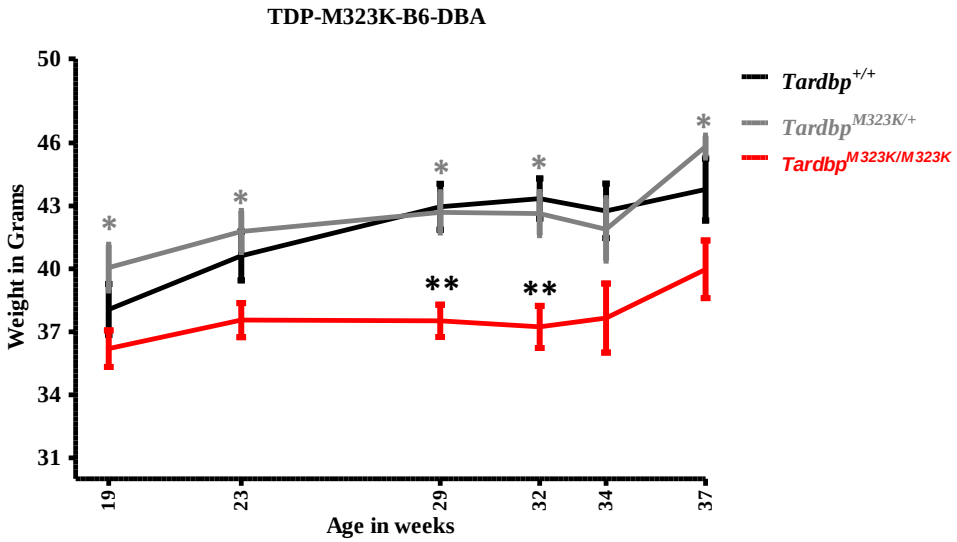

Figure 5.6: Weights of TDP-M323K-B6-DBA males.
Tardbp$^{M323K/M323K}$ mice have lower bodyweight from 19 weeks of age onwards, a difference which becomes significant at 29 and 32 weeks (**p=0.005) when compared to wild type controls (refer to table 5.5 for mean weights and statistical analysis). Remarkably, *Tardbp*$^{M323K/+}$ display a trend towards higher body weight when compared to *Tardbp*$^{+/+}$ which fails to reach statistical significance but is sufficient to determine that the difference towards *Tardbp*$^{M323K/M323K}$ animal weights is significantly statistically different at more time points, specifically every time point measured except at 34 weeks (*p<0.05 for comparison between mean weights of *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$; refer to table 5.5 for statistical analysis). Black asterisks denote statistically significant differences between *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/M323K}$ and grey asterisks statistically significant differences between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$. Error bars in all graphs are ±SEM.

Age (weeks)	Genotype	Mean Weight (Grams)	SEM	n	p (Anova & posthoc Bonferroni)	p (FDR corrected)
19	*Tardbp* $^{M323K/M323K}$	36.20	0.87	13		
	Tardbp $^{+/+}$	38.07	1.22	9	0.634	0.634
	Tardbp $^{M323K/+}$	40.06	1.08	10	0.033	**0.039**
23	*Tardbp* $^{M323K/M323K}$	37.56	0.81	16		
	Tardbp $^{+/+}$	40.62	1.16	12	0.104	0.154
	Tardbp $^{M323K/+}$	41.78	0.98	15	0.008	**0.016**
29	*Tardbp* $^{M323K/M323K}$	37.53	0.77	16		
	Tardbp $^{+/+}$	42.96	1.10	12	0.001	**0.005**
	Tardbp $^{M323K/+}$	42.69	0.95	18	0.001	**0.003**
32	*Tardbp* $^{M323K/M323K}$	37.24	1.00	10		
	Tardbp $^{+/+}$	43.35	0.96	10	0.002	**0.005**
	Tardbp $^{M323K/+}$	42.64	1.02	16	0.002	**0.006**
34	*Tardbp* $^{M323K/M323K}$	37.66	1.64	7		
	Tardbp $^{+/+}$	42.76	1.30	8	0.107	0.154
	Tardbp $^{M323K/+}$	41.88	1.48	11	0.180	0.180
37	*Tardbp* $^{M323K/M323K}$	39.98	1.37	5		
	Tardbp $^{+/+}$	43.78	1.48	6	0.129	0.154
	Tardbp $^{M323K/+}$	45.83	0.51	6	0.012	**0.018**

Table 5.5: Mean body weights of TDP-M323K-B6-DBA males and statistical analysis
Please note that the p-values given are for the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$ and the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{M323K/+}$ which are given in the table row corresponding to the wild type and heterozygous animals. Hence, the line in the table referent to the homozygous genotype (*Tardbp*$^{M323k/M323K}$) does not have a p value; p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the age and p-value numbers in red.

In summary, the *Tardbp*$^{F210I/+}$ females have significant lower body weight than wild type controls, which is attributable to lower fat body mass at least at 31 weeks (figure 5.2 table 5.1).

Tardbp$^{M323K/+}$ males, both on the C57BL/6J and the hybrid background, tend to have a higher body weight when compared to wild type controls, which at the time points tested failed to reach statistical significance (figure 5.5, table 5.4, figure 5.6 and table 5.5).

Tardbp$^{M323K/M323K}$ males have lower body weight when compared to *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$, which is statistically significant at least at two time points measured, and will possible reveal itself significant at additional time points with further longitudinal testing and analysis (figure 5.6 and table 5.5).

5.1.2 Brain mass: TDP-F210I-B6 and TDP-M323K-B6

TDP43 has been shown to regulate the expression of a wide range of genes (Polymenidou et al. 2011; Tollervey et al. 2011) and reduced protein levels lead, *in vivo*, to microcephaly (Yang et al. 2014). At a late stage for the *Tardbp*^{M323K} and *Tardbp*^{F210I} mutants, with ages between 24-26 months, at the time when the animals were sacrificed to harvest tissues, the whole bodies were weighted as were the brains immediately after dissection.

The three main events determining brain mass consist of neurogenesis, which occurs by enlarge prenatally, myelination, which occurs in the early post-natal period, and age-related degeneration. Brain mass measurements in individual animals can be expected to be a more reliable and relevant parameter of a genotype to phenotype effect (*Tardbp* mutations and brain size) at the age of 24-26 months than brain mass to weight ratio, given that weight is very variable at this advanced age in a laboratory mouse's lifespan, as demonstrated in figures 5.2 to 5.6.

5.1.2.1 TDP-F210I-B6

Tardbp$^{F210I/+}$ females had similar brain masses to wild type controls (figure 5.7 and table 5.6, refer to table for statistical analysis). The brain mass to body weight ratio was, however, significantly statistically different, being higher in the heterozygous animals (1.14±0.79 in *Tardbp*$^{+/+}$and 1.72±0.74 in *Tardbp*$^{F210I/+}$, n=5 for both genotypes; p=0.041, Anova).

Nevertheless, the body weights of the animals tested were markedly different with *Tardbp*$^{F210I/+}$ females weighting significantly less (mean body weight 34.35±1.61 in *Tardbp*$^{+/+}$and 28.02±0.83 in *Tardbp*$^{F210I/+}$, n=5 for both genotypes; p=0.015, Anova). Thus, the interpretation of these results cannot conclude that there are significant differences in the brain mass of TDP-F210I-B6 females at 24-26 months, with the observed difference in the ratio of brain to body mass being determined by the difference in weight between heterozygous and wild type animals.

Tardbp$^{F210I/+}$ males have comparable brain masses and ratios between brain mass and body weights to *Tardbp*$^{+/+}$ (figure 5.7 and table 5.6, refer to table for statistical analysis).

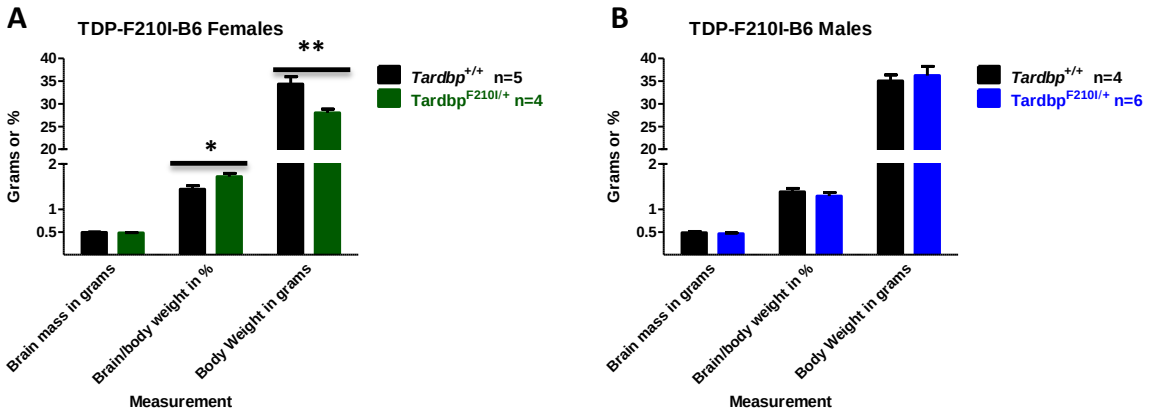

Figure 5.7: Brain mass of TDP-F210I-B6 animals 24-26 months of age
A) *Tardbp*$^{F210I/+}$ females had identical absolute brain mass, in grams to wild type controls. Body weights were significantly different (**p=0.015) causing the difference in the brain mass to body weight ratio to become significant between the two genotypes (*p=0.041). Refer to main text and table 5.6 for mean values of the parameters measured and statistical analysis.
B) In *Tardbp*$^{F210I/+}$ males all parameters measured were identical to wild type controls (refer to table 5.6 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{F210I/+}$			
		Mean	SEM	n	Mean	SEM	n	p (Anova)
Females	Brain Mass	0.492	0.007	5	0.48	0.009	4	0.334
	Brain Mass/Body weight	1.444	0.079	5	1.72	0.074	4	**0.015**
	Body weight	34.356	1.61	5	28.02	0.828	4	0.041
Males	Brain Mass	0.485	0.029	4	0.463	0.019	6	0.529
	Brain Mass/Body weight	1.385	0.077	4	1.293	0.078	6	0.679
	Body weight	35.045	1.349	4	36.228	2.035	6	0.446

Table 5.6: Brain mass of TDP-F210I-B6 animals
Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

Regrettably power could not be increased by combining the statistical analysis of males and females as body weight between sexes were significantly different (mean weight 31.54±1.44 for females, n=9, and 35.76±1.29 for males, n=10, p=0.043) and, consequently, brain mass to body weight ratio, was also significantly different (1.57±0.071 for females, n=9, and 1.33±0.055 for males, n=10, p=0.016, Anova).

5.1.2.2 TDP-M323K-B6

Tardbp$^{M323K/+}$ of both sexes, had comparable brain masses and ratios between brain mass and body weights to *Tardbp*$^{+/+}$ (figure 5.8 and table 5.7, refer to table for statistical analysis).

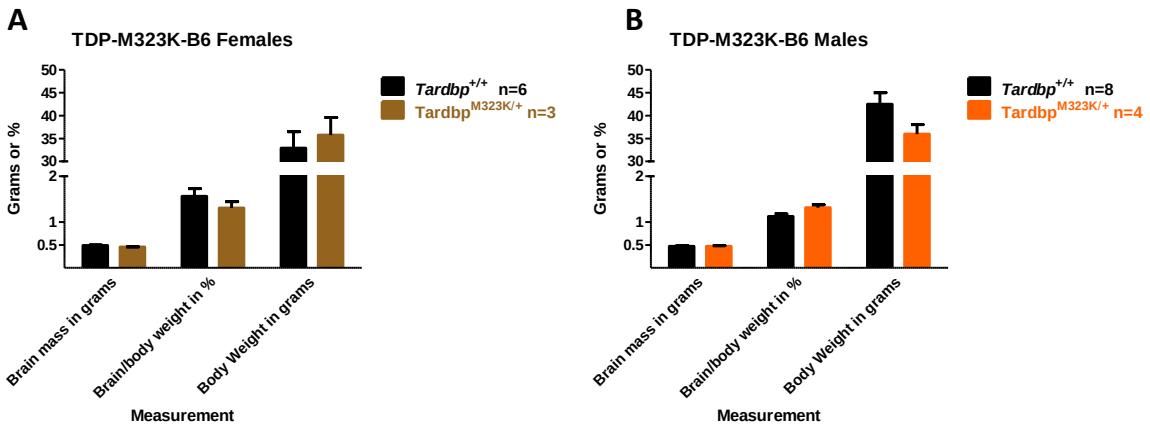

Figure 5.8: Brain mass of TDP-M323K-B6 animals
Tardbp$^{M323K/+}$ females (A) and males (B) had identical brain mass and brain mass to body weight ratios to wild type controls (refer to table 5.7 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{M323K/+}$			
		Mean	SEM	n	Mean	SEM	n	p (Anova)
Females	Brain Mass	0.487	0.011	6	0.457	0.003	3	0.111
	Brain Mass/Body weight	1.563	0.167	6	1.307	0.139	3	0.642
	Body weight	32.917	3.594	6	35.747	3.845	3	0.358
Males	Brain Mass	0.47	0.015	8	0.47	0.014	4	1
	Brain Mass/Body weight	1.124	0.052	8	1.315	0.062	4	0.128
	Body weight	42.476	2.532	8	35.96	2.102	4	0.05

Table 5.7: Brain mass of TDP-M323K-B6 animals

Analogously to the *Tardbp*^{F210I} mutants, it was not possible to increase the power of the calculation for TDP-M323K-B6 cohorts by analysing the animals together, given that testing the different parameters when comparing the sex of the animals, revealed that ratio of brain to body weights is significantly different between sexes (1.18±0.47 for males, n=10 and 1.48±0.12 for females, n=12; p=0.024, Anova).

The fact that the measurements were taken at a very late time point in the lifespan of the laboratory mouse and that no significant differences were observed between *Tardbp*$^{F210I/+}$ or *Tardbp*$^{M323K/+}$ and their wild type controls (*Tardbp*$^{+/+}$) can only be interpreted in two ways.

Firstly, that in heterozygosity, on the C57BL/6J background, these mutations do not affect brain growth detectable above the threshold sensitivity of the common laboratory scales used.

Secondly, that the mutations could affect early brain development which would then become indistinguishable from the wild type controls due to different rates of atrophy.

Unless overt and obvious differences are observed during behavioural testing, by far the most likely interpretation of the results obtained for the brain masses of the animals is that the F210I and M323K mutations do not lead to microcephaly or macrocephaly.

5.1.3 Lifespan/longevity

The effects of the F210I and M323K mutations in the lifespan of laboratory mice were assessed by performing survival analysis in the animals from the experimental cohorts. The analysis included "death from all causes", without exclusions and therefore animals found dead or euthanized by welfare reasons according to the humane end points in the project license (briefly described earlier; e.g. 20% weight loss) and decisions from the named care and welfare officer were all included. The Kaplan Meier survival analysis can therefore determine the influence of the mutations in longevity without giving any information for the possible underlying causes of different survival times. The cut off for longevity was determined at the late time point when most animals were sacrificed to harvest tissues (up to 27 months of age).

5.1.3.1 TFP-F210I-B6

In the TDP-F210I-B6 testing cohorts, no differences were observed in the longevity of either *Tardbp*$^{F210I/+}$ or *Tardbp*$^{+/+}$ females or males (figure 5.9).

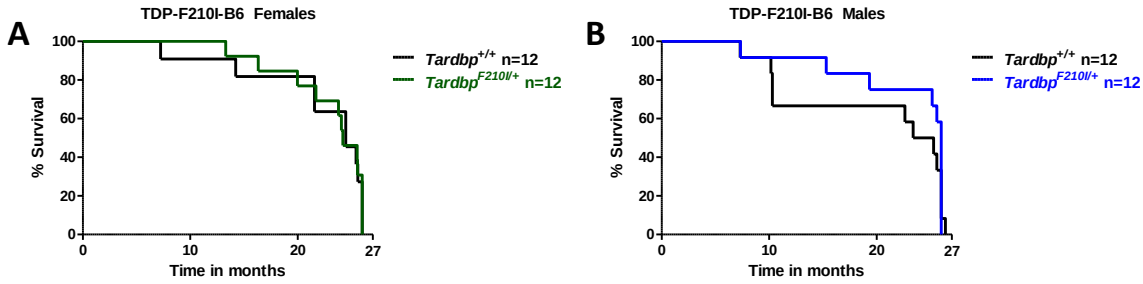

Figure 5.9: Kaplan Meier survival analysis of the effect of *Tardbp*^{F210I} in the longevity of C57BL/6J laboratory mice
No significant differences were observed in the longevity of *Tardbp*$^{F210I/+}$ (A) females or (B) males and wild type controls.

5.1.3.2 TDP-M323K-B6

In the TDP-M323K-B6 testing cohorts, no differences were observed in the longevity of neither *Tardbp*$^{M323K/+}$ or *Tardbp*$^{+/+}$ females or males (figure 5.10).

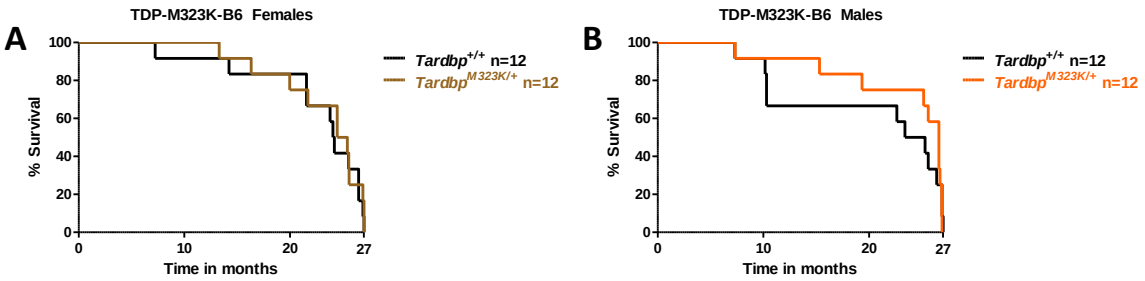

Figure 5.10: Kaplan Meier survival analysis of the effect of *Tardbp*^{M323K} in the longevity of C57BL/6J laboratory mice
No significant differences were observed in the longevity of *Tardbp*$^{M323K/+}$ (A) females or (B) males and wild type controls.

5.1.4 Modified SHIRPA procedure

The modified SHIRPA procedure consists of a battery of tests that aim to identify phenotypes particularly associated with neurological disorders. It is a fast and efficient process which takes approximately five minutes per animal when the researcher is experienced and, critically, it is highly operator dependent. Therefore the SHIRPA procedure should always be performed by the same operator when repeated at multiple time points for the characterisation of a specific mouse line.

Briefly, the SHIRPA procedure starts with the observation of the mouse on a grid floor, screening for tremors. Subsequently the animal is transferred into a rectangular arena where the startle reflex and gait are assessed. From the arena, the animal is suspended by the tail, observing for limb grasping, its body tone palpated, grip strength and limb tone tested, and the self-righting reflex and ability to walk in an inverted position on a grid (negative geotaxis) also assessed.

The modified SHIRPA procedure thus constitutes a non-specific but sensitive test for neurological dysfunction-associated phenotypes. A more detailed description of the modified

SHIRPA procedure is given in chapter 2, where a picture of the main equipment used is available (figure 2.5).

All elements of the modified SHIRPA procedure were tested in TDP-F210I-B6 and TDP-M323K-B6 cohorts. However, no phenotypes were identified associated with heterozygosity for *Tardbp*^{F210I} or *Tardbp*^{M323K} alleles in comparison to wild type controls. Therefore, no statistical analysis could be performed on the elements of the SHIRPA procedure.

Additionally, grip strength was tested separately using a quantitative methodology, and the results are described in the next section.

5.2 Motor function Phenotyping of *Tardbp*^{F210I} and *Tardbp*^{M323K} animals

In this section, the results of behavioural tests aimed at detecting motor deficits in the *Tardbp* ENU mutants will be described.

5.2.1 Open field: total distance travelled

Starting at 17 weeks, TDP-F210I-B6 and TDP-M323K-B6 animals entered the phenotyping pipeline and the open field test, despite not constituting a specific test for motor function, can give an indication of possible deficits by measuring the total distance travelled in the arena during the test period.

The open field system used constitutes of an automated system where animals are placed in a square arena (44x44cm), which is essentially a square Perspex box, and tracked through video recording by proprietary *Ethovision* software for the duration of the test (typically 20 or 30 minutes) while the mice are left to move freely and explore the arena. Central and peripheral areas to the area can be programmed into the software and the output measurements includes

total distance moved and the time spent in the different areas of the arena. The time spend in the centre and periphery of the arenas is given in section 5.3.1.

5.2.1.1 TDP-F210I-B6

At 17 weeks, naive heterozygous *Tardbp*$^{F210/+}$ animals of both sexes travelled similar distances to their wild type controls in the open field arenas during testing (figure 5.11, refer to figure legend for statistical analysis).

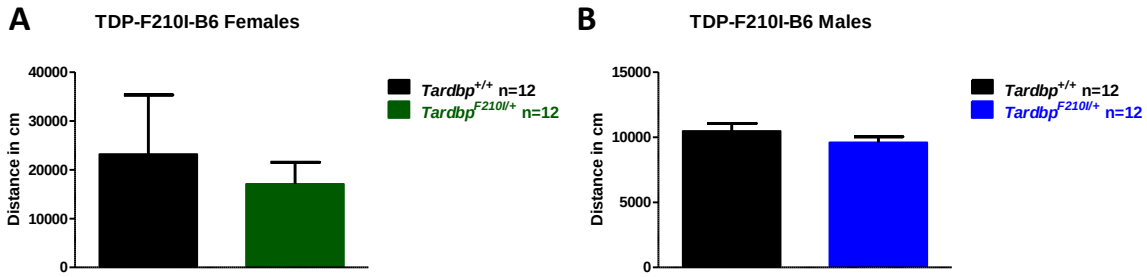

Figure 5.11: Total distance travelled during the open field test by TDP-F210I-B6
The mean total distance travelled by (A) *Tardbp*$^{F210I/+}$ females was comparable to the total distance travelled by wild type controls (mean total distance travelled in centimetres 23192.36±12116.96 for *Tardbp*$^{+/+}$ and 17096.70±4405.60 for *Tardbp*$^{F210I/+}$; p=0.641, Anova) as was the total distance travelled by (B) *Tardbp*$^{F210I/+}$ males and controls (mean total distance travelled in centimetres 10472.78±578.55 for *Tardbp*$^{+/+}$ and 9597.32±1511.94 for *Tardbp*$^{F210I/+}$; p=0.240, Anova). Error bars in all graphs are ±SEM.

5.2.1.2 TDP-M323K-B6

At 17 weeks, naive heterozygous *Tardbp*$^{M323K/+}$ and animals both sexes travelled similar distances to their wild type controls in the open field arenas during testing (Figures 5.12, refer to figure legend for satitstical analysis).

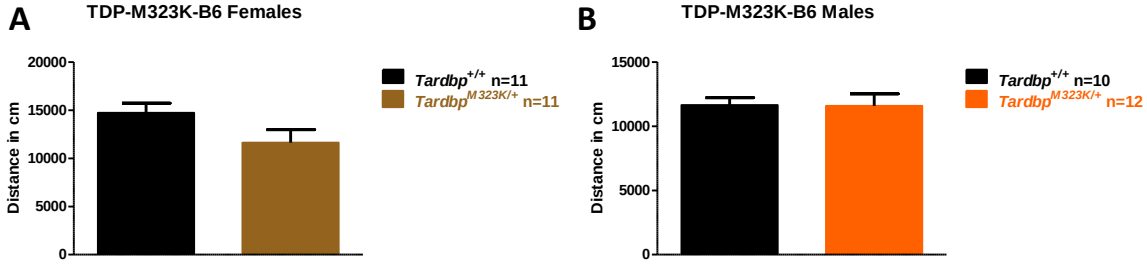

Figure 5.12: Total distance travelled during the open field test by TDP-M323K-B6
The mean total distance travelled by (A) *Tardp*$^{M323K/+}$ females was comparable to the total distance travelled by wild type controls (mean total distance travelled in centimetres 14732.19±998.57 for *Tardbp*$^{+/+}$ and 11621.79±1347.08 for *Tardbp*$^{M323K/+}$; p=0.078, Anova) as was the total distance travelled by (B) *Tardbp*$^{M323K/+}$ males and controls (mean total distance travelled in centimetres 11649.19±566.61 for *Tardbp*$^{+/+}$ and 11585.72±937.79 for *Tardbp*$^{M323K/+}$; p=0.957, Anova). Error bars in all graphs are ±SEM.

5.2.1.3 TDP-M323K-B6-DBA

In the TDP-M323K-B6-DBA cohort, naïve male animals were tested between the ages of 21 and 32 weeks and the distance travelled in the open field arenas was similar across all genotypes (figure 5.13, refer to figure legend for statistical analysis).

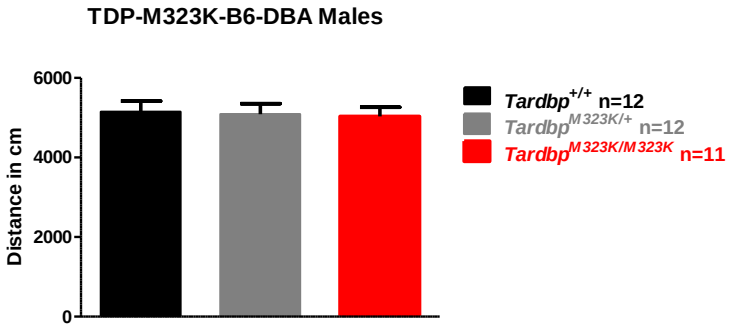

Figure 5.13: Total distance travelled during the open field test by TDP-M323K-B6-DBA mice during test time
In TDP-M323K-B6-DBA cohort, mice of all genotypes travelled identical distances in the open filed arenas during test time (mean total distance travelled in centimetres 5146.35±271.70 for *Tardbp*$^{+/+}$, 5089.14±264.60 for *Tardbp*$^{M323K/+}$ and 5031.78±232.25 for *Tardbp*$^{M323K/M323K}$; p=1 for all comparisons, Anova and posthoc Bonferroni). Error bars in all graphs are ±SEM.

5.2.2 Grip Strength

Forelimb grip strength is useful as a screening test for neuromuscular weakness. However it can also be influenced by sensory deficits as it requires that the animals grab a wire grid with the forepaws. It is performed by suspending the animal by its tail and placing its forelimbs on a grid, enabling it to grab it with its paws. Traction is applied on the tail and the force the animal applies on gripping the grid is measured in grams.

The animal's body weight will influence the grip strength and, additionally, the results are highly influenced by operator-dependent technique and therefore repeated testing is best performed by the same operator.

Measurements were taken at several time points for the TDP-F210I-B6 and TDP-M323K-B6 cohorts and the grip strength was adjusted to body weight and given as a ratio of grip strength to body weight.

5.2.2.1 TDP-F210I-B6

Tardbp$^{F210I/+}$ females had similar grip strength to controls (Figure 5.14, table 5.8; refer to table for statistical analysis), as did *Tardbp*$^{F210I/+}$ males (figure 5.14, table 5.9; refer to table for statistical analysis).

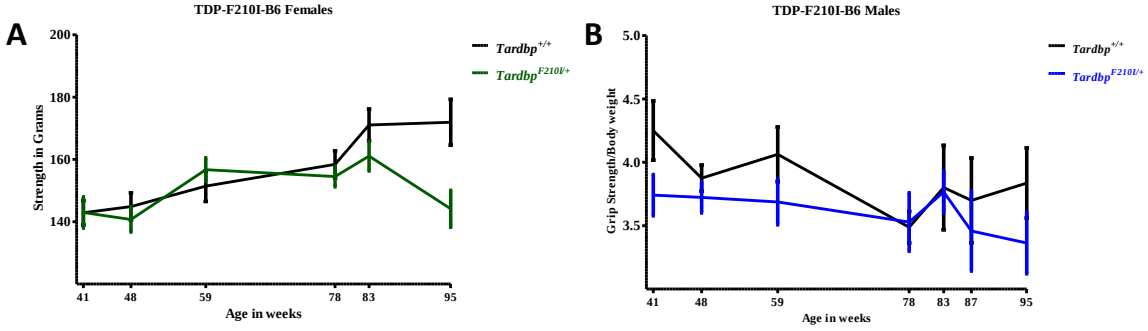

Figure 5.14: Grip strength of TDP-F210I-B6 animals
Grip strength of *Tardbp*$^{F210I/+}$ was equivalent to the grip strength from *Tardbp*$^{+/+}$ in both (A) females (refer to table 5.8 for statistical analysis) and (B) males (refer to tables 5.9 for statistical analysis). A trend towards lower grip strength is observed in (A) *Tardbp*$^{F210I/+}$ at 95 weeks but still fails to reach statistical significance. Error bars in all graphs are ±SEM.

Age (Weeks)	Genotype	Mean strength(Grams)/body Weight(Grams)	SEM	n	p (Anova)	p (FDR corrected)
41	*Tardbp* $^{+/+}$	4.54	0.14	11	0.095	0.208
	Tardbp $^{F210I/+}$	5.05	0.25	12		
48	*Tardbp* $^{+/+}$	4.18	0.12	11	0.169	0.254
	Tardbp $^{F210I/+}$	4.46	0.16	12		
59	*Tardbp* $^{+/+}$	4.36	0.23	11	0.089	0.208
	Tardbp $^{F210I/+}$	4.85	0.14	10		
78	*Tardbp* $^{+/+}$	4.15	0.13	10	0.722	0.823
	Tardbp $^{F210I/+}$	3.98	0.42	11		
83	*Tardbp* $^{+/+}$	4.48	0.17	10	0.823	0.823
	Tardbp $^{F210I/+}$	4.53	0.15	10		
95	*Tardbp* $^{+/+}$	4.23	0.15	5	0.104	0.208
	Tardbp $^{F210I/+}$	3.93	0.08	6		

Table 5.8: Mean grip strength (represented as grip strength in grams to body weight ratio) and statistical analysis of TDP-F210I-B6 females
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing.

Age (Weeks)	Genotype	Mean strength(Grams)/body Weight(Grams)	SEM	n	p (Anova)	p (FDR corrected)
41	*Tardbp* $^{+/+}$	4.25	0.23	8	0.081	0.515
	Tardbp $^{F210I/+}$	3.74	0.16	11		
48	*Tardbp* $^{+/+}$	3.87	0.10	7	0.382	0.669
	Tardbp $^{F210I/+}$	3.72	0.12	10		
59	*Tardbp* $^{+/+}$	4.06	0.22	6	0.202	0.515
	Tardbp $^{F210I/+}$	3.69	0.18	7		
78	*Tardbp* $^{+/+}$	3.49	0.12	6	0.893	0.921
	Tardbp $^{F210I/+}$	3.53	0.23	9		
83	*Tardbp* $^{+/+}$	3.80	0.33	5	0.921	0.921
	Tardbp $^{F210I/+}$	3.77	0.16	7		
87	*Tardbp* $^{+/+}$	3.70	0.33	6	0.613	0.858
	Tardbp $^{F210I/+}$	3.46	0.31	8		
95	*Tardbp* $^{+/+}$	3.84	0.28	6	0.221	0.515
	Tardbp $^{F210I/+}$	3.36	0.24	7		

Table 5.9: Mean grip strength (represented as grip strength in grams to body weight ratio) and statistical analysis of TDP-F210I-B6 males
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing.

5.2.2.2 TDP-M323K-B6

Tardbp$^{M323K/+}$ females had equivalent grip strength when compared to control animals throughout its lives (figure 5.15, table 5.10; refer to table for statistical analysis).

Tardbp$^{M323K/+}$ males displayed a trend towards lower grip strength throughout most of its lives (figure 5.15, table 5.11; refer to table for statistical analysis). Notwithstanding, the trend displayed by *Tardbp*$^{M3232K/+}$ males failed to achieve statistical significance, despite being close to being significant at 41 weeks, the first time point at which it was measured (table 5.10). Therefore, it is concluded that the grip strength of heterozygous males was equivalent to the grip strength of controls (figure 5.15, table 5.11; refer to table for statistical analysis).

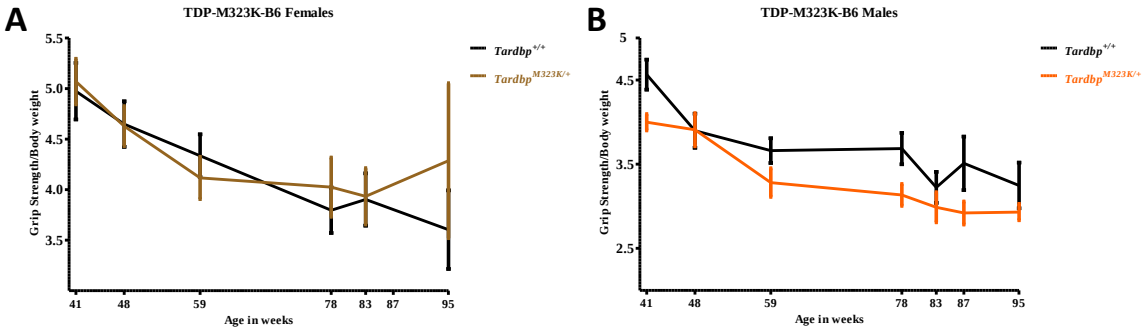

Figure 5.15: Grip Strength of TDP-M323K-B6 animals
Grip strength of *Tardbp*$^{M323K/+}$ was equivalent to the grip strength from *Tardbp*$^{+/+}$ in both (A) females (refer to table 5.10 for statistical analysis) and (B) males (refer to table 5.11 for statistical analysis), given that the trend towards lower grip strength in the latter never achieved statistical significance. Error bars in all graphs are ±SEM.

Age (Weeks)	Genotype	Mean strength(Grams)/body weight(Grams)	SEM	n	p (Anova)	p (FDR corrected)
41	*Tardbp* $^{+/+}$	4.97	0.28	11	0.796	0.944
	Tardbp $^{M323K/+}$	5.07	0.23	10		
48	*Tardbp* $^{+/+}$	4.65	0.22	9	0.944	0.944
	Tardbp $^{M323K/+}$	4.63	0.20	10		
59	*Tardbp* $^{+/+}$	4.34	0.21	11	0.470	0.944
	Tardbp $^{M323K/+}$	4.12	0.21	11		
78	*Tardbp* $^{+/+}$	3.80	0.22	10	0.539	0.944
	Tardbp $^{M323K/+}$	4.02	0.29	9		
83	*Tardbp* $^{+/+}$	3.90	0.26	10	0.936	0.944
	Tardbp $^{M323K/+}$	3.93	0.28	9		
95	*Tardbp* $^{+/+}$	3.60	0.39	5	0.415	0.944
	Tardbp $^{M323K/+}$	4.29	0.77	2		

Table 5.10: Mean grip strength and statistical analysis of TDP-M323K-B6 females
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing.

Age (Weeks)	Genotype	Mean strength(Grams)/body weight(Grams)	SEM	n	p (Anova)	p (FDR corrected)
41	*Tardbp* $^{+/+}$	4.56	0.18	9	0.008	0.056
	Tardbp $^{M323K/+}$	4.00	0.10	12		
48	*Tardbp* $^{+/+}$	3.90	0.20	4	0.974	0.974
	Tardbp $^{M323K/+}$	3.91	0.20	8		
59	*Tardbp* $^{+/+}$	3.66	0.15	5	0.185	0.273
	Tardbp $^{M323K/+}$	3.28	0.17	11		
78	*Tardbp* $^{+/+}$	3.69	0.19	6	0.026	0.091
	Tardbp $^{M323K/+}$	3.13	0.13	12		
83	*Tardbp* $^{+/+}$	3.22	0.18	9	0.369	0.431
	Tardbp $^{M323K/+}$	2.99	0.18	9		
87	*Tardbp* $^{+/+}$	3.51	0.32	6	0.064	0.149
	Tardbp $^{M323K/+}$	2.92	0.14	11		
95	*Tardbp* $^{+/+}$	3.25	0.27	3	0.195	0.273
	Tardbp $^{M323K/+}$	2.93	0.10	7		

Table 5.11: Mean grip strength and statistical analysis of TDP-M323K-B6 males
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the age and p-value numbers in red.

5.2.3 Rotarod

The rotarod is another commonly used test to assess motor function. Animals were tested in the accelerating rotarod (with increasing rotational speed of the rolling drum over time) after training which involved placing the animals back in the rolling drum every time they fell. By conditioning the animals before actually testing the latency to fall, it was aimed to reinforce that the animals would actually continue on the rolling drum until physical causes disabled the mice from doing so, minimising the influence of motivation in the variable measured (latency to fall in seconds).

Measurements were performed in three sessions per day, repeatedly for three days, and each session consisted of at least two tests, ensuring the consistency of the results. An average of all tests was taken as the latency to fall for each individual animal.

5.2.3.1 TDP-F210I-B6

TDP-F210I-B6 animals were tested in the rotarod at 42 and 65 weeks and no significant differences in the latency to fall were found between genotypes (figure 5.16, refer to figure legend for statistical analysis).

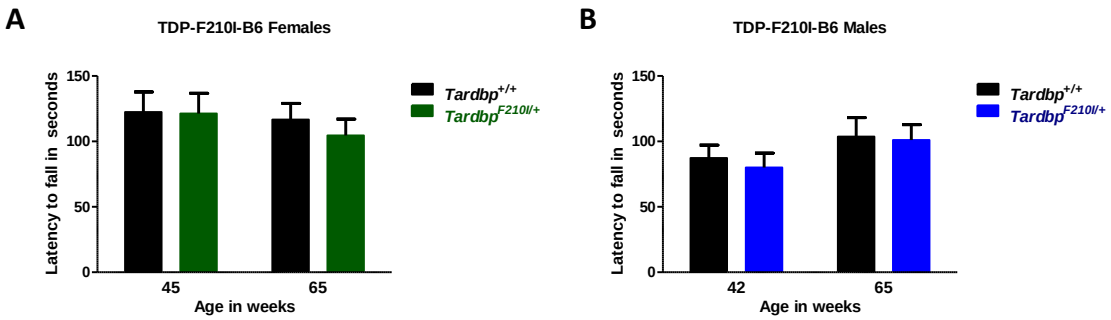

Figure 5.16: Latency to fall during the rotarod test of TDP-F210I-B6 animals
A) Latency to fall was not significantly different in *Tardbp*$^{F210I/+}$ females and controls a 42 weeks (mean latency to fall in seconds 122.38±15.44 for *Tardbp*$^{+/+}$, n=11 and 121.36±12.21 for *Tardbp*$^{F210I/+}$, n=12; p=0.959, Anova) and 65 weeks (mean latency to fall in seconds 116.65±12.46, n=10 for *Tardbp*$^{+/+}$ and 104.59±12.46 for *Tardbp*$^{F210I/+}$, n=11; p=0.503, Anova).
B) In *Tardbp*$^{F210I/+}$ males latency to fall was also similar to controls at 45 weeks (mean latency to fall in seconds 87.35±9.84 for *Tardbp*$^{+/+}$, n=11 and 80.12±10.99 for *Tardbp*$^{F210I/+}$, n=11; p=0.630, Anova) and 65 weeks (mean latency to fall in seconds 103.66±14.52 for *Tardbp*$^{+/+}$, n=8 and 101.11±11.74 for *Tardbp*$^{F210I/+}$, n=11 p=0.892, Anova). Error bars in all graphs are ±SEM.

5.2.3.2 TDP-M323K-B6

TDP-M323K-B6 were tested at 40 and 65 weeks, and no significant differences were observed in the latency to fall between heterozygous animals and wild type controls at both time points (figure 5.17, refer to figure legend for statistical analysis).

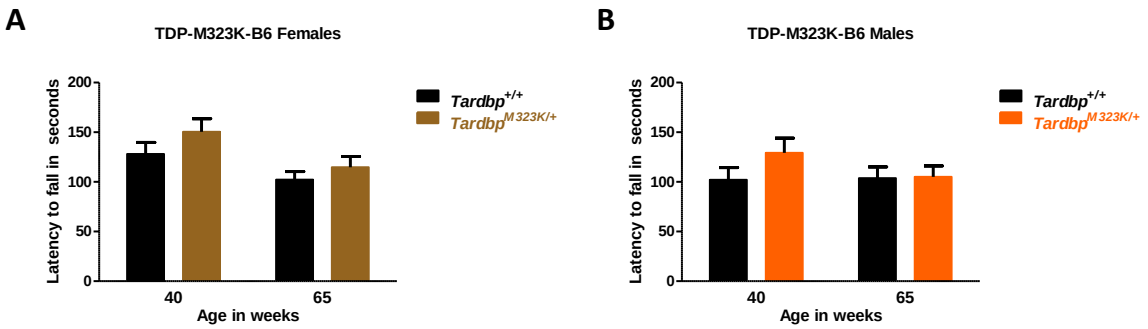

Figure 5.17: Latency to fall during the rotarod test of TDP-M323K-B6 animals
A) Latency to fall was not significantly different in *Tardbp*$^{M323K/+}$ females and controls at 40 weeks (mean latency to fall in seconds 128.15±11.52 for *Tardbp*$^{+/+}$, n=11 and 150.35±13.29 for *Tardbp*$^{M323K/+}$, n=11 ; p=0.221, Anova) and 65 weeks (mean latency to fall in seconds 102.25±8.21 for *Tardbp*$^{+/+}$, n=11 and 114.75±10.87 for *Tardbp*$^{M323K/+}$, n=10 ; p=0.365, Anova).
B) In *Tardbp*$^{M323K/+}$ males latency to fall was also similar to controls at 42 weeks (mean latency to fall in seconds 102.00±12.38 for *Tardbp*$^{+/+}$, n=9 and 129.26±14.87 for *Tardbp*$^{M323K/+}$, n=12; p=0.195, Anova) and 65 weeks (mean latency to fall in seconds 103.78±11.33 for *Tardbp*$^{+/+}$, n=9 and 105.13±11.01 for *Tardbp*$^{M323K/+}$, n=12 ; p=0.934, Anova). Error bars in all graphs are ±SEM.

5.2.4 Automated home cage wheel-running system

The automated home cage wheel-running system was recently validated as a reliable test in identifying motor deficits (Mandillo et al. 2014) and allows singly housed mice to run freely, automatically measuring speed, distance and time spent running, amongst other variables.

The test was developed at the MRC Harwell by Ines Heise, from the neurobehavioural group, and the mice tested with her guidance and advice. Only female animals were tested given that male animals would not tolerate being re-housed together in the same cage after singly housed for a week, which would lead to copious fighting amongst the animals and, inevitably, wounds which would, in turn, be likely to require that the animals were sacrificed due to welfare concerns.

In contrast to the rotarod, the home-cage wheel-running system is dependent on voluntary running by the animals tested and thus, the data obtained is especially influenced by motivation. Particular relevance for the parameters measured were given to total distance travelled and average maximum speed, given that these are reliable indicators of the animal's performance (Mandillo et al. 2014).

The animals were placed in its individual cages for seven days and allowed to run freely in the wheels. When analysing the total distance travelled, the analysis included even the animals which did not run in a particular night, thus giving an output influenced by motivation. The analysis of the average maximum speed included only the animals which ran in each individual night, which explains the different experimental numbers (n) in the data tables and graphs relative to this test.

5.2.4.1 TDP-F210I-B6

At testing ages between 83 and 84 weeks, *Tardbp*$^{F210I/+}$ females travelled a total distance comparable to wild type controls and also had similar average maximum speed to *Tardbp*$^{+/+}$ females (figure 5.18 and table 5.12. refer to table for statistical analysis).

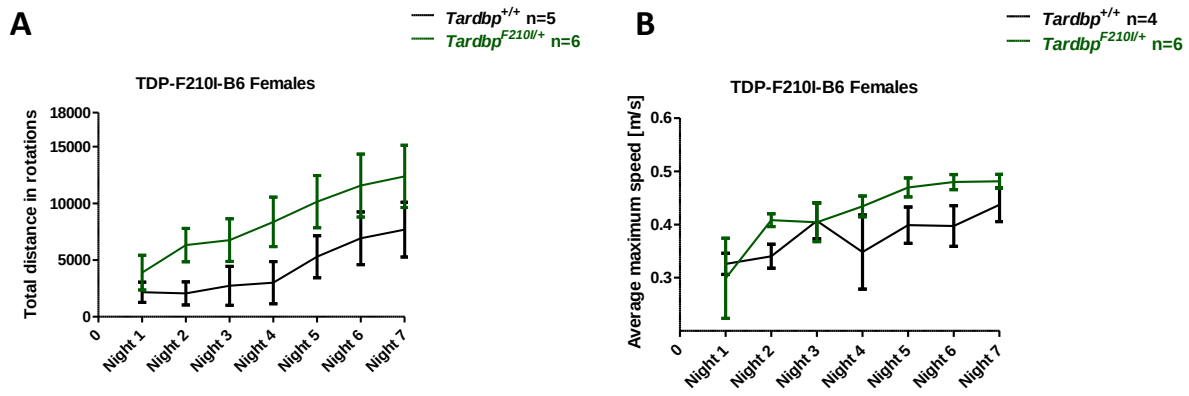

Figure 5.18: Total distance travelled and average maximum speed in the home cage automated wheel-running test of TDP-F210I-B6 females
Tardbp$^{F210I/+}$ females displayed a tendency to travel higher distances and have (B) higher average maximum speed, but the difference failed to reach statistical significance in any of the nights (refer to table 5.12 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{F210I/+}$				
		Mean	SEM	n	Mean	SEM	n	p (Anova)	p (FDR corrected)
Total Distance ran (rotations)	Night1	2163	895.26	5	3895.5	1527.78	6	0.379	0.38
	Night2	2063.2	1013.53	5	6326	1473.71	6	0.048	0.263
	Night3	2731.4	1715.24	5	6763.67	1885.51	6	0.155	0.263
	Night4	3010.6	1861.23	5	8366.5	2183.9	6	0.102	0.263
	Night5	5299.6	1853.54	5	10153.83	2306.96	6	0.146	0.263
	Night6	6917.8	2322.65	5	11575	2777.5	6	0.242	0.28
	Night7	7690	2418.89	5	12382	2741.49	6	0.241	0.28
Average Maximum Speed (meters/seconds)	Night1	0.326	0.02	3	0.299	0.075	6	0.799	0.932
	Night2	0.34	0.023	3	0.408	0.012	6	0.025	0.161
	Night3	0.407	0.034	3	0.404	0.036	6	0.97	0.97
	Night4	0.348	0.07	3	0.434	0.02	6	0.155	0.252
	Night5	0.399	0.034	4	0.47	0.018	6	0.077	0.18
	Night6	0.397	0.038	4	0.48	0.014	6	0.046	0.161
	Night7	0.437	0.032	4	0.481	0.013	6	0.18	0.252

Table 5.12: Total distance travelled and maximum average speed in the home cage automated wheel-running test of TDP-F210I-B6 females
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

5.2.4.2 TDP-M323K-B6

Tardbp$^{M323K/+}$ females, tested at ages between 97 and 100 weeks, displayed a tendency to have a lower speed and travel less distance. However, the difference never reached statistical significance for the average maximum speed and was only significant for the total distance ran in night 2 (mean total distance ran in rotations 2157.4±214.10 for *Tardbp*$^{+/+}$, n=5 and 564.29±265.8 for *Tardbp*$^{M323K/+,}$ n=7, p=0.01, Anova and Benjamini Hochberg correction) (figure 5.19 and table 5.13).

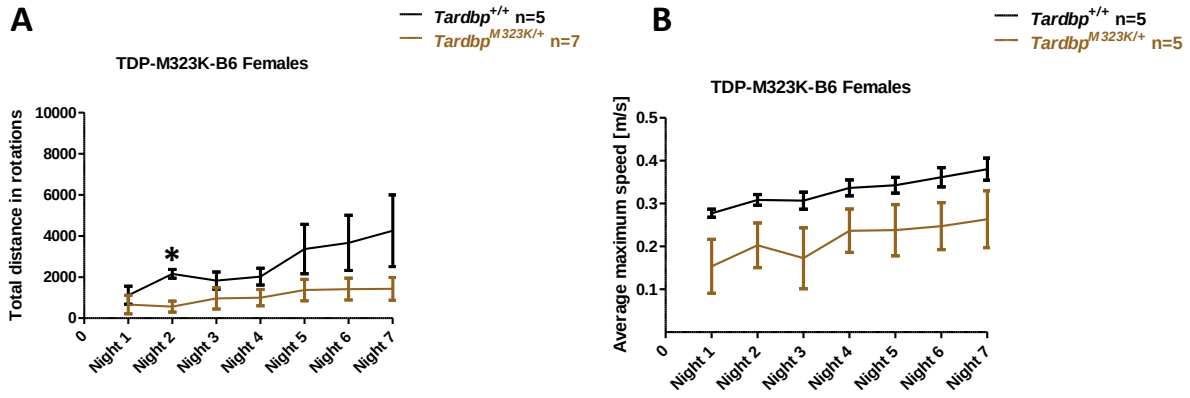

Figure 5.19: Total distance travelled and average maximum speed in the home cage automated wheel-running test of TDP-M323K-B6 females
Tardbp$^{M323K/+}$ females displayed a tendency to (A) travel less distance, which reached statistical significance in the second night (*p=0.01) and have a (B) lower average maximum speed, which failed to reach statistical significance in any of the nights (refer to table 5.13 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{M323K/+}$				
		Mean	SEM	n	Mean	SEM	n	p (Anova)	p (FDR corrected)
Total Distance ran (rotations)	Night1	1114.4	435.68	5	658.57	449.8	7	0.499	0.499
	Night2	2157.4	214.1	5	564.29	265.81	7	**0.001**	**0.007**
	Night3	1827.6	415.88	5	965.43	522.59	7	0.256	0.299
	Night4	2021.6	406.59	5	998.86	396.24	7	0.11	0.169
	Night5	3360.4	1199.18	5	1370.71	520.75	7	0.121	0.169
	Night6	3663.4	1340.12	5	1413.86	530.42	7	0.109	0.169
	Night7	4255.4	1739.96	5	1426.43	553.58	7	0.105	0.169
Average Maximum Speed (meters/seconds)	Night1	0.28	0.01	5	0.15	0.06	5	0.088	0.138
	Night2	0.31	0.01	5	0.2	0.05	5	0.084	0.138
	Night3	0.31	0.02	5	0.17	0.07	5	0.107	0.138
	Night4	0.34	0.02	5	0.24	0.05	5	0.121	0.138
	Night5	0.34	0.02	5	0.24	0.06	5	0.134	0.138
	Night6	0.36	0.02	5	0.25	0.05	5	0.108	0.138
	Night7	0.38	0.03	5	0.26	0.07	5	0.138	0.138

Table 5.13: Home cage automated-wheel running mean total distance ran; mean maximum average speed and statistical analysis in TDP-M323K-B6
Please note that the p (FDR corrected) is the p-value obtained after the Benjamini Hoechberg correction of multiple testing. Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

The results from the home cage automated-wheel running test for *Tardbp*^{F210I} and *Tardbp*^{M323K} displayed in figures 5.18 and 5.19 show overall a non-significant trend for total distance ran and average maximum speed in opposite directions between these two ENU mutants when compared to wild type controls, possibly suggesting that *Tardbp*$^{M323K/+}$ animals are less motivated to run (when compared to wild type) whereas *Tardbp*$^{F210I/+}$ animals are more motivated to run (when compared to wild type controls).

It is of course highly speculative to suggest that motivation to run is the factor responsible for the trends observed. Nevertheless, the data, particularly taken together with the data from rotarod testing, does not support the existence of significant neuromuscular degeneration or motor dysfunction in the heterozygous mutant animals, at least at the ages tested.

5.2.5 Locotronic/Horizontal ladder

Fine motor and sensory function was tested using the automated locotronic system. In this test, mice are placed in a horizontal ladder and as the animals walk, beams record each event its paws miss the ladder step and fall below the step level as errors, until the animal reaches the opposite end of the ladder.

5.2.5.1 TDP-F210I-B6

At 81 weeks of age, *Tardbp*$^{F210I/+}$ females tended to make more errors with both the forelimbs and hind limbs when compared to *Tardbp*$^{+/+}$ but these fell below statistical significance (figure 5.20 and table 5.14, refer to table for statistical analysis).

Male *Tardbp*$^{F210I/+}$ animals (81 weeks of age) also tended to make more errors than wild type controls (Figure 5.20 and table 5.14) and when all errors (forelimb and hind limb errors) were analysed together, the difference was statistically significant (mean number of errors 7.73±0.94, n=8 for *Tardbp*$^{+/+}$ and 12.72±1.87 for *Tardbp*$^{F210I/+}$, n=10; p=0.042, Anova)

Moreover, given that no statistical significance was identified in the number of errors given by males or females (p>0.68 for all parameters measured), the data was analysed together to increase power. When males and females were analysed together, *Tardbp*$^{F210I/+}$ made significantly higher number of hind limb error (3.55±0.65, n=20 for *Tardbp*$^{+/+}$ and 7.53±1.57, n=18 for *Tardbp*$^{F210I/+}$; p=0.30, Anova) and combined forelimb and hind limb errors (7.26±1.71, n=20 for *Tardbp*$^{+/+}$ and 12.73±1.86, n=18 for *Tardbp*$^{F210I/+}$; p=0.021, Anova) (figure 5.20 and table 5.14) compared to wild type controls.

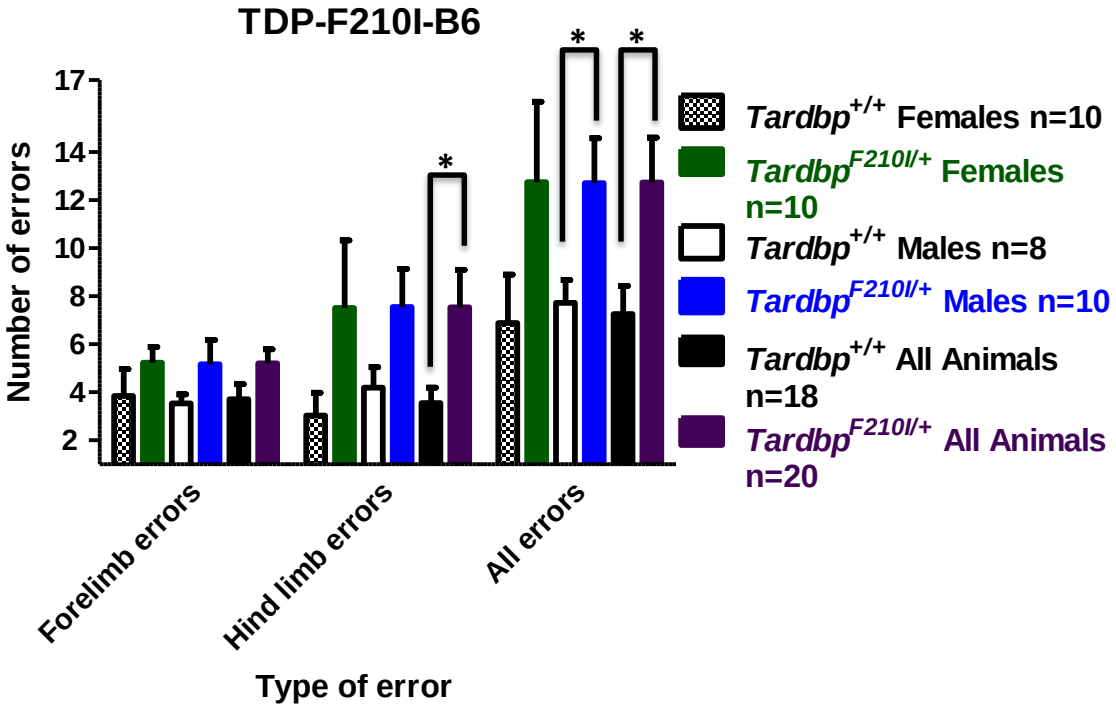

Figure 5.20: Errors made by TDP-F210I-B6 in the locotronic test
In the locotronic test, male *Tardbp*$^{F210I/+}$ made significantly more total number of errors when compared to controls (*p<0.05; refer to table 5.14 for statistical analysis). Additionally, when female and male animals were analysed together, *Tardbp*$^{F210I/+}$ of both sexes made significantly more hind limb and total number of errors more when compared to controls (*p<0.05; refer to table 5.14 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{F210I/+}$			
		Mean	SEM	n	Mean	SEM	n	p (Anova)
Females	Forelimb errors	3.85	1.11	10	5.23	0.66	10	0.299
	Hind limb erros	3.03	0.94	10	3.03	0.94	10	0.148
	All errors	6.88	2.02	10	6.88	2.02	10	0.151
Males	Forelimb errors	3.54	0.38	8	5.17	1.01	10	0.189
	Hind limb erros	4.19	0.86	8	7.55	1.6	10	0.104
	All errors	7.73	0.95	8	12.72	1.87	10	**0.042**
All animals	Forelimb errors	3.71	0.63	18	5.2	0.59	20	0.091
	Hind limb erros	3.55	0.65	18	7.53	1.57	20	**0.03**
	All errors	7.26	1.17	18	12.73	1.87	20	**0.021**

Table 5.14: Locotronic test results and statistical analysis for TDP-F210I-B6
Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

5.2.5.2 TDP-M323K-B6

At 83 weeks of age, *Tardbp*$^{M323K/+}$ males and females made similar numbers of errors in the locotronic, which continued to be similar when all animals were analysed together (figure 5.21, table 5.15).

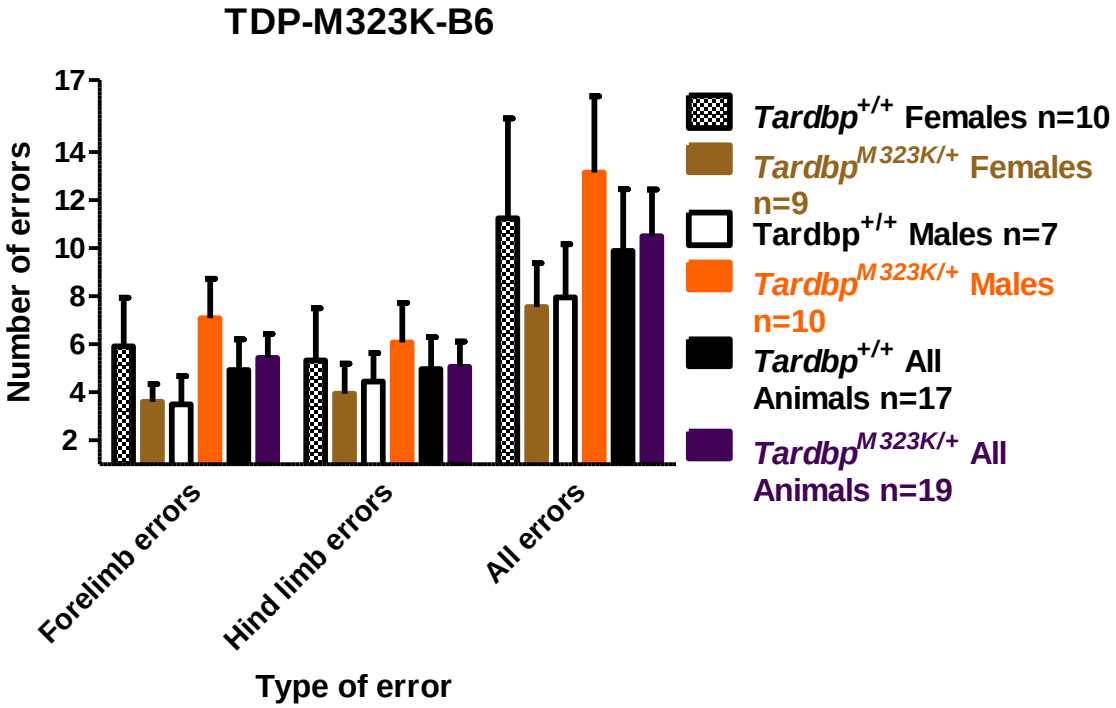

Figure 5.21: Errors made by TDP-M323K-B6 in the locotronic test
Tardbp$^{M323K/+}$ made equivalent number of errors to *Tardbp*$^{+/+}$ on locotronic testing (refer to table 5.15 for statistical analysis). Error bars in all graphs are ±SEM.

		Tardbp$^{+/+}$			*Tardbp*$^{M323K/+}$			
		Mean	SEM	n	Mean	SEM	n	p (Anova)
Females	Forelimb errors	5.92	2.03	10	3.61	0.72	9	0.32
	Hind limb erros	5.33	2.17	10	3.94	1.25	9	0.597
	All errors	11.25	4.15	10	7.56	1.82	9	0.444
Males	Forelimb errors	3.5	1.17	7	7.08	1.65	10	0.126
	Hind limb erros	4.45	1.17	7	6.07	1.66	10	0.48
	All errors	7.95	2.22	7	13.15	3.17	10	0.239
All animals	Forelimb errors	4.92	1.29	17	5.44	1	19	0.75
	Hind limb erros	4.97	1.33	17	5.06	1.06	19	0.957
	All errors	9.89	2.57	17	10.5	1.94	19	0.85

Table 5.15: Locotronic test results and statistical analysis for TDP-323K-B6

The results of tests and measurements aimed at identifying biological dysfunction involving multiple systems (weights, brain mass, lifespan and SHIRPA) and of tests directed towards testing motor function, albeit not exclusively testing motor function but also influenced by possible sensory deficits, (distanced travelled in open field, grip strength, rotarod, home cage wheel-running and locotronic) have been described for *Tardbp*^{F210I} and *Tardbp*^{M323K} mutants in the two sections above.

Motor testing has failed to identify phenotypes associated with neuromuscular degeneration in *Tardbp*^{M323K} or *Tardbp*^{F210I} animals. However, in the latter mutant mouse line, the results from locotronic/horizontal ladder testing suggests that a discreet fine motor or sensory deficit may be present in *Tardbp*$^{F210I/+}$.

Despite no overt motor phenotypes having been identified in the mice during motor function testing, full histopahological assessment of the brain and spinal cord harvested at the end point of the study (25-27 months) will be performed for both mutant lines, particularly as histopahological changes including TDP43 aggregation and astrogliosis have been reported to precede behavioural deficits in mice (Swarup et al. 2011).

The results from tests aimed at identifying endophenotypes of neuropsychiatric diseases, and deficits in learning and memory will be subsequently described. Behavioural tests aimed at assessing these functions can overlap and therefore, the results will be described together, and its interpretations given in the text.

5.3 Endophenotypes of neuropsychiatric diseases and learning and memory

As described above, starting at 17 weeks for TDP-F210I-B6 and TDP-M323K-B6 cohorts entered the phenotyping pipeline, starting with the open field test. At ages between 21 and 32 weeks the TDP-M323K-B6-DBA cohort was also tested in the open field. This was the first test performed, taking advantage of the animals being naive and never having been handled before except for breeding and identification purposes (ear punch-whole marking was performed).

The open field test, together with the Light/Dark box, is used as an anxiety paradigm.

5.3.1 Open Field and Light/Dark box

A brief description of the open field test is given in the previous section. The percentage of time an animal spends in the centre of an open field arena during the test, comparing with periphery, gives and output related to anxiety. Mice which spend more time in the centre of the arena are proposed to have a reduction in anxiety, given that this is the effect observed in wild type mice treated with anxiolytics (Bouwknecht et al. 2004).

5.3.1.1 TDP-F210I-B6

At 17 weeks, in the TDP-F210I-B6 cohorts no significant difference was observed between the time the animals spent in the centre or the periphery of the open field arena across genotypes (figure 5.22, refer to figure legend for statistical analysis).

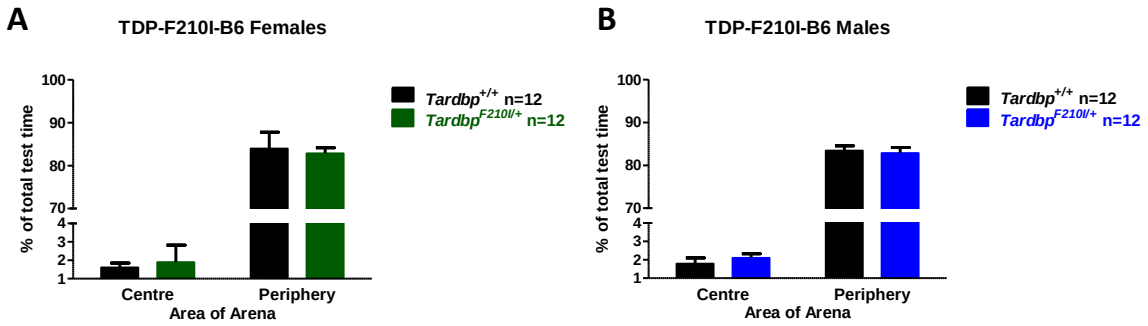

Figure 5.22: Percentage of test time spent in the centre and periphery of the arena during open field testing by TDP-F210I-B6 animals
A) *Tardbp*$^{F210I/+}$ females spent similar proportions of the open field test time in the centre of the arena (mean percentage of time spend in the centre 1.60±0.5 for *Tardbp*$^{+/+}$ and 1.90±0.93 for *Tardbp*$^{F210I/+}$; p=0.763, Anova) and the periphery of the arena (mean percentage of time spend in the periphery 83.97±3.83 for *Tardbp*$^{+/+}$ and 82.84±1.31 for *Tardbp*$^{F210I/+}$; p=0.594, Anova).
B) *Tardbp*$^{F210I/+}$ males also spent similar proportions of the open field test time in the centre of the arena (mean percentage of time spend in the centre 1.79±0.32 for *Tardbp*$^{+/+}$ and 2.10±0.22 for *Tardbp*$^{F210I/+}$; p=0.426, Anova) and the periphery of the arena (mean percentage of time spend in the periphery 83.40±1.16 for *Tardbp*$^{+/+}$ and 79.42±7.48 for *Tardbp*$^{F210I/+}$; p=0.752, Anova). Error bars in all graphs are ±SEM.

5.3.1.2 TDP-M323K-B6

At 17 weeks for the TDP-M323K-B6 cohorts, no significant difference was observed between the time the animals spent in the centre or the periphery of the open field arena across genotypes (figure 5.23, refer to figure legend for statistical analysis).

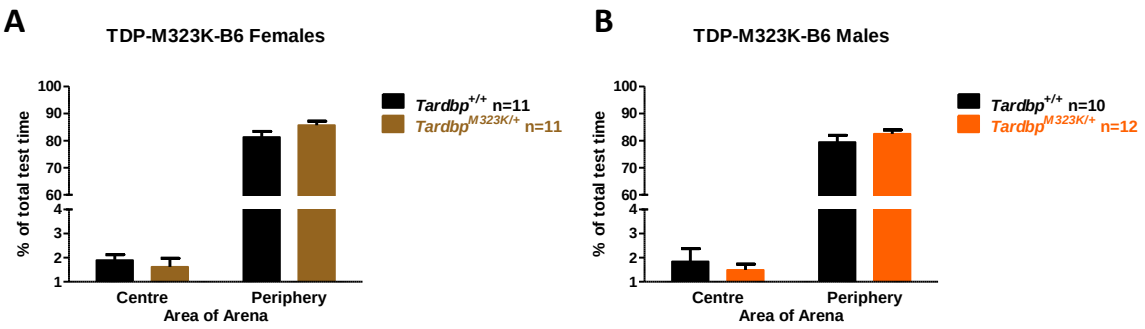

Figure 5.23: Percentage of test time spent in the centre and periphery of the arena during open field testing by TDP-M323K-B6 animals.
A) *Tardbp*$^{M323K/+}$ females spent similar proportions of the open field test time in the centre of the arena (mean percentage of time spend in the centre 1.89±0.23 for *Tardbp*$^{+/+}$ and 1.62±0.36 for *Tardbp*$^{M323K/+}$; p=0.526, Anova) and the periphery of the arena (mean percentage of time spend in the periphery 81.31±2.10 for *Tardbp*$^{+/+}$ and 85.65±1.54 for *Tardbp*$^{M323K/+}$; p=0.111, Anova).
B) *Tardbp*$^{M323K/+}$ males also spent similar proportions of the open field test time in the centre of the arena (mean percentage of time spend in the centre 1.84±0.54 for *Tardbp*$^{+/+}$ and 1.49±0.24 for *Tardbp*$^{F210I/+}$; p=0.537, Anova) and the periphery of the arena (mean percentage of time spend in the periphery 79.43±2.54 for *Tardbp*$^{+/+}$ and 82.50±1.51 for *Tardbp*$^{M323K/+}$; p=0.283, Anova). Error bars in all graphs are ±SEM.

5.3.1.3 TDP-M323K-B6-DBA

At ages between 21 and 32 weeks, in TDP-M323K-B6-DBA no significant difference was observed between the time the animals spent in the centre arena or the periphery of the open field across genotypes (figure 5.25, refer to figure legend for statistical analysis).

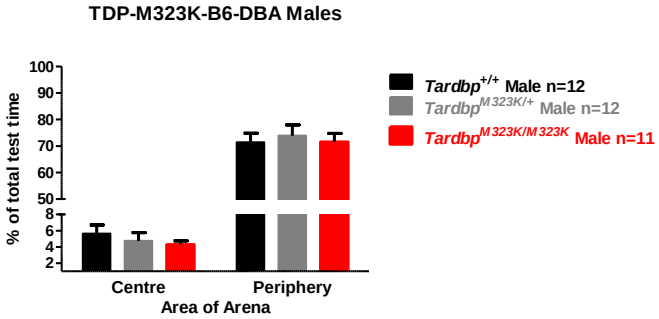

Figure 5.24: Percentage of test time spend in the centre and periphery of the arena during open field testing by TDP-M323K-B6-DBA animals
In the TDP-M323K-B6-DBA cohort, animals across all genotypes spent equivalent percentages of test time in the centre of the arena (mean percentage of test time spent in centre 5.64±1.06 for *Tardbp*$^{+/+}$, 4.74±1.02 for *Tardbp*$^{M323K/+}$and 4.30±0.48 for *Tardbp*$^{M323K/M323K}$; p=1 for comparison between wild type and heterozygous animals and p=0.919 for comparison between wild type and homozygous animals) and in the periphery (mean percentage of test time spent in the periphery 71.37±3.47 for *Tardbp*$^{+/+}$, 73.88±4.09 for *Tardbp*$^{M323K/+}$ and 71.55±10.64 for *Tardbp*$^{M323K/M323K}$; p=1 for all comparisons, Anova and poshoc bonferroni). Error bars in all graphs are ±SEM.

The Light/Dark box test was also performed in an automated system in which mice are tracked by video camera through *Ethovision* software. The animals were placed in a

rectangular box which is divided between an open area, exposed to light, and a closed area, with opaque walls and roof, to which the animal can gain access from the lighted area through a small opening.

Mice's innate behaviour in the novel environment of the Light/Dark box is to seek shelter in the dark area, spending, in general, most of the test time in the dark. Thus, increased time spent in the lighted area is associated with reduced anxiety (Bouwknecht et al. 2004).

5.3.1.4 TDP-F210I-B6

The Light/Dark box test was performed at 83 weeks of age in the TDP-F210I-B6 and no significant differences between the percentages of test time spend in the areas of the box between mice of the two genotypes was observed (figure 5.26, refer to figure legend for statistical analysis).

Tardbp^{F210I} males displayed a trend towards spending less time in the lighted area of the box (figure 5.25, refer to figure legend for statistical analysis). However this observation was likely due to the increased variability observed in the wild type controls. Moreover, the difference still failed to reach statistical significance (mean percentage of time spent in the lighted area of the box 37.19±11.03, n=8 for *Tardbp*$^{+/+}$ and 17.17±4.59, n=9 for *Tardbp*$^{F210I/+,}$ p=0.087, Anova).

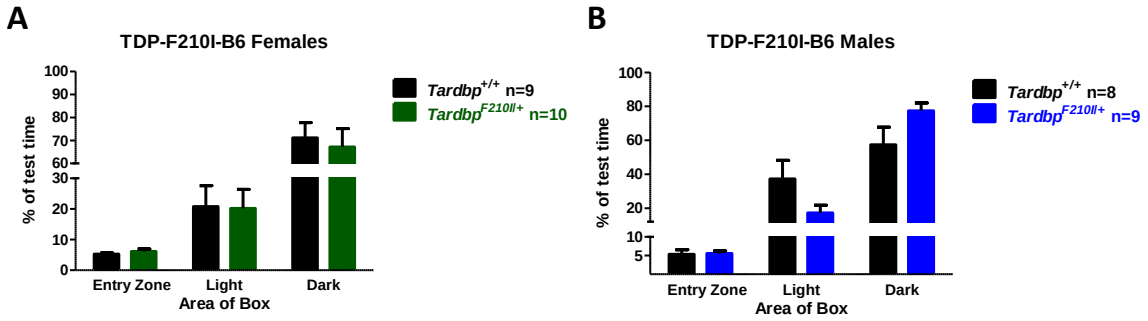

Figure 5.25: Percentage of test time spent in the different areas of the Light/Dark box by TDP-F210I-B6

Wild type and heterozygous (A) females spent equivalent percentages of the test time in the light and dark areas of the box (mean percentage of test time spent in the light 20.79±6.81 for *Tardbp*$^{+/+}$ and 20.20±6.18 for *Tardbp*$^{F210I/+}$; p=0.950, Anova; mean percentage of test time spent in the dark 71.10±6.63 for *Tardbp*$^{+/+}$ and 67.14±8.01 for *Tardbp*$^{F210I/+}$; p=0.711, Anova) as did (B) males (mean percentage of test time spent in the light 37.19±11.02 for *Tardbp*$^{+/+}$ and 17.17±4.59 for *Tardbp*$^{F210I/+}$; p=0.101, Anova; mean percentage of test time spent in the dark 57.33±10.44 for *Tardbp*$^{+/+}$ and 77.39±4.61 for *Tardbp*$^{F210I/+}$; p=0.087, Anova). Error bars in all graphs are ±SEM.

5.3.1.5 TDP-M323K-B6

The Light Dark box test was performed at 84 weeks of age in the TDP-M323K-B6 cohorts and no significant differences between the percentages of test time spent in the areas of the box between mice of the two genotypes was observed (figure 5.26, refer to figure legend for statistical analysis).

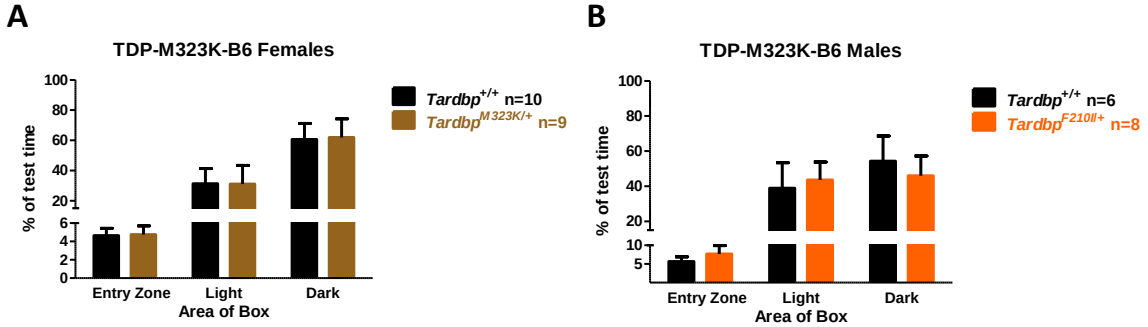

Figure 5.26: Percentage of test time spent in the different areas of the Light/Dark box by TDP-M323K-B6

Wild type and heterozygous (A) females spent equivalent percentages of the test time in the light and dark areas of the box (mean percentage of test time spent in the light 31.29±9.99 for *Tardbp*$^{+/+}$ and 31.01±12.34 for *Tardbp*$^{M323K/+}$; p=0.986, Anova; mean percentage of test time spent in the dark 60.61±10.59 for *Tardbp*$^{+/+}$ and 61.90±12.30 for *Tardbp*$^{M323K/+}$; p=0.937, Anova) as did (B) males (mean percentage of test time spent in the light 38.91±14.50 for *Tardbp*$^{+/+}$ and 43.54±10.25 for *Tardbp*$^{M323K/+}$; p=0.793, Anova; mean percentage of test time spent in the dark 54.32±14.29 for *Tardbp*$^{+/+}$ and 45.97±11.27 for *Tardbp*$^{M323K/+}$; p=0.650, Anova). Error bars in all graphs are ±SEM.

5.3.2 Contextual and cued fear conditioning

The fear conditioning protocol used has been described in the materials and methods section. Briefly, animals are conditioned to associate a mild foot shock delivered by a grid cage with that specific environment and, additionally, a loud high frequency acoustic stimulus (cue), after baseline measurements of freezing are taken. Three foot shocks are delivered preceded by the acoustic cue in the first day. The day following conditioning, the time the animals spend freezing when being placed in the same cage (contextual fear conditioning) or when the loud noise is played with the mouse placed in a different container (cued fear conditioning) is measured. Freezing times are determined by the tracking software; proprietary *Any Maze* software.

Freezing constitutes a natural reaction to a new environment, or fear, in mice and thus "freezing times" should be increased when animals successfully couple the innocuous context (cage environment) or cue (auditory tone) with the noxious shock.

Transgenic animal models of neurodegenerative diseases, including FTD (Cook et al. 2014) and AD (España et al. 2010) have been shown to exhibit dysfunctional fear conditioning responses. The fear conditioning results for TDP43 transgenics is not consensual with some studies being unable to find deficits (Medina et al. 2014) and others claiming deficits exist (Tsai et al. 2010).

Moreover, statistical analysis of fear conditioning has been performed in many studies comparing the difference in time freezing in response to the cue or context between experimental and control groups without subtracting the baseline freezing time, therefore not actually representing the difference in the response to context or cue after conditioning and being biased by the baseline freezing measurements.

All parameters from the contextual and fear conditioning test were analysed, including:

- “Tone-Noxius stimulus” (i.e. freezing time with each shock-cue conditioning event), which provides a measurement of whether mice from different genotypes have conditioned appropriately (day 1).
- “Context Baseline”, which gives a measurement of the time the mice spend freezing in the new environment (cage) before being submitted to any foot shocks (day 1).
- “Context Freezing” (i.e. time freezing in context), which gives an output of the time the mice spend freezing in the same cage where conditioning took place on the second day, thus giving an output of the contextual memory (day 2).
- Context-(Context Baseline) (i.e. the difference between the baseline freezing time and the time freezing in context), which enables to dissect the increase in freezing which is in fact associated with the contextual memory of the shock, eliminating any bias that can arise from the differences in context baseline freezing between experimental and control groups.
- “Cue Baseline”, which gives a measurement of the time the mice spend freezing in the new container before being submitted to the cue, thus enabling to correct the influence of the new environment in the measurement (percentage of time freezing, measured in day 2).
- “Cue freezing”, which is the measurement of time freezing when mice are exposed to the cue only after conditioning (day 2).
- Cue-(Cue Baseline) (i.e. the difference between the baseline freezing time and the time freezing in response to the cue), which enables to dissect the increase in freezing which is in fact associated with the memory of the association between the cue and the shock, eliminating any bias that can arise from the differences in cue baseline freezing between experimental and control groups.

5.3.2.1 TDP-F210I-B6

TDP-F210I-B6 females were tested in the fear conditioning automated set up between the ages of 24 and 26 months and the males between the ages of 24 and 27 months.

Tardbp$^{F210I/+}$ females displayed a similar increase in freezing time with the coupling of each tone-noxious stimulus (foot shock) to *Tardbp*$^{+/+}$ animals (figure 5.27, table 5.16; refer to table statistical analysis).

All other parameters measured in females were also comparable, without any statistically significant differences being observed (figure 5.27, table 5.16).

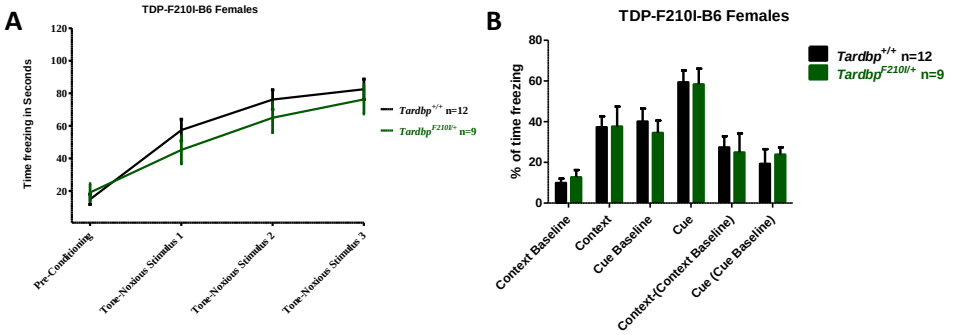

Figure 5.27: Fear conditioning in TDP-F210I-B6 females
A) Both heterozygous and controls associated the auditory cue (tone) to the noxious stimulus thus conditioning appropriately, as demonstrated by the increasing amount of freezing time with each tone-noxious stimulus event (refer to table 5.16 for statistical analysis).
B) No differences were observed between *Tardbp*$^{F210I/+}$ and *Tardbp*$^{+/+}$ females for all parameters measured (refer to table 5.16 for statistical analysis). Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean time freezing	SEM	n	p (Anova)
Pre-conditioning Baseline (seconds)	*Tardbp* $^{+/+}$	14.88	3.18	12	0.483
	Tardbp $^{F210I/+}$	19.06	5.28	9	
Tone-Noxious Stimulus 1 (seconds)	*Tardbp* $^{+/+}$	57.43	6.64	12	0.264
	Tardbp $^{F210I/+}$	45.23	8.48	9	
Tone-Noxious Stimulus 2 (seconds)	*Tardbp* $^{+/+}$	76.16	6.03	12	0.301
	Tardbp $^{F210I/+}$	65.00	9.11	9	
Tone-Noxious Stimulus 3 (seconds)	*Tardbp* $^{+/+}$	82.51	6.28	12	0.564
	Tardbp $^{F210I/+}$	76.30	8.90	9	
Context Baseline (%)	*Tardbp* $^{+/+}$	9.92	2.12	12	0.483
	Tardbp $^{F210I/+}$	12.70	3.52	9	
Context (%)	*Tardbp* $^{+/+}$	37.28	5.36	12	0.975
	Tardbp $^{F210I/+}$	37.61	9.84	9	
Cue Baseline (%)	*Tardbp* $^{+/+}$	40.07	6.33	12	0.546
	Tardbp $^{F210I/+}$	34.53	6.07	9	
Cue (%)	*Tardbp* $^{+/+}$	59.39	5.76	12	0.913
	Tardbp $^{F210I/+}$	58.36	7.67	9	
Context-(Context Baseline) (%)	*Tardbp* $^{+/+}$	27.36	5.45	12	0.813
	Tardbp $^{F210I/+}$	24.91	9.32	9	
Cue-(Cue-Baseline) (%)	*Tardbp* $^{+/+}$	19.32	7.19	12	0.619
	Tardbp $^{F210I/+}$	23.83	3.57	9	

Table 5.16: Fear conditioning test results and statistical analysis of TDP-F210I-B6 females

Tardbp$^{F210I/+}$ male animals displayed a trend to freeze for a shorter time in each tone-noxious stimulus event when compared to *Tardbp*$^{+/+}$, however the difference never reached statistical significance (figure 5.28, table 5.17, refer to table for statistical analysis)

In all other parameters measured, there were no significant differences between *Tardbp*$^{F210I/+}$ animals and wild type controls (figure 5.28, table 5.17)

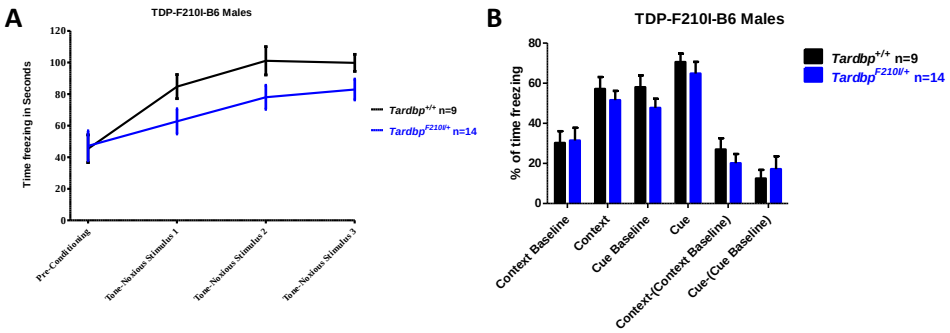

Figure 5.28: Fear conditioning in TDP-F210I-B6 males
A) Both heterozygous and controls associated auditory cue (tone) to the noxious stimulus thus conditioning appropriately, as demonstrated by the increasing amount of freezing time with each tone-noxious stimulus event (refer to table 5.17 for statistical analysis).
B) No differences were observed between *Tardbp*$^{F210I/+}$ and *Tardbp*$^{+/+}$ males for all parameters measured (refer to table 5.17 for statistical analysis). Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean time freezing	SEM	n	p (Anova)
Pre-conditioning Baseline (seconds)	*Tardbp* $^{+/+}$	45.34	8.80	9	0.895
	Tardbp $^{F210I/+}$	47.18	9.44	14	
Tone-Noxious Stimulus 1 (seconds)	*Tardbp* $^{+/+}$	84.70	7.63	9	0.072
	Tardbp $^{F210I/+}$	62.70	7.89	14	
Tone-Noxious Stimulus 2 (seconds)	*Tardbp* $^{+/+}$	101.04	8.94	9	0.066
	Tardbp $^{F210I/+}$	78.00	7.58	14	
Tone-Noxious Stimulus 3 (seconds)	*Tardbp* $^{+/+}$	99.77	5.37	9	0.083
	Tardbp $^{F210I/+}$	82.88	6.56	14	
Context Baseline (%)	*Tardbp* $^{+/+}$	30.23	5.87	9	0.895
	Tardbp $^{F210I/+}$	31.45	6.29	14	
Context (%)	*Tardbp* $^{+/+}$	57.18	5.92	9	0.461
	Tardbp $^{F210I/+}$	51.58	4.60	14	
Cue Baseline (%)	*Tardbp* $^{+/+}$	58.04	5.87	9	0.175
	Tardbp $^{F210I/+}$	47.68	4.56	14	
Cue (%)	*Tardbp* $^{+/+}$	70.57	4.19	9	0.480
	Tardbp $^{F210I/+}$	64.85	5.76	14	
Context-(Context Baseline) (%)	*Tardbp* $^{+/+}$	26.95	5.63	9	0.361
	Tardbp $^{F210I/+}$	20.13	4.60	14	
Cue-(Cue-Baseline) (%)	*Tardbp* $^{+/+}$	12.53	4.26	9	0.601
	Tardbp $^{F210I/+}$	17.17	6.41	14	

Table 5.17: Fear conditioning test results and statistical analysis of TDP-F210I-B6 males

5.3.2.2 TDP-M323K-B6

TDP-M323K-B6 males and females were tested at the age of 24 months; however the experimental numbers for these cohorts were suboptimal. The results of the fear conditioning test in the TDP-M323K-B6 cohorts, with the caveat that experimental numbers are suboptimal, are given below.

Tardbp$^{M323K/+}$ females conditioned to the coupling of the innocuous and noxious stimuli appropriately (figure 5.29, table 5.18, refer to table for statistical analysis). Upon actual fear conditioning testing, the only parameter that was significantly different from *Tardbp*$^{+/+}$ controls was the "Cue-(Cue Baseline)" freezing time (mean difference between cue baseline and freezing in response to the cue 31.65±2.46 for *Tardbp*$^{+/+}$, n=6 and 19.69±4.7 for *Tardbp*$^{M323K/+,}$ n=6; p=0.041, Anova) (figure 5.29 and table 5.19), which is a classical phenotype associated with amygdala dysfunction.

However, given that the heterozygous animals have a higher cue baseline, albeit falling below statistical significance (mean cue baseline 29.90±8.37 for *Tardbp*$^{+/+}$, n=6 and 41.14±9.28 for *Tardbp*$^{M323K/+,}$ n=6; p=0.071, Anova), as displayed in figure 5.29, and a similar percentage of time freezing in response to the cue, the statistical significance observed in the "Cue-(Cue Baseline)" between the genotypes could be a reflection of the higher cue baseline and underpowered sample size and thus not necessarily representing an actual dysfunction in the fear response (figure 5.29 and table 5.19).

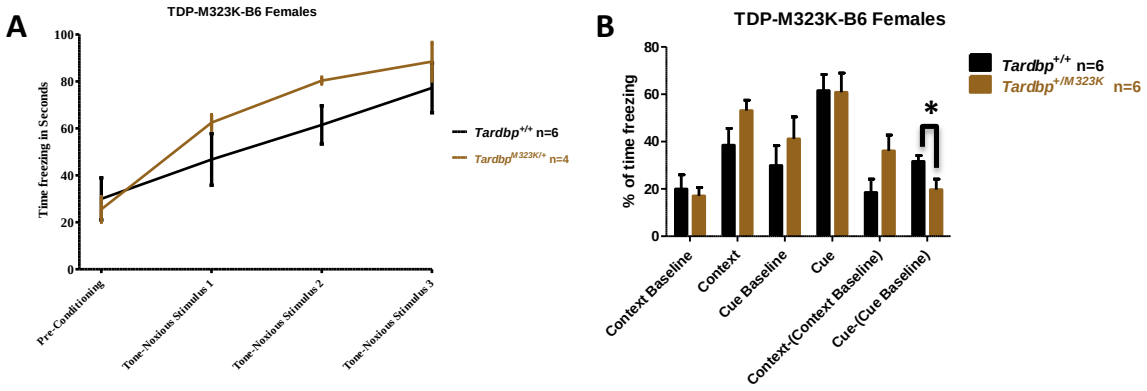

Figure 5.29: Fear conditioning in TDP-M323K-B6 females
A) Both heterozygous and controls associated auditory cue (tone) to the noxious stimulus thus conditioning appropriately, as demonstrated by the increasing amount of freezing time with each tone-noxious stimulus event (refer to table 5.19 for statistical analysis).
B) Of all parameters measured the only statistical significant difference between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ females was observed in the Cue-(Cue baseline) parameter (*p=0.041), which is unlikely to represent a dysfunction in the fear conditioning response (see main text for further interpretation of this difference). No significant differences were observed between genotypes for all parameters measured (refer to table 5.19 for statistical analysis). Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean time freezing	SEM	n	p (Anova)
Pre-conditioning Baseline (seconds)	*Tardbp* $^{+/+}$	29.97	8.94	6	0.678
	Tardbp $^{M323K/+}$	25.50	5.38	6	
Tone-Noxious Stimulus 1 (seconds)	*Tardbp* $^{+/+}$	46.77	11.00	6	0.295
	Tardbp $^{M323K/+}$	62.48	3.34	4	
Tone-Noxious Stimulus 2 (seconds)	*Tardbp* $^{+/+}$	61.48	8.13	6	0.103
	Tardbp $^{M323K/+}$	80.33	1.40	4	
Tone-Noxious Stimulus 3 (seconds)	*Tardbp* $^{+/+}$	77.22	10.54	6	0.467
	Tardbp $^{M323K/+}$	88.43	8.13	4	
Context Baseline (%)	*Tardbp* $^{+/+}$	19.98	5.96	6	0.678
	Tardbp $^{M323K/+}$	17.00	3.59	6	
Context (%)	*Tardbp* $^{+/+}$	38.49	7.09	6	0.110
	Tardbp $^{M323K/+}$	53.11	4.40	6	
Cue Baseline (%)	*Tardbp* $^{+/+}$	29.90	8.37	6	0.389
	Tardbp $^{M323K/+}$	41.15	9.28	6	
Cue (%)	*Tardbp* $^{+/+}$	61.56	6.83	6	0.948
	Tardbp $^{M323K/+}$	60.84	8.13	6	
Context-(Context Baseline) (%)	*Tardbp* $^{+/+}$	18.51	5.67	6	0.071
	Tardbp $^{M323K/+}$	36.11	6.64	6	
Cue-(Cue-Baseline) (%)	*Tardbp* $^{+/+}$	31.66	2.46	6	**0.041**
	Tardbp $^{M323K/+}$	19.69	4.47	6	

Table 5.18: Fear conditioning test results and statistical analysis of TDP-M323K-B6 females
Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

Tardbp$^{M323K/+}$ males also conditioned appropriately, with *Tardbp*$^{M323K/+}$ freezing for significantly longer time in the third coupling event of the tone with the noxious stimulus, which is of ambiguous biological relevance given that it only occurred at one time point and that the objective of the conditioning session, which is the successful coupling of the noxious and non-noxious stimulus, was achieved in the heterozygous males and wild type controls (figure 5.30 and table 5.19, refer to table for statistical analysis)

No statistically significant differences were observed in all other parameters measured (figure 5.30 and table 5.19, refer to table for statistical analysis).

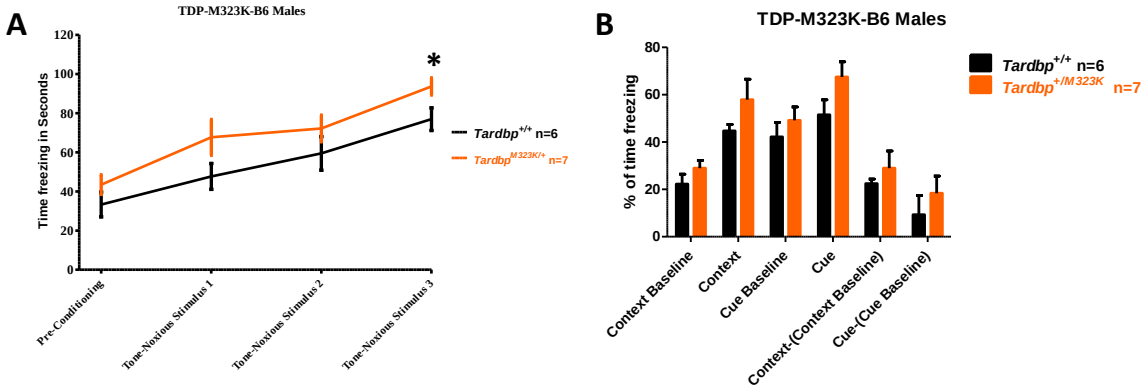

Figure 5.30: Fear conditioning in the TDP-M323K-B6 males
A) Both heterozygous and controls associated auditory cue (tone) to the noxious stimulus thus conditioning appropriately, as demonstrated by the increasing amount of freezing time with each tone-noxious stimulus event.
In the third coupling event, *Tardbp*$^{M323K/+}$ males spent significantly more time freezing (*p=0.037), which is of ambiguous biological relevance (see main text for further interpretation of this difference). Refer to table 5.19 for statistical analysis.
B) No differences were observed between *Tardbp*$^{M323K/+}$ and *Tardbp*$^{+/+}$ males for all other parameters measured (table 5.20). Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean time freezing	SEM	n	p (Anova)
Pre-conditioning Baseline (seconds)	*Tardbp* $^{+/+}$	33.33	6.27	6	0.227
	Tardbp $^{M323K/+}$	43.43	4.95	7	
Tone-Noxious Stimulus 1 (seconds)	*Tardbp* $^{+/+}$	47.67	6.61	6	0.113
	Tardbp $^{M323K/+}$	67.64	9.10	7	
Tone-Noxious Stimulus 2 (seconds)	*Tardbp* $^{+/+}$	59.43	8.59	6	0.261
	Tardbp $^{M323K/+}$	72.17	6.71	7	
Tone-Noxious Stimulus 3 (seconds)	*Tardbp* $^{+/+}$	76.93	5.78	6	**0.037**
	Tardbp $^{M323K/+}$	93.66	4.28	7	
Context Baseline (%)	*Tardbp* $^{+/+}$	22.22	4.18	6	0.227
	Tardbp $^{M323K/+}$	28.95	3.30	7	
Context (%)	*Tardbp* $^{+/+}$	44.68	2.76	6	0.199
	Tardbp $^{M323K/+}$	57.93	8.61	7	
Cue Baseline (%)	*Tardbp* $^{+/+}$	42.19	6.11	6	0.422
	Tardbp $^{M323K/+}$	49.15	5.67	7	
Cue (%)	*Tardbp* $^{+/+}$	51.50	6.35	6	0.107
	Tardbp $^{M323K/+}$	67.48	6.43	7	
Context-(Context Baseline) (%)	*Tardbp* $^{+/+}$	22.46	1.88	6	0.438
	Tardbp $^{M323K/+}$	28.98	7.27	7	
Cue-(Cue-Baseline) (%)	*Tardbp* $^{+/+}$	9.31	8.10	6	0.425
	Tardbp $^{M323K/+}$	18.33	7.33	7	

Table 5.19: Fear conditioning test results and statistical analysis in TDP-M323K-B6 males
Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

5.3.2.3 TDP-M323K-B6-DBA

In order to validate any heterozygous phenotypes and test homozygous animals TDP-M323K-B6-DBA males were tested in the fear conditioning automated set up at ages between 29 and 40 weeks.

Association between the cue and the noxious stimulus worked well across all genotypes without any significant differences detected in freezing times at each event measured (figure 5.31, table 5.20, refer to table for statistical analysis).

Remarkably, *Tardbp*$^{M323K/M323K}$ had significantly higher difference in freezing times in response to the cue when compared to the baseline [Cue-(Cue Baseline] when compared to wild type controls (mean difference in percentage of time freezing in response to the cue when compared to baseline 26.93±3.19, n=12 for *Tardbp*$^{+/+}$ and 42.24±4.03, n=11 for *Tardbp*$^{M323K/M323K}$; p=0.005).

Moreover, heterozygous animals, *Tardbp*$^{M323K/+}$, also displayed a similar trend to freeze for longer in response to the cue (mean difference in percentage of time freezing in response to the cue when compared to baseline 26.93±3.19, n=12 for *Tardbp*$^{+/+}$ and 36.44±1.86, n=12 for *Tardbp*$^{M323K/+}$; p=0.106), which did not reach statistical significance (figure 5.31 and table 5.21) but suggests the effect could be dose-dependent on the *Tardbp*^{M323K} allele. The same trend is observed in *Tardbp*$^{M323K/+}$ males in the C57BL/6J background, despite a greater amplitude in the SEM due to the reduced sample size (figure 5.30, table 5.19).

Additionally, the results from testing this cohort, despite the sex difference, strengthen the probability that the statistical difference seen in the in the Cue-(Cue Baseline) in *Tardbp*$^{M323K/+}$ females (C57BL/6J background) and wild type controls (figure 5.29, table 5.18) results from the difference in Cue Baseline and the effect of the suboptimal sample size.

However a conclusive answer can only be obtained by testing further animals in the TDP-M323K-B6 cohort.

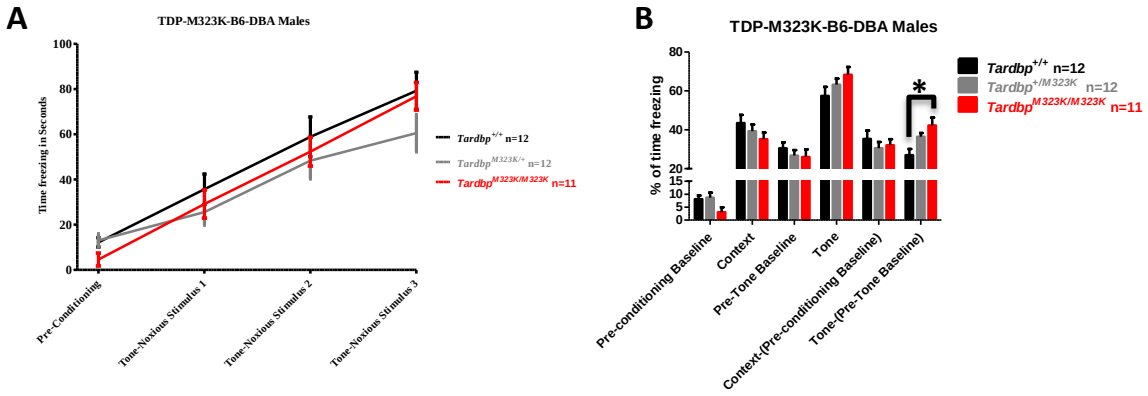

Figure 5.31: Fear conditioning in TDP-M323K-B6-DBA males
A) Animals of all *Tardbp*^{M323K} genotypes associated auditory cue (tone) to the noxious stimulus thus conditioning appropriately, as demonstrated by the increasing amount of freezing time with each tone-noxious stimulus event, thus conditioning appropriately.
B) An increase in the difference between the time spend freezing after the cue and the baseline [Cue-(Cue Baseline)] is seen in *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ when compared to controls, reaching statistical significance between homozygous and control animals (*p=0.005). No additional statistically significant differences were seen in all other parameters measured. Refer to table 5.20 for statistical analysis. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean time freezing	SEM	n	p (Anova & posthoc Bonferroni)
Pre-conditioning Baseline (seconds)	*Tardbp* $^{M323K/M323K}$	4.58	2.86	11	
	Tardbp $^{+/+}$	12.18	2.10	12	0.153
	Tardbp $^{M323K/+}$	13.03	2.88	12	0.093
Tone-Noxious Stimulus 1 (seconds)	*Tardbp* $^{M323K/M323K}$	29.15	6.24	11	
	Tardbp $^{+/+}$	35.67	6.70	12	1
	Tardbp $^{M323K/+}$	25.57	5.72	12	1
Tone-Noxious Stimulus 2 (seconds)	*Tardbp* $^{M323K/M323K}$	52.30	6.36	11	
	Tardbp $^{+/+}$	58.88	8.77	12	1
	Tardbp $^{M323K/+}$	48.33	8.02	12	1
Tone-Noxious Stimulus 3 (seconds)	*Tardbp* $^{M323K/M323K}$	76.78	6.10	11	
	Tardbp $^{+/+}$	79.24	8.17	12	1
	Tardbp $^{M323K/+}$	60.50	8.24	12	0.437
Context Baseline (%)	*Tardbp* $^{M323K/M323K}$	3.05	1.91	11	
	Tardbp $^{+/+}$	8.12	1.40	12	0
	Tardbp $^{M323K/+}$	8.69	1.92	12	0.093
Context (%)	*Tardbp* $^{M323K/M323K}$	35.15	3.41	11	
	Tardbp $^{+/+}$	43.44	4.21	12	0.389
	Tardbp $^{M323K/+}$	39.28	3.46	12	1
Cue Baseline (%)	*Tardbp* $^{M323K/M323K}$	26.03	3.89	11	
	Tardbp $^{+/+}$	30.51	2.96	12	1
	Tardbp $^{M323K/+}$	26.73	2.81	12	1
Cue (%)	*Tardbp* $^{M323K/M323K}$	68.26	3.99	11	
	Tardbp $^{+/+}$	57.44	4.60	12	0.192
	Tardbp $^{M323K/+}$	63.17	3.14	12	1
Context-(Context Baseline) (%)	*Tardbp* $^{M323K/M323K}$	32.10	3.03	11	
	Tardbp $^{+/+}$	35.32	4.22	12	1
	Tardbp $^{M323K/+}$	30.59	3.10	12	1
Cue-(Cue-Baseline) (%)	*Tardbp* $^{M323K/M323K}$	42.24	4.03	11	
	Tardbp $^{+/+}$	26.93	3.19	12	**0.005**
	Tardbp $^{M323K/+}$	36.44	1.86	12	0.596

Table 5.20: Fear conditioning test results and statistical analysis in TDP-M323K-B6-B6 males
Please note that the p-values given are for the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$ and the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{M323K/+}$ which are given in the table row corresponding to the wild type and heterozygous animals. Hence, table rows referent to the homozygous genotype (*Tardbp*$^{M323k/M323K}$) do not have a p value. Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

The observed increase in freezing time in response to the cue is not compatible with a learning and memory deficit; the opposite effect would be expected (i.e. a reduction in the freezing time in response to the cue) if memory of the non-noxious and noxious stimuli had not been consolidated.

Thus, the biological significance of the phenotype observed in *Tardbp*$^{M323K/M323K}$ male mice is not immediately apparent. However, the inferences from the increase in "freezing time" in *Tardbp*$^{M323K/M323K}$ mice in response to the cue lead to the formulation of three hypotheses:

1. The observed phenotype is part of the intrinsic variation in the fear response and does not have *de facto* biological significance. Testing further cohorts should enable the determination of whether this is a robust phenotype that was not identified as part of the test variability.
2. The observed phenotype is a consequence of the shock being perceived as more painful to *Tardbp*$^{M323K/M323K}$ animals when compared to controls, suggesting alterations in nociception or higher integration of nociceptive information. However, the failure to observe a difference in the contextual fear conditioning makes this hypothesis less likely.
3. The observed phenotype is associated with dysfunction in brain areas associated with emotional responses (e.g. amygdala) and therefore there is an enhanced fear response to a cue associated with painful stimuli. Interestingly, a similar effect has been observed in AD animal models (España et al. 2010) which provides support to this hypothesis. Additional behavioural testing, including the novel object test and the free exploratory paradigm, will be done to further test this hypothesis.

5.3.3 Pre-pulse inhibition and acoustic startle response

The acoustic startle response consists of a reflex behaviour triggered by an unexpected acoustic stimulus. In the system used, the acoustic startle stimulus is a high frequency tone of different amplitudes, reaching a maximum amplitude of 110 decibels (dB).

Pre-pulse inhibition (PPI) consists on the attenuation of the acoustic startle response when a pulse of white noise precedes the acoustic stimulus, thus causing a reduction in the amplitude of the startle response. The PPI response was assessed with pulse of white noise with the

amplitudes of 60, 65 and 75 decibels preceding the acoustic startle stimulus and these are referred to PPI60, PPI70 and PPI75, respectively.

Pre-pulse inhibition measures how sensory information (in this case an acoustic stimulus) is transmitted to the motor system. It specifically measures the effect of pre-pulse of white noise in suppressing of the amplitude of the motor response, thus constituting a reflection of the influence of the activity of inhibitory interneurones over the sensory neurones relaying the information (Nusbaum & Contreras 2004). Changes in the PPI therefore reflect alterations in the transmission of the neuronal networks involved in sensorimotor gating.

A deficit in the pre-pulse inhibition is seen in patients diagnosed with schizophrenia (but is not pathognomonic or prognostic), in whom the attenuation of the acoustic startle, when preceded by a pulse of white noise, is partly inhibited (Braff et al. 1992; Braff et al. 2005; Swerdlow et al. 2006; Kumari et al. 2014) and in several other neuropsychiatric disorders characterised by synaptic dysfunction (Kohl et al. 2013).

Moreover, different PPI deficits are also consistently found in mice with lesions (e.g. transgenic, mutants, knockouts) in genes associated with the pathophysiology of neuropsychiatric diseases, including autism spectrum disorder and schizophrenia (Heldt et al. 2004; Gómez-Sintes et al. 2014). A review of the literature has found it to be a reliable endophenotype of neuronal dysfunction and therefore useful in the identification of relevant animal models and in evaluating the efficiency of therapeutic interventions (Swerdlow et al. 2008).

TDP-F210I-B6 and TDP-M323K-B6 cohorts were tested in the automated startle response and pre-pulse inhibition system at 22 weeks.

In the system used, mice are placed in perforated transparent plastic tubes in individually acoustic-isolated cages and the acoustic startle response, either on its own or preceded by pulses of white noise, measured by proprietary software, enabling the PPI to be calculated.

5.3.3.1 TDP-F210I-B6

In TDP-F210I-B6 of both sexes the startle response to a tone of 110 decibels and PPI caused by pulses of white noise (with the amplitudes of 65, 70 and 75 dB) preceding the acoustic stimulus in the heterozygous (*Tardbp*$^{F210I/+}$) animals was comparable to wild type controls (*Tardbp*$^{+/+}$), without any significant differences being detected (figures 5.32 and 5.33; tables 5.21 and 5.22, refer to tables for statistical analysis).

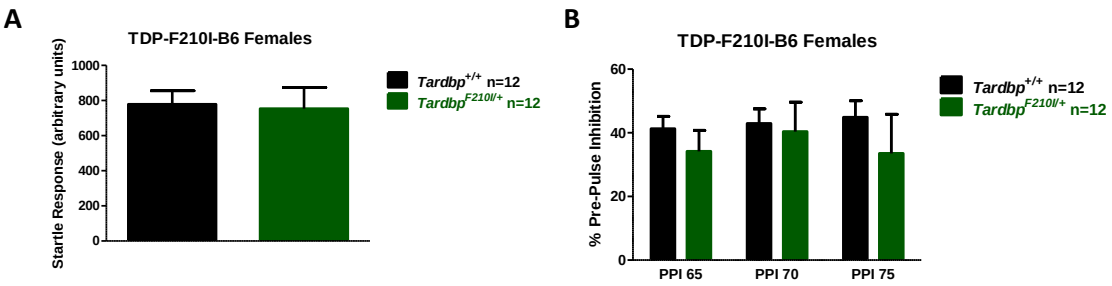

Figure 5.32: Startle response and PPI in TDP-F210I-B6 females
No significant differences were identified between *Tardbp*$^{F210I/+}$ females and controls in the (A) acoustic startle response or (B) PPI when the acoustic stimulus was preceded by pulses of white noise. Refer to table 5.21 for statistical analysis. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean	SEM	n	p (Anova)
Startle 110 dB (arbitrary units)	*Tardbp* $^{+/+}$	778.79	76.41	12	0.864
	Tardbp $^{F210I/+}$	754.08	120.23	12	
PPI 65 dB (%)	*Tardbp* $^{+/+}$	41.28	3.90	12	0.365
	Tardbp $^{F210I/+}$	34.19	6.60	12	
PPI 70 dB (%)	*Tardbp* $^{+/+}$	42.92	4.63	12	0.807
	Tardbp $^{F210I/+}$	40.35	9.26	12	
PPI 75 dB (%)	*Tardbp* $^{+/+}$	44.85	5.20	12	0.406
	Tardbp $^{F210I/+}$	44.85	5.20	12	

Table 5.21: Results and statistical analysis of the startle response and pre-pulse inhibition in TDP-F210I-B6 females.

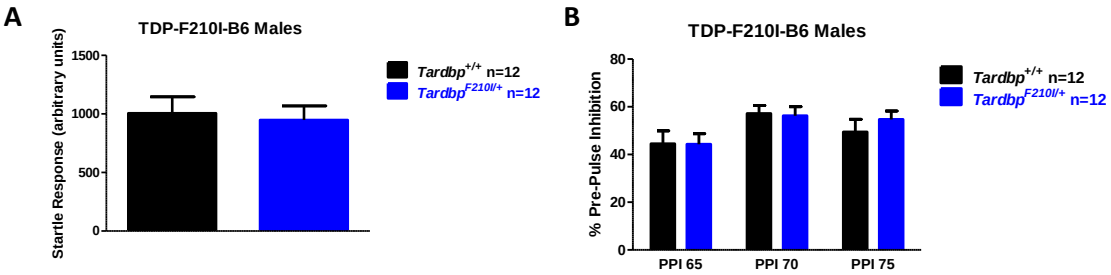

Figure 5.33: Startle response and PPI in TDP-F210I-B6 males
No significant differences were identified between *Tardbp*$^{F210I/+}$ males and controls in the (A) acoustic startle response or (B) PPI when the acoustic stimulus was preceded by pulses of white noise. Refer to table 5.22 for statistical analysis. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean	SEM	n	p (Anova)
Startle 110 dB (arbitrary units)	*Tardbp* $^{+/+}$	1005.36	140.79	12	0.758
	Tardbp $^{F210I/+}$	947.67	120.17	12	
PPI 65 dB (%)	*Tardbp* $^{+/+}$	44.49	5.44	12	0.975
	Tardbp $^{F210I/+}$	44.27	4.45	12	
PPI 70 dB (%)	*Tardbp* $^{+/+}$	57.19	3.41	12	0.853
	Tardbp $^{F210I/+}$	56.22	3.84	12	
PPI 75 dB (%)	*Tardbp* $^{+/+}$	49.42	5.35	12	0.424
	Tardbp $^{F210I/+}$	54.67	3.59	12	

Table 5.22: Results and statistical analysis of the startle response and pre-pulse inhibition in TDP-F210I-B6 males.

5.3.3.2 TDP-M323K-B6

In TDP-M323K-B6 the startle response to an acoustic stimulus of 110 decibels and pre-pulse inhibition caused by white noise pulse of 65, 70 and 75 decibels in the heterozygous (*Tardbp*$^{M323K/+}$) animals was comparable to wild type controls (*Tardbp*$^{+/+}$), without any significant differences being detected (figures 5.34 and 5.35; tables 5.23 and 5.24, refer to tables for statistical analysis).

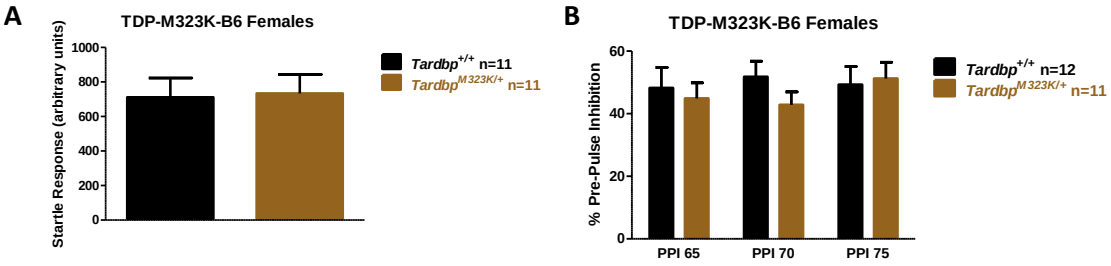

Figure 5.34: Startle response and PPI in TDP-M323K-B6 females
No significant differences were identified between *Tardbp*$^{M323K/+}$ females and controls in the (A) acoustic startle response or (B) PPI when the acoustic stimulus was preceded by pulses of white noise. Refer to table 5.23 for statistical analysis. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean	SEM	n	p (Anova)
Startle 110 dB (arbitrary units)	*Tardbp* $^{+/+}$	711.15	111.62	11	0.886
	Tardbp $^{M323K/+}$	733.83	109.57	11	
PPI 65 dB (%)	*Tardbp* $^{+/+}$	48.22	6.57	11	0.689
	Tardbp $^{M323K/+}$	44.87	5.00	11	
PPI 70 dB (%)	*Tardbp* $^{+/+}$	51.78	4.99	11	0.183
	Tardbp $^{M323K/+}$	42.83	4.16	11	
PPI 75 dB (%)	*Tardbp* $^{+/+}$	49.28	5.79	11	0.807
	Tardbp $^{M323K/+}$	51.21	5.21	11	

Table 5.23: Results and statistical analysis of the startle response and pre-pulse inhibition in TDP-M323K-B6 females.

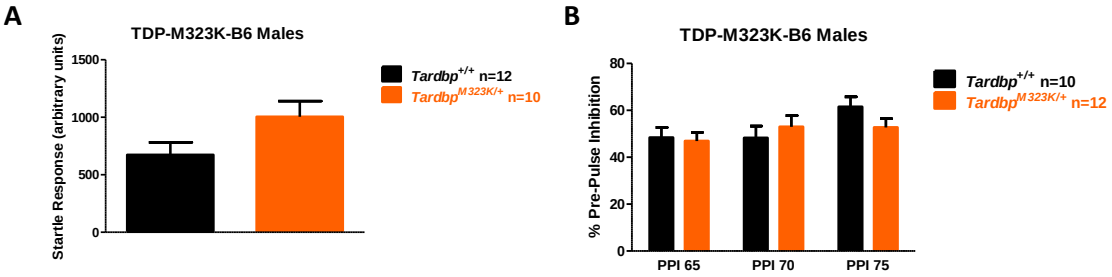

Figure 5.35: Startle response and PPI in TDP-M323K-B6 males
No significant differences were identified between *Tardbp*$^{M323K/+}$ males and controls in the (A) acoustic startle response or (B) PPI when the acoustic stimulus was preceded by pulses of white noise of white noise. Refer to table 5.24 for statistical analysis. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean	SEM	n	p (Anova)
Startle 110 dB (arbitrary units)	*Tardbp* $^{+/+}$	672.57	109.55	10	0.080
	Tardbp $^{M323K/+}$	1004.45	136.44	12	
PPI 65 dB (%)	*Tardbp* $^{+/+}$	48.36	4.35	10	0.798
	Tardbp $^{M323K/+}$	46.89	3.64	12	
PPI 70 dB (%)	*Tardbp* $^{+/+}$	48.23	5.09	10	0.510
	Tardbp $^{M323K/+}$	52.95	4.84	12	
PPI 75 dB (%)	*Tardbp* $^{+/+}$	61.48	4.30	10	0.141
	Tardbp $^{M323K/+}$	52.73	3.78	12	

Table 5.24: Results and statistical analysis of the startle response and pre-pulse inhibition in TDP-M323K-B6 males.

5.3.3.3 TDP-M323K-B6-DBA

TDP-M323K-B6-DBA male animals were tested at ages between 32 and 49 weeks.

The startle response did not have any significant differences across genotypes (figure 5.35 and table 5.25, refer to table for statistical analysis). Notwithstanding, pre-pulse inhibition was decreased in *Tardbp*$^{M323K/+}$ animals and *Tardbp*$^{M323K/M323K}$ animals when compared to wild type controls (figure 5.35 and table 5.25).

The attenuation in pre-pulse inhibition did not reach statistical significance in the heterozygous animals at any of the white noise amplitudes when compared to controls (figure 5.35 and table 5.25). Nevertheless, the reduction in pre-pulse inhibition was significant in *Tardbp*$^{M323K/M323K}$ males (figure 5.35 and table 5.25) when pulses of white noise of 70 dB (mean percentage of pre-pulse inhibition 61.91±7.33, n=12 for *Tardbp*$^{+/+}$ and 20.38±12.53, n=11 for *Tardbp*$^{M323K/M323K}$; p=0.015, Anova and post-hoc Bonferroni) and 75 dB (mean percentage of pre-pulse inhibition 59.28±7.83, n=12 for *Tardbp*$^{+/+}$ and 17.51±9.58, n=11 for *Tardbp*$^{M323K/M323K}$; p=0.011, Anova and post-hoc Bonferroni) preceded the acoustic startle stimulus.

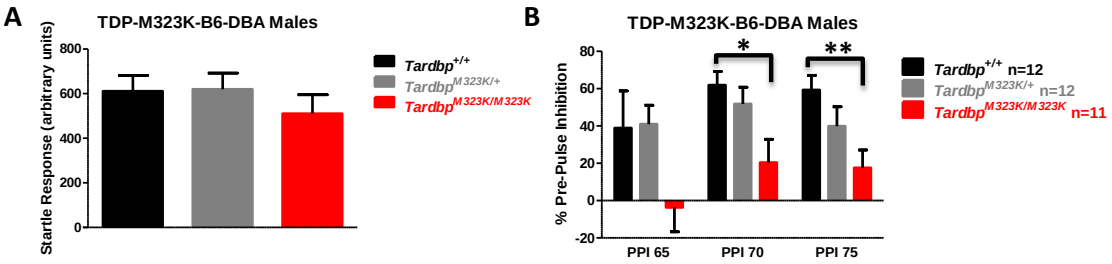

Figure 5.36: Startle response and PPI in TDP-M323K-B6 males
A) No significant differences were found in the startle response across genotypes in *Tardbp*^{M323K} males in the hybrid C57BL/6J-DBA background (refer to table 5.25 for statistical analysis).
B) *Tardbp*$^{M323K/+}$ have a non-significant reduction in pre-pulse inhibition when compared to *Tardbp*$^{+/+}$ (table 5.25). *Tardbp*$^{M323K/M323K}$ have a significant reduction in the pre-pulse inhibition compared to controls when pulses of white of 70 dB (*p=0.015, Anova and posthoc Bonferroni) and 75 dB (**p=0.011, Anova and posthoc Bonferroni) precede the acoustic startle stimulus. Error bars in all graphs are ±SEM.

Measurement (units)	Genotype	Mean	SEM	n	p (Anova & posthoc Bonferroni)
Startle 110 dB (arbitrary units)	*Tardbp* $^{M323K/M323K}$	510	84.67	11	
	Tardbp $^{+/+}$	611	70.25	12	1.000
	Tardbp $^{M323K/+}$	620	72.77	12	0.953
PPI 65 dB (%)	*Tardbp* $^{M323K/M323K}$	-4	13.11	11	
	Tardbp $^{+/+}$	39	19.94	12	0.172
	Tardbp $^{M323K/+}$	41	10.22	12	0.141
PPI 70 dB (%)	*Tardbp* $^{M323K/M323K}$	20	12.53	11	
	Tardbp $^{+/+}$	62	7.33	12	**0.015**
	Tardbp $^{M323K/+}$	52	8.93	12	0.088
PPI 75 dB (%)	*Tardbp* $^{M323K/M323K}$	18	9.58	11	
	Tardbp $^{+/+}$	59	7.83	12	**0.011**
	Tardbp $^{M323K/+}$	40	10.45	12	0.310

Table 5.25: Results and statistical analysis of the startle response and pre-pulse inhibition in TDP-M323K-B6-DBA males.
Please note that the p-values given are for the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{+/+}$ and the comparison between *Tardbp*$^{M323K/M323K}$ and *Tardbp*$^{M323K/+}$ which are given in the table row corresponding to the wild type and heterozygous animals. Hence, table rows referent to the homozygous genotype (*Tardbp*$^{M323k/M323K}$) do not have a p value. Statistical significance is highlighted in the table by colouring the parameter measured and p-value numbers in red.

Having tested the TDP-M323K-B6-DBA animals within age ranges of almost 4 months was suboptimal. Nevertheless, given the time constraints and impossibility to establish a phenotyping cohort within the time frame of this DPhil, testing the startle response and pre-pulse inhibition in the available animals was a justified undertaking.

Furthermore, the mean ages across genotypes were comparable and no significant differences existed (Table 5.26), which enables making inferences on the results obtained from testing these animals.

				95% Confidence Interval			
Genotype	n	Mean Age	SEM	Lower Bound	Upper Bound	Minimum	Maximum
Tardbp$^{+/+}$	12	41.25	1.75	37.40	45.10	32	49
Tardbp$^{M323K/+}$	12	42.26	1.47	39.02	45.51	35.86	48.86
Tardbp$^{M323K/M323K}$	11	40.83	1.89	36.63	45.03	32.14	49

Table 5.26: Statistical analysis of age distribution in the TDP-M323K-B6-DBA animal cohort at the time of Acoustic Startle and pre-pulse inhibition testing

The significant reduction in pre-pulse inhibition in *Tardbp*$^{M323K/M323K}$ animals, if confirmed with repeated testing, is suggestive of dysregulation of synaptic transmission in neuronal circuits associated with schizophrenia and other neuropsychiatric disorders.

5.3.4 Circadian phenotyping: TDP-F210I-B6, TDP-M323K-B6

Sleep disturbances and alterations in the circadian rhythms are often associated with cognitive dysfunction, including frontotemporal dementia (Anderson et al. 2009), and therefore mice were phenotyped in order to assess the influence of the F210I and M323K mutations in circadian rhythmicity.

TDP-F210I-B6 and TDP-M323K-B6 female animal cohorts were phenotyped, at the ages of 85 and 86 weeks respectively, in the automated circadian system set up at the MRC Harwell in collaboration with Gareth Banks, in the neurobehaviour group.

In this set up, mice are singly housed in individual cages, and activity monitored on a running wheel; being initially entrained to a twelve hour light/dark cycle and allowed to run freely on its individual wheels. After the entrainment period, wheel running is also recorded in constant darkness and constant light conditions.

Given that the animals were necessarily individually caged, analogously to the home cage wheel-running test, only females were tested as male animals would not tolerate being re-caged together after being individually housed.

The activity of the animal on the running wheels is recorded and automatically integrated with the light/dark conditions and the output is given as a specialised activity over time graph, designated “actogram”, where the height of the vertical bars in the y axis represent activity and the x axis time, with the shading representing “lights on” or “lights off”. Measurements taken include “*Tau”*, which constitutes the period of the animal’s internal circadian clock in constant conditions, “counts”, which constitute the quantification of wheel running activity and “Alpha”, which constitutes the duration of the active phase during the light and dark periods.

In addition, the "amplitude", which constitutes the difference between the peak (or trough) and the mean value of the rhythm, and the "phase angle of entrainment", which constitute the difference between the rhythm of the entraining light onset pulse and the onset of activity are also measured.

A comprehensive review of the circadian testing system used for *Tardbp*^{F210I} and *Tardbp*^{M323K} is available in (Banks & Nolan 2011).

5.3.4.1 TDP-F210I-B6

Tardbp$^{F210I/+}$ females had increased proportion of activity in the light phase (p=0.040, refer to "% light counts" in table 5.27) in the wheels and had a significantly higher "Alpha" in constant darkness (p=0.0156, refer to "Alpha DD" in table 5.27) and during the light/dark entrainment period (p=0.024, refer to "Alpha LD" in table 5.27) when compared to *Tardbp*$^{+/+}$ controls (figure 5.37 and table 5.27).

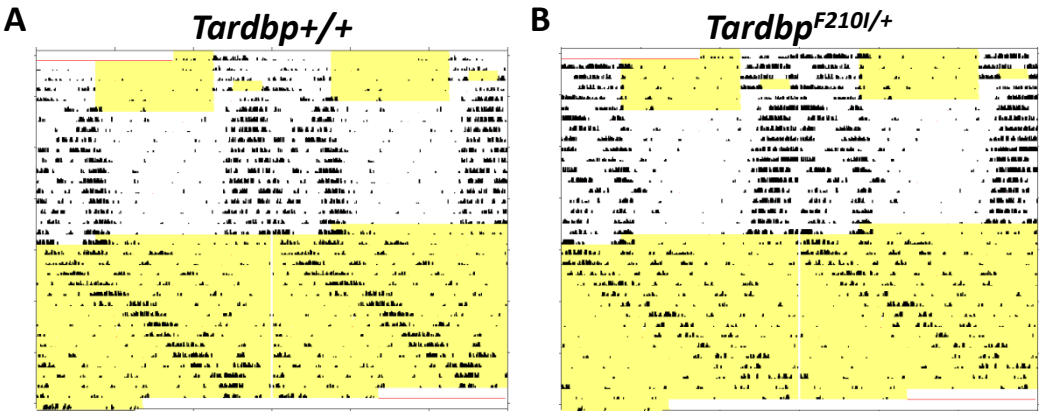

Figure 5.37: Representative actograms of TDPF210I-B6 females.
Graphical representation of the circadian phenotyping data from (A) *Tardbp*$^{+/+}$ and (B) *Tardbp*$^{F210I/+}$. Yellow shaded areas representing "lights on" and the "white shaded areas" representing "lights off". The vertical black bars represent wheel running activity. No differences were identified on the circadian rhythm of heterozygous animal when compared to controls. Refer to table 5.27 for statistical analysis.

					LD Activity					DD activity	LL activity	Amplitude	Alpha					Phase angle
	Tau			LP	Light_counts	Dark_Counts	Total_Counts	% Light_counts	Avg_Counts	Avg_Counts	Avg_Counts	LD	DD	LL	LD	DD	LL	
	DD	LL	LL-DD		Counts/day	Counts/day	Counts/day	%Light_counts	Counts/min	Counts/min	Counts/min							
Tardbp$^{+/+}$	24	26	1.65	36.7	104.56	1125.5	1230.06	9.429	0.8542	0.9589	0.4172	500.59	874.515	748.57	11.538	11.743	6.016	178.95
Tardbp$^{F210I/+}$	24	26	1.68	34	295.1333	1862.156	2157.289	13.6444444	1.498	1.97089	0.57433	532.32	897.836	700.5	12.309	12.989	7.4567	178.5
Tardbp$^{+/+}$ SEM	0	0	0.06	17.1	36.19569	414.8468	442.5431	1.00659968	0.30729	0.28969	0.10054	24.595	71.4042	83.186	0.2233	0.33	0.4494	1.703673
Tardbp$^{F210I/+}$ SEM	0	0.1	0.05	14.8	108.3412	567.8097	666.9237	1.66569614	0.46311	0.48547	0.1784	14.86	54.2774	108.96	0.214	0.3231	0.7348	1.394433
p (Anova)	0.3	0.3	0.76	0.91	0.099454	0.30283	0.253922	**0.04049067**	0.25396	0.08424	0.44149	0.2981	0.80142	0.7271	**0.024**	**0.016**	0.1052	0.842623

Table 5.27: Results and statistical analysis of circadian phenotyping in TDP-F210I-B6 females

The results are not suggestive of a disturbance in the circadian rhythm of *Tardbp*$^{F210I/+}$ animals. However the increased in the Alpha in constant darkness and light/dark cycle ("Alpha DD" and "Alpha LD") could be suggestive that the heterozygous females are active for a longer period in the running wheels, which would be coherent with the trend displayed by animals in the same genotype to run for longer distances (albeit failing to reach statistical significance) in the Home Cage automated-wheel running system (section 5.2.4). The increased activity in the running wheels (% light counts in table 5.28) can also therefore also be a reflection of the animals running for longer periods.

5.3.4.2 TDP-M323K-B6

For TDP-M323K-B6 female animals, no significant differences were found in any of the parameters measured (figure 5.38 and table 5.28).

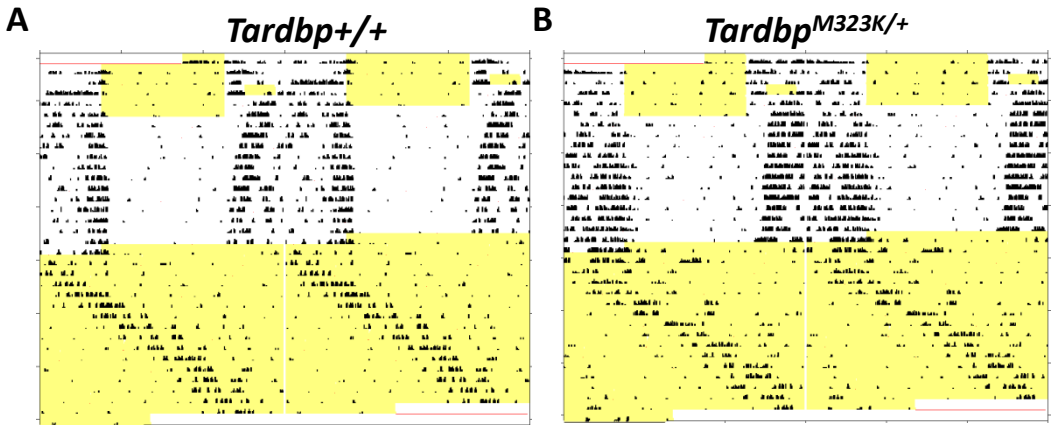

Figure 5.38: Representative actograms of TDP-M323K-B6 females.
Graphical representation of the circadian phenotyping data from (A) *Tardbp*$^{+/+}$ and (B) *Tardbp*$^{F210I/+}$. Yellow shaded areas representing "lights on" and the "white shaded areas" representing "lights off". The vertical black bars represent wheel running activity. No differences were identified on all parameters measured during circadian phenotyping between heterozygous animal when compared to controls. Refer to table 5.28 for statistical analysis.

					LD Activity					DD activity	LL activity	Amplitude	Alpha					Phase Angle
	Tau			LP	Light_counts	Dark_Counts	Total_Counts	% Light_counts	Avg_Counts	Avg_Counts	Avg_Counts	LD	DD	LL	LD	DD	LL	
	DD	LL	LL-DD		Counts/day	Counts/day	Counts/day	%Light_counts	Counts/min	Counts/min	Counts/min							
Tardbp$^{+/+}$	24	26	1.67	35.1	132.7111	877.6889	1010.4	11.5567778	0.70144	0.79467	0.33078	430.67	722.693	578.52	11.281	11.481	6.4389	176.3333
Tardbp$^{M323K/+}$	24	26	1.65	23.9	110.4333	878.6667	989.1	9.8705	0.68683	0.99083	0.31233	468.65	850.533	613.89	11.917	11.202	6.7033	176.3333
Tardbp$^{+/+}$ SEM	0	0	0.05	15.3	64.22547	342.0374	401.7921	1.81151082	0.27899	0.32168	0.1372	15.169	80.7244	56.02	0.2701	0.2986	0.6347	2.949812
Tardbp$^{M323K/+}$ *SEM*	0	0	0.05	10.6	60.60527	413.161	471.7133	1.8997086	0.32753	0.50558	0.16299	36.699	109.993	86.943	0.2771	0.3579	0.5447	2.740641
p (Anova)	0.4	0.4	0.78	0.59	0.814957	0.998577	0.973344	0.54502371	0.97367	0.73551	0.93277	0.2962	0.35519	0.7249	0.1377	0.5609	0.7735	1

Table 5.28: Results and statistical analysis of circadian phenotyping in TDP-M323K-B6 females

5.4 Summary of the *in vivo* characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K} mutants

In the previous chapter, the opposing effects of the F210I and M323K mutations on TDP43-dependent splicing were described with the support of the experimental results.

In vivo characterisation of the mutant animals therefore presented a unique opportunity to investigate whether the "loss of normal splicing function", caused by the F210I mutation, "the augmented alternative exon selection", caused by the M323K mutation, or both, lead to neurodegeneration, neuronal dysfunction or other relevant phenotypes.

Within the limitations of early lethality caused by homozygosity for these mutations in the C57BL/6J background, comprehensive behavioural testing was performed. Only a reduced number of tests were performed in *Tardbp*^{M323K} mice on the C57BL6/J-DBA hybrid background (which lead to viable *Tardbp*$^{M323K/M323K}$) due to time constraints. New cohorts (of TDP-M323K-B6-DBA) are being established to confirm the findings identified and for further testing.

The results from the experimental tests and measurements which were performed and constitute original work presented in this thesis suggest that:

- *Tardbp*$^{F210I/+}$ females have lower body weight and fat mass composition when compared to controls (figure 5.2, table 5.1).
- *Tardbp*$^{M323K/+}$ males display a non-significant trend towards higher body weight (in both the C57BL/6J background and C57BL/6J-DBA hybrid background) whilst *Tardbp*$^{M323K/M323K}$ (C57BL/6J-DBA hybrid background) have significantly lower body weights when compared to *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ (figures 5.5 and 5.6, tables 5.4 and 5.5).

- In heterozygosity, neither the F210I mutation nor the M323K mutation affects brain growth (tables 5.6 and 5.7) or longevity (figures 5.9 and 5.10).
- Comprehensive motor function testing failed to detect any signs of neuromuscular degeneration in *Tardbp*$^{M323K/+}$ (C57BL/6J background).
- In *Tardbp*$^{F210I/+}$ (C57BL/6J background) the horizontal ladder/locotronic test suggests that fine motor or sensory function may be discreetly impaired in the heterozygous animals (figure 5.20 and table 5.14) but no other significant phenotypes associated with neuromuscular degeneration were identified.
- In *Tardbp*$^{M323K/+}$ (C57BL/6J background) no signs of neuromuscular degeneration were identified with the tests performed.
- Test for learning and memory failed to detect phenotypes in *Tardbp*$^{F210I/+}$ mice.
- *Tardbp*$^{M323K/M323K}$ have a significant deficit in cued fear conditioning (C57BL/6J-DBA hybrid background) which is consistent with neophobic behaviour (figure 5.31 and table 5.20) and a classical PPI deficit (figure 5.36 and table 5.25) suggesting that the M323K mutation causes neuronal dysfunction in homozygosity.

The experimental results of the molecular characterisation and *in vivo* characterisation will be discussed in the next chapter, together with the inferences that can be made towards a better understanding of TDP43 biology and neurodegeneration.

Chapter 6.

Discussion and conclusions

In this chapter, the experimental results will be critically discussed and integrated in order to enable inferences on the role of TDP43 in neurodegeneration to be made. Future work considered to be relevant for the characterisation of these mutations will also be outlined and discussed.

TDP43 is highly conserved, ubiquitously expressed protein with multiple roles in RNA metabolism including regulation of gene expression and alternative exon splicing (Ou et al. 1995; Wang et al. 2004; Buratti et al. 2001). Additionally, TDP43 dysfunction is a pivotal factor in pathophysiology of ALS and FTLD-TDP. Firstly, hyperphosphorylated, polyubiquitinated cytoplasmic fragments of TDP43 are histopathological hallmarks in ALS and FTLD-TDP (Neumann et al. 2006; Davidson et al. 2007; Pamphlett et al. 2009) and secondly, TDP43 mutations, the vast majority of which located in the C-terminal, are causative of these diseases (Rutherford et al. 2008; Van Deerlin et al. 2008; Chiò et al. 2010; Borroni et al. 2009).

Moreover, histopathological cytoplasmic TDP43 aggregates have also been identified in Alzheimer's Disease (Tremblay et al. 2011; Wilson et al. 2011; Josephs et al. 2014) and in the primary muscle disease sporadic Inclusion Body Myositis by immunoblot (Hernandez Lain et al. 2011), which involve inflammatory changes and degeneration of the muscle fibres. Therefore, the association of pathological TDP43 features in additional diseases involving brain and muscle degeneration raise the possibility that TDP43 dysfunction is widely involved in degenerative processes of neurones and muscle that lead to different clinical diseases.

Notwithstanding, TDP43's biological functions and its role in neurodegeneration remain incompletely understood given that both over-expression of human wild type or mutant TDP43 and its conditional deletion or knockdown have been shown to lead to neurodegeneration in animal models of TDP43 proteinopathies (Wegorzewska et al. 2009;

Tsai et al. 2010; Wu et al. 2012; Yang et al. 2014). Consequently, it remains unclear how TDP43 dysfunction leads to neurodegeneration even in the context of TDP43 mutations.

Studying the effects of mutations in neuronal dysfunction and degeneration constitutes the most direct path to start dissecting the link between TDP43 dysfunction and neurodegeneration. Once robust scientific data supports specific pathways leading from TDP43 dysfunction to neurodegeneration in the context of mutations, the same pathways can be assessed in cases of sporadic disease with TDP43 histopathology (e.g. sALS, FTLD-TDP, AD), endeavouring to create a roadmap for therapeutic interventions.

The research presented in this thesis was performed with the purpose of gaining further understanding of TDP43's biological functions and the consequences of its biological dysfunction, through the characterisation of ENU mouse mutants, specifically *Tardbp*^{F210I} and *Tardbp*^{M323K}. Moreover, it had a particular interest in translating findings across species and contribute towards a better understanding of the role of TDP43 in human neurodegeneration.

The F210I mutation constitutes of a Phenylalanine to Isoleucine amino acid change at residue 210 of TDP43, located it the RMM2. *Tardbp*^{F210} has been previously molecularly characterised and shown to lead to a "loss of normal splicing function" at specific gene target levels, including *Sort1* and *DnajC5*, without the alteration of TDP43 protein levels (Ricketts 2012). The biological mechanism responsible for the splicing phenotype was investigated through qualitative EMSA methodology and suggested to be due to the reduced affinity between the TDP43-F210I mutant protein and RNA (Ricketts 2012).

The M323K mutation, constitutes a Methionine to Lysine substitution at amino acid residue 323, which is located in the C-terminal of TDP43, where the vast majority of ALS causing mutations are also located (Sreedharan et al. 2008; Rutherford et al. 2008; Van Deerlin et al. 2008; Corrado et al. 2009; Chiang et al. 2012). *Tardbp*^{M323K} has not been studied previously

and the results presented in this thesis are the first reported characterisation of this mouse line. Molecular characterisation of the M323K mutation revealed it leads to "augmented alternative exon selection" in TDP43-dependent splicing.

6.1 The *Tardbp*^{F210I} and *Tardbp*^{M323K} mutations are functional during embryonic development and interact with genetic background of different laboratory mouse strains

In homozygosity, on the C57BL/6J background, the F210I mutation leads to prenatal lethality around 15.5dpc (Ricketts 2012). However, on a C57BL/6J-DBA hybrid background viability is extended until at least 18.5dpc (table 3.4).

The M323K mutation leads to perinatal lethality in the C57BL/6J background, possibly between postnatal days 1 and 5 as indicated by preliminary data (table 3.2). The lethality phenotype in *Tardbp*$^{M323K/M323K}$ is rescued by a C57BL/6J-DBA hybrid background, with the *Tardbp*^{M323K} mutant allele being inherited in a Mendelian pattern by the offspring of heterozygous matings at least at weaning age (table 3.3).

The influence of genetic background in the phenotype of transgenic animals is well documented (Kerr et al. 2013; Tanabe et al. 2012) and genetic background has inclusively been shown to influence the phenotype of most commonly used ALS mouse model, the transgenic SOD1-G93A (Mancuso et al. 2012), which prompted us to breed the *Tardbp* ENU mutants with the DBA strain in addition to C57BL/6J (as described earlier mice were backcrossed at least five generations in order to segregate additional ENU mutations from the TDP43 mutations, refer to section 3.1 in chapter 3). The partial rescue of the lethality phenotype in both, *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{M323K/M323K}$ mutants on hybrid background C57BL/6J-DBA mice thus suggests that mutant TDP43 proteins interact with the genetic

background pool, modifying TDP43 function to the extent of enabling the mutant animals to live longer.

The early lethality observed in the C57BL/6J background in *Tardbp*$^{F210I/F210I}$ and *Tardbp*$^{M323K/M323K}$ demonstrates that these mutations are functional and affect embryonic development. This interpretation is also supported by the very early embryonic lethality which has been described in TDP43 knockout mice (Sephton et al. 2010; Chiang et al. 2010) and ENU mutants which are homozygous for the *Tardbp*^{Q101X} ENU mutation, which leads to an early termination of TDP43 at Glutamine 101, resulting in a functional null allele (Ricketts et al. 2014). In addition, data from different TDP43 genetically manipulated mouse models also clearly indicates that TDP43 plays important roles in embryonic development, such as in the proliferation of the blastocyst's inner cell mass (Wu et al. 2010).

The investigation into the developmental consequences of the F210I and M323K mutations are currently being undertaken in collaboration with colleagues at the Institute of Child Health. Whole-foetus CT scan images at 16.5dpc and 18.5dpc for *Tardbp*^{M323K} mutants in the C57BL/6J background and *Tardbp*^{F210I} mutants in the C57BL/6J-DBA hybrid background will be analysed, aiming to identify structural abnormalities that may explain the early lethality.

If relevant structural abnormalities which are associated with human or animal disease are detected, particularly on *Tardbp*$^{M323K/M323K}$ mutants, in the C57BL/6J background, which are absent in homozygous animals in the hybrid background, a modifier screen could be performed. *Tardbp*$^{M323K/+}$ mice (C57BL/6J background) can be mated *Tardbp*$^{M323K/+}$ mice (DBA background) and the offspring of those matings again mated, producing animals (referred to as F2s) with chromosomal segregations which enable quantitative trait loci mapping and statistical analysis to be performed in order to identify modifier genes of the abnormal development phenotype and associated lethality (Mott & Flint 2013). Thus, the

identification of modifier genes could lead to the identification of biochemical pathways susceptible to pharmacological manipulation and thus a route to correcting TDP43 dysfunction affecting development.

At the start of the experimental work presented in this thesis, TDP43's known biological functions encompassed multiple roles in RNA metabolism, including alternative exon splicing and regulation of gene expression, functions which continued to be studied, characterised and further communicated through scientific publication by several groups (Polymenidou et al. 2011; Tollervey et al. 2011; Bhardwaj et al. 2013; Ricketts et al. 2014). In addition, TDP43 was suggested to autoregulate its own levels (Ayala et al. 2011; Polymenidou et al. 2011; Avendaño-Vázquez et al. 2012; Bembich et al. 2013). In the next sections, the discussion will focus on the results from the molecular characterisation from the *Tardbp*^{M323K}, and further characterisation of *Tardbp*^{F210I} and its implications for the known functions of TDP43 biology.

6.2 The splicing phenotypes of TDP43-F210I and TDP43-M323K and the underlying biological mechanism

At the splicing level, the F210I mutation leads to a "loss of normal splicing function" (Ricketts 2012) which replicates the effects of TDP43 knockdown (Polymenidou et al. 2011). The biological mechanism leading to the "loss of normal splicing function" caused by the F210I mutation is suggested to consist of the reduction in the affinity between mutant protein and RNA as demonstrated experimentally by a qualitative EMSA methodology by Thomas Ricketts (Ricketts 2012). The disruption of TDP43-RNA binding by a mutation in the RRM2 constituted in itself a relevant contribution to the field of TDP43 biology given that it demonstrated, for the first time, that the RRM2 is pivotal in the protein-RNA interactions *in vivo*.

Strikingly, the M323K mutation leads to the opposite effect, an "augmented alternative exon selection", which is observed at the level of specific genes, including *Sort1* and *Dnajc5* (figure 4.2) and to the majority of genes which are significantly alternatively spliced by the two mutant proteins when compared to wild type (figure 4.6). The mechanism leading to the "augmented alternative exon selection" caused by the M323K mutation has, however, proven more elusive.

Given the opposite effects observed in genome wide splicing by the F210I and M323K mutations, a simple, yet plausible hypothesis, could be formulated:

"*TDP43-M323K leads to augmented alternative exon selection as a consequence of increased affinity for RNA thus being more frequently in contact with target transcripts and performing its normal function of alternative exon splicing more frequently*"

The experimental results used to test this hypothesis are discussed below.

6.2.1 Mutant M323K protein and RNA interactions

The affinity of mutant TDP43-M323K protein for RNA was investigated through qualitative and semi-quantitative EMSA and no significant differences were found in its affinity for RNA when compared to wild type TDP43 protein (figures 4.12 and 4.13).

The EMSA assays were performed using synthetic UG_6 oligonucleotides, a sequence for which wild type TDP43 has high affinity (Buratti & Baralle 2001). Consequently, the EMSA assays as performed will have higher sensitivity for detecting a decrease in affinity between the TDP proteins and UG_6 RNA than for detecting an increased affinity and therefore a different assay design may have to be utilised to enable detection of increased affinity of mutant M323K protein for RNA.

Recent *in vivo* studies (Polymenidou et al. 2011; Tollervey et al. 2011) have demonstrated that TDP43 binds a wide range of sequences with different affinities and the interaction does not require the UG_6 sequence as a minimal binding sequence, as initially proposed (Buratti & Baralle 2001). Two additional experiments could therefore be employed in the future to investigate the affinity of the mutant M323K protein for RNA and overcome the limitation of the EMSA using UG_6 oligonucleotides as the binding target.

One would consist of the repetition of the semi-quantitative EMSA using known TDP43 binding targets with loosely conserved UG/GU nucleotides interspersed within the sequence. TDP43 wild-type protein has been shown to have lower affinity for these sequences, when compared to UG_6 RNA (Bhardwaj et al. 2013), and therefore the assay may be more sensitive in detecting an increase in protein-RNA affinity.

An alternative experiment would consist of performing RNA immunoprecipitation with TDP43-specific antibodies from embryonic heads, MEFs or adult brain harvested from all M323K genotypes followed by real time PCR of transcripts known to be more frequently alternatively spliced by the M323K protein (e.g. *Sort1*, *Dnajc5*) thus analysing target RNA abundance after immunoprecitation. Using this methodology would enable to assess, *in vivo*, whether TDP43-M323K was more frequently bound to the target RNAs in which an "augmented alternative exon selection" is observed. Experimentally, if target RNAs were found to be significantly more abundant in immunoprecipitates from $Tardbp^{M323K/M323K}$, compared to immunoprecipitates from $Tardbp^{+/+}$, it would be suggestive that the affinity of TDP43-M323K for specific target RNA is increased *in vivo*.

In addition to binding RNA, TDP43 is known to interact with other proteins. Moreover, the C-terminal of TDP43 has been shown to play an essential role in its interactions with other proteins (Buratti et al. 2005; D'Ambrogio et al. 2009) and TDP43 has also been shown to homodimerise (Kuo et al. 2009; Zhang et al. 2013). Integrity of the C-terminal domain is

essential for splicing (Wang et al. 2004) and increasing homodimerisation is positively correlated with TDP43's splicing activity (Zhang et al. 2013). Therefore, particularly as no differences were found between the affinity of wild type TDP43 or mutant M323K protein for RNA, an alternative hypothesis for the underlying biological mechanism leading to a "augmented alternative exon selection" observed in the mutant M323K protein can be formulated:

"*TDP43-M323K leads to augmented alternative exon selection as a consequence of changes in the nature of the interactions with other proteins or itself, including its homodimerisation properties, thus becoming more efficient in its splicing activity*"

6.2.2 Mutant M323K protein and interactions with other proteins and itself

The known interactions of TDP43 with other proteins, in which the C-terminal of the protein plays a pivotal role (Buratti et al. 2005; D'Ambrogio et al. 2009), were investigated qualitatively using the GST overlay/far western methodology.

The homodimerisation properties of TDP43-M323K were assessed through covalent crosslinking protein in cultured MEFs, followed by SDS gel electrophoresis and immunoblotting.

As described in chapter 4, the experiments outlined above did not reveal any differences between the protein interactions of wild type and mutant M323K protein (figure 4.14) or in the homodimerisation properties of the wild type and TDP43-M323K (figure 4.15).

The GST overlay/ far western methodology is limited to providing qualitative information concerning the interactions between proteins (refer to section 4.3.6 in chapter 4). Therefore, despite the lack of detectable differences in the subset of mammalian proteins with which wild type and mutant M323K proteins interact with, inferences cannot be made on the affinity of the interactions between wild type and mutant M323K protein and the subset of

mammalian proteins it interacts. Thus, the possibility that mutant TDP43 has different affinity for specific protein interactors, specifically hnRNPA2/B1, remains a hypothetical mechanism which could lead to the "augmented alternative exon selection" phenotype.

The affinity of the interactions between TDP43 and other proteins can be studied by co-immunoprecipitation (Co-IP) with a TDP43-specific antibody followed by semi-quantitative immunoblotting with antibodies against relevant proteins with hnRNPA2/B1 constituting a particular interesting candidate protein as it interacts with the C-terminal region affected by the M323K mutation and influences the splicing activity of TDP43 (Buratti et al. 2005). Therefore, future work will include Co-IP from *Tardbp*^{M323K} MEFs, embryo heads and adult brains followed by hnRNPA2/B1 immunoblotting.

The homodimerisation experiment could also be executed by performing native polyacrylamide gel electrophoresis (PAGE) as an alternative to chemical cross-linking and SDS-PAGE, which may be expected to confirm the results of the cross-linking experiment given that this methodology has proven to be reliable in crosslinking protein complexes, including TDP43 homodimers (Bartels et al. 2011; Lo et al. 2006; Zhang et al. 2013).

Another experimental strategy to study the homodimerisation properties of TDP43 consists of expressing exogenous wild type TDP43 and mutant TDP43 in cell systems followed by either chemical cross-linking and SDS-PAGE or native PAGE. While over-expression of exogenous constructs may be more sensitive for detecting any existing differences in homodimerisation due to higher protein levels, the results may not be physiologically relevant given that alteration of TDP43 protein levels could lead to non-physiological changes in homodimerisation properties itself.

6.2.3 Summary of the investigations into the biological mechanism leading to "augmented alternative exon selection" in the M323K protein and future work

The biological mechanism leading to "augmented alternative exon selection" in mutant TDP43-M323K remains unknown.

Within the limitations of the methodologies used, no significant differences in the affinity of the M323K mutant protein for RNA, its interactions with other proteins or homodimerisation properties have been seen in comparison to wild type TDP43.

In view of the limitations of the methodologies used, future work includes:

- Semi-quantitative EMSA with oligonucleotide sequences known to bind to TDP43 but without the canonical UG_6 repeat sequences to reassess the affinity between the mutant protein and RNA with increased sensitivity.
- RNA-IP (with a TDP43-specific antibody) followed by real time PCR of target transcripts to assess whether higher amounts of mRNA are bound to mutant TDP43.
- Co-IP (with a TDP43-specific antibody) followed by immunoblotting with antibodies against known interactors (e.g. hnRNPA1/B2) in order to assess the affinity of the interactions semi-quantitatively.

6.3 Effects of the M323K mutation on TDP43 protein levels and RNA levels; implications for the autoregulation process

TDP43 protein levels have been shown to be equivalent in embryo heads, MEFs and adult brains across *Tardbp*^{M323K} genotypes (figure 4.8), which could suggest that autoregulation is preserved in the presence of the M323K mutation.

In addition to protein levels, total RNA levels were also found to be comparable across *Tardbp*^{M323K} genotypes (figure 4.11). However, a dose dependent effect was found in the usage of an alternative polyadenylation site Poly(A)4, accompanied by a reduction in the usage of the canonical polyadenylation site Poly(A)1, which reached significance in MEFs harvested from *Tardbp*$^{M323K/M323K}$ embryos (figure 4.11).

Two models of TDP43 autoregulation have been proposed, both involving binding of TDP43 to the 3' UTR of its own transcript and splicing of the cryptic intron 7 (figure 1.9).

The model proposed by the Cleveland research group suggests that alternative splicing of the 3'UTR of its own transcript (i.e. the cryptic intron 7) deposits exon-junction complexes 3' of the stop codon thus marking the RNA for non-sense mediated decay (Polymenidou et al. 2011).

The model proposed by the research group of Emanuele Buratti and Francisco Baralle suggests that excess TDP43 protein binds the 3' UTR of its own transcript and promotes splicing of the cryptic intron 7 from the TDP43 mRNA, thus removing the canonical polyadenylation site Poly(A)1. According to the "Buratti and Baralle model", polyadenylation would then occur at the alternative Poly(A)2 site, which would disrupt cytoplasmic transport, leading to nuclear retention of the transcripts, thus disabling its translation (Ayala et al. 2011; Bembich et al. 2013).

The literature is sparse regarding Poly(A)4 usage in *Tardbp* transcripts, however polyadenylation at this alternative site is suggested to lead to mature *Tardbp* mRNAs which are normally exported and translated (Avendaño-Vázquez et al. 2012). Nevertheless, the increased usage of Poly(A)4, when interpreted together with the "augmented alternative exon selection" observed in the TDP43 protein, suggests that increased splicing of intron 7 from the *Tardbp* transcript by the M323K protein could be responsible for the change observed in

the polyadenylated site usage. It can therefore be hypothesised that the biological mechanism leading to the change in polyadenylation site usage may be the same as the one leading to the "augmented alternative exon selection".

However, since neither TDP43 protein levels nor total *Tardbp* RNA levels are reduced in *Tardbp*$^{M323K/M323K}$ when compared to wild type, but levels of canonically polyadenylated *Tardbp* at Poly(A)1 are reduced in *Tardbp*$^{M323K/M323K}$ and levels of alternatively polyadenylated *Tardbp* at Poly(A)4 are increased, the relationship between this alternative polyadenylation event and autoregulation is not evident. If alternative polyadenylation at the Poly(A)4 site led to nuclear retention of the *Tardbp* transcript, or, alternatively, non-sense mediated decay, reduced protein levels would be expected in *Tardbp*$^{M323K/M323K}$ MEFs when compared to *Tardbp*$^{+/+}$. Given that protein levels are not significantly different across genotypes, the relationship, if any, between usage of the alternative Poly(A)4 site and autoregulation remains obscure.

Furthermore, despite protein levels in tissue from *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$ being comparable to wild type, it has not yet been demonstrated whether this is due to a functional autoregulation process or any other event. Wild type mouse TDP43 has been shown to autoregulate its protein levels in the presence of the non-functional mutant Q101X allele, given that tissue from *Tardbp*$^{Q101X/+}$ animals, including brain and spinal cord, have comparable TDP43 levels to wild type controls (Ricketts et al. 2014). As *Tardbp*$^{M323K/Q101X}$ animals are viable (table 3.5), protein and RNA extraction from its tissues would enable the assessment of whether the mutant M323K protein retains its ability to autoregulate its protein levels when it is forced to do so *in vivo* by the presence of the null Q101X allele.

Further assessment of the polyadenylation site usage in *Tardbp* transcripts, including Poly(A)4 but also the canonical Poly(A)1 and Poly(A)2 can be achieved by RT-PCR with an oligoDT primer at the 3' end and a primer upstream from intron 7 in the *Tardbp* transcript,

which would confirm the findings of the Northern blot and, given that it involves amplification steps, may identify alternatively polyadenylated isoforms that are below the threshold of sensitivity of the Northern Blot.

6.3.1 Summary of implications of the experimental results of mutant M323K protein levels and RNA studies for the TDP43 autoregulation process and future work

Protein and total RNA levels are equivalent in tissue of all three *Tardbp*^{M323K} genotypes (*Tardbp*$^{+/+}$, *Tardbp*$^{M323K/+}$ and *Tardbp*$^{M323K/M323K}$).

Nevertheless, it is unclear whether the similar protein levels seen across *Tardbp*^{M323K} genotypes result from the integrity of the autoregulation process in the mutant protein or are a consequence of a different biological process. Despite an increase in the usage of the alternative Poly(A)4 site being observed, its biological relevance in the context of TDP43 autoregulation remains undetermined.

Future work includes:

- Assessment of the TDP43 protein levels in tissue from *Tardbp*$^{Q101X/M323K}$ animals in order to determine whether the mutant M323K protein retains the autoregulation properties inherent to wild type TDP43.
- Further assessment of alternative polyadenylation site usage promoted by the M323K mutant protein by RT-PCR and integration of results with the additional data from the Northern Blot and RNA-seq in order to determine the possible contribution of alternative Poly(A) site usage to TDP43 autoregulation.

6.4 *Tardbp*^{M323K} molecular characterisation

The "augmented alternative exon selection" and opposite effects in splicing, at the genome-wide level of the M323K and F210I mutations have already been discussed above, as were the results concerning protein and RNA levels and the implications for autoregulation.

Making inferences on the consequences of TDP43 mutations in TDP43 biological dysfunction and its association with the pathophysiology of ALS and FTLD-U, and possibly other neurodegenerative processes, including Alzheimer's Disease, was one of the of the aims of the research presented in this thesis. The molecular characterisation therefore included the assessment of whether histopathological hallmarks of TDP43 proteinopathies were recapitulated by the M323K mutation.

TDP43 subcellular distribution was not altered in MEFs harvested from embryos of all *Tardbp*^{M323K} genotypes when investigated by immunoblotting (figure 4.9) and immunofluorescence (figure 4.10). Additionally, inclusions were not seen in MEFs on immunofluorescence, which despite not being an expected finding, if present would suggest that mutant M323K protein would be prone to aggregation. As part of the characterisation into histopathological hallmarks of TDP43 proteinopathies, additional experiments should be performed to assess the aggregation properties of the M323K protein.

TDP43 aggregation can be assessed by extraction of the soluble and insoluble protein fractions followed by immunoblotting with a TDP43-specific antibody and quantification of the TDP43 bands in the blot. This methodology has been previously used in our group for the characterisation of other TDP43 ENU mutants and found to be reliable (Ricketts et al. 2014) and will be used in parallel with TDP43 immunohistochemistry in brain and spinal cord tissue.

Additional important experiments for future work include the validation of splicing changes, RNA levels and polyadenylation of the *Tardbp* transcript that has been performed using MEFs and embryonic heads in adult brain and spinal cord tissue of $Tardbp^{M323K/M323K}$ animals. The validation of the molecular characterisation in adult tissue from these animals was unfortunately not possible in time for inclusion in this thesis since homozygous M323K mutants were identified to be viable at an adult age only at a late phase of the DPhil research project. The protein levels in adult brains remain, however, comparable to wild type controls (figure 4.8), and the remaining molecular characterisation will be performed in tissue from adult animals in future experiments, including RNA-seq and full histopathological assessment for TDP43 aggregation and inclusions in the brain and spinal cord.

6.4.1 Summary of $Tardbp^{M323K}$ molecular characterisation and future work

The results from the molecular characterisation of $Tardbp^{M323K}$ mutation in regards to splicing changes, RNA and protein levels and the implications for autoregulation have been discussed in the previous sections. In this section, attention was given to experiments that have not been performed and would be informative for the molecular characterisation of this mutation.

Future work includes:

- Assessment of the effects of the M323K mutation in the solubility of TDP43 by performing immunoblotting with fractionated soluble and insoluble protein extracted from M323K tissue. The results would enable correlations to be made with the histopathological hallmarks of TDP43 proteinopathies.
- Extensive validation of the results from the molecular characterisation in adult CNS tissue of $Tardbp^{M323K}$ mice, given that $Tardbp^{M323K/M323K}$ are viable in a C57BL/6J-DBA background, thus ensuring the results represent a biological activity of the mutant protein which is conserved across tissues and genetic backgrounds.

- Full histopathological assessment of adult brains and spinal cords of *Tardbp*^{M323K} mutants to assess for TDP43 mislocalisation, aggregation or inclusion body formation, which constitute the histopathological hallmarks of ALS and FTLD-TDP. With the identification of adult *Tardbp*$^{M323K/M323K}$ homozygous animals in the C57BL/6J-DBA hybrid background, this experiment was postponed so that all genotypes could be assessed and not only *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$.

6.5 *Tardbp*^{F210I} leads to a unique effect in the expression of long intron genes: implications for TDP43 biology and neurodegeneration

In the introduction the advantages of investigating TDP43 biology through the characterisation of ENU mouse mutants were reviewed. Succinctly, taking into consideration that TDP43 levels are controlled within a narrow physiological range, as discussed above, ENU mutants constitute tools to study the effects of TDP43 mutations at isogenic levels of expression, which is not possible using other transgenic models currently available.

The partial recapitulation of the splicing effects caused by TDP43 knockdown in animals with the F210I mutation and the opposite pattern seen with the M323K mutation has been discussed in section 4.2. Remarkably, with the caveat that validation by real time PCR is pending, an opposite effect between TDP43 knockdown and the F210I mutation is seen in the regulation of the expression of long intron genes, which are downregulated in the former and upregulated in the latter (figure 4.7).

Genes with long introns whose expression is affected by TDP43 levels are predominantly involved in synaptic transmission (Polymenidou et al. 2011). Herein, lays a possible link between TDP43 biology and synaptic dysfunction. The opposite effects caused in the expression of long intron genes by the F210I mutation and TDP43 knockdown, which

contrast with the similar effects in splicing, constitute important “proof of principle” for the fact that TDP43 mutations can have multiple effects that cannot be modelled by changing protein levels. The M323K mutation does not have an effect on the expression of long intron genes (figure 4.7), at least in embryonic heads. It remains to be determined whether it has an effect in the expression of these genes in the adult brain.

Based on TDP43 knockdown, it has been proposed that TDP43 maintains the levels of expression from genes with long introns by binding its transcripts (Polymenidou et al. 2011). However, the upregulation of the expression of genes with long introns by the F210I mutant protein, which has reduced affinity for RNA, defy that model, suggesting that the regulation of the expression in this subset of genes is occurring through a mechanism independent of direct RNA binding.

Alternatively, and perhaps a simple but elegant hypothesis, is that due to its decreased RNA affinity, the F210I protein is more frequently unbound to RNA, therefore leading to increased levels of RNA-unbound mutant F210I protein when compared to wild type TDP43 (which could not be detected in immunoblotting). Higher levels of RNA-unbound TDP43-F210I would be available to participate in the interactions required to upregulate expression of long intron genes through a mechanism which is independent from direct RNA binding, but required TDP43 to be unbound to RNA in order to be involved in those specific interactions (e.g. binding a different protein complex which is itself in contact with RNA or DNA thus causing histone modification or promoting translation).

The alternative hypothesis described above, could explain the opposite patterns seen in TDP43 knockdown and the F210I mutation in the expression of genes with long introns and similar effects seen in alternative exon splicing. However, as described in chapter 4, the validation of the effects of the F210I mutation in the expression of genes with long introns by

real time PCR is currently in planning and confirmation of the experimental results is essential before the suggestions made in this discussion can be considered relevant.

The multiple effects in TDP43 biological functions that result from a single mutation have further implications for the dissection of the causative mechanism of neurodegeneration associated with TDP43 dysfunction.

In the event that *Tardbp*$^{M323K/M323K}$ animals develop phenotypes, establishing a cohort of *Tardbp*$^{F210I/M323K}$, which have been shown to be viable at weaning age (table 3.3), would constitute a possible form of further studying the link between TDP43 dysfunction and neurodegeneration.

Molecular characterisation of compound heterozygous (*Tardbp*$^{F210I/M323K}$) mice would determine whether the splicing phenotypes are rescued by the opposite effects caused by the two mutant proteins in TDP43-dependent splicing, even in the absence of a wild type *Tardbp*. Additionally, it would be possible to study the consequences of the absence of a wild type allele at the whole animal level through longitudinal *in vivo* phenotyping and correlate it with the molecular phenotype. Thus the unpredictable results from characterising *Tardbp*$^{F210I/M323K}$ mice could contribute towards further understanding of the functional consequences of TDP43 mutations.

6.5.1 Summary of the effect of the F210I mutation in the expression of genes with long introns and future work

The F210I mutation leads to upregulation of the expression of genes with long introns in embryonic heads. The effect on the expression levels of the genes with long introns is opposite to the one observed by TDP43 knockdown in the striatum of adult mouse.

Future work includes:

- Validation of the changes seen in RNA-seq by real time PCR in both embryonic head tissue and adult brain tissue and spinal cord, to confirm that the effect is in fact biologically significant.
- Breeding *Tardbp*$^{M323K/F210I}$ mice and molecular characterisation of splicing changes and the effect on long intron gene expression in CNS tissue from these mice. It is possible that the splicing phenotypes are rescued but that the effects of the F210I protein in the expression of the long intron genes are not, therefore providing a valuable opportunity to try and segregate these two molecular phenotypes and assess its effects at the whole-animal level.

The results from the *in vivo* characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K} will be discussed in the next section and integrated with the molecular phenotypes.

6.6 *In vivo* characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K}

Comprehensive *in vivo* characterisation of heterozygous *Tardbp*^{F210I} and *Tardbp*^{M323K} mouse mutants, on the C57BL/6J background has been performed. Cohorts of male and female heterozygous animals and its wild type controls were tested separately (to account for any sexual dimorphism in the resulting phenotypes) given that adult homozygous were not viable for any of the two mutations on this genetic background. Since it was impossible to know, *a priori*, whether any phenotypes would develop in the mutant animals, a comprehensive phenotyping pipeline was developed, which included tests aimed to identify deficits in motor function, learning and memory and endophenotypes of neuropsychiatric disease (Figure 5.1).

6.6.1 In vivo Characterisation of Tardbp$^{F210I/+}$

Tardbp$^{F210I/+}$ females were found to have reduced body weight throughout most of its lives when compared to controls and reduced fat mass, at least at 31 weeks of age. The difference in body composition found between the heterozygous and wild type females can be the result

of a general effect of the F210I mutation due to changes in multiple cellular processes or, alternatively, an effect in a specific process.

In mice, TDP43 conditional deletion has been shown to downregulate *Tbc1d1* in skeletal muscle of adult animals (Chiang et al. 2010) and TDP43 overexpression leads to upregulation of *Tbc1d1*, increased fat deposition and impaired insulin mediated glucose uptake (Stallings et al. 2013). Given that the F210I mutation lead to decreased affinity of TDP43 for RNA and, as a consequence to a "loss of normal splicing function", it is possible that the decreased body weight and fat mass of *Tardbp*$^{F210I/+}$ results from *Tbc1d1* downregulation. The fact that the body weight phenotype is only seen in females could be explained by sexual dimorphism playing a role in the downregulation of *Tbc1d1* (e.g. due to the different steroid hormone environment in males and females), an effect that has been observed in the study of metabolism-related genes in other genetic mouse lines (Garg et al. 2011). Moreover, since the wild type *Tardbp* allele is still present in these animals, any effect on body weight may not be as pronounced as in a conditional deletion model. The expression levels of *Tbc1d1* will therefore be investigated in skeletal muscle of *Tardbp*^{F210I} mice by real time PCR.

In addition to changes in body weight, *Tardbp*$^{F210I/+}$ animals also made more errors with their hind limbs and cumulative errors with forelimbs and hind limbs when compared to controls in the horizontal ladder/locotronic test. This test has been used to assess motor coordination in transgenic mice and mice exposed to neurotoxic chemicals as well as to test the effect of therapeutic agents in those mice (Rousselet et al. 2003; Chort et al. 2013), and increased number of errors made by rodents have been shown to result from damage to the corticospinal tracts (Carmel & Martin 2014) and peripheral nerves (Khuong et al. 2014).

The number of hind limb and hind limb and forelimb combined errors made by *Tardbp*$^{F210I/+}$ animals (males and females analysed together) are significantly higher than the errors made by wild type controls (figure 5.20 and table 5.14), thus suggesting that axonal dysfunction in

peripheral nerves or centrally, at the level of the cortiscopinal tracts or motor cortex/cerebellar system, could be present. Nevertheless caution is warranted in attributing biological relevance to the locotronic deficit in isolation, given that no significant deficits were identified in the other motor tests performed (e.g. open field, grip strength, rotarod, home cage-automated wheel running).

6.6.1.1 Summary of in vivo phenotyping of *Tardbp^{F210I}* mutants and future work

The *in vivo* characterisation of *Tardbp$^{F210I/+}$* cohorts, on the C57BL/6J background resulted in the observation of a body weight and possibly motor coordination phenotypes. Future work should aim to identify the underlying mechanism for the weight phenotype and confirm the motor phenotype suggested by the locotronic test and its anatomical and pathological correlation.

Future work includes:

- Further testing of TDP-F210I-B6 cohorts in the locotronic test at different ages
- Histopathological assessment of the CNS in the mutant mouse line, including TDP43 immunoreactivity and structural assessment of the motor cortex, cerebellum and spinal cord, particularly if the motor deficit is confirmed (n.b. the locotronic phenotype was identified in older animals at 19 months of age). At one year of age TDP43 immunohistochemistry was performed in the brain and spinal cord of *Tardbp$^{F210I/+}$* mice (and controls) by the group of Professor Sebastian Brandner and no abnormalities were identified.

6.6.2 In vivo Characterisation of Tardbp^{M323K}

In the cohorts of *Tardbp$^{M323K/+}$* mutants, in the C57BL/6J background, no significant phenotypes were identified with the behavioural tests and measurements that were performed. However, despite the time constraints, it was possible to perform limited testing in animals on

a C57BL/6J-DBA hybrid background which enabled homozygous mice (*Tardbp*$^{M323K/M323K}$) to be tested, which proved fruitful in identifying phenotypes in the latter.

Tardbp$^{M323K/M323K}$ weighed significantly less than wild type controls and *Tardbp*$^{M323K/+}$ (figure 5.6 and table 5.5).

Moreover, a fear conditioning phenotype was also identified; *Tardbp*$^{M323K/M323K}$ had increased freezing time in response to the auditory cue when compared to *Tardbp*$^{+/+}$ (figure 5.31 and table 5.20).

Classically, when a fear conditioning paradigm is used as a tool for assessing learning and memory, deficits manifest as a reduction in the percentage of time freezing in context and/or cue. At the time of testing, the experimental hypothesis dictated that if the M323K mutation affected learning and memory, a reduction in freezing times, in context, cue, or both, would be observed in *Tardbp*$^{M323K/M323K}$ animals and possibly *Tardbp*$^{M323K/+}$. Notwithstanding, the opposite result was observed; the M323K mutation revealed a dose dependent effect with increased freezing seen in association with the cue, which reached statistical significance in *Tardbp*$^{M323K/M323K}$ (figure 5.31 and table 5.20).

Interestingly, a similar phenotype has been seen associated with amygdala dysfunction in mouse models of Alzheimer's disease. The APP transgenic models (carrying mutant Amyloid Precursor Protein gene) of AD and the triple transgenic AD model (carrying mutant APP, Presinilin and Tau genes) exhibit increased freezing on contextual fear conditioning, which partly coincides with the histological evidence of intracellular accumulation of β-amyloid peptide in the amygdala (España et al. 2010). This previous study suggests that the increased freezing phenotype, termed enhanced neophobia, observed in *Tardbp*$^{M323K/M323K}$ could be associated with amygdala dysfunction, with the caveat of being seen in cued fear conditioning but not in context.

Additionally, pre-pulse inhibition of the acoustic startle response, when a pulse of white noise of 70 or 75 decibels precedes the acoustic stimulus, is significantly supressed in *Tardbp*$^{M323K/M323K}$ (figure 5.36, table 5.25). This finding suggests that synaptic dysfunction is present in *Tardbp*$^{M323K/M323K}$ and leads to sensorimotor gating deficits analogous to those which are commonly seen in schizophrenia (Braff et al. 1992; Swerdlow et al. 2006; Kumari et al. 2014).

TDP43 knockdown has been shown to downregulate the expression of long intron genes involved in synaptic transmission (Polymenidou et al. 2011) and the F210I mutation to upregulate the same genes as already discussed. At the genome-wide level no overall effect was seen in the expression of long intron genes when RNA-seq was performed in embryo heads of *Tardbp*$^{M323K/M323K}$ (figure 4.7), however this has not yet been tested in adult brain tissue. It is therefore reasonable to hypothesise that the PPI deficit seen in the adult *Tardbp*$^{M323K/M323K}$ animals may be associated with abnormal expression or splicing of specific long intron genes involved in the synaptic activity, which requires further investigation. Furthermore, the PPI phenotype found by testing animals in the TDP-M323K-B6-DBA with ages between 32 and 49 weeks (table 5.26) raises the possibility that *Tardbp*$^{M323K/M323K}$ may develop working memory deficits, since a correlation between lower PPI scores and reduced scores in working memory tests has been described in inbred C57BL/6J mice (Singer et al. 2013).

Preliminary data for the characterisation the M323K mutant animals in the C57BL/6J-DBA hybrid background, specifically the results from fear conditioning testing and PPI of the acoustic startle response, is therefore supportive that this mutation leads to TDP43-mediated synaptic dysfunction, resulting in identifiable behavioural deficits in *Tardbp*$^{M323K/M323K}$. Moreover, the failure to identify significant phenotypes in *Tardbp*$^{M323K/+}$ mice (C57BL/6J background) suggests that the information that can be obtained from *Tardbp* ENU mutants is

limited if adult homozygous animals cannot be characterised; a limitation which could not be overcome in the case of *Tardbp*^{F210I}, given that *Tardbp*$^{F210I/F210I}$ die before weaning age even in a hybrid C57BL/6J-DBA background (table 3.4).

Given the phenotypes identified in *Tardbp*$^{M323K/M323K}$ so far, a full characterisation of *in vivo* phenotypes, full molecular adult CNS tissue characterisation as well as histopathology of adult CNS tissue from TDP-M323K-B6-DBA cohorts needs to be completed.

6.6.2.1 Summary of *in vivo* phenotyping of Tardbp^{M323K} *mutants and future work*

Applying rigorous statistical analysis, no significant phenotypes were identified in *Tardbp*$^{M323K/+}$ of either the C57BL/6J background or hybrid C57BL/6J-DBA hybrid background.

In a hybrid C57BL/6J-DBA background, which leads to viable adult homozygous animals, statistically significant body weight phenotypes and a classic PPI deficit associated with neuronal dysfunction in schizophrenia was seen in the homozygous animals. Additionally, on cued fear conditioning, *Tardbp*$^{M323K/M323K}$ do not display the classical phenotype associated with learning and memory deficits but do have a phenotype previously described in the literature as enhanced neophobia, suggested to be associated with amygdala dysfunction (España et al. 2010). The preliminary data thus suggests that the M323K mutation leads to neuronal dysfunction above the threshold necessary to detect behavioural deficits.

Future work includes:

- Comprehensive molecular and *in vivo* characterisation *Tardbp*^{M323KM323K} animals (together with littermate *Tardbp*$^{+/+}$ and *Tardbp*$^{M323K/+}$ mice) on the C57BL/6J-DBA background, which will enable the detection of any motor deficits as well as deficits in learning and memory, endophenotypes of psychiatric disease and effects on general health.

- Molecular characterisation of adult brain, employing RNA-seq, which will enable the assessment of genome wide splicing changes and gene expression, including genes associated with synaptic activity which may be affected and therefore lead to the phenotypes in PPI and fear conditioning.
- Full histopathological analysis to assess whether the animals develop pathological hallmarks of ALS and FTLD-TDP.

6.7 From Mouse to Man: implications of the characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K} on the role of TDP43 in neurodegeneration

One of the aims from the work presented in this thesis was to further contribute to the understanding on how TDP43 biological dysfunction is involved in the pathophysiology of neurodegenerative diseases, and particularly ALS and FTLD-TDP.

An increasing body of published evidence supports the conclusion that TDP43 has a pivotal role as a splicing factor, particularly in alternative exon splicing, which constitutes a very common process in the mammalian transcriptome, being estimated that around 95% of multiexon genes in the human transcriptome are alternatively spliced (Pan et al. 2008). Furthermore, alternative exon splicing, leading to multiple transcript and protein isoforms coded by the same gene, exponentially increases the diversity and complexity of the mammalian transcriptome. The fundamental role of alternative exon splicing in the complexity of living organisms can be illustrated by the contrast between the genomes and transcriptomes of the Human and Chimpanzee species. The human genome is estimated to have 98% homology in its coding genes to the chimpanzee (Mikkelsen et al. 2005). Therefore, alternative splicing exponentially increases the diversity between the two species'

transcriptional landscape thus contributing to the degree of complexity and difference observed across the two species.

Studies using nervous tissue from human patients who suffered from ALS and FTLD-U harvested post-mortem have shown that TDP43-dependent splicing is altered when compared to nervous human tissue harvested from individuals that did not suffer from these diseases (Tollervey et al. 2011; Yang et al. 2014; Highley et al. 2014). Additionally, changes in the splicing of target genes and differential expression of long intron genes have also been shown in muscle cells of patients with sporadic inclusion body myosistis with TDP43 histopathological changes (Cortese et al. 2014). Taken together, the observations form the cited studies are highly suggestive that widespread RNA dysfunction in TDP43-depedent processes is inherent to TDP43 proteinopathies. However, a general conclusion regarding whether the RNA changes observed in nervous tissue from patients constitute a gain or loss of splicing function cannot be made as the splicing of specific target genes are altered in one direction and others on the opposite direction (e.g. *RDWW1* splicing changes towards the direction observed in "augmented alternative exon selection" and *POLDIP3* towards loss of function). Furthermore, the studies describing widespread RNA dysregulation in human CNS tissue (Tollervey et al. 2011; Yang et al. 2014; Highley et al. 2014) cannot provide an answer on whether dysfunction in gene expression and exon splicing are a direct cause of downstream neurodegeneration or effects which are associated with TDP43 dysfunction but not causative of neurodegeneration.

Abnormal splicing has also been shown to be inherent to TDP43 dysfunction in mice. Animals expressing mutant *TARDBP* transgenes have been shown to develop neurodegeneration and have marked splicing changes, with at least one model (carrying a human TDP43 transgene with the Q331K mutation) exhibiting predominantly a bias towards promoting exon exclusion in target transcripts thus suggesting an "augmented alternative

exon selection" (Arnold et al. 2013). TDP43 knockdown in transgenic mice with inducible antisense RNA specific to target *Tardbp* also develop neurodegeneration and exhibit splicing changes towards a reduction in exon exclusion thus suggesting a "loss of normal splicing function".

Therefore, in transgenic TDP43 mice splicing changes are also a prominent feature of neurodegeneration in the context of TDP43 dysfunction, in harmony to the data from studies using tissue from human patients. Thus, the question of whether a "loss of function", a "gain of function", including "augmented alternative exon selection", or a combination of the former and the latter, in TDP43-dependent splicing is necessary and/or sufficient to cause neurodegeneration remains unanswered both in Man and Mouse.

Through the characterisation of *Tardbp* ENU mouse mutants, relevant contributions towards a better understanding of the consequences of *Tardbp* mutations in TDP43-dependent splicing and its implications for neurodegeneration can be made.

In this thesis, it has been described that $Tardbp^{M323K}$, when expressed at isogenic levels, leads to widespread transcriptome changes due to a dose-dependent "augmented alternative exon selection" in the TDP43-M323K mutant protein, through a yet unidentified mechanism. In contrast, experimental work predating this thesis has shown that $Tardbp^{F210I}$ leads to a dose-dependent "loss of splicing function" due to reduced affinity between the TDP43-F210I protein and RNA (Ricketts 2012). Moreover, analysis of RNA-seq from these two mutants has shown that the changes in TDP43-dependent alternative exon splicing occur in opposite directions in $Tardbp^{F210I}$ and $Tardbp^{M323K}$ mutants at a genome-wide scale. Comprehensive *in vivo* characterisation of the two mutant lines ($Tardbp^{F210I}$ and $Tardbp^{M323K}$) failed to identify robust phenotypes associated with neurodegeneration, at least in heterozygous mutants ($Tardbp^{F210I/+}$ and $Tardbp^{M323K/+}$) on the C57BL/6J background.

Within the limitations of having only fully characterised heterozygous animals, the results from the molecular and *in vivo* characterisation of the F210I and M323K mutant mouse lines suggests that an "augmented alternative exon selection", a "loss of normal splicing function" or partial loss of RNA binding affinity by the TDP43 protein (in the case of *Tardbp*$^{F210I/+}$) is not sufficient to cause neurodegeneration in the mouse.

Nevertheless, it is possible that an additional insult is necessary to promote neurodegeneration in the context of mutation-associated TDP43 dysfunction, a model commonly referred to as "the two hit hypothesis" mostly based evidence from *ex vivo* cell studies (Nonaka et al. 2009; Pesiridis et al. 2011). In view of this hypothesis, the consequences of TDP43 dysfunction in the ENU mutants in the nervous system of mice submitted to additional insults could be modelled by performing axotomies and quantifying both motor recovery and the extent of motor neurone loss, given that TDP43 dysfunction in transgenic mice has been shown to impair recovery after axotomy (Swarup et al. 2012). In fact, axotomies in *Tardbp*$^{F210I/+}$ have already been performed in collaboration with Dr Bernadett Kalmar in the research group of Professor Linda Greensmith and the spinal cord samples are currently being prepared for histological analysis.

Links have started to be established between TDP43 dysfunction and splicing changes observed in both mouse models of TDP43 proteinopathies and in patients suffering from ALS and FTLD-TDP as a consequence of *TARDBP* mutations. Despite splicing changes being found in transgenic TDP43 mice which develop neurodegeneration and tissue from human patients, the data from *Tardbp*^{F210I} and *Tardbp*^{M323K} mutants suggests that TDP43 dysfunction leading to genome-wide "loss of normal splicing function" or "augmented alternative exon selection" is not sufficient to cause neurodegeneration in the mouse; at least during its normal lifespan. Nevertheless, translating between splicing changes seen in mouse models of TDP43

dysfunction and those seen in tissue from patients with ALS or FTLD-TDP needs to be done with thoughtful consideration in order to avoid precipitated, erroneous, conclusions.

Beyond the obvious limitation of the difference in complexity between the human and mouse genomes, with an estimated number of over twenty thousand coding genes for both species (Kim et al. 2014; Waterston et al. 2002) which lead to a wide diversity in the transcript and protein isoforms, additional obstacles need to be overcome, in order to allow a direct comparison of animal and human data.

These include the current difficulty in obtaining good quality RNA from post-mortem human tissue and the identification and annotation of species specific factors which also influence the outcome of TDP43-dependant splicing.

Extracting good quality RNA from post-mortem human tissue is a challenge given that as the patient passes away, the cellular molecules start to degrade; an effect which is minimised in experimental animals by instantly freezing the tissue of interest. This obstacle is now starting to be overcome by the use of fibroblasts harvested from living patients with TDP43 mutations to study changes in RNA metabolism (Raman et al. 2014; Highley et al. 2014).

The influence of species specific factors in TDP43 dependent splicing is well demonstrated with the case of *Sort1*. Sort1 is the receptor from Progranulin and an obvious link between TDP43 splicing dysfunction and neurodegeneration given that both mutations in TDP43 and Progranulin cause FTLD (Chiang et al. 2013). The M323K and F210I mutations in TDP43 both lead to splicing changes in mouse *Sort1* as described earlier in section 4.1.

However, a variation between the intronic sequences of human and mouse Sort1 determines that in humans the cassette exon of Sort1 is almost always excluded as opposed to the pattern seen in mouse Sort1 where the flanking intronic sequences act, in *Cis*, to promote inclusion of the cassette exon (Prudencio et al. 2012). Moreover, associated with the difference in

inclusion and exclusion of the cassette exon from Sort1, is the fact that the cassette exon carries a premature stop codon in the genomic sequence of humans and primates but not mice (Prudencio et al. 2012), suggesting that the evolution of the flanking intronic sequences is associated with this event in the exonic sequence. Consistent with the description above is the observation that changes in Sort1 splicing are not observed in human derived lymphoblastoid cells carrying disease-causative TDP43 mutations (Belzil et al. 2012).

Together with Pietro Fratta in the research group of Professor Elizabeth Fisher, fibroblasts from patients with ALS caused by TDP43 mutations have been ethically obtained. We plan to perform the CFTR minigene splicing assay in these fibroblasts, thus aiming to partially overcome the limitation of the species-specific factors which influence TDP43-dependant splicing and beginning to make direct comparisons between the splicing changes seen in *Tardbp* mutants and human cells with *TARDBP* mutations. With this experiment we hope to make inferences on whether a "gain of splicing function" or "loss of splicing function" is particularly associated with in TDP43 dysfunction and ALS caused by *TARDBP* mutations.

Limitations regarding our knowledge of the variations in the mouse and human genomes that influence TDP43-dependent splicing will be progressively overcome by further research within the frameworks of the human and mouse genome projects. Moreover, despite the evidence that in the mouse splicing changes associated with TDP43 dysfunction are not sufficient to cause neurodegeneration, such conclusion cannot be extrapolated to the case of TDP43 dysfunction and human disease. For all the advantages that modelling a complex human disease in the mouse offer, limitations also exist, including the size and complexity of the CNS of the mouse compared to humans and the reduced life span of the former, which may preclude clinical manifestations from underlying pathological dysfunction from occurring.

Therefore, it can be hypothesised that TDP43 dysfunction interacts with the changes in the transcriptome which occur naturally during the human life cycle, from early development through juvenile growth, middle age and senescence, and/or precipitated by external factors (e.g. environmental toxins, viral illnesses). Thus, by interacting with time and situation dependent transcriptome changes, TDP43 dysfunction could lead to pivotal alterations in splicing and/or gene expression levels which precipitate neuronal dysfunction and, ultimately, neurodegeneration (figure 6.1).

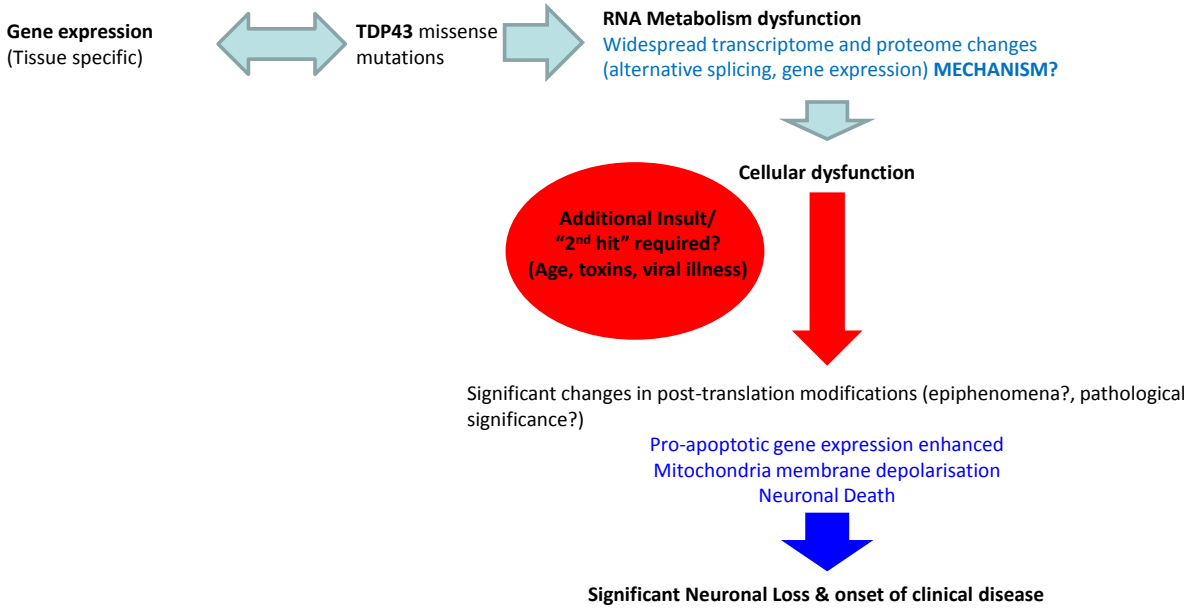

Figure 6.1: Hypothesis of the biological pathways leading to neurodegeneration caused by TDP43 dysfunction
The observation of widespread RNA metabolism changes caused by TDP43 dysfunction enables the postulation of hypotheses of TDP43's role in neurodegeneration. It can be hypothesised that changes in the transcriptome and proteome that occur during the human lifecycle interact with TDP43 dysfunction, particularly in the context of *TARDBP* mutations, leading to cellular dysfunction and downstream neurodegeneration. It is possible that either aging itself or additional environmental factors further influence changes in the transcriptome and proteome and precipitate the key events leading to neurodegeneration in genetically susceptible individuals. Better understanding of species specific factors influencing TDP43 dependent splicing and an increase in the quality of RNA extracted from human tissue are critical factors in further investigation this hypothesis.

Consequently, the knowledge from animal and cellular models of TDP43 dysfunction, allied with a better understanding of species specific genetic factors that influence TDP43's biological functions, and steady progress in the characterisation of the molecular pathophysiology of the neurodegenerative disease processes in humans, leave hope for

identification of specific pathways of disease which can be manipulated to change disease progression.

6.8 Concluding remarks

The work presented in this thesis consisted of the characterisation, both molecular and behavioural, of *Tardbp* ENU mutants. It has described how mutations in *Tardbp* lead to widespread transcriptome changes, with two specific mutations in different regions of TDP43 leading to opposite effects in its splicing activity.

It was aimed to correlate the molecular effects with deficits, at the whole animal level, in motor function, memory and learning or phenotypes of neuropsychiatric disease, which would point towards specific TDP43 biological dysfunction pathways leading to neurodegeneration. Whole animal phenotyping was limited to heterozygous animals due to early lethality of homozygous animals for both mutations and no firm phenotypes associated with neurodegeneration were identified. The results therefore suggest that splicing changes associated with TDP43 dysfunction are not sufficient to cause neurodegeneration in the mouse.

However, much remains to be gained regarding the understanding of TDP43 biology with the study of ENU mouse mutants and future work has been discussed in detail. Moreover, on a hybrid C57BL/6J-DBA background *Tardbp*$^{M323K/M323K}$ adult animals are viable. Further molecular characterisation of adult animals will be performed as will comprehensive *in vivo* phenotyping. In the event of *Tardbp*$^{M323K/M323K}$ developing neurodegeneration, the already established molecular phenotype will enable inferences with relevant implications for human disease to be made, particularly if future work also unravels the underlying biological mechanism leading to the “augmented alternative exon selection” in the mutant M323K protein.

The characterisation of *Tardbp*^{F210I} and *Tardbp*^{M323K} has generated limited answers but, more importantly, forces us to ask more questions about the role of TDP43 in ALS and FTLD-U and challenge the current paradigms and hypotheses in the quest for a better understanding of the biological mechanisms of neurodegeneration.

Chapter 7.

Bibliography

Abel, O. et al., 2013. Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Database (ALSoD). *JMIR mhealth and uhealth*, 1(2), p.e18.

Acevedo-Arozena, A. et al., 2008. ENU Mutagenesis, a Way Forward to Understand Gene Function. *Annual Review of Genomics and Human Genetics*, 9(1), pp.49–69.

Al-Chalabi, A. et al., 2012. The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta neuropathologica*, 124(3), pp.339–52.

Al-Chalabi, A. & Hardiman, O., 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nature reviews. Neurology*, 9(11), pp.617–28.

Andersen, P.M. & Al-Chalabi, A., 2011. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nature reviews. Neurology*, 7(11), pp.603–15.

Anderson, K.N. et al., 2009. Disrupted sleep and circadian patterns in frontotemporal dementia. *European journal of neurology : the official journal of the European Federation of Neurological Societies*, 16(3), pp.317–23.

Arnold, E.S. et al., 2013. 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), pp.E736–45.

Avendaño-Vázquez, S.E. et al., 2012. Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes & development*, 26(15), pp.1679–84.

Ayala, Y.M. et al., 2008. Structural determinants of the cellular localization and shuttling of TDP-43. *Journal of cell science*, 121(Pt 22), pp.3778–85.

Ayala, Y.M. et al., 2011. TDP-43 regulates its mRNA levels through a negative feedback loop. *The EMBO journal*, 30(2), pp.277–88.

Ayala, Y.M., Pagani, F. & Baralle, F.E., 2006. TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS letters*, 580(5), pp.1339–44.

Banks, G.T. & Nolan, P.M., 2011. Current Protocols in Mouse Biology J. Auwerx et al., eds. *Current Protocols in Mouse Biology*, 1, pp.369–381.

Bartels, T., Choi, J.G. & Selkoe, D.J., 2011. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature*, 477(7362), pp.107–110.

Bäumer, D. et al., 2010. Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations. *Neurology*, 75(7), pp.611–8.

Bei Hu et al., 2010. Off-label medication use in frontotemporal dementia. *American journal of Alzheimer's disease and other dementias*, 25(2), pp.128–33.

Belzil, V. V et al., 2012. Analysis of the SORT1 gene in familial amyotrophic lateral sclerosis. *Neurobiology of aging*, 33(8), pp.1845.e7–9.

Bembich, S. et al., 2013. Predominance of spliceosomal complex formation over polyadenylation site selection in TDP-43 autoregulation. *Nucleic acids research*, pp.1–10.

Bensimon, G. et al., 1994. A controlled trial of Riluzole in Amyotrophic Lateral Sclerosis. *The New England journal of medicine*, 330, pp.584–591.

Le Ber, I. et al., 2009. Chromosome 9p-linked families with frontotemporal dementia associated with motor neuron disease. *Neurology*, 72(19), pp.1669–76.

Bhardwaj, A. et al., 2013. Characterizing TDP-43 interaction with its RNA targets. *Nucleic acids research*, 41(9), pp.5062–74.

Bielas, J.H. & Heddle, J. a, 2000. Proliferation is necessary for both repair and mutation in transgenic mouse cells. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), pp.11391–6.

Bird, T. et al., 2003. Epidemiology and genetics of frontotemporal dementia/Pick's disease. *Annals of neurology*, 54 Suppl 5, pp.S29–31.

Boeve, B.F., 2007. Links between frontotemporal lobar degeneration, corticobasal degeneration, progressive supranuclear palsy, and amyotrophic lateral sclerosis. *Alzheimer disease and associated disorders*, 21(4), pp.S31–8.

Borghero, G. et al., 2011. A patient carrying a homozygous p.A382T TARDBP missense mutation shows a syndrome including ALS, extrapyramidal symptoms, and FTD. *Neurobiology of aging*, 32(12), pp.2327.e1–5.

Borroni, B. et al., 2009. Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. *Human mutation*, 30(11), pp.E974–83.

Bose, J.K. et al., 2008. TDP-43 overexpression enhances exon 7 inclusion during the survival of motor neuron pre-mRNA splicing. *The Journal of biological chemistry*, 283(43), pp.28852–9.

Bourke, S.C. et al., 2006. Effects of non-invasive ventilation on survival and quality of life in patients with amyotrophic lateral sclerosis: a randomised controlled trial. *Lancet neurology*, 5(2), pp.140–7.

Bouwknecht, J.A. et al., 2004. Effects of repeated testing in two inbred strains on flesinoxan dose-response curves in three mouse models for anxiety. *European journal of pharmacology*, 494(1), pp.35–44.

Bowser, R., Turner, M.R. & Shefner, J., 2011. Biomarkers in amyotrophic lateral sclerosis: opportunities and limitations. *Nature reviews. Neurology*, 7(11), pp.631–8.

Braff, D.L. et al., 2005. Female schizophrenia patients have prepulse inhibition deficits. *Biological psychiatry*, 57(7), pp.817–20.

Braff, D.L., Grillon, C. & Geyer, M. a, 1992. Gating and habituation of the startle reflex in schizophrenic patients. *Archives of general psychiatry*, 49(3), pp.206–15.

Brandmeir, N.J. et al., 2008. Severe subcortical TDP-43 pathology in sporadic frontotemporal lobar degeneration with motor neuron disease. *Acta neuropathologica*, 115(1), pp.123–31.

Buratti, E. et al., 2001. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *The EMBO journal*, 20(7), pp.1774–84.

Buratti, E. et al., 2010. Nuclear factor TDP-43 can affect selected microRNA levels. *The FEBS journal*, 277(10), pp.2268–81.

Buratti, E. et al., 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), pp.37572–84.

Buratti, E. & Baralle, F.E., 2001. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *The Journal of biological chemistry*, 276(39), pp.36337–43.

Burrell, J.R., Kiernan, M.C., et al., 2011. Motor neuron dysfunction in frontotemporal dementia. *Brain : a journal of neurology*, 134(Pt 9), pp.2582–94.

Burrell, J.R., Vucic, S. & Kiernan, M.C., 2011. Isolated bulbar phenotype of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 12(4), pp.283–9.

Byrne, S. et al., 2011. Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. *Journal of neurology, neurosurgery, and psychiatry*, 82(6), pp.623–7.

Caccamo, A., Majumder, S. & Oddo, S., 2012. Cognitive decline typical of frontotemporal lobar degeneration in transgenic mice expressing the 25-kDa C-terminal fragment of TDP-43. *The American journal of pathology*, 180(1), pp.293–302.

Cairns, N.J. et al., 2007. TDP-43 in familial and sporadic frontotemporal lobar degeneration with ubiquitin inclusions. *The American journal of pathology*, 171(1), pp.227–40.

Cannon, A. et al., 2012. Neuronal sensitivity to TDP-43 overexpression is dependent on timing of induction. *Acta neuropathologica*, 123(6), pp.807–23.

Carmel, J.B. & Martin, J.H., 2014. Motor cortex electrical stimulation augments sprouting of the corticospinal tract and promotes recovery of motor function. *Frontiers in integrative neuroscience*, 8(June), p.51.

Che, M.-X. et al., 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), pp.2344–53.

Chen-Plotkin, A.S., Lee, V.M.-Y. & Trojanowski, J.Q., 2010. TAR DNA-binding protein 43 in neurodegenerative disease. *Nature reviews. Neurology*, 6(4), pp.211–20.

Chiang, H. et al., 2012. Novel TARDBP mutations in Nordic ALS patients. *Journal of Human Genetics*, 57, pp.316–319.

Chiang, H.-H. et al., 2013. Novel progranulin mutations with reduced serum-progranulin levels in frontotemporal lobar degeneration. *European journal of human genetics : EJHG*, 21(11), pp.1260–5.

Chiang, P.-M. et al., 2010. Deletion of TDP-43 down-regulates Tbc1d1, a gene linked to obesity, and alters body fat metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 107(37), pp.16320–4.

Chiò, A. et al., 2010. Amyotrophic lateral sclerosis-frontotemporal lobar dementia in 3 families with p.Ala382Thr TARDBP mutations. *Archives of neurology*, 67(8), pp.1002–9.

Chiò, A., 1999. ISIS Survey: an international study on the diagnostic process and its implications in amyotrophic lateral sclerosis. *Journal of Neurology*, 246(S3), pp.III1–III5.

Chiò, A. et al., 2009. Prognostic factors in ALS: A critical review. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 10(5-6), pp.310–23.

Chort, A. et al., 2013. Interferon β induces clearance of mutant ataxin 7 and improves locomotion in SCA7 knock-in mice. *Brain : a journal of neurology*, 136(Pt 6), pp.1732–45.

Colombrita, C. et al., 2012. TDP-43 and FUS RNA-binding proteins bind distinct sets of cytoplasmic messenger RNAs and differently regulate their post-transcriptional fate in motoneuron-like cells. *The Journal of biological chemistry*, 287(19), pp.15635–47.

Cook, C. et al., 2014. Severe amygdala dysfunction in a MAPT transgenic mouse model of frontotemporal dementia. *Neurobiology of aging*, 35(7), pp.1769–77.

Corcia, P. et al., 2008. Causes of death in a post-mortem series of ALS patients. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 9(1), pp.59–62.

Corcia, P. et al., 2012. Phenotype and genotype analysis in amyotrophic lateral sclerosis with TARDBP gene mutations. *Neurology*, 78(19), pp.1519–26.

Corrado, L. et al., 2009. High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Human mutation*, 30(4), pp.688–94.

Cortese, A. et al., 2014. Widespread RNA metabolism impairment in sporadic inclusion body myositis TDP43-proteinopathy. *Neurobiology of aging*, 35(6), pp.1491–8.

Cudkowicz, M.E. et al., 2013. Dexpramipexole versus placebo for patients with amyotrophic lateral sclerosis (EMPOWER): a randomised, double-blind, phase 3 trial. *Lancet neurology*, 12(11), pp.1059–67.

D'Ambrogio, A. et al., 2009. Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic acids research*, 37(12), pp.4116–26.

Davidson, Y. et al., 2007. Ubiquitinated pathological lesions in frontotemporal lobar degeneration contain the TAR DNA-binding protein, TDP-43. *Acta neuropathologica*, 113(5), pp.521–33.

Van Deerlin, V.M. et al., 2008. TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet neurology*, 7(5), pp.409–16.

DeJesus-Hernandez, M. et al., 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*, 72(2), pp.245–56.

Dickson, D.W., Josephs, K. a & Amador-Ortiz, C., 2007. TDP-43 in differential diagnosis of motor neuron disorders. *Acta neuropathologica*, 114(1), pp.71–9.

Dreumont, N. et al., 2010. Antagonistic factors control the unproductive splicing of SC35 terminal intron. *Nucleic acids research*, 38(4), pp.1353–66.

Elamin, M. et al., 2011. Executive dysfunction is a negative prognostic indicator in patients with ALS without dementia. *Neurology*, 76(14), pp.1263–9.

España, J. et al., 2010. Intraneuronal beta-amyloid accumulation in the amygdala enhances fear and anxiety in Alzheimer's disease transgenic mice. *Biological psychiatry*, 67(6), pp.513–21.

Fiesel, F.C. et al., 2010. Knockdown of transactive response DNA-binding protein (TDP-43) downregulates histone deacetylase 6. *The EMBO journal*, 29(1), pp.209–21.

Finsterer, J. & Burgunder, J.-M., 2014. Recent progress in the genetics of motor neuron disease. *European journal of medical genetics*, 57(2-3), pp.103–12.

Fragkouli, a et al., 2006. Sexually dimorphic effects of the Lhx7 null mutation on forebrain cholinergic function. *Neuroscience*, 137(4), pp.1153–64.

Fujita, Y. et al., 2008. Anterior horn cells with abnormal TDP-43 immunoreactivities show fragmentation of the Golgi apparatus in ALS. *Journal of the neurological sciences*, 269(1-2), pp.30–4.

Gamazon, E.R. & Stranger, B.E., 2014. Genomics of alternative splicing: evolution, development and pathophysiology. *Human genetics*, 133(6), pp.679–87.

Garg, N. et al., 2011. High fat diet induced insulin resistance and glucose intolerance are gender-specific in IGF-1R heterozygous mice. *Biochemical and biophysical research communications*, 413(3), pp.476–80.

Geser, F. et al., 2010. Pathological 43-kDa transactivation response DNA-binding protein in older adults with and without severe mental illness. *Archives of neurology*, 67(10), pp.1238–50.

Goldman, J.S. et al., 2005. Comparison of family histories in FTLD subtypes and related tauopathies. *Neurology*, 65(11), pp.1817–9.

Gómez-Sintes, R. et al., 2014. Mice with a naturally occurring DISC1 mutation display a broad spectrum of behaviors associated to psychiatric disorders. *Frontiers in behavioral neuroscience*, 8(July), p.253.

Guo, W. et al., 2011. An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nature structural & molecular biology*, 18(7), pp.822–30.

Guo, Y. et al., 2012. HO-1 induction in motor cortex and intestinal dysfunction in TDP-43 A315T transgenic mice. *Brain research*, 1460, pp.88–95.

Halliday, G. et al., 2012. Mechanisms of disease in frontotemporal lobar degeneration: gain of function versus loss of function effects. *Acta neuropathologica*, 124(3), pp.373–82.

Hardiman, O., van den Berg, L.H. & Kiernan, M.C., 2011. Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nature reviews. Neurology*, 7(11), pp.639–49.

Haverkamp, L.J., Appel, V. & Appel, S.H., 1995. Natural history of amyotrophic lateral sclerosis in a database population. Validation of a scoring system and a model for survival prediction. *Brain : a journal of neurology*, 118 (Pt 3), pp.707–19.

Heldt, S. a, Green, A. & Ressler, K.J., 2004. Prepulse inhibition deficits in GAD65 knockout mice and the effect of antipsychotic treatment. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 29(9), pp.1610–9.

Hernandez Lain, A. et al., 2011. Abnormal TDP-43 and FUS proteins in muscles of sporadic IBM: similarities in a TARDBP-linked ALS patient. *Journal of neurology, neurosurgery, and psychiatry*, 82(12), pp.1414–1416.

Highley, J.R. et al., 2014. Loss of nuclear TDP-43 in ALS causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurons. *Neuropathology and applied neurobiology*, pp.670–685.

Hodges, J.R. et al., 2003. Survival in frontotemporal dementia. *Neurology*, 61(3), pp.349–54.

Hornberger, M. & Piguet, O., 2012. Episodic memory in frontotemporal dementia: a critical review. *Brain : a journal of neurology*, 135(Pt 3), pp.678–92.

Huey, E.D., Putnam, K.T. & Grafman, J., 2006. A systematic review of neurotransmitter deficits and treatments in frontotemporal dementia. *Neurology*, 66(1), pp.17–22.

Igaz, L.M. et al., 2011. Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. The Journal of clinical investigation, 121(2), pp.726–38.

Igaz, L.M. et al., 2009. Expression of TDP-43 C-terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies. *The Journal of biological chemistry*, 284(13), pp.8516–24.

Ishikawa, T. et al., 2010. High-resolution melting curve analysis for rapid detection of mutations in a Medaka TILLING library. *BMC molecular biology*, 11, p.70.

Josephs, K. a et al., 2014. TDP-43 is a key player in the clinical features associated with Alzheimer's disease. *Acta neuropathologica*, pp.811–824.

Joyce, P.I. et al., 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), pp.420–48.

Justice, M.J. et al., 2000. Effects of ENU dosage on mouse strains. *Mammalian genome*, 11, pp.484–488.

Justice, M.J. et al., 1999. Mouse ENU mutagenesis. *Human molecular genetics*, 8(10), pp.1955–63.

Kawahara, Y. & Mieda-Sato, A., 2012. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), pp.3347–52.

Keays, D. a, Clark, T.G. & Flint, J., 2006. Estimating the number of coding mutations in genotypic- and phenotypic-driven N-ethyl-N-nitrosourea (ENU) screens. *Mammalian genome : official journal of the International Mammalian Genome Society*, 17(3), pp.230–8.

Keren, H., Lev-Maor, G. & Ast, G., 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nature reviews. Genetics*, 11(5), pp.345–55.

Kerr, T.M. et al., 2013. Genetic background modulates phenotypes of serotonin transporter Ala56 knock-in mice. *Molecular autism*, 4(1), p.35.

Khuong, H.T. et al., 2014. Skin derived precursor Schwann cells improve behavioral recovery for acute and delayed nerve repair. *Experimental neurology*, 254, pp.168–79.

Kiernan, M.C. et al., 2011. Amyotrophic lateral sclerosis. *Lancet*, 377(9769), pp.942–55.

Kim, H.J. et al., 2013. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*, 495(7442), pp.467–73.

Kim, M.-S. et al., 2014. A draft map of the human proteome. *Nature*, 509(7502), pp.575–81.

Kohl, S. et al., 2013. Prepulse inhibition in psychiatric disorders--apart from schizophrenia. *Journal of psychiatric research*, 47(4), pp.445–52.

Kondo, N. et al., 2010. DNA damage induced by alkylating agents and repair pathways. *Journal of nucleic acids*, 2010, p.543531.

Kornblihtt, A.R. et al., 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature reviews. Molecular cell biology*, 14(3), pp.153–65.

Kumari, V. et al., 2014. Prepulse Inhibition of the Startle Response in Men With Schizophrenia. *Archives of general Psychiatry*, 57(57), pp.609–614.

Kuo, P.-H. et al., 2009. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic acids research*, 37(6), pp.1799–808.

Kuo, P.-H. et al., 2014. The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic acids research*, 42(7), pp.4712–22.

Kurland, L.T & Mulder, D.W 1955. Epidemiologic Investigations of Amyotrophic Lateral Sclerosis: 2. Familial Aggregations Indicative of Dominant Inheritance Part I. *Neurology*, 5:182-96

Kwiatkowski, T.J. et al., 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 323(5918), pp.1205–8.

Lacomblez, L. et al., 1996. Articles Dose-ranging study of riluzole in amyotrophic lateral sclerosis. *Lancet*, 347, pp.1425–1431.

Liu, X. et al., 2012. Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. *The EMBO journal*, 31(23), pp.4415–27.

Liu, Y.-C., Chiang, P.-M. & Tsai, K.-J., 2013. Disease animal models of TDP-43 proteinopathy and their pre-clinical applications. *International journal of molecular sciences*, 14(10), pp.20079–111.

Lo, K.W., Kan, H. & Kevin, K., 2006. Identification of a Novel Region of the Cytoplasmic Dynein Intermediate Chain Important for Dimerization in the Absence of the Light Chains. Journal of biological chenistry, 281, pp.95552–59.

Logroscino, G. et al., 2010. Incidence of amyotrophic lateral sclerosis in Europe. *Journal of neurology, neurosurgery, and psychiatry*, 81(4), pp.385–90.

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

Maekawa, S. et al., 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), pp.672–83.

Maessen, M. et al., 2014. Euthanasia and physician-assisted suicide in amyotrophic lateral sclerosis: a prospective study. *Journal of neurology*, pp.1–8.

Majounie, E. et al., 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), pp.323–30.

Mancuso, R. et al., 2012. Effect of genetic background on onset and disease progression in the SOD1-G93A model of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis :*

official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases, 13(3), pp.302–10.

Mandillo, S. et al., 2014. Early motor deficits in mouse disease models are reliably uncovered using an automated home-cage wheel-running system: a cross-laboratory validation. *Disease models & mechanisms*, 7(3), pp.397–407.

McGoldrick, P. et al., 2013. Rodent models of amyotrophic lateral sclerosis. *Biochimica et biophysica acta*, 1832(9), pp.1421–36.

Medina, D.X., Orr, M.E. & Oddo, S., 2014. Accumulation of C-terminal fragments of transactive response DNA-binding protein 43 leads to synaptic loss and cognitive deficits in human TDP-43 transgenic mice. *Neurobiology of aging*, 35(1), pp.79–87.

Mercado, P.A. et al., 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), pp.6000–10.

Mikkelsen, T.S. et al., 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, 437(7055), pp.69–87.

Miller, R., Mitchell, J. & Moore, D., 2012. Riluzole for amyotrophic lateral sclerosis (ALS)/ motor neuron disease (MND) (Review). *The Chochrane Database of Systematic Reviews*, (3), p.CD001447.

Mioshi, E. et al., 2013. Cortical atrophy in ALS is critically associated with neuropsychiatric and cognitive changes. *Neurology*, 80(12), pp.1117–23.

Mioshi, E. et al., 2014. Neuropsychiatric changes precede classic motor symptoms in ALS and do not affect survival. *Neurology*, 82(2), pp.149–55.

Mitchell, J.D. & Borasio, G.D., 2007. Amyotrophic lateral sclerosis. *Lancet*, 369(9578), pp.2031–41.

Mori, F. et al., 2008. Maturation process of TDP-43-positive neuronal cytoplasmic inclusions in amyotrophic lateral sclerosis with and without dementia. *Acta neuropathologica*, 116(2), pp.193–203.

Morita, M. et al., 2006. A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology*, 66(6), pp.839–44.

Mott, R. & Flint, J., 2013. Dissecting quantitative traits in mice. *Annual review of genomics and human genetics*, 14, pp.421–39.

Neary, D., Snowden, J. & Mann, D., 2005. Frontotemporal dementia. *Lancet neurology*, 4(11), pp.771–80.

Neumann, M. et al., 2006. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 314(5796), pp.130–3.

Niksic, M. et al., 1999. Functional analysis of cis-acting elements regulating the alternative splicing of human CFTR exon 9. *Human molecular genetics*, 8(13), pp.2339–49.

Nonaka, T. et al., 2009. Phosphorylated and ubiquitinated TDP-43 pathological inclusions in ALS and FTLD-U are recapitulated in SH-SY5Y cells. *FEBS letters*, 583(2), pp.394–400.

Noveroske, J.K., Weber, J.S. & Justice, M.J., 2000. The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mammalian Genome*, 483, pp.478–483.

Nusbaum, M.P. & Contreras, D., 2004. Sensorimotor gating: startle submits to presynaptic inhibition. *Current biology : CB*, 14(6), pp.R247–9.

O'Neill, J.P., 2000. DNA damage, DNA repair, cell proliferation, and DNA replication: how do gene mutations result? *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), pp.11137–9.

O'Tuathaigh, C.M. et al., 2006. Sexually dimorphic changes in the exploratory and habituation profiles of heterozygous neuregulin-1 knockout mice. *Neuroreport*, 17(1), pp.79–83.

Onyike, C.U. & Diehl-Schmid, J., 2013. The epidemiology of frontotemporal dementia. *International review of psychiatry (Abingdon, England)*, 25(2), pp.130–7.

Ou, S.H. et al., 1995. Cloning and characterization of a novel cellular protein , TDP-43 , that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs *Journal of Virology*, 69(6), pp.3584–3586.

Pamphlett, R. et al., 2009. TDP-43 neuropathology is similar in sporadic amyotrophic lateral sclerosis with or without TDP-43 mutations. *Neuropathology and applied neurobiology*, 35(2), pp.222–5.

Pan, Q. et al., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics*, 40(12), pp.1413–5.

Pesiridis, G.S. et al., 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. *The Journal of biological chemistry*, 286(21), pp.18845–55.

Philips, T. & Rothstein, J.D., 2014. Glial cells in amyotrophic lateral sclerosis. *Experimental neurology*, Advanced publication (10.1016/j.expneurol.2014.05.015)

Phukan, J., Pender, N.P. & Hardiman, O., 2007. Cognitive impairment in amyotrophic lateral sclerosis. *Lancet neurology*, 6(11), pp.994–1003.

Polymenidou, M. et al., 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of. *Nature Publishing Group*, 14(4), pp.459–468.

Prudencio, M. et al., 2012. Misregulation of human sortilin splicing leads to the generation of a nonfunctional progranulin receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 109(52), pp.21510–15.

Raman, R. et al., 2014. Gene expression signatures in motor neuron disease fibroblasts reveal dysregulation of metabolism, hypoxia-response and RNA processing functions. *Neuropathology and applied neurobiology*, pp.1–48.

Renton, A.E. et al., 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72(2), pp.257–68.

Ricketts, T. et al., 2014. A nonsense mutation in mouse Tardbp affects TDP43 alternative splicing activity and causes limb-clasping and body tone defects. *PloS one*, 9(1), p.e85962. A

Ricketts, T., 2012. *Assessment of an allelic series of mouse TDP43 mutations*. PhD Tesis, University College London.

Rosen, D. et al., 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, 362(6415), pp.56–62.

Rousselet, E. et al., 2003. Behavioral changes are not directly related to striatal monoamine levels, number of nigral neurons, or dose of parkinsonian toxin MPTP in mice. *Neurobiology of Disease*, 14(2), pp.218–228.

Rutherford, N.J. et al., 2008. Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS genetics*, 4(9), p.e1000193.

Sayed, D. & Abdellatif, M., 2011. MicroRNAs in development and disease. *Physiological reviews*, 91(3), pp.827–87.

Sephton, C.F. et al., 2010. TDP-43 is a developmentally regulated protein essential for early embryonic development. *The Journal of biological chemistry*, 285(9), pp.6826–34.

Shrivastav, N., Li, D. & Essigmann, J.M., 2010. Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis*, 31(1), pp.59–70.

Singer, P. et al., 2013. Prepulse inhibition predicts working memory performance whilst startle habituation predicts spatial reference memory retention in C57BL/6 mice. *Behavioural brain research*, 242, pp.166–77.

Sleegers, K., Cruts, M. & Van Broeckhoven, C., 2010. Molecular pathways of frontotemporal lobar degeneration. *Annual review of neuroscience*, 33, pp.71–88.

Smith, B.N. et al., 2013. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *European journal of human genetics : EJHG*, 21(1), pp.102–8.

Sreedharan, J. et al., 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 319(5870), pp.1668–72.

Stallings, N.R. et al., 2010. Progressive motor weakness in transgenic mice expressing human TDP-43. *Neurobiology of disease*, 40(2), pp.404–14.

Stallings, N.R. et al., 2013. TDP-43, an ALS linked protein, regulates fat deposition and glucose homeostasis. *PloS one*, 8(8), p.e71793.

Strong, M.J. et al., 2007. TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. *Molecular and cellular neurosciences*, 35(2), pp.320–7.

Strong, M.J., 2008. The syndromes of frontotemporal dysfunction in amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 9(6), pp.323–38.

Sureau, a et al., 2001. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *The EMBO journal*, 20(7), pp.1785–96.

Swarup, V. et al., 2012. Abnormal regenerative responses and impaired axonal outgrowth after nerve crush in TDP-43 transgenic mouse models of amyotrophic lateral sclerosis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(50), pp.18186–95.

Swarup, V. et al., 2011. Pathological hallmarks of amyotrophic lateral sclerosis/frontotemporal lobar degeneration in transgenic mice produced with TDP-43 genomic fragments. *Brain : a journal of neurology*, 134(Pt 9), pp.2610–26.

Swerdlow, N.R. et al., 2008. Realistic expectations of prepulse inhibition in translational models for schizophrenia research. *Psychopharmacology*, 199(3), pp.331–88.

Swerdlow, N.R. et al., 2006. Startle Gating Deficits in a Large Cohort of Patients With Schizophrenia. *Archives of general psychiatry*, 63(63), pp.1325–1335.

Synofzik, M. et al., 2014. Targeted high-throughput sequencing identifies a TARDBP mutation as a cause of early-onset FTD without motor neuron disease. *Neurobiology of aging*, 35(5), pp.1212.e1–5.

Takahasi, K.R., Sakuraba, Y. & Gondo, Y., 2007. Mutational pattern and frequency of induced nucleotide changes in mouse ENU mutagenesis. *BMC molecular biology*, 8, p.52.

Tan, R.H. et al., 2013. Classification of FTLD-TDP cases into pathological subtypes using antibodies against phosphorylated and non-phosphorylated TDP43. *Acta neuropathologica communications*, 1(1), p.33.

Tanabe, L.M., Martin, C. & Dauer, W.T., 2012. Genetic background modulates the phenotype of a mouse model of DYT1 dystonia. *PloS one*, 7(2), p.e32245.

Tollervey, J.R. et al., 2011a. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature neuroscience*, 14(4), pp.452–8.

Tremblay, C. et al., 2011. Accumulation of transactive response DNA binding protein 43 in mild cognitive impairment and Alzheimer disease. *Journal of neuropathology and experimental neurology*, 70(9), pp.788–98.

Tsai, K.-J. et al., 2010. Elevated expression of TDP-43 in the forebrain of mice is sufficient to cause neurological and pathological phenotypes mimicking FTLD-U. *The Journal of experimental medicine*, 207(8), pp.1661–73.

Uchida, A. et al., 2012. Non-human primate model of amyotrophic lateral sclerosis with cytoplasmic mislocalization of TDP-43. *Brain : a journal of neurology*, 135(Pt 3), pp.833–46.

Vance, C. et al., 2006. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain : a journal of neurology*, 129(Pt 4), pp.868–76.

Vance, C. et al., 2009. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*, 323(February), pp.1208–1211.

Wang, H.-Y. et al., 2004. Structural diversity and functional implications of the eukaryotic TDP gene family. *Genomics*, 83(1), pp.130–139.

Wang, I.-F. et al., 2012. The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. *Nature communications*, 3(may 2011), p.766.

Waterston, R.H. et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–62.

Wegorzewska, I. et al., 2009. TDP-43 mutant transgenic mice develop features of ALS and frontotemporal lobar degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 106(44), pp.18809–14.

Wijesekera, L.C. et al., 2009. Natural history and clinical features of the flail arm and flail leg ALS variants. *Neurology*, 72(12), pp.1087–94.

Wils, H. et al., 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), pp.3858–63.

Wilson, A.C. et al., 2011. TDP-43 in aging and Alzheimer's disease - a review. *International journal of clinical and experimental pathology*, 4(2), pp.147–55.

Wollerton, M.C. et al., 2004. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Molecular cell*, 13(1), pp.91–100.

Wu, L.-S. et al., 2010. TDP-43, a neuro-pathosignature factor, is essential for early mouse embryogenesis. *Genesis (New York, N.Y. : 2000)*, 48(1), pp.56–62..

Wu, L.-S., Cheng, W.-C. & Shen, C.-K.J., 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. *The Journal of biological chemistry*, 287(33), pp.27335–44.

Xiao, S. et al., 2011. RNA targets of TDP-43 identified by UV-CLIP are deregulated in ALS. *Molecular and cellular neurosciences*, 47(3), pp.167–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21421050 [Accessed June 16, 2014].

Xu, Y.-F. et al., 2011. Expression of mutant TDP-43 induces neuronal dysfunction in transgenic mice. *Molecular neurodegeneration*, 6(1), p.73.

Xu, Y.-F. et al., 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), pp.10851–9.

Yang, C. et al., 2014. Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proceedings of the National Academy of Sciences of the United States of America*.

Van der Zee, J. et al., 2013. A pan-European study of the C9orf72 repeat associated with FTLD: geographic prevalence, genomic instability, and intermediate repeats. *Human mutation*, 34(2), pp.363–73.

Zhang, Y.-J. et al., 2013. The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Human molecular genetics*, 22(15), pp.3112–22.

APPENDIX 1

RNA-seq data table of significantly differentially spliced exons in *Tardbp*$^{M323K/M323K}$ embryo heads compared to controls

Ensemb ID	Gene Name	Chr	Gene Start bp	Gene End bp	Exon ID	Dispersion	p value	p adjust	Mean Base	log2fold(Hom/Ctl)
ENSMUSG00000042208	0610010F05Rik	11	23564961	23633639	E001	0.001399089	3.36E-05	0.012965557	1167.121839	-0.094815241
ENSMUSG00000030663	1110004F10Rik	7	116093301	116105210	E005	0.000948311	0.000190343	0.046161498	2469.846005	-0.053990732
ENSMUSG00000020441	2310033P09Rik	11	59208321	59210738	E006	0.011602083	0.000134761	0.035939467	90.20563297	-0.318787788
ENSMUSG00000026319	2310035C23Rik	1	105663861	105755131	E029	0.002495178	8.23E-05	0.024419234	511.3286861	-0.135181286
ENSMUSG00000024726	2410127L17Rik	19	18670780	18704788	E008	0.006269978	1.31E-06	0.000994392	438.5863644	-0.235972454
ENSMUSG00000022195	6030458C11Rik	15	12808177	12824657	E007	0.006782443	6.57E-06	0.003730812	159.9009841	0.289167826
ENSMUSG00000057230	Aak1	6	86849517	86991864	E021	0.001269454	8.17E-05	0.024402877	2862.661834	-0.056433871
ENSMUSG00000026782	Abi2	1	60409619	60481162	E009	0.003152506	6.79E-06	0.003804348	2597.621251	-0.110105394
ENSMUSG00000030861	Acadsb	7	131410601	131446211	E011	0.001379638	8.91E-05	0.02560778	1194.303291	-0.079816459
ENSMUSG00000015143	Actn1	12	80167542	80260371	E001	0.002316442	6.45E-10	2.01E-06	562.9052417	-0.209967926
ENSMUSG00000021076	Actr10	12	70937857	70964718	E012	0.001977725	5.18E-05	0.017990312	695.9339432	0.123611957
ENSMUSG00000026341	Actr3	1	125392905	125435727	E001	0.001762962	2.86E-08	4.48E-05	3159.218204	-0.117022018
ENSMUSG00000054693	Adam10	9	70678997	70780229	E016	0.001644233	3.32E-05	0.012890623	2457.497647	-0.085093143
ENSMUSG00000025026	Add3	19	53140445	53247399	E017	0.001192882	7.61E-06	0.004166898	1538.279486	-0.08607234
ENSMUSG00000051149	Adnp	2	168180986	168207112	E001	0.001177107	5.03E-06	0.002989501	6873.634316	-0.044300024
ENSMUSG00000015961	Adss	1	177762962	177796511	E001	0.001575407	3.07E-07	0.000294109	967.5150027	-0.129598823
ENSMUSG00000030232	Aebp2	6	140622449	140678472	E011	0.001193931	1.56E-08	2.88E-05	1535.795919	-0.098922291
ENSMUSG00000026159	Agfg1	1	82839483	82896275	E013	0.002115362	0.000168068	0.042076135	973.5547511	-0.102814349
ENSMUSG00000042410	Agps	2	75832177	75931350	E020	0.001009773	6.06E-05	0.019986346	2143.61563	-0.061978638
ENSMUSG00000019986	Ahi1	10	20952547	21080428	E025	0.001838461	6.70E-06	0.003764876	770.8321328	-0.130366981
ENSMUSG00000060935	AI597468	10	85102627	85117747	E003	0.002497204	5.57E-05	0.018821163	2407.023345	-0.067273861
ENSMUSG00000061603	Akap6	12	52699383	53151015	E014	0.002066388	4.85E-05	0.017015024	655.3908042	-0.115921165
ENSMUSG00000028291	Akirin2	4	34550937	34566908	E005	0.00165889	7.18E-06	0.003958169	1186.096022	-0.10597654
ENSMUSG00000028291	Akirin2	4	34550937	34566908	E001	0.001561897	3.25E-05	0.01267193	980.3618118	0.101602733
ENSMUSG00000022636	Alcam	16	52248996	52454074	E001	0.002508976	5.05E-11	2.27E-07	2285.335204	-0.165041663
ENSMUSG00000021238	Aldh6a1	12	84430723	84450950	E001	0.002324716	1.79E-05	0.007782798	560.2891246	-0.121706553

ENSMUSG00000025135	Anapc11	11	120598421	120608198	E005	0.002100375	0.0002013	0.048115342	1200.46509	0.08463761
ENSMUSG00000036777	Anln	9	22331214	22389206	E001	0.002588649	4.82E-06	0.002897022	487.9482151	-0.161417899
ENSMUSG00000015749	Anp32e	3	95929246	95947390	E007	0.002228831	8.29E-10	2.41E-06	3260.269351	-0.133890827
ENSMUSG00000015749	Anp32e	3	95929246	95947390	E004	0.001571951	1.76E-05	0.00771837	970.7695202	0.10759121
ENSMUSG00000074238	Ap1ar	3	127807264	127837492	E001	0.001338859	8.74E-06	0.004586916	1255.609504	-0.089472304
ENSMUSG00000024480	Ap3s1	18	46741917	46790826	E006	0.003315738	4.35E-08	6.50E-05	1359.365214	-0.171489118
ENSMUSG00000031996	Aplp2	9	31149557	31211815	E001	0.001439809	0.000165247	0.041595571	2369.053943	-0.075179237
ENSMUSG00000022892	App	16	84954440	85173707	E001	0.000806031	2.22E-11	1.05E-07	3813.286151	-0.095783963
ENSMUSG00000001127	Araf	X	20797814	20860519	E007	0.005503013	3.26E-07	0.000309042	748.3087365	-0.219585294
ENSMUSG00000040459	Arglu1	8	8666577	8690537	E001	0.002533596	7.11E-06	0.003945174	1737.673444	-0.110796118
ENSMUSG00000041225	Arhgap12	18	6024448	6136098	E001	0.001978933	0.000195004	0.047085328	881.9861606	-0.098385053
ENSMUSG00000007880	Arid1a	4	133679008	133756769	E001	0.001534305	0.000171451	0.042767831	3204.736071	-0.069482936
ENSMUSG00000069729	Arid1b	17	4994332	5347656	E020	0.004708504	0.0001613	0.040663797	1399.424257	-0.138455835
ENSMUSG00000033237	Arid2	15	96287518	96404992	E021	0.001325879	2.15E-09	5.58E-06	1276.46745	-0.13103659
ENSMUSG00000039219	Arid4b	13	14063232	14199603	E024	0.002658407	0.000183462	0.044886579	649.8437955	-0.121603067
ENSMUSG00000031776	Arl2bp	8	94666755	94674417	E008	0.001100564	1.71E-06	0.001243458	1793.643748	-0.085412193
ENSMUSG00000026960	Arl6ip6	2	53191726	53219220	E004	0.002028071	8.05E-08	0.000101944	672.3174566	-0.132978233
ENSMUSG00000074794	Arrdc3	13	80883384	80896035	E008	0.001035128	8.13E-05	0.02433398	2032.847953	-0.058082812
ENSMUSG00000021388	Aspn	13	49544443	49567557	E009	0.002414704	3.21E-14	4.44E-10	890.0598223	-0.223308407
ENSMUSG00000021388	Aspn	13	49544443	49567557	E005	0.002696614	5.09E-05	0.017714401	463.4697535	0.148138048
ENSMUSG00000022360	Atad2	15	58094047	58135082	E001	0.002686042	1.08E-08	2.11E-05	465.7577954	-0.210338825
ENSMUSG00000027104	Atf2	2	73816509	73892639	E001	0.005162719	3.44E-05	0.013157973	1589.307522	-0.155422537
ENSMUSG00000021066	Atl1	12	69893105	69964085	E015	0.004510066	9.02E-06	0.004675299	306.7297452	-0.209909527
ENSMUSG00000037400	Atp11b	3	35754134	35856274	E001	0.011397324	0.000156349	0.039717378	91.9075234	0.3299552
ENSMUSG00000029467	Atp2a2	5	122442702	122502225	E002	0.002377016	9.84E-05	0.027867868	1684.496269	-0.106493655
ENSMUSG00000025428	Atp5a1	18	77773729	77782869	E012	0.001072379	1.74E-05	0.00769456	3482.116389	-0.072765465
ENSMUSG00000031229	Atrx	X	105797615	105929397	E001	0.001733254	8.64E-14	9.55E-10	2370.447064	-0.163068466
ENSMUSG00000029673	Auts2	5	131437335	132543344	E001	0.000887919	3.25E-06	0.002114027	2904.127375	-0.064509863

ENSMUSG00000045767	B230219D22Rik	13	55693124	55703499	E001	0.003273937	7.80E-05	0.023699694	365.4395821	0.154468471
ENSMUSG00000045767	B230219D22Rik	13	55693124	55703499	E002	0.001546341	7.80E-05	0.023699694	5684.262527	-0.034069122
ENSMUSG00000026987	Baz2b	2	59899363	60209839	E001	0.00232052	3.39E-05	0.012978241	561.6129019	-0.137975173
ENSMUSG00000033799	BC016423	13	3566036	3611108	E001	0.004862061	3.01E-06	0.001972873	231.0216403	-0.250077633
ENSMUSG00000037608	Bclaf1	10	20312469	20342501	E014	0.000961837	1.40E-14	2.11E-10	2389.809198	-0.127734845
ENSMUSG00000040363	Bcor	X	12036740	12160355	E001	0.003213157	0.000198067	0.047569962	373.76248	-0.152044953
ENSMUSG00000049658	Bdp1	13	100017994	100104070	E001	0.002993107	4.88E-06	0.002919695	407.3510236	-0.175380789
ENSMUSG00000024073	Birc6	17	74528295	74703773	E073	0.002415624	7.38E-09	1.55E-05	533.068331	-0.206872876
ENSMUSG00000026739	Bmi1	2	18677018	18686629	E010	0.001120144	2.70E-08	4.38E-05	1732.639426	-0.094775196
ENSMUSG00000040481	Bptf	11	107033081	107132127	E001	0.002168976	1.88E-05	0.008016463	614.0030913	-0.137515552
ENSMUSG00000026918	Brd3	2	27445579	27507662	E009	0.002780609	0.000101757	0.028658253	446.060765	0.149072593
ENSMUSG00000022914	Brwd1	16	95992092	96082526	E039	0.04254252	6.95E-05	0.021925498	1635.353438	-0.437199701
ENSMUSG00000022914	Brwd1	16	95992092	96082526	E040	0.021571417	0.000173856	0.043204232	353.1120326	-0.33551502
ENSMUSG00000025103	Btbd1	7	81792075	81829431	E001	0.002189821	4.98E-08	7.12E-05	2075.8616	-0.130499414
ENSMUSG00000066979	Bub3	7	131560391	131571898	E008	0.001854952	5.79E-05	0.019293165	761.1316892	-0.11513595
ENSMUSG00000038658	C030046E11Rik	19	29522282	29606829	E026	0.003455381	0.000112193	0.031017887	800.5588601	-0.134353904
ENSMUSG00000033031	C330027C09Rik	16	48994185	49019709	E021	0.003006868	3.53E-06	0.002228567	416.7601915	-0.181976846
ENSMUSG00000040118	Cacna2d1	5	15934691	16374511	E041	0.001172555	2.05E-05	0.008683706	1588.063436	-0.078122469
ENSMUSG00000001175	Calm1	12	100199435	100209806	E006	0.000646856	0.000130648	0.035280238	9740.829863	-0.035433111
ENSMUSG00000036438	Calm2	17	87433412	87446935	E001	0.00173973	1.12E-06	0.000876698	5252.255965	-0.089579785
ENSMUSG00000041570	Camsap2	1	136268123	136346104	E001	0.001066263	8.77E-05	0.025396654	1911.549727	-0.067622714
ENSMUSG00000020368	Canx	11	50293961	50325673	E001	0.000732238	3.93E-09	9.45E-06	5311.790551	-0.07274265
ENSMUSG00000001794	Capns1	7	30186942	30195164	E001	0.001013502	2.60E-05	0.010713799	2126.571307	0.072286873
ENSMUSG00000001794	Capns1	7	30186942	30195164	E010	0.003968945	0.000180646	0.044393962	291.2727761	-0.182097653
ENSMUSG00000027184	Caprin1	2	103762941	103797649	E002	0.00166968	0	0	1114.112598	-0.350111785
ENSMUSG00000015733	Capza2	6	17636234	17666972	E010	0.002093284	0.00014721	0.038095992	2516.934542	-0.085936987
ENSMUSG00000028745	Capzb	4	139192899	139291818	E012	0.043405079	7.76E-08	9.98E-05	23.27218255	-1.021901038
ENSMUSG00000060227	Casc4	2	121866970	121936220	E009	0.004075719	1.78E-07	0.000193586	1800.970682	-0.16620382

ENSMUSG00000021585	Cast	13	74694286	74807921	E001	0.007126548	0.000104842	0.029426978	151.5415332	-0.255935045
ENSMUSG00000031885	Cbfb	8	105170674	105217986	E007	0.002282291	2.66E-11	1.23E-07	2915.253027	-0.142650779
ENSMUSG00000020074	Ccar1	10	62743928	62792368	E008	0.003564318	3.33E-10	1.13E-06	330.299924	-0.286581048
ENSMUSG00000020074	Ccar1	10	62743928	62792368	E007	0.010182897	8.74E-05	0.025396654	198.7746004	-0.277181765
ENSMUSG00000020074	Ccar1	10	62743928	62792368	E003	0.002508531	0.000151385	0.038875945	507.852533	0.138832808
ENSMUSG00000071855	Ccdc112	18	46282151	46311928	E001	0.006436833	4.69E-05	0.016651581	169.27979	-0.242077502
ENSMUSG00000027160	Ccdc34	2	110017817	110173360	E004	0.009928746	1.34E-07	0.000150009	192.9414023	-0.359696045
ENSMUSG00000027160	Ccdc34	2	110017817	110173360	E003	0.011263623	2.26E-05	0.009431922	171.4019891	-0.308541647
ENSMUSG00000032740	Ccdc88a	11	29373658	29510808	E032	0.001586451	8.61E-12	5.10E-08	1146.141762	-0.16730257
ENSMUSG00000029385	Ccng2	5	93267257	93276231	E008	0.001230313	8.22E-05	0.024419234	1454.328194	-0.076832782
ENSMUSG00000024286	Ccny	18	9314044	9450150	E001	0.001857261	9.57E-05	0.027132861	2242.586868	-0.083083568
ENSMUSG00000019818	Cd164	10	41519500	41531042	E006	0.001708929	2.21E-06	0.001542576	4607.262818	-0.07678137
ENSMUSG00000061665	Cd2ap	17	42792951	42876424	E001	0.005530971	0.000116918	0.032217057	1075.933126	-0.151762192
ENSMUSG00000024304	Cdh2	18	16588877	16809246	E001	0.001289759	1.66E-05	0.007428624	2207.837828	-0.083291572
ENSMUSG00000036510	Cdh8	8	99024471	99416471	E001	0.011045805	1.51E-05	0.006853945	94.98401264	-0.364221961
ENSMUSG00000020015	Cdk17	10	93160876	93241342	E017	0.003445175	2.79E-05	0.011254407	1100.541391	-0.138519335
ENSMUSG00000029635	Cdk8	5	146231230	146302874	E013	0.001742033	3.72E-05	0.01390765	832.8989025	-0.106762426
ENSMUSG00000003031	Cdkn1b	6	134920401	134925513	E003	0.002050627	1.32E-08	2.49E-05	1104.723221	-0.127281946
ENSMUSG00000032803	Cdv3	9	103353102	103365780	E001	0.001736261	2.15E-05	0.00901757	3913.97378	-0.076246865
ENSMUSG00000032803	Cdv3	9	103353102	103365780	E005	0.003693502	0.00018748	0.045600579	316.7500133	0.170683961
ENSMUSG00000005506	Celf1	2	90940382	91019497	E015	0.002791417	2.85E-05	0.01140733	827.3627572	-0.137946533
ENSMUSG00000002107	Celf2	2	6539705	7081302	E001	0.00255203	9.66E-06	0.004869375	4524.428478	-0.09780645
ENSMUSG00000033671	Cep350	1	155844966	155972890	E001	0.001969432	0.000128806	0.034855718	845.2223704	-0.10320167
ENSMUSG00000054604	Cggbp1	16	64851996	64859507	E003	0.000760803	4.66E-06	0.002833515	4610.472046	-0.032176995
ENSMUSG00000038467	Chmp4b	2	154651705	154694785	E005	0.001234111	7.35E-08	9.57E-05	1446.317984	-0.109100911
ENSMUSG00000028419	Chmp5	4	40948407	40965303	E008	0.00358615	7.37E-08	9.57E-05	916.8550954	-0.188173562
ENSMUSG00000001017	Chtop (2500003M10Rik)	3	90498956	90509498	E002	0.003053612	6.25E-07	0.00055141	397.5282684	0.204211268
ENSMUSG00000001017	Chtop (2500003M10Rik)	3	90498956	90509498	E003	0.001685778	0.000113789	0.031406864	873.9517362	0.103723926

ENSMUSG00000028165	Cisd2	3	135406412	135423925	E001	0.002791417	4.59E-09	1.04E-05	1044.873793	-0.136599303
ENSMUSG00000000605	Clcn4-2	7	7282316	7299511	E001	0.001098132	0.000152025	0.03897724	1801.522506	-0.064879156
ENSMUSG00000006169	Clint1	11	45852051	45910625	E012	0.002041488	7.53E-06	0.004136341	1620.069245	-0.108591494
ENSMUSG00000034390	Cmip (4933407C03Rik)	8	117257019	117461503	E021	0.004795575	6.37E-08	8.89E-05	234.6347525	-0.294833881
ENSMUSG00000028719	Cmpk1	4	114959336	114987241	E001	0.004374768	2.31E-07	0.000232646	1695.637842	-0.178323691
ENSMUSG00000020166	Cnot2	10	116485161	116581511	E001	0.002791417	0.000186688	0.045474542	1209.127167	-0.113265804
ENSMUSG00000034724	Cnot6l	5	96070333	96164171	E001	0.004613864	1.01E-08	2.02E-05	348.2841985	-0.258462876
ENSMUSG00000031601	Cnot7	8	40492540	40515847	E001	0.001666648	1.42E-07	0.00015851	1654.906532	-0.112733688
ENSMUSG00000027966	Col11a1	3	114030540	114220718	E067	0.003784765	0.000172701	0.043014962	2414.46514	-0.123193797
ENSMUSG00000029661	Col1a2	6	4504814	4541543	E052	0.000735291	8.28E-06	0.004387667	9840.006485	-0.050497289
ENSMUSG00000026837	Col5a1	2	27886425	28039514	E066	0.000839927	2.51E-05	0.010392901	3640.770124	-0.063817855
ENSMUSG00000026042	Col5a2	1	45374321	45503282	E001	0.001559982	0.000144046	0.037468527	2624.327877	-0.079363067
ENSMUSG00000026147	Col9a1	1	24177610	24252684	E038	0.001137933	3.30E-07	0.000310934	1680.705382	-0.102434712
ENSMUSG00000037852	Cpe	8	64592558	64693040	E001	0.002264633	2.19E-10	8.24E-07	2421.196425	-0.157881338
ENSMUSG00000039782	Cpeb2	5	43233463	43289724	E012	0.005900374	1.51E-05	0.006853945	752.1507842	-0.170934464
ENSMUSG00000055531	Cpsf6	10	117344673	117380015	E005	0.001472341	8.58E-05	0.025202906	1074.982418	0.095019586
ENSMUSG00000055531	Cpsf6	10	117344673	117380015	E001	0.003615465	8.63E-05	0.025294346	3612.739319	-0.108666796
ENSMUSG00000038002	Cramp1l	17	24961228	25015230	E001	0.001505491	6.98E-05	0.021925498	1037.902345	-0.083423091
ENSMUSG00000024074	Crim1	17	78200248	78376592	E017	0.006574817	1.43E-05	0.006594644	520.8219628	-0.205492328
ENSMUSG00000017776	Crk	11	75679259	75706908	E003	0.001519345	1.25E-05	0.005998719	4291.705322	-0.062600526
ENSMUSG00000068823	Csde1	3	103020546	103058189	E020	0.001213131	4.75E-08	6.93E-05	3189.175821	-0.098770953
ENSMUSG00000024576	Csnk1a1	18	61555274	61590061	E010	0.000991923	4.63E-06	0.002825917	3521.247927	-0.071305076
ENSMUSG00000024576	Csnk1a1	18	61555274	61590061	E002	0.001474061	0.000119499	0.032710735	1072.993566	0.09690922
ENSMUSG00000022433	Csnk1e	15	79417856	79443919	E002	0.003835926	7.31E-05	0.02267455	352.1751786	-0.180342558
ENSMUSG00000073563	Csnk1g3	18	53862113	53955684	E013	0.005214905	3.02E-05	0.011964291	1314.141063	-0.162695979
ENSMUSG00000030970	Ctbp2	7	132987563	133124354	E001	0.007120413	2.64E-05	0.010829284	1274.716676	-0.191498187
ENSMUSG00000005698	Ctcf	8	105636568	105682922	E012	0.001132021	1.77E-05	0.00775201	1697.618133	-0.080216685
ENSMUSG00000033411	Ctdspl2	2	121956001	122013642	E013	0.00140934	2.76E-08	4.38E-05	1153.288481	-0.115172998

ENSMUSG00000006932	Ctnnb1	9	120929216	120960507	E017	0.000646048	1.48E-05	0.006771308	9818.294969	-0.043904235
ENSMUSG00000031446	Cul4a	8	13105621	13147940	E020	0.001791283	2.55E-06	0.001751113	1137.904959	-0.12081516
ENSMUSG00000025898	Cwf19l2	9	3404085	3479236	E018	0.005352714	6.17E-06	0.003554776	207.4473265	-0.240842819
ENSMUSG00000022865	Cxadr	16	78301496	78359785	E008	0.002401891	3.36E-06	0.002159658	4726.572351	-0.092450114
ENSMUSG00000040651	D14Abb1e	14	27428847	27483552	E025	0.001679616	9.89E-05	0.027960125	878.6956842	-0.100155562
ENSMUSG00000037119	D15Ertd621e	15	58415468	58457801	E024	0.001014836	5.52E-05	0.018784507	2120.543818	-0.06739122
ENSMUSG00000041360	D19Bwg1357e	19	27388702	27429820	E001	0.0023951	1.23E-11	6.43E-08	538.980022	-0.232331114
ENSMUSG00000057858	D19Ertd737e	19	60198586	60226697	E001	0.006964132	1.18E-05	0.005736923	222.9638512	-0.249761654
ENSMUSG00000037266	D4Wsu53e	4	134923592	134927671	E004	0.001586176	1.33E-05	0.006253233	957.5126286	-0.109880376
ENSMUSG00000055639	Dach1	14	97786853	98169765	E001	0.006979107	9.65E-06	0.004869375	682.6041024	-0.203944852
ENSMUSG00000027797	Dclk1	3	55242526	55539064	E021	0.001202422	3.46E-05	0.013201303	3790.846017	-0.064335921
ENSMUSG00000019929	Dcn	10	97479500	97518162	E009	0.001056415	6.83E-08	9.21E-05	1948.319341	-0.098618256
ENSMUSG00000027708	Dcun1d1	3	35892105	35937445	E001	0.002699903	0.000122245	0.033407213	1298.846813	-0.096014975
ENSMUSG00000000787	Ddx3x	X	13280970	13294052	E005	0.0017247	6.34E-05	0.020779852	845.1308976	0.11467815
ENSMUSG00000000787	Ddx3x	X	13280970	13294052	E017	0.001607201	0.000106651	0.029783509	7262.810847	-0.063549901
ENSMUSG00000069045	Ddx3y	Y	1260715	1286613	E001	0.005766994	1.24E-13	1.29E-09	787.4603489	-0.450363464
ENSMUSG00000032097	Ddx6	9	44604892	44640731	E016	0.000669775	2.36E-06	0.00163745	7959.389337	-0.046320651
ENSMUSG00000029169	Dhx15	5	52150209	52190519	E001	0.005781925	1.81E-05	0.00783867	985.7036959	-0.190568161
ENSMUSG00000042699	Dhx9	1	153455758	153487660	E001	0.001644365	1.06E-08	2.09E-05	906.8571258	-0.160734323
ENSMUSG00000061689	Dlgap4	2	156613705	156764363	E015	0.001022293	1.62E-05	0.007312331	2087.450652	-0.072863426
ENSMUSG00000037416	Dmxl1	18	49832997	49965473	E043	0.005787113	5.81E-05	0.019314392	405.7000043	-0.202141946
ENSMUSG00000029131	Dnajb6	5	29735688	29786478	E012	0.004064665	1.06E-07	0.000124116	1301.094082	-0.190919636
ENSMUSG00000021905	Dph3	14	32080566	32085692	E001	0.00197289	5.71E-05	0.019101954	1063.503064	-0.073505882
ENSMUSG00000032393	Dpp8	9	65032458	65082650	E021	0.001205381	0.000196625	0.047407814	1756.208949	-0.072754603
ENSMUSG00000022048	Dpysl2	14	66802864	66868688	E001	0.002348559	0.000144108	0.037468527	2751.914181	-0.086850027
ENSMUSG00000024914	Drap1	19	5422805	5424979	E006	0.002125604	5.64E-05	0.018897894	630.8457953	-0.128175779
ENSMUSG00000024914	Drap1	19	5422805	5424979	E003	0.001838052	6.84E-05	0.021830634	771.0753697	0.114276422
ENSMUSG00000027012	Dync1i2	2	71211706	71263303	E016	0.002414027	1.18E-05	0.005736923	1565.382845	-0.122041676

ENSMUSG00000064061	Dzip3	16	48924232	48994165	E001	0.001467433	1.19E-05	0.005745632	1080.698818	-0.105141712
ENSMUSG00000010476	Ebf3	7	137193673	137314445	E001	0.002016904	7.60E-05	0.023343017	3427.232317	-0.084043284
ENSMUSG00000027699	Ect2	3	27097222	27153878	E001	0.002845078	2.32E-09	5.82E-06	433.5612388	-0.227581141
ENSMUSG00000034488	Edil3	13	88821472	89323223	E011	0.002167862	2.77E-08	4.38E-05	958.5418298	-0.139150652
ENSMUSG00000001707	Eef1e1	13	38645698	38659028	E001	0.002349696	3.14E-07	0.000299152	638.4771317	-0.138788773
ENSMUSG00000003070	Efna2	10	80179482	80190010	E004	0.001527013	5.09E-05	0.017714401	1015.167776	-0.076213071
ENSMUSG00000031987	Egln1	8	124908596	124949254	E005	0.001633381	3.05E-06	0.001993384	916.004535	0.112686211
ENSMUSG00000031987	Egln1	8	124908596	124949254	E001	0.002263165	6.95E-05	0.021925498	1922.145651	-0.092364758
ENSMUSG00000074656	Eif2s2	2	154871410	154892935	E007	0.00358842	8.93E-05	0.025625675	327.6846384	-0.17450115
ENSMUSG00000024991	Eif3a	19	60761116	60790693	E001	0.002526472	6.68E-05	0.021584238	2180.122721	-0.106338737
ENSMUSG00000022884	Eif4a2	16	23107444	23114136	E012	0.000806104	0.000135747	0.036090435	3812.229845	-0.051583145
ENSMUSG00000026254	Eif4e2	1	87213914	87240488	E010	0.001551701	2.58E-08	4.26E-05	990.285793	-0.13765058
ENSMUSG00000045983	Eif4g1	16	20668313	20692884	E032	0.001599896	2.63E-05	0.010819061	945.0658616	-0.116790386
ENSMUSG00000005610	Eif4g2	7	111067985	111083030	E002	0.000867412	3.84E-07	0.000349985	3088.534519	-0.083477302
ENSMUSG00000005610	Eif4g2	7	111067985	111083030	E001	0.001468726	7.21E-07	0.000616074	11058.32939	-0.075516925
ENSMUSG00000005610	Eif4g2	7	111067985	111083030	E009	0.000840972	4.72E-05	0.016714324	3363.932324	0.063931324
ENSMUSG00000040731	Eif4h	5	134619880	134639348	E003	0.003003467	0	0	405.6348155	-0.372687905
ENSMUSG00000040731	Eif4h	5	134619880	134639348	E002	0.001269437	0.000183142	0.044874382	1375.84344	0.080902086
ENSMUSG00000026083	Eif5b	1	37998010	38055579	E005	0.00755462	3.22E-12	2.13E-08	142.2877225	-0.498433023
ENSMUSG00000026083	Eif5b	1	37998010	38055579	E006	0.008057284	5.01E-07	0.000444411	132.7676201	-0.367422318
ENSMUSG00000058070	Eml1	12	108370957	108539617	E023	0.001922541	4.85E-05	0.017015024	723.8019035	-0.114174149
ENSMUSG00000069495	Epc2	2	49451486	49551948	E014	0.004838463	1.66E-05	0.007428624	744.8054155	-0.176807654
ENSMUSG00000028289	Epha7	4	28813131	28967499	E010	0.015693258	7.97E-05	0.024110973	65.844166	-0.39862457
ENSMUSG00000021709	Erbb2ip	13	103818787	103920514	E001	0.003294256	4.97E-08	7.12E-05	1913.792621	-0.162853537
ENSMUSG00000022034	Esco2	14	65819038	65833994	E001	0.004442233	0.000126535	0.034353417	255.904883	-0.177963737
ENSMUSG00000021171	Esyt2	12	116281222	116373096	E022	0.001454863	1.42E-05	0.006569114	1095.620482	-0.093453269
ENSMUSG00000032314	Etfa	9	55454508	55512243	E001	0.005467111	6.27E-05	0.020609969	202.6264734	-0.232695477
ENSMUSG00000009079	Ewsr1	11	5069689	5099266	E001	0.001446655	7.76E-06	0.004190909	1105.587909	-0.113135938

ENSMUSG00000038646	Fam103a1	7	81762925	81769491	E003	0.00222457	1.05E-07	0.000124116	684.8112824	-0.1180282
ENSMUSG00000047368	Fam108b	19	21653309	21685637	E004	0.00196976	9.23E-07	0.000765177	699.8226532	-0.118160332
ENSMUSG00000041040	Fam117b	1	59913006	59985346	E008	0.003676888	1.33E-05	0.006253233	2489.361608	-0.118871245
ENSMUSG00000038014	Fam120a	13	48879219	48968017	E001	0.002919814	4.83E-05	0.017015024	2558.736554	-0.111587076
ENSMUSG00000026153	Fam135a	1	24010758	24100341	E001	0.002894305	7.53E-12	4.62E-08	424.4783935	-0.262275963
ENSMUSG00000043542	Fam164a	3	7503426	7553848	E009	0.002657297	9.99E-10	2.81E-06	1374.383329	-0.159950169
ENSMUSG00000037503	Fam168b	1	34813219	34843065	E001	0.001576255	3.29E-06	0.002128609	10159.37167	-0.063316312
ENSMUSG00000051185	Fam174a	1	95313628	95335284	E003	0.005752728	3.81E-05	0.014180944	191.5146191	-0.203570946
ENSMUSG00000027751	Fam48a	3	54693105	54716837	E020	0.007421022	3.66E-07	0.000337238	145.0520936	-0.346524911
ENSMUSG00000022378	Fam49b	15	63929097	64060448	E001	0.001675087	4.07E-06	0.002506016	1377.269456	-0.106159852
ENSMUSG00000037808	Fam76b	9	13827716	13846541	E010	0.005230917	8.74E-05	0.025396654	936.9579625	-0.147599516
ENSMUSG00000027349	Fam98b	2	117249739	117271540	E008	0.002109613	1.13E-05	0.005508341	637.2908152	-0.132139665
ENSMUSG00000030759	Far1	7	113513834	113571511	E014	0.007143568	0.000126917	0.034400726	1744.563789	-0.164489485
ENSMUSG00000074505	Fat3	9	15910205	16378231	E001	0.002308237	0.000132411	0.035541494	879.508919	-0.097801689
ENSMUSG00000022124	Fbxl3	14	103080239	103099566	E001	0.002175269	1.89E-07	0.00020199	1209.742638	-0.112742625
ENSMUSG00000027180	Fbxo3	2	104027721	104063240	E012	0.001589765	0.000153633	0.039259135	954.2250424	-0.092092085
ENSMUSG00000041685	Fcho2	13	98723407	98815449	E001	0.005353995	3.32E-08	5.15E-05	885.8211936	-0.228454228
ENSMUSG00000037712	Fermt2	14	45458792	45530118	E001	0.001706244	8.96E-07	0.00074724	1072.174606	-0.124003944
ENSMUSG00000036053	Fmnl2	2	52857860	53134202	E027	0.001242038	5.18E-09	1.12E-05	1429.882012	-0.122016651
ENSMUSG00000000838	Fmr1	X	68678541	68717963	E017	0.003067391	6.93E-10	2.06E-06	1807.134556	-0.176969807
ENSMUSG00000026193	Fn1	1	71585523	71653171	E001	0.001891445	0.000158647	0.04017823	2959.504103	-0.086825343
ENSMUSG00000033487	Fndc3a	14	72537946	72710003	E001	0.001255183	3.71E-05	0.01390765	1661.364851	-0.080284601
ENSMUSG00000039286	Fndc3b	3	27416162	27710439	E001	0.001273992	1.05E-05	0.005197839	1367.253136	-0.083742972
ENSMUSG00000035992	Fnip1	11	54438199	54518235	E018	0.007915733	0.000166873	0.041877795	548.3147326	-0.193861262
ENSMUSG00000039275	Foxk2	11	121259990	121309896	E009	0.00125612	8.56E-05	0.025167908	2680.343824	-0.064978853
ENSMUSG00000026843	Fubp3	2	31572651	31617526	E018	0.004024225	8.22E-06	0.004371892	886.8222314	-0.162379124
ENSMUSG00000022800	Fyttd1	16	32877500	32908963	E010	0.001481895	2.31E-07	0.000232646	2486.799372	-0.093796478
ENSMUSG00000035293	G2e3	12	51348061	51376986	E017	0.001256506	6.43E-07	0.000561366	1400.829811	-0.088937708

ENSMUSG00000031714	Gab1	8	80764434	80880479	E001	0.001522373	4.34E-05	0.015627978	1019.984948	-0.089401514
ENSMUSG00000024462	Gabbr1	17	37045966	37075067	E023	0.001833584	6.31E-06	0.003607612	1264.319256	-0.116021333
ENSMUSG00000008976	Gabpa	16	84834925	84863779	E010	0.001920167	2.60E-06	0.001769173	1650.612176	-0.093594967
ENSMUSG00000029432	Gbas	5	129725063	129758327	E010	0.001411159	0.000176769	0.043570284	1150.867695	-0.081022907
ENSMUSG00000022707	Gbe1	16	70313949	70569716	E016	0.014441135	1.04E-05	0.005155953	71.77688093	-0.419727926
ENSMUSG00000020530	Ggnbp2	11	84832361	84870817	E003	0.012390736	1.42E-05	0.006569114	209.3546056	-0.326823219
ENSMUSG00000041028	Ghitm	14	37120444	37135322	E001	0.001200057	2.36E-07	0.00023366	1521.443437	-0.102271432
ENSMUSG00000021552	Gkap1	13	58233351	58274188	E001	0.006426254	1.85E-05	0.007953109	169.5842516	-0.267962176
ENSMUSG00000019715	Gle1	2	29935426	29960371	E016	0.002301816	8.80E-05	0.025431003	567.5902749	-0.130049718
ENSMUSG00000028020	Glrb	3	80843599	80913660	E001	0.004221489	4.98E-06	0.002971345	271.2676682	-0.188144316
ENSMUSG00000028901	Gmeb1	4	132221025	132261602	E001	0.001522886	2.48E-08	4.22E-05	1019.449817	-0.107039852
ENSMUSG00000057614	Gnai1	5	18265135	18360413	E001	0.001544954	1.11E-16	2.63E-12	996.9640932	-0.174378333
ENSMUSG00000031748	Gnao1	8	93809966	93969388	E011	0.001203241	8.44E-13	6.37E-09	1514.091242	-0.146424806
ENSMUSG00000031748	Gnao1	8	93809966	93969388	E008	0.004603602	0.000135948	0.036090435	2470.952815	0.129966466
ENSMUSG00000029064	Gnb1	4	155491361	155559269	E002	0.00509259	3.11E-05	0.0122428	516.4599657	-0.195524512
ENSMUSG00000034109	Golim4	3	75876183	75956949	E001	0.001488399	9.07E-05	0.025896572	1056.69595	-0.088885218
ENSMUSG00000014959	Gorasp2	2	70661576	70712636	E002	0.003836003	4.39E-05	0.015764692	303.0369466	-0.190940197
ENSMUSG00000032745	Gpbp1	13	111425680	111490111	E012	0.001985725	1.34E-06	0.001014726	1147.670395	-0.129095461
ENSMUSG00000027346	Gpcpd1	2	132529082	132587729	E001	0.003160906	1.87E-07	0.000201016	797.5599503	-0.178681482
ENSMUSG00000031517	Gpm6a	8	54954728	55060877	E007	0.001434311	6.54E-10	2.01E-06	5426.787195	-0.098378552
ENSMUSG00000033981	Gria2	3	80684936	80802791	E001	0.015331958	0.000168171	0.042076135	1081.430366	-0.253357552
ENSMUSG00000044221	Grsf1	5	88659448	88676171	E001	0.001238673	4.98E-05	0.017384177	1436.813116	-0.083381816
ENSMUSG00000056870	Gulp1	1	44551511	44796833	E011	0.004307822	8.15E-09	1.67E-05	497.6570835	-0.224434214
ENSMUSG00000060743	H3f3a	1	180800832	180813943	E001	0.004407761	1.02E-05	0.005087832	4983.70423	-0.128983003
ENSMUSG00000004698	Hdac9	12	34371503	34528889	E001	0.004638076	3.83E-05	0.014195063	243.6621355	-0.174808674
ENSMUSG00000021109	Hif1a	12	73901375	73947530	E015	0.002729863	0	0	1661.809166	-0.34305464
ENSMUSG00000008730	Hipk1	3	103739815	103791563	E001	0.001261517	2.81E-05	0.011324959	1969.997693	-0.070704832
ENSMUSG00000027177	Hipk3	2	104426481	104494446	E001	0.001163908	0.00017518	0.043307312	1610.231516	-0.065927524

ENSMUSG00000066551	Hmgb1	5	149046702	149053513	E001	0.001989785	8.87E-05	0.025543684	7267.852217	-0.07057711
ENSMUSG00000054717	Hmgb2	8	57511843	57515999	E005	0.000943399	5.57E-05	0.018821163	2500.258923	-0.064494088
ENSMUSG00000034518	Hmgxb4	8	74993356	75031978	E011	0.001770738	8.85E-05	0.025535152	828.5604881	-0.094455706
ENSMUSG00000046434	Hnrnpa1	15	103240432	103244955	E008	0.001497418	9.39E-06	0.004811851	1046.694839	0.114575552
ENSMUSG00000007850	Hnrnph1	11	50376990	50386528	E014	0.002311621	1.97E-12	1.42E-08	4137.298034	-0.160088172
ENSMUSG00000021546	Hnrnpk	13	58391142	58403343	E001	0.002028995	2.12E-05	0.008941711	3782.460631	-0.094977113
ENSMUSG00000059208	Hnrnpm	17	33646233	33686860	E001	0.001552474	2.05E-06	0.001454633	1384.168446	-0.115215842
ENSMUSG00000066037	Hnrnpr	4	136310942	136340681	E011	0.001081296	0.000201036	0.048115342	2275.699126	-0.065555213
ENSMUSG00000071659	Hnrnpul2	19	8819401	8834142	E014	0.000916951	8.15E-06	0.004359127	2677.782982	-0.070314844
ENSMUSG00000024095	Hnrpll	17	80029487	80062334	E001	0.00178003	0	0	807.2852138	-0.240709461
ENSMUSG00000007617	Homer1	13	93303757	93404129	E012	0.002374424	8.94E-06	0.004646506	545.0696767	-0.130440368
ENSMUSG00000021270	Hsp90aa1	12	110690605	110702728	E007	0.000831918	5.85E-05	0.019357414	3469.882467	-0.060467809
ENSMUSG00000020361	Hspa4	11	53259814	53300457	E001	0.000897702	2.38E-05	0.009878031	2823.700115	-0.067468891
ENSMUSG00000020361	Hspa4	11	53259814	53300457	E007	0.002232935	4.74E-05	0.016780071	590.7448562	0.140024756
ENSMUSG00000048583	Igf2	7	142650766	142666816	E005	0.000810707	1.03E-05	0.005112714	3746.31998	0.059274626
ENSMUSG00000019975	Ikbip	10	91082940	91102607	E003	0.005208356	7.89E-05	0.023928874	649.1378547	-0.139217651
ENSMUSG00000032178	Ilf3	9	21368019	21405361	E019	0.001461554	8.56E-06	0.004507776	1281.161206	-0.106706448
ENSMUSG00000066324	Impad1	4	4762484	4793355	E005	0.003939826	4.66E-05	0.016581797	430.7893897	0.164132005
ENSMUSG00000030662	Ipo5	14	120911194	120948042	E026	0.001093197	2.82E-06	0.00188609	1817.725305	-0.092855729
ENSMUSG00000030662	Ipo5	14	120911194	120948042	E024	0.001670603	0.000139988	0.036801219	885.7287656	0.107517236
ENSMUSG00000051243	Islr2	9	58196298	58204319	E002	0.020530763	1.34E-14	2.11E-10	49.90718138	-0.971549687
ENSMUSG00000027598	Itch	2	155133509	155226855	E025	0.001218685	7.28E-05	0.022603324	1479.408414	-0.07790815
ENSMUSG00000031703	Itfg1	8	85717557	85840949	E001	0.001337545	1.35E-05	0.006319912	1257.690289	-0.10066132
ENSMUSG00000027087	Itgav	2	83724397	83806916	E030	0.001402815	3.99E-09	9.46E-06	1162.054421	-0.1195362
ENSMUSG00000031239	Itm2a	X	107397099	107403376	E001	0.002030434	2.80E-07	0.000273228	3398.160289	-0.112420766
ENSMUSG00000024384	Iws1	18	32067734	32104331	E014	0.003710753	5.52E-06	0.003233407	1611.309414	-0.140646469
ENSMUSG00000038518	Jarid2	13	44729474	44921643	E019	0.002033631	0.000150105	0.038664145	669.8074705	-0.115473575
ENSMUSG00000037876	Jmjd1c	10	67096125	67256326	E026	0.003272762	2.49E-08	4.22E-05	365.5969441	-0.235978343

ENSMUSG00000028790	Khdrbs1	4	129703164	129742303	E001	0.001179779	2.89E-10	1.04E-06	4395.325536	-0.099992062
ENSMUSG00000021693	Kif2a	13	106958996	107022126	E001	0.00125544	1.48E-07	0.000164214	1402.930426	-0.110514806
ENSMUSG00000006740	Kif5b	18	6201002	6242174	E001	0.00176212	3.21E-13	2.80E-09	2190.61503	-0.162422581
ENSMUSG00000006740	Kif5b	18	6201002	6242174	E016	0.004662143	4.35E-07	0.000392076	242.2380325	-0.26916607
ENSMUSG00000006740	Kif5b	18	6201002	6242174	E003	0.002581918	4.80E-05	0.016949846	489.5601496	0.151332327
ENSMUSG00000006740	Kif5b	18	6201002	6242174	E017	0.004599404	7.05E-05	0.022050951	245.9859473	-0.210346775
ENSMUSG00000025597	Klhl4	X	114474333	114560829	E011	0.006372219	3.64E-06	0.002270083	171.1566198	-0.259379042
ENSMUSG00000054920	Klhl5	5	65107568	65168188	E011	0.002234857	2.83E-07	0.000274731	590.0732352	-0.160987224
ENSMUSG00000021929	Kpna3	14	61365186	61439947	E001	0.003893013	4.05E-05	0.014808336	2218.206007	-0.129960087
ENSMUSG00000021843	Ktn1	14	47663756	47736563	E042	0.004354874	3.50E-07	0.000326346	319.54882	-0.246805287
ENSMUSG00000016534	Lamp2	X	38401357	38456454	E002	0.001644721	6.87E-05	0.02184577	2284.329943	-0.079083775
ENSMUSG00000023025	Larp4	15	99970074	100016351	E018	0.001199881	4.47E-09	1.03E-05	1521.853509	-0.098303831
ENSMUSG00000033499	Larp4b	13	9093905	9173086	E018	0.001332347	2.78E-05	0.011254407	1265.988494	-0.085545845
ENSMUSG00000004880	Lbr	1	181815316	181842401	E001	0.001788429	3.87E-06	0.00240338	911.7164511	-0.116335141
ENSMUSG00000048661	Lemd3	10	120923413	120979332	E001	0.002382852	9.98E-08	0.000121789	542.5708455	-0.158740413
ENSMUSG00000000247	Lhx2	2	38339281	38369733	E005	0.001317452	1.00E-06	0.000812816	1290.38316	-0.098784721
ENSMUSG00000019920	Lims1	10	58323466	58424691	E010	0.004729032	0.000105728	0.029625669	2214.95637	-0.121650455
ENSMUSG00000027162	Lin7c	2	109890853	109901003	E005	0.002081579	2.98E-06	0.001966598	4539.449753	-0.078946533
ENSMUSG00000021484	Lman2	13	55343833	55362783	E001	0.005214691	3.81E-05	0.014180944	2565.491385	-0.147097935
ENSMUSG00000024590	Lmnb1	18	56707813	56753424	E011	0.001286287	8.94E-05	0.025625675	2019.340023	-0.079598853
ENSMUSG00000024529	Lox	18	52516069	52529867	E001	0.00091543	1.18E-08	2.28E-05	4030.596359	-0.071758349
ENSMUSG00000031290	Lrch2	X	147470375	147554081	E001	0.002414564	2.33E-13	2.14E-09	533.3703778	-0.247055823
ENSMUSG00000066568	Lsm14a	7	34344646	34389540	E001	0.005323991	0.000168823	0.042175676	1421.797109	-0.148580181
ENSMUSG00000039108	Lsm14b	2	180024987	180035465	E009	0.001174235	1.80E-13	1.76E-09	1583.825357	-0.14661336
ENSMUSG00000036446	Lum	10	97565501	97572703	E003	0.001688585	2.90E-06	0.001933722	2522.373314	-0.086558497
ENSMUSG00000025903	Lypla1	1	4807788	4886770	E009	0.004130367	9.43E-06	0.004811851	1057.265649	-0.152752077
ENSMUSG00000025151	Maged1	X	94535474	94542143	E001	0.001267684	0.000118248	0.032475554	1379.179299	-0.088273451
ENSMUSG00000024242	Map4k3	17	80580512	80728093	E001	0.001566831	5.90E-06	0.003421814	975.6302415	-0.115543961

ENSMUSG00000026074	Map4k4	1	39900913	40026310	E031	0.001390396	7.94E-08	0.000101376	3281.056607	-0.101679776
ENSMUSG00000063358	Mapk1	16	16983382	17047453	E009	0.005687907	9.20E-05	0.02616364	223.8294136	-0.222316109
ENSMUSG00000027479	Mapre1	2	153741274	153773310	E007	0.000751142	4.46E-09	1.03E-05	5937.020675	-0.061946879
ENSMUSG00000026977	March7	2	60209887	60250676	E010	0.002361428	2.95E-10	1.04E-06	548.9683453	-0.210056063
ENSMUSG00000007411	Mark3	12	111574510	111656227	E017	0.001505399	5.43E-05	0.018578837	1038.001633	-0.101122285
ENSMUSG00000042032	Mat2b	11	40679314	40695203	E001	0.001917066	0.000145917	0.037879499	726.6887325	-0.100088164
ENSMUSG00000020583	Matn3	12	8947929	8972028	E008	0.008455875	2.37E-06	0.00163745	126.0785599	-0.300086691
ENSMUSG00000037236	Matr3	18	35562158	35592045	E015	0.004536054	2.60E-08	4.26E-05	4155.313172	-0.177163818
ENSMUSG00000037236	Matr3	18	35562158	35592045	E001	0.007017457	1.93E-06	0.001388949	235.7756134	-0.287596303
ENSMUSG00000030678	Maz	7	127022143	127026479	E005	0.001969699	0.000110205	0.030621508	1293.900791	0.098934412
ENSMUSG00000022139	Mbnl2	14	120275669	120431697	E009	0.00313373	1.03E-09	2.84E-06	1624.74115	-0.162634532
ENSMUSG00000038612	Mcl1	3	95658721	95663176	E003	0.002045987	0.000141126	0.036983073	3604.40286	-0.067966292
ENSMUSG00000034297	Med13	11	86267033	86357602	E001	0.002018792	0.000144034	0.037468527	1654.586227	-0.090959898
ENSMUSG00000026150	Mff	1	82724890	82752394	E009	0.002053225	0.000154651	0.039406671	1367.945697	-0.097680118
ENSMUSG00000033943	Mga	2	119897228	119969581	E025	0.001500091	3.03E-05	0.011964291	1543.87459	-0.08727186
ENSMUSG00000025220	Mgea5	19	45750261	45783520	E001	0.002049359	0.000198158	0.047569962	3127.734704	-0.081029502
ENSMUSG00000038056	Mll3	5	25271798	25498783	E001	0.002075734	1.88E-06	0.001363567	651.3908111	-0.147878654
ENSMUSG00000062270	Morf4l1	9	90091671	90114820	E001	0.004606227	8.18E-05	0.024402877	2979.293826	-0.130952505
ENSMUSG00000069769	Msi2	11	88339382	88718513	E001	0.001408418	0.000204511	0.048742421	4098.543501	-0.055029071
ENSMUSG00000052915	Msl1	11	98795516	98807859	E009	0.001560356	6.09E-06	0.003519632	2396.810024	-0.088199352
ENSMUSG00000032479	Mtap4	9	109931458	110083947	E010	0.002255484	3.72E-05	0.01390765	582.9595785	0.141840413
ENSMUSG00000024012	Mtch1	17	29332072	29347904	E012	0.00130824	3.34E-05	0.012911877	1305.945845	0.097434582
ENSMUSG00000029267	Mtf2	5	108065674	108109004	E015	0.001249864	1.44E-06	0.00107744	1414.019164	-0.093890349
ENSMUSG00000031918	Mtmr2	9	13748410	13806481	E015	0.001743176	0.000147095	0.038095992	832.1042403	-0.102879109
ENSMUSG00000020900	Myh10	11	68691559	68816632	E041	0.0008934	0.000153249	0.039230433	2858.514404	-0.061594704
ENSMUSG00000024048	Myl12a	17	70993656	71002878	E004	0.002891298	3.36E-08	5.16E-05	425.0222706	-0.216135278
ENSMUSG00000024048	Myl12a	17	70993656	71002878	E002	0.001332618	6.30E-06	0.003607612	1265.552551	0.094111279
ENSMUSG00000010505	Myt1	2	181763332	181827797	E025	0.002904175	6.91E-05	0.021883493	704.2037393	-0.131750386

ENSMUSG00000036282	Naa30	14	49172226	49191031	E005	0.001537667	1.59E-06	0.001171695	1408.343309	-0.075108819
ENSMUSG00000036282	Naa30	14	49172226	49191031	E002	0.003435516	6.61E-06	0.003730812	345.0154492	0.182462678
ENSMUSG00000044155	Naa38	6	18848579	18855751	E004	0.002351602	0.000166217	0.041776315	551.95334	-0.110007636
ENSMUSG00000022698	Naa50	16	44139830	44163366	E005	0.001301034	1.91E-05	0.008110224	3788.873575	-0.055716316
ENSMUSG00000022253	Nadk2 (1110020G09Rik)	15	9071260	9110487	E013	0.00195425	2.44E-08	4.22E-05	844.8288916	-0.142784187
ENSMUSG00000020572	Nampt	12	32820335	32853369	E011	0.002000031	2.08E-07	0.000215942	685.2691422	-0.147900999
ENSMUSG00000058799	Nap1l1	10	111473192	111498142	E015	0.001324402	2.17E-05	0.009070537	1278.884265	0.102020804
ENSMUSG00000028693	Nasp	4	116601052	116627941	E010	0.001104362	3.72E-07	0.00034118	1781.479553	0.100852511
ENSMUSG00000020647	Ncoa1	12	4247363	4477182	E001	0.001821724	7.23E-05	0.022499041	780.933012	-0.10366621
ENSMUSG00000005886	Ncoa2	1	13139105	13374083	E001	0.001194585	5.56E-05	0.018821163	1534.250502	-0.075424279
ENSMUSG00000056234	Ncoa4	14	32159865	32179855	E011	0.001506429	7.79E-05	0.023699694	1036.889722	-0.079935773
ENSMUSG00000018501	Ncor1	11	62316426	62458541	E032	0.009713809	1.50E-06	0.001117434	108.7819656	-0.385645798
ENSMUSG00000018501	Ncor1	11	62316426	62458541	E001	0.001506134	8.04E-05	0.024188387	1037.208178	-0.100768346
ENSMUSG00000062661	Ncs1	2	31245823	31295989	E008	0.00170046	1.12E-06	0.000876698	3731.548706	-0.083292237
ENSMUSG00000019988	Nedd1	10	92684746	92722420	E001	0.010473899	2.10E-06	0.001476885	383.2341812	-0.294952799
ENSMUSG00000032216	Nedd4	9	72662557	72749842	E029	0.000591351	7.10E-06	0.003945174	21270.15493	-0.036141851
ENSMUSG00000020982	Nemf	12	69311544	69357176	E001	0.010031939	7.76E-05	0.023699694	388.763835	-0.236456191
ENSMUSG00000032340	Neo1	9	58874679	59036441	E001	0.00128375	5.77E-06	0.003369696	1938.777253	-0.09227585
ENSMUSG00000028565	Nfia	4	97772734	98118874	E012	0.005009998	1.05E-05	0.005197839	2385.073606	-0.142727801
ENSMUSG00000020248	Nfyb	10	82748701	82764144	E001	0.001697716	2.36E-08	4.13E-05	1044.877085	-0.127733851
ENSMUSG00000020936	Nmt1	11	103028190	103068912	E001	0.009758881	3.03E-05	0.011964291	108.249859	-0.333404677
ENSMUSG00000020936	Nmt1	11	103028190	103068912	E002	0.0062519	3.33E-05	0.012900266	174.7648061	-0.257311089
ENSMUSG00000041923	Nol4	18	22693181	23041653	E001	0.002532669	5.83E-05	0.019327811	501.6868246	-0.126333265
ENSMUSG00000026020	Nop58	1	59685006	59711510	E015	0.004626027	1.11E-15	2.30E-11	244.381448	-0.429784899
ENSMUSG00000026020	Nop58	1	59685006	59711510	E014	0.003777711	3.70E-14	4.72E-10	308.5003624	-0.35869488
ENSMUSG00000021047	Nova1	12	46694517	46818775	E001	0.00162074	3.41E-06	0.002173757	2881.935024	-0.071013414
ENSMUSG00000057113	Npm1	11	33152287	33163206	E001	0.003668043	2.87E-05	0.011463842	3152.717835	-0.125199765
ENSMUSG00000069171	Nr2f1	13	78188973	78199757	E001	0.001022366	8.27E-08	0.000103151	2087.130313	-0.079930901

ENSMUSG00000030551	Nr2f2	7	70351946	70366735	E001	0.003049093	2.41E-07	0.000236684	2385.915386	-0.129851504
ENSMUSG00000053510	Nrd1	4	109000655	109061777	E007	0.003643595	5.21E-05	0.018032832	321.8507542	0.187493152
ENSMUSG00000058440	Nrf1	6	30047988	30153458	E017	0.002438125	2.34E-07	0.00023366	609.338062	-0.160827633
ENSMUSG00000024109	Nrxn1	17	90033644	91092805	E001	0.003501401	1.45E-05	0.006676316	1215.29987	-0.129589777
ENSMUSG00000026434	Nucks1	1	131910458	131936321	E008	0.001924162	1.71E-05	0.007585737	5919.300709	-0.070749555
ENSMUSG00000037857	Nufip2	11	77686155	77712741	E004	0.002791417	3.71E-05	0.01390765	1549.877004	-0.09432701
ENSMUSG00000016619	Nup50	15	84923428	84942955	E008	0.001136208	2.39E-05	0.009897668	1943.251174	-0.06710519
ENSMUSG00000030934	Oat	7	132557475	132576398	E001	0.001583486	8.43E-05	0.024938422	1336.5499	-0.092227365
ENSMUSG00000031561	Odz3	8	48227682	48674690	E001	0.000988217	3.57E-05	0.01354792	2247.744992	-0.06935696
ENSMUSG00000034160	Ogt	X	101640060	101684351	E022	0.000813036	9.18E-05	0.026152786	3713.827557	-0.054956168
ENSMUSG00000022307	Oxr1	15	41447482	41861048	E020	0.002248623	8.05E-05	0.024188387	1125.45428	-0.099936935
ENSMUSG00000022283	Pabpc1	15	36595661	36608973	E001	0.000949728	5.74E-08	8.06E-05	2461.211047	-0.096400845
ENSMUSG00000020745	Pafah1b1	11	74673949	74724670	E001	0.001592305	1.08E-06	0.000859359	6315.317283	-0.080737651
ENSMUSG00000025451	Paip1	13	119428601	119458218	E012	0.001580158	1.11E-16	2.63E-12	963.0766676	-0.216726134
ENSMUSG00000037058	Paip2	18	35598617	35617186	E004	0.001321861	7.74E-05	0.023683339	2665.364681	-0.066538712
ENSMUSG00000018846	Pank3	11	35769484	35791285	E007	0.002582222	9.54E-05	0.027110089	3871.418914	-0.079377217
ENSMUSG00000036779	Papd5	8	88199213	88259722	E014	0.001771864	8.06E-05	0.024188387	812.6559454	-0.094033805
ENSMUSG00000021111	Papola	12	105784694	105838944	E001	0.003903304	2.77E-05	0.011254407	373.9738226	0.191239691
ENSMUSG00000020273	Papolg	11	23862646	23895253	E001	0.003416565	0.000100545	0.028365025	347.2919233	-0.159965425
ENSMUSG00000028032	Papss1	3	131564768	131643670	E012	0.001467078	3.37E-05	0.012965557	1081.114749	-0.101038139
ENSMUSG00000042323	Pbrm1	14	31019138	31121592	E033	0.002717905	4.55E-05	0.016255336	2029.339144	-0.111798179
ENSMUSG00000049100	Pcdh10	3	45378398	45434566	E007	0.00421502	1.80E-09	4.73E-06	627.6623916	-0.230117806
ENSMUSG00000051323	Pcdh19	X	133582863	133688993	E001	0.006097338	6.99E-05	0.021925498	1031.402196	-0.14076067
ENSMUSG00000023036	Pcdhga11	18	37674335	37841873	E005	0.005391691	6.33E-06	0.003607612	317.8668326	-0.233653111
ENSMUSG00000061601	Pclo	5	14514918	14863457	E026	0.013857092	1.51E-05	0.006853945	158.9970714	-0.353620411
ENSMUSG00000027342	Pcna	2	132249162	132253314	E001	0.001876953	1.24E-11	6.43E-08	1154.460825	-0.179906139
ENSMUSG00000027835	Pdcd10	3	75516490	75556856	E001	0.003633028	1.32E-07	0.000149122	322.951978	-0.230118489
ENSMUSG00000009030	Pdcl	2	37350074	37359332	E001	0.002896047	1.22E-06	0.000943609	1372.341673	-0.11297746

ENSMUSG00000026078	Pdcl3	1	38987814	38997236	E006	0.002362965	2.58E-05	0.010634713	583.859314	-0.12879555
ENSMUSG00000034021	Pds5b	5	150673739	150810690	E035	0.001441802	1.19E-05	0.005745632	1567.034396	-0.095757982
ENSMUSG00000020134	Peli1	11	21091291	21150323	E007	0.001371181	9.27E-07	0.000765177	1497.217139	-0.091772995
ENSMUSG00000020430	Pes1	11	3963975	3980004	E015	0.002374356	0.000149438	0.038552331	545.0900649	-0.12834201
ENSMUSG00000020003	Pex7	10	19859929	19907689	E001	0.006790392	2.82E-05	0.011332385	159.6974789	-0.251796009
ENSMUSG00000026409	Pfkfb2	1	130689043	130729253	E001	0.004076575	1.70E-05	0.007573421	282.3971366	-0.192691524
ENSMUSG00000006373	Pgrmc1	X	36598206	36606079	E003	0.00085298	9.02E-05	0.025785326	4854.25765	-0.042442567
ENSMUSG00000058318	Phf21a	2	92093117	92364666	E024	0.001279747	5.30E-08	7.52E-05	1385.627026	-0.111972295
ENSMUSG00000025626	Phf6	X	52912266	52956943	E011	0.000928418	6.63E-05	0.021482156	2597.809158	-0.056087363
ENSMUSG00000032253	Phip	9	82866159	82975489	E001	0.003142534	7.18E-06	0.003958169	3375.269259	-0.116997103
ENSMUSG00000032405	Pias1	9	62880077	62980879	E001	0.002113397	6.69E-05	0.021600649	635.7538513	-0.111771062
ENSMUSG00000025423	Pias2	18	77065208	77155708	E013	0.003621193	0.000185139	0.045230005	324.1942213	0.167158743
ENSMUSG00000041417	Pik3r1	13	101680761	101768217	E001	0.000982649	3.32E-06	0.002140939	2276.306268	-0.068686373
ENSMUSG00000024083	Pja2	17	64281005	64331916	E001	0.000847808	8.20E-06	0.004371377	3288.12483	-0.059924797
ENSMUSG00000028577	Plaa	4	94567514	94603244	E001	0.00191781	1.94E-06	0.001389368	1085.582147	-0.12278128
ENSMUSG00000002733	Plekha3	2	76675281	76696828	E008	0.003266995	1.02E-05	0.005087832	366.3713303	-0.16203612
ENSMUSG00000025758	Plk4	3	40800019	40816883	E016	0.008215443	0.000160833	0.040607677	130.0302433	-0.26878249
ENSMUSG00000032199	Polr2m	9	71478437	71485935	E001	0.001789402	2.67E-06	0.001800989	3839.295847	-0.089533241
ENSMUSG00000037519	Ppfia1	7	144476755	144553729	E001	0.002140339	3.88E-05	0.014335314	625.0209538	-0.131069256
ENSMUSG00000042133	Ppig	2	69722545	69754012	E014	0.001049013	6.43E-05	0.02095871	1976.902555	-0.061800767
ENSMUSG00000021096	Ppm1a	12	72761211	72794940	E006	0.001427435	3.21E-05	0.012531168	1129.653145	-0.091997098
ENSMUSG00000021096	Ppm1a	12	72761211	72794940	E001	0.003328837	8.01E-05	0.024146506	358.2341687	0.167322681
ENSMUSG00000061130	Ppm1b	17	84956741	85023991	E008	0.002613313	3.99E-05	0.01460729	482.131224	-0.14864457
ENSMUSG00000014956	Ppp1cb	5	32458974	32517433	E008	0.001210715	8.08E-11	3.44E-07	5290.402979	-0.097365374
ENSMUSG00000004455	Ppp1cc	5	122158278	122175273	E001	0.003483272	4.43E-05	0.015883026	339.4089244	0.178249404
ENSMUSG00000056612	Ppp1r14b	19	6975048	6977324	E001	0.00186016	0.000139368	0.03675453	2910.87436	0.078172193
ENSMUSG00000017843	Ppp2r5c	12	110485739	110583061	E015	0.001136918	2.81E-12	1.95E-08	1683.585545	0.136550935
ENSMUSG00000017843	Ppp2r5c	12	110485739	110583061	E013	0.00302884	1.42E-10	5.75E-07	401.4919405	-0.264256279

ENSMUSG00000021816	Ppp3cb	14	20499364	20546573	E001	0.000974427	2.36E-09	5.84E-06	2319.831635	-0.100370318
ENSMUSG00000026502	Pppde1	1	178187417	178252598	E005	0.002886033	3.39E-05	0.012978241	1477.181001	-0.087589371
ENSMUSG00000045302	Preb	5	30950853	30960330	E002	0.00582004	0.000140519	0.036882348	189.0710497	-0.224839693
ENSMUSG00000005034	Prkacb	3	146729579	146812960	E001	0.002738264	0.000141654	0.037062947	3839.700877	-0.087933613
ENSMUSG00000020612	Prkar1a	11	109649405	109669656	E011	0.002679588	5.07E-09	1.11E-05	6941.560309	-0.12838049
ENSMUSG00000002997	Prkar2b	12	31958476	32061296	E001	0.001722741	2.08E-08	3.67E-05	2112.621198	-0.115967087
ENSMUSG00000045038	Prkce	17	86167785	86657919	E015	0.001861856	4.59E-05	0.016386328	1030.957327	-0.09432405
ENSMUSG00000023110	Prmt5	14	54507187	54517525	E001	0.001692518	4.78E-06	0.002885473	868.820777	-0.125269729
ENSMUSG00000024735	Prpf19	19	10888156	10909559	E016	0.001580921	1.83E-05	0.007893049	962.3672923	-0.115499549
ENSMUSG00000021413	Prpf4b	13	34875494	34902877	E015	0.001430773	1.31E-08	2.49E-05	1125.399237	-0.132770889
ENSMUSG00000040225	Prrc2c	1	162673402	162740495	E001	0.003779171	6.81E-05	0.021830634	1782.508784	-0.134300906
ENSMUSG00000028484	Psip1	4	83455680	83486459	E001	0.002495776	1.81E-05	0.00783551	2189.376179	-0.112088741
ENSMUSG00000030751	Psma1	7	114264606	114276118	E008	0.002236302	7.98E-05	0.02412131	589.569074	-0.132812442
ENSMUSG00000060073	Psma3	12	70974621	70996347	E002	0.002317226	8.84E-06	0.004611513	562.6564224	-0.155110279
ENSMUSG00000060073	Psma3	12	70974621	70996347	E003	0.001525565	1.31E-05	0.006209736	1016.666491	-0.1125788
ENSMUSG00000060073	Psma3	12	70974621	70996347	E010	0.00170798	0.000179315	0.044132272	857.2754717	0.105466875
ENSMUSG00000021832	Psmc6	14	45329824	45349071	E014	0.004289126	6.96E-06	0.003885016	859.3718077	-0.179706778
ENSMUSG00000026229	Psmd1	1	86064387	86139151	E022	0.004423279	0	0	257.155404	-0.509322012
ENSMUSG00000026229	Psmd1	1	86064387	86139151	E021	0.003153528	1.59E-08	2.90E-05	382.3046078	-0.242349769
ENSMUSG00000026229	Psmd1	1	86064387	86139151	E023	0.002503024	5.64E-05	0.018897894	509.2804134	-0.148883754
ENSMUSG00000028134	Ptbp2	3	119718742	119783388	E001	0.002439598	2.22E-07	0.000228493	2650.602705	-0.128982099
ENSMUSG00000013663	Pten	19	32757497	32826160	E009	0.001367419	0.000111707	0.030935186	2903.574508	-0.063375252
ENSMUSG00000028788	Ptp4a2	4	129811219	129850003	E007	0.000752702	8.41E-07	0.000707864	4789.828261	-0.056131359
ENSMUSG00000028788	Ptp4a2	4	129811219	129850003	E001	0.003131728	6.41E-05	0.020929217	385.5258077	0.167649968
ENSMUSG00000028771	Ptpn12	5	20986645	21055911	E001	0.009456284	4.15E-05	0.015080722	507.935778	-0.242541422
ENSMUSG00000028399	Ptprd	4	75941238	78211961	E001	0.000967272	5.39E-11	2.35E-07	2359.091452	-0.106211063
ENSMUSG00000068748	Ptprz1	6	22875502	23052916	E030	0.011000351	9.43E-06	0.004811851	350.6108135	-0.295016112
ENSMUSG00000062078	Qk	17	10206471	10319361	E011	0.00355713	1.10E-07	0.000127928	635.0380249	0.205675689

ENSMUSG00000062078	Qk	17	10206471	10319361	E001	0.003351724	2.09E-06	0.001473941	4047.696631	-0.125272318
ENSMUSG00000062078	Qk	17	10206471	10319361	E002	0.003007359	1.54E-05	0.006967808	1575.801747	-0.130506337
ENSMUSG00000074994	Qser1	2	104754795	104816696	E001	0.001311234	6.84E-05	0.021830634	1300.847221	-0.077831613
ENSMUSG00000020671	Rab10	12	3247430	3309969	E001	0.003365256	1.76E-05	0.00771837	4553.638427	-0.109322033
ENSMUSG00000047187	Rab2a	4	8535644	8607778	E008	0.001458739	3.61E-05	0.013604597	3965.934931	-0.072553717
ENSMUSG00000030704	Rab6a	7	100607410	100641268	E009	0.002168758	7.19E-07	0.000616074	3960.727527	-0.101854441
ENSMUSG00000030704	Rab6a	7	100607410	100641268	E004	0.009928919	5.22E-05	0.018038459	106.2884786	0.326856236
ENSMUSG00000036943	Rab8b	9	66843672	66919705	E001	0.00256178	1.24E-05	0.005953431	1645.497416	-0.092000352
ENSMUSG00000026721	Rabgap1l	1	160219174	160792938	E001	0.002400483	1.91E-06	0.001376761	1244.239944	-0.113925663
ENSMUSG00000015087	RabI6 (B230208H17Rik)	2	25583018	25608521	E002	0.007305906	1.02E-06	0.000822592	147.5216439	-0.333643393
ENSMUSG00000022314	Rad21	15	51962605	51991760	E001	0.002702947	1.97E-07	0.000206369	3690.85683	-0.129415993
ENSMUSG00000028426	Rad23b	4	55350043	55392237	E010	0.001647535	9.88E-07	0.000807424	3429.874024	-0.09510723
ENSMUSG00000068798	Rap1a	3	105727267	105801330	E001	0.005253841	3.32E-07	0.000310934	654.4655159	-0.213557653
ENSMUSG00000050029	Rap2c	X	51003912	51018018	E001	0.00457824	1.99E-05	0.008457586	1980.527499	-0.130840575
ENSMUSG00000021549	Rasa1	13	85214780	85289029	E001	0.001808306	8.52E-10	2.44E-06	789.2234338	-0.176935436
ENSMUSG00000025907	Rb1cc1	1	6206197	6276648	E024	0.002455683	7.56E-07	0.000639512	521.8952696	-0.172567162
ENSMUSG00000027641	Rbl1	2	157145893	157204534	E001	0.006619889	0.000181816	0.044615368	402.9092671	-0.188864853
ENSMUSG00000022119	Rbm26	14	105106751	105177327	E001	0.003626719	1.18E-11	6.43E-08	323.6130236	-0.305691184
ENSMUSG00000024491	Rbm27	18	42275353	42341542	E020	0.001794559	7.55E-07	0.000639512	983.21284	-0.122677072
ENSMUSG00000031167	Rbm3	X	8138975	8145880	E002	0.003176437	5.11E-06	0.003027359	2066.586059	-0.134540897
ENSMUSG00000042396	Rbm7	9	48488709	48495330	E001	0.001566025	4.30E-05	0.015517854	976.4001372	-0.085399189
ENSMUSG00000038374	Rbm8a	3	96629928	96633779	E007	0.002265567	0.000110127	0.030621508	1194.561151	-0.10194396
ENSMUSG00000026970	Rbms1	2	60750193	60963192	E001	0.002006922	3.13E-05	0.012257881	682.0398633	-0.122851307
ENSMUSG00000032320	Rcn2	9	56041845	56061882	E007	0.001093215	1.08E-09	2.95E-06	1817.665643	-0.111058347
ENSMUSG00000032050	Rdx	9	52047150	52088737	E015	0.001785563	8.46E-05	0.024984101	803.685751	0.114392333
ENSMUSG00000032050	Rdx	9	52047150	52088737	E014	0.002705998	0.000160232	0.040517587	461.4577978	0.145638803
ENSMUSG00000039852	Rere	4	150281646	150621966	E023	0.002711983	1.14E-06	0.000882692	1699.864361	-0.134304588
ENSMUSG00000024712	Rfk	19	17394043	17401349	E004	0.001807196	4.13E-05	0.015065564	1385.104415	-0.078186978

ENSMUSG00000036202	Rif1	2	52072837	52122383	E036	0.003236639	3.23E-10	1.12E-06	370.5024407	-0.2608765
ENSMUSG00000028557	Rnf11	4	109451098	109484728	E001	0.003282452	1.29E-07	0.00014658	2510.724015	-0.131957169
ENSMUSG00000028557	Rnf11	4	109451098	109484728	E003	0.002076225	6.60E-06	0.003730812	651.1820643	0.124038028
ENSMUSG00000032217	Rnf111	9	70425429	70503725	E001	0.002056044	9.98E-09	2.02E-05	659.8758714	-0.166909882
ENSMUSG00000070426	Rnf121	7	102019137	102065469	E001	0.005854748	0.00014196	0.037084377	392.4433103	0.174116167
ENSMUSG00000024317	Rnf138	18	21001341	21028223	E008	0.004078862	7.07E-05	0.022058305	783.5548822	-0.135700051
ENSMUSG00000048234	Rnf149	1	39551296	39577405	E001	0.002458657	6.91E-07	0.000599984	521.0845006	-0.152700192
ENSMUSG00000026484	Rnf2	1	151469406	151500823	E001	0.004561516	5.50E-05	0.018766295	2356.655313	-0.132575216
ENSMUSG00000028309	Rnf20	4	49632006	49656887	E022	0.001848727	1.79E-05	0.007775314	764.7647732	-0.123873642
ENSMUSG00000035696	Rnf38	4	44126210	44233789	E001	0.002953525	3.37E-06	0.002159658	2969.029162	-0.1149228
ENSMUSG00000027981	Rnpc3	3	113605067	113630149	E002	0.008569868	5.42E-05	0.018577719	371.8842156	-0.240924041
ENSMUSG00000022883	Robo1	16	72663149	73046095	E029	0.001145384	3.01E-06	0.001972873	1659.865645	-0.093456359
ENSMUSG00000052516	Robo2	16	73892306	74411825	E001	0.002193081	9.20E-08	0.00011306	1360.904292	-0.14194056
ENSMUSG00000020580	Rock2	12	16894978	16988274	E033	0.001543927	1.00E-05	0.005038129	997.988201	-0.107924113
ENSMUSG00000028936	Rpl22	4	152325742	152334071	E004	0.001067263	6.87E-05	0.02184577	5321.72405	-0.051449522
ENSMUSG00000008683	Rps15a	7	118104378	118116186	E001	0.001239411	6.38E-05	0.020885863	2131.076792	-0.075357562
ENSMUSG00000025665	Rps6ka6	X	111388192	111537959	E001	0.001756731	1.42E-10	5.75E-07	822.8006182	-0.173488875
ENSMUSG00000030978	Rrm1	7	102441695	102469771	E013	0.002393632	0.000158331	0.040159416	539.4080591	0.134723556
ENSMUSG00000027304	Rtf1	2	119675068	119735407	E018	0.00252542	5.53E-05	0.018784507	1266.969001	-0.11487636
ENSMUSG00000021087	Rtn1	12	72211752	72409054	E001	0.000753892	1.07E-05	0.005259867	4762.610117	-0.053053886
ENSMUSG00000020458	Rtn4	11	29692947	29744331	E001	0.003326712	4.68E-08	6.92E-05	1762.529962	0.167037652
ENSMUSG00000020458	Rtn4	11	29692947	29744331	E008	0.001792578	0.000155429	0.03954423	799.1689945	-0.108186003
ENSMUSG00000006586	Runx1t1	4	13743436	13893649	E014	0.001043338	4.08E-06	0.002506016	1999.390401	-0.066668831
ENSMUSG00000031568	Rwdd4a	8	47533664	47552825	E008	0.002343685	5.64E-05	0.018897894	774.2956902	-0.107668842
ENSMUSG00000032547	Ryk	9	102834917	102908305	E016	0.001381818	6.82E-05	0.021830634	1191.194394	-0.092559682
ENSMUSG00000006763	Saal1	7	46686108	46710680	E001	0.00377981	8.55E-08	0.000105842	308.3001908	-0.234482745
ENSMUSG00000025240	Sacm1l	9	123529882	123592598	E020	0.005336513	4.54E-07	0.000407154	971.8617934	-0.207760165
ENSMUSG00000015305	Sash1	10	8722219	8886070	E001	0.001530296	7.93E-09	1.65E-05	1011.787524	-0.121796729

ENSMUSG00000000037	Scml2	X	161117193	161258213	E026	0.008161269	0.0001308	0.035280238	130.9550713	-0.247541956
ENSMUSG00000057182	Scn3a	2	65457118	65567627	E001	0.001633692	1.48E-05	0.006771308	915.7432938	-0.105007811
ENSMUSG00000022261	Sdc2	15	32920723	33034730	E005	0.002563066	2.27E-05	0.009453318	1353.972092	-0.104322437
ENSMUSG00000020986	Sec23a	12	58958383	59012017	E001	0.002270672	1.03E-07	0.000124116	1418.92877	-0.146986527
ENSMUSG00000027429	Sec23b	2	144556229	144590749	E020	0.004299217	0.000123929	0.03375626	265.6521574	-0.190690213
ENSMUSG00000036391	Sec24a	11	51692264	51763634	E001	0.002791417	2.88E-07	0.000277624	467.5442832	-0.173931736
ENSMUSG00000039234	Sec24d	3	123267496	123365635	E023	0.002579333	8.76E-07	0.000733965	490.182207	-0.178436021
ENSMUSG00000035325	Sec31a	5	100361650	100416234	E007	0.003336253	0.000173982	0.043204232	503.1741792	0.15049188
ENSMUSG00000079614	Seh1l	18	67774876	67795461	E010	0.001836459	0.000117473	0.032316043	1673.754007	-0.084033663
ENSMUSG00000075700	Selt (2810407C02Rik)	3	58576636	58593133	E006	0.001545612	3.67E-09	8.96E-06	2492.075934	-0.101419731
ENSMUSG00000058013	Sept11	5	93093457	93174958	E010	0.000675305	0.000106527	0.029783509	7623.006174	-0.037318803
ENSMUSG00000001833	Sept7	9	25252439	25308571	E014	0.00201456	7.88E-15	1.45E-10	2666.447681	-0.180479407
ENSMUSG00000041571	Sepw1	7	15917208	15922402	E001	0.000950326	7.89E-06	0.004250244	2457.587003	0.072973434
ENSMUSG00000036371	Serbp1	6	67266979	67289302	E010	0.003970177	6.46E-12	4.12E-08	3632.078383	-0.208049238
ENSMUSG00000054766	Set	2	30057378	30072577	E008	0.001371823	2.58E-06	0.001760283	3500.098869	-0.083835853
ENSMUSG00000024949	Sf1	19	6363690	6378030	E014	0.000971295	6.99E-05	0.021925498	2336.852905	-0.06871795
ENSMUSG00000002129	Sf3a1	11	4160350	4182541	E016	0.001340019	2.08E-05	0.008797931	1253.778676	-0.096816686
ENSMUSG00000006527	Sfmbt1	14	30714849	30822713	E024	0.00223779	9.81E-07	0.000805993	589.0510009	-0.145711654
ENSMUSG00000028820	Sfpq	4	127021324	127037013	E010	0.000978019	0.000133353	0.035621403	4677.64856	-0.05502141
ENSMUSG00000027996	Sfrp2	3	83766321	83774316	E003	0.0007275	7.01E-07	0.000605425	5449.288066	-0.049542823
ENSMUSG00000027996	Sfrp2	3	83766321	83774316	E001	0.000849963	0.000139017	0.03675453	3642.013064	0.048375671
ENSMUSG00000028248	Sfrs18	4	21847583	21876475	E012	0.004565591	3.58E-06	0.00225195	1048.652384	-0.168011338
ENSMUSG00000026039	Sgol2	1	57995971	58025899	E010	0.005217183	4.61E-06	0.002823416	221.0554557	-0.23534738
ENSMUSG00000024976	Shoc2	19	53944306	54033268	E010	0.001297449	2.23E-09	5.68E-06	1324.66101	-0.119023169
ENSMUSG00000090112	Shprh	10	11149427	11217595	E031	0.00274069	1.49E-08	2.77E-05	454.1685968	-0.209640827
ENSMUSG00000020063	Sirt1	10	63319005	63381704	E001	0.002230357	7.68E-06	0.004177898	591.6483473	-0.12397639
ENSMUSG00000055717	Slain1	14	103650228	103704907	E007	0.003407255	0.000123669	0.033740821	348.4213667	-0.151280502
ENSMUSG00000029221	Slc30a9	5	67306955	67358565	E018	0.001396257	1.48E-05	0.006771308	1171.001343	-0.099774623

ENSMUSG00000022462	Slc38a2	15	96687394	96699698	E001	0.001021522	3.63E-07	0.000336847	4703.792169	-0.070493703
ENSMUSG00000025986	Slc39a10	1	46807544	46854046	E001	0.002812912	0.000193606	0.046884395	2174.947692	-0.096626475
ENSMUSG00000024270	Slc39a6	18	24579881	24603817	E001	0.001352986	3.57E-05	0.01354792	1233.670964	-0.089945034
ENSMUSG00000025790	Slco3a1	7	74275419	74554780	E002	0.003571208	0.000203602	0.048595463	329.5480783	-0.144263001
ENSMUSG00000025060	Slk	19	47579678	47645246	E019	0.003445617	1.11E-05	0.005427428	1117.155804	-0.138042741
ENSMUSG00000031681	Smad1	8	79338395	79399518	E001	0.001742843	7.73E-06	0.004190416	866.039399	-0.11065057
ENSMUSG00000024515	Smad4	18	73639009	73703780	E001	0.002194441	1.05E-06	0.000846947	1581.01388	-0.123002253
ENSMUSG00000021540	Smad5	13	56703010	56742377	E010	0.00166409	2.94E-06	0.001949364	4392.179445	-0.070989263
ENSMUSG00000021540	Smad5	13	56703010	56742377	E006	0.003380875	7.65E-06	0.00417256	351.661858	0.188251493
ENSMUSG00000031099	Smarca1	X	47809368	47892974	E001	0.007505689	7.39E-08	9.57E-05	252.4156098	-0.324650374
ENSMUSG00000031715	Smarca5	8	80698507	80739497	E001	0.007995412	1.63E-11	7.93E-08	922.3111729	-0.351027746
ENSMUSG00000025369	Smarcc2	10	128459236	128490173	E029	0.005105743	1.36E-05	0.006345244	218.6795845	0.248953594
ENSMUSG00000028312	Smc2	4	52439243	52488260	E025	0.001432735	2.07E-10	7.98E-07	1122.913058	-0.150184762
ENSMUSG00000033419	Snap91	9	86765936	86880654	E001	0.004742797	3.18E-05	0.012446532	780.5663787	-0.167854523
ENSMUSG00000063511	Snrnp70	7	45376462	45395692	E001	0.001957442	3.94E-05	0.014498365	2188.203734	-0.095574923
ENSMUSG00000000948	Snrpn	7	59982495	60140219	E003	0.001582367	5.32E-05	0.018256978	961.027311	0.107180449
ENSMUSG00000006818	Sod2	17	13007839	13018119	E005	0.005525363	0.000189521	0.046029344	1245.152811	-0.146222939
ENSMUSG00000000567	Sox9	11	112782224	112787760	E003	0.00112554	6.54E-05	0.021261123	1716.549926	-0.051770214
ENSMUSG00000001280	Sp1	15	102406143	102436404	E006	0.001766381	3.60E-05	0.013604597	2892.690532	-0.061608216
ENSMUSG00000027109	Sp3	2	72936427	72980446	E001	0.003844359	5.26E-06	0.00310452	1752.5809	-0.144301499
ENSMUSG00000025323	Sp4	12	118234933	118301440	E001	0.01363206	0.000148258	0.038307452	810.444658	-0.231582239
ENSMUSG00000020315	Spnb2	11	30099395	30268175	E001	0.000984597	3.03E-05	0.011964291	2266.228972	-0.075085051
ENSMUSG00000020315	Spnb2	11	30099395	30268175	E005	0.012843911	6.84E-05	0.021830634	308.7005633	-0.29287129
ENSMUSG00000054162	Spock3	8	62951009	63357103	E012	0.003000215	6.81E-08	9.21E-05	406.1718591	-0.188235546
ENSMUSG00000027351	Spred1	2	117121374	117182279	E001	0.004945227	1.25E-06	0.000962435	226.6557921	0.259298633
ENSMUSG00000022351	Sqle	15	59315077	59331192	E009	0.004121504	8.68E-05	0.025361541	278.8501242	-0.192806034
ENSMUSG00000039218	Srrm2	17	23803187	23824739	E014	0.001886726	1.37E-05	0.006366122	743.1144008	-0.131127403
ENSMUSG00000039218	Srrm2	17	23803187	23824739	E011	0.001483795	8.53E-05	0.025125983	1061.875298	0.099506231

ENSMUSG00000028676	Srsf10	4	135855747	135869908	E006	0.000870379	7.70E-13	6.09E-09	3060.41433	-0.096117056
ENSMUSG00000028676	Srsf10	4	135855747	135869908	E004	0.00166751	9.52E-06	0.004827969	888.1679654	0.118029179
ENSMUSG00000055436	Srsf11	3	158010473	158036639	E003	0.00451001	1.66E-06	0.001210888	538.5339073	-0.207288357
ENSMUSG00000034120	Srsf2	11	116849901	116853094	E001	0.003010244	4.70E-10	1.50E-06	5280.791087	-0.144393409
ENSMUSG00000071172	Srsf3	17	29032673	29043366	E003	0.001167578	0.000199516	0.04782672	2078.530019	0.070484049
ENSMUSG00000016921	Srsf6	2	162931528	162937121	E006	0.000708299	1.78E-05	0.007775314	6087.90747	-0.044265387
ENSMUSG00000025862	Stag2	X	42149317	42277185	E034	0.002829487	1.87E-05	0.007993664	1429.510716	-0.129925554
ENSMUSG00000026718	Stam	2	14074100	14148338	E014	0.001510936	1.07E-05	0.005243431	1032.054668	-0.104125237
ENSMUSG00000028832	Stmn1	4	134468320	134473843	E004	0.000658861	0.000108125	0.030144573	8718.674701	0.037324362
ENSMUSG00000030224	Strap	6	137735078	137751932	E010	0.00397849	1.28E-05	0.006097251	2125.724967	-0.14368877
ENSMUSG00000020954	Strn3	12	51608541	51691914	E001	0.003333859	1.08E-06	0.000859359	1427.354402	-0.154738664
ENSMUSG00000020954	Strn3	12	51608541	51691914	E018	0.003094738	8.07E-06	0.004330588	391.1175978	0.184680065
ENSMUSG00000022110	Sucla2	14	73525319	73596142	E011	0.002779158	5.32E-06	0.00312684	446.3504597	-0.172955102
ENSMUSG00000032423	Syncrip	9	88449729	88482574	E001	0.002076033	8.77E-05	0.025396654	4165.86085	-0.082068808
ENSMUSG00000024261	Syt4	18	31437808	31447415	E001	0.003324938	1.06E-07	0.000124116	2305.483744	-0.13764541
ENSMUSG00000017291	Taok1	11	77529162	77607815	E001	0.003006478	7.47E-05	0.022976844	2974.745803	-0.093657197
ENSMUSG00000046985	Tapt1	5	44175164	44226606	E009	0.013224122	7.68E-05	0.023557587	78.66613239	-0.37292904
ENSMUSG00000041459	Tardbp	4	148612382	148627019	E002	0.008630385	0.000111505	0.030931007	648.3311212	-0.214478886
ENSMUSG00000020130	Tbc1d15	10	115197877	115251493	E001	0.008113375	0.000133289	0.035621403	451.1051062	-0.209609895
ENSMUSG00000071359	Tbpl1	10	22703879	22731938	E001	0.001706175	1.63E-06	0.001198509	1366.320246	-0.102541456
ENSMUSG00000033813	Tcea1	1	4857814	4897909	E012	0.000932116	4.12E-08	6.21E-05	2573.026977	-0.086857573
ENSMUSG00000051579	Tceal8	X	136168984	136172342	E001	0.005110603	1.74E-05	0.00767987	1974.491594	-0.13327909
ENSMUSG00000032228	Tcf12	9	71844252	72111819	E001	0.001667778	6.33E-07	0.000555939	3831.652151	-0.096282928
ENSMUSG00000053477	Tcf4	18	69344146	69687967	E022	0.002715179	2.53E-08	4.24E-05	5378.149471	-0.124224153
ENSMUSG00000024985	Tcf7l2	19	55741810	55933654	E019	0.005883714	1.86E-05	0.007962422	2314.257612	-0.164676372
ENSMUSG00000007613	Tgfbr1	4	47353222	47414931	E009	0.000832	5.29E-05	0.018217158	3468.895258	-0.044150117
ENSMUSG00000023885	Thbs2	17	14665500	14694262	E001	0.001394829	1.99E-06	0.001415321	1172.968517	-0.108468188
ENSMUSG00000037475	Thoc2	X	41794991	41920674	E001	0.005157067	5.24E-05	0.018058823	1203.016009	-0.15894415

ENSMUSG00000043962	Thrap3	4	126164082	126202760	E001	0.000922772	0.000132787	0.035584704	2636.579056	-0.060633578
ENSMUSG00000071337	Tia1	6	86404219	86433403	E014	0.000839088	4.03E-06	0.002493846	3385.448901	-0.060177978
ENSMUSG00000030846	Tial1	7	128439777	128461717	E001	0.002130055	1.60E-10	6.32E-07	1820.167672	-0.154621698
ENSMUSG00000030516	Tjp1	7	65296165	65371239	E001	0.001542992	3.86E-10	1.26E-06	998.9233176	-0.158506967
ENSMUSG00000025016	Tm9sf3	19	41210842	41264004	E014	0.001940374	5.87E-05	0.019393507	714.5548937	0.124658566
ENSMUSG00000025016	Tm9sf3	19	41210842	41264004	E011	0.003011669	0.000151396	0.038875945	404.2861752	-0.157444025
ENSMUSG00000038497	Tmco3	8	13288013	13322924	E013	0.004580743	0.000139672	0.036776389	247.1232445	-0.191616096
ENSMUSG00000063406	Tmed5	5	108106366	108132620	E001	0.001894706	1.28E-06	0.00098141	738.7229987	-0.089920802
ENSMUSG00000033184	Tmed7	18	46560235	46597535	E001	0.006104991	0.000136795	0.036248951	3503.758169	-0.128916066
ENSMUSG00000028347	Tmeff1	4	48585174	48663131	E010	0.004492492	2.37E-07	0.00023366	2397.303752	-0.176423442
ENSMUSG00000029571	Tmem106b	6	13069759	13089269	E009	0.002355925	8.40E-06	0.004435546	2276.555705	-0.093192926
ENSMUSG00000050912	Tmem123	9	7764041	7794333	E005	0.003368796	7.07E-05	0.022058305	1966.176778	-0.103464958
ENSMUSG00000038079	Tmem237	1	59100594	59120096	E002	0.011032782	9.46E-06	0.004811851	165.7987947	-0.325928046
ENSMUSG00000032328	Tmem30a	9	79768943	79793507	E001	0.001221181	3.64E-06	0.002270083	2550.705384	-0.072190371
ENSMUSG00000037720	Tmem33	5	67260565	67291461	E010	0.002841719	6.60E-08	9.05E-05	1795.183074	-0.146779274
ENSMUSG00000028221	Tmem55a	4	14864076	14915176	E007	0.002563998	6.95E-10	2.06E-06	1431.720428	-0.160423398
ENSMUSG00000028826	Tmem57	4	134802759	134853345	E001	0.001319775	2.39E-06	0.001641829	1286.516649	-0.099918081
ENSMUSG00000062373	Tmem65	15	58782269	58823530	E001	0.001089852	9.36E-06	0.004811851	2032.200725	-0.061480411
ENSMUSG00000030059	Tmf1	6	97155266	97179124	E001	0.003740946	4.64E-13	3.85E-09	312.0486507	-0.324541357
ENSMUSG00000058587	Tmod3	9	75497796	75559657	E001	0.001379797	5.72E-05	0.019101954	1784.846397	-0.076936522
ENSMUSG00000019961	Tmpo	10	91147571	91171582	E001	0.003481679	1.44E-06	0.00107744	5076.476054	-0.126150989
ENSMUSG00000041594	Tmtc4	14	122918971	122984035	E001	0.002973803	6.11E-05	0.020097024	410.5878187	-0.153121157
ENSMUSG00000024811	Tnks2	19	36834232	36893477	E027	0.000779226	1.51E-06	0.001118263	4248.665422	-0.064137473
ENSMUSG00000025571	Tnrc6c	11	117654289	117763439	E022	0.001957491	4.92E-07	0.000438518	1310.600042	-0.12644221
ENSMUSG00000022752	Tomm70a	16	57121714	57154530	E001	0.002357783	6.90E-05	0.0218746	550.0718613	0.139765713
ENSMUSG00000070544	Top1	2	160645888	160722764	E021	0.000958331	2.30E-07	0.000232646	2410.051197	-0.088831046
ENSMUSG00000017485	Top2b	14	16365179	16430787	E036	0.001725422	1.75E-08	3.15E-05	996.3900095	-0.158684609
ENSMUSG00000017485	Top2b	14	16365179	16430787	E031	0.001948939	1.31E-05	0.00621812	710.1973371	-0.138591694

ENSMUSG00000032555	Topbp1	9	103305215	103350432	E011	0.004361786	0.000139163	0.03675453	261.2978765	-0.194903319
ENSMUSG00000041272	Tox	4	6686353	6991557	E001	0.003202472	5.58E-05	0.018823018	375.2649707	-0.1530937
ENSMUSG00000060126	Tpt1	14	75845093	75848525	E005	0.001063589	9.18E-06	0.004745094	1921.395449	-0.085139514
ENSMUSG00000029817	Tra2a	6	49243921	49264052	E001	0.001338638	1.96E-07	0.000206369	1971.815668	-0.099191051
ENSMUSG00000022858	Tra2b	16	22244549	22266005	E001	0.002927543	7.16E-08	9.57E-05	3210.955988	-0.141359585
ENSMUSG00000021711	Trappc13 (2410002O22Rik)	13	104142149	104178469	E001	0.001762085	4.30E-05	0.015517854	819.1827208	-0.106758827
ENSMUSG00000029833	Trim24	6	37870811	37966296	E019	0.00216743	0.000175997	0.043444568	614.5880351	-0.121053314
ENSMUSG00000033014	Trim33	3	103279293	103358768	E020	0.001109117	1.71E-05	0.007585737	1766.476498	-0.075673939
ENSMUSG00000018548	Trim37	11	87127077	87220683	E024	0.001524964	2.03E-08	3.61E-05	1432.950801	-0.12944546
ENSMUSG00000038324	Trpc4ap	2	155634271	155692384	E015	0.00322198	2.99E-05	0.011929663	372.5308007	-0.179938571
ENSMUSG00000038324	Trpc4ap	2	155634271	155692384	E016	0.004047885	8.67E-05	0.025361541	284.7097032	-0.193475031
ENSMUSG00000027806	Tsc22d2	3	58415689	58466787	E004	0.003885934	4.20E-05	0.015230769	1064.082209	-0.11383565
ENSMUSG00000027806	Tsc22d2	3	58415689	58466787	E001	0.002659484	8.70E-05	0.025365043	533.32616	0.106728217
ENSMUSG00000040785	Ttc3	16	94370618	94469222	E013	0.005644082	0.000153835	0.039259135	195.5947775	-0.22880146
ENSMUSG00000072235	Tuba1a	15	98949841	98953551	E001	0.01391227	3.92E-05	0.014457358	13676.62664	0.166425293
ENSMUSG00000023004	Tuba1b	15	98931425	98934565	E002	0.003492389	3.85E-05	0.014269508	804.2451196	-0.139388457
ENSMUSG00000034377	Tulp4	17	6106437	6251128	E015	0.002972596	0.000174642	0.043277881	2786.662303	-0.093266334
ENSMUSG00000052997	Uba2	7	34140689	34169599	E001	0.001217072	3.43E-06	0.002181056	1482.955746	-0.099696124
ENSMUSG00000042520	Ubap2l	3	89999589	90052475	E007	0.004308377	3.09E-05	0.012166976	265.0056152	-0.212979762
ENSMUSG00000091896	Ube2d2	18	35771559	35807172	E007	0.000775584	9.72E-12	5.56E-08	4315.624147	-0.082229418
ENSMUSG00000091896	Ube2d2	18	35771559	35807172	E001	0.001265941	1.18E-07	0.000135152	1382.510946	0.115217396
ENSMUSG00000091896	Ube2d2	18	35771559	35807172	E002	0.002006535	2.67E-05	0.010895752	682.2206115	0.131759153
ENSMUSG00000078578	Ube2d3	3	135438297	135467178	E008	0.001019602	4.11E-08	6.21E-05	6028.935189	-0.073056547
ENSMUSG00000029203	Ube2k	5	65537233	65598988	E007	0.001661295	1.04E-07	0.000124116	2277.913889	-0.106379341
ENSMUSG00000036241	Ube2r2	4	41135743	41193380	E005	0.000919903	1.92E-07	0.000204647	2656.726857	-0.076251435
ENSMUSG00000036241	Ube2r2	4	41135743	41193380	E001	0.00193934	2.72E-05	0.011078242	1182.346747	0.110485116
ENSMUSG00000025326	Ube3a	7	59228752	59306727	E015	0.001316539	1.63E-05	0.007312331	1291.909111	-0.088980057
ENSMUSG00000059890	Ube4a	9	44923131	44965600	E001	0.001304089	4.76E-08	6.93E-05	1313.081824	-0.113415368

ENSMUSG00000001687	Ubl3	5	148504631	148552789	E001	0.002024391	1.18E-07	0.000135152	1171.978685	-0.120865187
ENSMUSG00000005312	Ubqln1	13	58176156	58215653	E001	0.002542154	2.67E-06	0.001800989	2563.235477	-0.119641672
ENSMUSG00000044308	Ubr3	2	69897246	70024013	E041	0.001657349	1.10E-06	0.000873379	896.2769493	-0.130304264
ENSMUSG00000029201	Ugdh	5	65413202	65435946	E001	0.003405541	2.08E-07	0.000215942	877.3909036	-0.179982713
ENSMUSG00000019951	Uhrf1bp1l	10	89744991	89819869	E021	0.002424904	7.41E-05	0.022841086	530.4376628	-0.130037981
ENSMUSG00000004798	Ulk2	11	61775649	61855073	E001	0.002791417	1.79E-07	0.000193586	1443.454498	-0.156401712
ENSMUSG00000036572	Upf3b	X	37091678	37110322	E004	0.008899548	0.000124634	0.033892754	119.3835123	-0.282572083
ENSMUSG00000019820	Utrn	10	12382188	12861735	E001	0.00197385	1.13E-06	0.000876698	697.8203622	-0.146257647
ENSMUSG00000024091	Vapa	17	65580056	65613555	E001	0.001104615	6.57E-08	9.05E-05	1780.674833	-0.097418717
ENSMUSG00000021614	Vcan	13	89655312	89742509	E001	0.001009365	4.75E-09	1.05E-05	4384.407101	-0.085114628
ENSMUSG00000029462	Vps29	5	122354369	122364984	E004	0.002152902	6.59E-05	0.021405455	620.1392761	-0.112857759
ENSMUSG00000021115	Vrk1	12	106010263	106077410	E015	0.003809455	8.81E-06	0.004607821	478.2722021	-0.178226244
ENSMUSG00000024283	Wac	18	7868832	7973547	E014	0.001611575	3.97E-05	0.014586248	1846.086969	-0.090246053
ENSMUSG00000041408	Wapal	14	34673928	34747983	E019	0.003481859	5.81E-06	0.003381953	412.2056218	-0.189049925
ENSMUSG00000019831	Wasf1	10	40883475	40938570	E009	0.002928886	4.21E-05	0.015232468	676.5070354	-0.135466522
ENSMUSG00000029684	Wasl	6	24613805	24665009	E001	0.00127417	2.11E-06	0.001477896	1366.919971	-0.085768683
ENSMUSG00000030216	Wbp11	6	136813654	136828233	E002	0.0039672	0.000102959	0.028947622	445.1940334	0.167294274
ENSMUSG00000038733	Wdr26	1	181173228	181211552	E001	0.002560072	8.37E-05	0.024778705	3602.089582	-0.092345442
ENSMUSG00000038733	Wdr26	1	181173228	181211552	E014	0.002048502	0.000197648	0.047569962	663.184651	0.117117995
ENSMUSG00000036769	Wdr44	X	23693051	23806025	E020	0.010668544	0.000132396	0.035541494	98.52346805	-0.304542637
ENSMUSG00000017677	Wsb1	11	79239372	79254671	E006	0.001379471	0.0001947	0.04708049	1194.541389	-0.087194845
ENSMUSG00000031930	Wwp2	8	107436398	107558594	E024	0.00281167	7.37E-05	0.022816885	1231.816845	-0.120207018
ENSMUSG00000025860	Xiap	X	42059679	42109656	E008	0.001586575	1.13E-09	3.01E-06	2655.406334	-0.094842975
ENSMUSG00000025860	Xiap	X	42059679	42109656	E003	0.003134244	0.0001748	0.043277881	468.5818449	0.140501543
ENSMUSG00000020290	Xpo1	11	23256041	23298249	E027	0.001047849	4.75E-09	1.05E-05	1981.474213	-0.112681922
ENSMUSG00000053110	Yap1	9	7931999	8004596	E001	0.000830537	5.58E-09	1.19E-05	3486.63181	-0.071142276
ENSMUSG00000028639	Ybx1	4	119277981	119294604	E001	0.002027068	8.26E-08	0.000103151	4007.856889	-0.116978512
ENSMUSG00000026775	Yme1l1	2	23156369	23199260	E019	0.001446609	6.81E-05	0.021830634	1463.918541	-0.090432265

ENSMUSG00000035851	Ythdc1	5	86804221	86836659	E016	0.003150038	0.000131016	0.035281103	382.8166922	0.157378411
ENSMUSG00000034653	Ythdc2	18	44828665	44889720	E030	0.003997914	7.25E-08	9.57E-05	288.829388	-0.244331606
ENSMUSG00000038848	Ythdf1	2	180904377	180920949	E001	0.003041437	2.74E-08	4.38E-05	1194.657684	-0.153985893
ENSMUSG00000040025	Ythdf2	4	132184912	132212303	E001	0.004087184	7.40E-05	0.022841086	815.6909177	-0.142519905
ENSMUSG00000047213	Ythdf3	3	16183183	16217035	E006	0.002759626	2.75E-06	0.001848042	2205.832316	-0.106150923
ENSMUSG00000018326	Ywhab	2	163994960	164018588	E006	0.000682494	1.43E-06	0.00107744	7225.991863	-0.0478627
ENSMUSG00000021264	Yy1	12	108792973	108816632	E005	0.003473359	1.64E-07	0.000180698	1372.311051	-0.158838233
ENSMUSG00000033964	Zbtb41	1	139422383	139453005	E011	0.001496449	0.000118916	0.032604939	1581.371722	-0.06751492
ENSMUSG00000026872	Zeb2	2	44983634	45117395	E001	0.006594349	3.61E-05	0.013604597	1916.985707	-0.167663223
ENSMUSG00000024750	Zfand5	19	21272278	21282289	E006	0.000918409	3.49E-10	1.16E-06	2667.33998	-0.088002067
ENSMUSG00000024750	Zfand5	19	21272278	21282289	E001	0.002222385	3.88E-07	0.000352083	675.0388299	0.161559311
ENSMUSG00000030629	Zfand6	7	84615054	84679361	E001	0.001315773	4.87E-05	0.0170416	1364.446965	-0.079318332
ENSMUSG00000038872	Zfhx3	8	108714644	108961636	E010	0.003289501	0.000185482	0.045247306	1321.75904	-0.094349261
ENSMUSG00000017421	Zfp207	11	80383279	80405733	E011	0.001219664	1.32E-05	0.006253233	1477.26489	0.093418859
ENSMUSG00000017421	Zfp207	11	80383279	80405733	E013	0.001218713	8.22E-05	0.024419234	1981.105954	-0.077624593
ENSMUSG00000036916	Zfp280c	X	48541625	48594504	E001	0.003094455	0.000135979	0.036090435	391.1611239	-0.146872673
ENSMUSG00000015597	Zfp318	17	46383731	46420920	E010	0.001532608	0.000206623	0.04917506	1009.420303	-0.080376944
ENSMUSG00000029290	Zfp326	5	105876565	105915818	E012	0.006976595	1.59E-11	7.93E-08	408.3314072	-0.356081813
ENSMUSG00000060206	Zfp462	4	54945048	55083563	E013	0.002497751	2.31E-10	8.53E-07	1461.679996	-0.168889013
ENSMUSG00000039841	Zfp800	6	28239931	28398005	E001	0.020941872	1.12E-06	0.000876698	85.45531021	-0.499345276
ENSMUSG00000022201	Zfr	15	12117831	12185683	E020	0.001051503	5.90E-14	6.99E-10	1967.197267	-0.140063928
ENSMUSG00000061524	Zic2	14	122475435	122479852	E003	0.001728974	3.46E-06	0.002192638	888.8909911	-0.089174637
ENSMUSG00000067860	Zic3	X	58022700	58041736	E003	0.00272824	4.77E-06	0.002885473	456.7577151	-0.121426682
ENSMUSG00000067860	Zic3	X	58022700	58041736	E001	0.004473711	2.79E-05	0.011254407	253.854821	0.181841521
ENSMUSG00000058446	Znrf2	6	54816916	54893500	E001	0.004853837	1.21E-05	0.005809488	231.4625193	0.224300916
ENSMUSG00000028180	Zranb2	3	157534160	157548390	E010	0.002552855	1.05E-07	0.000124116	2461.436353	-0.132253823

APPENDIX 2

RNA-seq data table of significantly differentially spliced exons in *Tardbp*$^{F210I/F210I}$ embryo heads compared to controls

Ensemb ID	Gene Name	Chr	Gene Start bp	Gene End bp	Exon ID	Dispersion	p value	p adjust	Mean Base	log2fold(Hom/Ctl)
ENSMUSG00000017843	Ppp2r5c	12	110485739	110583061	E015	0.001628858	0	0	1643.093307	-0.198639524
ENSMUSG00000024754	Tmem2	19	21778342	21858327	E002	0.02167731	0	0	48.31337848	1.197862287
ENSMUSG00000027082	Tfpi	2	84432855	84476775	E010	0.014356953	0	0	74.83449388	0.939576397
ENSMUSG00000027361	Gabpb1	2	126627442	126676337	E001	0.002088637	0	0	935.1644597	-0.236441433
ENSMUSG00000051243	Islr2	9	58196298	58204319	E002	0.009105289	0	0	329.1805811	0.935897942
ENSMUSG00000027523	Gnas	2	174284320	174346744	E003	0.009667694	1.54E-14	3.72E-10	1109.992086	0.41288257
ENSMUSG00000048910	2810453I06Rik	5	143548706	143564525	E002	0.011259577	5.10E-14	1.07E-09	97.47452989	0.647792524
ENSMUSG00000031967	Afg3l1	8	123477903	123503916	E017	0.00237734	2.12E-13	3.87E-09	736.0376023	0.218649164
ENSMUSG00000048910	2810453I06Rik	5	143548706	143564525	E003	0.001771742	2.30E-13	3.87E-09	1330.166785	-0.102784515
ENSMUSG00000027361	Gabpb1	2	126627442	126676337	E002	0.020385666	9.79E-13	1.50E-08	59.48034474	0.82138541
ENSMUSG00000051243	Islr2	9	58196298	58204319	E003	0.022749086	2.72E-12	3.53E-08	45.93018372	1.15250137
ENSMUSG00000022601	Rpl24	16	55966275	56008913	E017	0.003368822	1.17E-11	1.41E-07	694.7310827	0.252289609
ENSMUSG00000023411	Nfatc4	14	55824795	55833943	E010	0.001754768	5.70E-11	6.41E-07	1360.956995	0.146439439
ENSMUSG00000027668	Mfn1	3	32529465	32579239	E018	0.002242223	8.23E-11	8.67E-07	817.5061841	0.187024829
ENSMUSG00000026837	Col5a1	2	27886425	28039514	E066	0.001262084	1.37E-10	1.36E-06	7673.265019	0.090572927
ENSMUSG00000027082	Tfpi	2	84432855	84476775	E003	0.004041437	2.07E-10	1.94E-06	330.4542129	0.272762128
ENSMUSG00000020923	Ubtf	11	102304560	102319742	E014	0.003836073	2.73E-10	2.43E-06	354.5656264	0.282878179
ENSMUSG00000027184	Caprin1	2	103762941	103797649	E002	0.002758413	7.42E-10	6.26E-06	1563.885358	0.170407155
ENSMUSG00000058056	Palld	8	61513246	61902690	E019	0.010953689	1.20E-09	9.63E-06	148.2648092	0.48551711
ENSMUSG00000027523	Gnas	2	174284320	174346744	E002	0.02380324	1.61E-09	1.23E-05	2388.911635	0.436309257
ENSMUSG00000041360	D19Bwg1357e	19	27388702	27429820	E001	0.002325127	3.64E-09	2.67E-05	802.6657999	0.170918259
ENSMUSG00000016181	Diexf	1	193104403	193130251	E001	0.00284009	3.98E-09	2.80E-05	548.7492549	0.186901749
ENSMUSG00000029203	Ube2k	5	65537233	65598988	E007	0.001441781	4.97E-09	3.36E-05	2374.465147	-0.099221923
ENSMUSG00000032366	Tpm1	9	67022590	67049406	E003	0.002491174	8.43E-09	5.24E-05	727.1453724	0.183383908
ENSMUSG00000037857	Nufip2	11	77686155	77712741	E004	0.002339948	8.70E-09	5.24E-05	2494.945856	0.096816286
ENSMUSG00000027940	Tpm3	3	90072649	90100902	E001	0.063535187	1.26E-08	7.33E-05	15.9637381	1.35715426

ENSMUSG00000031816	Mthfsd	8	121095012	121108392	E002	0.006045182	2.48E-08	0.000137696	265.0388254	0.280888964
ENSMUSG00000034867	Ankrd27	7	35586247	35639237	E013	0.016226428	2.53E-08	0.000137696	65.63345198	0.591419042
ENSMUSG00000037503	Fam168b	1	34813219	34843065	E006	0.052921313	4.20E-08	0.000221329	19.22843364	1.150075084
ENSMUSG00000021140	Pcnx	12	81860030	82000924	E036	0.005786978	4.79E-08	0.0002447	1547.696936	0.212446717
ENSMUSG00000026150	Mff	1	82724890	82752394	E006	0.005304416	6.23E-08	0.000309218	233.0074098	0.291962215
ENSMUSG00000022483	Col2a1	15	97975602	98004695	E001	0.003479125	7.36E-08	0.000354667	6189.623446	0.121699569
ENSMUSG00000027523	Gnas	2	174284320	174346744	E009	0.004669934	1.62E-07	0.000757967	273.5286213	0.269804663
ENSMUSG00000027523	Gnas	2	174284320	174346744	E001	0.023264338	1.74E-07	0.000795168	364.5132558	0.471787691
ENSMUSG00000026879	Gsn	2	35279663	35307892	E001	0.012457352	2.04E-07	0.000906621	87.2651901	0.480362349
ENSMUSG00000028207	Asph	4	9448069	9669344	E023	0.017651666	2.41E-07	0.001040977	60.00853727	0.562319722
ENSMUSG00000027940	Tpm3	3	90072649	90100902	E002	0.077364874	4.57E-07	0.001926209	20.38465048	1.136891661
ENSMUSG00000022483	Col2a1	15	97975602	98004695	E002	0.001253525	5.23E-07	0.002154375	4300.971058	0.076232569
ENSMUSG00000028223	Decr1	4	15917240	15945507	E001	0.003516798	6.62E-07	0.002659916	399.9324777	0.190127886
ENSMUSG00000023764	Sfi1	11	3131850	3193463	E001	0.006302249	7.02E-07	0.002756022	188.9791266	0.300955954
ENSMUSG00000032366	Tpm1	9	67022590	67049406	E008	0.003116523	7.45E-07	0.002855783	476.3432459	-0.203102679
ENSMUSG00000032366	Tpm1	9	67022590	67049406	E004	0.002456425	1.04E-06	0.003822589	695.4712342	0.157729342
ENSMUSG00000029131	Dnajb6	5	29735688	29786478	E014	0.005849911	1.68E-06	0.00591509	798.839308	-0.203533177
ENSMUSG00000009394	Syn2	6	115134902	115282626	E014	0.014080465	1.68E-06	0.00591509	76.41890894	-0.378354594
ENSMUSG00000001472	Tcf25	8	123373824	123404173	E023	0.022643459	1.89E-06	0.006521374	61.94854108	0.572780494
ENSMUSG00000038914	Dido1	2	180657964	180709999	E001	0.003523652	2.03E-06	0.006844158	1903.901763	-0.132608103
ENSMUSG00000031577	Tti2	8	31150316	31164702	E009	0.008625125	2.18E-06	0.007162741	131.2468205	0.314895657
ENSMUSG00000040350	Trim7	11	48826140	48852209	E001	0.009041301	2.21E-06	0.007162741	134.5392829	-0.206944859
ENSMUSG00000027523	Gnas	2	174284320	174346744	E004	0.020835753	2.49E-06	0.007933061	102.26321	0.468235866
ENSMUSG00000000303	Cdh1	8	106603351	106670246	E016	0.001990553	2.57E-06	0.008045272	1029.817821	0.116633785
ENSMUSG00000017843	Ppp2r5c	12	110485739	110583061	E013	0.003783368	2.65E-06	0.008116434	627.0387841	0.179882469
ENSMUSG00000000881	Dlg3	X	100767722	100818410	E011	0.015411604	3.54E-06	0.01068109	69.34987627	0.466261712
ENSMUSG00000067786	Nnat	2	157560078	157562522	E002	0.001408626	4.03E-06	0.011931321	2577.817738	-0.061349638
ENSMUSG00000027823	Gmps	3	63976106	64022579	E016	0.002672109	4.19E-06	0.012201461	649.0636364	0.148169469

ENSMUSG00000032875	Arhgef17	7	100869746	100932161	E021	0.002854925	4.60E-06	0.013153435	1577.609692	-0.125406616
ENSMUSG00000024426	Atat1	17	35897595	35910075	E001	0.020710645	4.79E-06	0.013461792	297.6734952	-0.372562486
ENSMUSG00000054823	Whsc1l1	8	25601601	25719667	E026	0.001783645	5.84E-06	0.016155899	1505.642489	0.097772405
ENSMUSG00000028364	Tnc	4	63959785	64047015	E025	0.003095574	5.94E-06	0.016166863	481.1546456	-0.177728493
ENSMUSG00000090053	Palm2	4	57434247	57712016	E006	0.025959967	6.81E-06	0.018233832	866.5715679	0.379753891
ENSMUSG00000051243	Islr2	9	58196298	58204319	E001	0.006535263	6.92E-06	0.018233832	7354.656049	-0.144328846
ENSMUSG00000050315	Synpo2	3	123076519	123236149	E001	0.024888146	7.53E-06	0.019540132	384.1486589	-0.380085599
ENSMUSG00000010517	Faf1	4	109676588	109963960	E019	0.003103244	7.70E-06	0.019680451	479.381731	0.167434663
ENSMUSG00000033060	Lmo7	14	101729957	101934710	E010	0.02014706	8.15E-06	0.020532767	52.17896713	0.560179895
ENSMUSG00000025034	Trim8	19	46501648	46516444	E001	0.002194012	9.12E-06	0.022636813	1160.720483	-0.110875772
ENSMUSG00000039735	Fnbp1l	3	122538719	122619667	E007	0.122049134	9.46E-06	0.023076302	59.28700264	0.800322364
ENSMUSG00000019978	Epb4.1l2	10	25359798	25523519	E014	0.009441538	1.00E-05	0.023869554	118.5210388	0.341049883
ENSMUSG00000061689	Dlgap4	2	156613705	156764363	E003	0.00679586	1.03E-05	0.024140312	172.8245884	0.282161312
ENSMUSG00000038729	Akap2	4	57717657	57896984	E002	0.010467762	1.18E-05	0.027363873	105.6451462	0.342524486
ENSMUSG00000020173	Cobl	11	12236608	12464960	E004	0.004158096	1.75E-05	0.039922929	318.163615	-0.182786605
ENSMUSG00000025102	3110040N11Rik	7	81782182	81789478	E001	0.014639183	1.92E-05	0.043168784	73.28353061	0.398397785
ENSMUSG00000021427	Ssr1	13	37971401	37994190	E001	0.002336582	2.11E-05	0.046674354	6541.41808	0.074002812
ENSMUSG00000030306	Tmtc1	6	148232430	148444389	E001	0.005104364	2.16E-05	0.046674354	742.3240165	-0.113306936
ENSMUSG00000036550	Cnot1	8	95719451	95807462	E043	0.014134517	2.16E-05	0.046674354	121.6737973	0.317950973
ENSMUSG00000004565	Pnpla6	8	3515384	3544266	E015	0.010433995	2.23E-05	0.047134855	106.0241444	0.343271002
ENSMUSG00000026442	Nfasc	1	132564691	132741797	E001	0.001554409	2.25E-05	0.047134855	1872.634609	-0.066590618
ENSMUSG00000020315	Spnb2	11	30099395	30268175	E005	0.106585745	2.26E-05	0.047134855	368.0538369	-0.805018908
ENSMUSG00000004565	Pnpla6	8	3515384	3544266	E012	0.010651094	2.33E-05	0.047852448	103.6338323	0.333054493
ENSMUSG00000020661	Dnmt3a	12	3806007	3914443	E006	0.006733535	2.45E-05	0.049896659	174.7103079	0.269815602
ENSMUSG00000040479	Dgkz	2	91932824	91975864	E032	0.013846795	2.49E-05	0.049970807	77.81122226	0.387177872