# Creating and characterising physiological mouse models of SOD1- and C9orf72-ALS




**David Thompson**

Linacre College

Nuffield Department of Clinical Neurosciences

Medical Sciences Division

University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*

Word count: ~ 48,100

Michaelmas 2024

I, David Thompson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that it has been clearly indicated in the thesis.

# Abstract

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease characterised by the loss of motor neurons. Mutations in the superoxide dismutase 1 (*SOD1*) gene and an intronic repeat expansion in the *C9orf72* gene together account for ~50% of familial ALS cases.

Mouse models of SOD1- and C9orf72-ALS have been instrumental in improving our understanding of the disease; however, current models have a number of caveats which may limit the translatability of findings. This thesis focusses on genomically humanised mice, a novel class of mouse model in which mouse genomic regions are replaced with the orthologous human sequences. Genomically humanised mice avoid issues associated with transgene overexpression and the presence of intact mouse orthologues.

In this thesis I present the creation and characterisation of two novel humanised mouse models of SOD1-ALS – *hSOD1*$^{WT}$, which carries wild-type human *SOD1*, and *hSOD1*$^{A4V}$, which carries human *SOD1* with the ALS causative A4V mutation. I also present the creation of a *hSOD1*$^{G93A}$ mouse model, a humanised counterpart to the widely used transgenic *tgSOD1*$^{G93A}$ mouse, and a humanised *hC9orf72* mouse model.

Phenotyping of *hSOD1*$^{WT}$ animals demonstrated that humanisation of *SOD1* is well-tolerated in mice. Extensive behavioural and molecular phenotyping of *hSOD1*$^{A4V}$ mice revealed only mild metabolic phenotypes without overt neurodegeneration.

This work highlights the challenges associated with modelling late-onset neurodegenerative conditions, such as ALS, in mice. This work also highlights the importance of studying a range of mouse models to gain insights into different aspects of disease. The questions answerable with humanised *SOD1* mouse models relate to early-stage disease processes, and the study or pharmacological/genetic manipulation of human SOD1 *in vivo*.

# Acknowledgements

I would like to take this opportunity to express my heartfelt gratitude to all the people who have supported and contributed to the completion of this thesis. Firstly, I would like to thank my supervisors, Dr Tom Cunningham, Professor Lizzy Fisher and Professor Kevin Talbot for their unwavering encouragement, invaluable guidance, and continuous support which have been instrumental in making this research a reality. I would like to extend a special thanks to Dr Tom Cunningham for his mentorship throughout my DPhil, and his guidance during the writing of this thesis.

I would also sincerely like to thank Dr Jonathan Gilthorpe for sharing his vast knowledge and experience of all things SOD1, and his excitement and positivity towards data that was, at times, seemingly unexciting. I am immensely grateful to Chloe Williams for her endless motivation and enthusiasm, which astounds me to this day, and for the huge amount of work she has put into this project. I would also like to thank Dr Andrew Tosolini for his expert contribution to the axonal transport experiments in this thesis.

I'd like to thank the other members of the MMoN team – Dr Remya Nair, Dr Charlotte Tibbit, Georgia Price and Ross McLeod for their support and team spirit both in and out of the lab. A particularly wholehearted thanks must go to Georgia Price, my DPhil 'partner in crime' who has supported my project and myself in too many ways to mention. Georgia has been a pillar to lean on when things got tough and I owe her a great deal.

I would also like to extend my sincere thanks to all the staff at the Mary Lyon Centre that contributed to this project, who are too many to name. A particular thank you to the Ward 5 staff – including Jackie Harrison, Gemma Atkins, Dr Anju Paudyal and Emily Ireson – who were not only excellently helpful and accommodating, but were also a source of positivity and encouragement, and a joy to work with. I would also like to thank Dr Michelle Stewart for her unwavering support and

sound advice on all things mouse-related. I am also extremely grateful to the MLC phenotyping team, including Clare Norris and Sarah Taylor, for all their hard work to phenotype my mice. My deepest thanks also go to the MLC histology team – particularly Adele Austin – for their indispensable work and support in processing my samples, and the MLC genotyping and CRISPR teams for their comprehensive experimental support.

Finally, thank you to my mum, dad and brother for being the best. And thanks to my wonderful and patient girlfriend Anca, whose kindness and pragmatism prevented me from going mad during this write up.

# Table of contents

**Abstract** ..... **3**

**Acknowledgements** ..... **5**

**Table of contents** ..... **7**

**List of figures** ..... **13**

**List of tables** ..... **15**

**Abbreviations** ..... **17**

**Chapter 1: Introduction** ..... **20**

**1.1 Amyotrophic Lateral Sclerosis (ALS)** ..... **20**

1.1.1 Clinical disease course ..... 20

1.1.2 The ALS-FTD spectrum ..... 21

1.1.3 Genetics of ALS ..... 21

**1.2 General ALS pathomechanisms** ..... **23**

**1.3 SOD1** ..... **23**

1.3.1 SOD1-fALS ..... 24

1.3.2 SOD1 gain-of-function pathomechanisms ..... 25

1.3.3 SOD1 loss-of-function (LOF) ..... 33

1.3.4 Relationship between SOD1-ALS and other forms of ALS ..... 34

**1.4 C9orf72** ..... **35**

1.4.1 Loss-of-function mechanisms ..... 36

1.4.2 Gain-of-function mechanisms ..... 36

**1.5 Mouse models of ALS** ..... **40**

1.5.1 The utility of mouse models for ALS research ..... 40

1.5.2 Current mouse models of ALS ..... 42

1.5.3 Existing mouse models of SOD1-ALS ..... 48

1.5.4 Existing BAC transgenic mouse models of C9orf72-ALS/FTD ..... 54

**1.6 Research project aims** ..... **59**

1.6.1 Chapter 3 ..... 59

1.6.2 Chapters 4 and 5 ..... 59

1.6.3 Chapter 6 ..... 59

**Chapter 2: Materials and Methods** ..... **60**

**2.1 Materials and Equipment** ..... **60**

2.1.1 Equipment ..... 60

2.1.2 Consumables ..... 61
2.1.3 Reagents ..... 61
2.1.4 Software ..... 66
2.1.5 Work conducted by other individuals/groups ..... 66
**2.2 Mice ..... 67**
2.2.1 Ethics statement ..... 67
2.2.2 Housing conditions and husbandry ..... 68
2.2.3 Humane end points ..... 68
2.2.4 Abbreviated genotypes ..... 68
2.2.5 C57BL/6J mice ..... 69
2.2.6 129S9/SvEvH mice ..... 69
2.2.7 Humanised *SOD1* mice ..... 69
**2.3 Mouse experimental procedures ..... 73**
2.3.1 Genotyping ..... 73
2.3.2 Phenotyping ..... 74
**2.4 Histology, immunohistochemistry and immunofluorescence ..... 82**
2.4.1 MN counts ..... 82
2.4.2 NMJ analysis on lumbrical muscles ..... 83
2.4.3 NMJ analysis on TA and soleus muscles ..... 84
2.4.4 SOD1 immunofluorescence ..... 85
2.4.5 Misfolded/aggregated SOD1 immunohistochemistry ..... 85
2.4.6 Muscle fiber typing ..... 86
2.4.7 Determination of muscle fibre size and central nucleation ..... 87
**2.5 Molecular biology ..... 87**
2.5.1 DNA extraction from mouse tissues ..... 87
2.5.2 Protein extraction from mouse tissues ..... 88
2.5.3 Western blotting ..... 88
2.5.4 In-gel dismutase activity assays ..... 88
2.5.5 Spectrophotometric dismutase activity assays ..... 89
2.5.6 RNA extraction for quantitative PCR (qPCR) and RNA sequencing (RNA-seq) ..... 89
2.5.7 Quantitative PCR (qPCR) ..... 90
2.5.8 Quantseq 3' mRNA sequencing ..... 90
2.5.9 Xdrop enrichment and Nanopore sequencing of the *hSOD1*$^{WT}$ locus ..... 92
2.5.10 Targeted locus amplification and sequencing of the *hSOD1*$^{G93A}$ and *hC9orf72* loci ..... 92
2.5.11 Metabolomics ..... 93

2.5.12 Measurement of GSSG/GSH/ASA ..... 96
2.5.13 Measurement of amino acids ..... 97
2.5.14 Axonal transport experiments ..... 97
2.5.15 *In vivo* muscle isometric tension measurements ..... 97
**2.6 *In-vitro* molecular cloning ..... 98**
2.6.1 BAC and plasmid vectors ..... 98
2.6.2 Bacterial strains and culture conditions ..... 98
2.6.3 Transformation ..... 99
2.6.4 DNA isolation from bacterial cultures ..... 99
2.6.5 Agarose gel electrophoresis ..... 99
2.6.6 Primers ..... 99
2.6.7 PCR ..... 100
2.6.8 DNA Sequencing ..... 101
2.6.9 CRISPR/Cas9 and CRISPR/Cpf1 molecular cloning ..... 102
2.6.10 Red/ET recombineering ..... 103
2.6.11 Gibson assembly ..... 104
2.6.12 Restriction enzyme digests ..... 104
2.6.13 Ligation reactions ..... 104
2.6.14 Storage of DNA constructs ..... 104
**2.7 Mammalian cell culture and cell experimental procedures ..... 105**
2.7.1 Primary spinal cord cultures ..... 105
2.7.2 Assessment of mitochondrial respiration in primary spinal cord cultures ..... 105
2.7.3 Assessment of mitochondrial membrane potential in primary spinal cord cultures . 106
2.7.4 Statistics ..... 107
**Chapter 3: Characterisation of a genetically humanised mouse carrying wild-type human SOD1 ..... 108**
**3.1 Introduction ..... 108**
**3.2 Aims ..... 111**
**3.3 Results ..... 112**
3.3.1 hSOD1-WT Mendelian ratios ..... 112
3.3.2 *hSOD1$^{WT}$* Allele QC (TLA and long-read sequencing) ..... 112
3.3.3 *SOD1* mRNA expression ..... 117
3.3.4 SOD1 protein levels ..... 120
3.3.5 SOD1 protein levels following excision of duplicated exons ..... 122
3.3.6 SOD1 dismutase activity ..... 123

3.3.7 SOD1 Immunofluorescence ..... 124
3.3.8 hSOD1-WT phenotypic characterisation ..... 125
3.3.9 hSOD1-WT Histological characterisation ..... 135
**3.4 Discussion ..... 138**
**Chapter 4: Characterisation of a genetically humanised mouse carrying mutant human SOD1$^{A4V}$ ..... 140**
**4.1 Introduction ..... 140**
**4.2 Aims ..... 143**
**4.3 Results ..... 143**
4.3.1 hSOD1-A4V Mendelian ratios ..... 143
4.3.2 *SOD1-A4V* mRNA expression ..... 144
4.3.3 SOD1-A4V protein levels ..... 146
4.3.4 SOD1-A4V dismutase activity ..... 148
4.3.5 hSOD1-A4V molecular characterisation ..... 150
4.3.6 hSOD1-A4V phenotypic characterisation ..... 161
4.3.7 The effect of genetic background on phenotype ..... 195
4.3.8 hSOD1-A4V Histological characterisation ..... 204
**4.4 Discussion ..... 211**
4.4.1 SOD1 levels and activity in *hSOD1$^{A4V}$* mice ..... 211
4.4.2 There is a lack of clear motor and behavioural phenotypes in *hSOD1$^{A4V}$* mice ..... 211
4.4.3 Reduced TA NMJ functionality may be a phenotype of *hSOD1$^{A4V}$* mice ..... 212
4.4.4 Muscle pathology is absent in *hSOD1$^{A4V}$* mice ..... 214
4.4.5 Minimal transcriptional changes are detected in aging *hSOD1$^{A4V}$* mice ..... 214
4.4.6 Mild metabolic phenotypes are present in *hSOD1$^{A4V}$* mice ..... 215
4.4.7 No change in mitochondrial function in SC cultures from *hSOD1$^{A4V}$* mice ..... 217
4.4.8 Conclusion ..... 218
**Chapter 5: Creation and early characterisation of a genetically humanised mouse carrying mutant human SOD1$^{G93A}$ ..... 219**
**5.1 Introduction ..... 219**
**5.2 Aims ..... 220**
**5.3 Results ..... 221**
5.3.1 Generation of a targeting construct for the creation of an hSOD1$^{G93A}$ mouse line ..... 221
5.3.2 Generation of the *hSOD1$^{G93A}$* mouse line ..... 226
5.3.3 *hSOD1$^{G93A}$* allele quality control ..... 231
**5.4 Discussion ..... 236**

**Chapter 6: Progress towards the creation of a humanised mouse model of C9orf72-ALS/FTD ..... 239**

**6.1 Introduction ..... 239**

**6.2 Aims ..... 243**

**6.3 Results ..... 243**

6.3.1 Growing and sizing a large hexanucleotide repeat in the *C9orf72* gene within a BAC construct 243

6.3.2 Creation of a mouse *C9orf72* targeting construct ..... 248

6.3.3 Generation of a humanised *C9orf72* mouse by Dr Ben Davies ..... 268

6.3.4 Creation of a targeting construct to introduce a large *C9orf72* repeat expansion into iPSCs 270

**6.4 Discussion ..... 273**

6.4.1 Cloning ..... 273

6.4.2 Further use of constructs ..... 274

**Chapter 7: Discussion ..... 276**

**7.1 Summary of key findings ..... 276**

**7.2 Knock-in mouse models are not sufficient to model SOD1 GOF pathology ..... 277**

**7.3 Why isn't the *SOD1$^{A4V}$* mutation as severe in mice as in humans? ..... 279**

7.3.1 Age ..... 279

7.3.2 Species-specific biochemistry ..... 280

7.3.3 Anatomy ..... 281

**7.4 The suitability of mice as a model of SOD1 loss-of-function (LOF) ..... 282**

**7.5 Metabolic phenotypes precede neurodegeneration in *hSOD1$^{A4V}$* mice ..... 285**

**7.6 New methods for cloning repeat expansions and unstable DNA ..... 286**

**7.7 Future research into the phenotypes of humanised *hSOD1* and *hC9orf72* mice ..... 286**

7.7.1 *hSOD1$^{A4V}$* mice ..... 286

7.7.2 *hSOD1$^{G93A}$* mice ..... 288

7.7.3 *hC9orf72* mice ..... 288

**7.8 Potential applications of humanised *SOD1* mice ..... 289**

7.8.1 Approaches to increase the severity of phenotypes in humanised *SOD1* mice ..... 289

7.8.2 Potential uses of humanised *SOD1* mice in their current form ..... 291

**7.9 Conclusions ..... 292**

**References ..... 294**

**Appendices ..... 333**

8.1 Full lists of differentially expressed genes (DEGs) in *hSOD1$^{A4V}$* mice ..... 333

8.2 Full list of changed metabolites in plasma from *hSOD1$^{A4V}$* mice ..... 339

8.3 Representative immunofluorescence images from *hSOD1$^{A4V}$* embryonic spinal cord primary cultures ..... 343

8.4 Representative images of neuromuscular junction (NMJ) staining in aged *hSOD1$^{A4V}$* mice 344

8.5 Full sequence of a C9orf72 repeat expansion within the RP11-27J8 BAC construct..... 346

8.6 Full Western blot images ..... 348

# List of figures

Figure 1-1 | Schematic of the human *C9orf72* gene ..... 35
Figure 1-2 | Phenotypes in *tgSOD1$^{G93A}$* mice. ..... 53
Figure 2-1 | Predicted truncated SOD1 protein sequence in *hSOD1$^{-/-}$* mice ..... 73
Figure 2-2 | Subjective visual rating scale to assess misfolded/aggregated SOD1 immunohistochemistry ..... 86
Figure 2-3 | Representative images of TA muscle sections following ATPase staining ..... 87
Figure 2-4 | LC-MS internal standards ..... 94
Figure 2-5 | Example of normalisation of MS data ..... 96
Figure 3-1 | Schematic of the *hSOD1$^{WT}$* allele ..... 110
Figure 3-2 | Outline of experiments performed on *hSOD1$^{WT}$* mice. ..... 111
Figure 3-3 | TLA and long-read sequencing of the humanised *hSOD1$^{WT}$* locus ..... 114
Figure 3-4 | Strain identity of the humanised *hSOD1$^{WT}$* locus ..... 115
Figure 3-5 | Expression of *SOD1* RNA in *hSOD1$^{WT}$* mice by qRT-PCR ..... 119
Figure 3-6 | Levels of SOD1 protein in *hSOD1$^{WT}$* mice ..... 121
Figure 3-7 | Levels of SOD1 protein in *hSOD1$^{RDEL}$* mice. ..... 122
Figure 3-8 | SOD1 dismutase activity in *hSOD1$^{WT}$* mice ..... 124
Figure 3-9 | SOD1 immunoreactivity in *hSOD1$^{WT}$* spinal cord ..... 125
Figure 3-10 | Kaplan-Meier survival graphs for *hSOD1$^{WT}$* mice ..... 126
Figure 3-11 | Longitudinal open field test data from *hSOD1$^{WT}$* mice ..... 128
Figure 3-12 | Longitudinal grip strength data from *hSOD1$^{WT}$* mice ..... 129
Figure 3-13 | Longitudinal Locotronic data from *hSOD1$^{WT}$* mice ..... 130
Figure 3-14 | Longitudinal Rotarod data from *hSOD1$^{WT}$* mice ..... 131
Figure 3-15 | Longitudinal voluntary wheel running data from *hSOD1$^{WT}$* mice ..... 132
Figure 3-16 | Longitudinal bodyweight data from *hSOD1$^{WT}$* mice ..... 133
Figure 3-17 | Echo MRI data from *hSOD1$^{WT}$* mice ..... 135
Figure 3-18 | Spinal motor neuron counts from *hSOD1$^{WT}$* mice ..... 136
Figure 3-19 | Muscle fiber size and central nucleation in *hSOD1$^{WT}$* mice ..... 137
Figure 4-1 | Schematic of the *hSOD1$^{A4V}$* allele ..... 141
Figure 4-2 | Outline of experiments performed on *hSOD1$^{A4V}$* mice ..... 142
Figure 4-3 | Expression of *SOD1* RNA in *hSOD1$^{A4V}$* mice by qRT-PCR. ..... 145
Figure 4-4 | Levels of SOD1 protein in *hSOD1$^{A4V}$* mice ..... 147
Figure 4-5 | SOD1 dismutase activity in *hSOD1$^{A4V}$* mice ..... 149
Figure 4-6 | Retrograde axonal transport of signaling endosomes in TA motor neurons of 18-month-old male *hSOD1$^{A4V}$* mice ..... 151
Figure 4-7 | Longitudinal RNA sequencing data from the spinal cord of *hSOD1$^{A4V}$* mice ..... 157
Figure 4-8 | Longitudinal RNA sequencing data from the TA muscle of *hSOD1$^{A4V}$* mice ..... 161
Figure 4-9 | Kaplan-Meier survival graphs for *hSOD1$^{A4V}$* mice ..... 162
Figure 4-10 | Marble burying data from *hSOD1$^{A4V}$* mice ..... 166
Figure 4-11 | Longitudinal open field test data from *hSOD1$^{A4V}$* mice ..... 167
Figure 4-12 | Longitudinal grip strength data from *hSOD1$^{A4V}$* mice ..... 168
Figure 4-13 | Longitudinal gait analysis data from *hSOD1$^{A4V}$* mice ..... 171
Figure 4-14 | Longitudinal Locotronic data from *hSOD1$^{A4V}$* mice ..... 173
Figure 4-15 | Longitudinal Rotarod data from *hSOD1$^{A4V}$* mice ..... 174
Figure 4-16 | Longitudinal voluntary wheel running data from *hSOD1$^{A4V}$* mice ..... 176
Figure 4-17 | Longitudinal bodyweight data from *hSOD1$^{A4V}$* mice ..... 178

Figure 4-18 | Echo MRI data from *hSOD1$^{A4V}$* mice ..... 179
Figure 4-19 | Calorimetric cage data from *hSOD1$^{A4V}$* mice ..... 180
Figure 4-20 | Intraperitoneal glucose tolerance testing (IPGTT) in *hSOD1$^{A4V}$* mice ..... 182
Figure 4-21 | LC-MS analysis of plasma metabolites in *hSOD1$^{A4V}$* mice ..... 186
Figure 4-22 | GC-MS analysis of plasma metabolites in *hSOD1$^{A4V}$* mice ..... 188
Figure 4-23 | Plasma phenylalanine quantification in *hSOD1$^{A4V}$* mice ..... 189
Figure 4-24 | Multivariate analysis of *hSOD1$^{A4V}$* plasma metabolomics data. ..... 190
Figure 4-25 | Blood glutathione and ascorbate measurements from *hSOD1$^{A4V}$* mice. ..... 192
Figure 4-26 | Mitochondrial function in spinal cord cultures from *hSOD1$^{A4V}$* mice ..... 194
Figure 4-27 | Marble burying data from *hSOD1$^{A4V}$-129* mice ..... 196
Figure 4-28 | Open field test data from *hSOD1$^{A4V}$-129* mice ..... 197
Figure 4-29 | Motor function testing in *hSOD1$^{A4V}$-129* mice ..... 198
Figure 4-30 | Voluntary wheel running data from *hSOD1$^{A4V}$-129* mice ..... 200
Figure 4-31 | Longitudinal bodyweight data from *hSOD1$^{A4V}$-129* mice ..... 201
Figure 4-32 | Echo MRI data from *hSOD1$^{A4V}$-129* mice. ..... 202
Figure 4-33 | Calorimetric cage data from *hSOD1$^{A4V}$-129* mice ..... 203
Figure 4-34 | Intraperitoneal glucose tolerance testing (IPGTT) in *hSOD1$^{A4V}$-129* mice. ..... 204
Figure 4-35 | Spinal motor neuron counts from *hSOD1$^{A4V}$* mice ..... 204
Figure 4-36 | Analysis of neuromuscular junction (NMJ) denervation in *hSOD1$^{A4V}$* mice ..... 206
Figure 4-37 | Analysis of tibalis anterior myofibers in *hSOD1$^{A4V}$* mice. ..... 208
Figure 4-38 | Muscle physiology in *hSOD1$^{A4V}$* mice ..... 209
Figure 4-39 | Misfolded/aggregated SOD1 immunohistochemistry ..... 210
Figure 5-1 | Molecular cloning of a *hSOD1$^{G93A}$* ESC targeting construct ..... 225
Figure 5-2 | Nested PCR across the 5' homology arm ..... 228
Figure 5-3 | Targeted locus amplification (TLA) and short-read sequencing of *hSOD1$^{G93A}$* locus ..... 235
Figure 6-1 | Strategies for generation of a targeting construct to humanize mouse *C9orf72* and introduce a large hexanucleotide repeat ..... 242
Figure 6-2 | CRISPR/Cas9-based method to size *C9orf72* repeats ..... 244
Figure 6-3 | Optimisation of culture conditions for a repeat-containing human *C9orf72* BAC ..... 246
Figure 6-4 | Sanger sequencing into 5' and 3' ends of a patient-derived *C9orf72* repeat ..... 248
Figure 6-5 | Strategy for introduction of a mouse 3' homology arm (HA) into a patient-derived *C9orf72* BAC ..... 251
Figure 6-6 | CRISPR/Cpf1-based strategy to move a large *C9orf72* repeat from a patient-derived BAC into an intermediate vector ..... 256
Figure 6-7 | CRISPR/Cas9-based strategy to move a large *C9orf72* repeat from a patient-derived BAC into an intermediate vector ..... 258
Figure 6-8 | Strategy for removal of an AarI site from the pJAZZ vector ..... 261
Figure 6-9 | *De novo* GGGGCC repeat synthesis strategy ..... 263
Figure 6-10 | Screening of large GGGGCC repeats. ..... 265
Figure 6-11 | Strategy for the generation of a mouse *C9orf72* targeting construct, performed by Dr Nair ..... 268
Figure 6-12 | Humanisation of the mouse *C9orf72* locus by CRISPR/Cas9-assisted homology-directed repair (HDR) ..... 269
Figure 6-13 | Generation of a human iPSC *C9orf72* targeting construct ..... 272
Figure a-1 | Immunofluorescence images of neuronal/glial markers in spinal cord primary cultures from E13.5 *hSOD1$^{A4V}$* mice ..... 344

Figure a-2 | Representative images of neuromuscular junction (NMJ) staining in lumbrical muscles from 24-month-old *hSOD1$^{A4V}$* mice.....344
Figure a-3 | Representative images of neuromuscular junction (NMJ) staining in tibialis anterior (TA) muscles from 24-month-old *hSOD1$^{A4V}$* mice.....345
Figure a-4 | Representative images of neuromuscular junction (NMJ) staining in soleus muscles from 24-month-old *hSOD1$^{A4V}$* mice. .....346
Figure a-5 | Total protein staining images of Western blots of tissues from *hSOD1$^{WT}$* and *hSOD1$^{RDEL}$* mice .....348
Figure a-6 | Full images of Western blots of tissues from *hSOD1$^{A4V}$* mice .....349

## List of tables

Table 1-1 | Genes most commonly implicated in fALS.....22
Table 1-2 | Classes of mouse models and their key features.....43
Table 1-3 | SOD1 mouse models .....51
Table 1-4 | Published *C9orf72* BAC transgenic mouse models.....56
Table 2-1 | List of equipment .....61
Table 2-2 | List of consumables.....61
Table 2-3 | List of reagents.....66
Table 2-4 | List of software.....66
Table 2-5 | List of abbreviated strain names.....69
Table 2-6 | List of abbreviated genotypes.....69
Table 2-7 | PCR primers, sequencing primers, and probes for screening of *hSOD1$^{A4V}$* and hSOD1$^{-/-}$ mice.....72
Table 2-8 | Genotyping primers and probes .....73
Table 2-9 | Genotyping primers and probes against recombined *hSOD1$^{WT}$* allele.....73
Table 2-10 | List of gait analysis parameters.....78
Table 2-11 | List of qPCR primers .....90
Table 2-12 | Read counts from Quantseq 3' mRNA sequencing runs. mo, month-old.....91
Table 2-13 | Reaction mix for MyTaq Red PCR.....100
Table 2-14 | Cycling conditions for MyTaq Red PCR .....100
Table 2-15 | Reaction mix for Q5 PCR .....100
Table 2-16 | Cycling conditions for Q5 PCR .....100
Table 2-17 | Statistical tests for analysis of data from phenotyping and molecular biology experiments .....107
Table 3-1 | Genotypes of offspring born from intercrosses between heterozygous *mSod1/hSOD1$^{WT}$* mice.....112
Table 3-2 | Detection primers for Xdrop enrichment of the *hSOD1$^{WT}$* locus. .....117
Table 3-3 | Disease-related causes of death of *hSOD1$^{WT}$* mice .....126
Table 3-4 | SHIRPA phenotyping results from *hSOD1$^{WT}$* mice .....127
Table 4-1 | Genotypes of offspring born from intercrosses between heterozygous *hSOD1$^{WT}$/ hSOD1$^{A4V}$* mice.....143
Table 4-2 | Differentially expressed genes (DEGs) in the spinal cord of *hSOD1$^{A4V}$* mice.....154
Table 4-3 | GO biological pathways associated with DEGs in the spinal cord of *hSOD1$^{A4V}$* mice.....155
Table 4-4 | Differentially expressed genes (DEGs) in the TA muscle of *hSOD1$^{A4V}$* mice.....159
Table 4-5 | GO biological pathways associated with DEGs in the TA muscle of *hSOD1$^{A4V}$* mice.....160

Table 4-6 | Disease-related causes of death of *hSOD1$^{A4V}$* mice. ..... 163
Table 4-7 | Pathologies present at death of *hSOD1$^{A4V}$* mice ..... 164
Table 4-8 | SHIRPA phenotyping results from *hSOD1$^{A4V}$* mice ..... 165
Table 4-9 | Metabolic pathway analysis of significantly altered metabolites in plasma from 20-month-old *hSOD1$^{A4V}$* mice ..... 188
Table 5-1 | sgRNA guides used in the creation of hSOD1-G93A-pJAZZ. ..... 225
Table 5-2 | Sequencing primers used in the creation of hSOD1-G93A-pJAZZ. ..... 225
Table 5-3 | PCR Primers used in the creation of hSOD1-G93A-pJAZZ ..... 226
Table 5-4 | sgRNAs for CRISPR/Cas9-assisted editing of ESCs to generate *hSOD1$^{G93A}$* mice ..... 226
Table 5-5 | PCR primers for screening of targeted ESCs. ..... 229
Table 5-6 | Nested PCR cycling conditions ..... 229
Table 5-7 | Primers/probes used in ddPCR assays to screen for integration of the *hSOD1$^{G93A}$* transgene ..... 230
Table 5-8 | Summary of screening results for the *hSOD1$^{G93A}$* ESC clones. ..... 230
Table 5-9 | Detection primers for Xdrop enrichment of the *hSOD1$^{G93A}$* locus. ..... 236
Table 6-1 | Cas9 sgRNAs for sizing the *C9orf72* repeat ..... 244
Table 6-2 | Sequencing primers for sanger sequencing of the *C9orf72* repeat ..... 248
Table 6-3 | Primers used in the creation of hC9-BAC-Rpsl-HA ..... 252
Table 6-4 | Cpf1 crRNA sequences ..... 254
Table 6-5 | pSMART-cpf1 colony PCR primers. ..... 254
Table 6-6 | pSMART-cpf1 sequencing primers ..... 255
Table 6-7 | Primers for LR-PCR of Vector A ..... 259
Table 6-8 | Colony PCR primers to screen for 'Vector A + rep' construct ..... 259
Table 6-9 | Primers for sequencing pJAZZ-$\delta$AarI ..... 261
Table a-1 | Full list of differentially expressed genes (DEGs) in the spinal cord of male *hSOD1$^{A4V}$* mice ..... 334
Table a-2 | Full list of differentially expressed genes (DEGs) in the TA muscle of male *hSOD1$^{A4V}$* mice ..... 339
Table a-3 | Full list of significantly changed metabolites from the LC-MS analysis of plasma from *hSOD1$^{A4V}$* mice. ..... 341
Table a-4 | Full list of significantly changed metabolites from the GC-MS analysis of plasma from *hSOD1$^{A4V}$* mice. ..... 342

# Abbreviations

| Abbreviation | Full text |
|---|---|
| **AD** | Autosomal dominant |
| **ALS** | Amyotrophic lateral sclerosis |
| **AR** | Autosomal recessive |
| **ASA** | Ascorbate |
| **ASO** | Antisense oligonucleotide |
| **AUC** | Area under the curve |
| **BAC** | Bacterial artificial chromosome |
| **BDNF** | Brain-derived neurotrophic factor |
| **BSA** | Bovine serum albumin |
| **C9-BAC** | *C9orf72* BAC transgenic mice |
| **CN-PAGE** | Clear native polyacrylamide gel electrophoresis |
| **CNS** | Central nervous system |
| **crRNA** | CRISPR RNA |
| **CSF** | Cerebrospinal fluid |
| **DE** | Differential gene expression |
| **DEG** | Differentially expressed gene |
| **DENN** | Differentially expressed in normal and neoplasia |
| **DIV** | Days *in vitro* |
| **DPR** | Dipeptide repeat |
| **EDL** | Extensor digitorum longus |
| **ENU** | *N*-ethyl-*N*-nitrosourea |
| **ER** | Endoplasmic reticulum |
| **ERAD** | ER-associated degradation |
| **ESC** | Embryonic stem cell |
| **FACS** | Fluorescence-activated cell sorting |
| **fALS** | Familial amyotrophic lateral sclerosis |
| **FCS** | Foetal calf serum |
| **FDR** | False discovery rate |
| **FTD** | Frontotemporal Dementia |
| **FUS** | Fused in sarcoma |
| **GC-MS** | Gas chromatography–mass spectrometry |
| **GEF** | Guanine nucleotide exchange factor |
| **GI** | Gastrointestinal |
| **GLT** | Germline transmission |
| **GO** | Gene ontology |
| **GOF** | Gain-of-function |
| **GSEA** | Gene set enrichment analysis |
| **GSH** | Glutathione |
| **GSSG** | Glutathione disulfide |
| **HA** | Homology arm |
| **HBSS** | Hanks' Balanced Salt Solution |
| **$H_CT$** | Tetanus toxin heavy chain |
| **HDR** | Homology-directed repair |
| **HMW** | High molecular weight |
| **IGV** | Integrative genomics viewer |
| **IMS** | Inner membrane space of mitochondria |

| | |
|---|---|
| **IP** | Intraperitoneal |
| **IPGTT** | Intraperitoneal glucose tolerance test |
| **IVF** | *In vitro* fertilisation |
| **JALS** | juvenile amyotrophic lateral sclerosis |
| **KI** | Knock-in |
| **LCD** | Low complexity domain |
| **LC-MS** | Liquid chromatography–mass spectrometry |
| **LMN** | Lower motor neuron |
| **LOA** | Loss-of-allele |
| **LOF** | Loss-of-function |
| **LR-PCR** | Long range polymerase chain reaction |
| **MLC** | Mary Lyon Centre |
| **MN** | Motor neuron |
| **MPA** | *Meta*-phosphoric acid |
| **MRC** | Medical Research Council |
| **MS** | Mass spectrometry |
| **ND** | Not determined |
| **NHEJ** | Non-homologous end joining |
| **NMJ** | Neuromuscular junction |
| **OMM** | Outer mitochondrial membrane |
| **PCA** | Principal component analysis |
| **PCR** | Polymerase chain reaction |
| **PFGE** | Pulsed-field gel electrophoresis |
| **PLS-DA** | Partial least squares - discriminant analysis |
| **QC** | Quality control |
| **qPCR** | Quantitative polymerase chain reaction |
| **RBC** | Red blood cell |
| **RDL** | Recursive directional ligation |
| **RER** | Respiratory exchange ratio |
| **RIN** | RNA integrity number |
| **RNA-seq** | RNA sequencing |
| **ROS** | Reactive oxygen species |
| **sALS** | Sporadic amyotrophic lateral sclerosis |
| **SC** | Spinal cord |
| **SCS** | Space cadet syndrome |
| **sgRNA** | Cas9 single guide RNA |
| **SMA** | Spinal muscular atrophy |
| **SNP** | Single nucleotide polymorphism |
| **SOD1** | Superoxide dismutase 1 |
| **TA** | Tibialis anterior |
| **TDP-43** | TAR DNA binding protein 43 |
| **TLA** | Targeted locus amplification |
| **UMN** | Upper motor neuron |
| **UPR** | Unfolded protein response |
| **UTR** | Untranslated region |
| **VLE** | Very low expression |
| **WT** | Wild-type |
| **α-BTX** | α-bungarotoxin |

# Chapter 1: Introduction

## 1.1 Amyotrophic Lateral Sclerosis (ALS)

### 1.1.1 Clinical disease course

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease characterised by the progressive loss of upper and lower motor neurons (MNs), leading to muscle atrophy and eventually death. The incidence of ALS worldwide is approximately 2/100,000 population/year, equating to around 1/400 lifetime risk [1].

ALS is a highly heterogeneous disease and there is significant variability in clinical disease course between patients. The average age of onset of ALS is between 60 and 70 years of age [2], although this can vary substantially, with juvenile forms of ALS – most commonly associated with FUS mutations (FUS-JALS) – occurring much earlier in life [3,4]. Disease severity and progression also varies widely; average disease duration is 3-5 years after diagnosis, however some patients, such as FUS-JALS patients [3], exhibit more aggressive disease and shorter disease duration, whereas others display slower progression [2]. Approximately 15-20% of patients survive longer than 5 years, and a small percentage have a disease course of decades [5].

Initial symptoms of ALS include weakness of voluntary muscles, which may be accompanied by upper motor neuron (UMN) signs, including hyperreflexia and spasticity; lower motor neuron (LMN) signs, including hyporeflexia and muscle weakness; or both [2]. Cases are typically defined as bulbar onset, limb onset, or, rarely, diaphragmatic onset, depending on clinical presentation. Often, symptoms begin at a focal point of onset and spread to other areas as the disease progresses [6]. The hallmark of ALS is the presence of both UMN and LMN symptoms, accompanied by progressive motor weakness and a relative sparing of sensory function [2].

### 1.1.2 The ALS-FTD spectrum

ALS was originally believed to be purely a disorder of the motor system. However, detailed observation of ALS patients over a number of decades has revealed significant overlap with other neurodegenerative disorders, most notably frontotemporal dementia (FTD). FTD is characterised by loss of neurons in the frontal and temporal cortices of the brain, leading to symptoms of cognitive impairment, including behavioural disturbances, impaired judgement and executive deficits [7]. Approximately 50% of ALS patients present with cognitive impairment, with 15% of patients meeting the diagnostic criteria for FTD [8]. These observations, combined with the fact that ALS and FTD share many genetic risk factors and pathophysiological hallmarks, have led to the understanding that ALS and FTD lie on a disease spectrum [9,10]. ALS and FTD represent the extremes of the spectrum, which encompasses a continuum of neurodegenerative disorders with various degrees of motor and cognitive involvement.

### 1.1.3 Genetics of ALS

Whilst approximately 90% of ALS cases are sporadic (sALS; without a family history), the remaining 10% of cases are familial (fALS) [11]. Over 50 genes have been identified as causative or disease-modifying in ALS, accounting for approximately 15% and 70% of European sALS and fALS cases respectively [12]. Studies in identical twins estimate heritability in sALS to be around 60% [13] so it is likely that new ALS-linked genes will continue to be discovered and further our understanding of ALS disease mechanisms. Despite the large number of ALS-linked genes, just 4 genes are implicated in the majority of fALS cases – *C9ORF72, SOD1, TARDBP* and *FUS* (Table 1-1) [14–16].

There is significant overlap between the genetics of ALS and FTD [9,10]. Mutations in some genes, such as those encoding microtubule-associated protein tau (*MAPT*) [17] and progranulin (*PGRN*) [18,19] result in FTD but not ALS. Pathogenic variants of other genes, most notably *C9ORF72* [20,21], are causal for both ALS and FTD. Whereas mutations in a few genes, for example *SOD1* [22], are only causal for ALS and have not been linked to FTD.

| ALS locus number | Gene (protein) | Chromosomal location | Functions | Heredity | Onset | fALS % | sALS % | Other diseases | Original reference(s) |
|---|---|---|---|---|---|---|---|---|---|
| **ALS1** | *SOD1* (Superoxide dismutase 1) | 21q22.11 | Dismutates superoxide free radicals | AD (rarely AR) | Adult | 20 | 2 | | 22 |
| **FTD-ALS1** | *C9orf72* (Guanine nucleotide exchange C9orf72) | 9p21.2 | Autophagy; nucleocytoplasmic transport; protein homeostasis | AD | Adult | 25 | 5 | FTD | 20,21 |
| **ALS6** | *FUS* (Fused in sarcoma) | 16p11.2 | RNA processing | AD (rarely AR) | Adult, juvenile | 5 | <1 | FTD | 23,24 |
| **ALS10** | *TARDBP* (TAR DNA binding protein 43) | 1p36.22 | RNA processing | AD | Adult | 5 | <1 | FTD | 25,26 |

**Table 1-1 | Genes most commonly causative for ALS.** Adapted from Ling et al., 2013 [9]. AD, autosomal dominant; AR, autosomal recessive.

## 1.2 General ALS pathomechanisms

The aetiology of ALS is poorly understood. Despite an acceleration of ALS-linked gene discovery in recent years, there remains no unifying pathomechanism downstream of the disparate mutations to explain the selective death of motor neurons.

Examination of the relationships between genes implicated in ALS reveals several underlying themes. ALS-linked genes are involved in protein homeostasis, RNA metabolism, axon structure and function, mitochondrial function, vesicular transport, and a few other areas [12,14]. Accordingly, a wide range of cellular processes are disrupted in the disease; including those listed above, in addition to disruptions in DNA repair, nucleocytoplasmic transport and oxidative stress [14,27]. Additionally, it is clear that non-cell-autonomous mechanisms, including neuroinflammation, excitotoxicity and prion-like propagation of protein aggregates also play a role in ALS pathogenesis [27].

With so many pathways affected, the complexity of ALS makes establishing causality highly difficult. This complexity is underscored by significant phenotypic variability between patients, even between individuals carrying mutations in the same gene. However, aggregation of the RNA/DNA-binding protein TDP-43 is a common factor in some 97% of ALS cases regardless of the mechanisms of disease onset, highlighting perturbed RNA metabolism as one common mechanism [28].

## 1.3 SOD1

SOD1 (Superoxide Dismutase 1, or Cu/Zn Superoxide Dismutase) is a ubiquitously expressed metalloenzyme encoded by the 9.3kb *SOD1* gene located on chromosome 21. The 32kDa SOD1 holoenzyme is a homo-dimer of two identical 16kDa subunits, each consisting of 153 amino acids [29]. SOD1 requires binding to copper and zinc to carry out its enzymatic function. Each SOD1 subunit coordinates with zinc and copper ions, and has an intramolecular disulphide bond, resulting in extremely high protein stability [30].

The main known function of SOD1 is to metabolise superoxide radicals into less harmful hydrogen peroxide and molecular oxygen [31]. The superoxide scavenging activity of SOD1 provides protection against reactive oxygen species toxicity [32]; however, studies in yeast suggest that less than 1% of total SOD1 produced is required for this canonical function [33]. Accordingly, a number of other functions of SOD1 have been described, including roles in copper and zinc homeostasis [34,35], phosphatase activation [36,37], control of cellular respiration [38], and transcriptional control of oxidative stress resistance [39,40], among other roles (reviewed in reference [41]).

In mouse, the *Sod1* gene is located on chromosome 16 and is only 5.6kb in length due to shorter intronic sequences than human *SOD1*. *SOD1* is highly conserved among mammals [42] and the mouse *Sod1* orthologue shares 84% coding sequence homology and 83% amino acid homology with human *SOD1.* The 3D Structure of SOD1 is also highly conserved between humans and mice [43].

In humans, *SOD1* is expressed in all tissues, with highest expression observed in the central nervous system (CNS), kidney, liver and testes [44,45]. Within the mammalian CNS, *SOD1* is expressed particularly highly in astrocytes, and is also present at high levels in spinal MNs [46]. SOD1 is predominantly located in the cytosol and nucleoplasm [47]. All mammals also possess two other isoforms of SOD: the mitochondrial MnSOD (SOD2) and the extracellular Cu/ZnSOD (SOD3).

### 1.3.1 SOD1-fALS

Pathogenic variants of the *SOD1* gene were the first discovered genetic cause of ALS [22] and continue to account for 15-30% of fALS and approximately 2% of sALS cases [15]. Over 180 different SOD1 mutations have been identified – however, pathogenicity has not been established for every variant – and the vast majority of disease-causing variants follow an autosomal-dominant inheritance pattern [14,48,49]. Disease-causing variants are spread across all 5 exons.

ALS disease severity and progression varies greatly between different SOD1 variants. For example, patients with A4V, G85R and G93A variants generally display rapid disease progression, whereas

other mutations such as G93C and D90A are associated with slower degeneration and longer life expectancy [14].

A large number of mutant SOD1 isoforms have been examined *in vitro*, revealing considerable variations in their biochemical and biophysical attributes. As a result, pinpointing their common harmful characteristics has proven to be a challenging endeavour [50,51]. Many studies have tried to link the structure of SOD1 mutants to disease severity and progression [52–59], with proposed correlations between ALS severity and SOD aggregation propensity, dimer stability, misfolded monomeric structure, and heterodimerisation characteristics.

### 1.3.2 SOD1 gain-of-function pathomechanisms

The mechanisms through which SOD1 mutations cause ALS have not been fully elucidated. A large body of evidence supports toxic gain-of-function mechanisms, several of which have been proposed and comprehensively reviewed [41,60,61]. At present it is unclear which mechanisms are causative for SOD1-ALS and which are secondary to degeneration – although many mechanisms are likely both a cause and result of disease through participation in feedback loops of cell dysregulation and damage. The following is a brief description of the key proposed mechanisms in SOD1-ALS, with emphasis on those evidenced in both humans and mice.

#### *1.3.2.1 Formation of intracellular aggregates*

Large insoluble aggregates of SOD1 in MNs are a hallmark of SOD1-ALS, and have been consistently observed in tissues from SOD1-ALS patients [62–64] and mouse models [65]. However, the role of large SOD1 aggregates in ALS pathogenesis remains unclear.

It was initially believed that aggregated SOD1 was the major toxic species in SOD1-ALS. This hypothesis was supported by evidence from SOD1-ALS mouse models revealing a correlation between the presence of insoluble mutant SOD1 aggregates and the manifestation of phenotypes [66–71]. These findings aligned with observations in patients [72]. Furthermore, double transgenic mice co-

expressing both mutant and wild-type (WT) SOD1 typically display both increased SOD1 aggregation and more severe disease [71,73–75].

However, more recent studies cast doubt on this hypothesis, with some groups observing that SOD1 aggregation in the CNS is not correlated, or is inversely correlated, with disease [56,76]. Additionally, inclusions of aggregated SOD1 are not usually observed until after the mouse is symptomatic [66,67,77,78], suggesting that they may be involved in advanced disease, rather than a driver of early pathogenic processes. Instead, a growing body of *in vitro* and *in vivo* evidence suggests that soluble misfolded SOD1 is more toxic than insoluble aggregates [78–83]. This would parallel other neurodegenerative diseases such as prion disease and Alzheimer's disease, where soluble oligomers are suggested to be the most toxic species [84,85]. However, one study found a lack of correlation between soluble aggregate levels and disease duration [56].

Uncovering the toxic species in SOD1-ALS is complicated by the fact that most studies to date have used cell models or transgenic mice that overexpress mutant SOD1. In such models there is rampant SOD1 misfolding and aggregation, and an imbalance between mutant and wild-type SOD1 (almost all SOD1-ALS patients carry one wild-type and one mutant allele), making it difficult to evaluate different toxic species and the interactions between them. Further investigation using models with physiological levels of SOD1 expression are warranted to support earlier conclusions.

#### *1.3.2.2 Prion-like spreading of SOD1 pathology*

Prions ('proteinaceous infectious particles') are unconventional infectious agents that spread pathology in the absence of a nucleic acid component. The first prion disease described was scrapie, an infectious neurological disorder affecting sheep and goats, where the infectious particle was found to consist solely of misfolded prion protein (PrP) [86]. PrP exists in at least two isoforms; normal cellular prion protein denoted as $PrP^{C}$, and pathogenic misfolded prion protein denoted as $PrP^{SC}$; which share the same primary amino acid sequence but differ in secondary and tertiary structure. Crucially, misfolded $PrP^{SC}$ has been shown to catalyse the misfolding of benign $PrP^{C}$ molecules, thus

generating more copies of and triggering a chain-reaction of PrP misfolding and aggregation. $PrP^{SC}$ aggregates can fragment and spread within cells, throughout tissues, and between organisms [87].

Prion diseases, such as Creutzfeldt-Jakob disease, show stark similarities to other neurodegenerative disorders at the clinical level, with symptoms including dementia, ataxia, muscle incoordination, spasticity and weakness [87]. While the neuropathological features of prion diseases are variable, they usually include loss of neurons, gliosis, and accumulation of aggregated proteins [88]; common hallmarks present in many neurodegenerative diseases. Templated protein misfolding in a "prion-like" manner is increasingly implicated in other neurodegenerative diseases outside of the classical prion diseases (reviewed in reference [89]). The prion model is especially appealing when considering ALS, a disease often characterised by an initial focal onset that gradually spreads to contiguous areas of the brain and SC. Accumulating evidence suggests that prion-like propagation of SOD1 misfolding plays a role in SOD1-ALS [90].

Prion-like propagation is particularly common in amyloidogenic proteins; PrP is able to form amyloid-like aggregates, similar to the amyloid fibrils observed in AD and PD [91]. 'Amyloid' refers to distinctive protein assemblies with a highly stable cross-β quaternary structure. Amyloid fibrils grow by recruitment and self-templating of amyloidogenic proteins, and can fragment to seed amyloid formation at new sites within cells or tissues [92]. SOD1 has been shown to form amyloid-like fibrils *in vitro* [57,93–96] and SOD1 aggregates with an ordered molecular structure have been reported *in vivo* [97–100]. The structure of a recombinant amyloid fibril formed by SOD1 has recently been determined using cryo-EM [101].

A number of findings support prion-like propagation of SOD1: 1. Pre-aggregated SOD1 fibrils seed further SOD1 aggregation in neuronal cells *in vitro* [95,102]. 2. SOD1 aggregate seeding occurs between cells *in vitro* [103]. 3. Pre-aggregated $SOD1^{G93A}$ and SC homogenates from late-stage *SOD1*$^{G93A}$ transgenic mice can seed SOD1 pathology in healthy *SOD1*$^{G85R\text{-}YFP}$ transgenic mice [97,98,104]. 4. SOD1 aggregates spread through synaptically connected cells in the SC [98]. 5. Similar to PrP, different strains

of SOD1 aggregates with differing pathogenicity exist *in vivo* [99,100]. 6. SC homogenates from SOD1-ALS patients seed SOD1 aggregation in HEK cells [97,105] and in SOD1 transgenic mice [106,107].

Prion-like spreading of SOD1 pathology is a highly active area of research. Despite the above evidence, the question remains whether mutant SOD1 aggregates can seed aggregation of WT/mutant SOD1 expressed at physiological levels *in vivo.* Such experiments are difficult to perform with available transgenic mice, which generally overexpress SOD1 and display high SOD1 misfolding and aggregation at baseline. Physiological mouse models will be a valuable tool to address this question.

It is notable that TDP-43 and FUS also form amyloid-like fibrils [108–112], and TDP-43 is capable of seeding aggregation in cells and mice [113–117], suggesting that prion-like propagation of proteins may be a general mechanism applicable to many forms of ALS.

#### *1.3.2.3 Mitochondrial dysfunction*

MNs have high energy demands, and any disruption in mitochondrial function can lead to an energy deficit, contributing to MN degeneration. Mitochondrial structural abnormalities, such as vacuolisation, have been frequently observed in SOD1-fALS mouse models [118–122] and in various tissues from ALS patients, including liver, muscle, brain and SC [123–127]. Accordingly, mitochondrial function is severely compromised in MNs in SOD1-ALS [128].

It has been suggested that mutant SOD1 has an increased tendency to localise to the inner membrane space (IMS) of mitochondria, leading to impaired mitochondrial respiratory capacity [119,120,122,129–131]. Indeed, specific targeting of mutant SOD1 to the IMS leads to toxicity in neuronal cell lines [132,133] and mice – albeit with less severe neurological deficits than high copy SOD1-G93A mice, which have stronger transgene expression and no IMS targeting [134].

The importance of SOD1 localisation to the IMS is still debated, with other groups suggesting mutant SOD1 is deposited on the outer mitochondrial membrane (OMM) [135–138]. Mutant and WT SOD1

associate with different domains of OMM protein Bcl-2, with mutant SOD1 more prone to forming high molecular weight complexes with Bcl-2 [135]. Mutant SOD1 interacts with Bcl-2 to expose its BH3 death domain, leading to cellular toxicity [138].

SOD1 deposition in the IMS and OMM are not mutually exclusive, and it is possible that both play a role in mitochondrial dysfunction in SOD1-ALS [139]. Impaired mitochondrial function is detectable prior to symptom onset in many cases, suggesting it is an early mechanism in the disease cascade in SOD1-ALS.

#### *1.3.2.4 Oxidative stress*

Neurons are particularly susceptible to damage caused by oxidative stress [140]. SOD1 is the main antioxidant defence against superoxide radicals in the cytoplasm [141]. In SOD1-fALS patients, there is a consistent decrease of about 50% in SOD1 activity in tissues, regardless of the mutation's effect on SOD1 [48]. Therefore it follows that oxidative stress may play a role in the pathogenesis of SOD1-fALS. Indeed, oxidative damage of proteins, lipids, DNA, and mRNA is observed in the spinal cord of SOD1-fALS patients and some sporadic ALS cases post mortem. Similar markers are found in bodily fluids like urine, blood, and cerebrospinal fluid (CSF) [142,143]. Some oxidative stress biomarkers have been reported to correlate with disease severity [144–150].

Oxidative damage may be a direct result of loss of SOD1 or, more likely, a downstream effect of other disease processes as oxidative damage has been observed in mouse models that do not display SOD1 LOF [151–154]. Interestingly, changes in other antioxidant enzymes, such as peroxiredoxin (PRDX) family members and several isoforms of glutathione S-transferase (GST), have been observed in ALS mouse and cell models, suggesting compensatory mechanisms may be at play [155,156]. Oxidation of SOD1 has been observed in mouse models, which is important because oxidation of SOD1 enhances its monomerisation and subsequent aggregation [78,157–159], potentially forming a feedback loop of oxidative damage. Additionally, overexpression of the oxidation resistance 1 (OXR1) gene,

which protects against oxidative stress, is protective against neurodegeneration in *SOD1*$^{G93A}$ transgenic mice [160].

While the precise role of oxidative stress remains unclear, its involvement in various disease mechanisms underscores its importance in ALS.

#### *1.3.2.5 Endoplasmic reticulum (ER)-stress*

ER stress occurs when there is an imbalance between the protein-folding capacity of the ER and the load of unfolded or misfolded proteins entering the ER [161]. Misfolded mutant SOD1 can accumulate in the ER and trigger ER stress, resulting in the activation of two stress response pathways: the unfolded protein response (UPR) and the ER-associated degradation (ERAD) [41,162]. Molecular markers of UPR and ERAD activation have been observed in ALS patient tissues [163–166] and SOD1-fALS mice [167–170]. Mutant SOD1 has been proposed to induce ER stress through interaction with Derlin-1, resulting in ERAD dysfunction [171], and through disruption of ER-Golgi trafficking via interaction with COPII [172].

The UPR and ERAD ultimately function to reduce accumulation of misfolded proteins in the ER by reducing protein synthesis, enhancing protein-folding capabilities and targeting proteins for degradation. However, excessive ER stress and the prolonged activation of these pathways can contribute to cell dysfunction and neurodegeneration, as observed in ALS [162]. Notably, in transgenic mouse models of SOD1-fALS, the abundance of abnormal SOD1 protein due to overexpression could contribute to ER stress.

#### *1.3.2.6 Axonal transport defects*

Deficits in retrograde axonal transport have been reported in SOD1-fALS mouse models *in vivo* [173] – with defects detectable prior to symptom onset [174–176] – and in cultured neurons from SOD1-fALS mice [177–179]. A number of vesicle types are subject to impaired retrograde axonal transport in mouse models of SOD1-fALS, including late endosomes and acidic vesicles [178,179] and signalling endosomes containing TrkB – a receptor for Brain-derived neurotrophic factor (BDNF) [174]. Additonally,

neurotrophin-dependent modulation of axonal transport of signalling endosomes is impaired in MNs in *tgSOD1$^{G93A}$* mice [176,180].

Mutant SOD1 has been shown to directly interact with the dynein motor complex, providing a mechanistic link between SOD1 and impaired retrograde transport [173,181]. Blocking this interaction has been shown to reduce SOD1 toxicity *in vitro* [182]. Additionally, crossing SOD1-G93A transgenic mice with animals carrying a mutation in dynein heavy chain reduced neurodegeneration and prolonged lifespan [177], possibly by interrupting the SOD1-dynein interaction. Mutant SOD1 has also been shown to impair kinesin-dependent fast anterograde transport [183].

Defects in the axonal trafficking of mitochondria have also been reported in SOD1-fALS mice [174,175,184] and cellular models [185,186]. SOD1 may impair mitochondrial transport by disrupting the function of mitochondrial Rho GTPase 1 (Miro1) [186] – a regulator of mitochondrial transport [187,188] – although it is unclear if this results from a direct interaction. Mutant SOD1 has been shown to reduce the levels of Miro1; however, in a cellular model of SOD1-ALS, this transport impairment was successfully reversed through the overexpression of Miro1 [186]. Reduced Miro1 levels are observed in the spinal cord of ALS patients [189].

Impairments in axonal transport can result in various harmful consequences, such as abnormalities in the cytoskeleton and disruptions in mitochondrial function. It has been suggested that LMNs might be especially susceptible to these effects due to their unique morphology [190]. On the other hand, the absence of axonal transport deficits in *SOD1$^{G85R}$* mice – which display neurodegeneration – and presence in mice overexpressing WT human SOD1 – which do not display substantial neurodegeneration – suggest that axonal transport deficits are not necessary or sufficient to cause neurodegeneration [191].

#### 1.3.2.7 Glutamate excitotoxicity

Elevated levels of glutamate have been observed in the CNS of SOD1-ALS mouse models and in SOD1-fALS patients [192,193]. MNs are particularly vulnerable to glutamate excitotoxicity [194–196], with excess levels of glutamate leading to overstimulation of MNs and cellular toxicity. The cascade of toxic events within MNs is initiated by the prolonged activation of glutamate receptors, leading to the opening of mitochondrial ion channels, increased intracellular calcium levels, and the production of reactive oxygen species (ROS) [192,193].

The cause of increased extracellular CNS glutamate in ALS seems to be mainly due to reduced glutamate uptake. Decreased glutamate clearance, mediated by the glutamate transporter EAAT2, has been reported in the cortex and spinal cord of fALS and sALS patients [197–200], and SOD1-ALS mouse models in most [77,201–203], but not all [204], studies.

Notably Riluzole, the only approved ALS drug in the UK, acts to block glutamatergic neurotransmission by inhibiting glutamate release and modulating NMDA and AMPA receptors [205,206], providing strong evidence for a role of glutamate excitotoxicity in ALS.

#### 1.3.2.8 Non-cell autonomous mechanisms

Whilst MNs are the primary degenerating cell type in ALS, a large body of evidence suggests that other cells are also involved in disease pathogenesis. To identify participating cell types in SOD1-ALS researchers have generated mouse models with mutant SOD1 restricted to specific cell types: mouse models with neuron-specific expression of mutant SOD1 display mild phenotypes with varying degrees of MN degeneration [207–209]. Restricting expression of mutant SOD1 to astrocytes did not result in MN loss or motor phenotypes [210]. Restricting expression to skeletal muscle caused motor phenotypes, although studies did not agree on the presence of MN loss [211–213]. Additionally, the presence of WT microglia extended survival in *SOD1*$^{G93A}$ mice [214,215]. All of the above cell type-specific models display less severe phenotypes than mice with ubiquitous expression of SOD1-G93A, suggesting that different cell types conspire together to produce phenotypes in ALS.

Our understanding of the roles of different cell types in ALS is constantly evolving and a detailed discussion is beyond the scope of this thesis. This topic has been the subject of several excellent reviews [216–219].

### 1.3.3 SOD1 loss-of-function (LOF)

Whilst loss of SOD1 function is not causative for ALS, increasing evidence suggests that reduced SOD1 activity may modify phenotype in SOD1-ALS patients [48,220]. Not all SOD1 mutants cause a decrease in intrinsic enzyme activity, however the overwhelming majority of SOD1 mutant cell lines show reduced enzymatic activity resulting from a combination of reduced intrinsic activity, reduced protein stability and/or increased aggregation [48].

In rare cases where homozygous SOD1 truncating mutations have been described, patients display progressive loss of motor function as part of a distinct syndrome with some similarities to ALS [221–223]. This suggests that complete loss of SOD1 activity in humans results in a severe neurological phenotype, and supports data from *Sod1*$^{-/-}$ knockout mice that suggest SOD1 activity is important for maintaining motor neuron function [48]. It seems likely that both loss and gain of function mechanisms contribute to SOD1-ALS, and the heterogeneity seen among patients could possibly be explained by how individual mutations impact both the physiological functions and toxic potential of SOD1.

The impact of SOD1 LOF is relevant for antisense oligonucleotides (ASOs) targeting SOD1, the first of which, Tofersen, was recently granted accelerated approval by the FDA [224]. Tofersen has also been given a positive opinion by the European Medicines Agency (EMA) [225] and is under review by the National Institute for Health and Care Excellence (NICE) [226]. ASOs aim to reduce overall SOD1 levels and are not currently selective for the mutant protein. Thus, the effects of SOD1 LOF must be carefully considered. Heterozygous carriers of SOD1 truncating mutations have approximately half the SOD1 activity of wild-type controls and appear to be unaffected [221,222]. Thus, the minimum level of SOD1 required for full neurological function appears to be somewhere between 0% and 50% of normal levels. Current ASO silencing of SOD1 reduces CSF SOD1 levels by approximately 35% [227],

which appears to be well-tolerated in clinical trials [224,227,228]. More marked silencing of SOD1 is likely to increase adverse effects, and hence the development of ASOs selective for the mutant protein (e.g. $SOD1^{A4V}$) is highly desirable.

### 1.3.4 Relationship between SOD1-ALS and other forms of ALS

Multiple strands of evidence suggest that SOD1-ALS may be distinct from other forms of fALS and sALS. SOD1-ALS cases display a characteristic pattern of neuropathology that rarely overlaps with other forms of ALS. The majority of SOD1-ALS cases exhibit dorsal column degeneration accompanied by Lewy body-like and astrocytic hyaline inclusions, which are rare in other forms of ALS [64,229]. Additionally, SOD1-ALS cases have ubiquitinated skein-like inclusions that contain misfolded SOD1 but not TDP-43 [230]. TDP-43 aggregation is a hallmark of the disease and appears in almost all non-SOD1-ALS cases [230]. TDP-43 mislocalisation/aggregation is believed to be a central mechanism in ALS pathogenesis; hence, the conspicuous lack of TDP-43 pathology in SOD1-ALS would suggest that different disease mechanisms are at play.

However, the disease mechanisms in SOD1- and non-SOD1-ALS are unlikely to be mutually exclusive. Studies using conformation-specific antibodies have detected misfolded wild-type SOD1 in the spinal cord and CSF of non-SOD1-ALS patients [183,231–233], suggesting that SOD1 misfolding may contribute to pathology even in non-SOD1-ALS cases. There is also evidence to suggest that TDP-43 pathology may play a role in SOD1-ALS; pathological modification of TDP-43 has been observed in SOD1-ALS cell and mouse models, and in tissues from a $SOD1^{G17S}$ patient [234]. TDP-43 inclusions have also rarely been seen in SOD1-ALS mouse models [235] and SOD1-ALS patient tissues [236].

Whether SOD1 models are generally applicable to all forms of ALS remains heavily debated. While it is clear that models of SOD1-ALS do not faithfully recapitulate TDP-43 pathology, there are other convergent mechanisms between SOD1- and non-SOD1-ALS which may be accurately modelled. The majority of pre-clinical studies to date have been performed on SOD1-G93A mice, which may explain the lack of concordance between preclinical work and clinical trials.

## 1.4 C9orf72

In 2011, the landmark discovery was made that a hexanucleotide GGGGCC repeat expansion in the C9orf72 gene is the most common genetic cause of both ALS and FTD in western populations [20,21]. Repeat expansions in C9orf72 account for approximately 35% of fALS and 5% of sALS cases in Europe, but are substantially less common in Asian populations, being present in only 2% of fALS and less than 1% of sALS patients [15,237]. The majority of healthy individuals have ≤11 hexanucleotide repeats, while C9orf72-ALS/FTD patients carry repeat expansions ranging from around 30 repeats to, more commonly, hundreds or thousands of repeats [238].

The hexanucleotide repeat expansion is located within a non-coding sequence in the *C9orf72* gene (Figure 1-1), appearing either within intron 1, or within the promotor region, depending on the *C9orf72* transcript [238]. Where the repeats appear in intron 1 they are also present in the pre-mRNA.

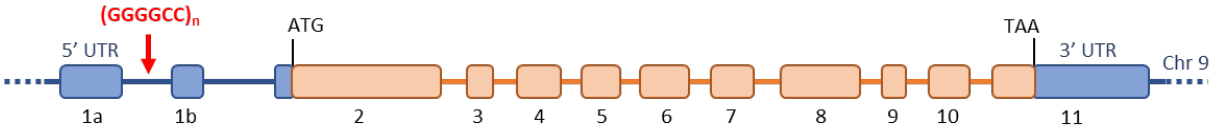



**Figure 1-1 | Schematic of the human *C9orf72* gene.** Coding exons are depicted in orange and non-coding exons are depicted in blue. The site of the GGGGCC hexanucleotide repeat expansion is indicated in red. UTR, untranslated region.

The physiological functions of the C9orf72 protein are not well understood, but evidence points to a role in autophagy. C9orf72 shares sequence homology with the Differentially Expressed in Normal and Neoplasia (DENN) family of proteins. DENN proteins act as Guanine nucleotide exchange factors (GEFs) for Rab-GTPases [239]. Similarly, C9orf72 is predicted to have GEF activity for a range of Rabs [240–243], many of which are involved in autophagy [240,244]. Rab-GTPases also play a critical role in the retrograde axonal transport [245,246]. C9orf72 has also been shown to regulate the activity of the ULK1 complex, an essential factor for autophagy initiation [242,243,247], and associate with p62, a well-characterized autophagy receptor [248]. Conflicting evidence exists as to whether C9orf72 also

regulates autophagy at the level of mammalian target of rapamycin (mTOR) signalling - which in turn regulates the ULK1 complex - with two studies finding that loss of C9orf72 led to reduced mTORC signalling, but reporting opposite effects on autophagic flux [249,250] and a third finding no effect [247].

Regardless of the precise molecular mechanisms, it is clear that C9orf72 is a regulator of autophagy. C9orf72 has also been linked to lysosomal function, endolysosomal trafficking and immune regulation [240,249,251–253].

Three main mechanisms have been proposed for how C9orf72 repeat expansions cause neurodegeneration – loss of function of the C9orf72 protein, toxicity caused by *C9orf72* repeat RNA, and toxicity from dipeptide repeat proteins (DPRs) produced from *C9orf72* repeat RNA.

### 1.4.1 Loss-of-function mechanisms

Examination of tissues from C9orf72-ALS/FTD patients has robustly shown that the repeat expansion reduces expression of *C9orf72* mRNA [254,255] and protein [256,257]. C9orf72 is also reported to be haploinsufficient in induced motor neurons from C9orf72-ALS patients, with haploinsufficiency suggested to contribute to neurodegeneration via several mechanisms: ineffective clearance of toxic DPR proteins produced from the repeat RNA, excitotoxicity caused by accumulation of glutamate receptors [253], and disruption of Ran-GTPase signalling an nucleocytoplasmic transport [258]. It is likely that *C9orf72* haploinsufficiency plays a modifying role in C9orf72-ALS/FTD rather than being a main driver of disease, as conditional knockout of the mouse C9orf72 ortholog in the CNS at embryonic day 10.5 did not cause neurodegeneration [259]. Whereas reduction in C9orf72 protein levels has been found to exacerbate disease phenotypes in mouse models expressing toxic *C9orf72* repeats [260,261].

### 1.4.2 Gain-of-function mechanisms

#### *1.4.2.1 RNA toxicity*

RNA toxicity plays a central role in the pathogenesis of other repeat expansion disorders such as myotonic dystrophy and fragile X-associated tremor, where RNA foci – which consist of repeat-

containing mRNAs in complex with RNA binding proteins – are observed in the nucleus and/or cytoplasm of affected cells [262,263]. Similarly, nuclear RNA foci are observed in the brain and SC of C9orf72-ALS/FTD patients [20,264–268], as well as in C9orf72-ALS/FTD patient induced pluripotent stem cells (iPSCs) [269], iPSC-derived MNs [264,270], and some C9orf72-ALS mouse models [271–275].

These RNA foci usually consist of a single species of repeat RNA transcript—sense or antisense—although co-occurrence, and rarely colocalisation, of sense and antisense RNA foci within neurons has been reported [266,267,276]. Cortical, frontal and hippocampal neurons display predominantly sense RNA foci, whereas cerebellar Purkinje cells and lower motor neurons show a predominance of antisense RNA foci [277,278]. *C9orf72* repeat RNA forms complex secondary structures, including hairpins, heterodimers and highly stable G-quadruplexes [279–281]. Antisense repeat RNA may also form i-motifs, although these structures have only been observed in *C9orf72* repeat RNA under conditions of low temperature and pH [279]. These secondary structures have been shown to bind and sequester RNA binding proteins (RBPs) [264,282].

Conflicting evidence exists regarding the pathogenicity of RNA foci in C9orf72-ALS/FTD [283]. Hexanucleotide GGGGCC repeat RNA is toxic in a number of *in vivo* models; degeneration of eye and motor neurons has been observed in a *Drosophila* model carrying 30 repeats [284], and in *zebrafish* injected with 35-90 GGGGCC repeat RNA [285]. Neither model had any detectable DPRs – although DPRs may have been present at levels below the threshold of detection. On the other hand, constructs designed to only allow repeat RNA but not DPR protein expression did not cause toxicity in a *Drosophila* model, even with large repeat sizes [286,287], strongly suggesting that RNA-foci alone do not cause degeneration.

Some studies have tried to correlate RNA foci burden with C9orf72-ALS/FTD clinical features. One small-scale study suggested that the appearance of RNA foci correlates with disease onset [276]. However, the largest study of C9orf72 RNA foci to date examined 63 C9orf72-ALS/FTD patients and did not find a positive correlation between RNA foci and clinico-pathological features [277].

The weight of evidence suggests that RNA foci alone are not toxic, however they may work in concert with other mechanisms in C9orf72-ALS. The key mechanism through which RNA foci may contribute to disease is via sequestration of RBPs, such as hnRNP H [288,289] and FUS [290], leading to perturbed RNA metabolism [283] – a central pathomechanism in many forms of ALS. Additionally, RNA foci may interfere with mRNA transport granules [291]. More data on RNA foci toxicity is needed in vertebrate models before solid conclusions can be drawn.

#### *1.4.2.2 DPR toxicity*

Since the GGGGCC repeats lie in the non-coding region of C9orf72, they are translated using an unconventional non-ATG translational process termed RAN translation. DPRs are generated from translation of the repeat in every reading frame, resulting in five distinct DPRs: poly-GA and poly-GR from the sense strand, poly-PA and poly-PR from the antisense strand, and poly-GP from both sense and antisense strands [238].

DPRs are widely believed to be the primary toxic species in C9orf72-ALS/FTD, exerting toxicity by directly binding to and sequestering other proteins. This interference disrupts various cellular processes including ribosome biogenesis, proteasome function, DNA repair and nucleocytoplasmic transport [292]. Nevertheless, the extent to which DPR-induced toxicity contributes to the progression of the ALS remains uncertain. Furthermore, the specific impact of each individual DPR on disease pathogenesis remains unclear, although this is beginning to be unravelled by the use of model systems that express individual DPRs. DPRs can be divided into arginine-rich DPRs and non-arginine DPRs, with multiple lines of evidence suggesting that arginine-rich DPRs are the most toxic species [238,293].

The arginine-rich DPRs, poly-GR and poly-PR, are promiscuous in their interactions with other proteins [294]. This is partly explainable by the positive charge and polarity conferred by abundant arginine residues. The arginine-rich DPRs are highly toxic both *in vitro* and *in vivo* [286,294–298], with a wide range of pathomechanisms proposed [292].

The liquid–liquid phase separation of stress granules is mediated via the low complexity domains (LCDs) of various proteins, including FUS. However, ALS-associated mutations in the FUS LCD increase its propensity to undergo pathological and irreversible liquid–solid phase transitions [299,300]. Arginine-rich DPRs have been shown to interact with the FUS LCD to similarly disrupt the dynamics of liquid–liquid phase separation in various membrane-less organelles [294,301] and the nucleolus [302]. Poly-PR has also been shown to bind to the LCD of nucleophosmin 1 (NPM1), sequestering it into large complexes and preventing it from carrying out its physiological functions of protein quality control and DNA damage repair [302–304]. Arginine-rich DPRs have also been shown to disrupt movement of various cellular cargo, including mitochondria and RNA granules, through their interaction with microtubule and motor proteins [305]. Arginine-rich DPRs can enter the nucleus – potentially by mimicking nuclear localisation sequences – impacting ribosomal RNA production and leading to reduced cell viability [306]. Poly-GR has also been shown to disrupt nucleocytoplasmic transport [296].

The non-arginine-rich DPRs, poly-GA, poly-PA and poly-GP, are uncharged and interact less readily with other proteins than the arginine-rich DPRs [294]. Among these, poly-GA is the most easily identifiable DPR within cytoplasmic inclusions [307–311] due to its higher translation efficiency and ability to aggregate into amyloid-like fibrils [312].

The degree to which poly-GA is toxic is heavily debated. Some studies have found poly-GA to be toxic *in vitro* [308,309,312] and in animal models [313,314], whereas others have not [286,294,295,297]. Poly-GA toxicity, to the extent at which it is present, seems to be largely attributable to its sequestration of other proteins such as the lipid-binding chaperone Unc119 [308] and the proteosomal subunit PSMC4 [315]. Poly-GA overexpression also modestly impairs DNA damage repair mechanisms *in vitro* [304,316].

Poly-GP and poly-PA appear to be the least toxic DPRs [292], they have a flexible coil structure and therefore cannot self-aggregate [294,317]. Poly-GP and poly-PA are non-toxic in a range of cell-based and *Drosophila* models [286,294,295,297].

The correlation between DPR pathology and neurodegeneration in humans is unclear. Despite being the most abundant DPR in C9orf72-ALS/FTD patient samples, poly-GA pathology does not appear to correlate with the degree of neurodegeneration [318,319]. An expanded study found no correlation between any DPR and neurodegeneration [310]. Furthermore, a systematic review found that DPR inclusions were less common in the spinal cord of C9orf72-ALS/FTD patients than in the brain, with inclusions commonly observed in the cerebellum – a brain region relatively spared in ALS/FTD [320]. On the other hand, more recent studies have found a correlation between Poly-GR pathology and neurodegeneration [321,322].

Some evidence suggests that DPR pathology precedes TDP-43 pathology, thereby playing a more important role in the earliest stages of disease [323,324]. Conversely, the patient samples examined for DPR pathology represent the latest disease stages and may not necessarily correlate with earlier disease processes. Animal models will be essential to address this issue as they can be examined throughout the entire disease course.

C9orf72-ALS/FTD patients will express a combination of all five DPR species, and thus it is important to understand how different DPRs interact together to give rise to disease, something that has rarely been investigated. Hence there is a need for *in vivo* models that express a combination of DPRs with physiologically relevant patterns and levels of expression.

## 1.5 Mouse models of ALS

### 1.5.1 The utility of mouse models for ALS research

For decades, mice have been the preferred animal model to study a range of neurodegenerative diseases, including ALS [325,326]. The utility of mice for scientific research lies primarily in our ability to edit their DNA; although it is possible to genetically manipulate other mammalian species, mice have often been the model organism in which genome editing approaches are pioneered and are currently the most genetically tractable mammal. More than 99% of human genes have a mouse

homologue [327] and the advent of modern gene editing tools such as CRIPSR allows us to edit mouse genes rapidly, with unprecedented precision, enabling large-scale study of gene function in health and disease. Furthermore the small size and fast generation time of mice allows for the creation of colonies of mutant mice on a controlled genetic background within a reasonable time-frame and budget.

Mouse models enable a systems-level approach to study neurodegenerative disease throughout the lifetime of the animal. It is well established that ALS is a multi-system disorder and that non-cell-autonomous mechanisms play a key role in the disease [60]. Mouse models enable the interplay between different cell types and tissues to be studied in the context of ALS, something that is not possible to the same extent in cell-based models.

Notwithstanding their advantages, it is clear that mice are not mini humans. Despite sharing homologues for the vast majority of genes, homologous mouse and human proteins share an average of 78% amino acid sequence identity [327] and crucial sequence differences exist in genes relevant to ALS. For example, human SOD1 forms aggregates more readily than mouse Sod1, a difference which can be attributed, at least in part, to the human-specific residues W32 and C111 [328–330].

Even more stark differences have been found between the regulation of gene expression in mouse and human tissues. The landscape of cis-regulatory elements differs substantially between our species [331,332] resulting in species-specific expression patterns of genes and isoforms involved in neurodegeneration. For example, the splicing of mouse and human TAU, encoded by *MAPT*, differs significantly and adult mice do not express the Tau 3R isoform [333]. This makes it impossible to examine Tau 3R:4R isoform ratios, which are believed to be dysregulated in FTD, in the context of mouse Tau [334].

Mice only live for around 2 years. This poses significant challenges for the investigation of late onset neurodegenerative conditions such as ALS, which typically arises around the 6th decade of life [2]. The development of ALS is thought to be a multistep process, with a gradual increase in insults to cellular function occurring years or decades prior to symptom onset [335,336]. Therefore time is a key risk factor for ALS, as the effect of any disease-causing genetic variant is likely to be cumulative. A direct conversion of 'mouse years' into 'human years' is not possible; a 70-year-old human and a 2-year-old mouse can both be considered aged, however the MNs of the 70-year-old human have had significantly more time to accrue cellular damage (e.g unfolded/aggregated proteins, DNA damage) than the MNs of the 2-year-old mouse.

How does one reconcile these differences and use mouse models to gain insight into human disease? The answer lies in selecting a mouse model that is tailored to answer the question of interest. Different mouse models have distinct characteristics and are poised to answer different questions. The failure of many studies stems from the fact that researchers expect a successful model to recapitulate every aspect of the disease; a one-size-fits-all approach that, due to the species-specific differences outlined above, is impossible to realise. As described by Fisher and Bannerman, the key is to "Know your question, know your mouse" [325].

### 1.5.2 Current mouse models of ALS

#### *1.5.2.1 Transgenic mouse models of ALS*

ALS is largely an autosomal-dominant disorder; hence the vast majority of existing mouse models of ALS were generated using insertional transgenesis to introduce mutant human fALS-associated genes into the mouse genome. Each transgenic mouse model of ALS has its own distinctive features arising from the nature of the genetic manipulation. The main factors which impact phenotypes in transgenic models are outlined below, and in Table 1-2.

| Type of model | Species of GOI | Copy number of GOI | Insertion/ mutation site | Regulatory sequences | Endogenous mouse gene intact? |
|---|---|---|---|---|---|
| **Transgenic** | Human | High | Random | Endogenous/exogenous human or mouse | Intact |
| **BAC transgenic** | Human | Low (~1-3 copies) | Random | Endogenous human | Intact |
| **ENU mutant** | Mouse | Physiological (1-2 copies) | Random | Endogenous mouse | Mutated |
| **Targeted point mutant** | Mouse | Physiological (1-2 copies) | Targeted | Endogenous mouse | Mutated |
| **KI to safe harbour locus** | Human | Low (~1-3 copies) | Targeted | Endogenous human | Intact |
| **Humanised mice** | Humanised | Physiological (1-2 copies) | Targeted | Endogenous human and/or mouse | Humanised |

**Table 1-2 | Classes of mouse models and their key features.** BAC, bacterial artificial chromosome; ENU, *N*-ethyl-*N*-nitrosourea; KI, knock-in.

#### 1.5.2.1.1 Transgene copy number

Transgene DNA tends to form concatemers which insert into the genome as multi-copy transgene arrays. At the extreme, concatemers of up to 1000 transgene copies have been reported following pronuclear injection of exogenous DNA [337]. ALS mouse models often carry many transgene copies, resulting in overexpression of the protein of interest.

For example, high-copy *SOD1*$^{G93A}$ transgenic mice (hereafter referred to as *tgSOD1*$^{G93A}$), the most commonly used model of ALS, carry ~25 copies of the human transgene and express mutant SOD1$^{G93A}$ protein at ~17-fold over the endogenous level. Accelerated protein deposition leads high-copy *tgSOD1*$^{G93A}$ mice to display a highly aggressive neurodegenerative phenotype [338]. In contrast, *tgSOD1*$^{G93Adl}$ mice, which carry only ~8-10 copies of the *SOD1*$^{G93A}$ transgene, display a milder phenotype and slower disease progression [339–341]. A similar reduction in phenotype severity is also observed in low-copy *tgSOD1*$^{G37R}$ mice [342].

Transgene overexpression can be advantageous as it can accelerate the neurodegenerative phenotype such that late-stage disease processes can be investigated within the lifespan of the mouse. However, overexpression makes it more difficult to study early disease processes –

mechanisms that are the root cause of later pathological changes and offer the most promising targets for disease modifying treatments.

Additionally, artefactual phenotypes can arise in transgenic models purely due to protein overexpression, and can be difficult to dissociate from the pathogenic effect of fALS mutations. For example, mice which overexpress WT human SOD1 at a similar level to the mutant protein in *tgSOD1$^{G93A}$* mice also develop an ALS-like phenotype [73,343] accompanied by cerebellar defects [344], suggesting that human SOD1 overexpression alone can give rise to neurodegeneration.

Another issue is that high-copy transgene arrays can be unstable and copy numbers may change during meiosis. Transgene copy number has a large impact on phenotype severity [339,340,342,345], so colonies of high-copy transgenic mice must be constantly monitored.

Issues regarding transgene copy number can be partially overcome by the use of bacterial artificial chromosome (BAC) transgenes, which are more likely to integrate as low copy transgene arrays (usually one to three copies) [346]. BAC transgenic models have been particularly widely employed for modelling C9orf72-ALS/FTD [238].

#### 1.5.2.1.2 Regulatory sequences

To ensure high levels of expression, many transgenes are engineered such that the gene of interest is under the control of a strong exogenous promotor such as Thy 1.2 or Prp [347]. However, ectopic temporal and/or spatial expression of transgenes can give rise to unexpected phenotypes. For example, mice expressing mutant TDP-43$^{A315T}$ under the control of the Prnp promotor displayed decreased survival due to severe gastrointestinal (GI) pathology resulting from TDP-43$^{A315T}$ overexpression in the lower GI tract, rather than MN degeneration [348,349].

Transgenes may also be placed under the control of the endogenous human promotor. This should theoretically lead to a more physiologically relevant pattern of expression compared to exogenous promotors. However, gene regulation varies greatly between humans and mice [331,332] and human

regulatory elements can drive different patterns of expression compared to their orthologous mouse counterparts.

#### 1.5.2.1.3 Transgene insertion site

In most transgenic models, injected transgene DNA randomly inserts into the mouse genome and can disrupt coding genes or essential elements. Transgene insertion sites are rarely reported, yet insertional mutagenesis and structural variation resulting from transgene integration are alarmingly common [350].

The SOD1$^{G93A}$ transgene was mapped to the distal end of chromosome 12 [351] with the exact integration site determined to be Chr12:97,165,800 in 2014 [352], two decades after the strain was first published.

#### 1.5.2.1.4 Orthologous mouse genes

In most ALS mouse models the transgene is integrated randomly into the genome and the orthologous mouse gene remains intact. This makes it difficult to investigate loss-of-function effects resulting from haploinsufficiency of fALS genes without a concurrent knockout of the mouse orthologue, something that is rarely attempted. Haploinsufficiency is believed to play a disease-modifying role in both SOD1-ALS [48] and C9orf72-ALS/FTD [253,261] and remains an important avenue of research.

### *1.5.2.2 Knock-in mouse models of ALS*

#### 1.5.2.2.1 Mutagenesis of mouse genes

One early method to edit the mouse genome was through the use of mutagens such as *N*-ethyl-*N*-nitrosourea (ENU), which produces near-random point mutations throughout the genome at a frequency of one mutation every 1–1.5 Mb [353]. In this technique, male mice are injected with ENU, causing them to produce mutant sperm. This sperm can be screened for interesting mutations, or the offspring assessed for interesting phenotypes. Despite the random nature of this method, an

ENU mutant was discovered which carries a *Sod1*$^{D83G}$ mutation identical to the fALS-associated gain-of-toxic-function *SOD1*$^{D83G}$ mutation [354]. Additionally, two *Tardbp* mutants (F210I and M323K) generated through ENU mutagenesis have been used to study TDP-43 function, aiding in the discovery of "skiptic exons" – a new category of splicing events controlled by TDP-43 [355].

Gene targeting approaches have been used to introduce specific mutations into the mouse orthologues of fALS-associated genes, including *Fus* and *Tardbp* [347]. One notable example is mice that recapitulate the human *TARDBP* Q331K mutation, which have revealed novel insights about TDP-43 misregulation in ALS/FTD [355,356].

Introduction of mutations into mouse genes can yield novel insights into ALS pathomechanisms. However, orthologous mouse and human genes often display unique gene structure, splicing patterns and protein biochemistry, which in turn impacts their potential pathogenicity. For this reason, gene targeting has rarely been used to introduce mutations into mouse *Sod1;* species-specific gene structure and protein biochemistry underpins SOD1-ALS, warranting the introduction of human *SOD1* in most cases.

#### 1.5.2.2.2 Knock-in into safe harbour loci

Knock-in (KI) models can be generated using homologous recombination to direct insertion of human genes to 'safe harbour' loci, such as the *Rosa26* locus [357]. This technique avoids issues associated with random insertion sites and high copy number in transgenic mice and can utilise BAC constructs carrying human introns and regulatory elements.

However, similar to transgenic models, targeted KI mice have intact orthologous mouse genes, preventing the study of gene LOF without further genetic manipulation. Furthermore the use of human promotors to express constructs out of a non-endogenous loci can lead to ectopic levels and/or patterns of expression. For example, *TARDBP* knock-in mice, generated by targeted

integration of a BAC containing the human *TARDBP* gene into the *Rosa26* locus, only express human TDP-43 at low levels [358].

#### 1.5.2.2.3 Genomically humanised mice

The advent of modern genome editing techniques such as CRISPR has allowed researchers to make precise changes to the mouse genome, including the replacement of mouse DNA with orthologous human sequences in a process termed genomic humanisation [359,360]. Mouse genes can be partially humanised – such as in FUSDelta14 mice, which express a partially humanised mutant *Fus* gene [361] – or fully humanised – such as the humanised *hSOD1, hTARDBP* and *hFUS* models recently published by our group [362]. Mouse regulatory sequences can be retained or replaced by human regulatory sequences, depending on the purpose of the model.

Genomically humanised mice have a unique combination of characteristics compared to other classes of mouse model:

1. They allow the study of proteins at physiological levels of expression, thereby avoiding artefacts arising due to transgene overexpression.
2. They allow the study of human proteins, with human biochemistry.
3. The orthologous mouse gene is replaced, allowing study of LOF phenotypes.

Genomically humanised mouse models also have a number of caveats:

1. They are more difficult to make.
2. Artefactual results may arise from altering mouse cellular pathways, necessitating the use of humanised control animals which express the wild-type human protein.
3. Physiological levels of expression may not lead to strong phenotypes as seen in some overexpression models.

### 1.5.3 Existing mouse models of SOD1-ALS

To date, almost 30 different SOD1 mouse models have been developed, including over 20 transgenic strains that overexpress mutant forms of human *SOD1* (Table 1-3)[347]. Whilst these mice display many phenotypic features associated with ALS, such as motor neuron degeneration, muscle denervation and progressive motor deficits, concerns remain regarding the ability of these models to accurately recapitulate human SOD1-ALS. Additionally, two knock-in strains with ALS mutations introduced into the endogenous mouse gene have been published (*Sod1*$^{D83G}$ and *Sod1*$^{G85R}$; Table 1-3).

| Strain name | Transgenic/ENU mutant/targeted mutant | Promotor | Protein | Disease progression | Survival (weeks) | Original reference |
|---|---|---|---|---|---|---|
| **WT hSOD1 (JAX 002297)** | Transgenic | hSOD1 | High hSOD1 expression with increased enzyme activity | Earlier signs: 30-50 weeks; MN loss: 20-30% at ~104 weeks | 75-104 | [338] |
| **WT hSOD1** | Transgenic | hSOD1 | hSOD1 expressed 50× higher than endogenous mSod1 but with low enzyme activity | Earlier signs: ~36 weeks; MN loss: ~41% at late stage | ~52 | [343] |
| **SOD1 A4V** | Transgenic | hSOD1 | hSOD1; low expression levels with no enzyme activity | Ab | Normal | [338] |
| **SOD1 G37R (JAX 008342)** | Transgenic | hSOD1 | hSOD1; low expression levels with no enzyme activity | Earlier signs: ~11-17 weeks; end stage paralysis | 25-29 | [363] |
| **SOD1 G37R** | Transgenic | hNFL | hSOD1 4× higher expression than endogenous SOD1. Expression restricted to neurons | Ab | Normal | [207] |
| **SOD1 H46R** | Transgenic | hSOD1 | hSOD1 | Earlier signs: 20 weeks | 24 | [364] |
| **SOD1 H46R/H48Q** | Transgenic | hSOD1 | hSOD1; high expression but inactive protein | Earlier signs: 17-26 weeks; end-stage paralysis: ~21 weeks | ND | [365] |
| **SOD1 H46R/H48Q/ H63G/H120G** | Transgenic | hSOD1 | hSOD1; high expression but inactive protein | Earlier signs: 35-52 weeks; end-stage paralysis: 34-52 weeks | 44-52 | [68] |
| **SOD1 L84V** | Transgenic | hSOD1 | hSOD1 | Earlier signs: 21-26 weeks; no end stage MN loss | 26-30 | [167] |
| **SOD1 G85R (JAX 008248)** | Transgenic (G85R/WT hSOD1) | hSOD1 | hSOD1; low expression | Earlier signs: 26-35 weeks; end-stage paralysis: 2 weeks after initial signs | 28-37 | [77] |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SOD1 Thy1.2-G85R (EGFP) bicistronic** | Transgenic | Thy 1.2 | hSOD1; high expression but low metabolic stability | Ab | Normal | [366] |
| **SOD1 G85R** | Transgenic | hSOD1 | hSOD1; high expression. Loss of endogenous SOD1 activity and mutant SOD1 | Earlier signs: 44-48 weeks; MN loss: significant; disease duration ~5-6 weeks | 50-54 | [74] |
| **Sod1 G86R (JAX 005110)** | Transgenic | mSOD1 | mSod1; increased endogenous expression | Earlier signs: 13-17 weeks | 17 | [367] |
| **Sod1 GFAP-G86R** | Transgenic | hGFAP | Increased mSod1 expression, correlates with copy number | Earlier signs: >70 weeks; no final MN loss | >70 | [363] |
| **SOD1 D90A** | Transgenic | hSOD1 | hSOD1; high expression and activity | Earlier signs: 52 weeks; 40% loss of ventral horn neurons; disease duration ~50 days | 61 | [368] |
| **SOD1 G93A (JAX 002726)** | Transgenic | hSOD1 | hSOD1; high expression | Earlier signs: 13-17 weeks; MN loss: 50% at ~17 weeks; end-stage paralysis: ~19 weeks | 17-26 | [338] |
| **SOD1 G93A (JAX 002300)** | Transgenic | hSOD1 | hSOD1; lower expression | Earlier signs: ~24-34 weeks; MN loss: >50%; end-stage paralysis: at 29-39 weeks | 29-39 | [338] |
| **SOD1 Thy1.2-G93A (EGFP) bicistronic** | Transgenic | Thy 1.2 | hSOD1; high expression | Ab | Normal | [366] |
| **SOD1 Thy1.2-G93A (JAX 008230)** | Transgenic | Thy 1.2 | hSOD1; low expression | Earlier signs: 54-104 weeks; MN loss: ~40% | 62-104 | [208] |
| **SOD1 G127X** | Transgenic | hSOD1 | hSOD1; low expression and no activity | Earlier signs: 35 weeks; 1/3 of mice show forelimb paralysis | 36 | [79] |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SOD1 L126Z** | Transgenic | hSOD1 | hSOD1; low expression | Earlier signs: 28-36 weeks; end-stage paralysis: ~39 weeks | ND | [369] |
| **SOD1 L126delTT** | Transgenic | hSOD1 | hSOD1; low expression | Earlier signs: 25-43 weeks; end stage paralysis | 26-45 | [370] |
| **SOD1 L126Z** | Transgenic | hSOD1 | hSOD1 | Earlier signs: 48 weeks | 51-57 | [71] |
| **SOD1 T116X** | Transgenic | hSOD1 | hSOD1; low expression | Earlier signs: 41 weeks; end stage paralysis | ~43 | [371] |
| **Sod1 D83G (JAX 020440)** | ENU point mutant | mSOD1 | mSod1; reduced protein levels and loss of activity | Earlier signs: 15 weeks; upper MN loss: 29 weeks ~23%; lower MN loss: ~23% by week 15-52 | ~88 | [354] |
| **Sod1 G85R (JAX 031343)** | Targeted point mutant | mSOD1 | mSod1; markedly reduced protein levels and complete loss of activity | Motor neuron loss not assessed; phenotype similar to Sod1-KO mice. | ~107 | [372] |
| **Sod1 A4V (JAX 031341)** | Targeted point mutant | mSOD1 | mSod1; expression nd | No overt phenotype | ND | Unpublished |
| **Sod1 G93A (JAX 031344)** | Targeted point mutant | mSOD1 | mSod1; expression nd | No overt phenotype | ND | Unpublished |
| **Sod1 H71Y (JAX 031342)** | Targeted point mutant | mSOD1 | mSod1; expression nd | Earlier signs: 16 weeks | ND | Unpublished |

**Table 1-3 | SOD1 mouse models.** Table of SOD1 mouse models adapted from De Giorgio et al., 2019 [347]. Ab, phenotype absent; MN, motor neuron; WT, wild type; ND, not determined; hSOD1, human SOD1; mSod1, mouse SOD1.

#### *1.5.3.1 tgSOD1$^{G93A}$ transgenic mice*

Shortly after the discovery of SOD1 as an ALS-linked gene [22] Gurney and colleagues created a transgenic *tgSOD1$^{G93A}$* mouse strain by introduction of a human SOD1 transgene carrying the fALS-associated G93A mutation. Four founder lines were established in the original work, with line G1 carrying ~18 copies of the transgene [338]. A subsequent expansion of the transgene array resulted in mice, designated G1H, carrying ~25 transgene copies [373]. G1H mice were maintained on a BL6/SJL background and reached end stage disease at around 139 days [373], compared with 170 days in the

original G1 mice [338]. G1H mice are often defined as Tg(SOD1*G93A)1Gur, hereafter termed *tgSOD1$^{G93A}$*, and are the most widely used mouse model of ALS.

*TgSOD1$^{G93A}$* mice have undergone an extraordinary amount of characterisation. A number of key behavioural and pathological milestones are outlined in Figure 1-2.

Some phenotypes are observed extremely early in *tgSOD1$^{G93A}$* mice; sensorimotor development is seemly impaired, with delayed righting and reduced hind-paw grasping observed in the first postnatal week [374]. Hyperexcitiability of hypoglossal MNs has been observed from postnatal day 4 (P4) [375]. Changes in neuromuscular junction (NMJ) structure have been observed as early as P14 [376].

A *Sod1$^{G93A}$* knock-in strain with the G93A mutation introduced into the mouse *Sod1* gene (JAX 031344) has been created but not published. This strain has no overt phenotype according to the JAX strain description.

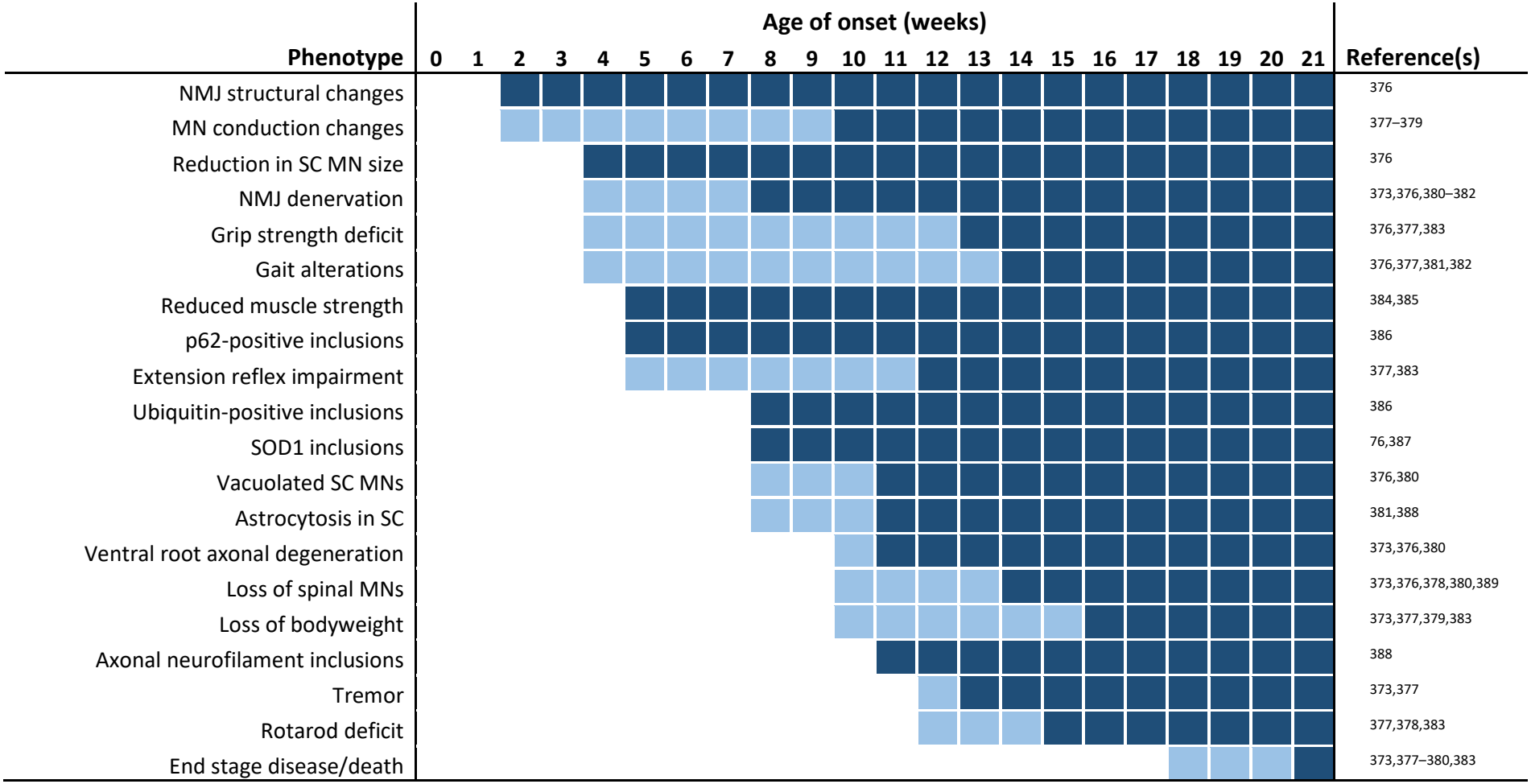



**Figure 1-2 | Phenotypes in *tgSOD1$^{G93A}$* mice.** Light blue indicates ages where phenotype was present in some studies. Dark blue indicates ages where phenotype was present in all studies referenced. MN, motor neuron; NMJ, neuromuscular junction; SC, spinal cord.

#### 1.5.3.2 *tgSOD1$^{A4V}$ transgenic mice*

A *tgSOD1$^{A4V}$* high-copy transgenic mouse strain was created alongside the *tgSOD1$^{G93A}$* strain by Gurney and colleagues [338]. However, very little work has been published using this strain and it reportedly has no overt phenotype [71,338]. Deng et al. report that *tgSOD1$^{A4V}$* transgenic mice do not develop disease even when aged to 24 months old, and do not have detectable SOD1 aggregation in the spinal cord. Crossing *tgSOD1$^{A4V}$* mice with mice overexpressing *SOD1$^{WT}$* reportedly triggered a disease phenotype and moderate reduction in lifespan to ~320 days [71]. However, this study suffered from a lack of *SOD1$^{WT}$* controls, especially as mice overexpressing wild-type human SOD1 have a similarly reduced lifespan [343].

The original study generated two *tgSOD1$^{A4V}$* founder lines, A1073 and A1074, which had 4.7 ± 0.4 and 3.2 ± 0.2 transgene copies respectively [338]. Both lines have substantially lower levels of mutant SOD1 protein compared to high copy line G1 *tgSOD1$^{G93A}$* mice, which may partially account for their lack of phenotype. Importantly, *tgSOD1$^{A4V}$* mice have not been evaluated using sensitive modern molecular biology techniques such as RNA-seq and proteomics. Hence it is possible that these mice have subtle molecular phenotypes that have hitherto gone undetected.

A *Sod1$^{A4V}$* knock-in strain with the A4V mutation introduced into the mouse *Sod1* gene (JAX 031341) has been created but not published. This strain has no overt phenotype according to the JAX strain description.

### 1.5.4 Existing BAC transgenic mouse models of C9orf72-ALS/FTD

Because C9orf72-ALS/FTD is caused by an intronic repeat expansion, rather than a mutation in a coding DNA sequence, the C9orf72 protein is not necessarily required to model the disease. Accordingly, a number of different classes of C9orf72-ALS/FTD mouse models have been made, some which express part or all of the C9orf72 protein plus the repeat [272–274,390], some which express only the GGGGCC repeat under the control of various promotors [271,275,391], and others which express individual DPRs through the use of codon-optimised expression constructs [298,314,392–394]. The diversity

of available mouse models is advantageous because, as described in section 1.5.1, different mouse models are suited to answering different questions. This section will focus on mouse models which express all or part of the human C9orf72 protein and hexanucleotide repeat, as these bear the closest resemblance to the humanised C9orf72 mouse model we are developing.

To date, four mouse models have been created through BAC transgenesis which express human *C9orf72* plus the repeat expansion (Table 1-4). The C9orf72 BAC transgenic models differ in the number of transgene copies, the size of the repeat expansion within each transgene, the size of the region of human C9orf72 incorporated, and the integration site. Therefore, it is perhaps unsurprising that phenotypic variability is seen between the strains. Intriguingly, C9-500 mice – which carry a single BAC transgene containing ~500 GGGGCC repeats – reportedly show a more severe phenotype than the three other lines, all of which carry multiple transgene copies.

| **Common name** | **Tg(C9orf72_3) line 112 (JAX 023099)** | **C9BAC (JAX 030222)** | **C9$^{450}$ (JAX 030581)** | **C9-500 (JAX 029099)** |
|---|---|---|---|---|
| **Full strain name** | C57BL/6J-Tg(C9orf72_i3) 112Lutzy/J; | B6J.Cg-Tg(C9orf72*)30 Rhbr/J | B6N.Cg-Tg(C9orf72*) 19Lagi/J | FVB/NJ-Tg(C9orf72) 500Lpwr/J |
| **Genetic background** | C57BL/6J | SJL/B6 | C57BL/6 | FVB/NJ |
| **Transgene integration site** | ND | Chr 4: 23,004,513-23,004,540 | ND | Chr 6: 114,939,853-114,939,873 |
| **Transgene copy number** | Multiple (exact no. unknown) | ~3-5 | Multiple (exact no. unknown) | 1 |
| **Number of GGGGCC repeats** | Each transgene carries ~100-1000 | Each transgene carries ~500 or ~300 | Each transgene carries ~450 or ~90 | ~500 |
| **C9ORF72 protein** | Entire hC9ORF72. ~1.75x level of mouse C9orf72. | Truncated hC9ORF72. Expression ND. | Truncated hC9ORF72. Expression: ~12x level of mouse C9orf72. | Entire hC9ORF72. Expression ND. |
| **RNA foci** | Sense, Y; Antisense, Y. | Sense, Y; Antisense, Y. | Sense, Y; Antisense, Y. | Sense, Y; Antisense, Y. |
| **DPR aggregates** | Poly(GP) | Poly(GP) | Poly(GA), poly(GP), poly(GR) | Poly(GA), poly(GP) |
| **Other inclusions/ aggregates** | TDP-43, N; p62, N; Ub, N. | TDP-43, N; p62, ND; Ub, ND. | TDP-43, N; p62, Y; Ub, ND. | TDP-43, Y; p62, Y; Ub, ND. |
| **Gliosis** | N | N | N | Y |

| | | | | |
|---|---|---|---|---|
| **NMJ denervation** | N | N | ND | Symptomatic mice: Y<br>Asymptomatic mice: N |
| **UMN loss** | ND | ND | N | Symptomatic mice: Y<br>Asymptomatic mice: N |
| **LMN loss** | N | ND | N | Y |
| **Survival (weeks)** | >78 | ~104 | >78 | Females: Overall ~70. Subset (~30%) <40.<br>Males: >80 |
| **Behavioural analysis** | No changes. | No changes. | Anxiety, spatial learning and memory deficits. | Gait abnormalities at 16 wks. Anxiety at >52 wks. Progressive kyphosis, reduced activity, seizures. |
| **Other phenotypes** | Nucleolar stress. | | Loss of neurons in dentate gyrus. | Reduced axon size in ventral lumbar SC roots. |
| **Original reference** | O'Rourke et al., 2015 | Peters et al., 2015 | Jiang et al., 2016 | Liu et al., 2016 |

**Table 1-4 | Published *C9orf72* BAC transgenic mouse models.** Y, phenotype present; N, phenotype absent; Ub, ubiquitin; ND, not determined; UMN, upper motor neuron; LMN, lower motor neuron; NMJ, neuromuscular junction; DPR, dipeptide repeat.

All four C9orf72 BAC transgenic strains display sense and antisense RNA foci, and aggregated poly(GP) DPRs [272–274,390]. However, characteristic ALS phenotypes have only been reported in C9-500 mice, including motor deficits, impaired survival, TDP-43 aggregation, gliosis, NMJ denervation and motor neuron loss [274,395]. Some mild cognitive phenotypes, including spatial learning and memory deficits, were observed in $C9^{450}$ mice [390], but these were markedly less severe than the phenotypes observed in C9-500 mice. The differences between the C9orf72 BAC transgenic strains are hard to reconcile and have been the subject of fierce debate. Two conflicting multi-centre studies describing the phenotype of C9-500 mice were published in 2020 [395,396].

Mordes et al. performed phenotyping on independent cohorts of C9-500 mice at Harvard University and Johns Hopkins University. They did not observe any difference in bodyweight, survival, grip strength or open field behaviour between C9-500 mice and non-transgenic controls up to one year old [396]. Additionally, the Jackson Laboratory did not detect disease-specific phenotypes in C9-500

mice generated at their facility (www.jax.org/strain/029099). In contrast, Nguyen et al collated findings from animals tested at the University of Florida (in the laboratory where the mice were originally characterised by Liu et al), the University of Bern and the University of Rochester, finding that C9-500 mice had reduced survival compared to non-transgenic littermates, and displayed ALS phenotypes at all centres [395].

One possible reason for the discrepancy between the two studies is that the housing environment substantially influenced the phenotype of C9-500 mice. Housing conditions and alterations in the microbiome has been shown to substantially impact the phenotype of laboratory mice [398–400], including mouse models of ALS [401].

Another possible explanation is that the FVB/N genetic background may have detrimentally affected the analysis. FVB/N mice have been associated with space cadet syndrome (SCS), which causes seizures and sudden death [402]. The original study by Liu et al subdivided their cohort into 'acute' vs 'slow progressive' animals, and in doing so may have selected for animals with SCS, which could have been more prevalent in the C9-500 group by chance. It is notable that phenotypes in C9-500 mice were more severe in female animals, as SCS is also more prevalent in females [402]. Seizures were also reported in C9-500 mice [274]; seizures are not typically associated with ALS but are associated with SCS, lending strength to the hypothesis that phenotypes in C9-500 mice are due to SCS.

On the other hand, Nguyen et al demonstrated that many of the phenotypes seen in C9-500 mice were ameliorated following reduction of poly(GA) levels using a GA1 antibody. Improvements were noted in survival, gait, open field behaviour, and MN loss following GA1 administration [403]; supporting the hypothesis that phenotypes in C9-500 mice are indeed due to the toxic properties of DPRs. Additionally, increased repeat length was associated with more severe phenotypes in C9-800 mice – a strain carrying 800 repeats, generated from a repeat expansion event in a C9-500 mouse – suggesting that the GGGGCC repeat is the pathogenic element [404]. Recent research has indicated that C9-500 animals have a neurodevelopmental phenotype, resulting in reduced thalamic and

cortical size in E18.5 embryos – although the impact of these changes on phenotypes in adult mice is unclear [405].

Currently, the above differences cannot be reconciled and further study of C9orf72-ALS/FTD mouse models is needed. It is possible that an interaction between the C9orf72 BAC transgene and centre-specific environmental conditions could explain the discrepancy. In light of the controversy, it is even more important to develop novel physiological models that accurately recapitulate C9orf72-ALS/FTD.

## 1.6 Research project aims

The above literature review highlights the need for novel ALS mouse models that accurately recapitulate aspects of human ALS, particularly early stage disease processes which are underway during the optimal window for therapeutic intervention. Genomically humanised mouse models, which express fALS genes with physiologically relevant patterns and levels of expression, represent a way to address current questions in the field while avoiding artefactual phenotypes arising from transgene overexpression.

This work will focus on the development and characterisation of 4 novel humanised fALS mouse models.

### 1.6.1 Chapter 3

| | | |
|---|---|---|
| **Primary research question** | – | Can human *SOD1* functionally replace mouse *Sod1*? |
| **Aim** | – | Characterise humanised *hSOD1*$^{WT}$ mice, an important control for mutant humanised *SOD1* mouse lines. |

### 1.6.2 Chapters 4 and 5

| | | |
|---|---|---|
| **Primary research question** | – | Can expression of mutant *hSOD1*$^{A4V}$ and *hSOD1*$^{G93A}$ at physiologically relevant levels recapitulate existing ALS phenotypes and reveal underlying mechanisms? |
| **Aim** | – | Characterise humanised *hSOD1*$^{A4V}$ mice, focussing on the development of ALS-relevant phenotypes. |
| | – | Develop and perform quality control of *hSOD1*$^{G93A}$ mice, for phenotyping at a later date. |

### 1.6.3 Chapter 6

| | | |
|---|---|---|
| **Primary research question** | – | Can physiological expression of *hC9orf72* plus a large GGGGCC expansion recapitulate existing ALS phenotypes and reveal underlying mechanisms? |
| **Aim** | – | Develop and perform quality control on *hC9orf72* mice, for phenotyping at a later date. |
| | – | Generate a targeting construct to introduce a large *C9orf72* repeat expansion into human iPSCs. |

# Chapter 2: Materials and Methods

## 2.1 Materials and Equipment

### 2.1.1 Equipment

| Equipment name | Manufacturer | Cat/part no. |
|---|---|---|
| **Phenotyping** | | |
| PhotoBeam Open Field | Omnitech Electronics | |
| Locotronic | Intellibio | |
| MouseWalker | Created in-house | |
| Rotarod | Ugo Basile | 47600 |
| Echo MRI | Echo Medical Systems | |
| Voluntary Running Wheels | TSE Systems | |
| Grip Strength Meter | Bioseb | BIO-GS3 |
| Phenomaster calormetric cages | TSE Systems | |
| Indirect Gas Calorimetry module | TSE Systems | |
| High-Speed Sensor Unit | Siemens | |
| AlphaTrak Glucometer | Zoetis | |
| Isometric force transducers | Dynamometer UFI Devices | |
| **Bacterial cell culture** | | |
| Innova 44 Incubator Shaker | New Brunswick Scientific | M1282-0002 |
| Electroporator. E.coli Pulser | BioRad | 165-2104 |
| Colourwave WPA colorimeter | Biochrom | CO7500 |
| **Mammalian cell culture** | | |
| Invitrogen™ Countess II Automated Cell Counter | Thermo Fisher Scientific | AMQAX1000 |
| Seahorse XFe96 Analyzer | Agilent | |
| FLUOstar Omega Microplate Reader | BMG Labtech | |
| **Molecular biology** | | |
| MicroCL 17R Microcentrifuge | ThermoFisher | 75002499 |
| Avanti JXN-26 High-Speed Centrifuge | Beckman Coulter | B37912 |
| Eppendorf ThermoMixer C | Eppendorf | 5382000031 |
| PowerPac Basic Power Supply | BioRad | 1645050 |
| Pippin Pulse electrophoresis power supply | Labtech International | PPI0200 |
| Mini-Sub Cell GT Electrophoresis Cell | BioRad | 1704468 |
| ChemiDoc XRS+ | BioRad | 1708265 |
| BluPad Blue & White LED Transilluminator | Cambridge Bioscience | BP001CU |
| 7500 Fast Real-Time PCR System | Applied Biosystems | 4351106 |
| 2100 Bioanalyzer | Agilent | G2939BA |
| NanoDrop 8000 Spectrophotometer | Thermo Fisher Scientific | ND-8000-GL |
| T100 Thermal Cycler | BioRad | 1861096 |
| iBlot Dry Blotting System | Invitrogen | IB1001 |

| | | |
|---|---|---|
| Odyssey® CLx | LI-COR | 9140 |
| **Histology** | | |
| CryoStar NX70 Cryostat | Thermo Fisher Scientific | |
| VT1000 S Vibrating blade microtome | Leica | 14047235613 |
| **Microscopy/imaging** | | |
| Axio Scan.Z1 Slide Scanner | Zeiss | |
| Nanozoomer slide scanner | Hamamatsu | |
| AxioObserver LSM 700 inverted confocal | Zeiss | |
| Plan-Apochromat 10x/0.3 M27 objective | Zeiss | |
| Plan-Apochromat 20x/0.8 M27 objective | Zeiss | |
| Plan-Apochromat 40x/1.4 Oil DIC M27 objective | Zeiss | |

**Table 2-1 | List of equipment**

### 2.1.2 Consumables

| **Consumable name** | **Manufacturer** | **Cat/part no.** |
|---|---|---|
| Semi Micro Cuvettes | Fisher Scientific | FB55147 |
| Electroporation Cuvettes, 1mm gap | ThermoFisher Scientific | P41050 |
| Gene Pulser/MicroPulser Cuvettes, 1mm gap | BioRad | 1652089 |
| SuperFrost Plus Adhesion slides | Fisher Scientific | 10149870 |
| Costar Assay plate | Costar/Corning | 3370 |
| Corning 96-Well Solid White Polystyrene Microplates | Fisher Scientific | 10022561 |
| Nunc 24 Well Multidish Nunclon Delta SI | SLS | 142475 |
| iBlot™ Transfer Stack, nitrocellulose, regular | Invitrogen | IB301001 |
| iBlot™ Transfer Stack, nitrocellulose, mini | Invitrogen | IB301002 |

**Table 2-2 | List of consumables**

### 2.1.3 Reagents

| **Reagent name** | **Manufacturer** | **Cat/part no.** |
|---|---|---|
| **General reagents** | | |
| Metaphosphoric acid | Merck | 239275 |
| Diethylenetriaminepentaacetic acid | Sigma-Aldrich | D6518 |
| Bovine serum albumin, Lyophilized powder | Merck | A9418 |
| Non fat dried milk from bovine | Sigma-Aldrich | M7409 |
| Triton X-100 | Fisher scientific | 11436631 |
| Tween 20 | Sigma-Aldrich | P9416 |
| EDTA | ThermoFisher Scientific | 17892 |
| 10% SDS | Fisher Scientific Ltd | 10265153 |
| Molecular Biology Grade Agarose | Severn Biotech Ltd | 30-10-50 |
| Pulsed Field Certified Agarose | BioRad | 1620137 |
| GelRed | Biotium | BT41003-5 |

| | | |
|---|---|---|
| GelGreen | Biotium | BT41005-5 |
| TURBO DNase I (2 U/µL) | ThermoFisher Scientific | AM2238 |
| Proteinase K Solution (20 mg/mL) | ThermoFisher Scientific | AM2546 |
| **Western blotting** | | |
| DC Protein Assay Reagents Package | BioRad | 500-0116 |
| cOmplete, EDTA-free Protease Inhibitor | Sigma-Aldrich | 11873580001 |
| NP-40 Surfact-Amps Detergent Solution | ThermoFisher Scientific | 85124 |
| Novex 14% Tris-Glycine Mini Gels, WedgeWell format | ThermoFisher Scientific | XP00145BOX |
| REVERT Total Protein Stain Kit | LI-COR | 926-11010 |
| Chameleon Duo Pre-stained Protein Ladder | LI-COR | 928-60000 |
| Protien loading buffer. 4x | LI-COR | 928-40004 |
| Novex Tris-Glycine SDS Running Buffer (10X) | ThermoFisher Scientific | LC2675 |
| Protein standard - 2 mg/vial BSA | Merck | P5619 |
| **SOD1 activity assays** | | |
| NativePAGE™ 3 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gel, 15-well | ThermoFisher Scientific | BN1003BOX |
| NativePAGE sample buffer | ThermoFisher Scientific | BN2003 |
| NativePAGE Running Buffer Kit | ThermoFisher Scientific | BN2007 |
| Potassium phosphate monobasic | Merck | P5655 |
| Nitrotetrazolium Blue chloride | Merck | N6876 |
| Riboflavin 5’-monophosphate sodium salt hydrate | Merck | F1392 |
| N,N,N’,N’-tetramethylethylenediamine (TEMED) | Merck | T9281 |
| Xanthine Oxidase from bovine milk | Sigma-Aldrich | X4500 |
| WST-1 | Abcam | ab155902 |
| Catalase from bovine liver | Sigma-Aldrich | C40 |
| Hypoxanthine | Sigma-Aldrich | H9377 |
| Sodium phosphate monobasic | Sigma-Aldrich | S8282 |
| Superoxide Dismutase bovine | Sigma-Aldrich | S9697 |
| **Muscle ATPase staining** | | |
| Sodium barbital | Sigma-Aldrich | B0500 |
| Sodium Acetate | Sigma-Aldrich | S2889 |
| Calcium Chloride | Sigma-Aldrich | C1016 |
| Cobalt(II) chloride | Sigma-Aldrich | 232696 |
| Ammonium sulfide, 20% solution | Sigma-Aldrich | A1952 |
| Adenosine 5'-Triphosphate Disodium Salt | Merck | A3377 |
| **RNA** | | |
| RNAlater Stabilization Solution | ThermoFisher Scientific | AM7020 |
| Rnase A (2.5ml 100mg/ml) | Qiagen | 19101 |

| | | |
|---|---|---|
| QIAzol Lysis Reagent | Qiagen | 79306 |
| UltraPure Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) | ThermoFisher Scientific | 15593031 |
| Fast SYBR Green Master Mix | ThermoFisher Scientific | 4385612 |
| Hi Scribe T7 high yield RNA syntheis Kit | NEB | E2040S |
| MEGAclear Transcription Clean-Up Kit | ThermoFisher Scientific | AM1908 |
| RNeasy Mini Kit | Qiagen | 74104 |
| RNeasy Lipid Tissue Mini Kit | Qiagen | 74804 |
| RNeasy Fibrous Tissue Mini Kit | Qiagen | 74704 |
| High-Capacity cDNA Reverse Transcription Kit | Applied Biosystems | 4368814 |
| Agilent RNA 6000 Nano Kit | Agilent | 5067-1511 |
| **DNA** | | |
| MyTaq Red 2x master mix | Meridian Bioscience | BIO-25043 |
| Q5 High fidelity polymerase 2 x master mix | NEB | M0492L |
| NEBuilder® HiFi DNA Assembly Master Mix | NEB | E2621S |
| QIAprep Spin Miniprep Kit | Qiagen | 27106 |
| HiSpeed Plasmid Maxi Kit | Qiagen | 12663 |
| Macherey Nagel NucleoBond Column BAC- Alternative Large Construct Kit | Machery-Nagel | 740579 |
| MyTaq Extract-PCR Kit | Meridian Bioscience | BIO-21127 |
| Nucleospin Gel and PCR Clean-up | Machery-Nagel | 740609.25 |
| Zymoclean™ Large Fragment DNA Recovery Kit | Zymo Research | D4045 |
| **IHC/IF** | | |
| 16% Formaldehyde (w/v), Methanol-free | ThermoFisher Scientific | 28908 |
| O.C.T. Compound | VWR | 361603E |
| Gallocyanine | Sigma-Aldrich | 124508 |
| Cresyl Violet acetate | Sigma-Aldrich | C5042 |
| α-Bungarotoxin, tetramethylrhodamine | Biotium | 12 |
| α-Bungarotoxin, Alexa Fluor 488 | ThermoFisher Scientific | B13422 |
| Mouse and Rabbit Specific HRP/DAB (ABC) Detection IHC kit | Abcam | ab64264 |
| Mouse on Mouse Polymer IHC Kit | Abcam | ab269452 |
| Aqua-Mount Mounting Medium | ThermoFisher Scientific | 13800 |
| ProLong Gold Antifade Mountant | ThermoFisher Scientific | P36930 |
| Prolong Glass Antifade Mountant with NucBlue | ThermoFisher Scientific | P36981 |
| **Bacterial cell culture** | | |
| Subcloning Efficiency DH5α Competent Cells | Thermo Fisher Scientific | 18265017 |
| NEB® 10-beta Electrocompetent E coli DH10B | NEB | C3020K |

| | | |
|---|---|---|
| ElectroMAX Stbl4 Competent Cells | Thermo Fisher Scientific | 11635018 |
| BigEasy®-TSA™ Electrocompetent Cells | Lucigen | 60224-2 |
| NEB® Stable Competent *E. coli* | NEB | C3040I |
| BigEasy® v2.0 Linear Cloning Kit (pJAZZ®-OK Blunt Vector) | Lucigen | 43036-1 |
| BigEasy® v2.0 Linear Cloning Kit (pJAZZ®-OC Blunt Vector) | Lucigen | 43018-1 |
| CopyRight PSMART BAC Kit | Lucigen | CB-18-03-45324 |
| Streptomycin Sulfate Salt | Sigma-Aldrich | S6501 |
| Chloramphenicol | Sigma-Aldrich | C0378 |
| Kanamycin Sulfate | ThermoFisher Scientific | 11815032 |
| Ampicillin | Sigma-Aldrich | A9393 |
| Tetracycline hydrochloride | Sigma-Aldrich | T3383 |
| Isopropyl-β-D-thiogalactoside | Sigma-Aldrich | 10724815001 |
| X-Gal Solution, ready-to-use-10 mL | ThermoFisher Scientific | R0941 |
| **Mammalian cell culture** | | |
| HBSS (10X) no calcium no magnesium no phenol red | ThermoFisher Scientific | 14185052 |
| B-27 Supplement (50X) serum free | ThermoFisher Scientific | 17504044 |
| Neurobasal Medium | ThermoFisher Scientific | 21103049 |
| Seahorse XFe96 FluxPak mini | Agilent | 102601-100 |
| Seahorse XF Cell Mito Stress Test Kit | Agilent | 103015-100 |
| Seahorse XF DMEM Medium, pH 7.4 | Agilent | 103575-100 |
| Seahorse XF 1.0 M Glucose Solution | Agilent | 103577-100 |
| Trypsin (2.5%) no phenol red | ThermoFisher Scientific | 15090046 |
| Bovine Serum Albumin solution | Merck | A1595 |
| Laminin Mouse Protein Natural | ThermoFisher Scientific | 23017015 |
| 500ML LEIBOVITZ L 15 MED | Fishersci | 15460554 |
| Neurobasal Medium | ThermoFisher Scientific | 21103049 |
| B-27 Supplement (50X) serum free | ThermoFisher Scientific | 17504044 |
| Horse serum heat inactivated | Merck | H1138 |
| Glutamax | ThermoFisher Scientific | 35050038 |
| 2-Mercaptoethanol (50 mM) | ThermoFisher Scientific | 31350010 |
| Penicillin-Streptomycin (5000 U/mL) | ThermoFisher Scientific | 15070063 |
| Rat CNTF | PeproTech | 450-50 |

| | | |
|---|---|---|
| Human/Murine/Rat BDNF | PeproTech | 450-02 |
| Murine GDNF | PeproTech | 450-44 |
| Potassium chloride | Sigma-Aldrich | 60130 |
| Calcium Chloride Dihydrate | Merck | C3881 |
| TMRE-Mitochondrial Membrane Potential Aβay Kit | Abcam | ab113852 |
| CyQUANT Cell Proliferation Assay for cells in cul | ThermoFisher Scientific | C7026 |
| CellTiter-Blue Cell Viability Assay | Promega | G8080 |
| CellTox Green Cytotoxicity Assay | Promega | G8741 |
| DNAse I,LYO., 100 MG | Merck | 11284932001 |
| Poly-L-Ornithine Hydrobromide, MW 30,000 | Merck | P3655 |
| Ammonium Chloride | Sigma-Aldrich | A9434 |
| Potassium Bicarbonate | Sigma-Aldrich | 237205 |
| Fetal Bovine Serum, qualified, heat inactivated | ThermoFisher Scientific | 10500064 |
| Dimethyl sulfoxide, Sterile-filtered | Merck | D2438 |
| **Enzymes** | | |
| T4 DNA Ligase 400,000 units/ml | NEB | M0202S |
| ElectroLigase | NEB | M0369S |
| CloneSmart DNA ligase | Lucigen | 43036-1 |
| AarI | ThermoFisher Scientific | ER1581 |
| XhoI | NEB | R0146S |
| XmaI | NEB | R0180S |
| BamHI - HF | NEB | R3136S |
| PpuMI | NEB | R0506S |
| SacI - HF | NEB | R3156S |
| XbaI | NEB | R0145S |
| FastDigest KflI | ThermoFisher Scientific | FD2164 |
| DpnI | NEB | R0176S |
| Guide-it Recombinant Cas9 (Electroporation-Ready) | Takara | 632641 |
| EnGen Lba Cas12a (Cpf1) | NEB | M0653T |
| **Primary Antibodies** | | |
| SV2-s Antibody Supernatant | DSHB | |
| 2H3-s Antibody Supernatant | DSHB | |
| Native SOD1 | Gift from Jonathan Gilthorpe | |
| Misfolded/aggregated SOD1 (mAb 85.11) | Gift from Jonathan Gilthorpe | |
| MAP2, RB X | Merck | AB5622 |
| Anti-Tau-1, clone PC1C6 | Merck | MAB3420 |
| Islet1 antibody | DSHB | 40.2D6-A |
| TUBULIN BETA III, CHKX | Merck | ab9354 |
| p75 NGF Receptor Polyclonal Antibody | Biosensis | M-1818-100 |

| | | |
|---|---|---|
| NeuN antibody | Merck | MAB377 |
| Anti-GFAP antibody | Abcam | ab7260 |
| **Secondary Antibodies** | | |
| Goat anti-Mouse IgG (H+L), Alexa Fluor Plus 488 | ThermoFisher Scientific | A32723 |
| Rabbit anti-Goat IgG (H+L), Alexa Fluor 488 | ThermoFisher Scientific | A11078 |
| IRDye 680RD Donkey anti-Mouse IgG (H+L) | LI-COR | 926-68072 |
| IRDye 800CW Goat anti-Mouse IgG (H+L) | LI-COR | 926-32210 |

**Table 2-3 | List of reagents**

### 2.1.4 Software

| **Name** | **Manufacturer/Reference** |
|---|---|
| 7500 Software v2.3 | Applied Biosystems |
| Image Lab v6.0.1 | BioRad |
| ZEN Black | Carl Zeiss software |
| ZEN Blue | Carl Zeiss software |
| GraphPad Prism 10 | Dotmatics |
| Empiria Studio | LI-COR |
| Image Studio Lite v5.2 | LI-COR |
| SPSS SamplePower v3.0 | IBM |
| ImageJ | National Institutes of Health |
| Fusion v6.5 SuperFlex Edition | Omnitech Electronics |
| Integrative Genomics Viewer v2.8.13 | Robinson et al., 2011 |
| Snapgene | Dotmatics |
| GSEA | Subramanian et al., 2005 |
| NDP.veiw 2 | Hamamatsu |
| Ilastik | Berg et al., 2019 |
| Picoscope v6 | Pico Technology |

**Table 2-4 | List of software**

### 2.1.5 Work conducted by other individuals/groups

Science is rarely a solo endeavour, and I have been privileged to work with a number of excellent collaborators. Below is an outline of the work described in this thesis which was conducted by other individuals/groups. In all cases, the work of others is highlighted in the relevant methods and/or results sections.

All animal husbandry and maintenance was performed by animal technicians at the MRC Harwell Institute. This included basic autopsy if an animal was found dead, or found sick and no tissues were

required. Trained animal technicians also performed the procedures for calorimetry, intraperitoneal glucose tolerance testing (IPGTT), automated wheel running, routine animal weight measurement, and blood sampling (not including processing of samples). All routine animal genotyping was performed by the GEMS team at the MRC Harwell Institute. Sectioning and staining of spinal cord samples for MN counting, and H&E staining of pre-sectioned muscle samples were performed by the histology team at the MRC Harwell Institute.

Axonal transport experiments were performed by Dr Andrew Tosolini and muscle isometric tension measurements by Dr Mhoriam Ahmed, tissues were extracted by me following these experiments.

Several molecular experiments were performed by external companies: Xdrop enrichment and Oxford NanoPore sequencing of the *hSOD1*$^{WT}$ locus was performed by Samplix; TLA and short read sequencing of the *hSOD1*$^{G93A}$ and *hC9orf72* loci were performed by Cergentis; 3' mRNA sequencing of *hSOD1*$^{A4V}$ samples were performed by Lexogen and data were analysed by Chloe Williams; PacBio sequencing of constructs was performed by CD genomics; Sanger sequencing was performed by Source Bioscience; metabolomics mass spectrometry experiments, including data pre-processing and annotation of metabolites, was performed by the Swedish Metabolomics Centre.

A number of cloning experiments described in Chapter 6 were performed by Dr Remya Nair, they are attributed to her in the results section describing the experiments.

## 2.2 Mice

### 2.2.1 Ethics statement

All animal husbandry and experiments were performed at the MRC Harwell Institute, and were in accordance with the Animals (Scientific Procedures) Act 1986, Home Office guidelines and the local ethics guidelines issued by the Medical Research Council. All reasonable measures were taken to minimise animal distress and suffering. Animal experiments were performed under Home Office Project License No. 20/0005.

### 2.2.2 Housing conditions and husbandry

Mice were housed in standard cages with *ad libitum* access to food and water, a 12h:12h light cycle with lights-on at 7am, and environmental temperature and humidity at 21±2$^{O}$C and 55±10% respectively. Mice were housed in cages of up to five mice, with single housing avoided wherever possible. All mice employed in the present work were maintained by animal technicians at the Mary Lyon Centre (MLC), MRC Harwell Institute. Detailed records of animal husbandry were stored on the Anonymus database (MRC Harwell Institute).

### 2.2.3 Humane end points

Animal health and wellbeing was assessed daily by qualified staff. Humane end points were defined as:

- Weight loss exceeding 15% of maximum bodyweight, or 20% of maximum bodyweight for mutants whose phenotypes include metabolic disturbances, not improving within 24 hours.
- Piloerection or other loss of condition not improving within 24 hours.
- Postural abnormalities not improving within a working day, including: Hunched appearance, tremors impinging on the ability to feed, ataxia impinging on the ability to feed or balance.
- Abnormal behaviour impinging on the welfare of the animal (e.g. lethargy, extreme hyperactivity), or excessive aggression to cage mates or handlers.
- Seizures not returning to normal within 2-3 minutes, or repeated seizures.
- Age-related loss of condition, such as that caused by age-related dermatitis, spontaneous tumours or reduced physiological function of vital organs.

### 2.2.4 Abbreviated genotypes

In the present work, mouse strains will be abbreviated as defined in Table 2-5 and mouse genotypes will be abbreviated as defined in Table 2-6.

| **Mouse strain** | **Abbreviated nomenclature** |
|---|---|

| | |
|---|---|
| Humanised *SOD1*$^{WT}$ | *hSOD1*$^{WT}$ |
| Humanised *SOD1*$^{A4V}$ | *hSOD1*$^{A4V}$ |
| Humanised *SOD1*$^{G93A}$ | *hSOD1*$^{G93A}$ |
| Transgenic *SOD1*$^{G93A}$ | *tgSOD1*$^{G93A}$ |
| Humanised *C9orf72* | *hC9orf72* |

**Table 2-5 | List of abbreviated strain names**

| Genotype | Abbreviated nomenclature |
|---|---|
| Wild-type mouse *Sod1* | *mSod1/mSod1* |
| Humanised *SOD1*$^{WT}$ heterozygous | *mSod1/hSOD1*$^{WT}$ |
| Humanised *SOD1*$^{WT}$ homozygous | *hSOD1* $^{WT}$*/hSOD1*$^{WT}$ |
| Humanised *SOD1*$^{WT}$ recombined allele heterozygous | *mSod1/hSOD1*$^{RDEL}$ |
| Humanised *SOD1*$^{WT}$ recombined allele homozygous | *hSOD1*$^{RDEL}$*/hSOD1*$^{RDEL}$ |
| Humanised *SOD1*$^{A4V}$ heterozygous | *hSOD1* $^{WT}$*/hSOD1*$^{A4V}$ |
| Humanised *SOD1*$^{A4V}$ homozygous | *hSOD1*$^{A4V}$*/hSOD1*$^{A4V}$ |
| Humanised *SOD1*$^{G93A}$ heterozygous with mouse *Sod1* | *mSod1/hSOD1*$^{G93A}$ |
| Humanised *SOD1*$^{G93A}$ heterozygous with human *SOD1*$^{WT}$ | *hSOD1*$^{WT}$*/hSOD1*$^{G93A}$ |
| Humanised *SOD1* homozygous knockout | *hSOD1*$^{KO}$*/hSOD1*$^{KO}$ |

**Table 2-6 | List of abbreviated genotypes**

### 2.2.5 C57BL/6J mice

An established C57BL/6J colony was maintained by animal technicians at the MRC Harwell Institute and routinely backcrossed to a Jackson Laboratory C57BL/6J stock strain (stock no. #000664) to prevent genetic drift. C57BL/6J was the background strain of experimental animals used for molecular work and longitudinal phenotyping.

### 2.2.6 129S9/SvEvH mice

An established 129S9/SvEvH colony was maintained by animal technicians at the MRC Harwell Institute. 129S9/SvEvH mice were used to investigate the effect of genetic background on the severity of the *hSOD1*$^{A4V}$ mutation.

### 2.2.7 Humanised *SOD1* mice

#### *2.2.7.1 Humanised hSOD1*$^{WT}$ *mice*

The official allele designation of the humanised *hSOD1*$^{WT}$ allele is Sod1$^{tm1.1(SOD1)Emcf}$ (MGI:6403195). The official strain name of the humanised *hSOD1*$^{WT}$ mice on a C57BL/6J background is B6.129-Sod1$^{tm1.1(SOD1)Emcf}$/H (MGI:6403204).

The targeting construct used for the generation of *hSOD1*$^{WT}$ mice was created by Dr Rosie Bunton-Stasyshyn and Dr Anny Devoy in the laboratory of Prof Elizabeth Fisher at UCL. Electroporation and screening of modified ESCs (embryonic stem cells) was performed in the laboratory of Dr Lydia Teboul at the MRC Harwell Institute. We have previously published a description of the generation of the *hSOD1*$^{WT}$ line [362], which is quoted verbatim below:

"A BAC targeting construct harbouring the mouse *Sod1* locus (BAC B6Ng01-068O19) was engineered to replace the mouse *Sod1* gene from the ATG start site (g.117) to the end of the 3'UTR (g.5583) with the orthologous human *SOD1* genomic sequence from a BAC harbouring the human *SOD1* locus (BAC RP11-535E10). LoxP sites were inserted into intron 3, and 1kb downstream of the inserted human 3'UTR. The sequence between the loxP sites was duplicated to allow conditional expression of a second copy of the sequence. An FRT-flanked Neomycin cassette was inserted between the first and second copies of the duplicated sequence. This construct was electroporated into the 129/SvJ mouse ESC line R1, humanising Sod1 in the mouse genome via homologous recombination. Correctly targeted clones were initially identified through Loss-Of-Allele (LOA) copy number qPCR assays. Mice were generated by injection of modified ESCs into BALB/cAnN donor blastocysts, with the resultant chimeric male offspring crossed to C57BL/6J females to obtain GLT. Following confirmation of GLT, hSOD1 heterozygous mice were backcrossed for one further generation to the C57BL/6J strain. Selection cassette removal was performed at generation N3 C57BL/6J backcrossed animals, via cytoplasmic injection of IVF derived 1-cell embryos with Flpo mRNA. The Neo-negative line was then backcrossed for one further generation onto C57BL6/J, followed by intercrossing to produce experimental animals at generation N4."

This line is freely available from the European Mouse Mutant Archive (EMMA, EM:13075).

### 2.2.7.2 *Humanised hSOD1$^{A4V}$ mice*

The official allele designation of the humanised *hSOD1$^{A4V}$* allele is Sod1$^{em1(SOD1)Emcf}$ (MGI:6856511). The official strain name of the humanised *hSOD1$^{A4V}$* mice on a C57BL/6J background is B6.129-Sod1$^{em1(SOD1)Emcf}$/H (MGI:6856512).

Editing and screening of hSOD1$^{A4V}$ zygotes was performed by Dr Lydia Teboul and colleagues at the MRC Harwell Institute. Briefly, the A4V point mutation (C->T at g.130) was introduced into the humanised *SOD1* allele via CRISPR/Cas9 editing in *hSOD1$^{WT}$* zygotes. A single-stranded oligodeoxynucleotide (ssODN) donor was used as a repair template. The CRISPR/Cas9 system components were delivered as ribonucleoprotein complexes into 1-cell stage hSOD1$^{WT}$ embryos by electroporation. Electroporated embryos were then re-implanted into CD1 pseudo-pregnant females. $F_0$ progeny were backcrossed for one generation onto C57BL6/J to obtain GLT, and the resulting $F_1$ offspring were screened for successful homology-directed repair (HDR) and integration of the A4V mutation by PCR followed by sanger sequencing (Table 2-7). To exclude off-target integration of the donor ssODN, copy counting of the donor sequence was carried out using ddPCR (Table 2-7). Following confirmation of successful targeting, founder hSOD1$^{A4V}$ mice were crossed to hSOD1$^{WT}$ mice to obtain compound heterozygotes (*hSOD1$^{WT}$/hSOD1$^{A4V}$*), which were then intercrossed to produce experimental animals. Experimental animals were considered to be on a mixed 129/SvJ x C57BL/6J background, being generation N4 on C57BL/6J. *hSOD1$^{A4V}$* mice were maintained on a C57BL/6J background.

The sequence of the sgRNA used in the generation of the *hSOD1$^{A4V}$* line was: 5'-GTCGCCCTTCAGCACGCACACGG-3'. The sequence of the ssODN donor was 5'-TTCCTGCGGCGCCTTCCGTCCGTCGGCTTCTCGTCTTGCTCTCTCTGGTCCCTCCGGAGGAGGCCGCCGCGCGTCTCCCGGGGAAGCATGGCGACGAAGG**T**CGTGTGCGTGCTGAAGGGCGACGGCCCAGTGCAGGGCATCATCAATTTCGAGCAGAAGGCAAGGGCTGGGACGGAGGCTTGTTTGCGAGGCCGCTCCCA-3' (mismatch to introduce C->T point mutation underlined).

| Assay | Primer/Probe | Sequence (5'→ 3') |
|---|---|---|
| PCR across A4V mutation | Forward primer | GCCGACGTCACAGTTAGAAGAC |
| | Reverse primer | ACCCTCAGGAACGAGAACAC |
| Sequencing across A4V mutation | Sequencing primer | CGGGAACTTTCTCAGTCCGC |
| | Sequencing primer | CTCAGCACTTGGGCACCG |
| ddPCR (mutation nonspecific) | Forward primer | CGTGTGCGTGCTGAAGG |
| | Reverse primer | AACAAGCCTCCGTCCCA |
| | 5' FAM labelled probe | CAATTTCGAGCAGAAGGCAAGGGC |
| ddPCR (specific for *hSOD1$^{A4V}$* allele) | Forward primer | TTCCGTCCGTCGGCTTC |
| | Reverse primer | CTTCAGCACGCACACGA |
| | 5' FAM labelled probe | TCGTCTTGCTCTCTCTGGTCCCTC |

**Table 2-7 | PCR primers, sequencing primers, and probes for screening of *hSOD1$^{A4V}$* and hSOD1$^{-/-}$ mice**

#### *2.2.7.3 Humanised hSOD1$^{A4V}$-129 mice*

The *hSOD1$^{A4V}$-129* strain was generated by crossing *hSOD1$^{A4V}$* mice (on a mixed 129/SvJ x C57BL/6J background) onto 129S9/SvEvH for 2 generations.

#### *2.2.7.4 Humanised hSOD1$^{G93A}$ mice*

The generation of the *hSOD1$^{G93A}$* line is described in detail in chapter 5. *hSOD1$^{G93A}$* mice were maintained on a C57BL/6N background.

#### *2.2.7.5 Humanised hSOD1$^{-/-}$ knockout mice*

Humanised *hSOD1$^{-/-}$* knockout mice were obtained by screening for nonsense mutations in F1 founder mice from the CRISPR targeting to introduce the A4V mutation. The screening was performed by Dr Lydia Teboul and colleagues at the MRC Harwell Institute. F1 animals were screened for indels by PCR followed by sanger sequencing using existing primers (Table 2-7). Copy counting of the donor sequence was carried out using ddPCR (Table 2-7).

A founder mouse was identified which carried a 28 nucleotide deletion in exon 1 of the humanised *hSOD1* gene (g.126-g.154) which was predicted to result in a frameshift and premature stop codon, yielding a non-functional 22aa polypeptide (Figure 2-1). These *hSOD1$^{-/-}$* mice are coisogenic with our *hSOD1$^{A4V}$* line.

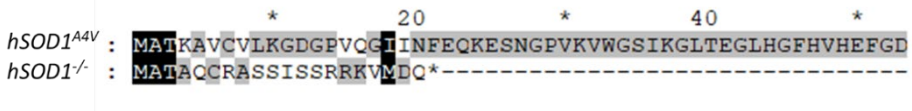



**Figure 2-1 | Predicted truncated SOD1 protein sequence in *hSOD1*$^{-/-}$ mice**

#### *2.2.7.6 Humanised hC9orf72 mice*

The generation of the *hC9orf72* line is described in detail in chapter 6. *hC9orf72* mice were maintained on a C57BL/6N background.

## 2.3 Mouse experimental procedures

### 2.3.1 Genotyping

All animal genotyping was performed on genomic DNA extracted from ear clip biopsies. Genotyping for the presence of the *hSOD1*$^{WT}$, *hSOD1*$^{A4V}$, *hSOD1*$^{G93A}$ and *hSOD1*$^{KO}$ alleles were performed by the GEMS team at the MRC Harwell Institute using Loss-Of-Allele copy number qPCR with the below primers and probes:

| Allele | Forward primer 5' -> 3' | Reverse primer 5' -> 3' | 5'-FAM labelled probe 5' -> 3' |
|---|---|---|---|
| *mSod1* | GTACCAGTGCAGGACCTCAT | AGCGTGCTGCTCACCTCT | AACATGGTGGCCCGGCGGATG |
| *hSOD1*$^{WT}$ | GTGCAGGTCCTCACTTTAATCC | CCAGAAAGCTATCGCCATTATTACAAG | CCAAAGGATGAAGAGAGGTAACAAGATGC |
| *hSOD1*$^{A4V}$ | TTCTCGTCTTGCTCTCTCTG | CCTTCTGCTCGAAATTGATGATG | ACGCACACGACCTTCG |
| *hSOD1*$^{G93A}$ | GTGACTGCTGACAAAGATGC | CAATGATGCAATGGTCTCCTG | ACAGAATCTTCAATAGACACATCGGCCAC |
| *hSOD1*$^{KO}$ | CCGTGTGCGTGCTGAAG | AACAAGCCTCCGTCCCA | TCAATTTCGAGCAGAAGGCAAGGGC |

**Table 2-8 | Genotyping primers and probes**

Genotyping for the presence of a recombined allele following Cre recombination was performed using PCR with the following primers:

| Recombined allele | Forward primer 5' -> 3' | Reverse primer 5' -> 3' | 5'-FAM labelled probe 5' -> 3' |
|---|---|---|---|
| *hSOD1*$^{RDEL}$ | GGCAATGTGACTGCTGACAAA | GCGGCCAATGATGCAATGGT | ATGGTGTGGCCGATGTGTCTATTGAA |

**Table 2-9 | Genotyping primers and probes against recombined *hSOD1*$^{WT}$ allele**

### 2.3.2 Phenotyping

All phenotyping was performed in the light phase (7am-7pm) – with the exception of tests that ran overnight. For all tests, mice were acclimatised to the testing room for at least 30 minutes prior to starting the procedure. Phenotyping was carried out on both male and female mice.

#### *2.3.2.1 Marble burying test*

Standard IVC cage bases were filled with Aspen shavings to approximately 5cm depth. Nine glass marbles were placed in an even 3 x 3 grid pattern on the shavings. A single mouse was placed in each cage for 30 minutes. After 30 minutes the mouse was removed and the number of moved/buried marbles was assessed. Marbles were counted as buried if more than 2/3 of the marble was covered by shavings.

#### *2.3.2.2 Open field test*

Mice were placed at the edge of a square arena measuring 50cm x 50cm (Omnitech Electronics) illuminated to 150-200 lux. Horizontal movement and rearing were recorded via beam breaks over a 20 minute period, following which the mice were returned to their home cage. The first 5 minutes of the test was excluded from the analysis to account for differences in handling. The periphery was defined as the area 10 cm from each wall.

#### *2.3.2.3 Combined SHIRPA and dysmorphology*

The SHIRPA and combined dysmorphology stage 1 observation screen was performed as described in Rogers et al [409]. Briefly, individual mice were removed from their home cage, placed into a clear Perspex viewing container with a wire mesh floor, and observed for approximately 2 minutes to assess natural behaviour and gross morphological abnormalities. The animal was then transferred to a 55 cm x 33 cm arena divided into 11 cm x 11 cm squares and locomotor activity – the number of squares fully entered by the animal – was recorded over 30 seconds. Auditory startle response to a high-frequency tone was assessed. Next, the animal was lifted by the base of the tail for 10 seconds to assess trunk curl and limb grasping, then gently lowered onto a platform to assess visual placing.

The animal was scruffed to assess morphological abnormalities of the paws, skin, mouth and genetalia, and to assess body and limb tone. Next the animal was placed facing downwards on a vertical wire grid approximately 15 cm above a platform to assess negative geotaxis – the ability of the animal to grip and climb down the mesh. Finally, the animal was placed into a clear Perspex tube and rotated $180^{O}$ to assess righting reflex. Overall, thirty-seven individual measurements were recorded for each animal, including measurements of behaviour, motor function, morphology and general wellbeing. The parameters recorded are listed in Table 2-10; parameters of particular relevance to neuromotor conditions are highlighted in bold.

| Parameter | Possible values |
|---|---|
| Tail morphology | as expected, not as expected |
| Coat | as expected, not as expected |
| Forelimb morphology | as expected, not as expected |
| Hindlimb morphology | as expected, not as expected |
| Vibrissae | as expected, not as expected |
| Head morphology | as expected, not as expected |
| Ear morphology | as expected, absent, large, small |
| Snout Size | as expected, decreased, increased |
| **Activity (body position)** | as expected, inactive, increased activity |
| **Tremor** | absent, present |
| Faecal pellet | as expected, hard, no data, soft |
| **Feet position in jar** | feet on bars, feet falling through |
| Transfer Arousal | as expected, extended freeze, immediate movement |
| **Locomotor Activity** | Numeric |
| **Gait** | fluid movement, lack of fluidity |
| **Tail Elevation** | as expected, dragging, no data, straub/elevated tail |
| Auditory Startle Response | present, no data, none |
| **Trunk Curl** | absent, no data, present |
| **Limb Grasp** | absent, no data, present |
| Visual Placing | upon vibrissae contact, no data, none, upon nose contact |
| **Body Tone** | as expected, flaccid, increased resistance |
| **Limb Tone** | as expected, increased, no data, reduced |
| Paw morphology | as expected, not as expected |
| Digit morphology | as expected, not as expected |
| Nails | as expected, not as expected |
| Skin | as expected, not as expected |
| Teeth | as expected, not as expected |
| Genitalia morphology | as expected, not as expected |
| Mouth Morphology | as expected, not as expected |
| Lip Morphology | as expected, not as expected |
| Salivation | none, saliva around margins of mouth, wet area around entire mouth |
| **Contact Righting** | present, absent, no data |
| **Negative Geotaxis** | as expected, falls, no data, slip/loss of grip |
| Vocalization | as expected, not as expected |
| Aggression | as expected, excessive aggression |
| Head Bobbing | absent, present |
| Unexpected Behaviours | none, circling, jumping, retropulsion |

**Table 2-10 | List of SHIRPA parameters**

#### 2.3.2.4 Locotronic test

The Locotronic test is a measure of motor coordination. The Locotronic test apparatus (IntelliBio) consists of a horizontal ladder above a series of optical beams. Limbs that fall between the rungs of the ladder are recorded as beam breaks. During the test, mice were placed at one end of the Locotronic apparatus and allowed to walk freely along the ladder to a covered dark zone at the opposite end. Leg errors were recorded as beam breaks under the rungs. Mice were tested 3 times over two consecutive days and the median number of leg errors calculated. Runs >30s long were excluded from the analysis.

#### 2.3.2.5 MouseWalker

The MouseWalker system is used to examine gait in freely walking rodents, similar to the footprint test [410]. The MouseWalker apparatus is described in Mendes et al., 2015 and consists of an illuminated acrylic glass catwalk with a high speed camera to record mouse foot placement.

The MouseWalker procedure was performed in the dark. Mice were placed at the end of the catwalk and allowed to walk freely along it for 1 minute. Videos captured by the camera were analysed frame by frame to quantify gait parameters (see Table 2-11 for list of gait parameters).

| Measure | Definition of measure | Units |
|---|---|---|
| Walking speed | The average velocity of the mouse during bouts of movement. | mm/s |
| Forelimb stance duration | The amount of time that each front paw is in contact with the ground; reported as an average across both forelimbs. | s |
| Forelimb step length | The distance between the foot-strike positions the two front paws on consecutive steps. | mm |
| Forelimb stride length | The distance that each front paw travels for a stride, from toe-off to foot-strike; reported as an average across both forelimbs. | mm |
| Forelimb stride width | The lateral distance between the foot-strike positions of the two front paws. | mm |
| Forelimb swing duration | The amount of time that each front paw is off the ground; reported as an average across both forelimbs. | s |
| Forelimb duty factor | The stance duration of a paw divided by the time taken for a full stride; reported as an average across both forelimbs. | none |
| Hindlimb stance duration | The amount of time that each hind paw is in contact with the ground; reported as an average across both hindlimbs. | s |
| Hindlimb step length | The distance between the foot-strike positions the two hind paws on consecutive steps. | mm |
| Hindlimb stride length | The distance that each hind paw travels for a stride, from toe-off to foot-strike; reported as an average across both hindlimbs. | mm |
| Hindlimb stride width | The lateral distance between the foot-strike positions of the two hind paws. | mm |
| Hindlimb swing duration | The amount of time that each hind paw is off the ground; reported as an average across both hindlimbs. | s |
| Hindlimb duty factor | The stance duration of a paw divided by the time taken for a full stride; reported as an average across both hindlimbs. | none |

**Table 2-11 | List of gait analysis parameters**

#### *2.3.2.6 Accelerating Rotarod test*

The accelerating Rotarod test (4–40 r.p.m. over 5 min; Ugo Basile) consisted of 6 trials performed on three days per week over two consecutive weeks. Mice were given one trial per day. The first 4 trials were considered training and no results were recorded. From the final two trials, the longest trial time was recorded. The end of the trial was classified as the time when the mouse fell from the rotarod or made three consecutive passive rotations. The maximum trial duration was 5 minutes.

#### 2.3.2.7 Grip strength

Grip strength was measured using a grip strength meter (Bioseb). To measure forelimb grip, mice were lifted gently by the base of the tail and held with their forelimbs in contact with the grid of the grip strength meter. Mice were then pulled horizontally away from the mesh in one smooth motion without excessive force and the resulting grip strength value was recorded in grams. To measure forelimb plus hindlimb grip the same procedure was applied, however mice were now allowed to grip the grid with all four limbs. Three measurements of forelimb grip and three measurements of forelimb plus hindlimb grip were collected from each mouse at each time point.

To minimise interoperator variability, at each phenotyping time point the same researcher performed all grip strength measurements. For forelimb, and forelimb plus hindlimb grip strength the median measurement from each animal was divided by the weight of the animal and the resulting value was used for analysis.

#### 2.3.2.8 Automated wheel-running

Automated home-cage wheel-running is a highly sensitive test which can detect very early motor deficits in mouse models of neurodegenerative disease [412]. This system has previously been used to detect subtle motor deficits in *tgSOD1*$^{G93A}$ mice at 12 weeks of age, approximately 3 months earlier than defects are detectable in grip strength analysis [412].

Automated wheel-running was measured using the PhenoMaster voluntary running wheel apparatus (TSE systems) by trained staff at the MLC. Mice were single housed in IVC cages containing a zero-resistance running wheel for 2 or 3 weeks. For phenotyping hSOD1$^{WT}$ animals at 6mo and 12mo, a 3-week protocol was used. For the first 2 weeks the wheel was a standard wheel with all rungs present. For the 3$^{rd}$ week, the standard wheel was replaced with a complex wheel which only has half the full complement of rungs, arranged in a consistent semi-random pattern. For phenotyping of *hSOD1*$^{WT}$ animals at 18mo, and phenotyping of *hSOD1*$^{A4V}$ animals, a modified 2-week protocol was used, where animals had access to the standard wheel for 1 week and the complex wheel for the 2$^{nd}$

week. Female mice were tested at all timepoints; male mice were only included at the final timepoint due to the requirement to single house the animals.

Only data from the dark phase (19:00-7:00) was included in the analysis. Parameters analysed were maximum speed, total distance run and average run duration.

#### *2.3.2.9 Echo MRI*

Echo MRI measures body composition in conscious animals; providing measurements of whole body fat mass, lean mass and water content. Echo MRI was performed using a mouse Echo MRI machine (Echo Medical Systems). Mice were then weighed and returned to their home cage following testing.

#### *2.3.2.10 Indirect calorimetry*

All calorimetry experiments were performed by trained staff at the MLC. Female animals were weighed and singly housed overnight in PhenoMaster calorimetric cages connected to an Indirect Gas Calorimetry module (TSE systems). Temperature was monitored and heat production normalised to the bodyweight of the animal. Oxygen consumption and carbon dioxide production were measured using a High-Speed Sensor Unit (Siemens) and used to calculate the respiratory exchange ratio (RER). Locomotor activity was measured by horizontal beam breaks.

#### *2.3.2.11 Intra-peritoneal glucose tolerance testing (IPGTT)*

IPGTT testing was performed by trained staff at the MLC. Mice were fasted for 6 hours (for *hSOD1$^{A4V}$-129* mice) or overnight (for *hSOD1$^{A4V}$* mice on a C57BL/6J background) prior to the test. EMLA cream was applied to the tail 30 minutes prior to the test. Blood was sampled from the lateral tail vein and glucose measurements performed using an AlphaTrak meter and AlphaTrak test strips. For the T0 baseline measurement, blood glucose was measured prior to glucose administration. Glucose solution was then administered via IP injection at a dose of 2 g/kg. Blood glucose was measured 15, 30, 60 and 120 minutes after glucose administration. Statistical analysis was

performed on data from the individual sampling time points and on the area under the curve (AUC) across all time points.

#### *2.3.2.12 Blood sampling for ascorbate measurement and metabolomics*

Mice were fasted for approximately 15 hours before blood sampling from the lateral tail vein. Blood sampling was performed by trained staff at the MLC and the volume of extracted blood did not exceed 15% total blood volume. EMLA cream was applied to the tail 30 minutes prior to blood sampling, then thoroughly removed with an ethanol wipe before a small incision was made at the base of the tail. Between 100ul and 200ul of blood was collected into heparin-coated tubes. 50ul of whole blood was immediately mixed with 50ul of *meta*-phosphoric acid (MPA) buffer (5% w/v *meta*-phosphoric acid, 2mM diethylenetriaminepentaacetic acid) and deprotonated on ice for 10 minutes. This whole blood/MPA buffer mix was used for measurement of glutathione and ascorbic acid. The remaining whole blood was left to coagulate on ice for 40 minutes, then centrifuged at 2000 x g for 10 minutes at $4^{O}C$. The plasma and cell pellet were separated and both were flash frozen in $LN_2$ and stored at $-80^{O}C$ before analysis.

#### *2.3.2.13 Monthly weights*

Mice were weighed monthly at approximately the same time of day (between 8am and 12am) by animal technicians at the MRC Harwell Institute. Mice that reached 10% weight loss compared to their individual maximum body weight were weighed more frequently (weekly or daily).

#### *2.3.2.14 Survival*

The humane end points for this study are defined in section 2.2.3. Mice which were culled for "non-ALS" causes (e.g. fighting, tumours) were not included in the survival analysis. Mice that did not develop ALS symptoms were culled by 27 months of age. Survival data were analysed by Kaplan-Meier. Statistical significance was determined using a log-rank (Mantel-Cox) test where $P < 0.05$ was considered significant.

#### *2.3.2.15 Harvesting of tissues*

Animals were terminally anesthetised with 1000 mg/kg of pentobarbitone administered via i.p. injection. Once sufficiently anesthetised all, or a subset, of the following tissues were harvested from each animal: brain, spinal cord, quadriceps, tibialis anterior (TA), Extensor digitorum longus (EDL), soleus, kidney, liver, sciatic nerve, lumbrical muscles and tail.

Further processing of fresh tissue differed depending on the experiment. For protein and RNA work, tissues were flash frozen in liquid nitrogen and stored at -80$^{O}$C until use. For SOD1 and CTIP2 immunofluorescence, brains and spinal cords were embedded in OCT and frozen in isopentane over liquid nitrogen. Embedded tissues were stored at -80$^{O}$C until use.

For muscle fibre size measurement and ATPase staining, tibialis anterior (TA) muscles were attached to cork using trigacanth gum, covered in OCT and flash frozen in isopentane over liquid nitrogen. Frozen muscles were stored at -80$^{O}$C until use.

For motor neuron counting, lumbar spinal cords were embedded in OCT and frozen in isopentane over dry ice. Embedded spinal cords were stored at -80$^{O}$C until use.

For the analysis of neuromuscular junctions (NMJs), hindlimb lumbrical muscles were extracted, fixed for 10 minutes in 4% PFA and stored in PBS until staining. TA and soleus muscles for NMJ analysis were extracted, pinned out, and fixed in 4% PFA for between 20 minutes and 3 hours. Fixed TA and soleus muscles were stored in PBS for up to one month before further processing.

## 2.4 Histology, immunohistochemistry and immunofluorescence

### 2.4.1 MN counts

Our protocol for motor neuron counting has previously been published [413]. Briefly, fresh-frozen lumbar spinal cords were embedded in OCT and sectioned using a cryostat. 20 µm thick transverse sections were cut spanning the L2-L6 region of the lumbar spinal cord and mounted on glass slides.

The sections were divided into 3 series such that the distance between sections within each series was 60µm. Sectioning was performed by histology staff at the MRC Harwell institute.

To visualise motor neurons, sections from *hSOD1$^{WT}$* mice were stained with gallocyanin and sections from *hSOD1$^{A4V}$* mice were stained with cresyl violet. Stained sections were mounted and imaged on a Nanozoomer slide scanner (Hamamatsu). For each animal, spinal motor neurons in the sciatic pool of the lumbar spinal cord were counted in a minimum of 40 sections spanning the L2-L6 region of the spinal cord. Spinal cord regions L2-L6 were identified from the spinal cord morphology. Motor neurons were differentiated from other cell types by their large size, dense staining pattern and the presence of multiple dentrites. Motor neuron counting was performed in the NDP.veiw 2 software (Hamamatsu). Analyses are presented as the average number of sciatic motor neurons per section in the L3-L5 region of the spinal cord.

### 2.4.2 NMJ analysis on lumbrical muscles

The protocol for staining and analysis of NMJs in lumbrical muscles was adapted from Sleigh et al., 2014. Fixed hindlimb lumbrical muscles in 24 well plates were washed once in PBS for 30 minutes. Muscles were then permeabilised in 2% triton X-100 (v/v) for 30 minutes, followed by a 30 minute incubation in blocking solution (4% BSA (w/v), 1% triton X-100 (v/v), PBS), and incubation in primary antibodies against neurofilament (2H3; 1:50) and synaptophysin (SV2; 1:100) in blocking solution at 4$^{O}$C overnight. Muscles underwent 3 x 30 minute washes in PBS, followed by incubation with anti-mouse IgG (1:500, Alexa Fluor 488, ThermoFisher) and α-Bungarotoxin (α-BTX; 1.5 µg/ml; tetramethylrhodamine; Biotium) in PBS for 2 hours at room temperature. Muscles were washed 3 x 30 minutes in PBS and mounted on glass slides in ProLong Gold mounting media (ThermoFisher).

Samples were imaged at 488 nm and 555 nm wavelengths on an LSM 700 confocal microscope at 20x magnification. NMJ analysis on Z-stack maximum intensity projections was performed in ImageJ. NMJs were recorded as innervated or denervated; partial denervation was not assessed due to high

non-specific staining with antibodies against 2H3 and SV2 in older animals. At least 40, but usually 100-200 NMJs were counted from each animal.

### 2.4.3 NMJ analysis on TA and soleus muscles

The TA and soleus muscles from 6 and 18-month-old *hSOD1$^{A4V}$* animals were injected with tetanus toxin heavy chain ($H_CT$) as part of axonal transport experiments; $H_CT$ labels neuronal endosomes and hence stains the axon/presynapse without the need for antibodies. The TA and soleus muscles from 24-month-old mice were labelled using antibodies against SV2/2H3.

The protocol for processing of TA and soleus muscles for NMJ analysis was adapted from Bolatto, Olivera-Bravo and Cerri, 2021. TA and soleus muscles were fixed for 30-180 minutes in 4% PFA, teased apart into bundles of approximately 5-20 fibres and transferred to PBS in 24 well plates. Fibre bundles were washed once in PBS for 20 minutes. Fibre bundles were then stained with α-BTX (1:500, Alexa Fluor 488 or tetramethylrhodamine) for 1 hour at room temperature. Fibre bundles from 24-month-old animals were permeabilised in 2% Triton X-100 (v/v) for 90 minutes, blocked in 4% BSA (w/v) and 1% Triton X-100 (v/v) for 30 minutes and incubated in primary antibodies against 2H3 (1:50) and SV2 (1:100) in blocking solution at 4$^{O}$C overnight. Fibre bundles from 24-month-old animals were washed 4 x 20 minutes in PBS before incubation in anti-mouse IgG (1:250, Alexa Fluor 488, ThermoFisher) for 1 hour at room temperature. All fibre bundles were then washed 4 x 20 minutes in PBS and mounted on glass slides in ProLong Gold mounting media (ThermoFisher).

Samples were imaged at 488 nm and 555 nm wavelengths on an LSM 700 confocal microscope at 20x magnification. NMJ analysis on Z-stack maximum intensity projections was performed in ImageJ. NMJs were recorded as innervated or denervated; partial denervation was not assessed due to poor quality of staining of the postsynaptic NMJ with α-BTX. At least 40 NMJs were counted from each muscle.

### 2.4.4 SOD1 immunofluorescence

Brains and lumbar spinal cords were dissected and embedded in OCT over isopentane over liquid nitrogen. 12 µm thick serial transverse sections were cut on a cryostat. Sections were washed 3 x 10 minutes in PBS before incubation in blocking solution (2% BSA (w/v), 0.1% Triton X-100 (v/v), TBS) for 1 hour. Sections were then incubated with a pan-native-SOD1 antibody (1:1000) – which recognises an epitope of native SOD1 that is conserved between the mouse and human proteins [416] – in blocking solution overnight at 4$^{O}$C. Sections were washed 3 x 10 minutes in PBS then incubated with anti-goat IgG (1:1500, Alexa Fluor 488) in blocking solution for 1 hour at room temperature. Finally, sections were washed 3 x 10 minutes in PBS and mounted on glass slides using Prolong Glass Antifade Mountant with NucBlue (ThermoFisher). Samples were imaged at 405 nm and 488 nm wavelengths on an LSM 700 confocal microscope.

### 2.4.5 Misfolded/aggregated SOD1 immunohistochemistry

SOD1 aggregation was examined in *hSOD1$^{A4V}$* mice using an aggregate-specific antibody (mAb 85.11; Lehmann et al., 2020). Lumbar spinal cords were dissected and embedded in OCT over isopentane over liquid nitrogen. 12µm thick serial transverse sections were cut on a cryostat. Sections were fixed in 4% PFA for 15 minutes and washed 1 x 10 minutes in TBS. Sections were then incubated in hydrogen peroxide block (Abcam; ab64264) for 10 minutes followed by 2 x 10 minute washes in TBS. Sections were then incubated in Fab'anti-mouse IgG(H+L) blocking reagent (Abcam; ab269452) for 60 minutes and washed 1 x 10 minutes in TBS. Sections were incubated in primary antibody (mAb 85.11; 0.5 µg/mL) for 40 minutes, followed by 3 x 10 minute washes in TBS. HRP polymer detector reagent (Abcam; ab269452) was applied for 15 minutes followed by a further 3 x 10 minute washes in TBS. Finally, sections were incubated in DAB chromogen/substrate mix (Abcam; ab64264) for 10 minutes, rinsed in 8 changes of water, and mounted on glass slides in ProLong Gold mountant (Thermofisher). Stained sections were digitalised on a Nanozoomer slide scanner (Hamamatsu) and

assessed using a subjective rating scale (Figure 2-2). *hSOD1*$^{KO}$ mice and P100 transgenic *tgSOD1*$^{G93A}$ mice were included as staining controls.

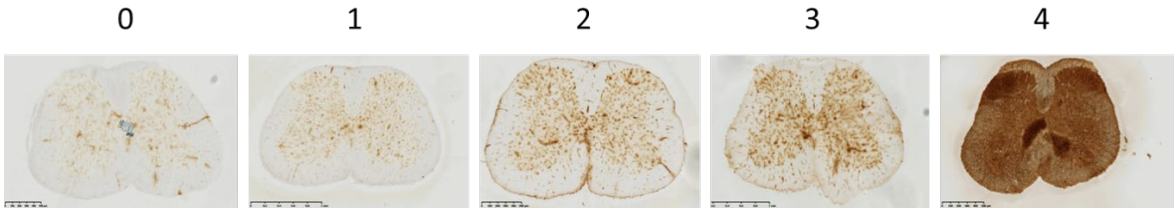



**Figure 2-2 | Subjective visual rating scale to assess misfolded/aggregated SOD1 immunohistochemistry**

### 2.4.6 Muscle fiber typing

Muscle ATPase staining was performed to differentiate type I, II and IIC muscle fibers. Flash frozen TA muscles were cut into 12 µm thick transverse sections on a cryostat. Sections were preincubated in sodium barbital solution at pH 4.5 (35 mM sodium barbital, 60 mM sodium acetate, HCl) or pH 10.8 (20 mM sodium barbital, 36mM $CaCl_2$, NaOH). Sections were rinsed once in distilled water followed by incubation in ATP solution (20 mM sodium barbital, 9 mM $CaCl_2$, 2.7 mM ATP; pH 9.5). Sections were then washed 3 x 3 minutes in 1% $CaCl_2$ (w/v) before incubation for 10 minutes in 2% $CoCl_2$ (w/v). Sections were washed in 5 changes of 5 mM sodium barbital. Sections were then incubated in 1% ammonium sulphide (v/v) for 30 seconds, washed in 6 changes of water and mounted using Aquamount (Life Technologies). Stained sections were digitalised on a Nanozoomer slide scanner (Hamamatsu). Muscle fiber type was assessed in a 750 µm x 750 µm area positioned at approximately the same location within each section (Figure 2-3). For each animal muscle fiber type was assessed at pH 4.5 and pH 10.8 on an average of 2 sections spaced 400 µm apart.

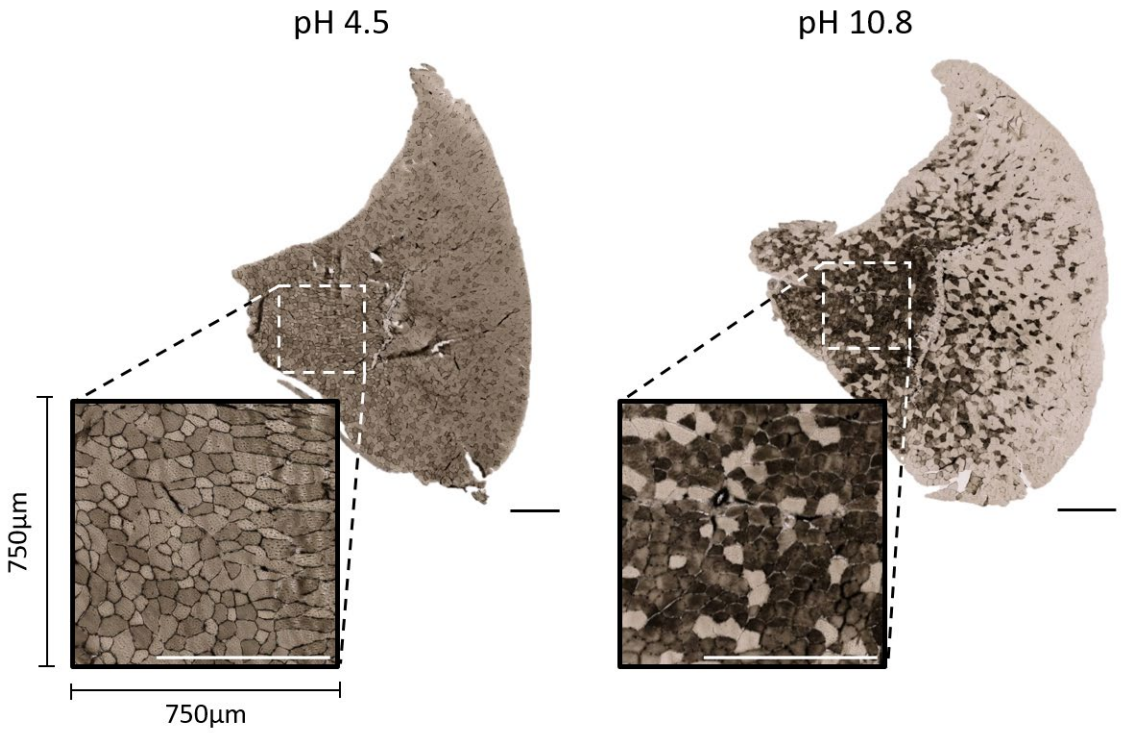



**Figure 2-3 | Representative images of TA muscle sections following ATPase staining.** Cutout shows the area of each section that was assessed for muscle fiber type. Scale bar = 500 µm

### 2.4.7 Determination of muscle fibre size and central nucleation

Fresh frozen TA muscles were sectioned on a cryostat at 8 µm thickness. Sections were transferred to glass slides, stained with Haemotoxylin and Eosin (H&E), and digitized using a slide scanner (Zeiss Axio Scan.Z1). For each muscle, three 500µm x 500µm regions were selected for determination of muscle fibre size. Segmentation of muscle fibres was performed in Ilastik [408] and ImageJ (National Institutes of Health), and checked manually. The minimum feret diameter for each fibre was calculated in ImageJ. Central nucleation was assessed manually in 500 fibres per muscle and expressed as a percentage.

## 2.5 Molecular biology

### 2.5.1 DNA extraction from mouse tissues

DNA was extracted from mouse tissues using a MyTaq extract kit following the manufacturers protocol.

### 2.5.2 Protein extraction from mouse tissues

Flash-frozen tissues, stored at -80°C, were thawed on ice and homogenised in 25 volumes of ice-cold lysis buffer (PBS, Complete Protease Inhibitor Cocktail (Roche), 0.5% NP-40). Lysates were centrifuged at 18,000 x g for 30 min at 4°C and the supernatants were collected and stored at −80°C until analysis. Protein concentration was measured using a DC protein assay kit (BioRad).

### 2.5.3 Western blotting

10-15ug of protein per sample was used for reducing SDS-PAGE on Novex 14% Tris-Glycine gels (ThermoFisher) followed by Western blotting. Blots were probed with a pan-SOD1 antibody raised against a peptide corresponding to amino acids 133-151 of the SOD1 protein, which are identical in human and murine SOD1, followed by IRDye 680RD conjugated secondary antibody (LI-COR). SOD1 immunofluorescence was normalised to the total protein content of each sample, measured using a Revert 700 Total Protein Stain (LI-COR). Images were taken using an Odyssey CLx imager (LI-COR).

### 2.5.4 In-gel dismutase activity assays

The protocol for in-gel dismutase activity assays was adapted from Weydert and Cullen, 2010. 15-30 ug of protein per sample was separated using clear native PAGE (CN-PAGE) on 3-12% NativePAGE Bis-Tris gels (ThermoFisher). Following CN-PAGE, gels were soaked in SOD1 stain solution (50 mM potassium phosphate, 336 µM nitro blue tetrazolium, 170 µM riboflavin, 21 mM TEMED) in the dark for 45 minutes. Gels were then rinsed with dd$H_2O$ twice and placed under a fluorescent light for approximately 1 hour, during which time clear bands of SOD1 activity began to appear. Gels were then washed 3 times with dd$H_2O$ and left under ambient light for 18-24 hours to allow for further band development. Gels were imaged using a ChemiDoc XRS+ imaging system (BioRad) with a white light conversion filter.

### 2.5.5 Spectrophotometric dismutase activity assays

The protocol for a high throughput spectrophotometric SOD dismutase activity assay was adapted from Peskin and Winterbourn, 2017. Spectrophotometric dismutase activity assays were optimised by myself and Chloe Williams, with the final assays performed by Chloe Williams. Optimal xanthine oxidase and WST-1 concentrations were selected in a pre-run to give an absorbance change of 0.04 AU per minute which remained linear over a 5 min experiment.

Protein samples were thawed on ice and dispensed into 96 well assay plates. Depending on the tissue, 1-5ug of protein per sample was used for dismutase activity assays. 200 µL of reagent mixture (100 mM pH 8.0 sodium phosphate, 100 µM DTPA, 10 mM hypoxanthine, 50 µM WST-1, 10 µg/mL catalase, 0.1 units/mL xanthine oxidase (Sigma)) was added to every well using a multichannel pipette. Plates were shaken for 30 seconds to mix then immediately absorbance was read at 438 nm wavelength every 40 seconds for 3 minutes at 25$^{O}$C. SOD1 activity was determined based on comparison to 12-point standard curve of bovine SOD1 ranging from 1-1000 ng/well. 1 unit of activity was defined as the amount of bovine SOD1 required for 50% inhibition of tetrazolium reduction (approximately 25 ng).

### 2.5.6 RNA extraction for quantitative PCR (qPCR) and RNA sequencing (RNA-seq)

Animals were euthanized by overdose of anaesthetic administered via i.p. injection. Tissues were collected and immediately immersed in ice-cold RNAlater solution (Sigma Aldrich). Tissues were stored at -20$^{O}$C in RNAlater until use.

RNA was extracted from CNS tissues using an RNeasy Lipid Tissue Mini Kit (Qiagen), from muscle using an RNeasy Fibrous Tissue Mini Kit (Qiagen), and from Kidney and liver using an RNeasy Mini Kit (Qiagen). RNA quantity was assessed using a NanoDrop Spectrophotometer (ThermoFisher). RNA quality was assessed using a 2100 Bioanalyzer (Agilent) where an RNA integrity number (RIN) of >7 was deemed acceptable for qPCR and RNA-seq. Extracted RNA was stored at -80$^{O}$C until use.

### 2.5.7 Quantitative PCR (qPCR)

cDNA was generated using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). 10 ng of cDNA was used to perform qPCR in triplicate, using 5 µl of cDNA, 10 µl of Fast SYBR Green Master Mix (Thermo Fisher) and 0.2 µM of each forward and reverse primer in a total reaction volume of 20 µl. qPCR was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems) using the following conditions: 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 1 min. The sequences of primers used in this study are listed in Table 2-12. Relative quantification of gene expression was performed using the $2^{-\Delta\Delta CT}$ method. *GAPDH* was used as a housekeeping gene.

| Primer Name | Primer Sequence (5'->3') |
|---|---|
| tSOD1 qPCR fwd 1 | TGAAGAGAGGCATGTTGGAGA |
| tSOD1 qPCR rev 1 | CCACCTTTGCCCAAGTCATC |
| tSOD1 qPCR fwd 2 | GGATGAAGAGAGGCATGTTGG |
| tSOD1 qPCR rev 2 | TTTCCACCTTTGCCCAAGTC |
| Gapdh-for | CGGCCGCATCTTCTTGTG |
| Gapdh-rev | CCGACCTTCACCATTTTGTCTAC |

**Table 2-12 | List of qPCR primers**

### 2.5.8 Quantseq 3' mRNA sequencing

Library preparation and Illumina sequencing were performed by Lexogen. cDNA libraries were prepared using QuantSeq 3' mRNA-Seq Library Prep FWD kits (Lexogen) and sequenced on an Illumina platform to generate single-end reads. Basic statistics are listed in Table 2-13.

| Sample group | Total read count (Median) | Uniquely mapped reads (Median) | Uniquely mapped reads % (Median) | Average input read length (Median) |
|---|---|---|---|---|
| 6 mo *hSOD1$^{A4V}$* spinal cord | 8200155 | 5369403 | 65.7 | 86 |
| 6 mo *hSOD1$^{A4V}$* TA muscle | 7379234 | 4984674 | 68.0 | 90 |
| 12 mo *hSOD1$^{A4V}$* spinal cord | 4860088 | 3198704 | 66.2 | 66 |
| 12 mo *hSOD1$^{A4V}$* TA muscle | 4767525 | 3146542 | 65.7 | 67 |
| 20 mo *hSOD1$^{A4V}$* spinal cord | 5150797 | 3571121 | 69.0 | 86 |
| 20 mo *hSOD1$^{A4V}$* TA muscle | 4609703 | 3419747 | 74.4 | 88 |
| 6 mo *hSOD1$^{KO}$* spinal cord | 4749682 | 3141176 | 65.5 | 66 |
| 6 mo *hSOD1$^{KO}$* TA muscle | 4867063 | 3202957 | 64.3 | 67 |

**Table 2-13 | Read counts from Quantseq 3' mRNA sequencing runs.** mo, month-old.

Differential gene expression (DE) analysis was performed by Chloe Williams. Briefly, quality control and read trimming was performed with Trimmomatic version 0.36. Trimmed reads were then aligned to the mm39 mouse reference genome with HISAT2 (Version 2.1.0). FeatureCounts was used to count exon spanning reads.

One *hSOD1$^{WT}$*/*hSOD1$^{WT}$* sample from the 6-month-old TA muscle dataset clearly deviated from all other samples in PCA and was excluded from downstream analyses.

DE analyses were conducted with DESeq2 (Version 1.40.1). Differentially expressed genes (DEGs) with an adjusted p value (padj) < 0.1 were considered significant. The Benjamini and Hochberg procedure implemented in DESeq2 was used for correction of p values for false discovery rate (FDR). DE figures were produced using RStudio (Version 2023.03.0), ggplot2 (Version 3.4.2) and EnhancedVolcano (Version 1.18.0). GO pathway analysis on DEGs was performed using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost).

EdgeR's TMM method for normalisation was applied to raw counts data prior to running gene set enrichment analysis (GSEA; Subramanian et al., 2005). GSEA against the mouse "hallmarks" gene sets (MSigDB) was performed using GSEA 4.3.2 (Broad institute).

### 2.5.9 Xdrop enrichment and Nanopore sequencing of the *hSOD1$^{WT}$* locus

High molecular weight (HMW) DNA was extracted from homozygous *hSOD1$^{WT}$*/ *hSOD1$^{WT}$* spleen by mechanical dissociation and incubation in lysis buffer (10 mM Tris-Cl, 100 mM NaCl, 25 mM ETDA, 0.5% SDS (w/v), 20 µg/ml RNAse A) at 37$^{O}$C for 1 hour. Followed by further incubation in lysis buffer plus proteinase k (100 µg/ml) at 50$^{O}$C for 3 hours. HMW DNA was purified by phenol/chloroform extraction and ethanol precipitation, and eluted in 10 mM Tris-Cl. HMW DNA was handled with care to avoid shearing of DNA fragments. Size range of the extracted fragments was estimated by PFGE.

Xdrop enrichment (Samplix; Madsen et al., 2020) and Nanopore sequencing (Oxford Nanopore Technologies) were performed by Samplix Services. We have published a detailed description of the enrichment and sequencing of the *hSOD1$^{WT}$* locus [362]. FASTQ sequencing data were aligned to the C57BL/6J (GRCm38.p6), 129S1 SvImJ (GCA_001624185.1) and human (hg38) reference genomes, and a reference file containing the expected engineered allele, using Minimap2 (Li, 2018). Samtools was used to sort and index alignment data for visualization in Integrative Genomics Viewer (IGV; Robinson et al., 2011).

### 2.5.10 Targeted locus amplification and sequencing of the *hSOD1$^{G93A}$* and *hC9orf72* loci

Splenocytes were isolated from male heterozygous *hSOD1$^{G93A}$* and *hC9orf72* animals. First, spleens were mechanically dissociated in 10% foetal calf serum (FCS) in PBS. Red blood cells (RBCs) were lysed by incubation in RBC lysis buffer (154 mM $NH_4Cl$, 10 mM $KHCO_3$, 193 µM EDTA, $H_2O$) for 5 minutes at room temperature. Splenocytes were pelleted, counted, and frozen in cryovials with $10^8$ viable cells/vial in 10% FCS, 10% DMSO in PBS.

Targeted locus amplification (TLA) and short-read Illumina sequencing were performed by Cergentis as previously described [422]. To provide coverage across the entire loci, two TLA PCR primer pairs were designed against each humanised allele. The resulting PCR products were purified, and DNA libraries prepped using the Nextera DNA Flex Library Prep protocol (Illumina). Libraries were pooled and sequenced on an Illumina NGS system. Base calling and demultiplexing was performed using the bcl2fastq Conversion Software (Illumina). FASTQ sequencing data were aligned to the C57BL/6J (GRCm38.p6) and human (hg38) reference genomes, and a reference file containing the expected engineered allele, using Minimap2 (Li, 2018). Samtools was used to sort and index alignment data for visualization in IGV (Robinson et al., 2011).

### 2.5.11 Metabolomics

#### *2.5.11.1 Preparation of samples*

Frozen plasma samples from hSOD1$^{A4V}$ animals were thawed at room temperature and diluted 1/10 with 90% MeOH. The samples were then precipitated at -20°C for 2 hours and centrifuged for 10 min at 17,000 x g at 4°C. The supernatants were transferred to glass vials and dried in a SpeedVac concentrator. Quality control samples were generated by pooling a small amount of all experimental samples. Internal standards of known concentration were added to each sample.

#### *2.5.11.2 LC-MS*

Liquid chromatography–mass spectrometry (LC-MS) experiments were performed by the Swedish Metabolomics Centre. Samples were run on an ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UHPLC/Q-TOF-MS) instrument in both positive and negative mode. To minimise the effect of instrumental drift, samples from each time point were run in a random order. No run-order effect was observed in negative mode (Figure 2-4 A, B). A slight run-order effect was observed in positive mode (Figure 2-4 C, D); however, as the effect was small and the samples from each time point were run together, it was not necessary to normalise the positive mode data. Using a targeted data processing approach, 212 metabolites were annotated in samples

from 4-13 month-old animals, and 195 metabolites were annotated in samples from 20-month-old animals.

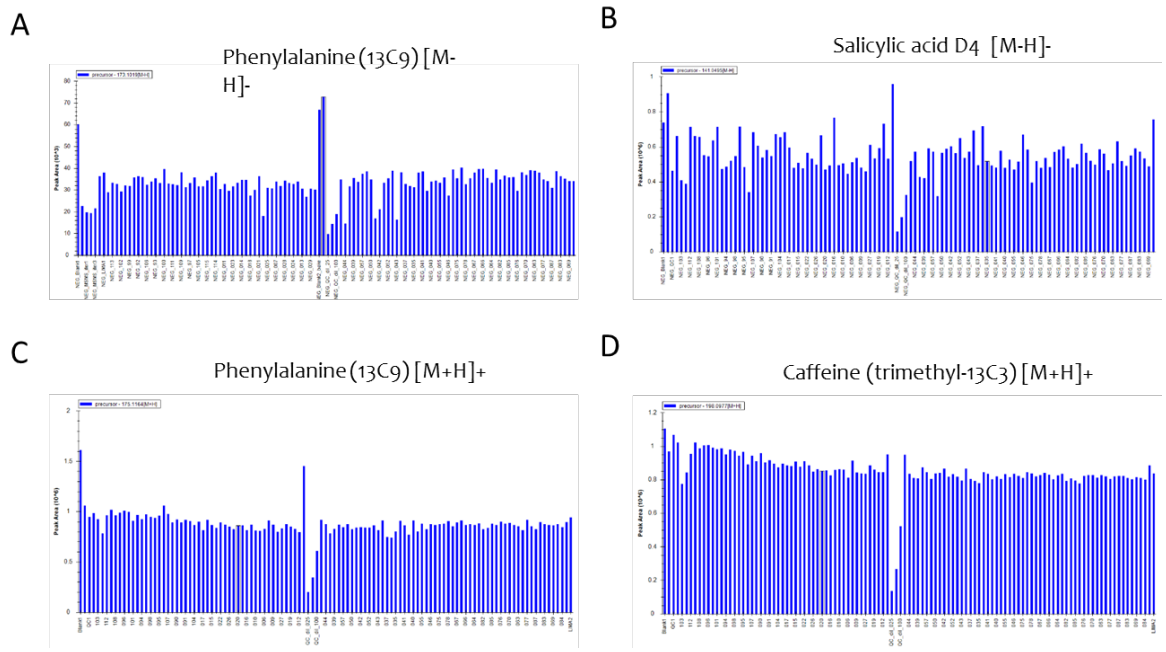



**Figure 2-4 | LC-MS internal standards. (A,B)** Internal standards from LC-MS in negative mode. **(C,D)** Internal standards from LC-MS in positive mode. X axis shows experimental samples and standards in run order. Data shown is from samples from 4-13 month-old animals. A similar pattern was observed in samples from 20-month-old animals.

### *2.5.11.3 GC-MS*

Gas chromatography–mass spectrometry (GC-MS) experiments were performed by the Swedish Metabolomics Centre. Samples were run on a gas chromatography time-of-flight mass spectrometry (GC-TOF-MS) instrument. Samples from each time point were run in a random order. Using a targeted data processing approach, 94 metabolites were annotated in samples from 4-13 month-old animals, and 82 metabolites were annotated in samples from 20-month-old animals.

### *2.5.11.4 Metabolomics data analysis*

Data pre-processing and annotation of metabolites was performed by Elin Näsström at the Swedish Metabolomics Centre. Further analysis was performed using MetaboAnalyst 5.0 [423]. The input for statistical analyses was a data matrix of annotated compound names with associated peak (mz/rt) intensities. MS data were normalised using the 'auto scaling' method implemented in

MetaboAnalyst (Figure 2-5). Data from each time point were analysed separately using Linear Models for Microarray Data (limma) with sex as a covariate, padj < 0.05 was considered significant. PCA was performed in MetaboAnalyst using the 'prcomp' R package. PCA results are presented as pairwise score plots - providing an overview of the various separation patterns among the most significant PCs, scree plots, and loadings plots.

Partial least squares - discriminant analysis (PLS-DA) was performed in MetaboAnalyst using the 'plsr' function from the R package 'R pls'. Classification and cross-validation were performed using the corresponding wrapper function from the R package 'caret'. The significance of class discrimination in the PLS-DA models was assessed using a permutation test which provides test statistics based on the prediction accuracy during training. These test statistics, $R^2$ and $Q^2$, can be interpreted as the quality of the prediction model when applied to the calibration data and novel test data, respectively [424]. PLS-DA analysis was considered significant when $R^2$ and $Q^2$ were similar in size and both greater than 0.

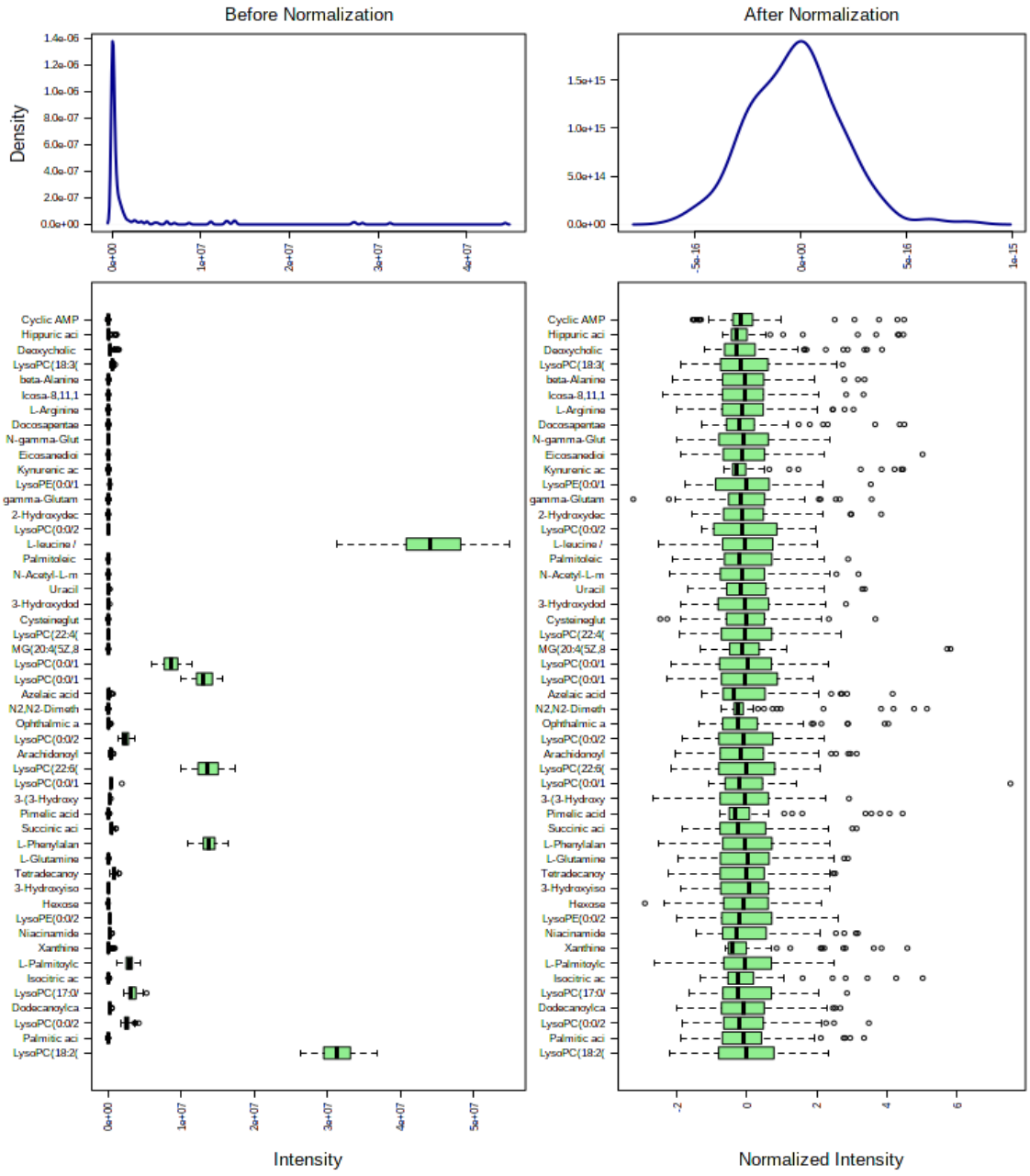



**Figure 2-5 | Example of normalisation of MS data.** Box plots and kernel density plots show LC-MS data before and after normalization. The boxplots show the first 50 metabolites due to space limit. The density plots are based on all samples.

### 2.5.12 Measurement of GSSG/GSH/ASA

Reduced glutathione (GSH), oxidised glutathione (GSSG) and ascorbic acid (ASA) levels were measured in whole blood from hSOD1$^{A4V}$ animals using LC-MS, performed by the Swedish Metabolomics Centre. Before GSH/GSSG analysis, 25 µL of each sample was mixed with 75 µL of extraction solution (1.67% MPA, 3.33 µM GSH as internal standard). Before ASA analysis, 25 µL of

each sample was mixed with 75 µL of extraction solution (3.1% MPA, 3.33 µM ASA as internal standard). GSH, GSSG and ASA were annotated in the MS data using a targeted processing approach.

### 2.5.13 Measurement of amino acids

Quantification of amino acids in whole blood samples was performed by the Swedish Metabolomics Centre using liquid chromatographic-triple quadrupole tandem mass spectrometry (LC-QqQ-MS/MS).

### 2.5.14 Axonal transport experiments

Axonal transport experiments were performed by Dr Andrew Tosolini at UCL as previously described [176]. Briefly, mice were anesthetised and the TA and soleus muscles were exposed following small incisions in the overlying tissue. The TA and soleus muscles on opposite hindlimbs were then injected with a fluorescently labelled (AlexaFluor555) fragment of the tetanus neurotoxin ($H_CT$). The skin was sutured and the mice were allowed to recover for 4-8 hours. Mice were then re-anesthetised and the sciatic nerve of each hindlimb were exposed. Parafilm was placed under the sciatic nerve to aid imaging. The mouse was placed in an environmental chamber – maintained at 37$^{O}$C – attached to an LSM780 confocal microscope (Zeiss). Imaging was performed using a Plan-Apochromat 40x/1.3 Oil DIC objective (Zeiss) at 2.5-3.5 frames per second.

Endosome tracking was performed by Dr Andrew Tosolini using ImageJ (NIH) and the TrackMate plugin [425]. The speed of $H_CT$-555-positive retrograde signalling endosomes was measured in 3-5 axons per muscle per animal. Data were analysed and compiled into figures by Dr Andrew Tosolini.

### 2.5.15 *In vivo* muscle isometric tension measurements

Muscle tension measurements were performed by Dr Mhoriam Ahmed at UCL as previously described [426]. Briefly, mice were deeply anesthetised and the distal tendons of the TA and EDL muscles were dissected free and attached to isometric force transducers (Dynamometer UFI Devices). The sciatic nerves were exposed, sectioned and stimulated via wire electrodes using 0.02

ms pulses at frequencies from 40-100 Hz to elicit isometric contractions. The number of functional motor units innervating the EDL muscles were determined by stimulation in stepwise increments of increasing intensity to recruit progressively more motor axons. Stepwise increases in force generation were counted to determine the number of functional motor units. Time to peak was determined from the time delay between the onset of stimulation and the recording of peak muscle twitch tension. Parameters were measured using Picoscope v6 software (Pico Technology).

## 2.6 *In-vitro* molecular cloning

Procedures used in the cloning of targeting vectors for *hSOD1*$^{G93A}$ and *hC9orf72* are described in detail in the relevant chapters. General methods used are described below.

### 2.6.1 BAC and plasmid vectors

pSMART BAC vector was purchased from Lucigen as part of a CopyRight v2.0 BAC Cloning Kit. pJAZZ-OC and pJAZZ-OK vectors were purchased from Lucigen as part of BigEasy v2.0 Linear Cloning System kits. The genomically-derived BAC constructs used in the generation of the *hSOD1*$^{G93A}$ and *hC9orf72* mouse strains are described in the relevant chapters.

### 2.6.2 Bacterial strains and culture conditions

The *E. coli* strains used in this project were NEB 10-beta Competent *E. coli* (New England Biolabs) ElectroMAX Stbl4 Competent Cells (ThermoFisher) and BigEasy-TSA Electrocompetent *E. coli* (Lucigen; used for cloning with pJazz). Bacteria containing large or unstable constructs were grown at room temperature or 30$^{O}$C, whereas bacteria containing small and/or stable constructs were grown at 37$^{O}$C. All bacteria were grown in LB Lennox media/plates with antibiotic selection. Liquid cultures were continuously shaken at 200-225 rpm in aerated containers. Unless otherwise stated, antibiotics were used for selection on plates and in liquid media at the following concentrations: chloramphenicol, 12.5 µg/mL; kanamycin, 15 ug/mL or 30 µg/mL; ampicillin, 100 µg/mL;

tetracycline, 5 µg/mL; streptomycin, 50 or 15 µg/mL. Blue-white screening was performed on LB Lennox plates containing 1 mM IPGT and 20 µg/mL x-gal.

### 2.6.3 Transformation

Bacteria were transformed solely by electroporation using an E. coli Pulser (Bio-Rad) and 0.1 cm gap cuvettes (Bio-Rad) with electroporation conditions: voltage 1-2 kV, resistance 20 Ω, time constant >4.5 ms. Ligation reaction mixes were cleaned by ethanol precipitation prior to transformation, or transformed directly when ElectroLigase (NEB) was used.

### 2.6.4 DNA isolation from bacterial cultures

QIAprep Spin Miniprep (Qiagen), HiSpeed Plasmid Maxi (Qiagen) and Nucleobond Large Construct kits (Macherey Nagel) were used for the isolation of plasmid/BAC constructs from bacterial cultures.

### 2.6.5 Agarose gel electrophoresis

Standard agarose gel electrophoresis used 1-2% agarose gels with 1X GelRed or GelGreen (Biotium) and 1X TBE running buffer at 80-150V running voltage. Pulsed-field gel electrophoresis (PFGE) used 1% PFGE agarose (Biorad) gels and 0.5X TBE running buffer run on a Pippin Pulse electrophoresis power supply (Sage Science) using the pre-set "5-80 kb" protocol. PFGE gels were post-stained in 3X GelRed or GelGreen for 1 hour. Bands were visualised on ChemiDoc XRS+ imaging system (BioRad), or a blue light transilluminator (FastGene) if DNA fragment(s) were to be isolated.

Where fragment(s) of specific sizes were required following a restriction or Cas9 digest, these were separated using standard or pulsed-field agarose gel electrophoresis, cut out of the gel, and isolated using a Nucleospin Gel and PCR Clean-up Kit (Macherey-Nagel) or Zymoclean Large Fragment DNA Recovery Kit (Zymo Research).

### 2.6.6 Primers

Primers for PCR, Gibson assembly and Sanger sequencing were designed using Primer3 (https://primer3.ut.ee/) [427] and SnapGene (Dotmatics). All primers were purchased from IDT.

### 2.6.7 PCR

Short-range screening/genotyping PCRs where 100% faithful amplification of the template was not required were performed using MyTaq Red 2x master mix (Meridian Bioscience).

| **Reagent** | **Volume (µL)** |
|---|---|
| MyTaq Red 2x master mix | 25 |
| Forward primer (20 µM) | 1 |
| Reverse primer (20 µM) | 1 |
| $ddH_2O$ | To 50 |
| Template | PCR specific |
| | **Total volume 50 µL** |

**Table 2-14 | Reaction mix for MyTaq Red PCR**

| **Step** | **Temperature ($^{O}$C)** | **Time (s)** | **Cycles** |
|---|---|---|---|
| Initial denaturation | 95 | 60 | 1 |
| Denaturation | 95 | 15 | 25-35 |
| Annealing | Primer specific | 15 | |
| Extension | 72 | 10 | |
| Final extension | 72 | 120 | 1 |

**Table 2-15 | Cycling conditions for MyTaq Red PCR**

All other PCRs – including long-range and nested PCRs – were performed using Q5 2x master mix (NEB).

| **Reagent** | **Volume (µL)** |
|---|---|
| Q5 2x master mix | 10 |
| Forward primer (10 µM) | 1 |
| Reverse primer (10 µM) | 1 |
| $ddH_2O$ | To 20 |
| Template | PCR specific |
| | **Total volume 20 µL** |

**Table 2-16 | Reaction mix for Q5 PCR**

| **Step** | **Temperature ($^{O}$C)** | **Time (s)** | **Cycles** |
|---|---|---|---|
| Initial denaturation | 98 | 30 | 1 |
| Denaturation | 98 | 7 | 25-35 |
| Annealing | Primer specific | 15 | |
| Extension | 72 | 30 seconds/kb | |
| Final extension | 72 | 120 | 1 |

**Table 2-17 | Cycling conditions for Q5 PCR**

'No DNA template' controls were run for all PCR reactions. Positive controls for amplification were run where possible.

### 2.6.8 DNA Sequencing

#### *2.6.8.1 Sanger sequencing*

Sanger sequencing was performed by Source Bioscience. dGTP chemistry provided by Source Bioscience was used for sequencing templates with a high degree of secondary structure (e.g. G-C rich templates, tandem repeats etc.).

#### *2.6.8.2 PacBio long-read sequencing*

PacBio SMRT long-read sequencing of the hSOD1-G93A-pJAZZ targeting construct was performed by CD Genomics. For library preparation, 5 µg of input plasmid DNA was fragmented using g-TUBEs (Covaris). DNA fragments were treated with a DNA-repair mix to generate blunt ends, followed by hairpin adapter ligation. A SMRTbell DNA template library containing fragments >10 kb was isolated using a BluePippin system (Sage Science). Library quality was assessed using a Qubit Fluorometer (ThermoFisher) and average fragment size was estimated using agarose gel electrophoresis. The library was then sequenced using a PacBio Sequel II platform. Approximately 150,000 subreads were obtained with an average length of 9061 bp. The SMRT LINK 8.0 software was used for raw data processing. *De novo* construct assembly was performed using Flye [428] and Mecat2.

PacBio long-read sequencing of the C9orf72 repeat-carrying BAC was performed by the Centre for Genomic Research at the University of Liverpool. We have previously published a detailed description of the sequencing of this construct [429].

#### *2.6.8.3 Illumina short-read sequencing*

Illumina short-read sequencing of the hC9orf72-rep-mHA-pJAZZ mouse targeting construct was performed by Source Bioscience. Input DNA sample quality was assessed using a Qubit Fluorometer (ThermoFisher) and the DNA library was prepared using a gDNA library preparation kit. The pooled library was sequenced on a MiSeq sequencing platform (Illumina) generating 150 bp paired-end reads. Base calling and demultiplexing was performed using the bcl2fastq Conversion Software (Illumina). Sequencing generated a total of 1.38 million reads with an average trimmed length of 120

bp. Raw FASTQ files were adapter- and quality- trimmed using bbduk. Further quality control of FASTQ reads was performed using FASTQC. Reads were aligned to a reference file containing the expected construct sequence using BWA-mem and duplicate reads were removed.

### 2.6.9 CRISPR/Cas9 and CRISPR/Cpf1 molecular cloning

#### *2.6.9.1 sgRNA/crRNA synthesis*

Cas9 single guide RNAs (sgRNAs) and Cpf1 crispr RNAs (crRNAs) were designed using CRISPOR [430] and Guidescan [431], and were either purchased from Sigma or generated in-house by reverse transcription from a DNA template. For sgRNAs and crRNAs generated in-house, a HiScribe T7 High Yield RNA Synthesis Kit (NEB) was used in the following reaction: 1.5 µl 10X reaction buffer, 1.5 µl ATP, 1.5 µl, CTP 1.5 µl, GTP 1.5 µl UTP, 1.5 µl T7 RNA polymerase mix, 5 µl DNA template, and 6 µl $H_2O$, incubated at 37 °C for 16 h. The DNA templates used for *in vitro* reverse transcription were gBlock Gene Fragments purchased from IDT. sgRNAs/crRNAs from the above reaction were then treated with Turbo DNAse I (ThermoFisher) at 37 °C for 15 minutes. sgRNAs/crRNAs were purified using a Megaclear Kit (Ambion), quantified and stored at −80 °C until use. To assess RNA size and integrity, 500 ng of RNA was heated to 70 °C for 5 minutes, snap cooled on ice, and run on a 1% agarose gel.

sgRNAs/crRNAs were tested using *in vitro* CRISPR/Cas9 or CRISPR/Cpf1 reactions on synthetic gBlock DNA (IDT). Cutting efficiency was estimated from the ratio of cut/uncut DNA.

#### *2.6.9.2 In vitro CRISPR/Cas9 reactions*

Purified DNA was digested with CRISPR/Cas9 in the following reaction: For ribonucleoprotein assembly, a mix of 0.5 µl Cas9 (3 ug/µl) (Takara Bio), 2 µl sgRNA guide (150 ng/µl), 2 µl Cas9 buffer (NEB), 5.5 µl MilliQ $H_2O$ was preincubated at 37 °C for 5 min. Next, target DNA (typically 1-2 µg) and MilliQ $H_2O$ were added to a total reaction volume of 20 µl and incubated at 37 °C for 2 hours. Cas9 was inactivated by digestion with 1 µl proteinase k (20 mg/mL) at 65°C for 15 minutes.

#### *2.6.9.3 In vitro CRISPR/Cpf1 reactions*

Purified DNA was digested with CRISPR/Cpf1 in the following reaction: For ribonucleoprotein assembly, a mix of 3 µl EnGen Lba Cpf1 (3 ug/µl) (NEB), 9 µl crRNA guide (300nM) and 4 µl NEBuffer 2.1 (NEB) was preincubated at 37 °C for 10 min. Next, target DNA (typically 1-4 µg) and MilliQ $H_2O$ were added to a total reaction volume of 40-50 µl and incubated at 37 °C for 30 minutes. Cpf1 was inactivated by digestion with 1 µl proteinase k (20 mg/mL) at 65°C for 15 minutes.

### 2.6.10 Red/ET recombineering

Our recombineering protocol was adapted from Hofemeister et al., 2011 [432]. First, bacteria containing the BAC construct of interest were transformed with the inducible Red/ET expression plasmid (pSC101-BAD-Redαβγ): Cells were streaked out from a glycerol stock onto LB plates with antibiotics and grown overnight. Individual colonies were picked and grown overnight in 1 ml cultures of LB media with antibiotics. 30 µl of the overnight cultures were transferred to new microfuge tubes containing 1.4 ml of LB medium with antibiotics and grown for 2 hours at 30°C, followed by 40 minutes at 37°C. Cells were then pelleted and washed twice with ice cold 10% glycerol. Cells underwent electroporation with 50 ng of the Red/ET expression plasmid, followed by recovery in LB media without antibiotics for 70 minutes at 30 °C. Transformed cells were plated onto LB plates with antibiotics – including 5 µg/ml tetracycline to select for the Red/ET plasmid – and grown at 30°C overnight.

Next, cells were transformed with linear donor DNA for recombination: Individual colonies were picked and grown overnight in 1 ml cultures of LB media with antibiotics at 30°C. 50 µl of the overnight cultures were transferred to new microfuge tubes containing 1.4 ml of LB medium with antibiotics and grown for 2 hours at 30°C. Red/ET expression was induced by addition of 20 µl of 10% L-arabinose (w/v) and incubation for 40 minutes at 37°C. A negative control was also performed without the addition of L-arabinose. Cells were then pelleted and washed twice with ice cold 10% glycerol. Cells underwent electroporation with ~300 ng of the linear donor DNA, followed by

recovery in LB media without antibiotics for 3 hours at 37 °C. Transformed cells were plated onto LB plates with antibiotics and grown at 30$^{O}$C overnight. Colonies were picked and screened using colony PCR against the recombined BAC.

### 2.6.11 Gibson assembly

Gibson assembly was performed using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). Reactions were performed as follows: 10µl HiFi DNA Assembly Master Mix, 20-75 fmol of each DNA fragment, MilliQ H2O to a total reaction volume of 20 µl, incubated at 50$^{O}$C for 25-60 minutes. The ratio of fragments (pJazz long arm : pJazz short arm : insert) was 1:1:3 for assembly of pJAZZ-$\delta$AarI and 1:1:1 for assembly of hSOD1-G93A-pJAZZ. 'No insert' negative controls were also performed.

### 2.6.12 Restriction enzyme digests

All restriction enzymes were purchased from New England Biolabs with the exception of AarI which was purchased from ThermoFisher. Restriction enzyme digests were carried out in the appropriate reaction buffer for between 15 minutes and 16 hours depending on the efficiency of the digest.

### 2.6.13 Ligation reactions

Ligation reactions were performed using T4 DNA Ligase (NEB), CloneSmart DNA Ligase (Lucigen) or ElectroLigase (NEB) following the manufacturer's protocols.

### 2.6.14 Storage of DNA constructs

All DNA constructs were stored at 4$^{O}$C in TE buffer for short periods and in glycerol stocks of transformed bacteria at -80$^{O}$C for long periods.

## 2.7 Mammalian cell culture and cell experimental procedures

### 2.7.1 Primary spinal cord cultures

Embryos for primary spinal cord cultures were harvested at day E13.5. A Pregnant female was culled by cervical dislocation and the embryos transferred into Hanks' Balanced Salt Solution (HBSS)(Sigma) on ice. For each embryo, the head and tail were removed and the spinal cord was extracted using curved forceps. The meninges was then separated from the spinal cord using two sets of straight forceps. The lateral section of the neural tube, which gives rise to the dorsal horn, was removed using a scalpel. The remaining ventral spinal cord was cut into 1 mm segments and transferred to a tube containing 1ml of HBSS.

Under sterile conditions, the HBSS was replaced with fresh sterile HBSS containing 0.025% trypsin. The spinal cord segments were dissociated at 37$^{O}$C for 10 minutes, triturated, and the cell suspension was transferred to a tube containing Leibovitz's L-15 Medium (L-15; ThermoFisher), 4% bovine serum albumin (BSA; Merck) and 100ug/ml DNase I (Merck). The cell suspension was incubated for 1 minute at room temperature, then triturated and transferred to a tube containing L-15, 4% BSA and 20ug/ml DNase I. 1 ml of 4% BSA was added directly to the bottom of the tube and the cells were centrifuged at 400 x g for 5 minutes. The supernatant was removed and the cells were resuspended in motor neuron media (Neurobasal media, Glutamax, 25 µM 2-Mercaptoethanol, Penicillin-Streptomycin, B-27 supplement (all ThermoFisher), 2% v/v horse serum (Merck), 1 ng/ml BDNF, 0.1 ng/ml GDNF and 10 ng/ml CNTF (all PeproTech)). Cells were counted using a Countess II automated cell counter and plated in tissue culture plates pre-coated with 15 µg/ml poly-L-ornithine (Merck) and 3 µg/ml Laminin (ThermoFisher). Cells were cultured at 37$^{o}$C, 5% $CO_2$ for between 2-7 days, depending on the experiment. The cell culture media was not changed during this period.

### 2.7.2 Assessment of mitochondrial respiration in primary spinal cord cultures

Mitochondrial respiration in primary spinal cord cultures was assessed using the Seahorse XF Cell Mito Stress Test on a Seahorse XFe96 Analyzer (Agilent). Primary spinal cord cells were plated at a

density of $1.5x10^5$ cells/well in Seahorse XF96 cell culture microplates and assayed after 2 days *in vitro* (DIV). The Mito Stress Test was performed using the standard protocol according to the manufacturer's instructions. The optimal concentrations of Seahorse compounds were first determined in a titration experiment; the final compound concentrations chosen for subsequent experiments were 1.5 µM oligomycin, 1 µM FCCP, and 0.5 µM rotenone/antimycin A, which were found to elicit the maximal effect on respiration in primary spinal cord cultures. A minimum of 4 technical replicates were performed for each embryo. Data from the Mito Stress Test was Analysed in Wave (Agilent). Data were normalised to the total number of cells in each well, which was determined using a CyQUANT Cell Proliferation Assay (ThermoFisher) performed following the manufacturer's instructions.

### 2.7.3 Assessment of mitochondrial membrane potential in primary spinal cord cultures

Mitochondrial membrane potential in primary spinal cord cultures was measured using a TMRE-Mitochondrial Membrane Potential Assay Kit (Abcam). Primary spinal cord cells were plated at a density of $1x10^5$ cells/well in white-walled 96-well cell culture plates and assayed after 5 DIV. FCCP, which uncouples of oxidative phosphorylation in mitochondria, was used as a negative control.

The assay was conducted as follows. FCCP was added to the negative control wells at a final concentration of 20 µM and incubated for 10 minutes at 37°C. TMRE was then added to all experimental wells at a final concentration of 500 nM and the cells were incubated for 20 minutes at 37°C. Cells were then washed twice with PBS and fluorescence was measured at Ex/Em = 549/575 nm on a FLUOstar Omega Microplate Reader (BMG Labtech). To account for differences in cell number, TMRE fluorescence was normalised to DAPI fluorescence. For normalisation, the cells were fixed in 4% PFA for 10 minutes, then washed once with PBS and incubated in 1 µg/ml DAPI in PBS for 10 minutes. DAPI fluorescence was measured at Ex/Em = 405/450 nm. No-TMRE and no-cell controls were also conducted on each plate.

### 2.7.4 Statistics

Statistical analyses were performed in GraphPad Prism (Dotmatics). Table 2-18 outlines the statistical tests used in the analysis of all phenotyping and molecular experiments, with the exception of RNA-seq and metabolomics experiments. Statistical significance was defined as $P \leq 0.05$ for analysis of phenotyping and molecular experiments. Statistical analysis of RNA-seq and metabolomics experiments are described in detail in the relevant methods sections. Kaplan-Meier survival data were analysed using the Log-rank (Mantel-Cox) test. Observed Mendelian ratios were compared to the expected ratios using the Chi-square test. Principal component analysis (PCA) was performed on the correlation matrix in GraphPad Prism 10.

Graphs were generated using GraphPad Prism 10, GSEA and R.

| **Number of groups** | **Number of time points** | **Multiple comparisons between groups** | **Repeated measures?** | **Statistical test used** |
|---|---|---|---|---|
| Two | One | - | - | Two-tailed unpaired t-test |
| Two | Two or more | - | Yes | Mixed effects analysis with Šídák's multiple comparisons test |
| Three or more | One | Experimental groups compared to control | - | One-way ANOVA with Dunnett's post hoc test |
| Three or more | One | All groups compared | - | One-way ANOVA with Tukey's post hoc test |
| Three or more | Two or more | All groups compared | No | Two-way ANOVA with Tukey's post hoc test |
| Three or more | Two or more | All groups compared | Yes | Two-way repeated measures ANOVA or mixed effects analysis with Tukey's post hoc test |

**Table 2-18 | Statistical tests for analysis of data from phenotyping and molecular biology experiments**

# Chapter 3: Characterisation of a genetically humanised mouse carrying wild-type human SOD1

## 3.1 Introduction

The structure and enzymatic function of SOD1 is highly conserved throughout mammalian evolution, with 112 of 153 residues being fully conserved across all mammalian species [42]. Human and mouse SOD1 display high amino acid sequence conservation, with 128 of 153 (84%) residues conserved. Accordingly, the fold and overall structure of human and mouse SOD1 are highly similar [43], and enzymatically active mouse-human SOD1 heterodimers and human SOD1 homodimers have been observed in SOD1-ALS mouse models [74,363,433].

Despite the structural similarity between the human and mouse SOD1 orthologues, human SOD1 mutants typically have a higher propensity for aggregation than the equivalent mouse SOD1 mutants [434]. Additionally, mouse SOD1 does not readily co-aggregate with human SOD1 in systems where both orthologues are expressed [43,435]. The difference in aggregation propensity is thought be driven, at least in part, by exposed tryptophan (W32) [104,330] and cysteine (C111) [434,436,437] residues which appear in human SOD1, but not mouse Sod1 (Ensembl).

Species-specific SOD1 aggregation properties highlight the need for mouse models of SOD1-ALS that express human SOD1, an approach that has been widely employed for generating existing mouse models. However, most existing mouse models of SOD1-ALS are high-copy models created by classical transgenesis, and have several drawbacks as discussed in section 1.5.2.1. Hence, there is a need for mouse models of SOD1-ALS that express human SOD1 at physiologically relevant levels to uncover ALS pathomechanisms.

Because of the high conservation of amino acid residues between human and mouse SOD1, and indeed the high conservation of SOD1 throughout evolution, we hypothesised that human SOD1 should be able to functionally replace mouse SOD1. To test this hypothesis, and to provide a

platform for the generation of further humanised models of SOD1-ALS, we developed the *hSOD1*$^{WT}$ mouse line, which is a genetically humanised mouse that expresses wild-type human *SOD1* from the endogenous mouse *Sod1* locus. The development of the gene targeting construct used for *SOD1* humanisation was undertaken by Dr Rosie Bunton-Stasyshyn and Dr Anny Devoy at UCL, and forms part of Dr Bunton-Stasyshyn's PhD thesis. The gene targeting was performed by Dr Lydia Teboul and colleagues at the MRC Harwell Institute, and is described in section 2.2.7.1**.**

A schematic of the humanised *SOD1* allele is presented in Figure 3-1. All mouse promotor sequences and the mouse *Sod1* 5' UTR were retained in an effort to maintain mouse-specific regulation of *Sod1* expression. The *SOD1* gene is humanised from the ATG start codon in exon 1 to the end of the 3' UTR, such that all 5 *SOD1* exons, all 4 *SOD1* introns and the *SOD1* 3' UTR are human sequence. The humanised *SOD1* allele contains a further conditional copy of *SOD1* exons 4 and 5 (hereafter designated exon 4' and 5'), which are expressed following excision of exons 4 and 5 by Cre recombinase. This conditional sequence enables Cre-dependant conditional expression or silencing of mutations introduced into exons 4' and 5', or exons 4 and 5, respectively.

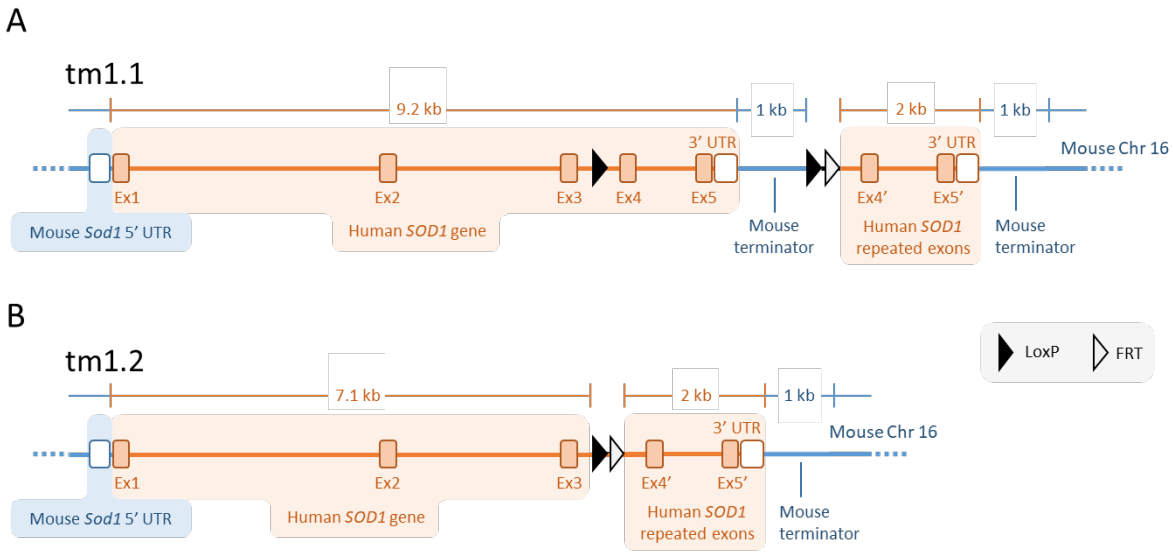



**Figure 3-1 | Schematic of the *hSOD1$^{WT}$* allele.** Humanised sequence is shown in orange, mouse sequence is shown in blue. **(A)** *hSOD1$^{WT}$* allele as present in *hSOD1$^{WT}$* mice. **(B)** *hSOD1$^{WT}$* allele following Cre-LoxP recombination to enable expression of conditional exons 4' and 5', as present in *hSOD1$^{RDEL}$* mice. Ex, exon; UTR, untranslated region.

The aim of this chapter is to characterise the newly developed *hSOD1$^{WT}$* mouse line, with particular emphasis on phenotypes related to SOD1 loss-of-function, which, if present, could indicate that human SOD1 and mouse Sod1 have discrete properties which prevent them from being interchangeable. An outline of the experiments performed is shown in Figure 3-2. This *hSOD1$^{WT}$* mouse line is coisogenic with our humanised *hSOD1$^{A4V}$* mouse model of SOD1-ALS, and is therefore an ideal control for experiments involving *hSOD1$^{A4V}$* mice.

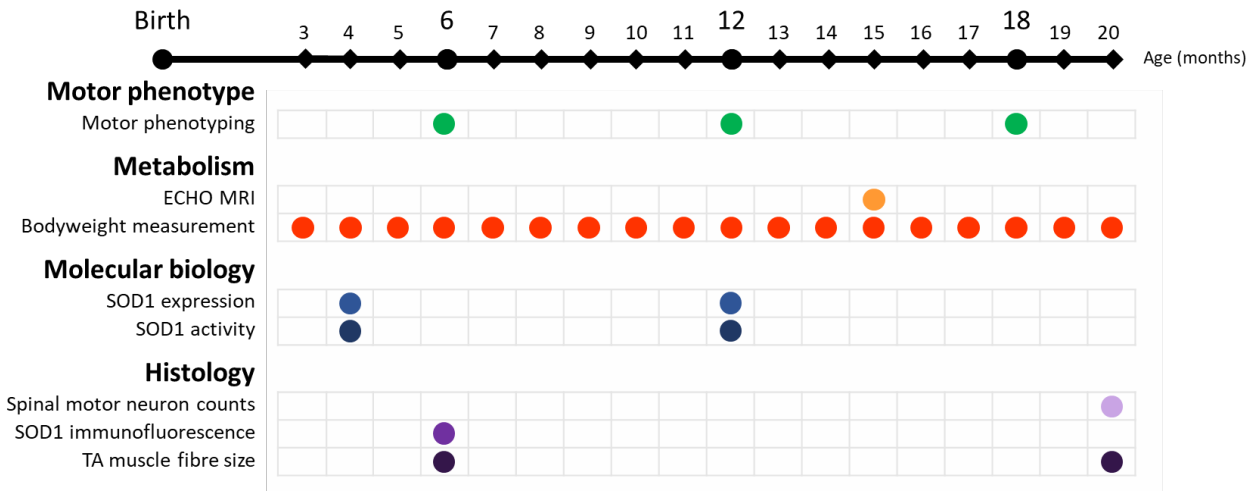



**Figure 3-2 | Outline of experiments performed on *hSOD1$^{WT}$* mice.**

## 3.2 Aims

1. To investigate the effect of *SOD1* humanisation in *hSOD1$^{WT}$* mice, we first examined *SOD1* RNA and protein expression, and SOD1 activity across several tissues. We then examined SOD1 localisation in the spinal cord of *hSOD1$^{WT}$* mice.
2. We performed longitudinal motor and behavioural phenotyping of *hSOD1$^{WT}$* mice at 6, 12 and 18 months of age.
3. Finally, we performed spinal motor neuron counts in aged *hSOD1$^{WT}$* mice.

## 3.3 Results

### 3.3.1 hSOD1-WT Mendelian ratios

To determine whether humanisation of *SOD1* affects the viability of *hSOD1$^{WT}$* mice, the number of offspring of the 3 potential genotypes was monitored in litters born from intercrosses between heterozygous *mSod1/hSOD1$^{WT}$* mice.

The ratio of animals of each genotype born from heterozygous *mSod1/hSOD1$^{WT}$* intercrosses did not differ significantly from the ratio predicted by Mendelian inheritance ($\chi^2(2)=0.29$, p=0.86) suggesting *SOD1* humanisation does not affect viability (Table 3-1).

| | **Male** | **Female** | **Sexes combined** | **Expected ratio** | **Observed ratio (sexes combined)** |
|---|---|---|---|---|---|
| *mSod1/mSod1* | 26 | 37 | 63 | 0.25 | 0.24 |
| *mSod1/hSOD1$^{WT}$* | 66 | 64 | 130 | 0.50 | 0.50 |
| *hSOD1$^{WT}$/hSOD1$^{WT}$* | 26 | 43 | 69 | 0.25 | 0.26 |
| Total | 118 | 144 | 262 | 1 | 1 |

**Table 3-1 | Genotypes of offspring born from intercrosses between heterozygous *mSod1/hSOD1$^{WT}$* mice.**

### 3.3.2 *hSOD1$^{WT}$* Allele QC (TLA and long-read sequencing)

To assess the integrity of the humanised *SOD1* locus we performed targeted locus amplification (TLA) followed by Nanopore long-read sequencing (Oxford Nanopore Technologies) using DNA from homozygous *hSOD1$^{WT}$/hSOD1$^{WT}$* mice. TLA was performed using Xdrop enrichment (Samplix; Madsen et al., 2020), which allows for enrichment and subsequent high-depth sequencing of a target sequence of interest – in this case the *SOD1* locus – enabling detection of unwanted recombination events or structural rearrangements. Briefly, this approach uses high molecular weight DNA encapsulated into droplets as an input for PCR with detection primers against the loci of interest. Amplification occurs only in droplets containing the locus of interest, which are then isolated by

fluorescence-activated cell sorting (FACS). Subsequently, droplet-based multiple displacement amplification generates enough DNA from the loci of interest for high coverage Nanopore sequencing.

We designed 3 pairs of detection primers against the *SOD1* locus, one within the humanised *SOD1 gene* and two within the 149.9 kb 3' homology arm (Table 3-2). This provided a ~200 kb island of sequencing coverage across the *SOD1* locus. Sequencing read depth ranged from 50x to >1000x, with read depth highest around the detection sequences.

Alignment (Minimap2; Li, 2018) of reads to the C57BL/6J (GRCm38.p6) reference genome showed the loss of the mouse *Sod1* gene (Figure 3-3, A). Conversely, alignment to the human (hg38) reference genome showed an island of sequencing coverage corresponding to the inserted human *SOD1* gene (Figure 3-3, B). The engineered duplication of human *SOD1* exons 4 and 5, and the duplicated mouse *Sod1* 3'UTR, were evident as a 2-fold increase in coverage in read alignment to the human and mouse reference genomes, respectively. Alignment to a reference file containing the expected engineered allele indicated seamless humanisation of the *SOD1* locus, with no deletion or insertion events detectable (Figure 3-3, C).

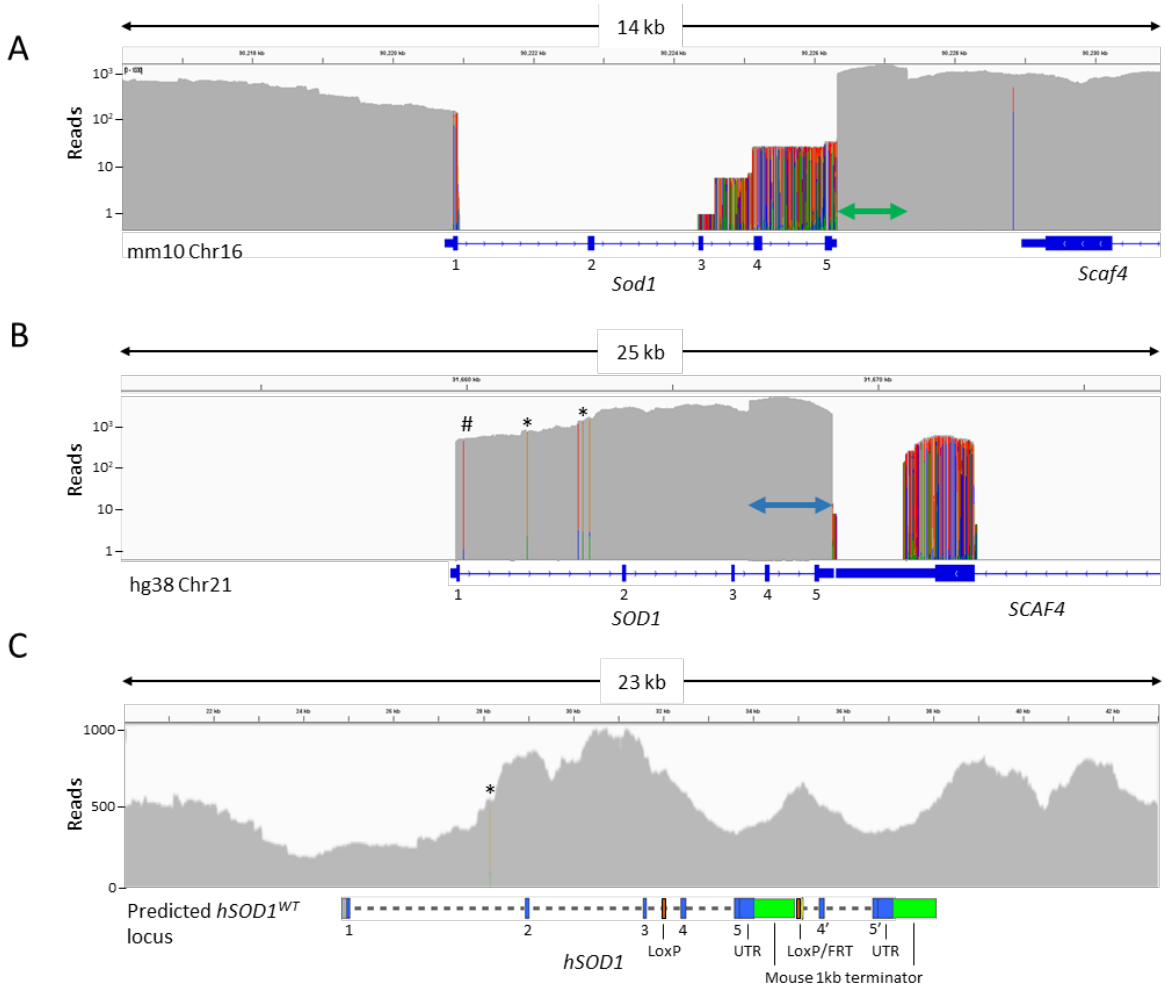



**Figure 3-3 | TLA and long-read sequencing of the humanised *hSOD1$^{WT}$* locus.** Alignment of Nanopore sequencing reads to **(A)** the mouse (mm10) reference genome **(B)** the human (hg38) reference genome and **(C)** the predicted *hSOD1$^{WT}$* locus. Green arrow indicates the duplicated mouse 1 kb terminator. Blue arrow indicates the duplicated region of human *SOD1*. Arrowhead indicates G93A mutation. Asterisks indicate flagged nucleotides that were checked manually. Hashtag symbol indicates site of a human intronic SNP. Figure adapted from Nair et al., 2020 [429].

The C57BL/6 and 129S1 genomes are polymorphic at a large number of loci. The mouse homology arms in the *SOD1* humanisation targeting vector were derived from C57BL/6N; however, the targeting was carried out in 129 ESCs. By aligning reads to the 129S1 SvImJ (GCA_001624185.1) and C57BL/6J reference genomes and looking for mismatches, we were able to estimate the extent of recombination of the 149.9 kb 3' mouse homology arm to be between 86-105 kb 3' of the humanised *SOD1* locus (Figure 3-4). The C57BL/6J reference genome was used rather than C57BL/6N because there are no known differences between the two strains at the *Sod1* locus, and the C57BL/6J reference genome is the most complete available.

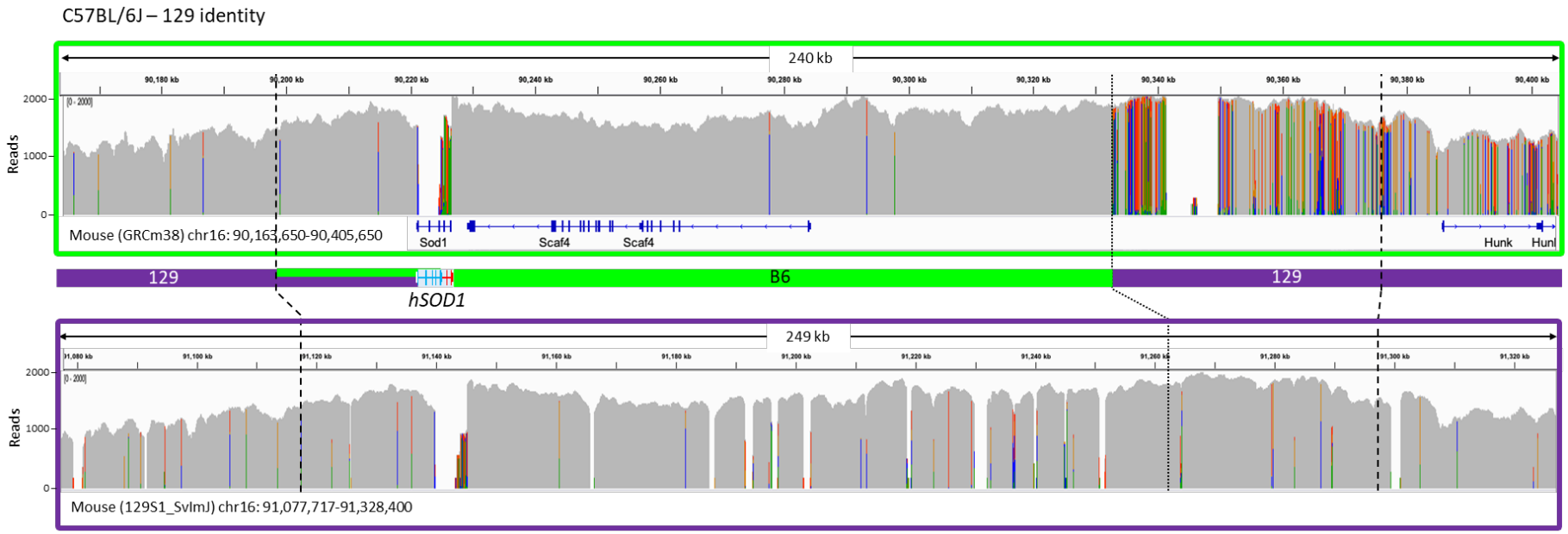



**Figure 3-4 | Strain identity of the humanised *hSOD1$^{WT}$* locus.** Comparison of alignment of Nanopore sequencing reads to the B6 and 129 genomes. Dashed lines indicate the boundaries of the mouse homology arms. Dotted line indicates the transition from B6 to 129 strain identity. Figure adapted from Nair et al., 2020 [429].

Our alignment seamlessly covered a region larger than the humanisation targeting construct, and thus we can be confident that the integration was correct and occurred solely at the *Sod1* locus. After accounting for the humanised region and strain differences, very few alignment imperfections remained; these misalignments were checked manually.

There was one misalignment with the human hg38 reference genome, a known human SNP situated in intron 1 of the *SOD1* gene; this SNP was present in the BAC RP11-535E10 BAC, from which the *hSOD1*$^{WT}$ targeting construct was derived. A number of nucleotides were flagged in our alignment to the C57BL/6J genome because they showed an increased error rate in the sequencing. Flagged nucleotides appeared exclusively in non-coding regions and mapped to polymorphic sequences – such as low complexity sequences or repetitive elements – which are prone to errors in nanopore sequencing [438]. Flagged nucleotides were checked manually and all had the expected base called in the majority of reads. A number of gaps were present in the 129S1 SvImJ strain alignment; however, the 129S1 SvImJ reference genome is less complete than C57BL/6J, and all alignment gaps corresponded to gaps in the reference 129 genome (denoted by strings of Ns).

| Assay | Primer | Sequence (5'→ 3') |
|---|---|---|
| Xdrop; dPCR enrichment; hSOD1 | forward primer | GAATGTCATCCCCTGGTGCA |
| | reverse primer | GTGCCCTTTACTTGGTGCAG |
| Xdrop; qPCR enrichment validation; hSOD1 | forward primer | TGATCCCAGGTGTTTGAGCT |
| | reverse primer | GCCTGTGTTTGGGGTGTAGA |
| Xdrop; dPCR enrichment; Sod1 3' homology arm proximal | forward primer | GCCCATGGCTTCCTTTACAC |
| | reverse primer | AAGCAAAAGTGGCAGGAAGC |
| Xdrop; qPCR enrichment validation; Sod1 3' homology arm proximal | forward primer | TATGTCTCCAGGCCAGGCT |
| | reverse primer | GCCTTGTCACGCCTTGTAAC |
| Xdrop; dPCR enrichment; Sod1 3' homology arm distal | forward primer | GCCCAACAACTCGGTAGGAA |
| | reverse primer | CGTAGCCTCAGTCAGCAAGT |
| Xdrop; qPCR enrichment validation; Sod1 3' homology arm distal | forward primer | AGTATGGGTTGCAGGCTGTC |
| | reverse primer | CGCTACAGTGGGATGGAATGT |

**Table 3-2 | Detection primers for Xdrop enrichment of the *hSOD1$^{WT}$* locus.**

### 3.3.3 *SOD1* mRNA expression

To determine whether humanisation affects *SOD1* expression, we examined *SOD1* mRNA expression in a range of tissues from 4-month-old and 12-month-old *hSOD1$^{WT}$* animals and wild-type littermates using qPCR. Initial experiments were conducted using brain, spinal cord and quadriceps muscle from 4-month-old male mice (Figure 3-5, A). Further expression analyses were conducted in 4-month-old female animals and 12-month-old animals (Figure 3-5 B, C) using an expanded range of tissues including tibialis anterior (TA) muscle, kidney and liver. 4-month old male and female animals showed similar expression of *SOD1,* so animals of both sexes were pooled for expression analysis at 12 months old.

The expression of *SOD1* mRNA did not significantly differ between *hSOD1$^{WT}$* animals and wild-type littermates in brain, spinal cord and kidney tissue at any time point. However, *SOD1* expression was consistently decreased in muscle tissue from *hSOD1$^{WT}$* animals compared to wild-type littermate

controls. A significant decrease in *SOD1* expression was observed in all muscles tested from homozygous *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals, apart from quadriceps from 4-month-old female animals, where variability among wild-type animals prevented statistical significance despite a strong trend being observed. *SOD1* expression was also significantly decreased in liver tissue from homozygous *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ mice.

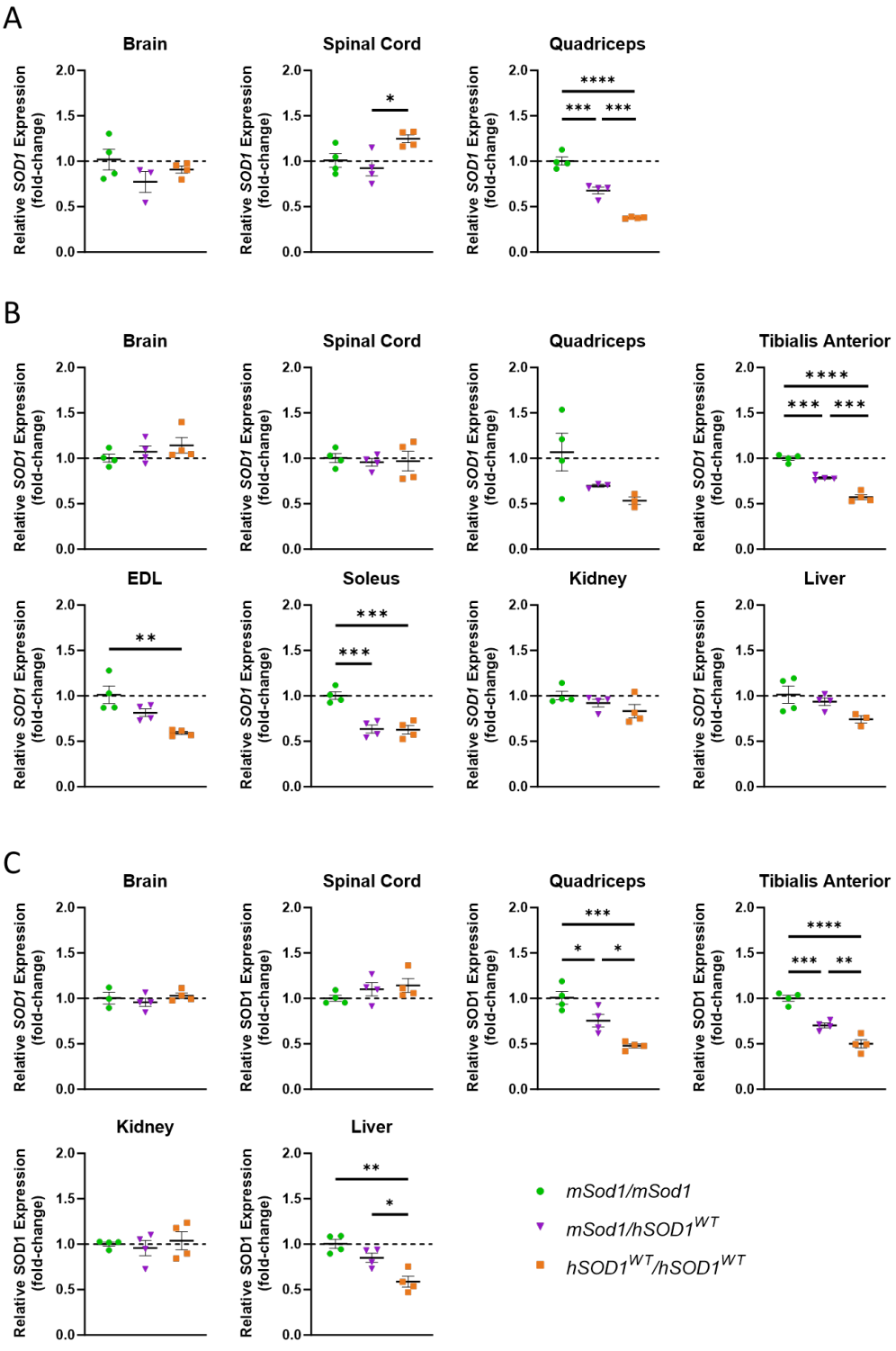



**Figure 3-5 | Expression of *SOD1* RNA in *hSOD1$^{WT}$* mice by qRT-PCR.** Expression of total *SOD1* mRNA using conserved mouse-human *SOD1* primers in tissues from **(A)** 4-month-old male animals, **(B)** 4-month-old female animals and **(C)** 12-month-old animals of both sexes. n=3-4 animals per genotype per tissue. *SOD1* expression was normalized to *GAPDH* expression as a housekeeping gene. All values are expressed as a fold change relative to the average *SOD1* expression in *mSod1/mSod1* control animals. Mean ± SEM. One-way ANOVA with Tukey's post hoc test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

### 3.3.4 SOD1 protein levels

In light of the tissue-specific decrease in *SOD1* mRNA in *hSOD1$^{WT}$* animals, SOD1 protein levels were assessed by Western blot in a range of tissues (Figure 3-5). Mouse Sod1 has higher mobility in PAGE than human SOD1, and thus mouse Sod1 appears as a lower band in Western blot. Total SOD1 protein levels were only directly compared in *mSod1*/*mSod1* and *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals, as the presence of 2 bands in heterozygous *mSod1*/*hSOD1$^{WT}$* impacted the densitometry analysis.

Consistent with the *SOD1* mRNA expression data, the level of SOD1 protein was unchanged in brain from 4-month-old and 12-month-old *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals. However, a small but significant (~25%) decrease in SOD1 level was observed in the spinal cord of 12-month-old *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals. The level of SOD1 protein was decreased in *hSOD1$^{WT}$*/*hSOD1$^{WT}$* liver tissue and in all *hSOD1$^{WT}$*/*hSOD1$^{WT}$* muscles examined at both time points. SOD1 protein was also slightly decreased in kidney from 4-month-old, but not 12-month-old, *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice.

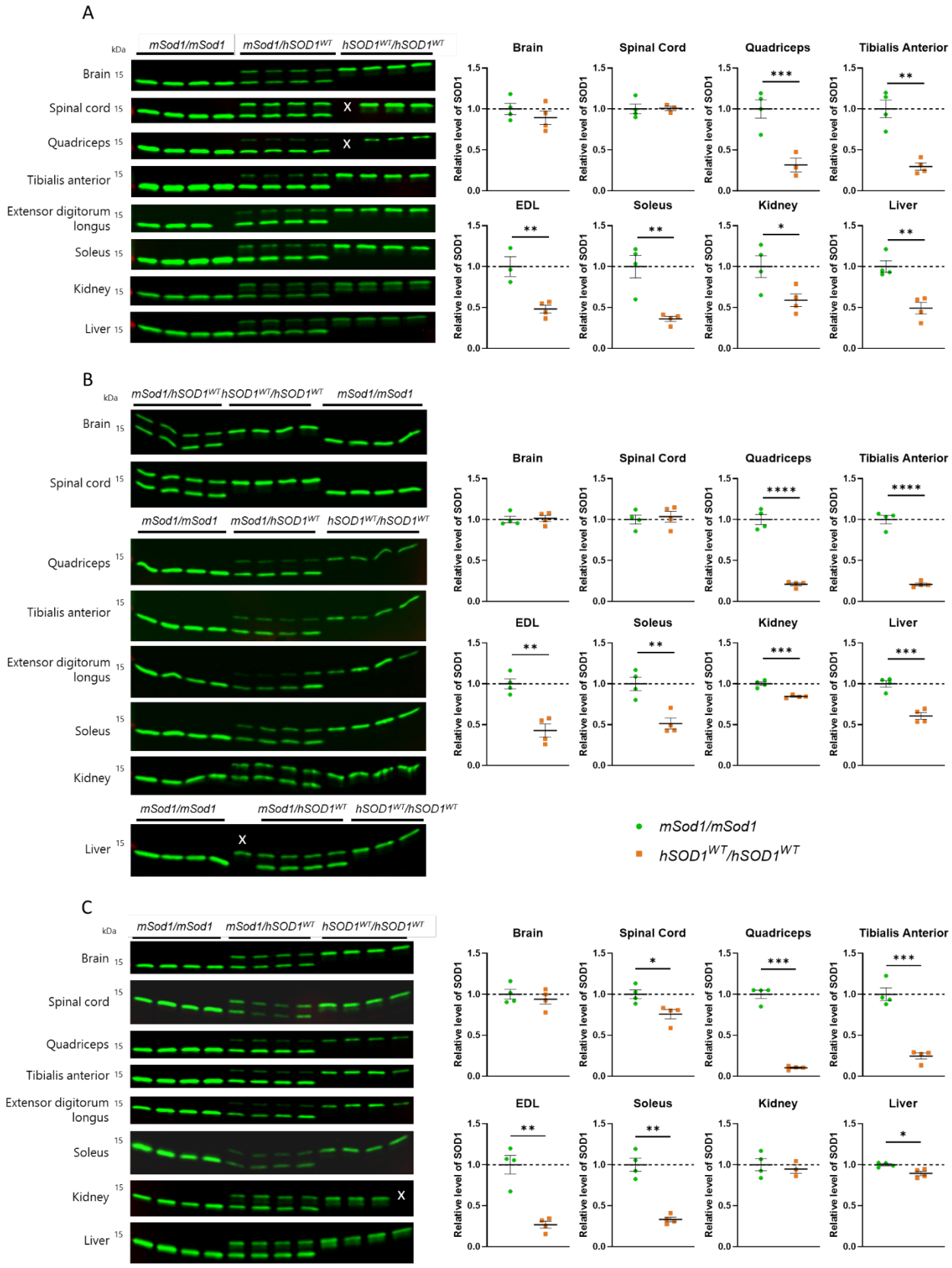



**Figure 3-6 | Levels of SOD1 protein in *hSOD1$^{WT}$* mice.** Western blots using a pan-mouse-human SOD1 antibody (aa131-153) in tissues from **(A)** 4-month-old male *hSOD1$^{WT}$* mice, **(B)** 4-month-old female *hSOD1$^{WT}$* mice and **(C)** 12-month-old *hSOD1$^{WT}$* mice of both sexes. Human SOD1 has lower mobility

than mouse Sod1 in SDS-PAGE and appears as a higher band. Blots normalized to total protein (Figure a-5). Unused lanes marked by x. All values are expressed as a fold change relative to the average Sod1 expression in *mSod1/mSod1* control animals. n=3-4 per genotype per tissue. Mean ± SEM, unpaired t-test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$

### 3.3.5 SOD1 protein levels following excision of duplicated exons

One possible cause for the decreased expression of *SOD1* in *hSOD1$^{WT}$* mice is transcriptional alterations caused by the conditional cassette containing duplicate exons 4' and 5'. To rule out this possibility, we examined SOD1 levels in mice *hSOD1$^{RDEL}$* mice – which carry a version of the *hSOD1$^{WT}$* allele where the conditional exons 4 and 5 have been excised by Cre recombinase (Figure 3-7).

The levels of SOD1 in 4-month-old male *hSOD1$^{RDEL}$* mice was decreased in a tissue-specific manner very similar to that seen in 4-month-old male *hSOD1$^{WT}$* mice. Specifically, SOD1 levels were unchanged in the brain, spinal cord and kidney of *hSOD1$^{RDEL}$* mice. However, SOD1 levels were decreased in skeletal muscle and liver tissue from homozygous *hSOD1$^{RDEL}$/hSOD1$^{RDEL}$* mice. These findings suggest the duplicated conditional sequence present in *hSOD1$^{WT}$* mice is not the cause of the observed tissue-specific reduction in *SOD1* transcription.

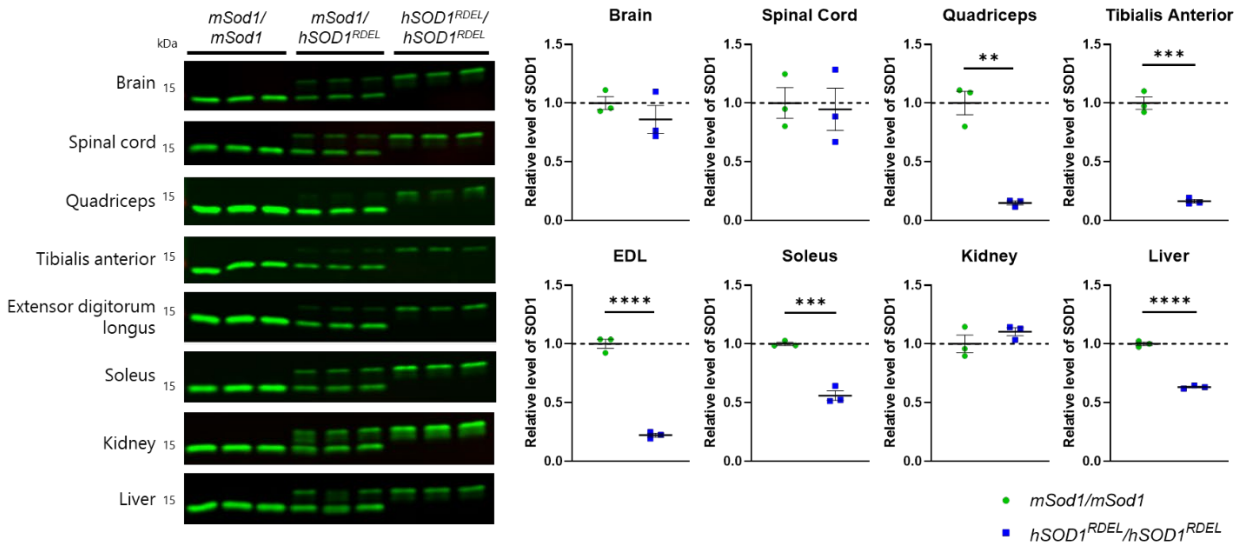



**Figure 3-7 | Levels of SOD1 protein in *hSOD1$^{RDEL}$* mice.** Western blots using a pan-mouse-human SOD1 antibody (aa131-153) in tissues from 4-month-old male *hSOD1$^{RDEL}$* mice. Blots normalized to total protein (Figure a-5). All values are expressed as a fold change relative to the average Sod1 expression in *mSod1/mSod1* control animals. n=3 per genotype per tissue. Mean ± SEM, unpaired t-test. **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

### 3.3.6 SOD1 dismutase activity

Although several non-canonical functions of SOD1 are emerging [41], the main function of SOD1 remains the dismutation of superoxide radicals. Hence, it is essential that human SOD1 should display dismutase activity in the context of the mouse cell if it is to functionally replace mouse Sod1. To confirm that human SOD1 is enzymatically active in *hSOD1$^{WT}$* mice we performed in-gel dismutase activity assays on a range of tissues (Figure 3-8).

In *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice, we found evidence of enzymatically active homodimers of human SOD1 in all investigated tissues. Furthermore, the dismutase activity of human SOD1 homodimers was approximately equivalent to mouse Sod1 homodimers in brain and spinal cord – where human SOD1 and mouse Sod1 are expressed at similar levels. In quadriceps muscle, SOD1 dismutase activity was approximately correlated with the level of SOD1 expression in the different genotypes. Finally, enzymatically active heterodimers of human and mouse SOD1 have previously been reported in SOD1 transgenic mice [338]; similarly, we observed functional mouse-human SOD1 heterodimers in *mSod1*/*hSOD1$^{WT}$* animals. These findings confirm that human SOD1 is enzymatically active in *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice.

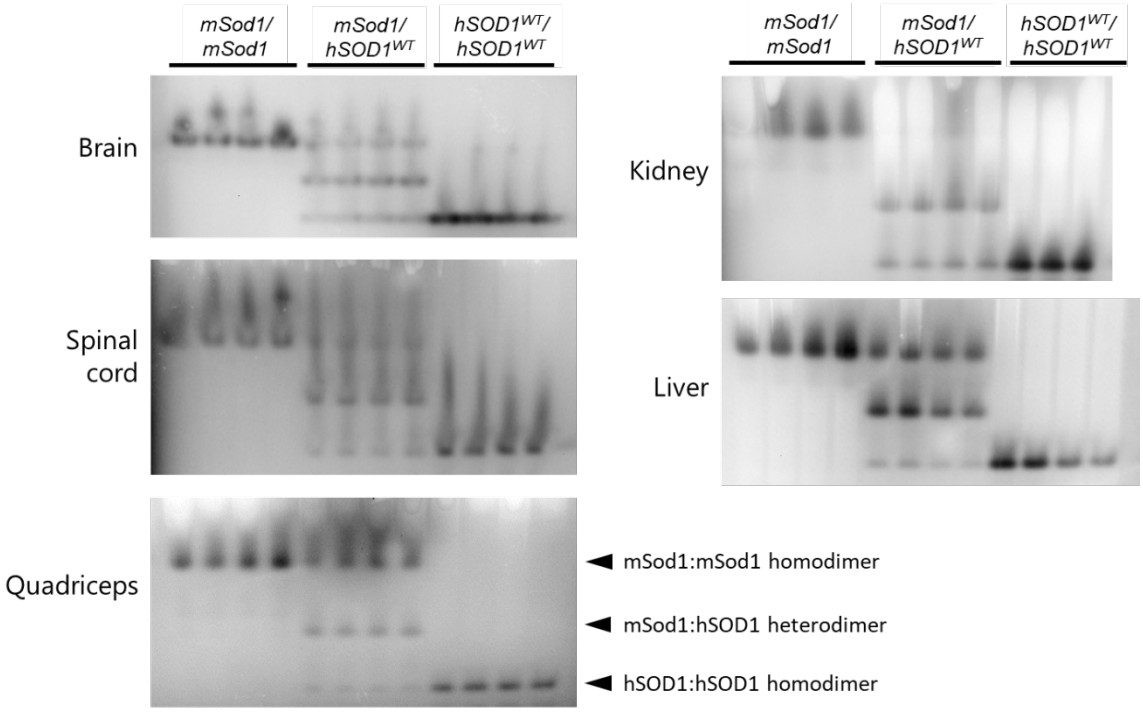



**Figure 3-8 | SOD1 dismutase activity in *hSOD1$^{WT}$* mice.** Colorimetric in-gel SOD1 dismutase activity assays in tissues from 12-month-old *hSOD1$^{WT}$* mice. Darker staining indicates higher dismutase activity. Human SOD1 has higher mobility than mouse Sod1 in clear-native PAGE and appears as a lower band.

### 3.3.7 SOD1 Immunofluorescence

We examined SOD1 localisation in the spinal cord of 6-month-old male *mSod1*/*mSod1* and *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice using a pan-SOD1 antibody (Figure 3-9). We did not observe any obvious difference in the distribution or subcellular localisation of SOD1 immunoreactivity in *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice vs littermate controls. In both genotypes SOD1 immunoreactivity was primarily cytoplasmic, with both diffuse and punctate staining patterns observed.

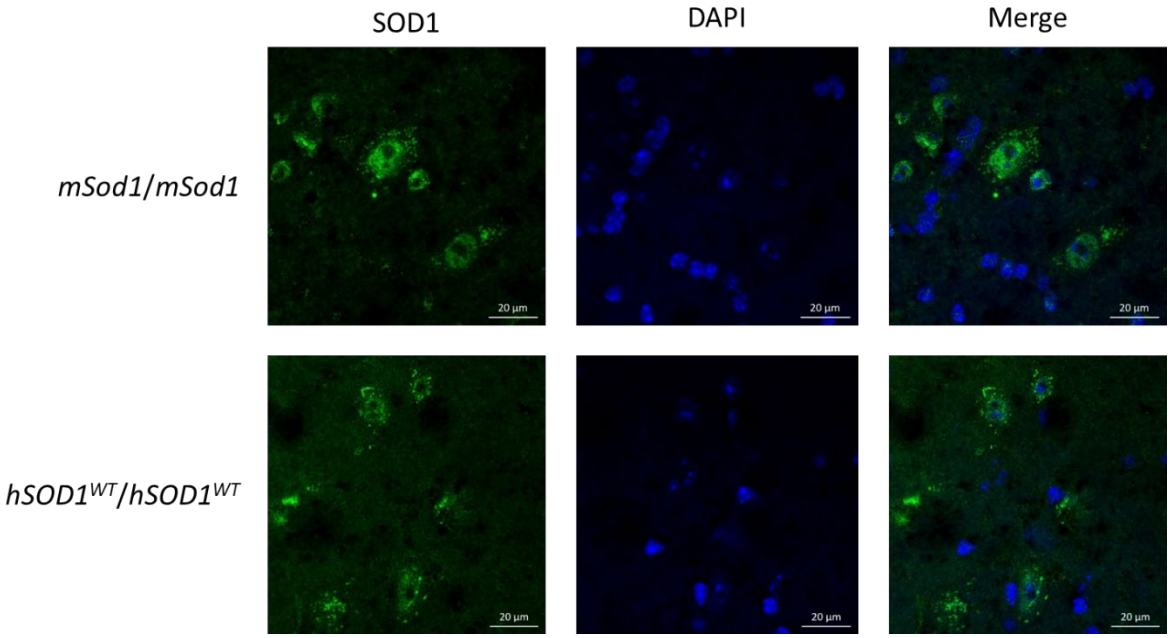



**Figure 3-9 | SOD1 immunoreactivity in *hSOD1$^{WT}$* spinal cord.** SOD1 immunoreactivity in lumbar spinal cord from male 6-month-old mice using a pan-mouse-human SOD1 antibody. Scale bar 20µm.

### 3.3.8 hSOD1-WT phenotypic characterisation

To assess the effect of *SOD1* humanisation on motor function in *hSOD1$^{WT}$* mice, we performed longitudinal phenotyping using a battery of phenotyping tests. Phenotyping was conducted at 6, 12 and 18 months of age.

#### *3.3.8.1 hSOD1-WT Kaplan–Meier survival curves*

Survival analysis was conducted on animals who reached a humane end point (defined in section 2.2.3) or who died unexpectedly. Animals who died due to non-disease related causes (e.g. fighting) were not included in the analysis.

There was no significant effect of genotype on survival in male ($\chi^2(2)=0.9044$, p=0.64) or female ($\chi^2(2)=0.71$, p=0.70) *hSOD1$^{WT}$* mice (Figure 3-10). The majority of mice in every group reached the study end-point of 18 months old. For animals culled for health concerns, the disease-related causes of death/humane end points reached are described in Table 3-3.

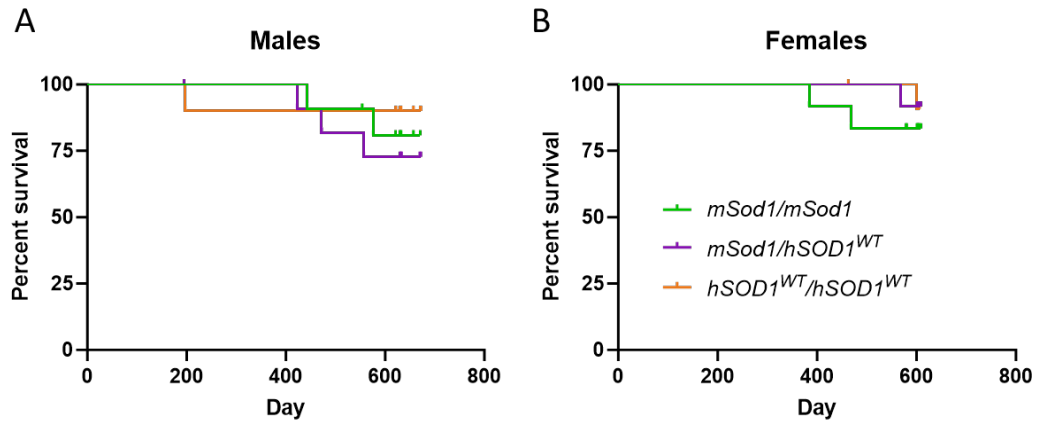



**Figure 3-10 | Kaplan-Meier survival graphs for *hSOD1$^{WT}$* mice.** Kaplan-Meier curves for **(A)** male (n=10-12 per genotype) and **(B)** female (n=12 per genotype) *hSOD1$^{WT}$* mice. Mice that died from non-disease related causes (e.g. mice culled from fight wounds, or for tissue collection) are censored at the time of death. Censored animals are indicated by tick marks. The study end point was approximately 650 days old for male animals and 600 days old for female animals.

| Cause of death | Percentage of animals (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *mSod1/ mSod1* | | | *mSod1/ hSOD1$^{WT}$* | | | *hSOD1$^{WT}$/ hSOD1$^{WT}$* | | | Total | | |
| | m | f | t | m | f | t | m | f | t | m | f | t |
| Mass in abdomen | - | 7.7 | 4.2 | 15.8 | 8.3 | - | - | 8.3 | 4.5 | 7.5 | 8.1 | 7.8 |
| Found dead | 9.1 | 7.7 | 8.3 | 5.3 | - | 3.2 | - | - | 0.0 | 5.0 | 2.7 | 3.9 |
| Weight loss | - | 7.7 | 4.2 | - | - | 0.0 | - | - | 4.5 | 2.5 | 2.7 | 2.6 |
| Swollen eye | 9.1 | - | 4.2 | - | - | 0.0 | - | - | 0.0 | 2.5 | 0.0 | 1.3 |
| Seizure | - | 7.7 | 4.2 | - | - | 0.0 | - | - | 0.0 | 0.0 | 2.7 | 1.3 |
| **Group n** | 11 | 13 | 24 | 19 | 12 | 31 | 10 | 12 | 22 | 40 | 37 | 77 |

**Table 3-3 | Disease-related causes of death of *hSOD1$^{WT}$* mice.** Mice culled due to fighting are excluded from this analysis.

### *3.3.8.2 hSOD1-WT motor and behavioural phenotyping*

#### 3.3.8.2.1 SHIRPA

We assessed *hSOD1$^{WT}$* mice for motor and behavioural abnormalities using the SHIRPA test (Table 3-4). We observed loss of movement fluidity in a slightly larger subset of *hSOD1$^{WT}$/hSOD1$^{WT}$* mice, compared to *mSod1/mSod1* animals, however this difference was not significant in the Chi-square test at any time point. Furthermore, individual animals were not consistently affected across different time points, suggesting the observed effect is unlikely to represent a genuine decline in motor function. There was no significant change in tremors, trunk curl, or the ability to walk down a vertical grate (negative geotaxis) in *hSOD1$^{WT}$* mice compared to littermate controls.

| Genotype | Sex | Sample size | Tremors | | Loss of movement fluidity | | Trunk curl | | Negative geotaxis | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Onset (months) | % | Onset (months) | % | Onset (months) | % | Onset (months) | % |
| ***mSod1/mSod1*** | Male | 11 | - | - | - | - | - | - | 12 | 10 |
| ***mSod1/hSOD1$^{WT}$*** | Male | 12 | 12 | 18 | 6 | 20 | - | - | 12 | 30 |
| ***hSOD1$^{WT}$/hSOD1$^{WT}$*** | Male | 10 | - | - | 18 | 30 | - | - | 18 | 20 |
| ***mSod1/mSod1*** | Female | 12 | - | - | 18 | 10 | - | - | - | - |
| ***mSod1/hSOD1$^{WT}$*** | Female | 12 | - | - | 12 | 16 | - | - | 18 | 8 |
| ***hSOD1$^{WT}$/hSOD1$^{WT}$*** | Female | 12 | - | - | 6 | 27 | - | - | 18 | 9 |

**Table 3-4 | SHIRPA phenotyping results from *hSOD1$^{WT}$* mice.** Only parameters related to motor function are shown.

#### 3.3.8.2.2 Open field

The open field test was used to measure locomotor and anxiety-like behaviour in *hSOD1$^{WT}$* mice (Figure 3-11). There was no difference between the genotypes in time spent in the centre vs the periphery of the open field apparatus at any time point, indicating that *hSOD1$^{WT}$* mice do not display altered anxiety-like behaviour when compared to wild-type littermates.

There was no difference in locomotor activity between the genotypes in male *hSOD1$^{WT}$* mice at any time point. However, female *mSod1*/*hSOD1$^{WT}$* mice displayed increased locomotor activity at 18 months old.

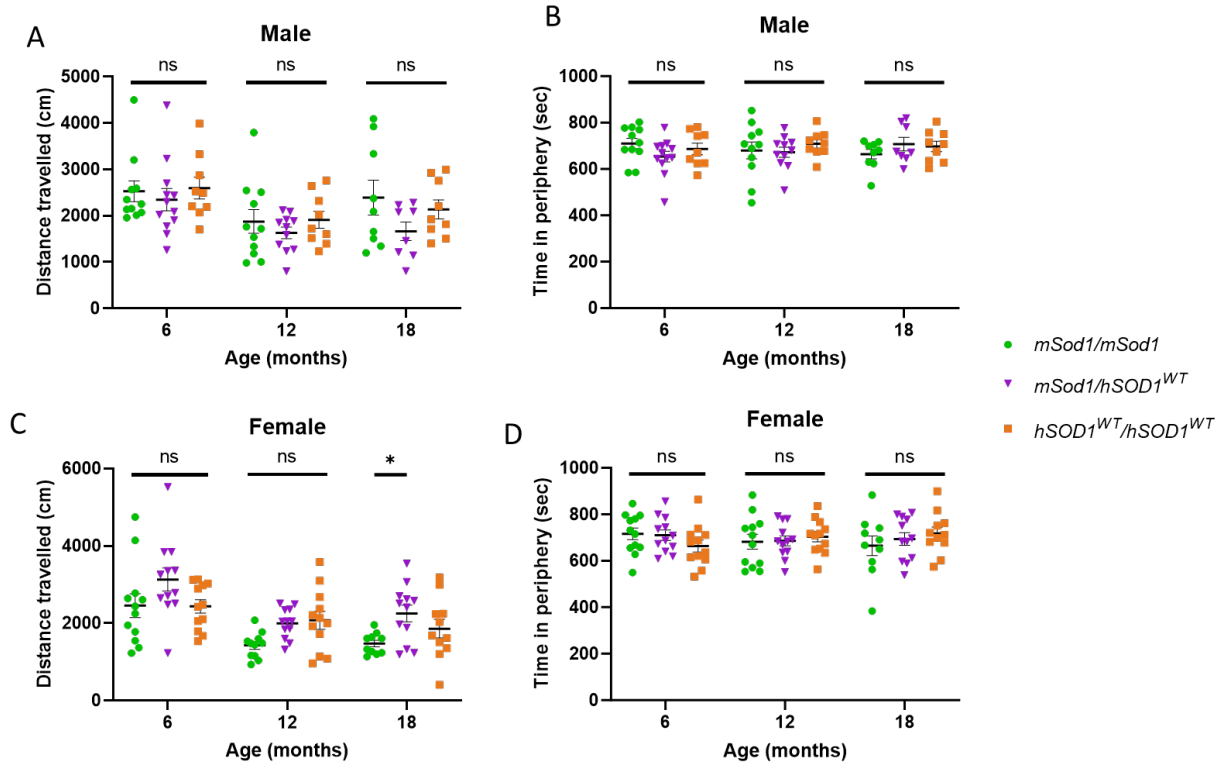



**Figure 3-11 | Longitudinal open field test data from *hSOD1$^{WT}$* mice.** Open field test data from **(A, B)** male (n=8-12 per genotype) and **(C, D)** female (n=10-12 per genotype) *hSOD1$^{WT}$* mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. *$P$ < 0.05. ns, not significant.

#### 3.3.8.2.3 Grip strength

There was no significant effect of genotype on grip strength in *hSOD1$^{WT}$* mice at any time point tested (Figure 3-12).

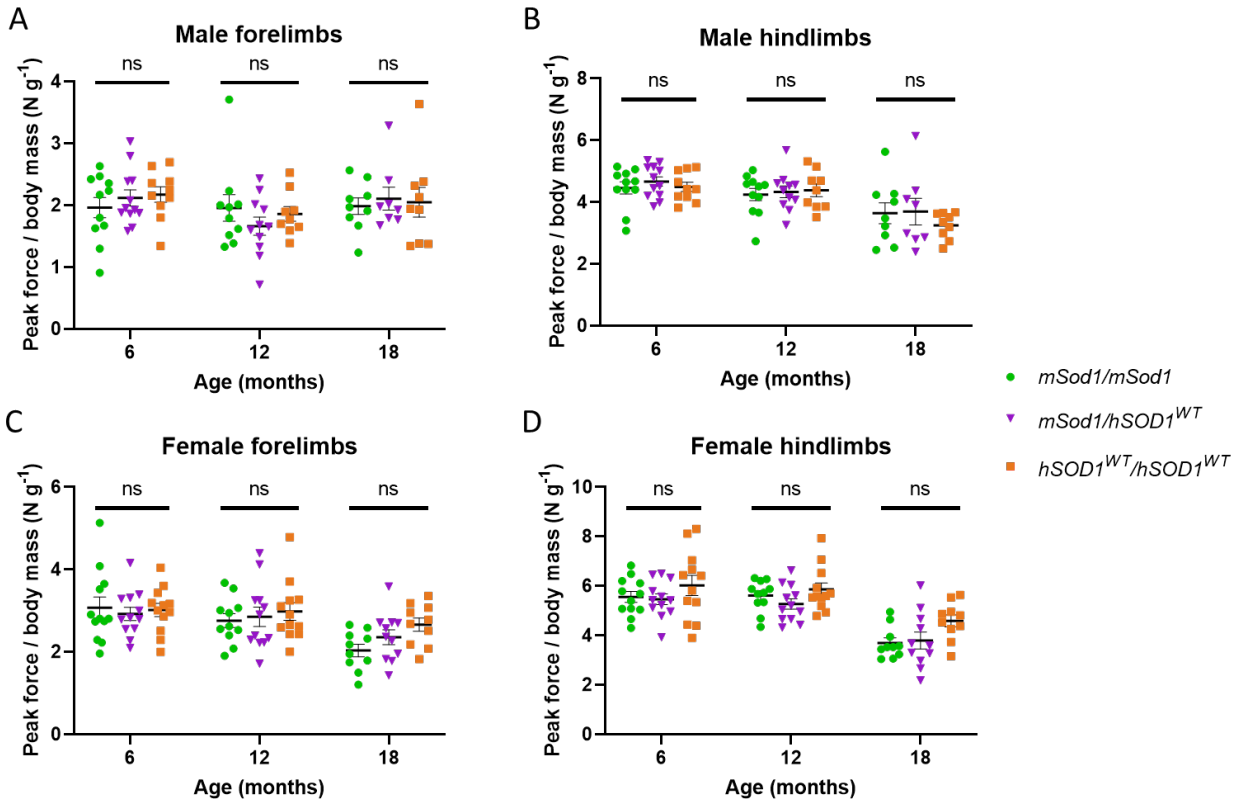



**Figure 3-12 | Longitudinal grip strength data from *hSOD1$^{WT}$* mice.** Grip strength data from **(A, B)** male (n=8-12 per genotype) and **(C, D)** female (n=10-12 per genotype) *hSOD1$^{WT}$* mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. ns, not significant.

#### 3.3.8.2.4 Gait analysis

We performed gait analysis on videos of freely walking *hSOD1$^{WT}$* mice obtained using the MouseWalker system [411]. Differences between the genotypes were assessed across 13 different gait parameters in 6, 12 and 18-month old mice using Bonferroni-corrected ANOVA. We did not observe any significant differences between the genotypes for any of the gait parameters measured at any time point.

#### 3.3.8.2.5 Locotronic

We examined motor coordination in *hSOD1$^{WT}$* mice using the Locotronic test (Figure 3-13). There was a significant decrease in the number of leg errors in 6-month-old female *hSOD1$^{WT}$/hSOD1$^{WT}$* animals when compared to *mSod1/hSOD1$^{WT}$* littermates. There was also a significant decrease in leg

error time in 6-month-old female *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals compared to WT and heterozygous *mSod1*/*hSOD1*$^{WT}$ littermates. These results would be indicative of an improvement in motor coordination at this time point; however, no significant differences were observed at later time points.

There was a significant increase in leg error time, but not leg error number, in 18-month-old male *mSod1*/*hSOD1*$^{WT}$ and *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ compared to wild-type littermates. This may be indicative of a mild impairment in motor coordination in these animals.

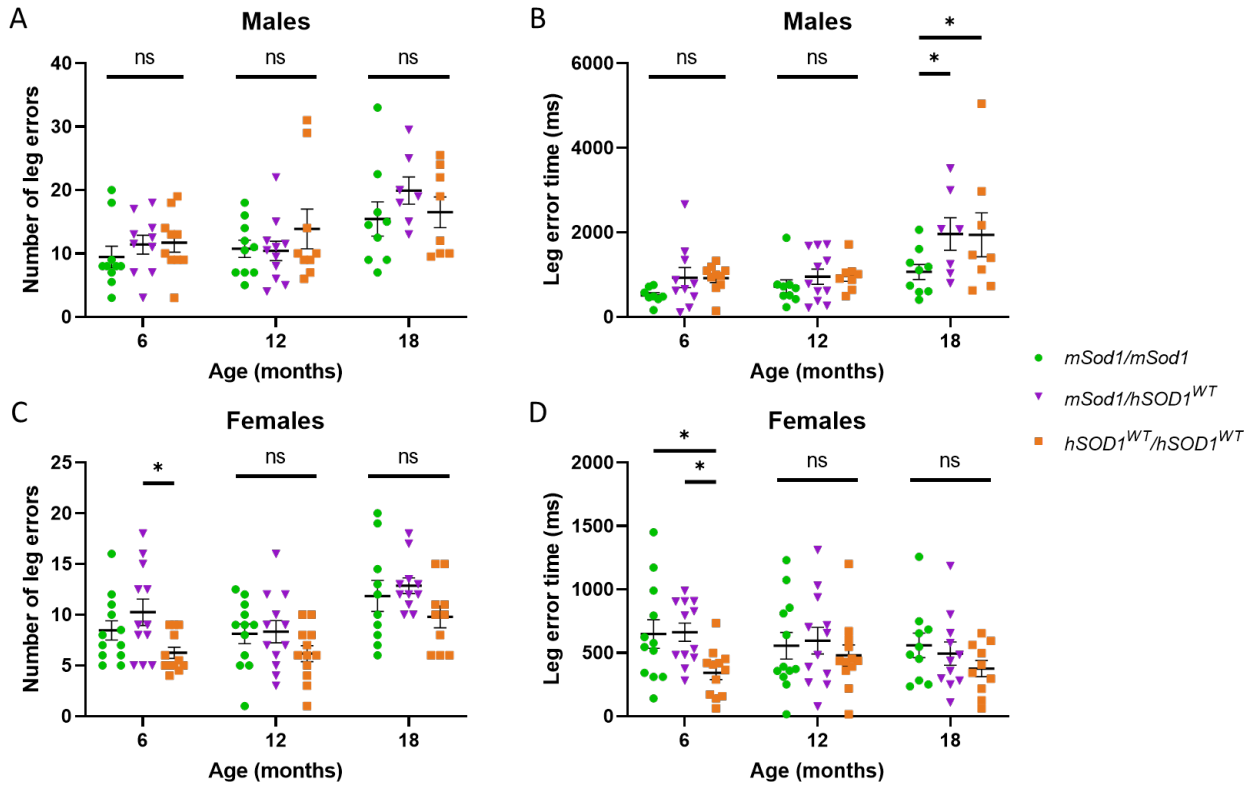



**Figure 3-13 | Longitudinal Locotronic data from *hSOD1*$^{WT}$ mice.** Locotronic data from **(A, B)** male (n=8-12 per genotype) and **(C, D)** female (n=10-12 per genotype) *hSOD1*$^{WT}$ mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. *$P < 0.05$. ns, not significant.

#### 3.3.8.2.6 Rotarod

There was no difference in rotarod performance between the genotypes of male *hSOD1*$^{WT}$ mice at any time point (Figure 3-14). However, female *mSod1*/*hSOD1*$^{WT}$ and *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ mice

displayed improved rotarod performance at 18 months old. This result was unexpected, as impaired motor function typically correlates with a decrease in performance in the rotarod test. This result also cannot be explained by SOD1 loss-of-function as *Sod1* knockout mice display impaired rotarod performance [439].

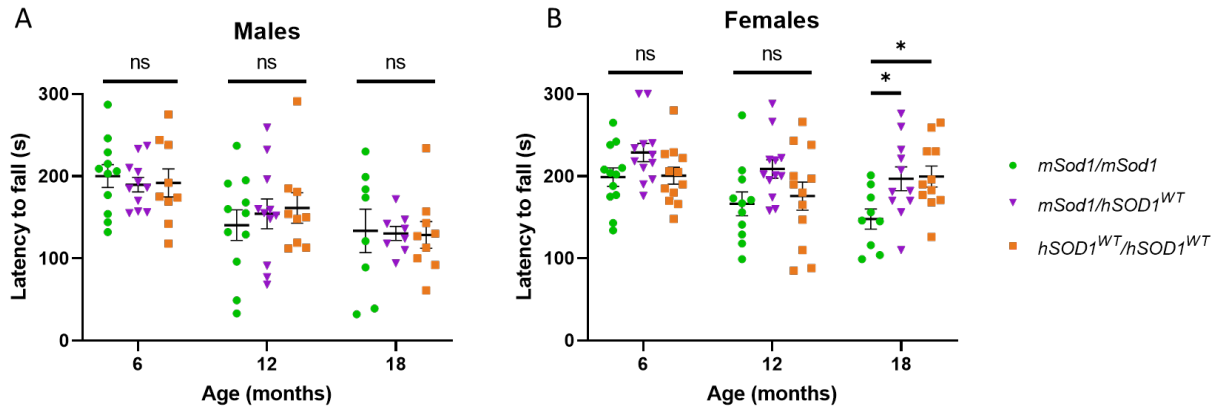



**Figure 3-14 | Longitudinal Rotarod data from *hSOD1$^{WT}$* mice.** Rotarod data from **(A)** male (n=8-11 per genotype) and **(B)** female (n=9-12 per genotype) *hSOD1$^{WT}$* mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. *$P < 0.05$. ns, not significant.

#### 3.3.8.2.7 Automated wheel running

We examined voluntary wheel running activity in 12-month-old female *mSod1*/*mSod1* and *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals, and 18-month-old male and female *mSod1*/*mSod1* and *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals (Figure 3-15). We measured maximum speed, distance travelled, time spent running and the number of bouts of activity. We did not observe any differences between the genotypes in any of these parameters for either sex at either time point.

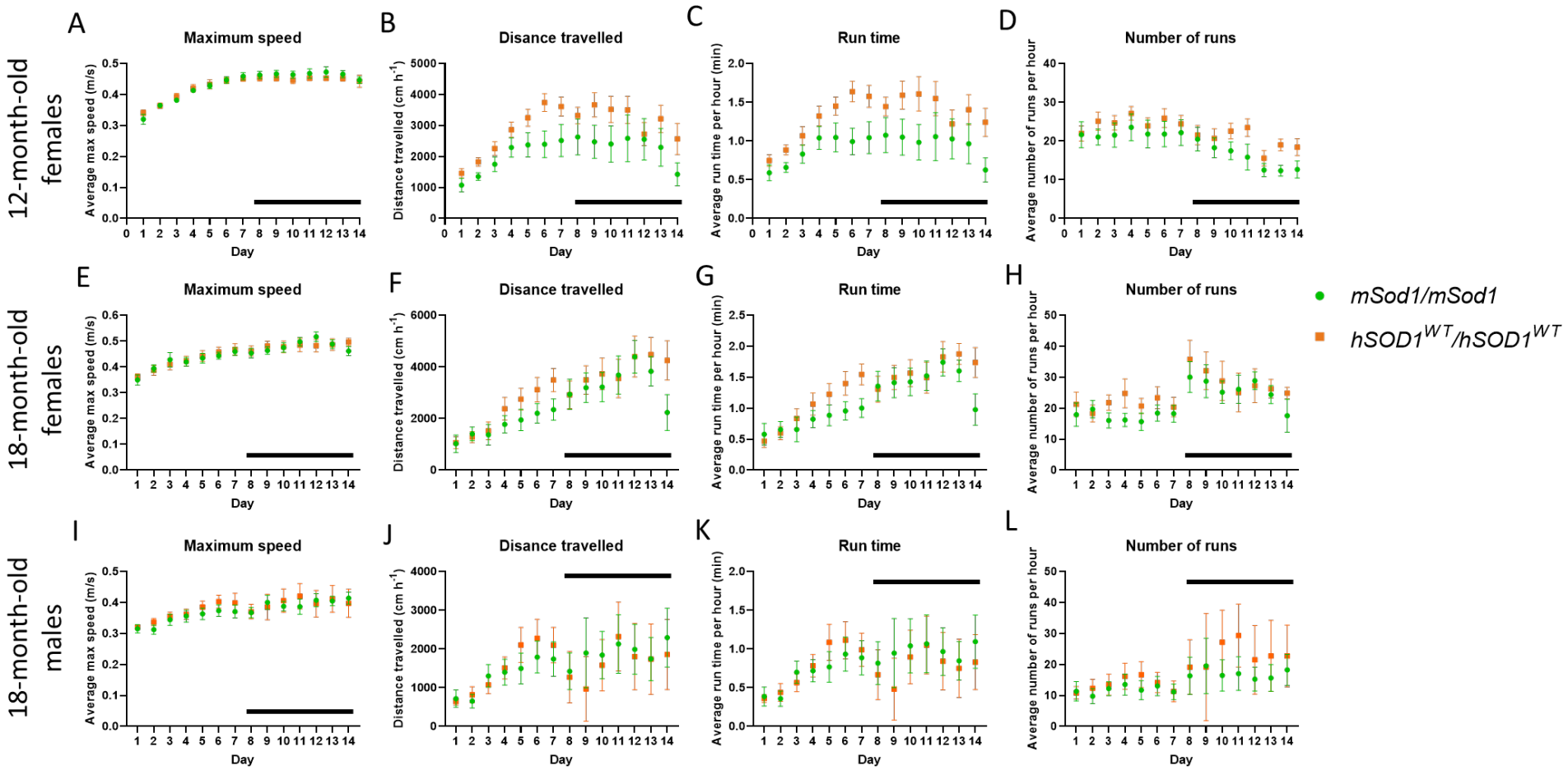



**Figure 3-15 | Longitudinal voluntary wheel running data from *hSOD1$^{WT}$* mice.** Voluntary wheel running data from **(A-D)** 12-month-old female (n=9 *mSod1/mSod1,* 10 *hSOD1$^{WT}$/hSOD1$^{WT}$*), **(E-H)** 18-month-old female (n=8 per genotype) and **(I-L)** 18-month-old male (n=8 per genotype) *hSOD1$^{WT}$* mice. No significant results found in mixed effects analysis followed by the Šídák multiple comparisons test. Mean ± SEM. Black bar indicates days using a complex wheel.

### 3.3.8.3 Bodyweight analysis

To investigate if *SOD1* humanisation has an effect on bodyweight, monthly weight measurements were conducted on *mSod1*/*mSod1* and *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals (Figure 3-16). The bodyweight of male *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ did not differ significantly from littermate *mSod1*/*mSod1* controls. However, female *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals were significantly lighter than littermate *mSod1*/*mSod1* controls (p=0.034), as determined by a linear mixed-effect model analysis. The lighter bodyweight in *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ female mice is unlikely to be caused by developmental factors as their weight was identical to *mSod1*/*mSod1* controls until approximately 16 weeks of age.

Pairwise comparisons conducted at each time point were not significant in either sex. The bodyweight of heterozygous *mSod1*/*hSOD1*$^{WT}$ mice were not measured longitudinally.

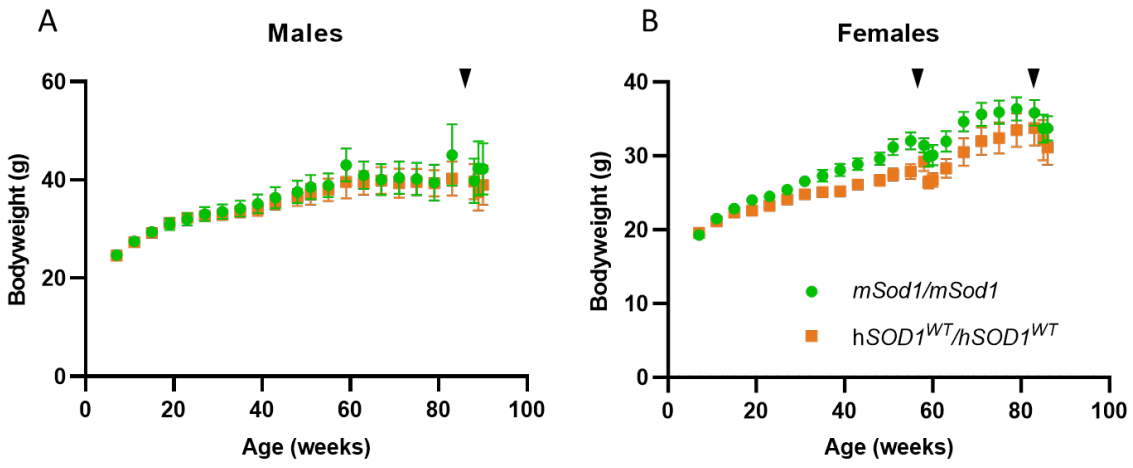



**Figure 3-16 | Longitudinal bodyweight data from *hSOD1*$^{WT}$ mice.** Monthly bodyweight measurement data from **(A)** male (n=10-11 per genotype) and **(B)** female (n=12-13 per genotype) *hSOD1*$^{WT}$ mice. Arrowheads indicate ages where automated wheel running experiments were performed. Mean ± SEM. Mixed-effects model analysis followed by the Šídák multiple comparisons test. No significant results were observed in multiple comparisons conducted at each time point. Main effects are reported in the thesis body.

#### 3.3.8.4 hSOD1-WT Echo-MRI

To assess genotype-specific differences in body composition, we performed Echo-MRI on 15-month-old *hSOD1$^{WT}$* animals (Figure 3-17). Male *hSOD1$^{WT}$*/*hSOD1$^{WT}$* and *mSod1*/*hSOD1$^{WT}$* animals did not display any significant changes in body composition compared to *mSod1*/*mSod1* controls. However, female *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice had significantly lower fat mass (p=0.035) and percentage body fat (p=0.009) than *mSod1*/*hSOD1$^{WT}$* littermates. Female *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals did not display a significant difference in lean mass, measured in grams, but did have significantly higher lean mass as a percentage of overall bodyweight when compared to *mSod1*/*hSOD1$^{WT}$* littermates (p=0.010) , an effect likely caused by the lower bodyweight in these animals. Body composition did not vary significantly between *hSOD1$^{WT}$*/*hSOD1$^{WT}$* and *mSod1*/*mSod1* mice, likely because *mSod1*/*hSOD1$^{WT}$* had slightly higher bodyweight and fat mass than *mSod1*/*mSod1* controls.

These observations are consistent with the lower bodyweight of female *hSOD1$^{WT}$*/*hSOD1$^{WT}$* animals; taken together, they suggest that the lower bodyweight in female *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice is due to a decrease in fat mass.

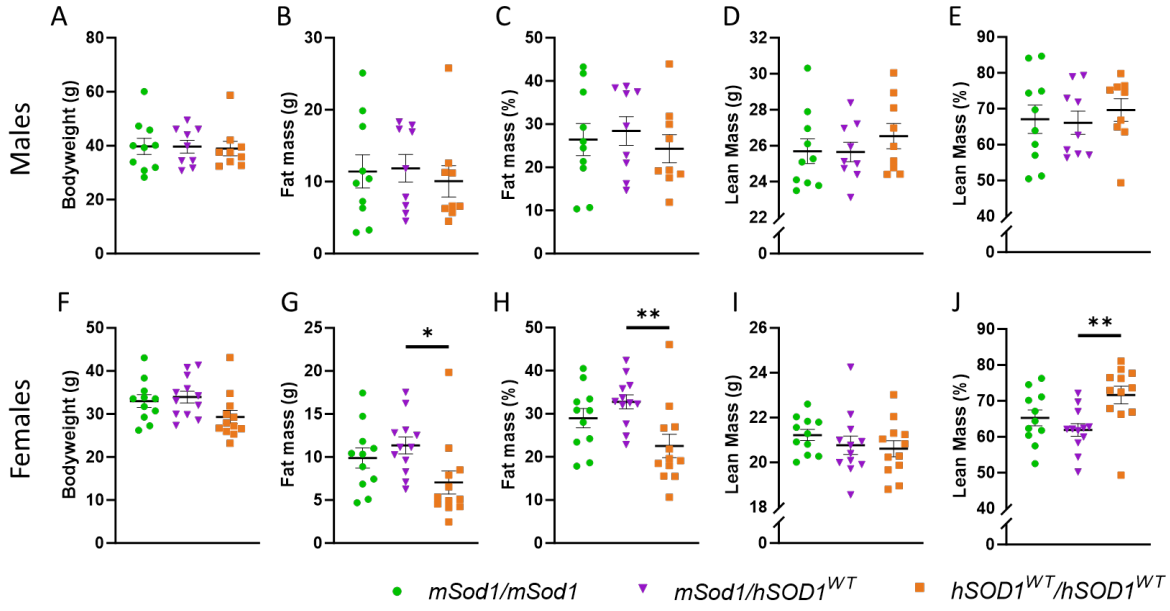



**Figure 3-17 | Echo MRI data from *hSOD1$^{WT}$* mice.** Total bodyweight, fat mass and lean mass in 15-month-old **(A-E)** male (n=9-10 per genotype) and **(F-J)** female (n=11-12 per genotype) *hSOD1$^{WT}$* animals. Mean ± SEM. One-way ANOVA with Tukey's post hoc test. *$P$ < 0.05, **$P$ < 0.01.

## 3.3.9 hSOD1-WT Histological characterisation

### *3.3.9.1 hSOD1-WT MN counts*

To investigate whether *SOD1* humanisation has an effect on motor neuron survival we counted the number of motor neurons in the sciatic pool of 18-month-old *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice and *mSod1*/*mSod1* controls (Figure 3-18). We did not observe any genotype-specific difference in the number of motor neurons in the L3-L5 region of the spinal cord, suggesting that *SOD1* humanisation does not detrimentally impact motor neuron survival.

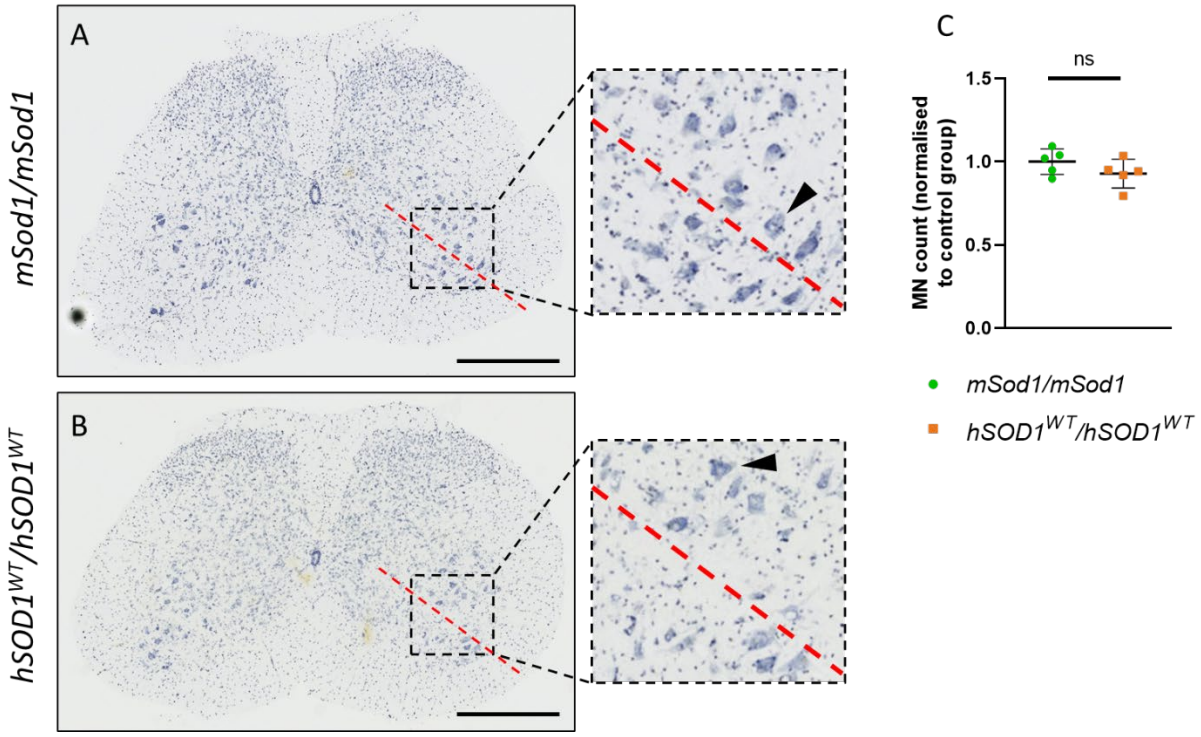



**Figure 3-18 | Spinal motor neuron counts from *hSOD1$^{WT}$* mice.** Representative images of lumbar spinal cord sections from 18-month-old **(A)** *mSod1/mSod1* and **(B)** *hSOD1$^{WT}$/hSOD1$^{WT}$* mice. Red dotted line indicates the lower boundary of the sciatic motor pool. Only motor neurons with cell bodies fully above this line were included in motor neuron counts. Arrowheads indicate cells displaying characteristic motor neuron features: polygonal shape, dense staining and a diameter greater than 20 µm. Scale bar = 500 µm. **(C)** Quantification of motor neurons within the sciatic pool of the lumbar spinal cord (L3-L5) of 18-month-old male *hSOD1$^{WT}$* animals. n=5 animals per genotype. Mean ± SEM. Unpaired t-test. ns, not significant.

### *3.3.9.2 hSOD1-WT muscle fibre size*

To examine the effect of *SOD1* humanisation on muscle atrophy we examined muscle fibre size and nuclear positioning in the tibialis anterior muscle from 4-month-old and 20-month-old male *hSOD1$^{WT}$/hSOD1$^{WT}$* mice and *mSod1/mSod1* controls. The tibialis anterior muscle is predominantly composed of fast twitch fibres and undergoes atrophy in animal models of SOD1-ALS [440], as well as in *Sod1* knockout animals [439].

We did not observe any difference in muscle fibre size or central nucleation between the genotypes at 4 or 20 months old (Figure 3-19).

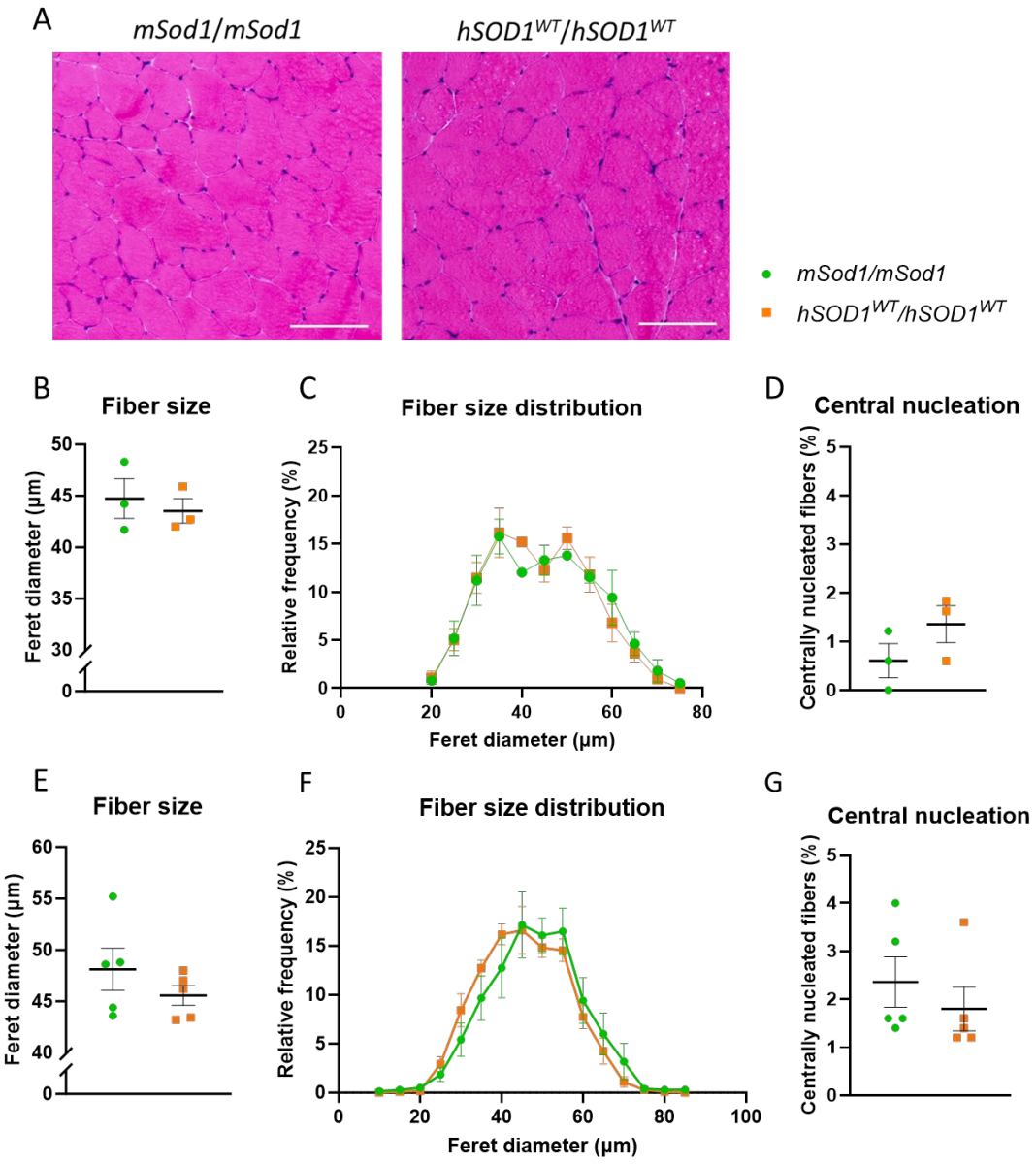



**Figure 3-19 | Muscle fiber size and central nucleation in *hSOD1$^{WT}$* mice. (A)** Representative images of tibialis anterior muscle from 4-month-old male *hSOD1$^{WT}$* mice stained with haematoxylin and eosin. Scale bar = 100µm. Quantification of **(B, C)** fiber size and **(D)** central nucleation in tibialis anterior muscle from 4-month-old male *hSOD1$^{WT}$* mice (n=3 per genotype). Quantification of **(E, F)** fiber size and **(G)** central nucleation in tibialis anterior muscle from 20-month-old male *hSOD1$^{WT}$* mice (n=5 per genotype). Mean ± SEM, unpaired t-test.

## 3.4 Discussion

To better understand the effect of SOD1 humanisation in *hSOD1*$^{WT}$ mice, we carried out immunohistochemical, biochemical, molecular and behavioural analyses in comparison to wild-type littermates.

We showed that *hSOD1*$^{WT}$ mice express human *SOD1* RNA and produce enzymatically active human SOD1 protein. As expected, *hSOD1*$^{WT}$ mice display normal viability, survival and motor function, and do not display motor neuron loss. Taken together, these data indicate that human SOD1 functionally replaces mouse Sod1 in *hSOD1*$^{WT}$ mice.

However, the decreased expression of human *SOD1* in muscle and liver was somewhat unexpected. Although the mouse *Sod1* promotor and 5'UTR were retained in *hSOD1*$^{WT}$ mice, the introns and 3' UTR were humanised, raising the possibility that these sequences contain transcriptional regulatory elements that differ between human and mouse. One possible site of species-specific regulation is the *SOD1* 3'UTR, which differs greatly between human and mouse. The human *SOD1* 3'UTR is 353bp in length, whereas the mouse *Sod1* 3'UTR is only 67bp. AUF-1 has been found to regulate human *SOD1* expression through binding to the *SOD1* 3'UTR [441], and the *SOD1* 3'UTR is a target of miRNAs in both human and mouse [442,443].

Additionally, the human *SOD1* gene contains two intronic ncRNAs that do not appear in the mouse *Sod1* gene (Ensembl). The function of these ncRNAs is unknown, but they could conceivably function to regulate *SOD1* expression.

The impact of the tissue-specific decrease in *SOD1* expression in *hSOD1*$^{WT}$ mice is still unclear. *Sod1*$^{-/-}$ null mice display a range of phenotypes, including reduced body weight from 4 months old [444,445] and declining motor performance from 9 months old, accompanied by muscle denervation and reduced muscle mass (reviewed in Saccon et al., 2013). The only phenotype we observed related to SOD1 loss of function (LOF) was lower bodyweight in female homozygous hSOD1$^{WT}$ mice. Importantly, muscle

pathology in *Sod1*$^{-/-}$ null mice appears to be a result of motor axonopathy, rather than loss of SOD1 in muscle tissue, as transgenic mice with *Sod1* expression restricted to neurons [446,447] or selectively deleted in muscle [448,449] do not display muscle pathology. Expression of *SOD1* in the CNS of *hSOD1*$^{WT}$ mice is normal, therefore we hypothesise that motor axonopathy does not occur. In this study we did not directly compare muscle denervation in *hSOD1*$^{WT}$ mice and *mSod1* controls; however, we did examine TA and lumbrical muscle denervation in *hSOD1*$^{WT}$ mice – as the control group for experiments involving *hSOD1*$^{A4V}$ mice – in chapter 4. We found denervation of both TA and lumbrical muscles to be very low in *hSOD1*$^{WT}$ mice, even at 24 months old (Figure 4-36), which is inconsistent with the progressive motor axonopathy observed in *Sod1*$^{-/-}$ mice from 2 months old [446,450,451]. These data, combined with the lack of any overt motor phenotype, suggest that motor axonopathy does not occur in *hSOD1*$^{WT}$ mice.

Further investigation into the causes of reduced human *SOD1* expression in certain tissues is not planned at this time. We have observed that the expression of human *SOD1* in hSOD1$^{A4V}$ mice follows the same pattern as in hSOD1$^{WT}$ mice, making *hSOD1*$^{WT}$ mice an essential control for future experiments.

SOD1 loss of function may play an important disease modulating role in SOD1-ALS [48]. Hence, future studies using the *hSOD1*$^{WT}$ mouse strain, and the closely related *hSOD1*$^{A4V}$ and *hSOD1*$^{G93A}$ strains, should be mindful of the decreased expression of *SOD1* in several tissues including, but likely not limited to, skeletal muscle and liver.

# Chapter 4: Characterisation of a genetically humanised mouse carrying mutant human SOD1$^{A4V}$

## 4.1 Introduction

The alanine to valine mutation at codon 4 (A4V) of the *SOD1* gene is the most common cause of SOD1-ALS in North America, with approximately 50% of U.S. SOD1-ALS patients carrying it [452,453]. The A4V mutation is one of the most severe *SOD1* mutations, with a disease duration of only 1.0-1.4 years [452–454]. The average age of disease onset in SOD1-A4V patients is 47 years old, which is comparable to other *SOD1* mutations, but still younger than the majority of non-SOD1-ALS cases [454]. Lower MN dysfunction typically dominates the clinical presentation of SOD1-A4V patients [453].

The A4V mutation sits within exon 1 of the *SOD1* gene, at a residue that forms part of the β1 strand and is involved in the dimer interface [42]. Accordingly, the A4V mutation destabilises the SOD1 dimer and promotes dimer dissociation [455], a process which is critical for aggregation [56,57,456]. A4V has high aggregation propensity, similar to G93A and G37R [51,55,56], and SOD1 aggregates are readily observed in the ventral spinal cord of SOD1-A4V patients [452].

The A4V mutation has a moderate impact on SOD1 dismutase activity, with intrinsic activity of the enzyme estimated to be ~45% that of wild-type [457]. However, in tissues from SOD1-A4V patients, who carry one WT and one A4V allele, the SOD1 activity is approximately 60% that of healthy controls [452,458], whereas tissue activity closer to 75% would be expected based on the intrinsic activity. This discrepancy can be explained by reduced levels of SOD1-A4V compared to WT SOD1; the A4V mutation is reported to negatively impact stability of the SOD1 holoenzyme and increase protein turnover [452,459–461].

As described in section 1.5.3.2, two SOD1-A4V mouse models have been created to date; one is a transgenic line harbouring 3-4 copies of the human SOD1-A4V transgene [338], and the other is an unpublished knock-in strain with the A4V mutation introduced into mouse *Sod1* (JAX strain 031341).

A lack of overt phenotypes have been reported in these models; the reason for this is currently unclear, especially in light of the aggressive disease in SOD1-A4V patients. The lack of phenotypes may result from a combination of low transgene copy numbers compared to other SOD1 transgenic strains, differences in mouse vs human SOD1 biochemistry (for knock-in mice), and a lack of SOD1 LOF (for transgenic mice).

New *in vivo* insights are needed to uncover how SOD1 causes ALS. Hence, the creation of novel mouse models of SOD1-ALS is warranted – particularly models of prevalent and severe mutations such as A4V. For this reason we applied our philosophy of creating humanised mouse models (described in section 1.5.2.2.3) to create a humanised *hSOD1*$^{A4V}$ mouse line.

The *hSOD1*$^{A4V}$ line described in this chapter was generated by Dr Lydia Teboul and colleagues by introduction of the A4V mutation into *hSOD1*$^{WT}$ embryos using CRISPR/Cas9. A diagram of the *hSOD1*$^{A4V}$ allele is shown in Figure 4-1. In this chapter I describe the behavioural, molecular and metabolic phenotyping of *hSOD1*$^{A4V}$ mice (an outline of the experiments performed is shown in Figure 4-2).

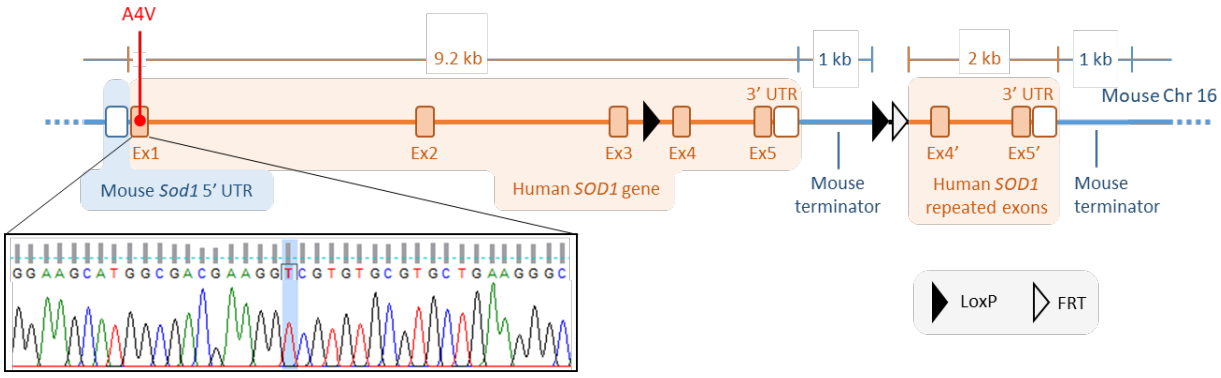



**Figure 4-1 | Schematic of the *hSOD1*$^{A4V}$ allele.** Humanised sequence is shown in orange, mouse sequence is shown in blue. Inset: Sequencing trace from a homozygous hSOD1A4V animal showing successful knock-in of the A4V point mutation (C->T; highlighted). Ex, exon; UTR, untranslated region.

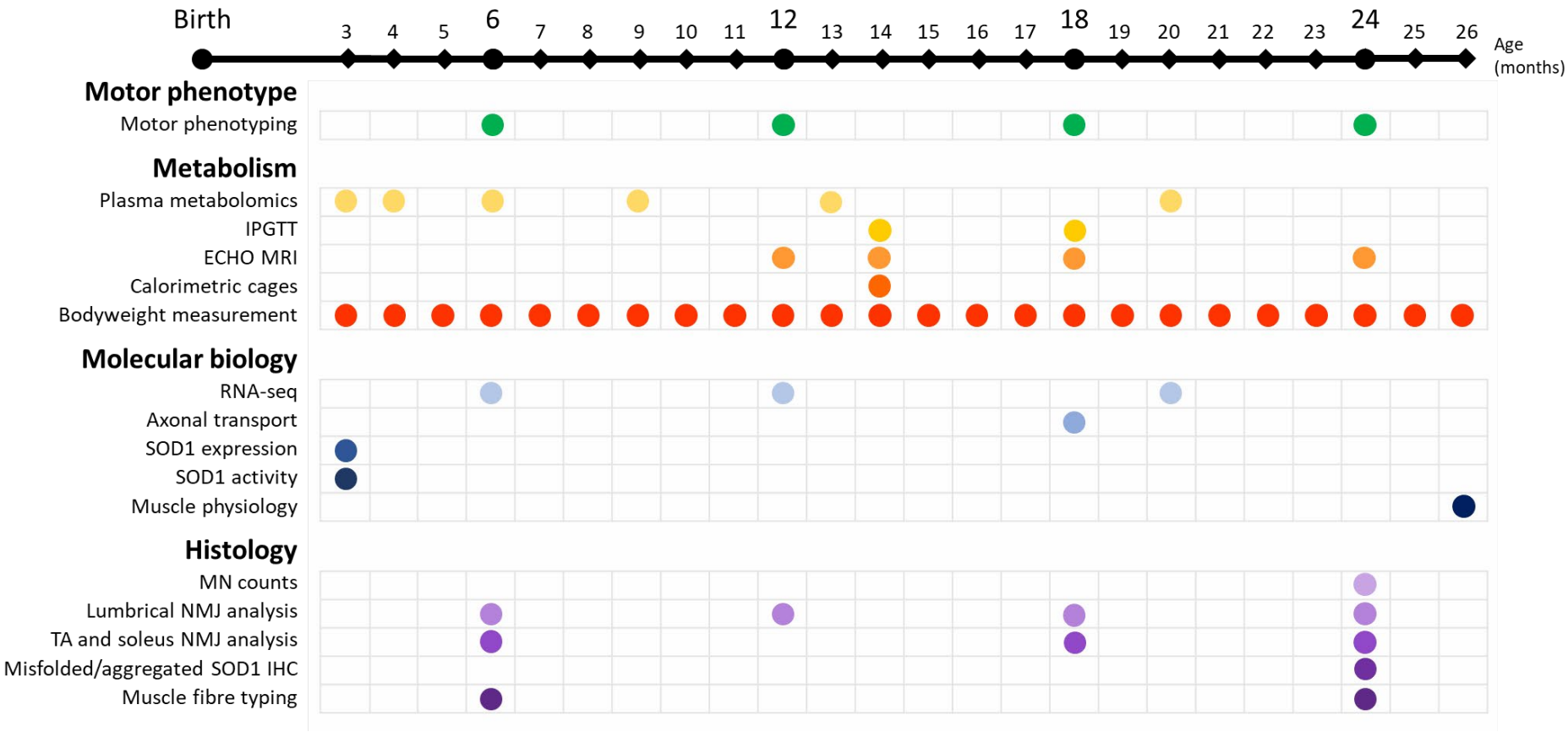



**Figure 4-2 | Outline of experiments performed on *hSOD1$^{A4V}$* mice.** MN, motor neuron; NMJ, neuromuscular junction; IHC, immunohistochemistry; IPGTT, intraperitoneal glucose tolerance testing.

## 4.2 Aims

1. Examine the expression and activity of SOD1-A4V in *hSOD1$^{A4V}$* mice
2. Perform behavioural phenotyping of *hSOD1$^{A4V}$* mice
3. Perform molecular and metabolic phenotyping of *hSOD1$^{A4V}$* mice
4. Perform additional behavioural phenotyping of *hSOD1$^{A4V}$* mice on a mixed C57BL/6J x 129S1/SvImJ background

## 4.3 Results

### 4.3.1 hSOD1-A4V Mendelian ratios

To determine whether the *SOD1$^{A4V}$* mutation affects the viability of *hSOD1$^{A4V}$* mice, the number of offspring of the 3 potential genotypes was monitored in litters born from intercrosses between heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice.

Although the ratio of animals of each genotype born from heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* intercrosses did not differ significantly from the ratio predicted by Mendelian inheritance, there was a trending decrease in the number of homozygous *hSOD1$^{A4V}$/hSOD1$^{A4V}$* animals born ($\chi^2(2)= 4.97$, p=0.08) (Table 4-1). This suggests that the *SOD1$^{A4V}$* mutation may have a mild impact on viability in homozygous animals.

| | **Male** | **Female** | **Sexes combined** | **Expected ratio** | **Observed ratio (sexes combined)** |
|---|---|---|---|---|---|
| *hSOD1$^{WT}$/hSOD1$^{WT}$* | 76 | 75 | 151 | 0.25 | 0.29 |
| *hSOD1$^{WT}$/hSOD1$^{A4V}$* | 131 | 124 | 255 | 0.50 | 0.49 |
| *hSOD1$^{A4V}$/hSOD1$^{A4V}$* | 58 | 58 | 116 | 0.25 | 0.22 |
| Total | 265 | 257 | 522 | 1 | 1 |

**Table 4-1 | Genotypes of offspring born from intercrosses between heterozygous *hSOD1$^{WT}$/ hSOD1$^{A4V}$* mice.**

### 4.3.2 *SOD1-A4V* mRNA expression

To determine whether the A4V mutation affects *SOD1* expression, we examined *SOD1* mRNA expression in a range of tissues from 3-month-old $hSOD1^{A4V}$ animals and $hSOD1^{WT}$ littermate controls using qPCR (Figure 4-3). The expression of *SOD1* mRNA did not significantly differ between $hSOD1^{A4V}$ animals and $hSOD1^{WT}$ littermate controls in the majority of tissues examined. There was a marginal (~25%) increase in *SOD1* mRNA expression in the TA muscle of male $hSOD1^{A4V}/hSOD1^{A4V}$ mice and the liver of male $hSOD1^{WT}/hSOD1^{A4V}$ mice, compared to $hSOD1^{WT}$ littermate controls.

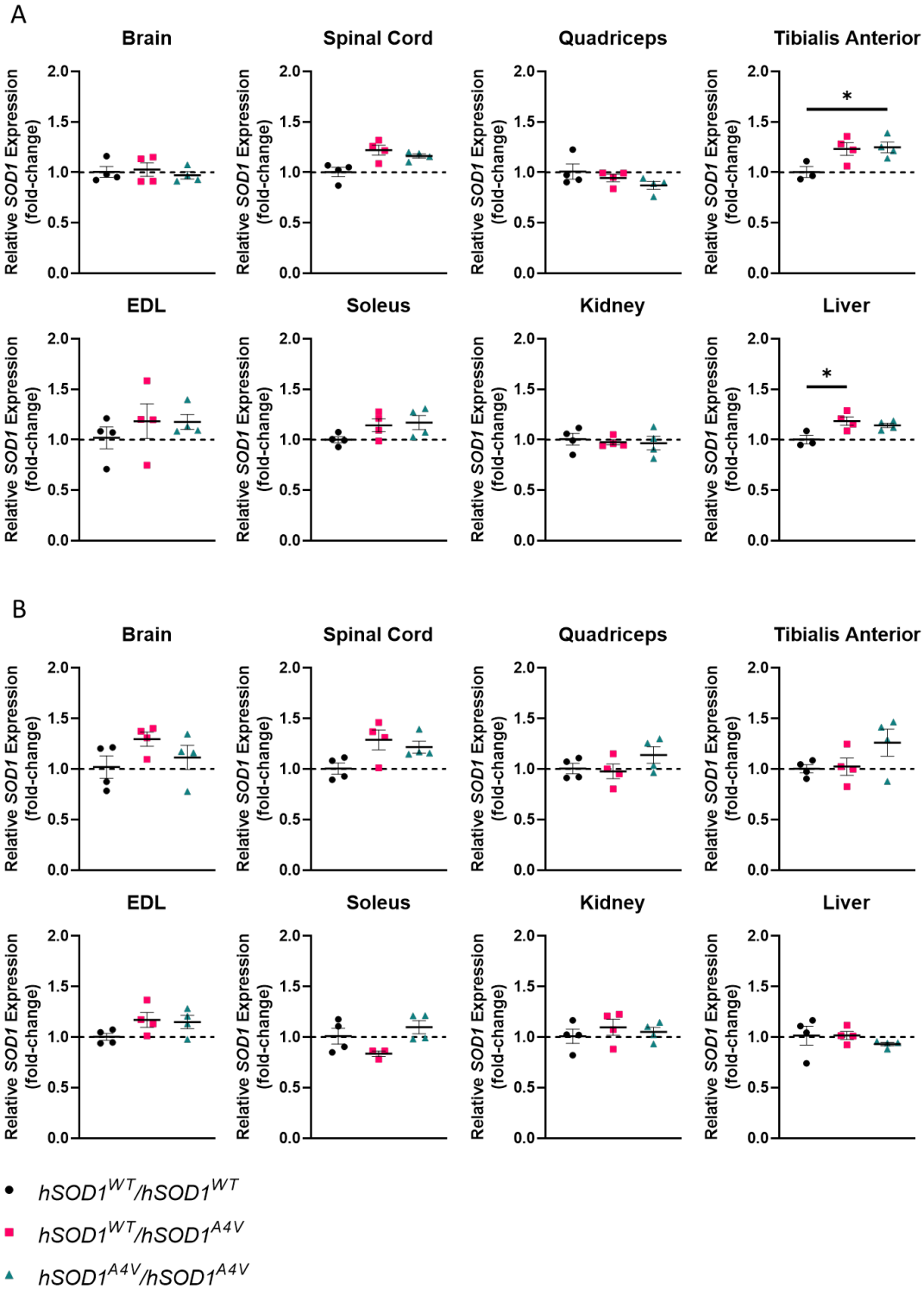



**Figure 4-3 | Expression of *SOD1* RNA in *hSOD1$^{A4V}$* mice by qRT-PCR.** Expression of total *SOD1* mRNA using conserved mouse-human *SOD1* primers in tissues from **(A)** 3-month-old male animals and **(B)** 3-month-old female animals. n=3-4 animals per genotype per tissue. *SOD1* expression was normalized to *GAPDH* expression as a housekeeping gene. All values are expressed as a fold change relative to the average *SOD1* expression in *hSOD1$^{WT}$/hSOD1$^{WT}$* control animals. Mean ± SEM. One-way ANOVA with Tukey's post hoc test. *$P < 0.05$.

### 4.3.3 SOD1-A4V protein levels

SOD1 protein levels were assessed by Western blot in a range of tissues from 3-month-old *hSOD1*$^{A4V}$ animals and *hSOD1*$^{WT}$ littermate controls (Figure 4-4). The level of SOD1 protein was significantly decreased in all but one tissue examined from homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice, and in most tissues from heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ mice, compared to *hSOD1*$^{WT}$ littermate controls. In tissues where SOD1 protein was significantly decreased, the level of SOD1 protein in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals ranged from half to approximately one tenth of the level found in *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ littermate controls.

The observed decrease in SOD1 protein without a decrease in *SOD1* mRNA strongly suggests that the A4V mutation impacts SOD1 protein levels via post-transcriptional mechanisms, likely by reducing the stability and/or increasing the turnover of mutant SOD1, although impaired translation of *SOD1-A4V* mRNA of cannot be ruled out.

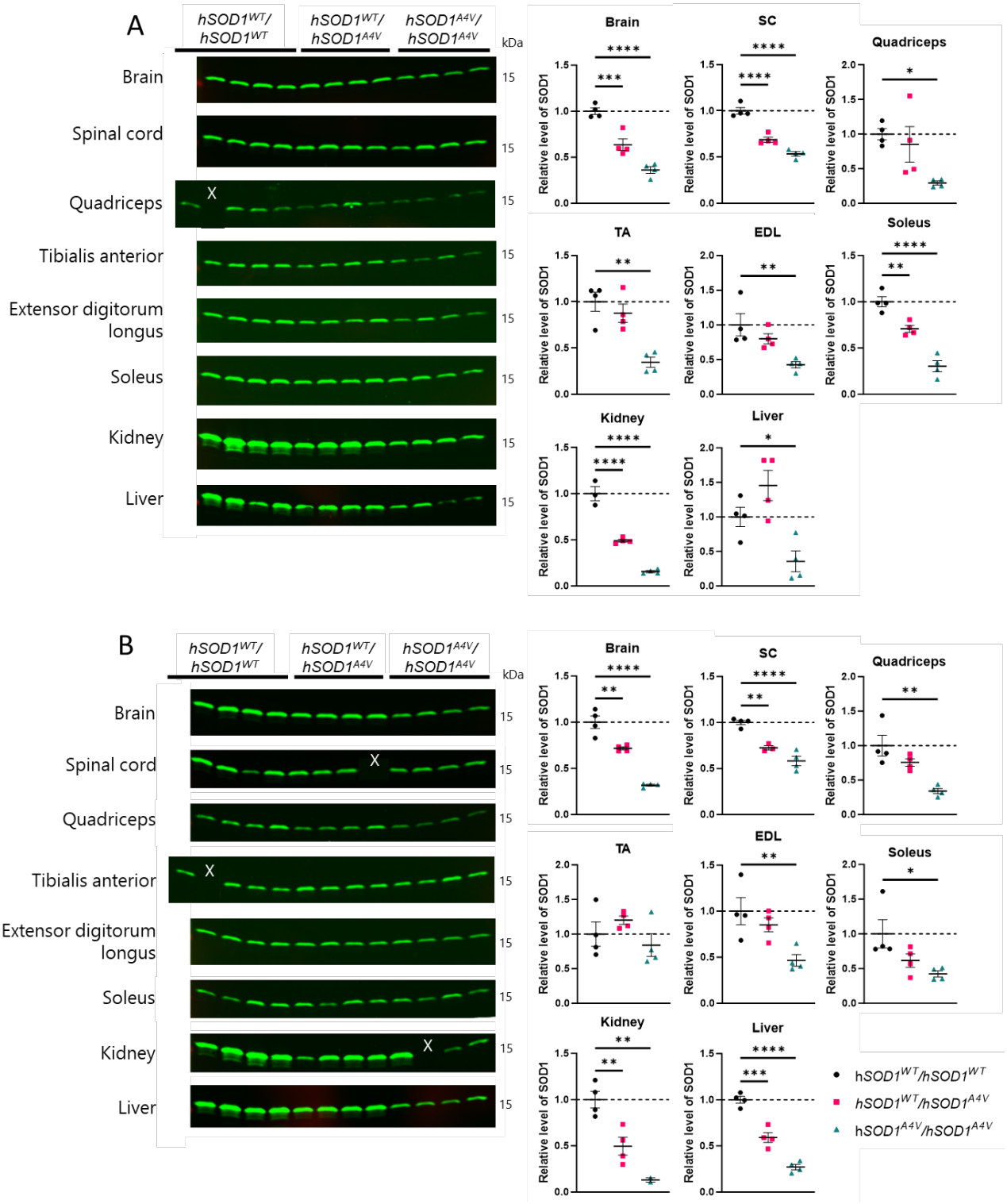



**Figure 4-4 | Levels of SOD1 protein in *hSOD1$^{A4V}$* mice.** Western blots using a pan-mouse-human SOD1 antibody (aa131-153) in tissues from **(A)** 3-month-old male animals and **(B)** 3-month-old female animals. Blots normalized to GAPDH (Figure a-6). Unused lanes marked by x. All values are expressed as a fold change relative to the average SOD1 level in *hSOD1$^{WT}$/hSOD1$^{WT}$* control animals. n=3-4 per genotype per tissue. Mean ± SEM, One-way ANOVA with Tukey's multiple comparisons test. For clarity, only comparisons to *hSOD1$^{WT}$/hSOD1$^{WT}$* animals are shown. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001.

### 4.3.4 SOD1-A4V dismutase activity

In line with the observed decrease in SOD1 protein, we observed decreased SOD1 dismutase activity in a range of tissues from 3-month-old *hSOD1*$^{A4V}$ animals compared to *hSOD1*$^{WT}$ littermate controls, measured using in-gel dismutase activity assays (Figure 4-5, A).

In-gel dismutase activity assays are qualitative and cannot be accurately assessed by densitometry. Hence, we also quantified SOD dismutase activity in a range of tissues using a spectrophotometric assay (Figure 4-5 B, C). SOD1 activity in heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ tissues was 40-80% of the activity observed in tissues from *hSOD1*$^{WT}$ littermate controls. SOD1 activity in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ tissues was 10-40% of the activity observed in tissues from *hSOD1*$^{WT}$ littermate controls.

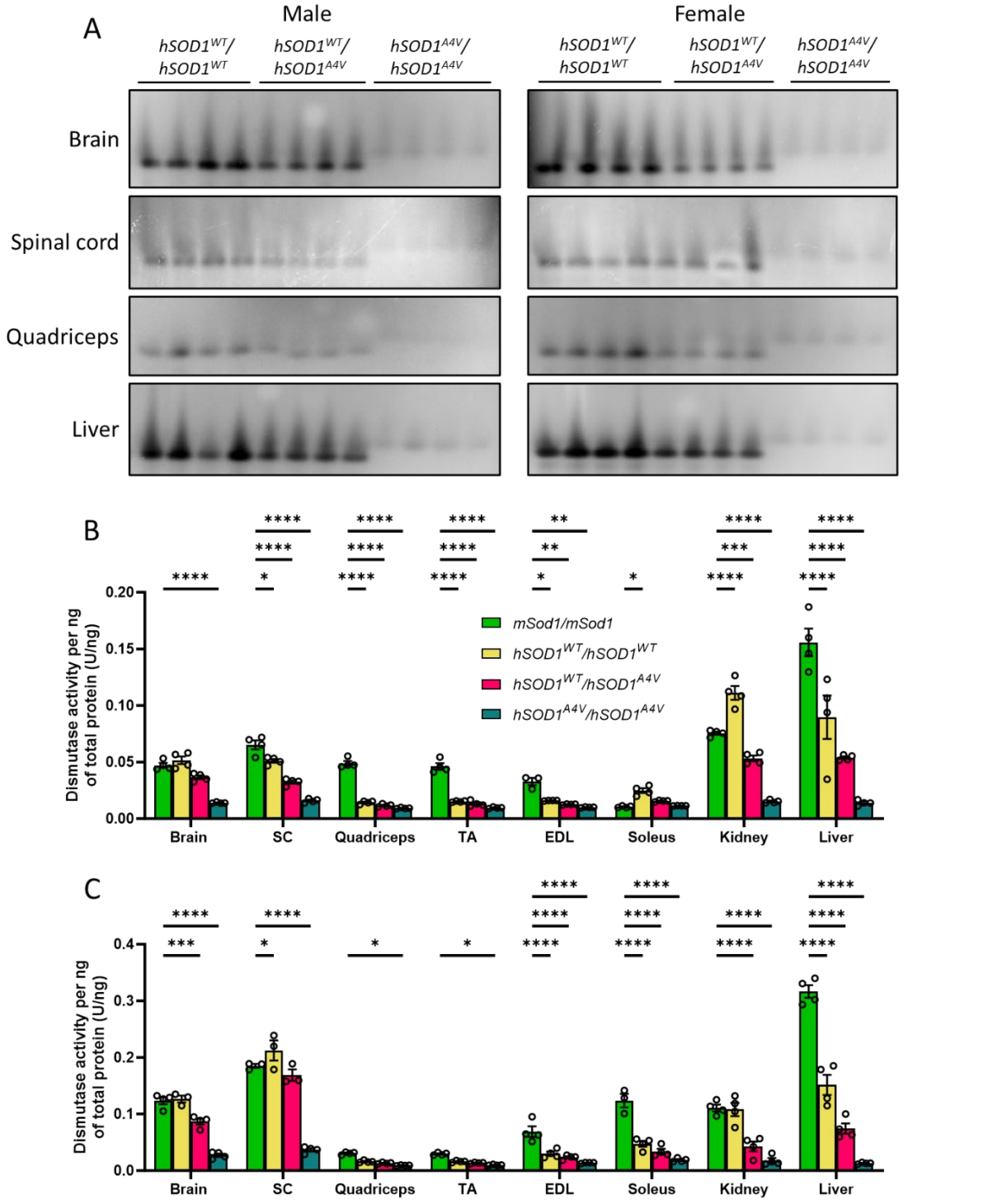



**Figure 4-5 | SOD1 dismutase activity in *hSOD1$^{A4V}$* mice. (A)** Colorimetric in-gel SOD1 dismutase activity assays in tissues from 3-month-old *hSOD1$^{A4V}$* mice. Darker staining indicates higher dismutase activity. SOD1-A4V has lower mobility than wild-type SOD1 in clear-native PAGE and appears as a higher band. **(B-C)** Spectrophotometric dismutase activity assays in tissues from **(B)** male and **(C)** female 3-month-old *hSOD1$^{A4V}$* mice and 4-month-old *mSod1/mSod1* controls. Mean ± SEM. One-way ANOVA with Tukey's post hoc test. For clarity, only significant comparisons to *mSod1* controls are shown. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001. ns, not significant; SC, spinal cord; TA, tibialis anterior; EDL, extensor digitorum longus.

### 4.3.5 hSOD1-A4V molecular characterisation

#### *4.3.5.1 Axonal transport*

Deficits in retrograde axonal transport are an early phenotype detectable in presymptomatic *tgSOD1$^{G93A}$* mice [176]. We examined retrograde axonal transport in the TA and soleus muscles of 18-month-old male *hSOD1$^{A4V}$* mice (Figure 4-6).

We did not observe any genotype-specific differences in the speed or pausing of endosome-mediated trafficking in the motor neurons innervating either muscle.

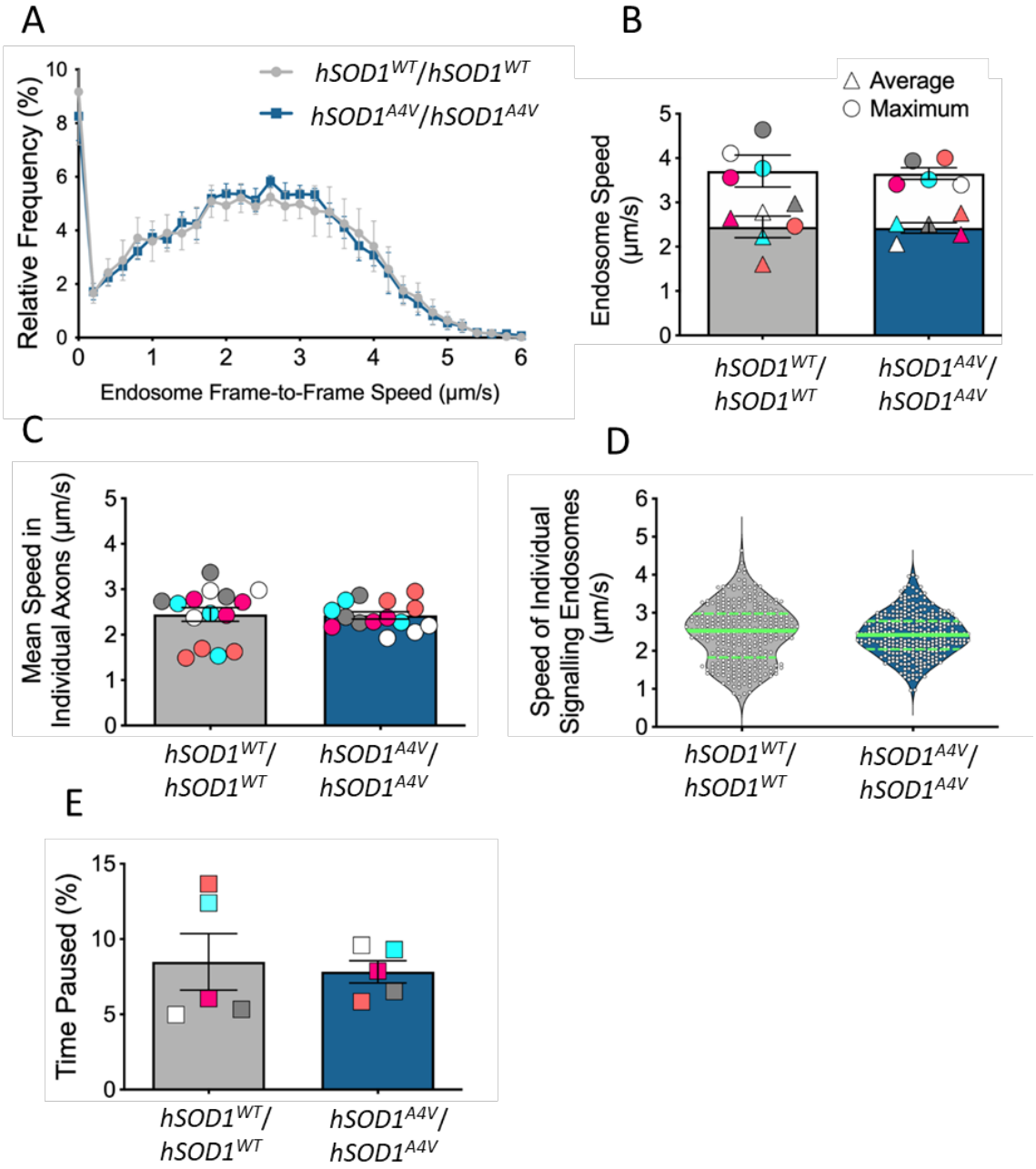



**Figure 4-6 | Retrograde axonal transport of signaling endosomes in TA motor neurons of 18-month-old male *hSOD1$^{A4V}$* mice. (A)** distribution curves of endosome frame-to-frame speed. **(B)** Average and maximum endosome speeds in individual animals. **(C)** Average endosome speeds in individual TA MN axons. **(D)** Average speed of individual endosomes. Solid green line represents the median and dashed green lines present the upper and lower quartiles. **(E)** Relative percentage of time endosomes spent paused. n=5 animals per genotype. Colours of points indicate individual animals. Mean ± SEM. Student's t test.

#### 4.3.5.2 RNA-seq

Transcriptional changes are widely reported in *in vitro* and *in vivo* models of ALS, and in cells and tissue samples from ALS patients. Disease-associated mutations in several RNA-binding proteins – most notably TDP-43 and FUS – have been implicated in ALS pathology, highlighting disrupted RNA homeostasis as a key pathomechanism in ALS [9]. Widespread transcriptional changes are also observed in models of SOD1-ALS, which lack a genetic insult to RNA processing [462,463].

We examined transcriptional changes in the lumbar SC and TA muscle of male *hSOD1$^{A4V}$* mice at 6mo, 12mo and 20mo using QuantSeq 3' mRNA sequencing (Lexogen). Initial RNA-seq analysis was performed by Lexogen using alignment to the mm10 mouse genome (2012). However, a more complete version of the mouse genome – mm39 – was made available in 2020. Therefore the RNA-seq data were re-analysed by Chloe Williams at Umeå University using alignment to the mm39 genome.

To assess differential gene expression at the level of individual genes we used a stringent analytical pipeline (adjusted p value (padj) < 0.1). To maximise the number of input genes for Gene Ontology (GO) analysis we included DEGs that met a significance threshold of unadjusted p value < 0.05. We also performed Gene Set Enrichment Analysis (GSEA) [407] to examine transcriptional profiles altered in *hSOD1$^{A4V}$* mice. As an exploratory analysis, we performed GSEA against the mouse hallmark gene sets in the MSigDB database (Broad Institute) – 50 gene sets that represent well-defined biological states. We chose a stringent significance threshold of FDR q-value < 0.05 for GSEA.

##### 4.3.5.2.1 SC results

We observed a number of differentially expressed genes (DEGs) in the SC of *hSOD1$^{A4V}$* animals (Table 4-2). We detected 47 DEGs in the SC of 6-month-old animals. The 2 most significantly altered genes were *Apod* (Apolipoprotein D), a lipid carrier, and *Plp1* (Proteolipid protein 1), a component of the myelin sheath – both were decreased in *hSOD1$^{A4V}$* animals. Both *Apod* and *Plp1* have been described in association with neurological disorders [464,465]. There were also several mitochondrially encoded

DEGs, however these had very small magnitude changes in expression. We found several DEGs encoding haemoglobin subunits, suggesting some minor contamination with blood in our samples.

We detected 6 DEGs in the SC of 12-month-old animals. However, 4 of these DEGs encoded haemoglobin subunits, again suggesting contamination with blood. The remaining 2 DEGs were *Pisd-ps3* (phosphatidylserine decarboxylase, pseudogene 3) and *Chchd2* (coiled-coil-helix-coiled-coil-helix domain containing 2).

Two DEGs were detected in 20-month-old animals – *Fasn* (fatty acid synthase) and *Ces1d* (carboxylesterase 1D) (Table 4-2). Both of these genes had decreased expression in *hSOD1$^{A4V}$* animals compared to controls. *Fasn* and *Ces1d* are both involved in lipid metabolism, being involved in lipogenesis and lipolysis respectively [466,467]. *Fasn* is downregulated in the liver of the "Taiwanese" SMA mouse model [468], but unchanged in the frontal cortex of ALS patients [469]. *Fasn*, but not *Ces1d* was identified as a DEG in symptomatic *tgSOD1$^{G93A}$* mice in a transcriptomic meta-analysis [463].

| Gene Symbol | Description | Base Mean | Log2 Fold Change | SE of Log2 Fold Change | p value | Adjusted p value | Significant comparison |
|---|---|---|---|---|---|---|---|
| **6-month-old *hSOD1$^{A4V}$* spinal cord (47 unique DEGs)** | | | | | | | |
| *Apod* | apolipoprotein D | 1908 | -0.58 | 0.09 | 5.78E-11 | 5.34E-07 | WT vs Hom |
| *Plp1* | proteolipid protein (myelin) 1 | 23050 | -0.49 | 0.08 | 6.97E-11 | 5.34E-07 | WT vs Hom |
| *Apod* | apolipoprotein D | 2022 | -0.49 | 0.08 | 2.23E-09 | 5.51E-06 | WT vs Het |
| *Cacng2* | calcium channel, voltage-dependent, gamma subunit 2 | 755 | -0.67 | 0.12 | 5.97E-08 | 3.25E-04 | WT vs Hom |
| *Rpl17* | ribosomal protein L17 | 643 | -0.58 | 0.11 | 7.07E-08 | 3.25E-04 | WT vs Hom |
| *Phip* | pleckstrin homology domain interacting protein | 176 | -0.90 | 0.17 | 2.07E-07 | 6.79E-04 | WT vs Hom |
| *Plp1* | proteolipid protein (myelin) 1 | 21692 | -0.32 | 0.06 | 2.29E-07 | 1.06E-03 | Het vs Hom |
| *mt-Rnr2* | mitochondrially encoded 16S rRNA | 155301 | 0.25 | 0.05 | 3.74E-07 | 4.61E-04 | WT vs Het |
| *Nfkbia* | NFKB inhibitor alpha | 386 | 0.62 | 0.13 | 8.16E-07 | 3.16E-03 | Het vs Hom |
| *Pcsk1n* | proprotein convertase subtilisin/kexin type 1 inhibitor | 1547 | -0.37 | 0.08 | 1.16E-06 | 9.54E-04 | WT vs Het |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Phip* | pleckstrin homology domain interacting protein | 167 | -0.76 | 0.16 | 4.14E-06 | 1.37E-02 | Het vs Hom |
| *Syp* | synaptophysin | 753 | -0.48 | 0.10 | 4.35E-06 | 9.36E-03 | WT vs Hom |
| *Cfl1* | cofilin 1, non-muscle | 886 | -0.47 | 0.10 | 4.48E-06 | 9.36E-03 | WT vs Hom |
| *Olig1* | oligodendrocyte transcription factor 1 | 898 | 0.41 | 0.09 | 5.52E-06 | 1.06E-02 | WT vs Hom |
| *Syne1* | spectrin repeat containing, nuclear envelope 1 | 861 | 0.40 | 0.09 | 8.14E-06 | 1.44E-02 | WT vs Hom |
| *Map1b* | microtubule-associated protein 1B | 6888 | -0.32 | 0.07 | 1.52E-05 | 2.49E-02 | WT vs Hom |
| *Vamp1* | vesicle-associated membrane protein 1 | 3015 | -0.25 | 0.06 | 1.75E-05 | 1.08E-02 | WT vs Het |
| *ND4* | mitochondrially encoded NADH dehydrogenase 4 | 56406 | -0.36 | 0.08 | 1.88E-05 | 2.88E-02 | WT vs Hom |
| *Arf5* | ADP-ribosylation factor 5 | 63 | -1.76 | 0.41 | 2.11E-05 | 6.11E-02 | Het vs Hom |
| *mt-Rnr1* | mitochondrially encoded 12S rRNA | 63020 | 0.24 | 0.06 | 2.65E-05 | 1.31E-02 | WT vs Het |
| **12-month-old *hSOD1$^{A4V}$* spinal cord (6 unique DEGs)** | | | | | | | |
| *Pisd-ps3* | phosphatidylserine decarboxylase, pseudogene 3 | 317 | -0.68 | 0.11 | 1.67E-10 | 1.33E-06 | WT vs Hom |
| *Chchd2* | coiled-coil-helix-coiled-coil-helix domain containing 2 | 1117 | -0.40 | 0.08 | 2.41E-06 | 1.43E-02 | Het vs Hom |
| **21-month-old *hSOD1$^{A4V}$* spinal cord (2 unique DEGs)** | | | | | | | |
| *Fasn* | fatty acid synthase | 696 | -1.53 | 0.29 | 1.52E-07 | 2.92E-03 | WT vs Het |
| *Ces1d* | carboxylesterase 1D | 119 | -2.36 | 0.46 | 2.78E-07 | 2.92E-03 | WT vs Het |

**Table 4-2 | Differentially expressed genes (DEGs) in the spinal cord of *hSOD1$^{A4V}$* mice.** Significance threshold was adjusted p value (padj) < 0.1. 47 DEGs were detected in the SC of 6-month-old animals; only the 20 most significant DEGs are shown. DEGs are listed in order of significance. Haemoglobin genes, likely resulting from contamination with blood, are not included in this table. A full list of DEGs can be found in appendix 8.1. SE, standard error; WT, *hSOD1$^{WT}$*/*hSOD1$^{WT}$*; Het, *hSOD1$^{WT}$*/*hSOD1$^{A4V}$*; Hom, *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$*.

We observed enrichment of a number of GO biological pathways in the SC of *hSOD1$^{A4V}$* animals (Table 4-3). Among these were several GO terms for metabolic processes, which were enriched in 20-month-old animals.

| GO biological pathway | GO term ID | Adjusted p value | GO term size | Percentage of DEGs overlapping with GO term |
|---|---|---|---|---|
| **6-month-old *hSOD1$^{A4V}$* spinal cord** | | | | |
| regulation of response to stress | GO:0080134 | 1.4E-03 | 1467 | 10.8 |
| regulation of DNA-binding transcription factor activity | GO:0051090 | 9.1E-03 | 426 | 4.5 |
| regulation of multicellular organismal process | GO:0051239 | 1.5E-02 | 3246 | 18.5 |
| regulation of I-kappaB kinase/NF-kappaB signaling | GO:0043122 | 4.2E-02 | 211 | 2.8 |
| **12-month-old *hSOD1$^{A4V}$* spinal cord** | | | | |
| DNA damage response | GO:0006974 | 4.6E-02 | 842 | 9.9 |
| **20-month-old *hSOD1$^{A4V}$* spinal cord** | | | | |
| cellular biosynthetic process | GO:0044249 | 8.1E-09 | 5474 | 30.2 |
| organonitrogen compound metabolic process | GO:1901564 | 9.9E-05 | 6489 | 32.1 |
| fat cell differentiation | GO:0045444 | 4.8E-04 | 259 | 3.1 |
| animal organ development | GO:0048513 | 8.6E-04 | 3326 | 18.2 |
| phosphorus metabolic process | GO:0006793 | 1.4E-03 | 2813 | 15.9 |
| tissue development | GO:0009888 | 2.5E-03 | 2099 | 12.5 |
| cellular response to lipoprotein particle stimulus | GO:0071402 | 2.5E-03 | 38 | 1.1 |
| cell surface receptor signaling pathway | GO:0007166 | 9.7E-03 | 2895 | 15.7 |
| developmental process | GO:0032502 | 1.7E-02 | 6865 | 32.0 |
| system development | GO:0048731 | 2.1E-02 | 4115 | 20.7 |
| cellular response to chemical stimulus | GO:0070887 | 2.1E-02 | 3173 | 16.7 |
| intracellular signal transduction | GO:0035556 | 3.9E-02 | 2607 | 14.1 |
| oxoacid metabolic process | GO:0043436 | 4.2E-02 | 1010 | 6.7 |

**Table 4-3 | GO biological pathways associated with DEGs in the spinal cord of *hSOD1$^{A4V}$* mice.** GO pathway analysis was performed on all DEGs with a significance threshold of $p < 0.05$. All GO pathways with a significance of $padj < 0.05$ are shown.

Using GSEA, we observed enrichment of the spermatogenesis hallmark gene set in the SC of 12-month-old animals (Figure 4-7 D), and enrichment of the myogenesis hallmark gene set in 20-month-old animals (Figure 4-7 E). Further investigation into these unexpected changes revealed that the spermatogenesis gene set was being driven by genes with very low, or no, expression in most SC samples. Furthermore, many of the genes annotated as spermatogenic have functions in other tissues. The enrichment of the myogenesis gene set was driven by increased expression of myogenic genes in one WT sample, indicating contamination with muscle cells in that sample (datum not shown).

We found very little overlap between DEGs ($p < 0.05$) at different time points (Figure 4-7 F), suggesting there is not ongoing perturbation of any specific genes.

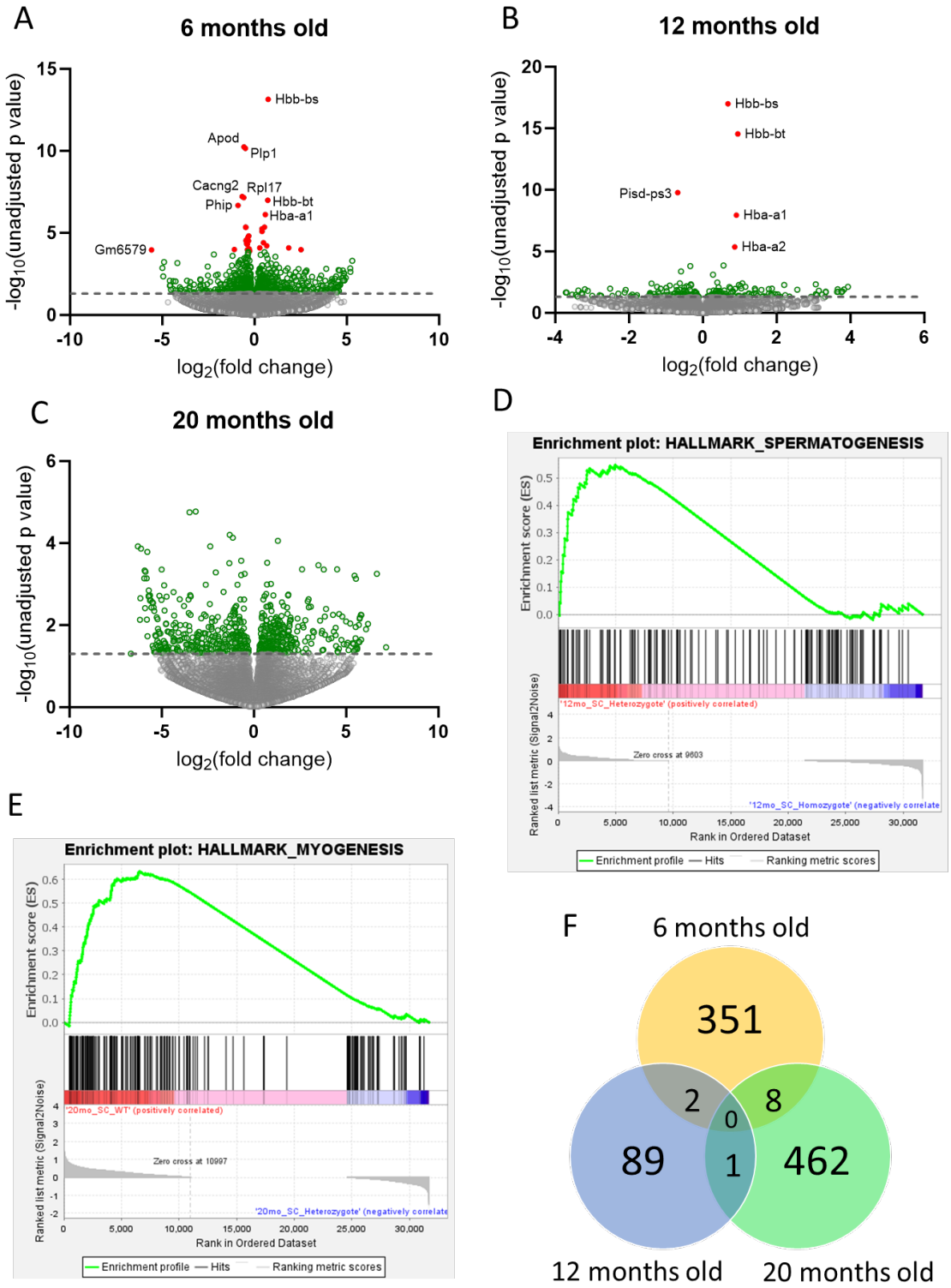



**Figure 4-7 | Longitudinal RNA sequencing data from the spinal cord of *hSOD1$^{A4V}$* mice. (A-C)** Volcano plots showing gene expression changes between *hSOD1$^{WT}$*/*hSOD1$^{WT}$* and *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* mice at different ages. Green points indicate p < 0.05, red points indicate padj < 0.1. **(D, E)** GSEA plots showing enrichment of (D) the spermatogenesis hallmark gene set in 12-month-old *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* vs *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* mice and **(E)** the myogenesis hallmark gene set in 20-month-old *hSOD1$^{WT}$*/*hSOD1$^{WT}$* vs *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* mice. **(F)** Venn diagram showing the overlap of DEGs (at p < 0.05) between *hSOD1$^{WT}$*/*hSOD1$^{WT}$* vs *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* comparisons at different time points.

#### 4.3.5.2.2 Muscle results

We observed a number of differentially expressed genes (DEGs) in the TA muscle of *hSOD1*$^{A4V}$ animals (Table 4-4). We detected 153 DEGs in muscle from 6-month-old animals. GO analysis of these DEGs revealed a significant enrichment of genes encoding ribosomal (padj = $4.59 \times 10^{-13}$) and mRNA binding proteins (padj = $3.78 \times 10^{-7}$) (Table 4-5).

In 12-month-old animals, 5 DEGs were detected. 3 of these DEGs encoded haemoglobin subunits. The remaining 2 DEGs were *Eda2r* (ectodysplasin A2 receptor) and *Klf2* (Kruppel-like factor 2). *Eda2r* RNA is upregulated with age in muscle [470], and EDA2R protein has been found to increase in plasma following ALS diagnosis [471]. *Klf2* has not been linked to ALS.

Only one DEG was detected in 20-month-old animals – *Myl3* (myosin, light polypeptide 3), which was decreased in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice. *MYL3* has been found to be decreased in ALS patient skeletal muscle [472]. Myl3 protein increased in skeletal muscle from 14-week-old *tgSOD1*$^{G93A}$ mice [473].

| Gene Name | Description | Base Mean | Log2 Fold Change | SE of Log2 Fold Change | P value | Adjusted p value | Significant comparison |
|---|---|---|---|---|---|---|---|
| **6-month-old hSOD1A4V muscle (153 unique DEGs)** | | | | | | | |
| *Des* | desmin | 5936 | 0.58 | 0.09 | 3.91E-10 | 6.77E-07 | WT vs Het |
| *Hspb8* | heat shock protein 8 | 882 | 0.78 | 0.13 | 4.99E-10 | 6.77E-07 | WT vs Het |
| *Ddit4l* | DNA-damage-inducible transcript 4-like | 868 | -0.86 | 0.16 | 5.24E-08 | 7.79E-05 | Het vs Hom |
| *Asb2* | ankyrin repeat and SOCS box-containing 2 | 1209 | -0.84 | 0.16 | 7.62E-08 | 7.79E-05 | Het vs Hom |
| *Ampd1* | adenosine monophosphate deaminase 1 | 1423 | -0.65 | 0.13 | 1.86E-07 | 1.27E-04 | Het vs Hom |
| *Eif4ebp2* | eukaryotic translation initiation factor 4E binding protein 2 | 239 | -1.01 | 0.20 | 2.53E-07 | 2.29E-04 | WT vs Het |
| *Trdn* | triadin | 10904 | 0.78 | 0.16 | 5.05E-07 | 2.95E-04 | WT vs Het |
| *mt-Rnr2* | mitochondrially encoded 16S rRNA | 252346 | -0.74 | 0.15 | 5.44E-07 | 2.95E-04 | WT vs Het |
| *Rpl13a* | ribosomal protein L13A | 996 | 0.67 | 0.13 | 6.81E-07 | 3.08E-04 | WT vs Het |
| *Nceh1* | neutral cholesterol ester hydrolase 1 | 154 | -1.13 | 0.23 | 1.16E-06 | 4.15E-04 | WT vs Het |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Aqp4* | aquaporin 4 | 82 | -1.27 | 0.26 | 1.23E-06 | 1.43E-02 | WT vs Hom |
| *Rps27a* | ribosomal protein S27A | 261 | 0.79 | 0.16 | 1.39E-06 | 4.15E-04 | WT vs Het |
| *Rps24* | ribosomal protein S24 | 1217 | 0.79 | 0.16 | 1.44E-06 | 4.15E-04 | WT vs Het |
| *ND2* | mitochondrially encoded NADH dehydrogenase 2 | 206386 | -0.55 | 0.12 | 1.53E-06 | 4.15E-04 | WT vs Het |
| *Ptrhd1* | peptidyl-tRNA hydrolase domain containing 1 | 132 | 1.18 | 0.25 | 2.06E-06 | 5.08E-04 | WT vs Het |
| *170004 2G15Rik* | RIKEN cDNA 1700042G15 gene | 181 | 6.28 | 1.34 | 2.76E-06 | 1.43E-02 | WT vs Hom |
| *Rbm25* | RNA binding motif protein 25 | 516 | 0.76 | 0.16 | 3.14E-06 | 7.09E-04 | WT vs Het |
| *Sh3rf2* | SH3 domain containing ring finger 2 | 59 | -1.66 | 0.36 | 3.58E-06 | 1.43E-02 | WT vs Hom |
| *Lmcd1* | LIM and cysteine-rich domains 1 | 918 | 0.86 | 0.19 | 4.79E-06 | 2.45E-03 | Het vs Hom |
| *Ndufa5* | NADH:ubiquinone oxidoreductase subunit A5 | 1953 | 0.64 | 0.14 | 4.92E-06 | 9.93E-04 | WT vs Het |
| **12-month-old hSOD1A4V muscle (5 unique DEGs)** | | | | | | | |
| *Eda2r* | ectodysplasin A2 receptor | 34 | 1.37 | 0.29 | 2.60E-06 | 1.29E-02 | Het vs Hom |
| *Klf2* | Kruppel-like factor 2 (lung) | 83 | 1.06 | 0.25 | 1.65E-05 | 6.55E-02 | Het vs Hom |
| **21-month-old hSOD1A4V muscle (1 unique DEG)** | | | | | | | |
| *Myl3* | myosin, light polypeptide 3 | 414 | -0.78 | 0.15 | 2.96E-07 | 5.59E-03 | Het vs Hom |

**Table 4-4 | Differentially expressed genes (DEGs) in the TA muscle of *hSOD1$^{A4V}$* mice.** Significance threshold was adjusted p value (padj) < 0.1. 153 DEGs were detected in the TA of 6-month-old animals; only the 20 most significant DEGs are shown. DEGs are listed in order of significance. Haemoglobin genes, likely resulting from contamination with blood, are not included in this table. A full list of DEGs can be found in appendix 8.1. SE, standard error; WT, *hSOD1$^{WT}$/hSOD1$^{WT}$*; Het, *hSOD1$^{WT}$/hSOD1$^{A4V}$*; Hom, *hSOD1$^{A4V}$/hSOD1$^{A4V}$*.

We observed enrichment of a number of GO biological pathways in the TA muscle of *hSOD1$^{A4V}$* animals (Table 4-5). However, most of these GO terms were broad terms related to general cellular function, rather than specific pathways. Using GSEA, we did not observe enrichment of any hallmark gene sets in our muscle dataset at any time point.

| GO biological pathway | GO term ID | Adjusted p value | GO term size | Percentage of DEGs overlapping with GO term |
|---|---|---|---|---|
| **6-month-old *hSOD1$^{A4V}$* muscle** | | | | |
| regulation of cellular process | GO:0050794 | 5.2E-06 | 12308 | 56.6 |
| biosynthetic process | GO:0009058 | 2.5E-04 | 5960 | 30.6 |
| multicellular organism development | GO:0007275 | 4.7E-03 | 4873 | 25.1 |
| regulation of nervous system development | GO:0051960 | 5.0E-03 | 551 | 5.0 |
| cell surface receptor signaling pathway | GO:0007166 | 1.3E-02 | 2895 | 16.3 |
| organonitrogen compound metabolic process | GO:1901564 | 1.4E-02 | 6489 | 31.4 |
| regulation of transcription by RNA polymerase II | GO:0006357 | 3.3E-02 | 2513 | 14.3 |
| positive regulation of interleukin-1 production | GO:0032732 | 4.0E-02 | 77 | 1.5 |
| DNA-templated transcription | GO:0006351 | 4.4E-02 | 3576 | 18.9 |
| **12-month-old *hSOD1$^{A4V}$* muscle** | | | | |
| negative regulation of cellular process | GO:0048523 | 6.7E-04 | 4975 | 32.9 |
| developmental process | GO:0032502 | 8.2E-04 | 6865 | 41.1 |
| cellular response to chemical stimulus | GO:0070887 | 1.0E-02 | 3173 | 22.7 |
| regulation of metabolic process | GO:0019222 | 1.9E-02 | 7262 | 40.6 |
| chordate embryonic development | GO:0043009 | 2.4E-02 | 796 | 9.2 |
| regulation of cell population proliferation | GO:0042127 | 2.7E-02 | 1780 | 15.0 |
| cell differentiation | GO:0030154 | 4.5E-02 | 4715 | 29.0 |
| regulation of cell death | GO:0010941 | 5.0E-02 | 1836 | 15.0 |
| **20-month-old *hSOD1$^{A4V}$* muscle** | | | | |
| regulation of multicellular organismal process | GO:0051239 | 3.6E-06 | 3246 | 25.6 |
| response to stress | GO:0006950 | 6.8E-03 | 4156 | 26.4 |
| cardiac atrium morphogenesis | GO:0003209 | 2.6E-02 | 34 | 2.0 |
| cell surface receptor signaling pathway | GO:0007166 | 3.4E-02 | 2895 | 19.6 |
| nucleobase-containing compound biosynthetic process | GO:0034654 | 3.4E-02 | 4007 | 24.8 |
| regulation of RNA metabolic process | GO:0051252 | 4.4E-02 | 3779 | 23.6 |

**Table 4-5 | GO biological pathways associated with DEGs in the TA muscle of *hSOD1$^{A4V}$* mice.** GO pathway analysis was performed on all DEGs with a significance threshold of $p < 0.05$. All GO pathways with a significance of padj < 0.05 are shown.

Similar to SC, we found very little overlap between muscle DEGs ($p < 0.05$) at different time points (Figure 4-8 D), suggesting there is not an ongoing perturbation of any specific genes.

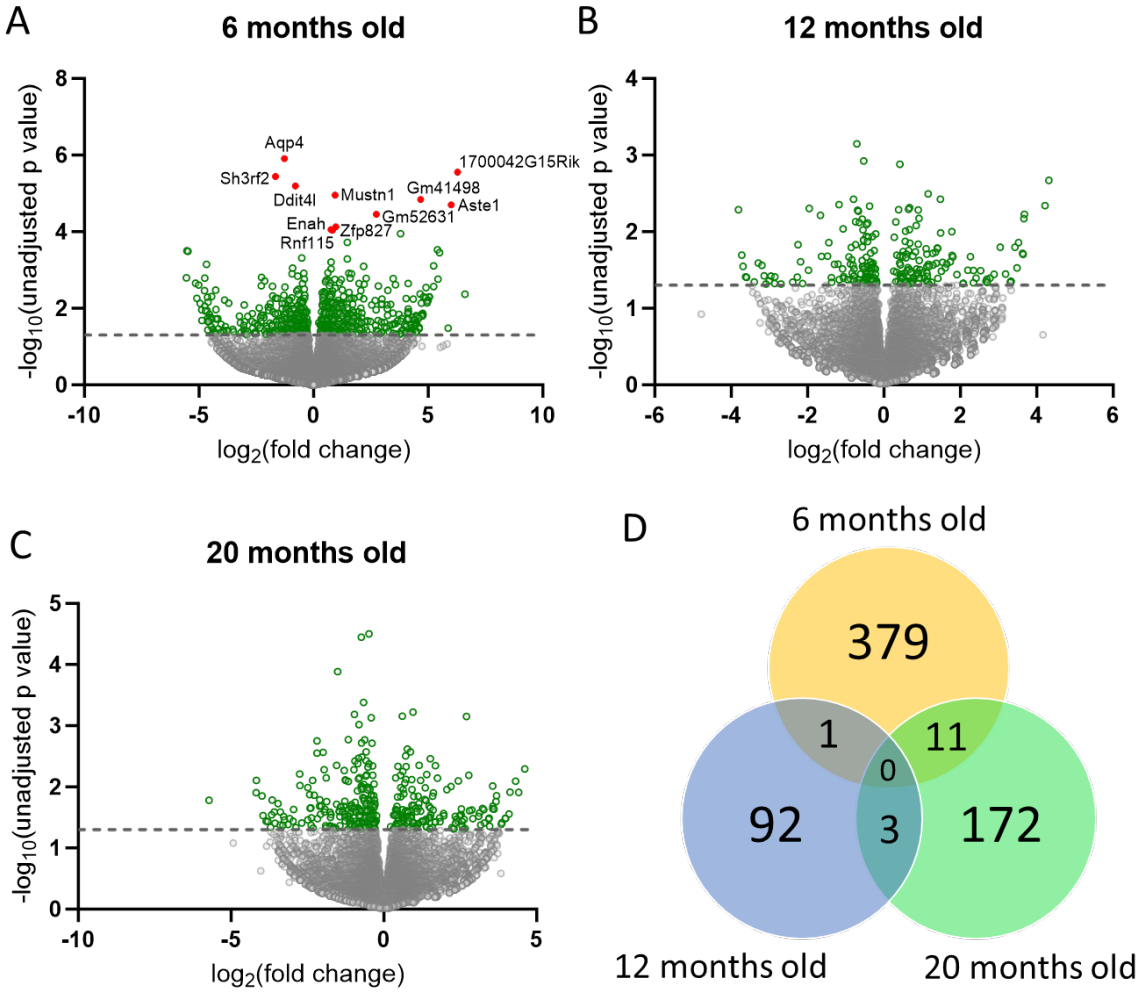



**Figure 4-8 | Longitudinal RNA sequencing data from the TA muscle of *hSOD1$^{A4V}$* mice. (A-C)** Volcano plots showing gene expression changes between *hSOD1$^{WT}$*/*hSOD1$^{WT}$* and *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* mice at different ages. Green points indicate $p < 0.05$, red points indicate padj < 0.1. **(D)** Venn diagram showing the overlap of DEGs (at $p < 0.05$) between *hSOD1$^{WT}$*/*hSOD1$^{WT}$* vs *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* comparisons at different time points.

Overall, these RNA-seq data show very minor disturbances in gene expression in the SC and TA muscle of *hSOD1$^{A4V}$* animals. The observed gene expression changes were inconsistent over time.

### 4.3.6 hSOD1-A4V phenotypic characterisation

To assess the impact of the *SOD1$^{A4V}$* mutation on motor function and behaviour in *hSOD1$^{A4V}$* mice, we performed longitudinal phenotyping using a battery of phenotyping tests. Phenotyping was conducted at 6, 12, 18 and 24 months of age.

### 4.3.6.1 *hSOD1-A4V Kaplan–Meier survival curves*

Survival analysis was conducted on animals who reached a humane end point defined in section 2.2.3, or who died unexpectedly. Animals who died due to non-disease related causes (e.g. fighting) were not included in the analysis.

There was no significant effect of genotype on survival in female *hSOD1$^{A4V}$* mice ($\chi^2(2)=1.583$, p=0.45) (Figure 4-9, A). However, there was a significant effect of genotype on survival in male *hSOD1$^{A4V}$* animals, determined using the log-rank (Mantel-Cox) test ($\chi^2(2)=6.226$, p=0.04) (Figure 4-9, B).

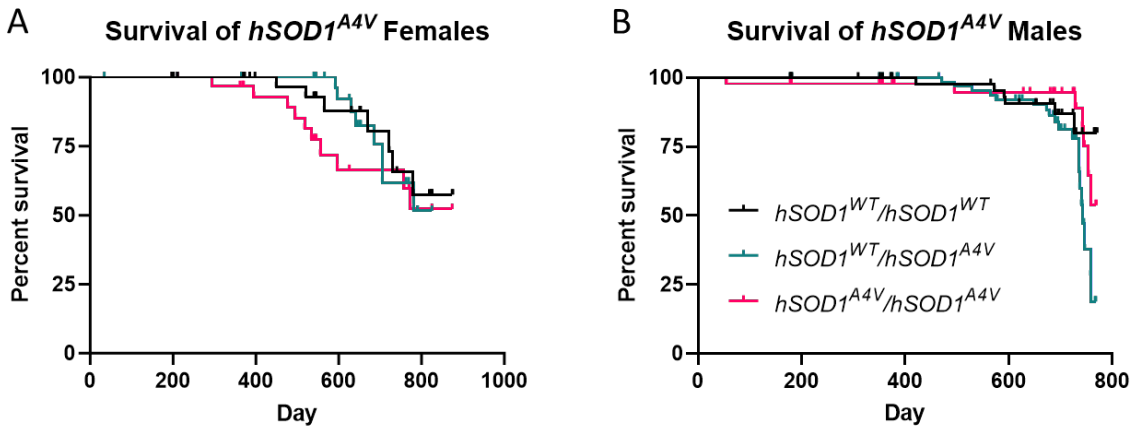



**Figure 4-9 | Kaplan-Meier survival graphs for *hSOD1$^{A4V}$* mice.** Kaplan-Meier curves for **(A)** female (n=36-44 per genotype) and **(B)** male (n=45-78 per genotype) *hSOD1$^{A4V}$* mice. Mice that died from non-disease related causes (e.g. mice culled from fight wounds, or for tissue collection) are censored at the time of death. Censored animals are indicated by tick marks. The study end point was approximately 750 days old for male animals and 850 days old for female animals.

This difference was driven by a rapid decline in the survival of male heterozygous and homozygous *hSOD1$^{A4V}$* mice at approximately 750 days old, whereas a similar decline was not observed in *hSOD1$^{WT}$/hSOD1$^{WT}$* littermates. It was not possible to continue the survival study past ~770 days due to health concerns related to the age of the animals.

The disease-related causes of death/humane end points reached are described in Table 4-6. Many animals had multiple pathologies at the point of death. The occurrence of different pathologies at the time of death are described in Table 4-7. Pathologies were detected following a basic autopsy

procedure – including examination of the heart, lungs, and abdomen – which will only detect gross pathologies. The two most common causes of death were masses in the abdomen and weight loss. The occurrence of masses in the abdomen ($\chi^2(2)=0.34$, p=0.84) and weight loss ($\chi^2(2)=0.32$, p=0.21) did not significantly differ between the genotypes.

| Cause of death | Percentage of animals (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *hSOD1*$^{WT}$/ *hSOD1*$^{WT}$ | | | *hSOD1*$^{WT}$/ *hSOD1*$^{A4V}$ | | | *hSOD1*$^{A4V}$/ *hSOD1*$^{A4V}$ | | | Total | | |
| | m | f | t | m | f | t | m | f | t | m | f | t |
| Mass in abdomen | 8.5 | 7.7 | 8.2 | 10.3 | 4.5 | 8.2 | 4.4 | 8.3 | 6.2 | 8.2 | 6.7 | 7.6 |
| Weight loss | - | 5.1 | 2.0 | 5.1 | 9.1 | 6.6 | 8.9 | 5.6 | 7.4 | 4.4 | 6.7 | 5.3 |
| Found dead | - | - | - | 2.6 | - | 1.6 | 2.2 | 2.8 | 2.5 | 1.6 | 0.8 | 1.3 |
| Hindlimb paralysis | - | 2.6 | 1.0 | - | 2.3 | 0.8 | - | 2.8 | 1.2 | - | 2.5 | 1.0 |
| Righting issues | - | - | - | - | 2.3 | 0.8 | - | 2.8 | 1.2 | - | 1.7 | 0.7 |
| Leg tumour | - | 2.6 | 1.0 | 1.3 | - | 0.8 | - | - | - | 0.5 | 0.8 | 0.7 |
| Neck tumour | - | - | - | 1.3 | - | 0.8 | - | 2.8 | 1.2 | 0.5 | 0.8 | 0.7 |
| Swollen eye | - | - | - | 2.6 | - | 1.6 | - | - | - | 1.1 | - | 0.7 |
| Pale extremities | - | - | - | - | - | - | - | 2.8 | 1.2 | - | 0.8 | 0.3 |
| Seizure | - | - | - | 1.3 | - | 0.8 | - | - | - | 0.5 | - | 0.3 |
| **Group n** | 59 | 39 | 98 | 78 | 44 | 122 | 45 | 36 | 81 | 182 | 119 | 301 |

**Table 4-6 | Disease-related causes of death of *hSOD1*$^{A4V}$ mice.** Mice culled due to fighting are excluded from this analysis.

| Pathologies present at death | Percentage of animals (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ | | | *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ | | | *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ | | | Total | | |
| | m | f | t | m | f | t | m | f | t | m | f | t |
| **Digestive system** | | | | | | | | | | | | |
| Enlarged spleen | 3.4 | 5.1 | 4.1 | 5.1 | 4.5 | 4.9 | - | 5.6 | 2.5 | 3.3 | 5.0 | 4.0 |
| Intestinal tumour(s) | 3.4 | 2.6 | 3.1 | 3.8 | 6.8 | 4.9 | - | - | - | 2.7 | 3.4 | 3.0 |
| Liver tumour(s) | - | - | - | 2.6 | - | 1.6 | 2.2 | - | 1.2 | 1.6 | - | 1.0 |
| Enlarged liver | - | - | - | - | - | - | - | 5.6 | 2.5 | - | 1.7 | 0.7 |
| Pale liver | 1.7 | 2.6 | 2.0 | - | - | - | - | - | - | 0.5 | 0.8 | 0.7 |
| Enlarged gall bladder | - | - | - | - | - | - | - | 2.8 | 1.2 | - | 0.8 | 0.3 |
| **Urinary system** | | | | | | | | | | | | |
| Enlarged kidney(s) | - | - | - | 1.3 | - | 0.8 | - | 2.8 | 1.2 | 0.5 | 0.8 | 0.7 |
| Pale kidney(s) | - | - | - | - | 2.3 | 0.8 | - | - | - | - | 0.8 | 0.3 |
| Polycystic kidney(s) | - | 2.6 | 1.0 | - | - | - | - | - | - | - | 0.8 | 0.3 |
| Swollen bladder | - | - | - | - | - | - | 2.2 | 2.8 | 2.5 | 0.5 | 0.8 | 0.7 |
| **Cardiovascular system** | | | | | | | | | | | | |
| Enlarged heart | - | - | - | 1.3 | 2.3 | 1.6 | - | - | - | 0.5 | 0.8 | 0.7 |
| **Respiratory system** | | | | | | | | | | | | |
| Pale lungs | 3.4 | 2.6 | 3.1 | 1.3 | 2.3 | 1.6 | - | 2.8 | 1.2 | 1.6 | 2.5 | 2.0 |
| **Reproductive system** | | | | | | | | | | | | |
| Swollen seminal vesicle(s) | 1.7 | - | 1.0 | 7.7 | - | 4.9 | 4.4 | - | 2.5 | 4.9 | - | 3.0 |
| **Other** | | | | | | | | | | | | |
| Hindlimb paralysis | - | 2.6 | 1.0 | - | 2.3 | 0.8 | - | 2.8 | 1.2 | - | 2.5 | 1.0 |
| Neck tumour(s) | - | - | - | 1.3 | - | 0.8 | - | 2.8 | 1.2 | 0.5 | 0.8 | 0.7 |
| Leg tumour(s) | - | 2.6 | 1.0 | 1.3 | - | 0.8 | - | - | - | 0.5 | 0.8 | 0.7 |
| Enlarged thymus | - | - | - | - | 2.3 | 0.8 | - | - | - | - | 0.8 | 0.3 |
| Swollen eye | - | - | - | 1.3 | - | 0.8 | - | - | - | 0.5 | - | 0.3 |
| **Group n** | 59 | 39 | 98 | 78 | 44 | 122 | 45 | 36 | 81 | 182 | 119 | 301 |

**Table 4-7 | Pathologies present at death of *hSOD1*$^{A4V}$ mice.** This table includes all disease-relevant pathologies noted at the time of death following a basic autopsy.

### *4.3.6.2 hSOD1-A4V motor and behavioural phenotyping*

#### 4.3.6.2.1 SHIRPA

We assessed *hSOD1*$^{A4V}$ mice for motor and behavioural abnormalities using the SHIRPA test. There was no significant change in tremors, movement fluidity, trunk curl, or the ability to walk down a vertical grate (negative geotaxis) in *hSOD1*$^{A4V}$ mice compared to littermate controls at any time point (Table 4-8).

| Genotype | Sex | Tremors | | | Loss of movement fluidity | | | Trunk curl | | | Negative geotaxis | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Onset (months) | At 24 months | | Onset (months) | At 24 months | | Onset (months) | At 24 months | | Onset (months) | At 24 months | |
| | | | # | % | | # | % | | # | % | | # | % |
| ***hSOD1$^{WT}$/ hSOD1$^{WT}$*** | M | 18 | 1/8 | 13 | 18 | 3/8 | 38 | 24 | 1/8 | 13 | - | 0/8 | 0 |
| ***hSOD1$^{WT}$/ hSOD1$^{WT}$*** | F | - | 0/11 | 0 | 18 | 2/11 | 18 | - | 0/11 | 0 | 24 | 3/10 | 30 |
| ***hSOD1$^{WT}$/ hSOD1$^{A4V}$*** | M | 24 | 1/10 | 10 | 18 | 2/10 | 20 | 24 | 3/10 | 33 | - | 0/10 | 0 |
| ***hSOD1$^{WT}$/ hSOD1$^{A4V}$*** | F | - | 0/9 | 0 | 18 | 4/9 | 44 | 24 | 1/9 | 11 | 24 | 2/9 | 22 |
| ***hSOD1$^{A4V}$/ hSOD1$^{A4V}$*** | M | 24 | 2/12 | 17 | 12 | 4/12 | 33 | 24 | 2/12 | 17 | 18 | 1/11 | 9 |
| ***hSOD1$^{A4V}$/ hSOD1$^{A4V}$*** | F | - | 0/10 | 0 | 18 | 3/10 | 30 | 18 | 2/10 | 20 | 24 | 2/9 | 22 |

**Table 4-8 | SHIRPA phenotyping results from *hSOD1$^{A4V}$* mice.** Only parameters related to motor function are shown. Onset describes the age that the majority of affected animals developed the phenotype.

#### 4.3.6.2.2 Marble burying

We performed marble burying to assess anxiety and obsessive–compulsive-like behaviours in 24-month-old *hSOD1$^{A4V}$* mice. There was no change in marble burying activity in *hSOD1$^{A4V}$* mice compared to littermate controls, as measured by the proportion of marbles which were fully buried (Figure 4-10).

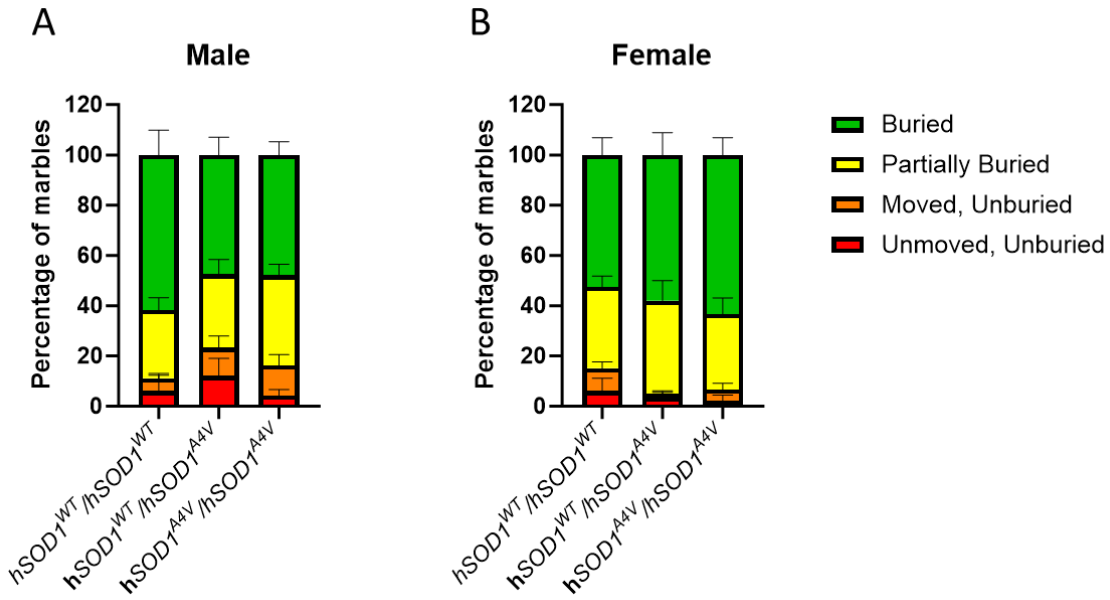



**Figure 4-10 | Marble burying data from *hSOD1$^{A4V}$* mice.** Marble burying data from 24-month-old **(A)** male (n=9-13 per genotype) and **(B)** female (n=9-11 per genotype) *hSOD1$^{A4V}$* mice. Mean ± SEM.

#### 4.3.6.2.3 Open field

The open field test was used to measure locomotor and anxiety-like behaviour in *hSOD1$^{A4V}$* mice. For each sex, two-way ANOVA tests were performed to analyse the effect of genotype and age on locomotor activity and time spent in the periphery. Simple main effects analysis showed that, for male animals, both age (p = <0.001) and genotype (p = <0.001) had a statistically significant effect on locomotor activity. Additionally, genotype had a significant effect on time spent in the periphery (p = <0.001). For female animals, both age (p = <0.001) and genotype (p = 0.010) had a significant effect on locomotor activity; however, there was no effect of age or genotype on time spent in the periphery.

Multiple pairwise comparisons conducted at each time point showed that 24-month-old heterozygous and homozygous male *hSOD1$^{A4V}$* mice had increased locomotor activity compared to *hSOD1$^{WT}$*/*hSOD1$^{WT}$* littermates (Figure 4-11 A). Heterozygous female *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* had increased activity at 6 months old, but this was not present at later time points. 12 and 24-month-

old heterozygous male *hSOD1*$^{A4V}$ mice spent more time in the centre of the open field arena compared to *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ littermates (Figure 4-11 B), indicating reduced anxiety in these animals.

These data suggest that there is increased locomotor activity in aged male, but not female, *hSOD1*$^{A4V}$ animals. There is also reduced anxiety in heterozygous male *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ mice. These data are not indicative of impaired motor function.

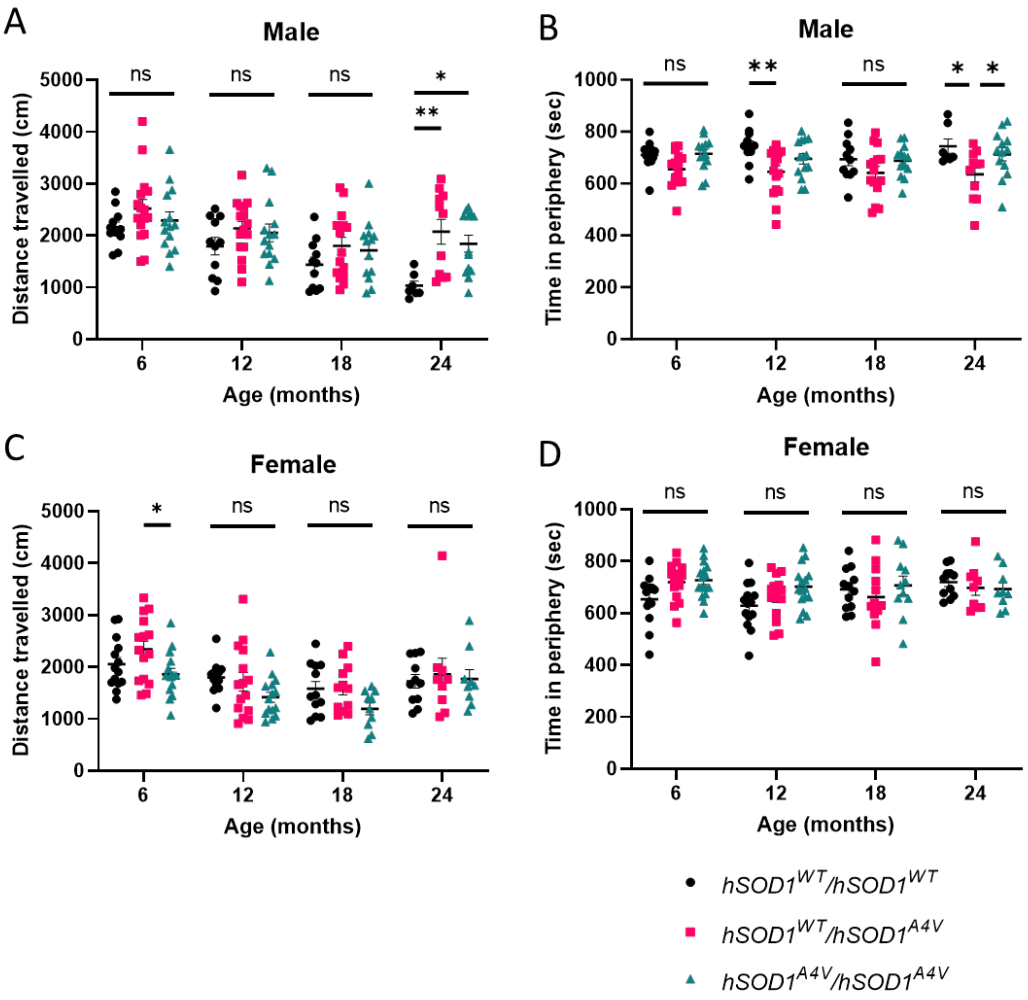



**Figure 4-11 | Longitudinal open field test data from *hSOD1*$^{A4V}$ mice.** Open field test data from **(A, B)** male (n=8-16 per genotype) and **(C, D)** female (n=9-16 per genotype) *hSOD1*$^{A4V}$ mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. *$P < 0.05$, **$P < 0.01$. ns, not significant.

#### 4.3.6.2.4 Grip strength

There was no significant effect of genotype on grip strength in *hSOD1*$^{A4V}$ mice at any time point tested (Figure 4-12). The grip strength values in Figure 4-12 were corrected for bodyweight; analysis of the uncorrected grip strength values also found no significant effect of genotype (data not shown).

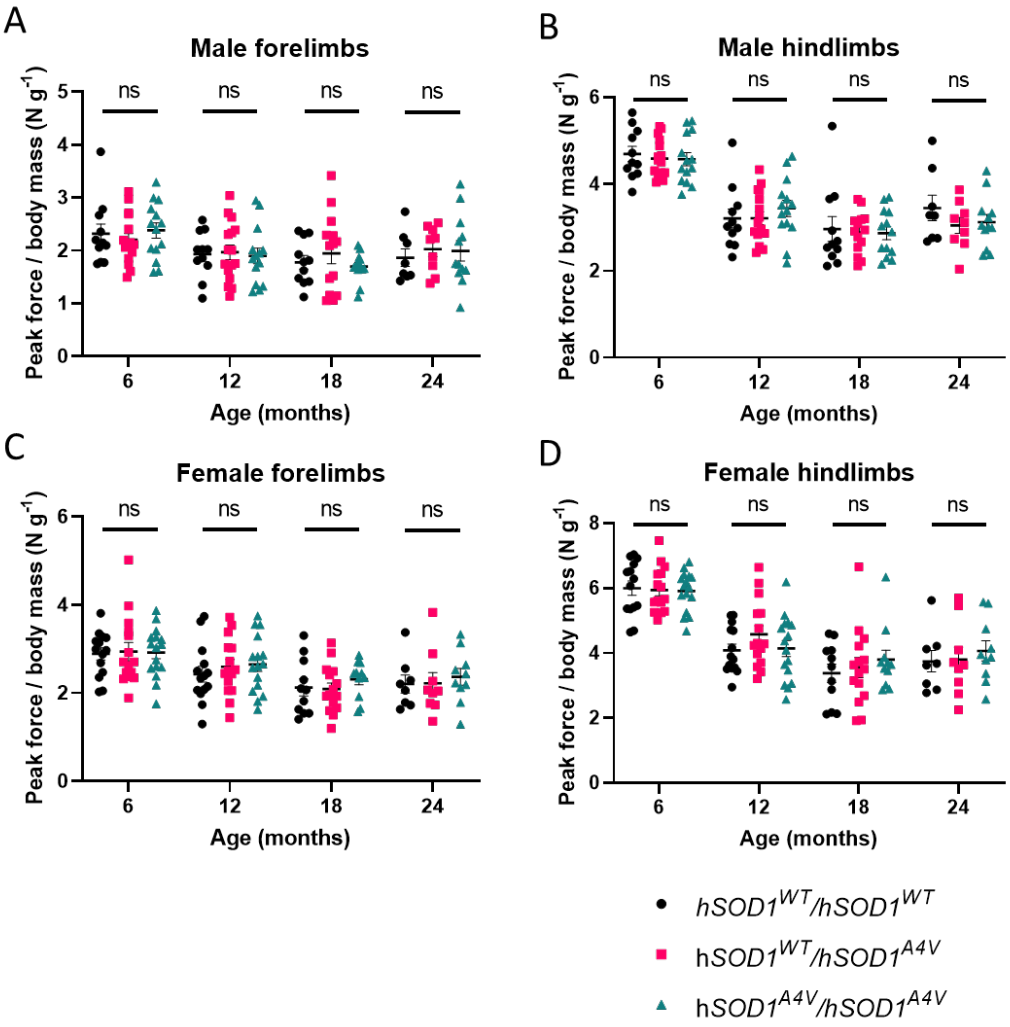



**Figure 4-12 | Longitudinal grip strength data from *hSOD1*$^{A4V}$ mice.** Grip strength data from **(A, B)** male (n=8-16 per genotype) and **(C, D)** female (n=9-16 per genotype) *hSOD1*$^{A4V}$ mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. ns, not significant.

#### 4.3.6.2.5 Gait analysis

Gait in freely moving animals was assessed across 13 measurement parameters. Of the 13 gait parameters analysed, only six were altered in *hSOD1*$^{A4V}$ animals compared to *hSOD1*$^{WT}$ littermates (Figure 4-13 A-F). A description of all gait parameters can be found in section 2.3.2.5. For each sex, two-way ANOVA tests were performed to analyse the effect of genotype and age on each gait parameter.

For male animals, simple main effects analysis revealed a significant effect of genotype on hindlimb stance duration ($p = <0.001$; Figure 4-13 A) and hindlimb duty factor ($p = <0.001$; Figure 4-13 B), and a significant interaction between age and genotype on forelimb swing duration ($p = 0.029$; Figure 4-13 C). Multiple pairwise comparisons conducted at each age found that hindlimb stance duration differed between the genotypes in 12-month-old animals only. Whereas hindlimb duty factor differed in 24-month-old animals and forelimb swing duration differed in 6- and 24-month-old animals.

For female animals, simple main effects analysis revealed a significant effect of genotype on forelimb stride width ($p = 0.006$; Figure 4-13 D), hindlimb stride width ($p = 0.011$ ; Figure 4-13 E) and hindlimb step length ($p = 0.012$ ; Figure 4-13 F). Multiple pairwise comparisons revealed a significant difference between the genotypes for forelimb stride width in 12-month-old animals, but no difference at any time point for hindlimb stride width and hindlimb step length.

We also sought to determine whether *hSOD1*$^{A4V}$ animals could be distinguished from wild-type littermates using a combination of all 13 parameters in principal component analysis. We observed clustering in PCA based on age (Figure 4-13 G, I) but not genotype (Figure 4-13 H, J) for both sexes.

Together these data suggest a limited effect of the A4V mutation on gait in *hSOD1*$^{A4V}$ mice. The most convincingly altered parameter was hindlimb duty factor in 24-month-old male *hSOD1*$^{A4V}$ animals, which is the fraction of the total step cycle for which a given foot is on the ground. However, other

hindlimb gait parameters were not changed at this age, including the closely associated gait parameters of hindlimb swing and stance duration. Furthermore, the reported significance values have not been corrected for multiple testing and therefore should only be considered exploratory. No pairwise comparisons met the threshold for significance once the Bonferroni correction was applied, and therefore any conclusions should be treated with caution.

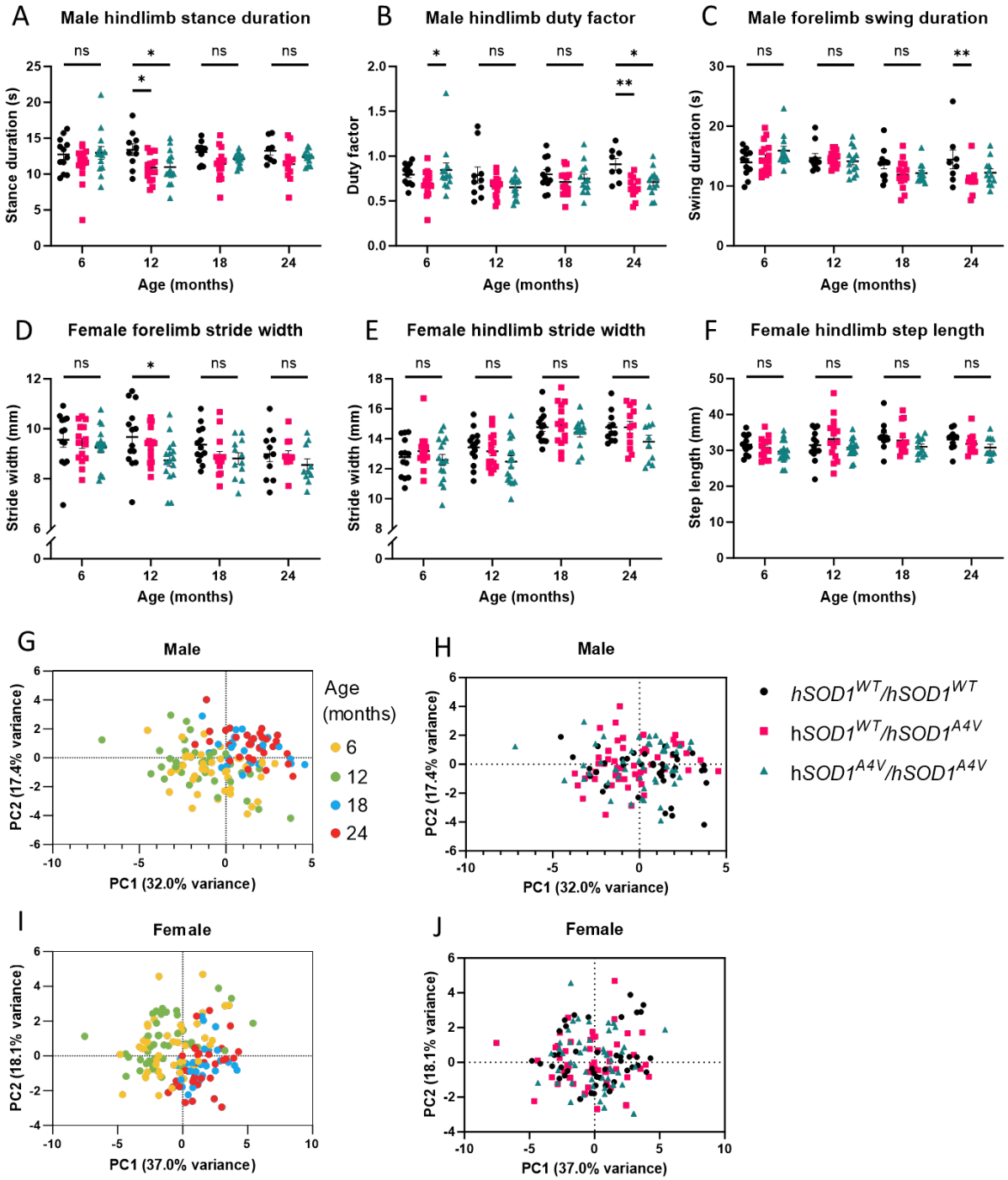



**Figure 4-13 | Longitudinal gait analysis data from *hSOD1$^{A4V}$* mice.** Significant MouseWalker gait analysis parameters from **(A-C)** male (n=8-16 per genotype) and **(D-F)** female (n=9-16 per genotype) *hSOD1$^{A4V}$* mice. The six parameters shown are the only ones found to be significant out of 13 parameters analysed. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. **(G-J)** Principal component analysis (PCA) using all 13 gait analysis parameters as variables in **(G, H)** male and **(I, J)** female *hSOD1$^{A4V}$* mice. **P* < 0.05, ***P* < 0.01. ns, not significant.

#### 4.3.6.2.6 Locotronic

We examined motor coordination in *hSOD1$^{A4V}$* mice using the Locotronic test. Two-way ANOVA with age and genotype as factors found a significant effect of genotype on the number of leg errors made by male *hSOD1$^{A4V}$* mice ($p = 0.002$) (Figure 4-14 A).

Multiple pairwise comparisons conducted at each time point found a significant decrease in the number of leg errors made by 6-month-old male *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* ($p = 0.047$)and 12-month-old male *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* ($p = 0.009$) and *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* ($p = 0.008$) animals compared to age-matched littermate controls (Figure 4-14 A). There was also a significant decrease in leg error time in 12-month-old male *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* ($p = 0.003$) and *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* ($p = 0.003$), and 24-month-old *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* ($p = 0.044$) animals compared to age-matched littermate controls (Figure 4-14 B). No differences between the genotypes were observed in female animals (Figure 4-14 C, D). These data may indicate an inconsistent improvement in motor coordination in male *hSOD1$^{A4V}$* animals.

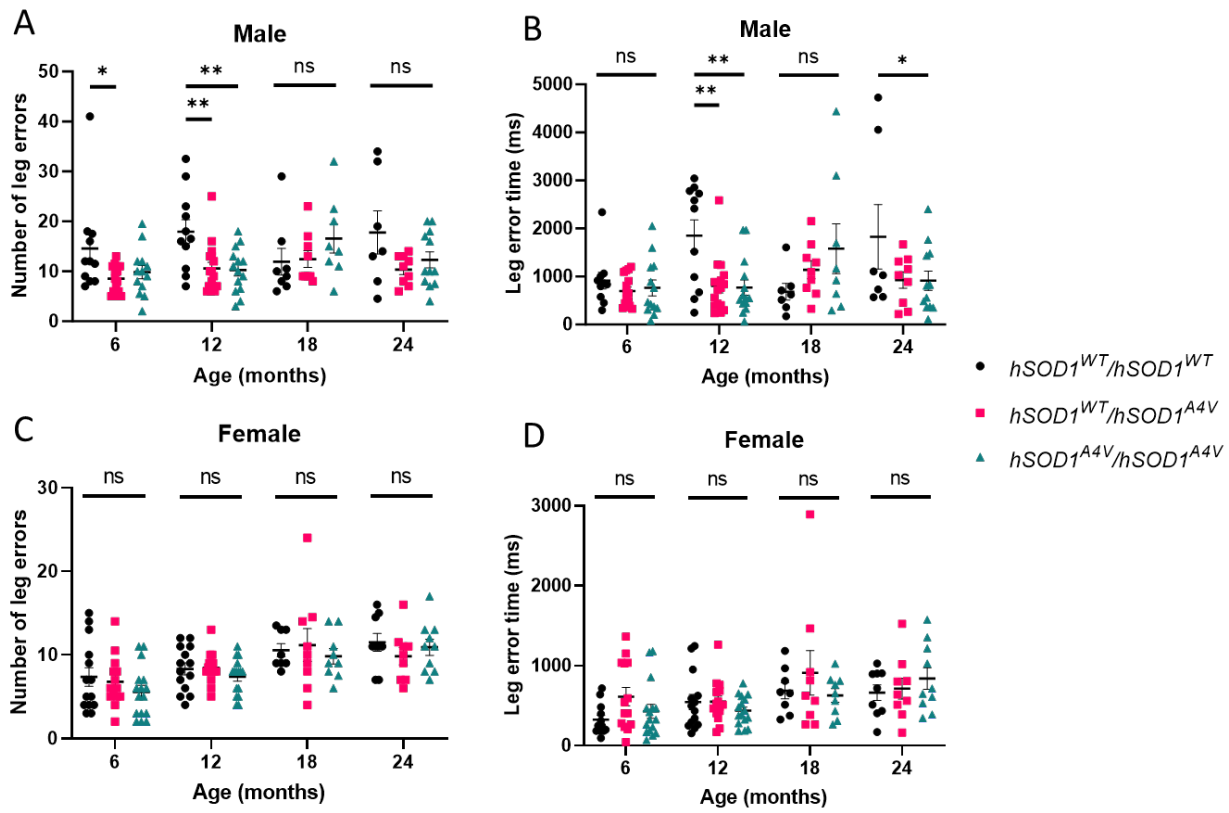



**Figure 4-14 | Longitudinal Locotronic data from *hSOD1$^{A4V}$* mice.** Locotronic data from **(A, B)** male (n=8-16 per genotype) and **(C, D)** female (n=9-16 per genotype) *hSOD1$^{A4V}$* mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. *$P$ < 0.05, **$P$ < 0.01. ns, not significant.

#### 4.3.6.2.7 Rotarod

Two-way ANOVA with age and genotype as factors found a significant effect of genotype on latency to fall from the Rotarod in male *hSOD1$^{A4V}$* mice (p = <0.001) (Figure 4-15, A). Multiple pairwise comparisons conducted at each time point found this effect to be driven by improved Rotarod performance in 18 and 24-month-old heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice. However, only three 24-month-old male heterozygous mice were tested, so the highly significant difference observed at this time point is not as robust as the findings in younger mice. No differences between the genotypes were observed in female animals (Figure 4-15, B). These data may indicate an improvement in motor function in aged heterozygous male *hSOD1$^{A4V}$* animals.

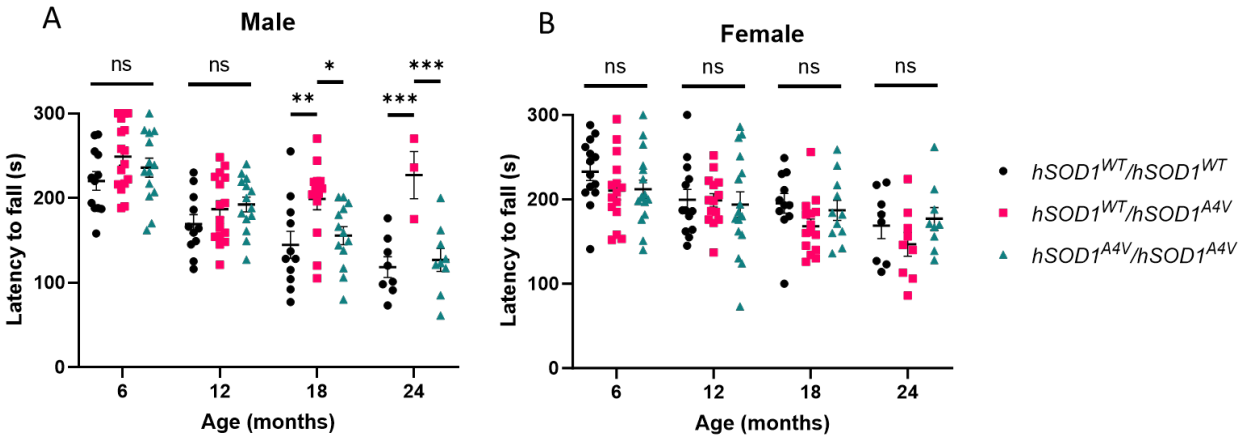



**Figure 4-15 | Longitudinal Rotarod data from *hSOD1$^{A4V}$* mice.** Rotarod data from **(A)** male (n=8-16 per genotype) and **(B)** female (n=9-16 per genotype) *hSOD1$^{A4V}$* mice. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. *$P$ < 0.05. **P < 0.01, ***P < 0.001. ns, not significant.

#### 4.3.6.2.8 Automated wheel running

We examined voluntary wheel running activity in 6, 12 and 18-month-old female *hSOD1$^{A4V}$* mice (Figure 4-16). Weight loss triggered by the stress of single cage housing precluded the use of this test in 24-month-old mice. We evaluated wheel running activity across four parameters: maximum speed, distance travelled, time spent running and the number of bouts of activity.

At each age, mixed effects analyses (similar to a repeated-measures ANOVA) were performed to examine the effect of genotype and test day on the four wheel running parameters. Simple main effects analysis revealed a significant effect of genotype on distance travelled in 12-month-old animals (p = 0.027; Figure 4-16 F), and a significant interaction between test day and genotype on distance travelled (p = 0.014) and time spent running (p = 0.010; Figure 4-16 G). Multiple pairwise comparisons conducted on each test day found an increase in distance travelled and time spent running in heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice compared to homozygous *hSOD1$^{A4V}$/hSOD1$^{A4V}$* only after replacement of the regular running wheel with a complex wheel (which had rungs missing in a semi-random pattern). Mixed effects analysis found no significant differences between the

genotypes at 6 or 18 months old (Figure 4-16 A-D, I-L). There was also no significant difference between the genotypes in maximum speed or the number of bouts of running at any age.

The complex wheel is designed to exacerbate mild motor phenotypes [412]. If *hSOD1$^{A4V}$* animals were to faithfully model ALS then we would expect motor phenotype to progressively worsen. The lack of a clear motor phenotype in 18-month-old animals would provide evidence against this. However, in 18-month-old animals weight loss due to the stress of single housing caused several animals to be removed before the addition of the complex wheel, bringing the group sizes down to n=4-8 animals per genotype. Hence, this experiment had lower power to detect a statistically significant difference between the genotypes. Considering this discrepancy, it is not possible to rule out a mild motor phenotype in 18-month-old *hSOD1$^{A4V}$* mice until this test is repeated with larger groups.

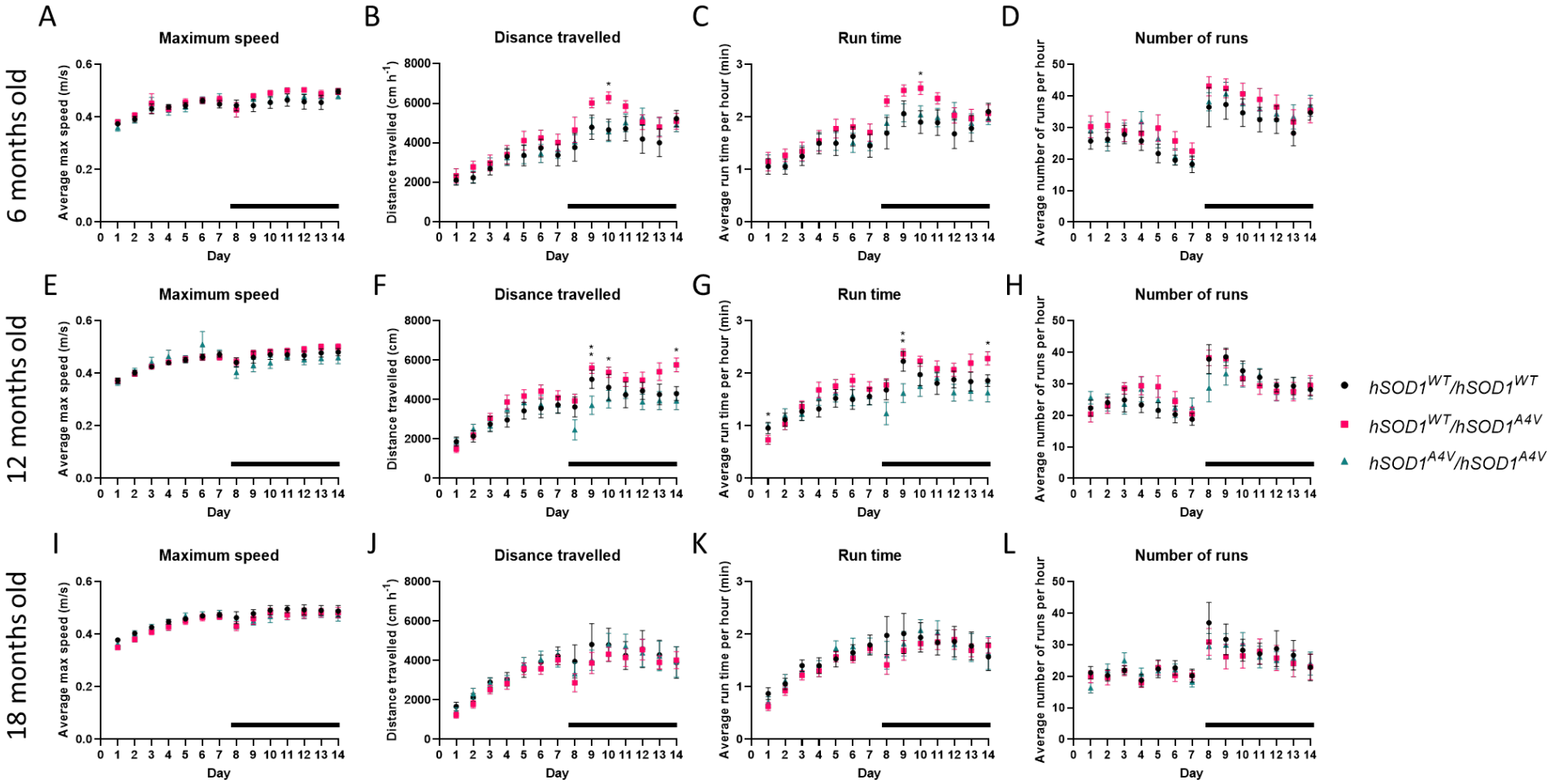



**Figure 4-16 | Longitudinal voluntary wheel running data from *hSOD1$^{A4V}$* mice.** Voluntary wheel running data from **(A-D)** 6-month-old (n=13-16 per genotype), **(E-H)** 12-month-old (n=14-15 per genotype) and **(I-L)** 18-month-old (n=10-14 per genotype) female *hSOD1$^{A4V}$* mice. Black bar indicates days using a complex wheel. Mixed effects analysis followed by the Tukey multiple comparisons test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. Mean ± SEM. *$P$ < 0.05. **P < 0.01.

#### *4.3.6.3 hSOD1-A4V metabolic phenotyping*

Metabolic alterations are consistently implicated in the pathophysiology of ALS [474–477], with metabolic phenotypes observed in a number of mouse models of ALS, including SOD1-G93A mice [463,478]. Hence, we sought to determine if metabolic phenotypes are present in *hSOD1$^{A4V}$* mice.

##### 4.3.6.3.1 hSOD1-A4V bodyweight data

To assess the impact of the A4V mutation on bodyweight, *hSOD1$^{A4V}$* animals were weighed monthly from 1 to 26 months of age (Figure 4-17). Mixed-effects analyses with age and genotype as factors were conducted separately for male and female animals.

Simple main effects analysis revealed a significant effect of time ($p = <0.001$), but no effect of genotype on bodyweight in male *hSOD1$^{A4V}$* animals (Figure 4-17 A). However, there was a significant effect of time ($p = <0.001$) and genotype ($p = 0.022$), and a significant interaction between time and genotype ($p = 0.003$), on bodyweight in female *hSOD1$^{A4V}$* animals (Figure 4-17 B). Pairwise comparisons conducted at each time point revealed this difference to be due to lower bodyweight in female homozygous *hSOD1$^{A4V}$/hSOD1$^{A4V}$* mice compared to age- and sex-matched *hSOD1$^{WT}$/hSOD1$^{WT}$* littermate controls. Pairwise comparisons on animals >18 months of age were not significant; however, the lack of statistical significance was likely due to increased variability in bodyweight and lower n numbers in older animals – a clear trend towards lower bodyweight in female *hSOD1$^{A4V}$/hSOD1$^{A4V}$* animals was consistently observed in mice >18 months of age.

Female *hSOD1$^{A4V}$/hSOD1$^{A4V}$* mice had similar bodyweight to *hSOD1$^{WT}$/hSOD1$^{WT}$* littermate controls until approximately 6 months of age, suggesting the observed difference in bodyweight was not due to impaired development of *hSOD1$^{A4V}$/hSOD1$^{A4V}$* animals.

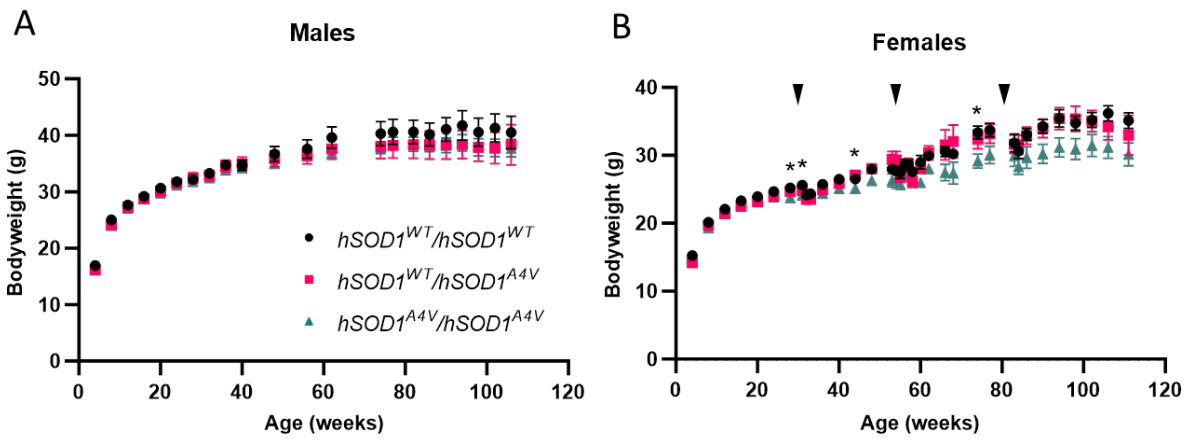



**Figure 4-17 | Longitudinal bodyweight data from *hSOD1$^{A4V}$* mice.** Monthly bodyweight measurement data from **(A)** male (n=12-16 per genotype at 4 weeks old, n=8-12 per genotype at 106 weeks old) and **(B)** female (n=14-16 per genotype at 4 weeks old, n=6-9 per genotype at 111 weeks old) *hSOD1$^{A4V}$* mice. Arrowheads indicate ages where automated wheel running experiments were performed. Mixed effects analysis followed by the Tukey multiple comparisons test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. Mean ± SEM. *$P < 0.05$.

#### 4.3.6.3.2 Echo MRI

We performed ECHO-MRI to examine whether reduced bodyweight in female *hSOD1$^{A4V}$* animals could be attributed to changes in fat mass or lean mass (Figure 4-18). ECHO-MRI was performed on 12, 18 and 24-month-old *hSOD1$^{A4V}$* mice. For each sex, two-way ANOVA tests were performed to analyse the effect of genotype and age on fat mass and lean mass.

Simple main effects analysis showed that, for male animals, neither age nor genotype had a significant effect on fat mass or lean mass as a percentage of total bodyweight (Figure 4-18 A, B). However, there was a significant effect of genotype, but not age, on percentage fat mass ($p = 0.002$) and percentage lean mass ($p = 0.002$) in female *hSOD1$^{A4V}$* animals (Figure 4-18 C, D).

Pairwise comparisons conducted at each age only found a significant increase in percentage lean mass in 24-month-old female *hSOD1$^{A4V}$/hSOD1$^{A4V}$* mice, compared to heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice (Figure 4-18 D). Pairwise comparisons did not reveal a significant change in percentage fat mass in female *hSOD1$^{A4V}$* mice (Figure 4-18 C). However, across all ages there was a general trend towards decreased percentage fat mass and increased percentage lean mass in female

homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice compared to age- and sex-matched *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ littermate controls. Finally, examination of data which was not normalised to bodyweight showed a similar trend towards decreased fat mass in female *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice without any change in lean mass (data not shown).

These data indicate that female *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice have lower fat mass without a loss of lean mass.

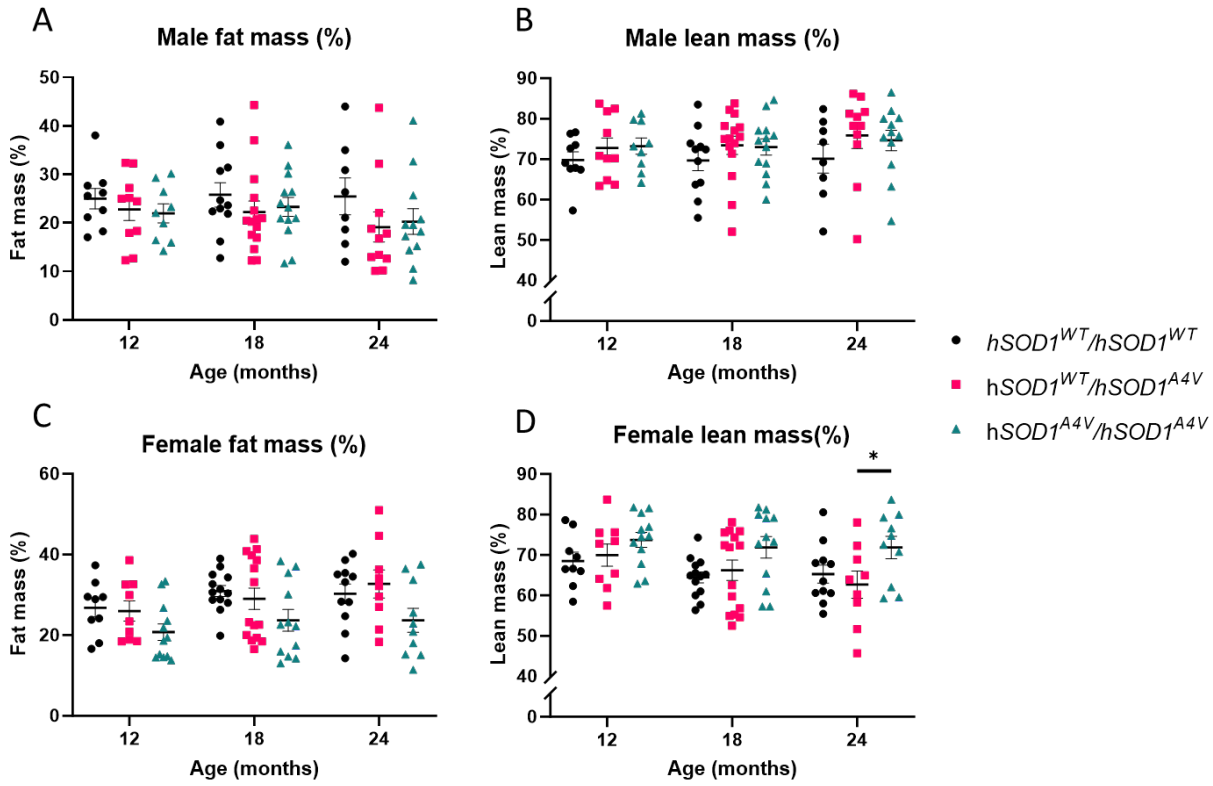



**Figure 4-18 | Echo MRI data from *hSOD1*$^{A4V}$ mice.** Fat mass and lean mass as a percentage of total bodyweight in **(A, B)** male (n=8-15 per genotype) and **(C, D)** female (n=9-15 per genotype) *hSOD1*$^{A4V}$ animals. Mean ± SEM. Two-way ANOVA with Tukey's post hoc test. Reported significance values are from Tukey's test. Main effects are reported in the thesis body. *$P < 0.05$.

#### 4.3.6.3.3 Calorimetric cages

Calorimetric cages (TSE Systems) were used to examine metabolism in 14-month-old female *hSOD1*$^{A4V}$ mice (Figure 4-19). We evaluated heat production, respiratory exchange ratio (RER) and

locomotor activity. Weight loss caused by the stress of single cage housing precluded the use of this test in older mice.

Two-way repeated measures ANOVA tests were performed to examine the effect of genotype and time point on each metabolic parameter. Simple main effects analysis revealed a significant effect of time on heat production ($p = <0.001$) and a significant interaction between time and genotype ($p = 0.007$; Figure 4-19 A). Multiple pairwise comparisons conducted at each time found an increase in heat production by heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ mice, but this was only significantly higher than *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ controls for 2 hours over the 24 hour test period. There was no significant effect of genotype on RER (Figure 4-19 B), suggesting there is no significant difference in fuel preference between *hSOD1*$^{A4V}$ mice and littermate controls.

There was a significant effect of genotype ($p = 0.040$) and time ($p = <0.001$) on locomotor activity and a significant interaction between time and genotype ($p = 0.024$; Figure 4-19 C). Multiple pairwise comparisons conducted at each time found decreased activity in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals compared to *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ and *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals.

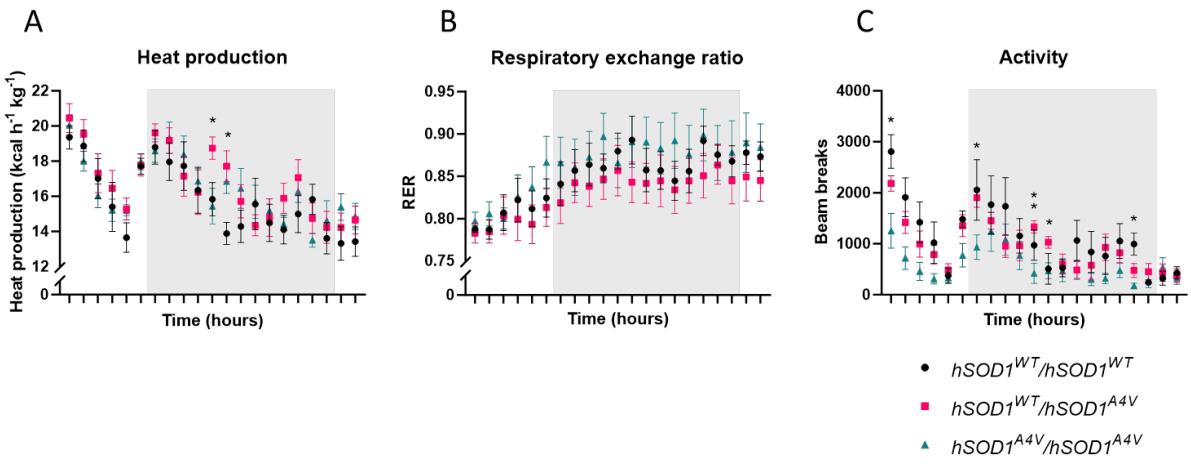



**Figure 4-19 | Calorimetric cage data from *hSOD1*$^{A4V}$ mice.** Data from 14-month-old female *hSOD1*$^{A4V}$ mice housed in Phenomaster calorimetric cages for 24hrs. **(A)** Heat production corrected for bodyweight. **(B)** Respiratory exchange ratio (RER). **(C)** Locomotor activity. Grey box indicates dark phase (19:00-7:00). n=7-8 animals per genotype. Two-way repeated measures ANOVA followed by the Tukey multiple comparisons test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. Mean ± SEM. *$P < 0.05$. **$P < 0.01$.

#### 4.3.6.3.4 IPGTT

Differences in glucose metabolism were investigated in 14 and 18-month-old *hSOD1*$^{A4V}$ mice using the intraperitoneal glucose tolerance test (IPGTT; Figure 4-20). A high prevalence of abdominal growths precluded the use of IPGTT in mice >18 months old. For each sex, two-way repeated measures ANOVA tests were used to investigate the effect of genotype and sampling time after glucose infusion on blood glucose levels.

In 14-month-old *hSOD1*$^{A4V}$ animals, simple main effects analysis revealed a trending effect of genotype on blood glucose levels in both male ($p = 0.071$) and female ($p = 0.076$) mice (Figure 4-20 A, B). Multiple pairwise comparisons conducted at each sampling time found a significant, or nearly significant, increase in blood glucose 15 and 30 minutes after glucose infusion in male homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals, compared to *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ littermate controls (Figure 4-20 A). In female mice, multiple pairwise comparisons found a significant increase in blood glucose 30 minutes after glucose infusion in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals compared to heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ animals (Figure 4-20 B). Examination of the area under the curve (AUC) across all sampling points revealed a trend towards increased blood glucose in homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice in both sexes (Figure 4-20 A, B). There was no difference in fasted blood glucose between the genotypes, as measured by blood glucose level at time T0.

We repeated the same analysis in 18-month-old *hSOD1*$^{A4V}$ animals (Figure 4-20 C, D). In 18-month-old animals simple main effects analysis did not find a significant effect of genotype on blood glucose levels in either sex. However, pairwise comparisons conducted at each sampling time found significant and trending increases in blood glucose in female *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals verses littermate controls at 15 and 30 minutes after glucose infusion (Figure 4-20 D). Visual examination of the data revealed similar patterns in 14- and 18-month-old animals. However in older animals, n numbers were lower and variability between animals was higher, reducing our power to detect statistically significant differences.

These data show that 14-month-old *hSOD1*$^{A4V}$ animals have an impaired early response to glucose, which is indicative of disturbed glucose metabolism. 18-month-old *hSOD1*$^{A4V}$ animals may also have an impaired response to glucose, but the data is less conclusive due to low n numbers.

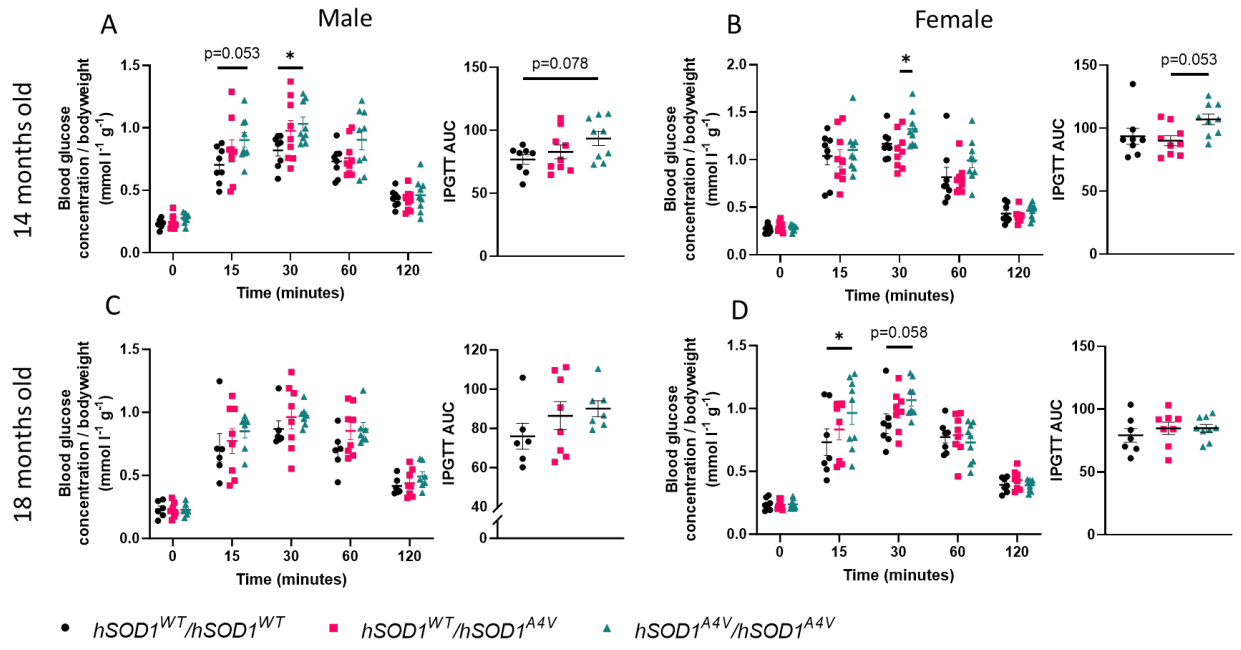



**Figure 4-20 | Intraperitoneal glucose tolerance testing (IPGTT) in *hSOD1*$^{A4V}$ mice.** IPGTT data from 14-month-old **(A)** male (n=8-9 per genotype) and **(B)** female (n=8-9 per genotype), and 18-month-old **(C)** male (n=6-8 per genotype) and **(D)** female (n=7-9 per genotype) *hSOD1*$^{A4V}$ mice. Changes in blood glucose over time were analysed using two-way repeated measures ANOVA followed by Tukey's test. Reported significance values are from Tukey's test. Main effects are reported in the thesis body. Area under the curve (AUC) for each graph was analysed using one-way ANOVA followed by Tukey's test. Mean ± SEM. *$P < 0.05$.

#### 4.3.6.3.5 Metabolomics

Metabolic changes have been widely observed in ALS mouse models, including at pre-symptomatic stages [463,476,479–481]. Hence we sought to establish if metabolism is perturbed in *hSOD1*$^{A4V}$ mice.

An analysis of blood metabolites in *hSOD1*$^{A4V}$ animals was carried out using blood samples collected at 4, 6, 9, 13 and 20 months of age. At each time point, a range of metabolites was measured in blood plasma using liquid chromatography–mass spectrometry (LC-MS) and gas chromatography–mass spectrometry (GC-MS). We also measured the levels of the antioxidant system components

reduced glutathione (GSH), oxidised glutathione (GSSG) and ascorbic acid (ASA) in whole blood. Samples from the 4- to 13-month-old animals were analysed together in one set of MS experiments, whereas samples from 20-month-old animals were analysed separately in a second set of MS experiments. A significant difference between the sexes was observed, hence pooled data from both sexes were analysed at each time point using Linear Models for Microarray Data (limma) with sex as a covariate.

MetaboAnalyst requires a minimum of 3 animals per group for analysis. Some groups in the LC-MS analysis had only two animals; therefore, to enable analysis of these groups an additional datum point was imputed based on the median value for that group. Imputed values were added for 4-month-old female homozygotes, 13-month-old female WTs, 13-month-old female homozygotes and 13-month-old male WTs. In these cases the statistical analysis is exploratory and should not be considered robust. Larger group sizes were used at the 20 month time point (n=6-8 per sex per genotype) so this data should be considered the most robust.

#### 4.3.6.3.5.1 MS analysis

When using a conservative significance threshold of padj < 0.05 only a single metabolite – xylulose – was significantly changed between the genotypes. Xylulose was found to be significantly increased in plasma from 20-month-old homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals compared to littermate controls (padj = $4x10^{-6}$; Figure 4-22 R). Xylulose is a monosaccharide intermediate in the pentose phosphate pathway and has not previously been reported as dysregulated in ALS. Xylulose levels were not assessed at earlier timepoints due to differences in the GC-MS spectrum annotation.

Because of the limited metabolic changes we relaxed the significance threshold to p < 0.05. At this threshold a number of metabolites were significantly changed (Figure 4-21; Figure 4-22). Appendix 8.2 contains a full list of significantly changed metabolites. Although there was little overlap in individual metabolites significantly changed between timepoints, a few classes of metabolites consistently changed across time points.

Different forms of carnitine were significantly changed in *hSOD1*$^{A4V}$ animals from 6 months old. Carnitine is an amino acid which plays an important role in mitochondrial lipid metabolism and is required for the transport of long-chain fatty acids into mitochondria [482,483]. Different forms of carnitine in our dataset represent carnitine bound to different cargo. Various forms of lysophosphatidylcholine (lysoPC) were also altered at every timepoint (Figure 4-21 F, J, N, T). LysoPCs are phospholipids derived from phosphatidylcholine [484]. The complex nature of lipid metabolism makes determining the causes and effects of these changes extremely challenging. Regardless, lipid dysregulation has been widely reported in ALS patients [485–487] and mouse models [463,480] and is increasingly recognised as central to ALS aetiology.

Plasma cholesterol was decreased in homozygous 20-month-old homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ animals. Serum cholesterol changes have been reported in ALS patients, although a clear consensus on the direction of change is lacking [488]. Upregulation of cholesterol synthesis pathways was recently reported in a physiological mouse model of SOD1-ALS [372].

The plasma concentration of a number of amino acids were found to be significantly changed in 20-month-old *hSOD1*$^{A4V}$ animals by GC-MS (Figure 4-22 S, T). Changes in plasma amino acids have been suggested as biomarkers in *tgSOD1*$^{G93A}$ mice [479].

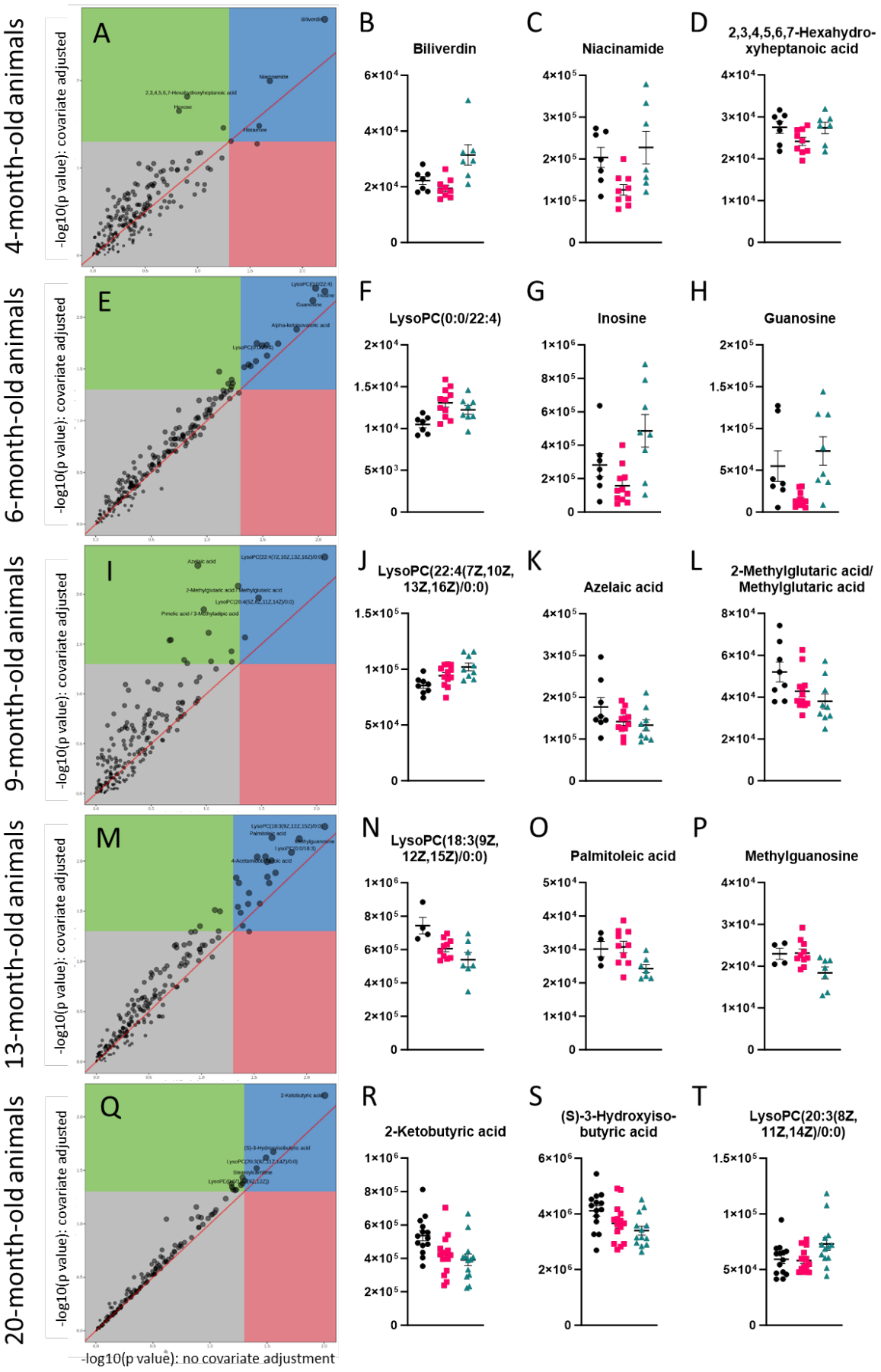

4-month-old animals
6-month-old animals
9-month-old animals
13-month-old animals
20-month-old animals
-log10(p value): covariate adjusted
-log10(p value): no covariate adjustment
A
B
Biliverdin
C
Niacinamide
D
2,3,4,5,6,7-Hexahydro-xyheptanoic acid
E
F
LysoPC(0:0/22:4)
G
Inosine
H
Guanosine
I
J
LysoPC(22:4(7Z,10Z, 13Z,16Z)/0:0)
K
Azelaic acid
L
2-Methylglutaric acid/ Methylglutaric acid
M
N
LysoPC(18:3(9Z, 12Z,15Z)/0:0)
O
Palmitoleic acid
P
Methylguanosine
Q
R
2-Ketobutyric acid
S
(S)-3-Hydroxyiso-butyric acid
T
LysoPC(20:3(8Z, 11Z,14Z)/0:0)

**Figure 4-21 | LC-MS analysis of plasma metabolites in *hSOD1$^{A4V}$* mice.** Analysis of plasma metabolites in **(A-D)** 4-month-old, **(E-H)** 6-month-old, **(I-L)** 9-month-old, **(M-P)** 13-month-old and **(Q-T)** 20-month-old male and female *hSOD1$^{A4V}$*. Graphs on the left of the figure show significant metabolites at each timepoint with and without covariate adjustment (p value cutoff = $p < 0.05$). Blue = significant. Green = significant when adjusted for sex as a covariate. Red = non-significant when adjusted for sex as a covariate. Graphs on the right of the figure show the top 3 most significantly changed metabolites at each timepoint. Limma analysis was performed with imputed values for some groups. Imputed values are not displayed in the graphs of individual metabolites to show the real data distribution. A full list of significantly changed metabolites, with p values, can be found in appendix 8.2.

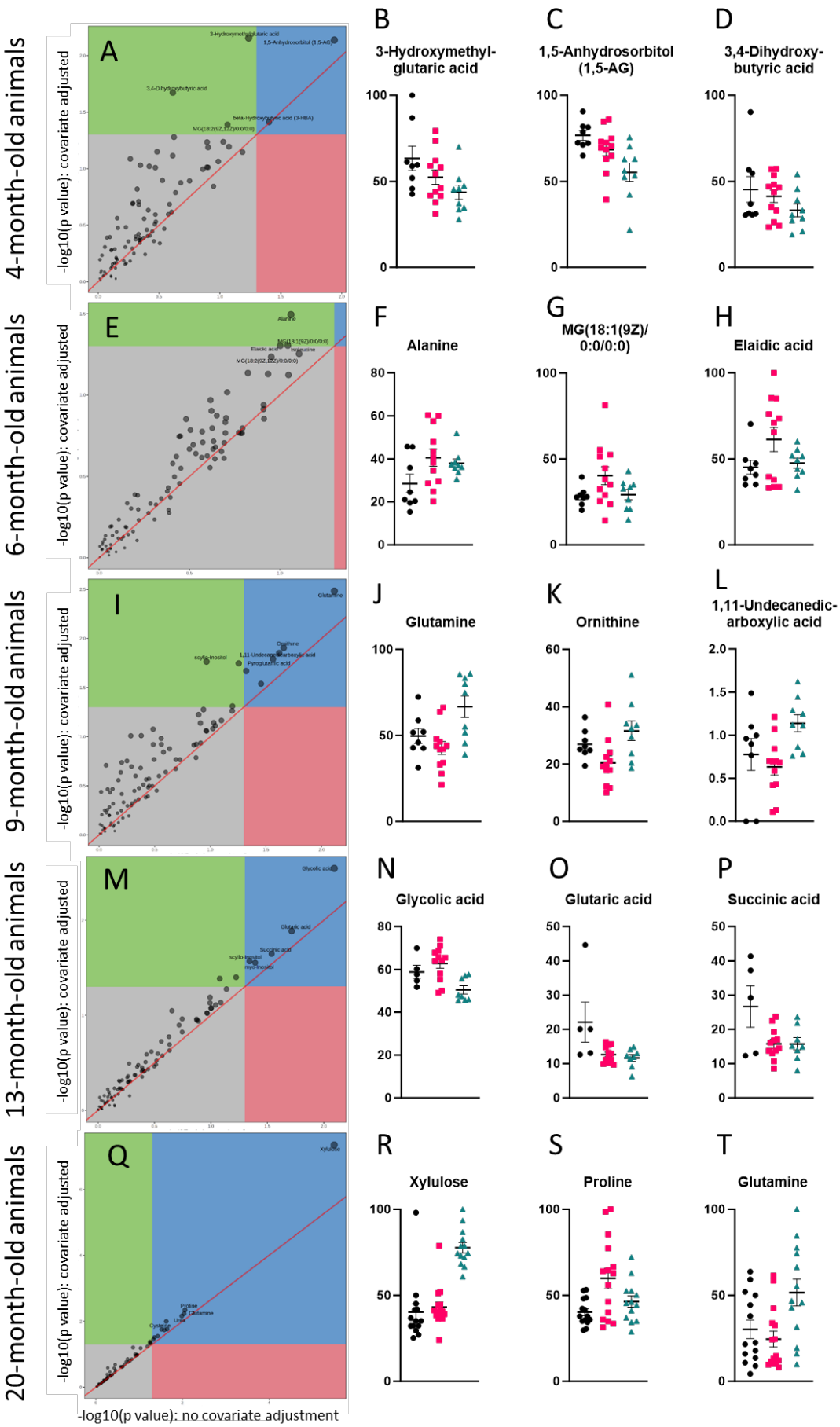

4-month-old animals
6-month-old animals
9-month-old animals
13-month-old animals
20-month-old animals
-log10(p value): covariate adjusted
-log10(p value): no covariate adjustment
A
B
3-Hydroxymethyl-glutaric acid
C
1,5-Anhydrosorbitol (1,5-AG)
D
3,4-Dihydroxy-butyric acid
E
F
Alanine
G
MG(18:1(9Z)/0:0/0:0)
H
Elaidic acid
I
J
Glutamine
K
Ornithine
L
1,11-Undecanedic-arboxylic acid
M
N
Glycolic acid
O
Glutaric acid
P
Succinic acid
Q
R
Xylulose
S
Proline
T
Glutamine

**Figure 4-22 | GC-MS analysis of plasma metabolites in *hSOD1$^{A4V}$* mice.** Analysis of plasma metabolites in **(A-D)** 4-month-old, **(E-H)** 6-month-old, **(I-L)** 9-month-old, **(M-P)** 13-month-old and **(Q-T)** 20-month-old male and female *hSOD1$^{A4V}$*. Graphs on the left of the figure show significant metabolites at each timepoint with and without covariate adjustment (p value cutoff = $p < 0.05$). Blue = significant. Green = significant when adjusted for sex as a covariate. Red = non-significant when adjusted for sex as a covariate. Graphs on the right of the figure show the top 3 most significantly changed metabolites at each timepoint. Limma analysis was performed with imputed values for some groups. Imputed values are not displayed in the graphs of individual metabolites to show the real data distribution. A full list of significantly changed metabolites, with p values, can be found in appendix 8.2.

To determine whether the metabolites altered in plasma from 20-month-old animals were enriched for specific metabolic pathways we ran a pathway analysis of our altered metabolites against the mus musculus KEGG pathway library using Metaboanalyst 5.0. Four pathways were significantly enriched at a significance threshold of $p < 0.05$ (Table 4-9), however these four pathways were largely driven by changes in the same metabolites, primarily amino acids, oxobutanoate, sarcosine, and pyruvate. Of these pathways, Glycine, serine and threonine metabolism was found to be altered in serum [489] and muscle [490] from ALS patients, and aminoacyl-tRNA biosynthesis was found to be altered in ALS patient muscle [490] and in iPSC-derived MNs from a SOD1-A4V patient [481].

| Pathway Name | p value | FDR | Overlap |
|---|---|---|---|
| **Glycine, serine and threonine metabolism** | 0.0007 | 0.04 | Glycine; Sarcosine; D-Glycerate; L-Cysteine; 2-Oxobutanoate; Pyruvate |
| **Cysteine and methionine metabolism** | 0.0045 | 0.12 | L-Methionine; L-Cysteine; 2-Oxobutanoate; Pyruvate |
| **Aminoacyl-tRNA biosynthesis** | 0.0100 | 0.18 | L-Phenylalanine; L-Glutamine; L-Cysteine; Glycine; L-Methionine; L-Proline |
| **Glyoxylate and dicarboxylate metabolism** | 0.0172 | 0.24 | Glycine; D-Glycerate; Pyruvate; L-Glutamine |

**Table 4-9 | Metabolic pathway analysis of significantly altered metabolites in plasma from 20-month-old *hSOD1$^{A4V}$* mice.**

#### 4.3.6.3.5.2 Amino acid quantification

To validate changes in amino acids in 20-month-old animals the Swedish Metabolomics Centre performed a more sensitive AccQ-based assay to quantify amino acids in the plasma samples. We observed a moderate effect of sex and a strong effect of cohort on amino acid levels, hence data were analysed using limma with sex and cohort as covariates.

In contrast to the GC-MS data, phenylalanine was the only amino acid whose levels differed significantly between the genotypes (p=0.029; Figure 4-23). This effect was driven by increased plasma phenylalanine in heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ animals. The reason for the discrepancy between the GC-MS and amino acid quantification is unclear.

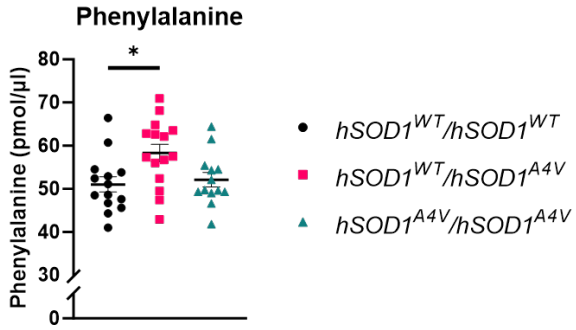



**Figure 4-23 | Plasma phenylalanine quantification in *hSOD1*$^{A4V}$ mice.** Data analysed by limma as part of a panel of amino acids. Reported significance value is from the Tukey multiple comparisons test. Mean ± SEM. **P* < 0.05.

#### 4.3.6.3.5.3 Multivariate MS analysis

To determine whether the genotypes could be distinguished by global changes across the spectrum of metabolites – i.e. "metabolic fingerprinting" [491] – we performed PCA and PLS-DA on male and female animals separately at every timepoint that did not have imputed values. In the LC-MS data, we observed clustering by genotype for 6-month-old female animals in PCA and PLS-DA (Figure 4-24 A, B). In the GC-MS data, we observed clustering by genotype for 4-month-old females in PCA (Figure 4-24 C), and for 9-month-old males and females in PLS-DA (Figure 4-24 C, D) – although these PLS-DA models were only weakly predictive due to high cross-validation error (data not shown). No clustering by genotype was observed for 20-month-old animals, which was the dataset with the highest number of replicates.

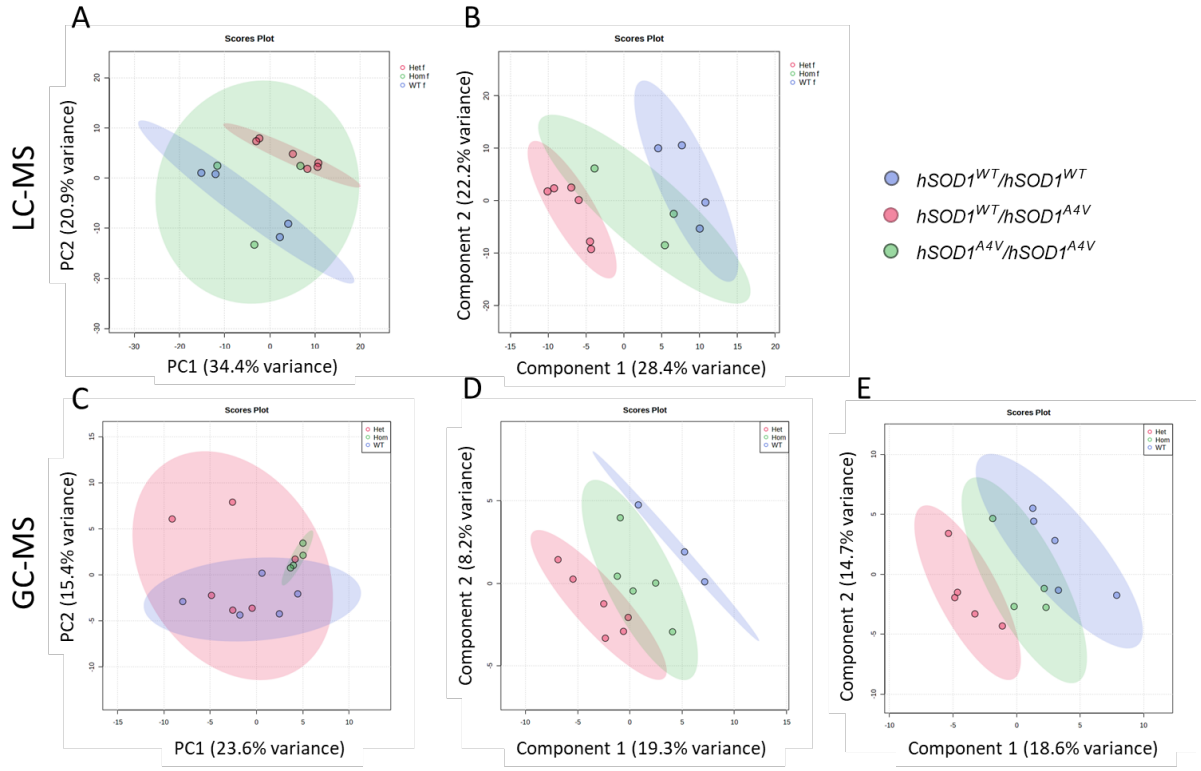



**Figure 4-24 | Multivariate analysis of *hSOD1$^{A4V}$* plasma metabolomics data. (A)** PCA and **(B)** PLS-DA score plots of LC-MS data from 6-month-old female *hSOD1$^{A4V}$* mice. **(C)** PCA score plot of GC-MS data from 4-month-old female *hSOD1$^{A4V}$* mice. PLS-DA score plots of GC-MS data from 9-month-old **(D)** male and **(E)** female *hSOD1$^{A4V}$* mice. The PLS-DA score plots **(D)** and **(E)** are only weakly predictive due to high cross-validation error. PCA and PLS-DA analyses were performed on MS data from 4, 6, 9, 13 and 20-month-old animals; only data with significant clustering by genotype is shown in this figure. PC, principal component.

#### 4.3.6.3.5.4 Glutathione and ascorbate quantification

Reduced glutathione (GSH) reacts with ROS – and is converted to its oxidised form (GSSG) in the process – forming a key part of the antioxidant defence [492]. Similarly, ascorbic acid (ASA; vitamin C) has an important role as a scavenger of ROS [493]. Both antioxidants play key roles in the CNS, with ASA being more abundant in neurons and GSH being more abundant in glia [494]. A disruption in SOD1 function has been shown to perturb other antioxidant systems [223,495], hence we examined the levels of GSH, GSSG and ASA in blood from *hSOD1$^{A4V}$* mice (Figure 4-25).

There was no difference in GSH/GSSG/ASA levels between the sexes, hence the sexes were pooled and data were analysed by one-way ANOVA. The 20-month-old blood samples were taken from two cohorts of animals in separate experiments. We observed a significant effect of cohort on the levels of GSH and GSSG. Hence, data at this timepoint was analysed by two-way ANOVA with genotype and cohort as factors. ASA was not measured in 20-month-old animals.

Levels of GSH were unchanged at all timepoints (Figure 4-25 A, E, I, M, Q). We observed a significant decrease in GSSG in 13- ($p=0.036$; Figure 4-25 N) and 20-month-old ($p<0.001$; Figure 4-25 R) $hSOD1^{A4V}$/$hSOD1^{A4V}$ animals compared to controls. Accordingly we observed an increase in the ratio of GSH-to-GSSG in 20-month-old $hSOD1^{A4V}$/$hSOD1^{A4V}$ animals ($p=<0.001$; Figure 4-25 S). An increased GSH-to-GSSG ratio is indicative of reduced oxidative stress in $hSOD1^{A4V}$/$hSOD1^{A4V}$ animals and is the opposite of what is seen in ALS patient blood [496] and CSF [497]. No changes in ASA levels were observed (Figure 4-25 D, H, L, P).

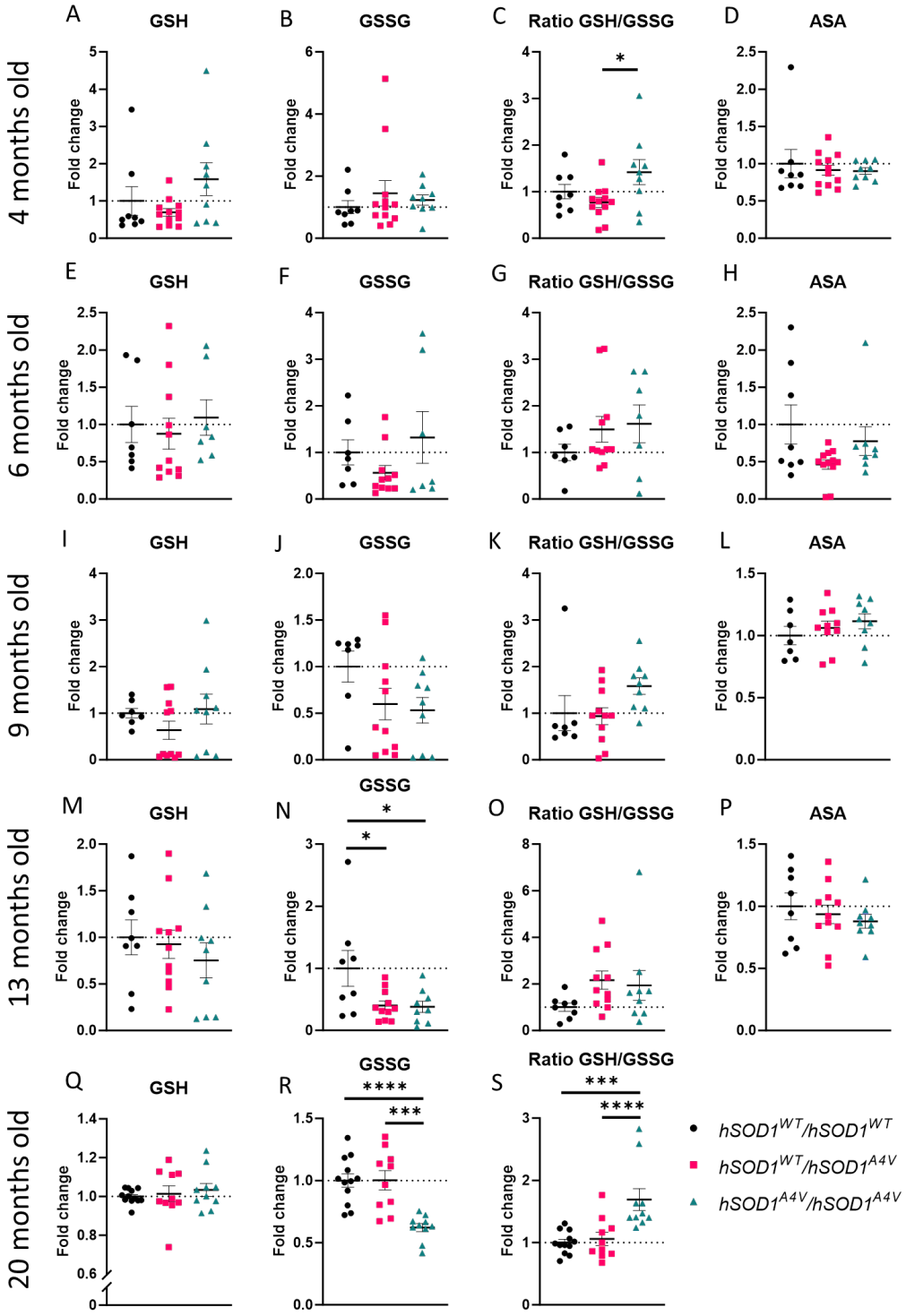



**Figure 4-25 | Blood glutathione and ascorbate measurements from *hSOD1$^{A4V}$* mice.** Quantification of reduced and oxidised glutathione, and ascorbic acid in whole blood from **(A-D)** 3-month-old, **(E-H)** 6-month-old, **(I-L)** 9-month-old and **(M-P)** 13-month-old *hSOD1$^{A4V}$* mice. **(Q-S)** Quantification of reduced and oxidised glutathione in whole blood from 20-month-old *hSOD1$^{A4V}$* mice. All data are displayed as fold change relative to *hSOD1$^{WT}$/hSOD1$^{WT}$* animals. 20-month-old data was pooled from two cohorts where each animal is displayed as fold change relative to *hSOD1$^{WT}$/hSOD1$^{WT}$* animals from the same cohort. Data from 4-13-month-old animals analysed by one-way ANOVA with Tukey's test. Data from 20-month-old animals data analysed by two-way ANOVA with Tukey's test. Mean ± SEM. *$P < 0.05$, ***$P < 0.001$, ****$P < 0.0001$. GSH, glutathione; GSSG, glutathione disulphide; ASA, ascorbic acid.

#### 4.3.6.3.6 Assessment of mitochondrial function in cells from hSOD1-A4V mice

Impairment of mitochondrial function is consistently implicated in the pathology of ALS [498,499]. Hence, we sought to evaluate mitochondrial function *in vitro* in cells isolated from the spinal cord of embryonic day (E) 13.5 *hSOD1$^{A4V}$* embryos. Similar to adult spinal cord, the level of SOD1 in the spinal cord of *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* and *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* embryos was found to be markedly reduced compared to *hSOD1$^{WT}$*/*hSOD1$^{WT}$* controls (Figure 4-26 A, B). Staining for neuronal/glial markers revealed neurons, astrocytes and microglia in our cultures (appendix 8.3); however, immunofluorescence staining for the MN markers ChAT and Islet1 did not identify any MNs, either due to poor immunofluorescence signal or a lack of viable MNs in the culture (data not shown).

*In vitro* mitochondrial function did not differ significantly between the genotypes in Seahorse Mito Stress Test assays (Agilent) under basal conditions, or when the cells were stressed 4 hrs prior to the assay by addition of 0.5 mM sodium arsenite to the culture media (Figure 4-26 C-H). Furthermore, no difference in TMRE fluorescence – an indicator of mitochondrial membrane potential – was observed between the genotypes (Figure 4-26 I).

The *SOD1-A4V* mutation has been found to impair the survival of human iPSCs and iPSC-MNs *in vitro* [462]. Additionally, cells from a human patient with a homozygous *SOD1* truncating mutation showed extreme oxygen sensitivity in culture [221]. However, in the context of *hSOD1$^{A4V}$* mice we did not find the *SOD1-A4V* mutation or the associated decrease in *SOD1* protein to impact cell survival in culture, as measured by cell survival assays performed at 7 days *in vitro* (DIV; Figure 4-26 J, K).

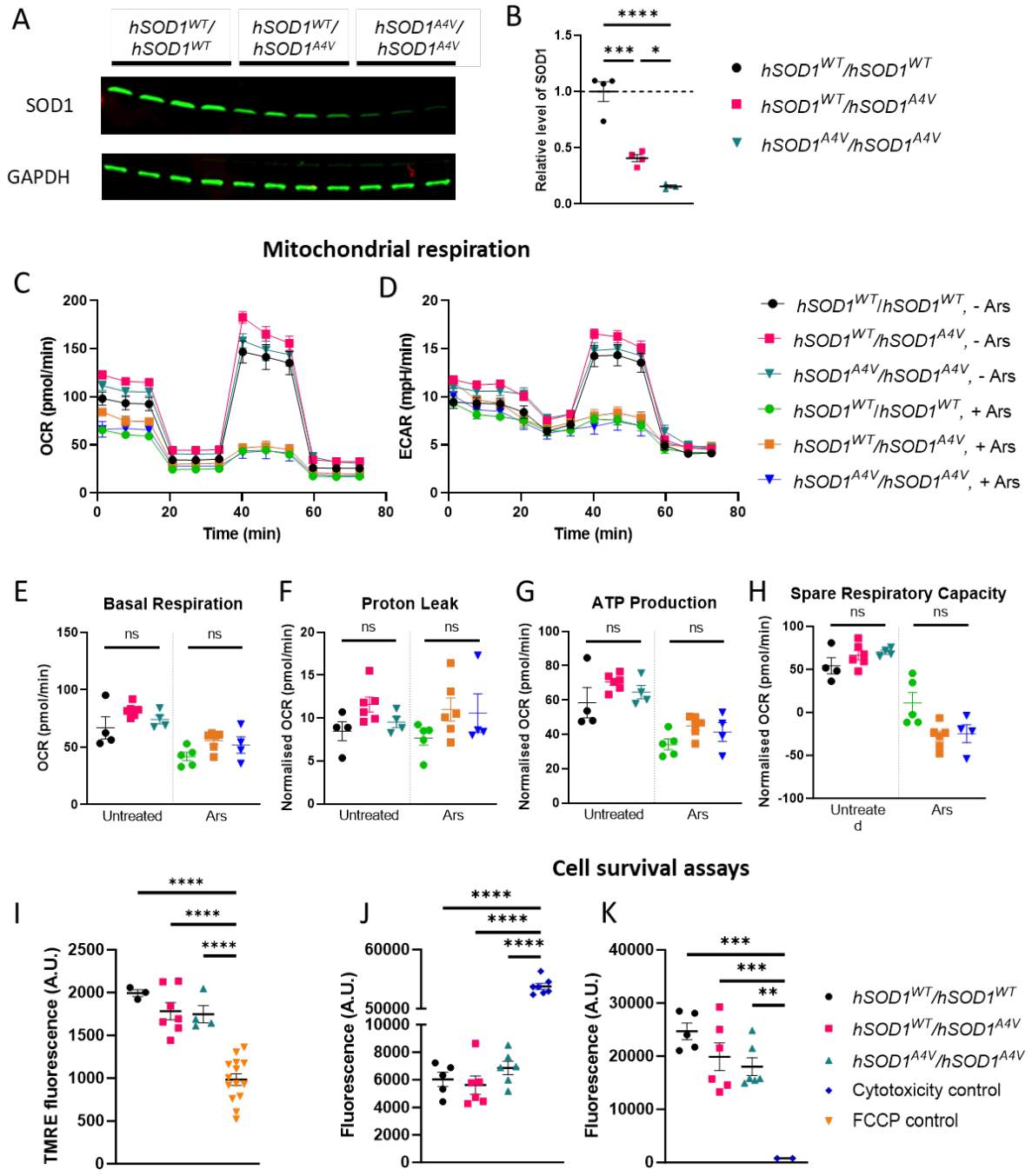



**Figure 4-26 | Mitochondrial function in spinal cord cultures from *hSOD1$^{A4V}$* mice. (A)** SOD1 Western blot performed on embryonic E13.5 spinal cord (n=3-4 per genotype). Blot normalised to GAPDH. **(B)** Quantification of SOD1 immunoreactivity in **(A). (C)** Oxygen consumption rate (OCR), **(D)** extracellular acidification rate (ECAR) and **(E-H)** mitochondrial respiration parameters – derived from OCR – from a Seahorse assay performed on primary spinal cord cells isolated from E13.5 embryos (n=4-6 embryos per genotype). Mitochondrial respiration was assessed under basal conditions and following treatment with 0.5 mM sodium arsenite. **(I)** TMRE mitochondrial membrane potential assay (n=3-7 embryos per genotype). FCCP eliminates mitochondrial membrane potential as a negative control for TMRE staining. **(J)** CellTox Green and **(K)** CellTiter Blue cell survival assays performed on primary spinal cord cells isolated from E13.5 embryos at 7 DIV (n=5-6 embryos per genotype). All data analysed by one-way ANOVA with Tukey's multiple comparisons test. Mean ± SEM. *$P < 0.05$, ***$P < 0.001$, ****$P < 0.0001$. Ars, sodium arsenite; A.U., arbitrary units.

### 4.3.7 The effect of genetic background on phenotype

Genetic background has a substantial impact on many of the phenotypes observed in genetically engineered mouse models [500]. Within the field of ALS, several studies have examined the modifying effect of genetic background on the phenotype of *tgSOD1$^{G93A}$* transgenic mice [501–504].

In the largest such study, which reported survival analysis of *tgSOD1$^{G93A}$* mice of 20 different background strains, transgenic mice with a C57BL/6J background survived the longest, suggesting the C57BL/6J genetic background is somewhat protective against SOD1-ALS [502]. All other experiments presented in this thesis were conducted on mice with a C57BL/6J (N4) background; however, because of the protective effect observed in this strain, we sought to determine whether crossing *hSOD1$^{A4V}$* onto a more vulnerable genetic background would reveal subtle ALS phenotypes.

Heiman-Patterson and colleagues found that *tgSOD1$^{G93A}$* mice with a 129 background displayed rapid disease progression, similar to that observed in B6 x SJL (F1) – the original background strain used for phenotypic analysis of *tgSOD1$^{G93A}$* mice. Several other strains exhibited even more rapid disease progression than 129 – for example SJL, FVB, and ALR [502] – however, we opted to use 129 due to the strain's amenability to motor and behavioural phenotyping tests and lack of confounding health concerns.

Notably, there is extensive genetic variability among 129 substrains [505,506]. Previous genetic background comparison studies have used 129S1/SvImJ [502] and 129S2/SvHsd [504,507–509], whereas the 129 substrain used in the present work was the related steel-derived substrain 129S9/SvEvH. *hSOD1$^{A4V}$* mice were backcrossed for two generations onto a 129S9/SvEvH background to generate the *hSOD1$^{A4V}$-129* strain. These mice are considered to have a mixed C57BL/6J x 129S1/SvImJ background.

#### *4.3.7.1 Motor and behavioural phenotyping of hSOD1$^{A4V}$-129 mice*

We performed motor and behavioural phenotyping of *hSOD1$^{A4V}$-129* mice at 12 months old.

#### 4.3.7.1.1 Marble burying

There was no change in marble burying activity in *hSOD1$^{A4V}$-129* mice compared to littermate controls, as measured by the proportion of marbles which were fully buried (Figure 4-27).

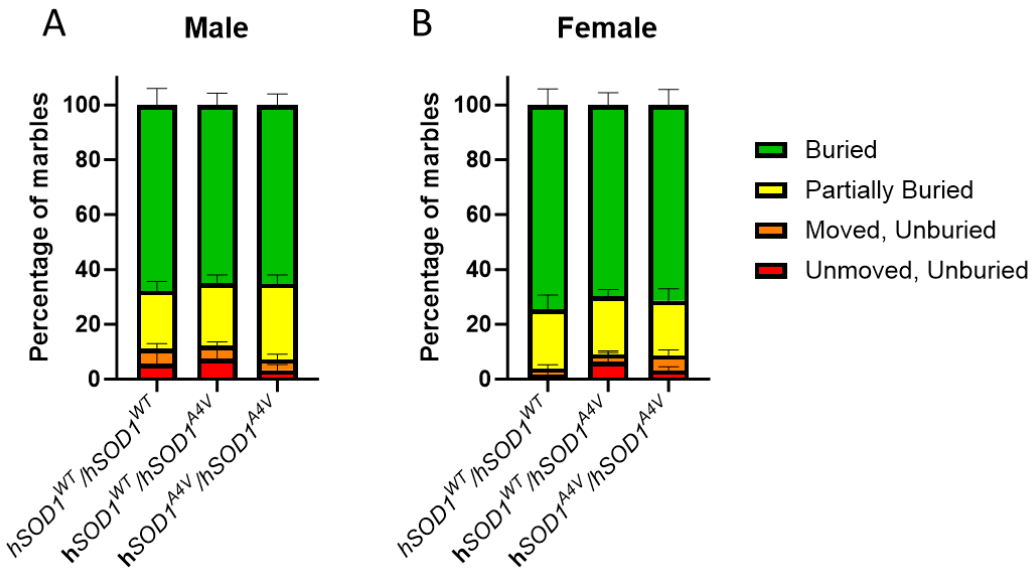



**Figure 4-27 | Marble burying data from *hSOD1$^{A4V}$-129* mice.** Marble burying data from 12-month-old **(A)** male (n=17-21 per genotype) and **(B)** female (n=14-21 per genotype) *hSOD1$^{A4V}$-129* mice. Mean ± SEM.

#### 4.3.7.1.2 Open field

There was no change in locomotor activity or time spent in the periphery in *hSOD1$^{A4V}$-129* mice compared to littermate controls (Figure 4-28). *hSOD1$^{A4V}$* mice on a 129 background displayed lower activity and spent more time in the periphery of the open field than mice on a C57BL/6J background, which is indicative of higher baseline anxiety in *hSOD1$^{A4V}$* mice on a 129 background.

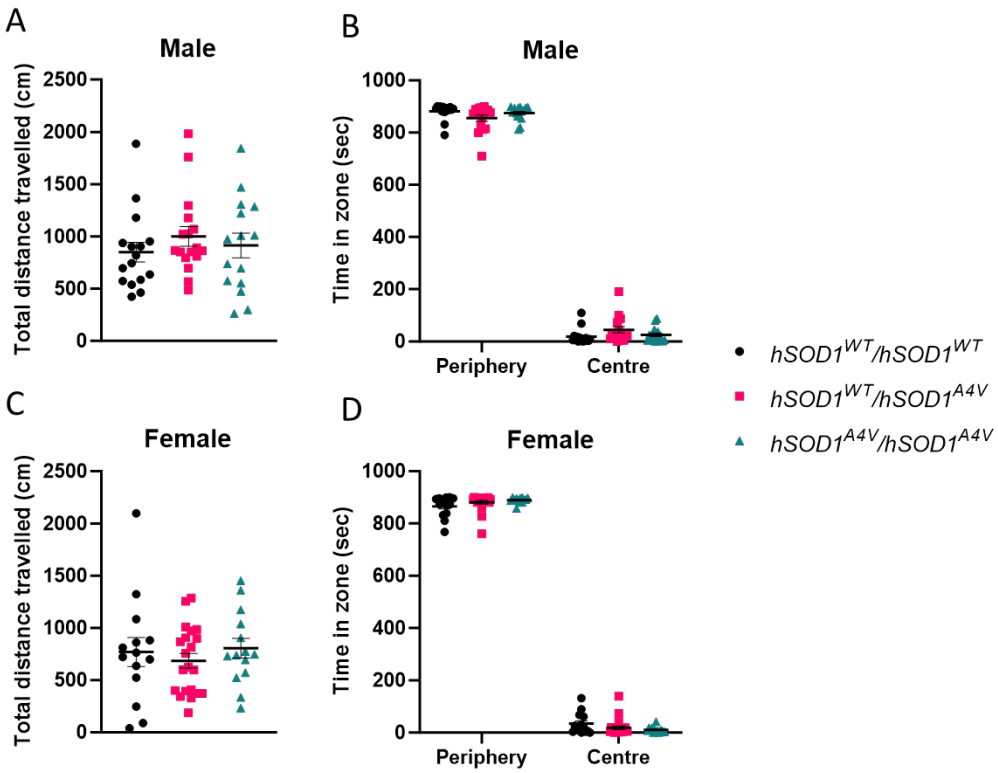



**Figure 4-28 | Open field test data from *hSOD1$^{A4V}$-129* mice.** Open field test data from 12-month-old **(A, B)** male (n=17-21 per genotype) and **(C, D)** female (n=14-21 per genotype) *hSOD1$^{A4V}$-129* mice. Mean ± SEM. One-way ANOVA with Tukey's post hoc test.

#### 4.3.7.1.3 Grip strength

There was no significant effect of genotype on grip strength in 12-month-old *hSOD1$^{A4V}$-129* mice (Figure 4-29 A, B). The grip strength values in Figure 4-29 were corrected for bodyweight; analysis of the uncorrected grip strength values also found no significant effect of genotype (data not shown).

#### 4.3.7.1.4 Locotronic

There was no significant effect of genotype on the number of leg errors or leg error time in the locotronic test in *hSOD1$^{A4V}$-129* mice (Figure 4-29 C, D).

#### 4.3.7.1.5 Gait analysis

Gait in freely moving *hSOD1$^{A4V}$-129* mice was assessed across 13 measurement parameters. Of the 13 gait parameters analysed, only two showed significant differences between the genotypes; front limb step length was significantly increased in male heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice

compared to *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ controls (Figure 4-29 E), and hind limb stride width was significantly decreased in female homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice compared to heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ littermates (Figure 4-29 F). A description of all gait parameters can be found in section 2.3.2.5. No differences were observed between *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ and *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ littermate controls in any parameter.

The reported significance values have not been corrected for multiple testing and therefore should only be considered exploratory. No comparisons met the threshold for significance once the Bonferroni correction was applied.

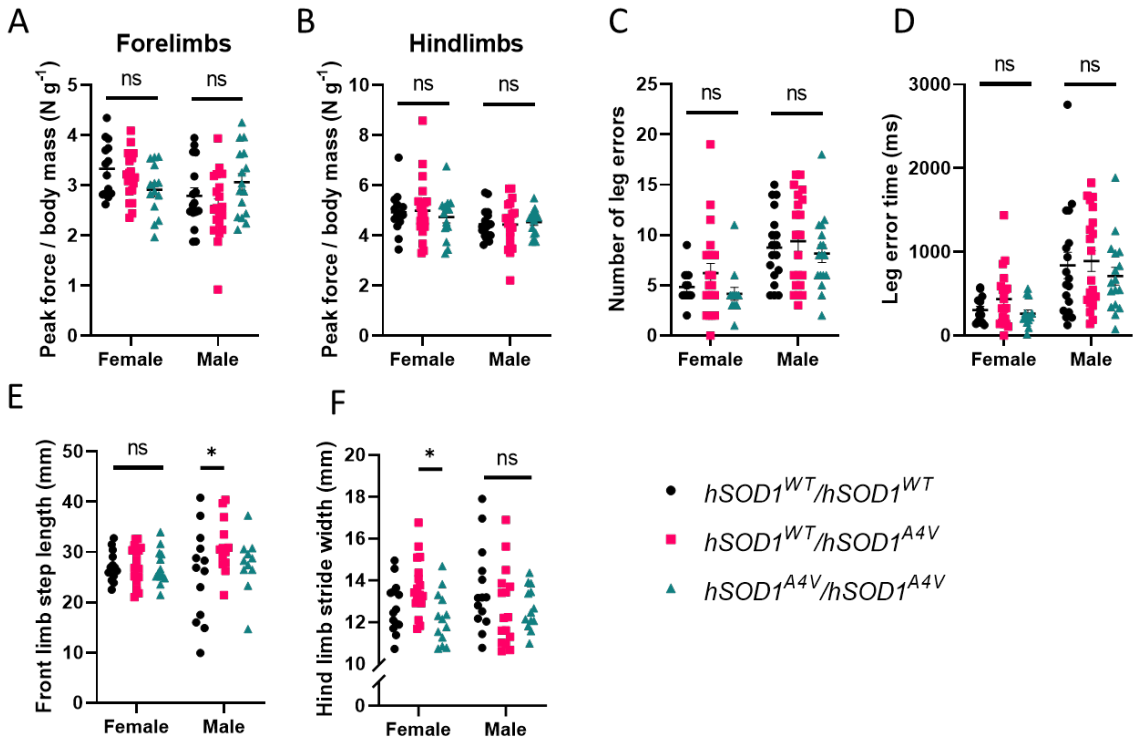



**Figure 4-29 | Motor function testing in *hSOD1*$^{A4V}$*-129* mice. (A, B)** Grip strength normalised to body weight, **(C, D)** locotronic data and **(E, F)** Mousewalker gait analysis data from 12-month-old *hSOD1*$^{A4V}$*-129* mice. The two gait analysis parameters shown were the only ones found to be significant out of 16 parameters analysed. n=17-21 per genotype for male animals; n=14-21 per genotype for female animals. One-way ANOVA with Tukey's post hoc test. Mean ± SEM. *$P < 0.05$. ns, not significant.

#### 4.3.7.1.6 Automated wheel running

Interestingly, homozygous *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ on a 129 background were significantly more active than *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ and *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ littermates in the automated wheel running test (Figure 4-30). *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice displayed a higher maximum speed ($p = 0.022$), ran further ($p = 0.010$) and spent more time on the running wheel ($p = 0.009$)(Figure 4-30 A-C). One important caveat is that most animals had very low running wheel activity; there was significant variability within the *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ group and the observed increase in activity was driven by a subset of animals (Figure 4-30 E). Running wheel activity was substantially lower in *hSOD1*$^{A4V}$ on a 129 background compared to a C57BL/6J background (Figure 4-16).

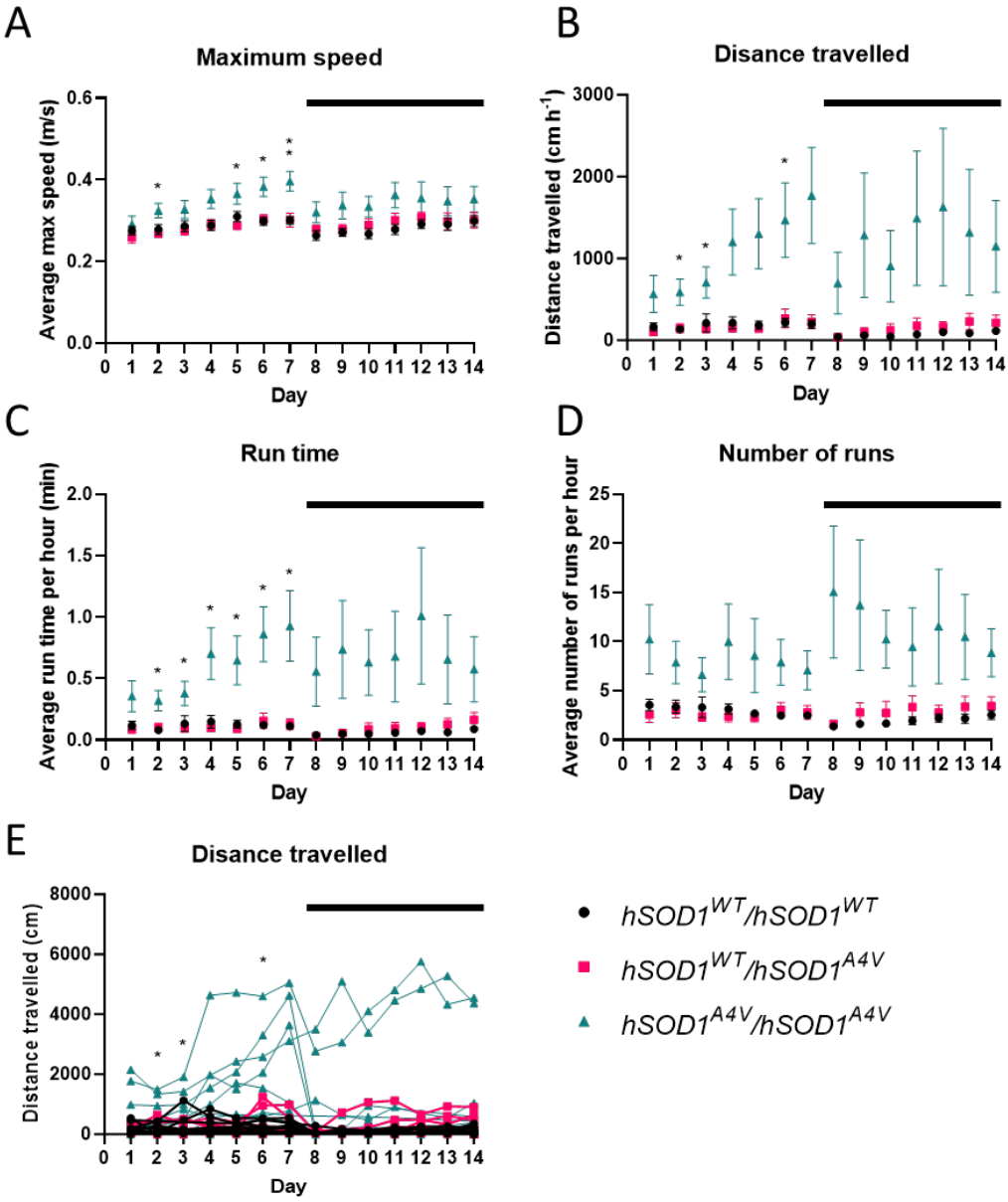



**Figure 4-30 | Voluntary wheel running data from *hSOD1$^{A4V}$-129* mice. (A-D)** Voluntary wheel running data from 13-month-old female *hSOD1$^{A4V}$-129* mice (n=12-13 per genotype) displayed as group mean ± SEM. **(E)** Distance travelled by individual animals. Black bar indicates days using a complex wheel. Mixed effects analysis followed by the Tukey multiple comparisons test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. *$P < 0.05$. **P < 0.01.

### *4.3.7.2 Metabolic phenotyping of hSOD1$^{A4V}$-129 mice*

#### 4.3.7.2.1 Bodyweight

To assess the impact of the A4V mutation on bodyweight, *hSOD1$^{A4V}$-129* animals were weighed monthly from 1 to 16 months of age. There was no effect of genotype on bodyweight in male or female *hSOD1$^{A4V}$-129* animals (Figure 4-31). This is in contrast to *hSOD1$^{A4V}$* mice on a C57BL/6J

background, where female *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice were significantly lighter than littermate controls (Figure 4-17 B).

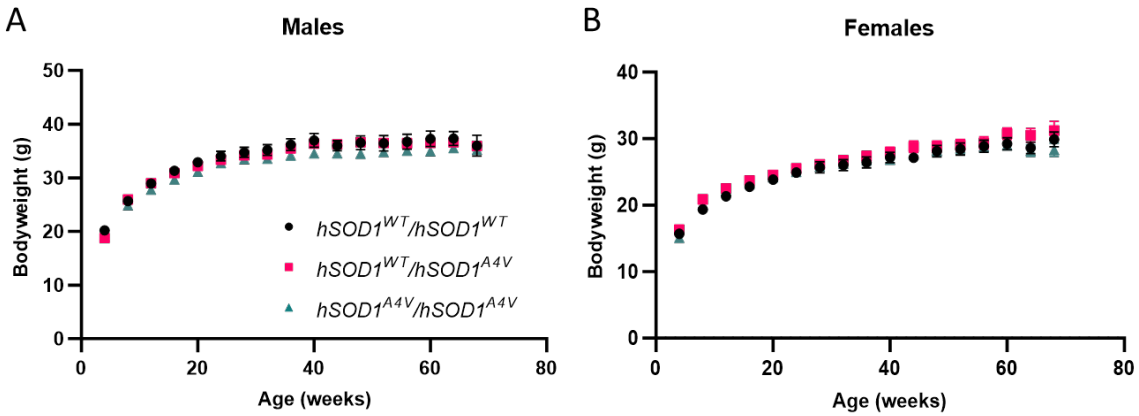



**Figure 4-31 | Longitudinal bodyweight data from *hSOD1*$^{A4V}$*-129* mice.** Monthly bodyweight measurement data from **(A)** male (n=17-34 per genotype) and **(B)** female (n=16-35 per genotype) *hSOD1*$^{A4V}$*-129* mice. Mixed effects analysis followed by the Tukey multiple comparisons test. Mean ± SEM.

#### 4.3.7.2.2 ECHO-MRI

There was no significant effect of genotype on whole body fat mass or lean mass in *hSOD1*$^{A4V}$*-129* mice (Figure 4-32).

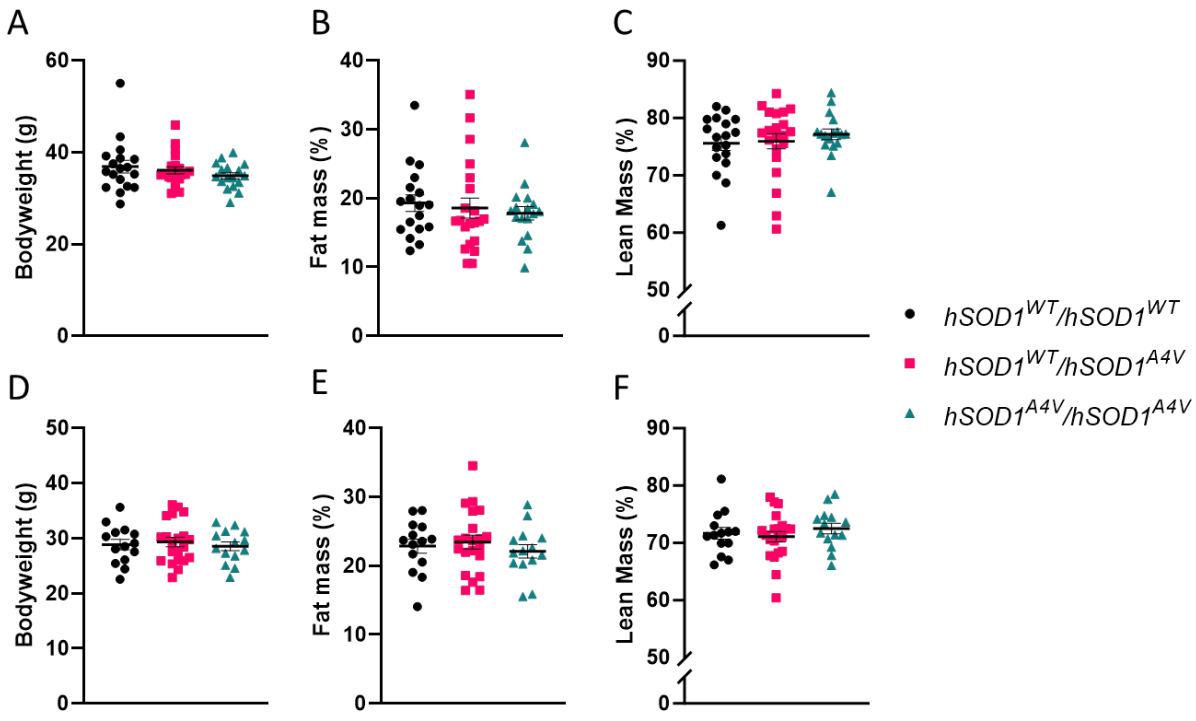



**Figure 4-32 | Echo MRI data from *hSOD1$^{A4V}$-129* mice.** Bodyweight, fat mass and lean mass in 12-month-old **(A-C)** male (n=17-21 per genotype) and **(D-F)** female (n=14-21 per genotype) *hSOD1$^{A4V}$-129* animals. Percentage fat/lean mass is expressed as a percentage of total bodyweight. Mean ± SEM. One-way ANOVA with Tukey's post hoc test. *$P < 0.05$.

#### 4.3.7.2.3 Calorimetric cages

Calorimetric cages (TSE Systems) were used to examine metabolism in 14-month-old female *hSOD1$^{A4V}$-129* mice (Figure 4-33). Two-way repeated measures ANOVA tests were performed to examine the effect of genotype and time on heat production, RER and locomotor activity. Simple main effects analysis revealed a significant effect of genotype on heat production ($p = 0.011$; Figure 4-33 A). Multiple pairwise comparisons conducted at each time found that the changes were driven by lower heat production in *hSOD1$^{WT}$/hSOD1$^{A4V}$* and *hSOD1$^{A4V}$/hSOD1$^{A4V}$* animals compared to *hSOD1$^{WT}$/hSOD1$^{WT}$* controls. Simple main effects analysis found no effect of genotype on RER or locomotor activity (Figure 4-33 B, C).

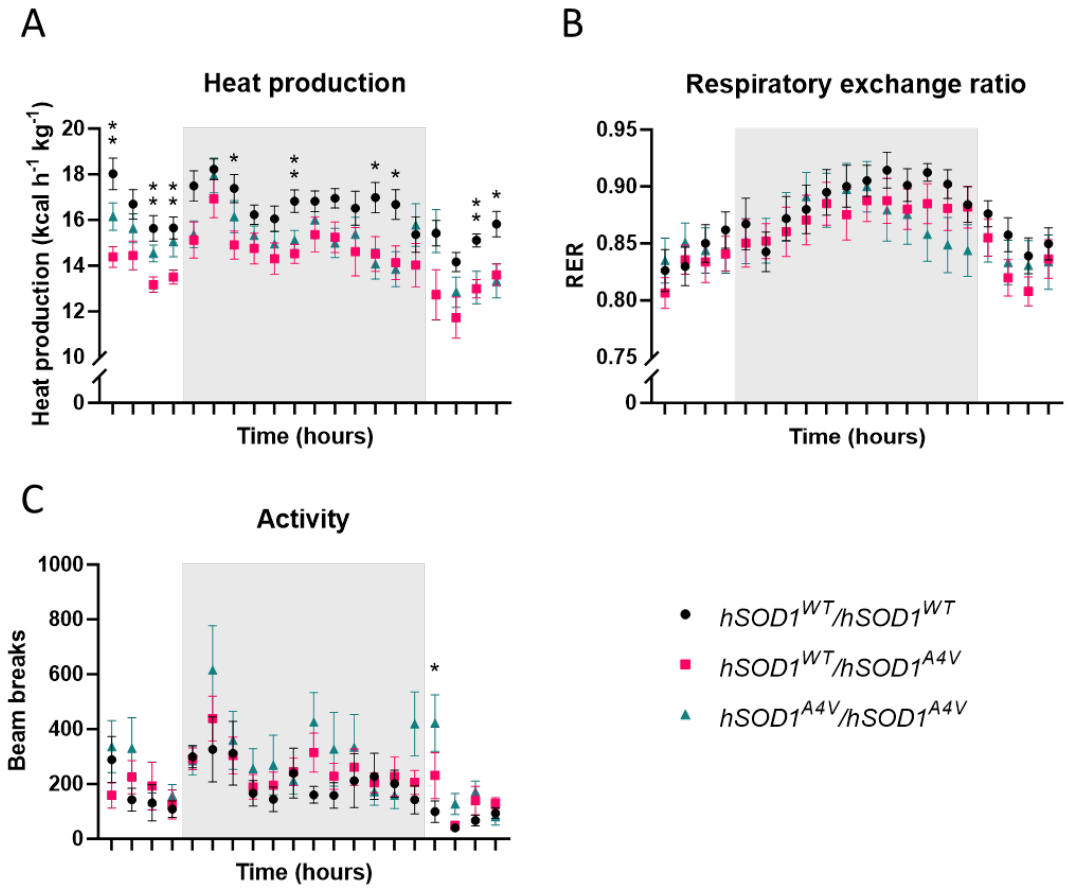



**Figure 4-33 | Calorimetric cage data from *hSOD1$^{A4V}$-129* mice**. Data from 14-month-old female *hSOD1$^{A4V}$-129* mice housed in Phenomaster calorimetric cages for 24hrs. **(A)** Heat production corrected for bodyweight. **(B)** Respiratory exchange ratio (RER). **(C)** Locomotor activity. Grey box indicates dark phase (19:00-7:00). n=9-10 animals per genotype. Two-way repeated measures ANOVA followed by the Tukey multiple comparisons test. Reported significance values are from the Tukey multiple comparisons test. Main effects are reported in the thesis body. Mean ± SEM. *$P < 0.05$. **P < 0.01.

#### 4.3.7.2.4 IPGTT

Differences in glucose metabolism were investigated in 14-month-old male *hSOD1$^{A4V}$-129* mice using IPGTT (Figure 4-34). Two-way repeated measures ANOVA, with genotype and sampling time after glucose infusion as factors, found no effect of genotype on blood glucose concentration. Additionally, the area under the curve (AUC) across all sampling points did not differ between the genotypes. These results contrast with findings from 14-month-old male *hSOD1$^{A4V}$* mice on a C57BL/6J background, which showed an impaired early response to glucose (Figure 4-20).

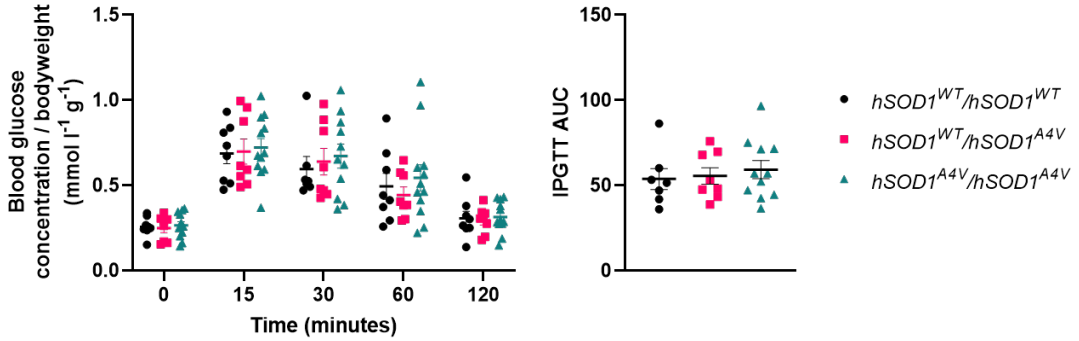



**Figure 4-34 | Intraperitoneal glucose tolerance testing (IPGTT) in *hSOD1$^{A4V}$-129* mice.** IPGTT data from 14-month-old male *hSOD1$^{A4V}$-129* mice (n=7-12 per genotype). Changes in blood glucose over time were analysed using two-way repeated-measures ANOVA followed by Tukey's test. Area under the curve (AUC) was analysed using one-way ANOVA followed by Tukey's test. Mean ± SEM.

## 4.3.8 hSOD1-A4V Histological characterisation

### *4.3.8.1 hSOD1-A4V spinal motor neuron counts*

Loss of spinal motor neurons is a hallmark of ALS. We counted the number of motor neurons in the sciatic pool of 24-month-old male *hSOD1$^{A4V}$* animals. We did not observe any genotype-specific difference in the number of motor neurons in the L3-L5 region of the spinal cord (Figure 4-35), suggesting that the *SOD1-A4V* mutation does not detrimentally impact motor neuron survival in *hSOD1$^{A4V}$* mice.

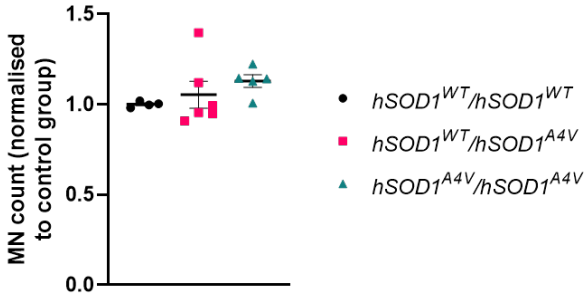



**Figure 4-35 | Spinal motor neuron counts from *hSOD1$^{A4V}$* mice.** Quantification of motor neurons within the sciatic pool of the lumbar spinal cord (L3-L5) of 24-month-old male *hSOD1$^{A4V}$* animals. n=4-6 animals per genotype. Mean ± SEM.

#### *4.3.8.2 hSOD1-A4V NMJ denervation*

Loss of functional neuromuscular junctions (NMJs) is among the earliest phenotypes observed in mouse models of ALS, and is apparent before the loss of motor neuron cell bodies [510]. NMJ denervation occurs primarily in "fast twitch" muscles (e.g. TA) – containing a higher proportion of type II muscle fibers – with "slow twitch" muscles (e.g. Soleus) – containing more a higher proportion of type I muscle fibers – being relatively spared [384,511]. We assessed NMJ innervation longitudinally in lumbrical, TA and soleus muscles from male *hSOD1$^{A4V}$* mice (Figure 4-36). We also examined NMJ innervation in the lumbrical muscles of 27-month-old female *hSOD1$^{A4V}$* mice (Figure 4-36 A).

Different methodologies were used to visualise the axons and presynaptic termini in the TA and soleus muscles at different ages; in muscles from 6-month-old and 18-month-old animals fluorescently labelled tetanus toxin heavy chain ($H_CT$), injected into the muscle of interest, was used for visualisation, whereas antibodies against synaptophysin/neurofilament were used in muscles from 24-month-old animals. For lumbrical muscles, antibodies against synaptophysin/neurofilament were used in all samples.

In male animals, for each tissue, two-way repeated measures ANOVA tests were used to investigate the effect of genotype and age on NMJ denervation. In lumbrical muscles (Figure 4-36 A), simple main effects analysis revealed a significant effect of age ($p = 0.003$) but no effect of genotype on NMJ denervation. In TA muscle (Figure 4-36 B), simple main effects analysis revealed a significant effect of age ($p = <0.001$) on NMJ denervation, and an interaction between age and genotype ($p = 0.044$). Multiple pairwise comparisons conducted at each age found that 18-month-old homozygous *hSOD1$^{A4V}$/hSOD1$^{A4V}$* and heterozygous *hSOD1$^{WT}$/hSOD1$^{A4V}$* mice displayed greater NMJ denervation compared to age-matched *hSOD1$^{WT}$/hSOD1$^{WT}$* littermate controls. Interestingly this effect was not observed in 24-month-old animals, which could be explained by: differences in methodology between 18-month-old and 24-month-old animals; increased denervation in *hSOD1$^{WT}$/hSOD1$^{WT}$*

control animals; the complete disappearance of denervated NMJ endplates in older *hSOD1*$^{A4V}$ mice – this could obscure an NMJ denervation phenotype because denervated NMJs were assessed as a percentage of total NMJs, an issue also observed in studies of a mouse model of SMA [512]; or a genuine recovery of NMJ innervation in older mice. In soleus muscle, simple main effects analysis found no effect of age or genotype on NMJ denervation (Figure 4-36 C).

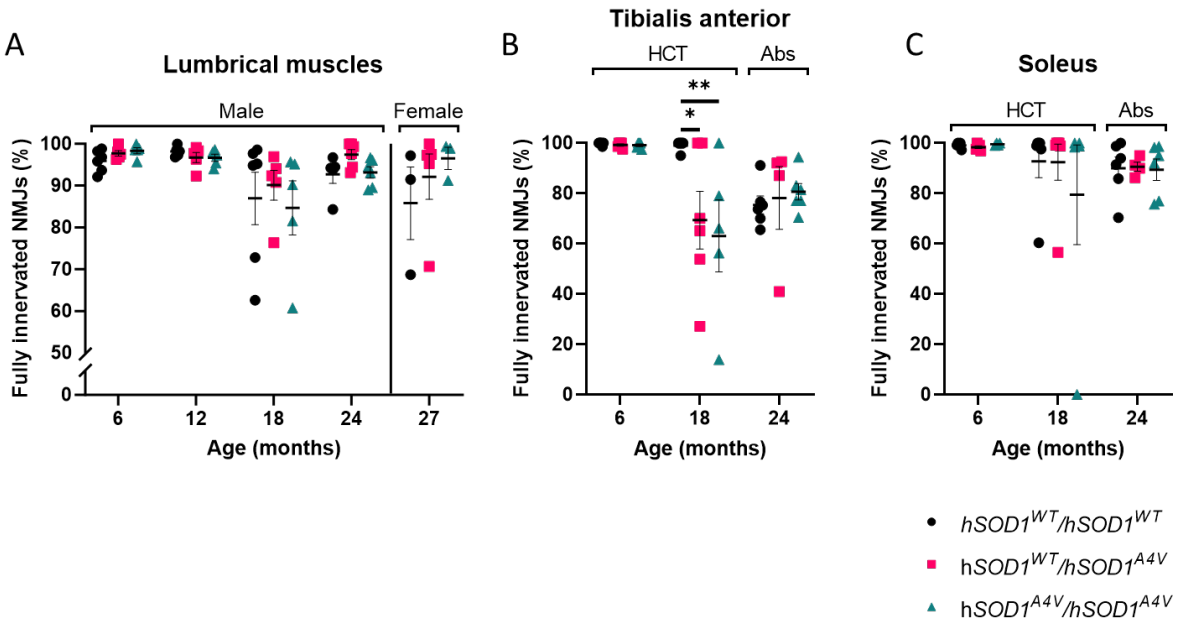



**Figure 4-36 | Analysis of neuromuscular junction (NMJ) denervation in *hSOD1*$^{A4V}$ mice. (A)** Quantification of fully innervated NMJs in lumbrical muscles from male (n=5-6 animals per genotype per time point) and female (n=3-5 animals per genotype) *hSOD1*$^{A4V}$ mice. The axons and presynaptic termini were visualised using antibodies against synaptophysin/neurofilament. **(B)** Quantification of fully innervated NMJs in tibialis anterior muscle from male *hSOD1*$^{A4V}$ mice (n=4-6 animals per genotype per time point). **(C)** Quantification of fully innervated NMJs in soleus muscle from male *hSOD1*$^{A4V}$ mice (n=4-6 animals per genotype per time point). **(B, C)** HCT and Abs indicate the method used for visualization of the axon and presynaptic terminal. 6-month-old and 18-month-old animals had fluorescently labeled tetanus toxin heavy chain (HCT) injected into the muscle of interest, which is internalized by peripheral nerve terminals and trafficked in the retrograde direction along the axon. 24-month-old animals had the axon and presynaptic terminal visualized using antibodies against synaptophysin/neurofilament. Two-way ANOVA followed by the Tukey multiple comparisons test. Reported significance values are from Tukey's test. Main effects are reported in the thesis body. Mean ± SEM. *$P$ < 0.05. **P < 0.01. HCT, tetanus toxin heavy chain; Abs, antibodies. Representative images of NMJ staining can be found in appendix 8.4.

#### *4.3.8.3 hSOD1-A4V muscle physiology*

Progressive muscle atrophy is a key phenotype in ALS, and preferentially affects fast twitch muscles such as the TA and extensor digitorum longus (EDL) [440]. Muscle atrophy is accompanied by fibre type switching from type 2 to type 1 fibres as neurons innervating type 2 fibres are lost [385].

##### 4.3.8.3.1 Tibialis anterior myofibre type and size

To investigate the impact of the *SOD1-A4V* mutation on muscle atrophy we examined myofiber type, size and nuclear positioning in the TA muscle of female *hSOD1$^{A4V}$* mice (Figure 4-37).

We found no evidence of fibre type switching in the TA muscle of *hSOD1$^{A4V}$* mice as the proportion of type I, type IIC and type II fibres did not differ between the genotypes at 6 or 27 months old (Figure 4-37 A-E). We also did not observe any difference in myofiber size or the proportion of centrally nucleated fibres in the TA muscle of 27-month-old *hSOD1$^{A4V}$* mice (Figure 4-37 F-H).

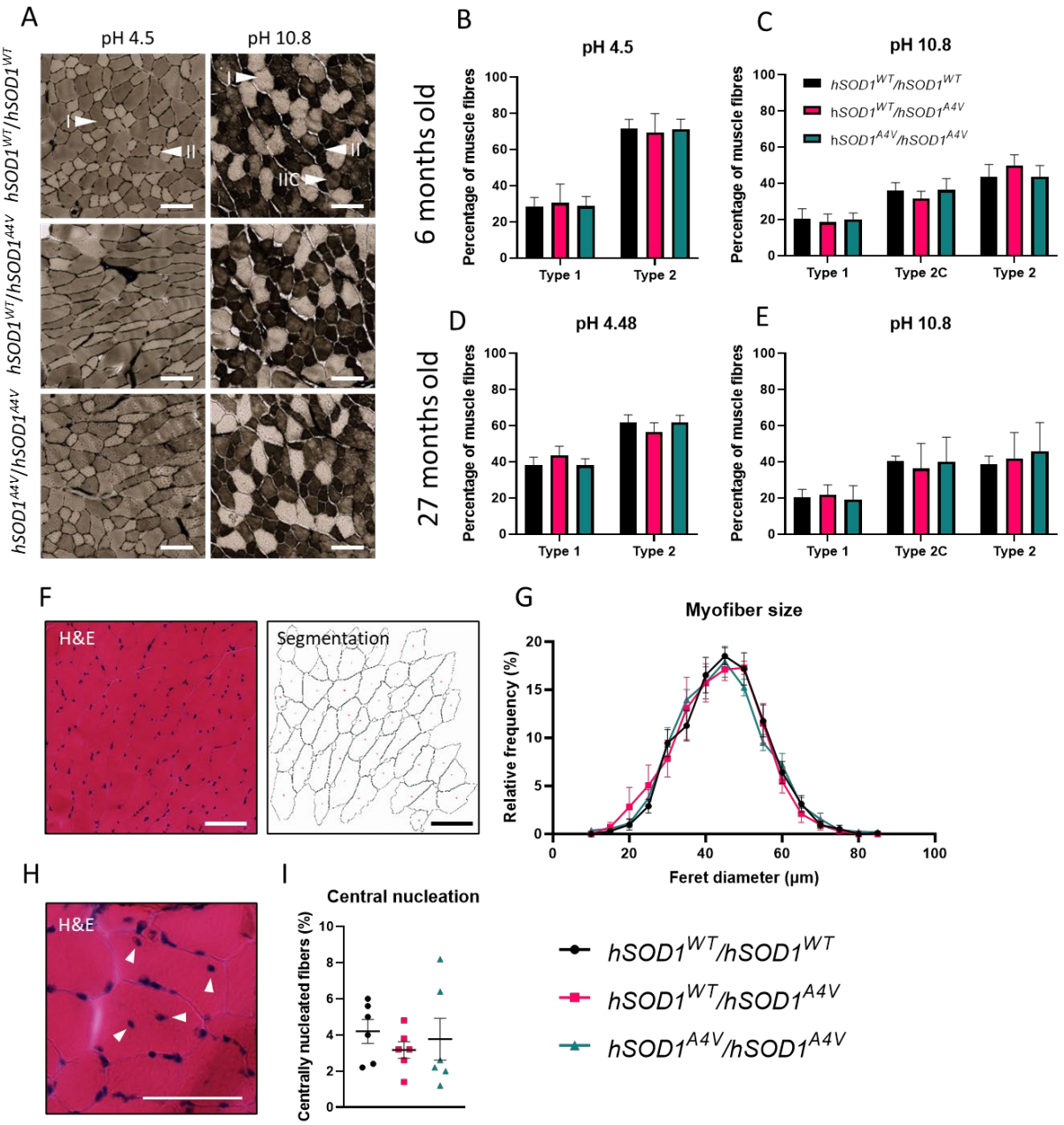



**Figure 4-37 | Analysis of tibalis anterior myofibers in *hSOD1$^{A4V}$* mice. (A)** Representative images of myosin-ATPase staining in tibialis anterior (TA) muscle from 6-month-old female *hSOD1$^{A4V}$* mice. Arrowheads indicate type I, type II and type IIC muscle fibers. Scale bar = 100µm. **(B, C)** Quantification of myofiber type in TA muscle from 6-month-old female *hSOD1$^{A4V}$* mice (n=5 per genotype). **(D, E)** Quantification of myofiber type in TA muscle from 27-month-old female *hSOD1$^{A4V}$* mice (n=5-7 per genotype). **(F)** Representative image of a H&E stained TA muscle section and the corresponding segmented image used for Feret diameter determination. Scale bar = 100µm. **(G)** Quantification of myofiber size in TA muscle from 27-month-old female *hSOD1$^{A4V}$* mice (n=5 per genotype), measured by the minimum feret diameter. **(H)** Example of centrally nucleated myofibers. Arrowheads indicate centralised nuclei. Scale bar = 50µm. **(I)** Percentage of centrally nucleated myofibers in TA muscle from 27-month-old female *hSOD1$^{A4V}$* mice (n=6 per genotype). **(B-E, I)** One-way ANOVA with Tukey's multiple comparisons test. **(G)** Mixed effects analysis with Tukey's post hoc test. Mean ± SEM shown in all analyses.

#### 4.3.8.3.2 Isometric muscle tension measurements

Muscle atrophy is associated with reduced contractile force. Hence, we performed isometric tension measurements on the TA and EDL muscles of 27-month-old female *hSOD1$^{A4V}$* mice. We did not observe any differences in the speed or force of contraction in the TA or EDL muscles of *hSOD1$^{A4V}$/hSOD1$^{A4V}$* mice compared to *hSOD1$^{WT}$/hSOD1$^{WT}$* controls. Furthermore, we did not detect any change in the number of functional motor units in the EDL muscle of *hSOD1$^{A4V}$/hSOD1$^{A4V}$* mice compared to *hSOD1$^{WT}$/hSOD1$^{WT}$* controls.

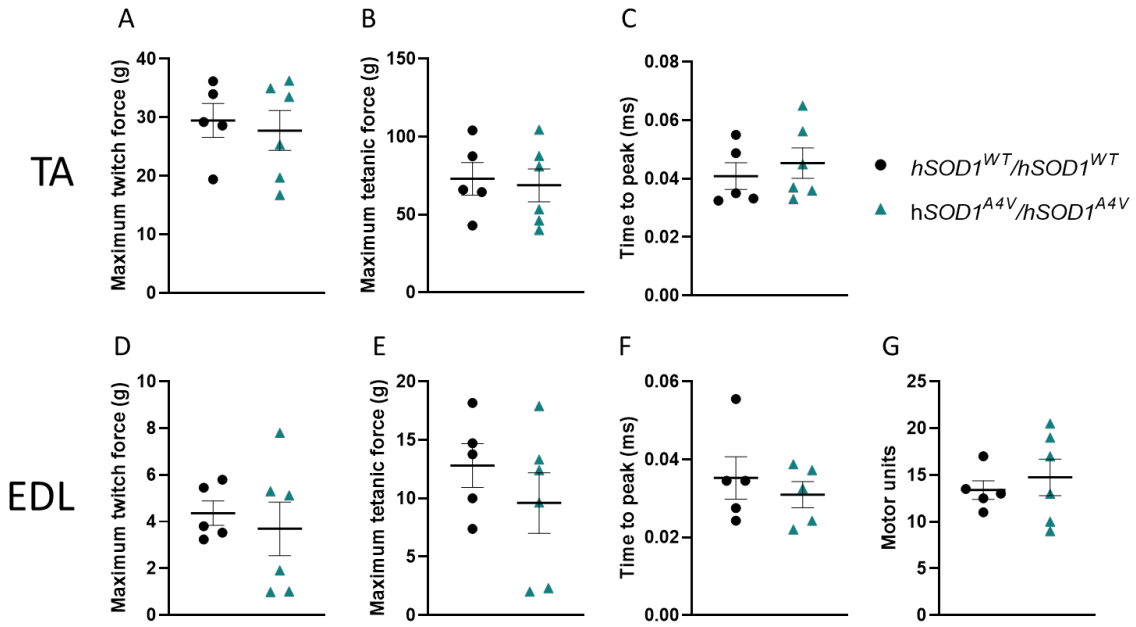



**Figure 4-38 | Muscle physiology in *hSOD1$^{A4V}$* mice.** Isometric tension measurements from the tibialis anterior (TA) and extensor digitorum longus (EDL) muscles of 27-month-old female *hSOD1$^{A4V}$* mice (n=5 animals per genotype). **(A)** Maximum twitch force, **(B)** maximum tetanic force and **(C)** time to peak contraction in TA muscle. **(D)** Maximum twitch force, **(E)** maximum tetanic force, **(F)** time to peak contraction and **(G)** number of functional motor units in EDL muscle. Student's t test. Mean ± SEM.

#### *4.3.8.4 hSOD1-A4V immunohistochemistry*

To assess the levels of misfolded and aggregated SOD1 in the lumbar spinal cord of 24-month-old male *hSOD1$^{A4V}$* mice we performed immunohistochemistry using an antibody that selectively recognises misfolded and aggregated SOD1 (mAb 85.11 [417]; Figure 4-39 A). We did not observe any difference in staining intensity between the genotypes using a subjective visual rating scale (Figure

4-39 B). All *hSOD1*$^{A4V}$ samples had a staining intensity greater than a *hSOD1*$^{KO}$ negative control but substantially less than a P100 transgenic *tgSOD1*$^{G93A}$ positive control.

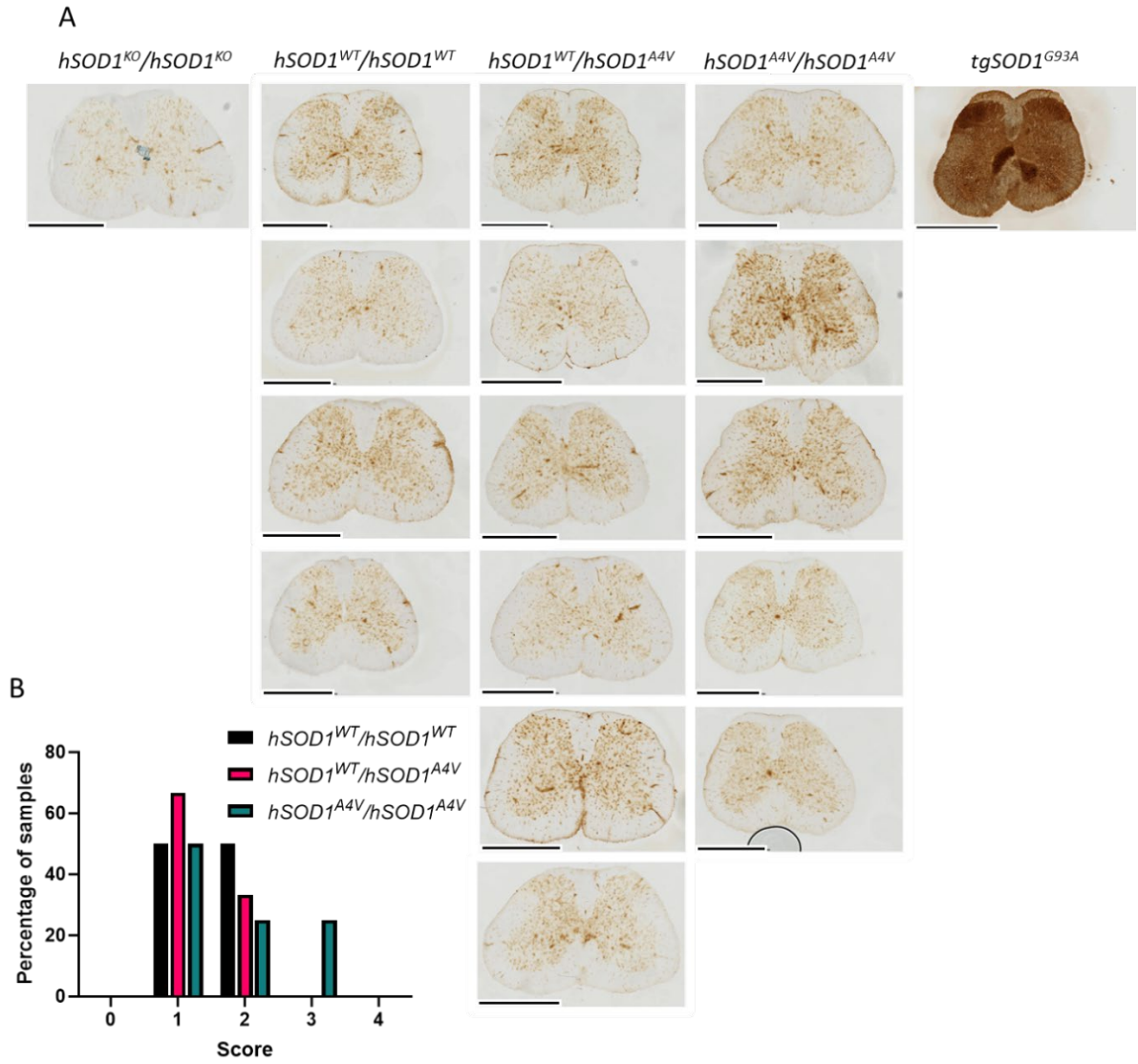



**Figure 4-39 | Misfolded/aggregated SOD1 immunohistochemistry. (A)** SOD1 IHC of lumbar spinal cord from 24-month-old male *hSOD1*$^{A4V}$ animals, and *hSOD1*$^{KO}$ and *tgSOD1*$^{G93A}$ controls. Each image is representative of staining from one animal. **(B)** Scoring of staining intensity in *hSOD1*$^{A4V}$ animals using a subjective rating scale from 0 (background staining) to 4 (strong staining). IHC was performed using an antibody that selectively recognises misfolded and aggregated SOD1 (mAb 85.11). Scale bar = 1 mm.

## 4.4 Discussion

### 4.4.1 SOD1 levels and activity in *hSOD1$^{A4V}$* mice

We observed reduced SOD1 levels and activity in *hSOD1$^{A4V}$* mice, which is consistent with results from SOD1-A4V patients [452,458]. We observed extremely low SOD1 activity in skeletal muscle of *hSOD1$^{A4V}$* mice, especially when compared to *mSod1*/*mSod1* mice. This was due to the compounded effects of lower SOD1 transcription in muscle (as seen in *hSOD1$^{WT}$* mice) plus the decreased stability of SOD1-A4V (vs SOD1-WT) in *hSOD1$^{A4V}$* mice. The minimum level of SOD1 required to prevent LOF phenotypes in mice is unclear, and it is possible that LOF phenotypes may be observed in *hSOD1$^{A4V}$* mice.

We did not observe SOD1 aggregation in the SC of *hSOD1$^{A4V}$* mice. However, SOD1 aggregation was only assessed in a qualitative manner by immunostaining of fresh-frozen 20 µm sections. A more sensitive ELISA assay to assess SOD1 misfolding is being performed by Chloe Williams at Umea University, the experiments are ongoing at the time of writing.

### 4.4.2 There is a lack of clear motor and behavioural phenotypes in *hSOD1$^{A4V}$* mice

The progressive development and worsening of motor symptoms is a hallmark of human ALS and a key phenotype in ALS mouse models. Most ALS mouse models that show phenotypes display deficits in grip strength, Rotarod and gait [73,378,513–515]. On the contrary, we did not observe any deficits in these tests in *hSOD1$^{A4V}$* mice, and actually saw improvements over baseline in some tests in male *hSOD1$^{WT}$*/*hSOD1$^{A4V}$* animals – although this was likely related to the observed hyperactivity in these animals, rather than a genuine increase in motor function.

We observed some differences between the genotypes in automated wheel running and gait analysis; however, the magnitude of changes observed was small and progression of phenotypes was not observed – although detection of phenotypes in aged animals was more challenging due to reduced n numbers.

There was an expected attrition of animal numbers due to non-ALS related illness. At the final 24 month time point the smallest group in the majority of tests contained 8 animals. With the within-group variances we observed, this gave us 80% power to detect a 35-40% difference in the group means in the grip strength and rotarod tests. Changes of this magnitude are seen in transgenic *tgSOD1$^{G93A}$* animals at 90-110 days old, when disease is already severe [173,515]. Therefore it is not wholly unexpected that no motor phenotype was detected in aged *hSOD1$^{A4V}$* mice, which are a substantially less severe model compared to *tgSOD1$^{G93A}$*.

### 4.4.3 Reduced TA NMJ functionality may be a phenotype of *hSOD1$^{A4V}$* mice

MN degeneration is a hallmark of ALS, and occurs concurrently with the onset of motor phenotypes in *tgSOD1$^{G93A}$* mice (Figure 1-2). Numerous studies have established that NMJ denervation occurs before the loss of MN cell bodies (reviewed in reference [510]), in support of a "dying back" hypothesis of ALS [516].

We did not observe loss of LMN cell bodies in aged *hSOD1$^{A4V}$* mice. However, we did observe NMJ denervation in the TA muscle of 18-month-old animals. TA NMJ denervation was not fully penetrant, with some animals showing marked denervation and others showing none. Intriguingly, this difference was not detected in 24-month-old animals.

The discrepancy between the 18- and 24-month-old TA NMJ data may be due to the different methodologies used for the two experiments. In 18mo mice, the neuronal processes and presynaptic NMJ were labelled using labelled tetanus heavy chain ($H_CT$) infused into the TA of living animals. $H_CT$ is only taken up by NMJs that are both intact and functional [517]. Whereas, in 24mo mice the neuronal processes and presynaptic NMJ were labelled using antibodies against synaptic vesicle glycoprotein 2A and neurofilament medium chain, which will label intact NMJs regardless of whether they are functional. It is possible that there is loss of NMJ functionality without denervation occurring in 18- and 24-month-old animals, but which is not detected using the antibody labelling approach.

It is also possible that the labelled $H_CT$ did not spread throughout the entire muscle, leaving some fibre bundles dark. This is unlikely as prior observations by Dr Tosolini suggest that $H_CT$ spreads widely throughout the injected area, even occasionally being observed in the contralateral hindlimb. This would also not explain why only *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ and *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice were affected.

It is possible that NMJ denervation was present in 24-month-old animals but was undetected. The antibody labelling approach produced immunofluorescence images with lower signal-to-noise ratios than the $H_CT$ labelling approach, particularly in older animals, meaning that the antibody labelling of some NMJs in 24-month-old animals were ambiguous. Signal-to-noise ratios were lowest in the centre of each fibre bundle where the staining was weakest. Poor staining may have contributed to the denervation observed in control *hSOD1*$^{WT}$/*hSOD1*$^{WT}$ animals, and in turn prevented detection of differences between the genotypes.

Another possible explanation is that motor endplates recorded as denervated in 18-month-old *hSOD1*$^{A4V}$ mice were completely lost in 24-month-old animals. As we only recorded denervated NMJs as a percentage of total NMJs, the disappearance of denervated endplates would result in an apparent recovery in NMJ innervation, as observed in our aged *hSOD1*$^{A4V}$ mice. A similar phenomenon has previously been reported in a mouse model of SMA [512]. It is extremely challenging to determine absolute NMJ numbers in the TA muscle as this requires imaging of every muscle fibre, which is impractical considering the large size of the muscle.

On balance, it is likely that NMJ denervation is not occurring in *hSOD1*$^{A4V}$ mice, however because of the technical difficulty involved in assessing NMJs in older animals this conclusion is not certain. This data raises the interesting possibility of reduced functionality of TA NMJs in *hSOD1*$^{A4V}$ mice without denervation. A recent study demonstrated that structural changes at the NMJ precede denervation [518]. NMJ structure and function are inextricably linked, hence it is likely that changes in NMJ function

also occur prior to NMJ denervation. The observed degeneration of TA NMJs but sparing of soleus NMJs is consistent with reports from other ALS mouse models [376,519,520].

### 4.4.4 Muscle pathology is absent in *hSOD1$^{A4V}$* mice

Our muscle physiology data suggest that no significant muscle atrophy in occurs in *hSOD1$^{A4V}$* mice. Although we cannot rule out atrophy of other muscles, we examined the TA and EDL – two of the most vulnerable muscles in ALS – and one would expect that if widespread muscle atrophy was occurring then these two muscles would be among those affected. No change in EDL motor units was observed, which would support a lack of MN degeneration.

### 4.4.5 Minimal transcriptional changes are detected in aging *hSOD1$^{A4V}$* mice

Very few DEGs were detected in our longitudinal RNA-seq analysis of *hSOD1$^{A4V}$* mice. Just three DEGs were detected across muscle and SC in 20-month-old animals. For contrast, 90-day-old presymptomatic *tgSOD1$^{G93A}$* mice have >1500 DEGs in the SC alone [463]. Counterintuitively, we observed more transcriptional changes in 6-month-old animals than in older animals. These changes are unlikely to result from SOD1 LOF, as there was very little overlap between transcriptional changes observed in 6-month-old *hSOD1$^{A4V}$* and *hSOD1$^{-/-}$* mice (data not shown).

Mild metabolic changes were observed in the SC of aged *hSOD1$^{A4V}$* animals, including two DEGs involved in lipid metabolism and an enrichment of GO terms related to metabolism. On the other hand, there was no significant enrichment of the “Fatty acid metabolism” hallmark gene set in GSEA.

MNs represent a tiny proportion of all cells in the lumbar SC. Bulk RNA-seq is sufficient to detect transcriptional changes in the SC of transgenic *tgSOD1$^{G93A}$* mice; however, as the phenotype in *hSOD1$^{A4V}$* mice is much less severe and any transcriptional changes likely to be subtle, it is likely that any transcriptional changes in MNs are masked by a lack of changes in other cell types. To uncover neuron-specific transcriptional changes we would need to do single-cell RNA-seq, or bulk RNA-seq of MNs isolated using laser capture microdissection as recently performed in knock-in *Sod1$^{G85R}$* mice [372].

### 4.4.6 Mild metabolic phenotypes are present in *hSOD1$^{A4V}$* mice

Several observations support a mild metabolic phenotype in *hSOD1$^{A4V}$* mice: 1. Female *hSOD1$^{A4V}$* mice are lighter than littermate controls and have a trending decrease in fat mass. 2. *hSOD1$^{A4V}$* mice have an impaired early response to glucose, indicating disturbed glucose metabolism. 3. *hSOD1$^{A4V}$* mice have changes in a number of plasma metabolites, including several lipids. 4. Aged *hSOD1$^{A4V}$* mice have decreased plasma GSSG and an increased ratio of GSH to GSSG. 5. Aged *hSOD1$^{A4V}$* mice have SC transcriptional changes related to metabolism.

Considering the lack of overt SOD1 GOF phenotypes, it is possible that metabolic changes in *hSOD1$^{A4V}$* mice may be due to SOD1 LOF. A few lines of evidence suggest this is the case:

1. Female *hSOD1$^{WT}$*/*hSOD1$^{WT}$* mice – the control group for experiments involving *hSOD1$^{A4V}$* mice – are significantly lighter than non-humanised *mSod1*/*mSod1* mice (Figure 3-16). *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* mice are lighter still, suggesting a compounding effect of both *SOD1* humanisation and the *SOD1-A4V* mutation on bodyweight. One common outcome resulting from both *SOD1* humanisation and the *SOD1-A4V* mutation was reduced levels of SOD1 protein. This suggests that reduced bodyweight in these animals may be a result of SOD1 loss of function. SOD1$^{-/-}$ mice also display reduced bodyweight [444,521,522].

2. Disturbed glucose metabolism has been observed in SOD1 knockout mice. SOD1$^{-/-}$ mice display impaired pancreatic β-cell function and reduced plasma insulin levels [444,523] as well as altered glycolysis and gluconeogenesis [524]. We observed impaired glucose tolerance in *hSOD1$^{A4V}$* mice; altered glucose tolerance has also been observed in SOD1$^{-/-}$ mice, although studies unexpectedly disagree on the direction of change [444,523].

3. SOD1$^{-/-}$ mice show disrupted lipid metabolism, with altered hepatic and plasma lipid profiles, and altered lipogenesis [524–526]. Visceral fat is also decreased in SOD1$^{-/-}$ mice, correlating with

lower bodyweight [521]. Similarly, we observed changes in plasma lipid profiles in *hSOD1$^{A4V}$* mice.

On the contrary, other evidence suggests that the metabolic changes in *hSOD1$^{A4V}$* mice could be due to SOD1 GOF: Several SOD1-ALS mouse models without SOD1 LOF display metabolic phenotypes, including reduced bodyweight, reduced fat mass, dyslipidaemia and reduced glucose utilisation (reviewed in Maksimovic et al., 2023).

It is important to note that, in addition to SOD1 GOF phenotypes, the majority of SOD1-ALS patients have some degree of SOD1 LOF, which seemingly has a more severe impact on neurological function in humans than in mice (see section 7.4). Therefore, both GOF and LOF phenotypes remain relevant for ALS disease modelling.

Our plasma metabolomics analysis was exploratory and, with the exception of xylulose, did not reveal any metabolites that were significantly changed at a stringent p value cutoff of padj<0.05. To replicate and expand on our findings, a more detailed examination of the metabolic changes in *hSOD1$^{A4V}$* mice is warranted. We have planned tissue metabolomics experiments to examine metabolites in the brain, spinal cord and liver of aged *hSOD1$^{A4V}$* mice, with the goal of uncovering tissue-specific metabolic dysregulation.

Reduced GSH reacts with ROS – being oxidised to GSSG in the process – protecting cells from oxidative damage [492]. The observed increase in the GSH-to-GSSG ratio in *hSOD1$^{A4V}$* mice was unexpected. This change was driven by a decrease in GSSG rather than an increase in GSH. The GSH-to-GSSG ratio is widely used as a marker of oxidative stress [156,492,496,497,527] and hence a lower GSH-to-GSSG ratio is indicative of reduced oxidative stress in *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$* animals; which is the opposite of what is seen in ALS patient blood [496] and CSF [497]. We did not examine changes in the expression/activity of glutathione reductase, which catalyses the reduction of GSSG to GSH, or glutathione peroxidase 1 (Gpx1), which catalyses the oxidation of GSH. Therefore it is difficult to say

whether the change in GSH-to-GSSG ratio represents decreased reactive oxidation of GSH, or increased reduction of GSSG to GSH. A few studies have shown a decrease in Gpx1 activity in SOD1$^{-/-}$ mice [495,528,529], which suggests that reduction in GSSG could be a SOD1 LOF phenotype related to reduced Gpx1 activity rather than a reduction in oxidative stress. A more thorough examination of the flux through this pathway, and oxidative stress in *hSOD1$^{A4V}$* mice in general, is necessary to better understand our observations.

Intriguingly, there was a significant effect of genetic background on metabolic phenotypes in *hSOD1$^{A4V}$* mice. In contrast to animals on a C57BL/6J background, *hSOD1$^{A4V}$* mice on a mixed C57BL/6J x 129S1/SvImJ background were not lighter and showed no change in fat mass or glucose tolerance. However, they did display decreased heat production without any change in activity or RER. The causes of these strain-specific differences remain unclear.

### 4.4.7 No change in mitochondrial function in SC cultures from *hSOD1$^{A4V}$* mice

Impairment of mitochondrial function has been implicated in the aetiology of ALS [498,499]. However, we did not observe any change in mitochondrial function in SC cultures derived from E13.5 *hSOD1$^{A4V}$* embryos.

One important caveat to our experiments is that these cultures represent a mix of spinal cord cells, where only a very small subset of cells will be MNs, and most cells will be other neuronal subtypes and glia. Our cultures were negative for the MN markers ChAT and Isl1 by immunostaining (data not shown). This presents a similar issue to bulk RNA-sequencing of SC; if changes in mitochondrial function are specific to MNs then any differences are likely to be masked by the overwhelming majority of cells that are unaffected. To avoid this issue we aimed to isolate MNs by immunopanning [530], however our attempts failed to yield viable cultures enriched in MNs (data not shown).

### 4.4.8 Conclusion

In summary, *hSOD1*$^{A4V}$ mice have a lack of overt SOD1 GOF phenotypes; however, they do have some mild metabolic phenotypes, which are more likely a result of SOD1 LOF than GOF. While the lack of GOF phenotypes is disappointing, it is not wholly unexpected due to the strong correlation between higher mutant SOD1 copy number and increased phenotype severity in transgenic models.

# Chapter 5: Creation and early characterisation of a genetically humanised mouse carrying mutant human SOD1$^{G93A}$

## 5.1 Introduction

The transgenic *tgSOD1$^{G93A}$* mouse, sometimes referred to as simply "the SOD1 mouse", is the most widely studied mouse model of ALS [338]. Thousands of studies using *tgSOD1$^{G93A}$* mice have greatly furthered our understanding of ALS, and have implicated wide-ranging mechanisms in the pathophysiology of the disease [531]. High-copy *tgSOD1$^{G93A}$* mice carry ~25 copies of the human transgene and express mutant SOD1$^{G93A}$ protein at ~17-fold over the level of endogenous mouse Sod1. As a result of mutant SOD1 overexpression, *tgSOD1$^{G93A}$* mice display a highly aggressive neurodegenerative phenotype with onset of motor symptoms at approximately 60 days of age [378,381,382], followed by progressive clinical weakness, paralysis, and death by 140 days [338,532].

The severity and rapid progression of disease in high-copy *tgSOD1$^{G93A}$* mice makes them an invaluable model for the study of mid-to-late-stage disease processes. However, early disease processes are more difficult to model in *tgSOD1$^{G93A}$* mice. Cellular pathology is observed from early postnatal ages, including vacuolization of mitochondria [532], fragmentation of Golgi [533], formation of insoluble SOD1 inclusions [66,387], endoplasmic reticulum (ER) stress [170], and muscle denervation [376,511].

It is unclear how well observations of disease in early postnatal *tgSOD1$^{G93A}$* mice relate to disease processes in ALS patients, who are typically older adults. This point is underscored by the repeated failure to translate promising preclinical findings from *tgSOD1$^{G93A}$* mice into successful clinical trials [534]. Furthermore, SOD1 loss of function, which is believed to play a role in the pathology of ALS [48], cannot be studied in the context of *tgSOD1$^{G93A}$* mice due to the presence of intact mouse *Sod1*. For these reasons we sought to apply our philosophy of creating humanised mouse models (described in section 1.5.2.2.3) to create a humanised *hSOD1$^{G93A}$* mouse line.

In this chapter I describe the molecular cloning of a targeting construct for humanisation of the mouse *Sod1* locus and introduction of a conditional *SOD1*$^{G93A}$ allele. I then describe the embryonic stem cell (ESC) targeting, performed by Dr Ben Davies and colleagues at The Wellcome Centre for Human Genetics, and the subsequent screening of putative positive clones. I finally describe the generation of the humanised *hSOD1*$^{G93A}$ mouse line by Dr Ben Davies and colleagues, and the quality control of the humanised *hSOD1*$^{G93A}$ allele to confirm the integrity of the targeted locus.

## 5.2 Aims

1. To create a humanised *hSOD1*$^{G93A}$ mouse model.
2. To perform quality control to ensure the integrity of the humanised *SOD1*$^{G93A}$ locus.

## 5.3 Results

### 5.3.1 Generation of a targeting construct for the creation of an hSOD1$^{G93A}$ mouse line

The duplicated region of *hSOD1* in our *hSOD1$^{WT}$* mice precluded the use of a simple CRISPR/Cas9-based strategy to introduce the SOD1-G93A mutation directly into *hSOD1$^{WT}$* ESCs or zygotes in a similar manner to SOD1-A4V. The corresponding G93 genomic sequence appears in both exon 4 and conditional exon 4' of the humanised *hSOD1* gene; therefore, CRISPR-mediated targeting of this sequence would cut the humanised gene twice, likely resulting in deletion or translocation of the sequence between the cut sites. Given this difficulty, we opted to create a new *hSOD1$^{G93A}$* mouse targeting construct consisting of the entire humanised *SOD1* gene – including the duplicated region, mouse homology arms (HAs), and the SOD1-G93A mutation. Sequence homology between the repeated sets of exons prevented the use of a recombineering based cloning strategy. Therefore, we used a cloning strategy based on CRISPR/Cas9 cleavage, PCR mutagenesis, and Gibson assembly.

The starting material for this cloning was a large 197 kb bacterial artificial chromosome (BAC) vector generated by Dr Devoy and Dr Rosie Bunton-Stasyshyn at UCL as part of a previous project (hSOD1-L126delTT BAC). hSOD1-L126delTT BAC was similar to the construct used for humanisation of *hSOD1$^{WT}$* mice (described in section 2.2.7.1) but harboured a 2 bp deletion in exon 5, a familial ALS mutation termed L126delTT [535]. The generation of this construct is described in detail in Dr Rosie Bunton-Stasyshyn's thesis. The goal of this cloning was to correct the L126delTT mutation in exon 5 and introduce the G93A point mutation into exon 4'.

Cloning with large BAC vectors is challenging, hence we first transferred a 30.4 kb region of the hSOD1-L126delTT BAC, consisting of the hSOD1-L126delTT cassette plus ~8 kb mouse HAs, into a pJAZZ linear vector backbone (Figure 5-1 A, step 1). We designed CRISPR/Cas9 sgRNA guides against mouse sequences flanking the humanised *SOD1* gene (Table 5-1); sgRNAs 1 and 2 (7.7 kb 5' of the humanised *SOD1* gene) and sgRNA 3 and 4 (8.8 kb 3' of the humanised *SOD1* gene). The hSOD1-L126delTT BAC was digested with Cas9 using sgRNAs 1, 2, 3 and 4. The 30.4 kb DNA fragment

containing the humanised SOD1 gene was identified by size selection on PFGE and isolated by gel purification. The fragment was ligated into a pre-digested pJAZZ-OC linear vector at a 1:1 vector:insert ratio using electroligase. The resulting 42.8 kb construct (hSOD1-L126delTT-pJAZZ) was transformed into Bigeasy TSA cells. Transformants were plated on LB plates with 12.5 µg/mL chloramphenicol, 30 µg/mL kanamycin, IPGT (1 mM) and x-gal (20 µg/mL) for blue-white screening. White colonies were selected and screened for the correct insert by PFGE (to determine total construct size) and sanger sequencing across the breakpoints (Table 5-2; Figure 5-1 B).

Next, hSOD1-L126delTT-pJAZZ was used as a template for PCR with extended mutagenic primers (Table 5-3; Figure 5-1 A, step 2). The resulting 3.5 kb PCR product spanned the region between *SOD1* exon 5 and *SOD1* exon 4'. The PCR product had the L126delTT mutation in *SOD1* exon 5 corrected (insertion of TT) and the G93A mutation introduced into *SOD1* exon 4' (G→C point mutation). The PCR product also included additional sequence outside of the corrected sites to be used as homologous sequence for Gibson assembly. The PCR product was identified by size selection on agarose gel electrophoresis and isolated by gel purification.

We then designed CRISPR/Cas9 sgRNA guides against *SOD1* exon 5 (sgRNA A) and *SOD1* exon 4' (sgRNA B), with cut sites less than 10 bp away from the site of PCR mutagenesis in each exon (Table 5-1). hSOD1-L126delTT-pJAZZ was digested with Cas9 plus sgRNA A and Cas9 plus sgRNA B in two separate reactions (Figure 5-1 A, steps 3 and 4). The resulting 26.8 kb fragment from Cas9 digestion with sgRNA A – containing the pJAZZ long arm and sequences 5' of *SOD1* exon 5 – and 12.5 kb fragment from Cas9 digestion with sgRNA B – containing the pJAZZ short arm and sequences 3' of *SOD1* exon 4' – were identified by size selection on PFGE and isolated by gel purification.

Finally, the DNA fragments isolated in steps 3 and 4 were combined with the PCR product from step 2 at an equimolar ratio in a 'Hifi' Gibson assembly reaction, resulting in a 42.8 kb assembled construct (hSOD1-G93A-pJAZZ; Figure 5-1 A, step 5). The Gibson assembly mix was used to transform Bigeasy TSA cells. Transformants were plated on LB plates with 12.5 µg/mL

chloramphenicol and 30 μg/mL kanamycin. Colonies were screened by colony PCR using a reverse primer that binds specifically to wild-type *SOD1* exon 5, not the L126delTT variant (Table 5-3). DNA from putative positive clones was sequenced using Sanger sequencing across the Gibson assembly breakpoints, which confirmed the introduction of G93A and correction of L126delTT (Table 5-2, Figure 5-1 C). The construct was further screened by XhoI digest followed by PFGE, which resulted in the expected fragment sizes (data not shown). The entire construct, including the vector backbone, was then sequenced using PacBio long-read sequencing (CD genomics) which confirmed the entire humanised sequence was as expected, and identical to the *hSOD1*$^{WT}$ allele with the exception of the G93A mutation in exon 4'.

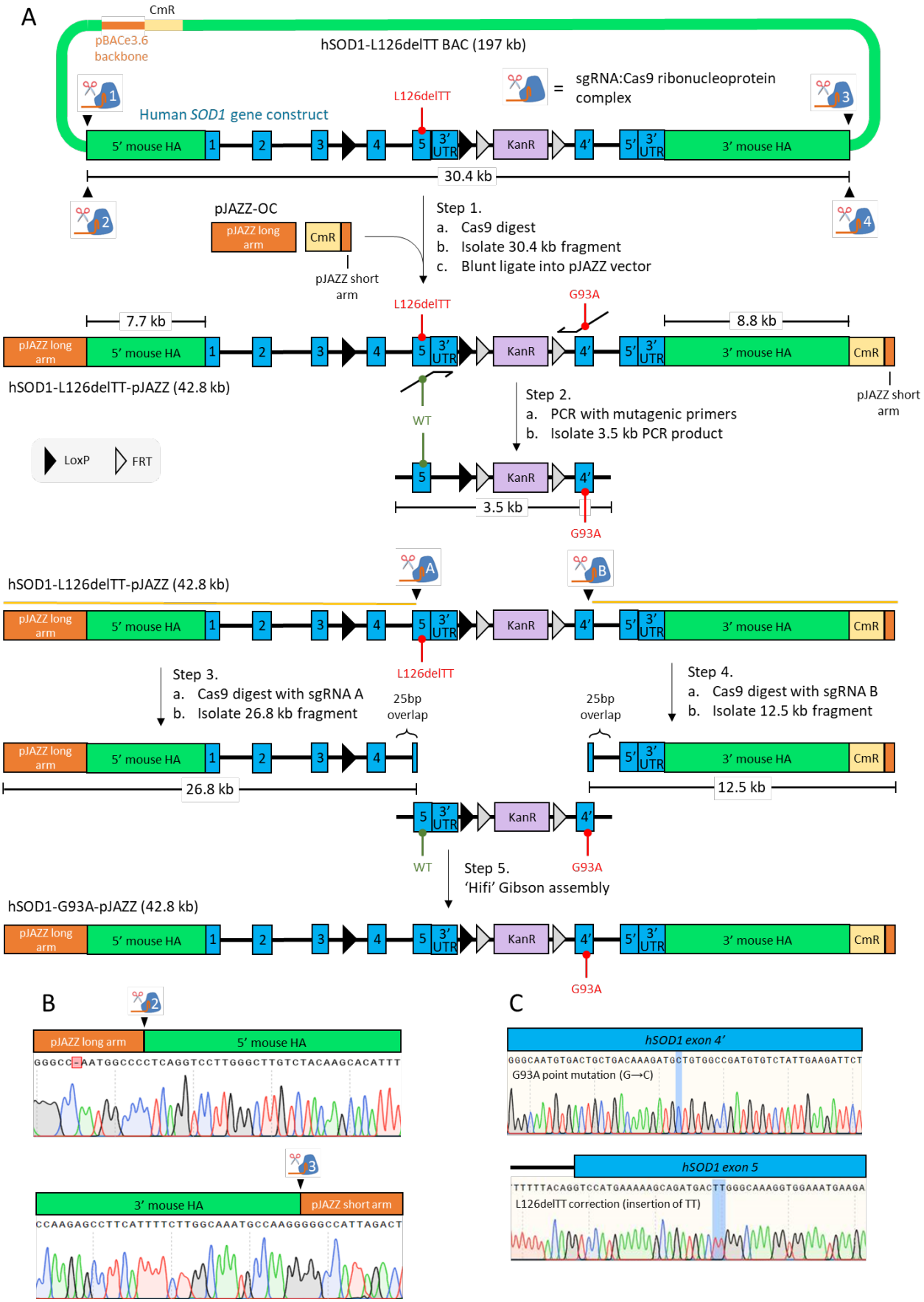

A
CmR
pBACe3.6 backbone
hSOD1-L126delTT BAC (197 kb)
= sgRNA:Cas9 ribonucleoprotein complex
Human SOD1 gene construct
L126delTT
5' mouse HA
3' UTR
KanR
3' mouse HA
30.4 kb
pJAZZ-OC
pJAZZ long arm
CmR
pJAZZ short arm
Step 1.
a. Cas9 digest
b. Isolate 30.4 kb fragment
c. Blunt ligate into pJAZZ vector
G93A
7.7 kb
8.8 kb
hSOD1-L126delTT-pJAZZ (42.8 kb)
WT
Step 2.
a. PCR with mutagenic primers
b. Isolate 3.5 kb PCR product
3.5 kb
LoxP
FRT
Step 3.
a. Cas9 digest with sgRNA A
b. Isolate 26.8 kb fragment
Step 4.
a. Cas9 digest with sgRNA B
b. Isolate 12.5 kb fragment
25bp overlap
26.8 kb
12.5 kb
Step 5.
'Hifi' Gibson assembly
hSOD1-G93A-pJAZZ (42.8 kb)
B
C
hSOD1 exon 4'
G93A point mutation (G→C)
hSOD1 exon 5
L126delTT correction (insertion of TT)

**Figure 5-1 | Molecular cloning of a *hSOD1$^{G93A}$* ESC targeting construct. (A)** Diagram of the cloning strategy used to generate the *hSOD1$^{G93A}$* targeting construct. Step 1 involved Cas9 digestion of the hSOD1-L126delTT BAC and ligation of the resulting 30.4 kb fragment into the pJAZZ linear vector backbone. Step 2 involved PCR using mutagenic primers and hSOD1-L126delTT-pJAZZ as a template to generate a PCR product containing wild-type hSOD1 exon 5 and a G93A point mutation in hSOD1 exon 4'. Steps 3 and 4 involved Cas9 single digests of hSOD1-L126delTT-pJAZZ with an sgRNA guide that binds 5' of the corrected L126TT site, or 3' of the G93A site, respectively. Step 5 involved a Gibson assembly reaction between the PCR product from step 2 and the hSOD1-L126delTT-pJAZZ fragments isolated in steps 3 and 4, which yielded the final hSOD1-G93A-pJAZZ targeting construct. Not drawn to scale. **(B)** Sanger sequencing traces across the 5' and 3' breakpoints between the *hSOD1$^{G93A}$* construct and the pJAZZ vector. Sites of Cas9 digestion of the hSOD1-L126delTT BAC, and the sgRNA guides used, are indicated. **(C)** Sanger sequencing traces across the L126delTT correction and G93A point mutation sites showing successful seamless cloning. WT, wild-type; HA, Homology arm; CmR, chloramphenicol resistance gene; KanR, kanamycin resistance gene.

| sgRNA name | Guide sequence (5'→ 3') | Binding site |
|---|---|---|
| sgRNA 1 | CCCAGAACAACCATCCCAAC | 7.7 kb 5' of humanised *SOD1* gene |
| sgRNA 2 | TTCCATCAGGAATCGAACTC | 7.7 kb 5' of humanised *SOD1* gene |
| sgRNA 3 | AAACAAAGTGCGTGAGCCCT | 8.8 kb 3' of humanised *SOD1* gene |
| sgRNA 4 | AAAACAGCGAAGTACTCAGG | 8.8 kb 3' of humanised *SOD1* gene |
| sgRNA A | GTCCATGAAAAAGCAGATGA | 1bp 3' of L126delTT site in *SOD1* exon 5 |
| sgRNA B | AGAATCTTCAATAGACACAT | 8bp 5' of G93A site in *SOD1* exon 4' |

**Table 5-1 | sgRNA guides used in the creation of hSOD1-G93A-pJAZZ.**

| Primer name | Sequence (5'→ 3') | Binding site | Target of sequencing |
|---|---|---|---|
| SL1 | CAGTCCAGTTACGCTGGAGTC | pJAZZ vector long arm MCS | 5' breakpoint between mouse HA and pJAZZ vector |
| NZ RevC | AAATGGTCAGTTAATCAGTTCT | pJAZZ vector short arm MCS | 3' breakpoint between mouse HA and pJAZZ vector |
| hSOD1_PCR3_F | TCCAGACTGCCTTGGGAAAA | KanR cassette | G93A mutation in *hSOD1* exon 4' |
| Neo_Cas_5b | GGGGAACTTCCTGACTAGGG | KanR cassette | G93A mutation in *hSOD1* exon 4' |
| hSOD1_PCR4_R | GGGCCTCAGACTACATCCAA | hSOD1 3' UTR (and duplicated 3' UTR) | L126delTT correction in *hSOD1* exon 5 |

**Table 5-2 | Sequencing primers used in the creation of hSOD1-G93A-pJAZZ.** MCS, multiple cloning site; HA, homology arm.

| PCR | Primer name | Sequence (5'→ 3') | Binding site | PCR product size |
|---|---|---|---|---|
| PCR mutagenesis | hSOD1 L126 to WT fwd | TTTTACAGGTCCATGAAAAAGCAGATGACTTGGGCAAAGGTGGAAATGAAGA | hSOD1 exon 5 | 3.5 kb |
| | hSOD1 WT to G93A Rev | GAGATCACAGAATCTTCAATAGACACATCGGCCACAGCATCTTTGTCAGCAGTCACATTGCCCAAGTCTCCAAC | hSOD1 exon 4' | |
| Colony PCR to screen for hSOD1-G93A-pJAZZ | hSOD1 genotyping Fwd | CAGTACTGTCAACCACTAGCA | hSOD1 intron 3 | 1742 bp |
| | hSOD1 WT Rev 2 | TCATTTCCACCTTTGCCCAAG | hSOD1 exon 5 (wild-type allele) | |

**Table 5-3 | PCR Primers used in the creation of hSOD1-G93A-pJAZZ.** The TT insertion to correct L126delTT is shown in green. The G→C point mutation to introduce G93A is shown in red.

### 5.3.2 Generation of the *hSOD1$^{G93A}$* mouse line

#### *5.3.2.1 ESC editing*

Editing of naïve ESCs to introduce the *hSOD1$^{G93A}$* allele was performed by Dr Ben Davies and colleagues at The Wellcome Centre for Human Genetics. Editing was achieved by CRISPR/Cas9 assisted homology directed repair (HDR) using the hSOD1-G93A-pJAZZ construct as a DNA donor.

Briefly, hSOD1-G93A-pJAZZ and CRISPR-Cas9 RNPs (Table 5-4) were introduced into the C57Bl/6NTac mouse ESC line JM8.F6 by electroporation. Putative positive transformants were selected for neomycin resistance by growth in ESC media supplemented with neomycin.

| sgRNA name | Guide sequence (5'→ 3') | Binding site |
|---|---|---|
| sgRNA 5' UTR | GGGAAGCATGGCGATGAAAG | Junction between mouse *Sod1* 5' UTR and mouse exon 1 |
| sgRNA 3' UTR | ATTACAGTTTAATGGTTTGA | Mouse *Sod1* 3' UTR |

**Table 5-4 | sgRNAs for CRISPR/Cas9-assisted editing of ESCs to generate *hSOD1$^{G93A}$* mice**

#### *5.3.2.2 Screening of targeted ESCs*

##### 5.3.2.2.1 Long-range PCR across the 3' junction

Correctly targeted ESC clones were initially identified through long-range PCR (LR-PCR) across the 3' junction, performed by Dr Ben Davies. The expected PCR amplicon size from a correctly targeted clone was 9.2 kb (Table 5-5). Genomic DNA from 13 clones yielded correctly sized PCR products in

this 3' junction screening, as identified by size selection on agarose gel electrophoresis (data not shown).

#### 5.3.2.2.2 Long-range nested PCR across the 5' junction

Genomic DNA from the 13 putative positive ESC clones was screened for correct 5' integration of the *hSOD1*$^{G93A}$ transgene by nested LR-PCR and sequencing across the 5' breakpoint. Amplification of the 5' HA was attempted using several different primer sets and annealing temperatures without success. PCR amplification either resulted in many PCR products, due to non-specific primer annealing, or no/very low concentration of PCR product, likely due to poor reaction efficiency and low quantity of input genomic DNA (data not shown). Hence, a nested PCR protocol was developed to improve selective amplification of the 5' HA while minimising non-specific amplification of other regions of genomic DNA (Figure 5-2 A).

The nested PCR protocol consisted of two PCR reactions using Q5 polymerase (Table 5-5, Table 5-6). The first reaction used 150 ng of genomic ESC DNA as a template and resulted in a 9.6 kb amplicon. The second reaction used 1 µl (1/20$^{th}$) of the PCR mix from the first reaction as a template, and primers internal to the 9.6 kb amplicon, to generate a 9.2 kb amplicon. This nested PCR protocol resulted in a correctly sized 9.2kb amplicon in 10 of the 13 putative positive ESC clones, as identified by size selection on agarose gel electrophoresis (Figure 5-2 B). This PCR reaction was repeated 3 times and clones positive in two or more reactions were considered to have correct 5' integration of the hSOD1$^{G93A}$ transgene (Table 5-8).

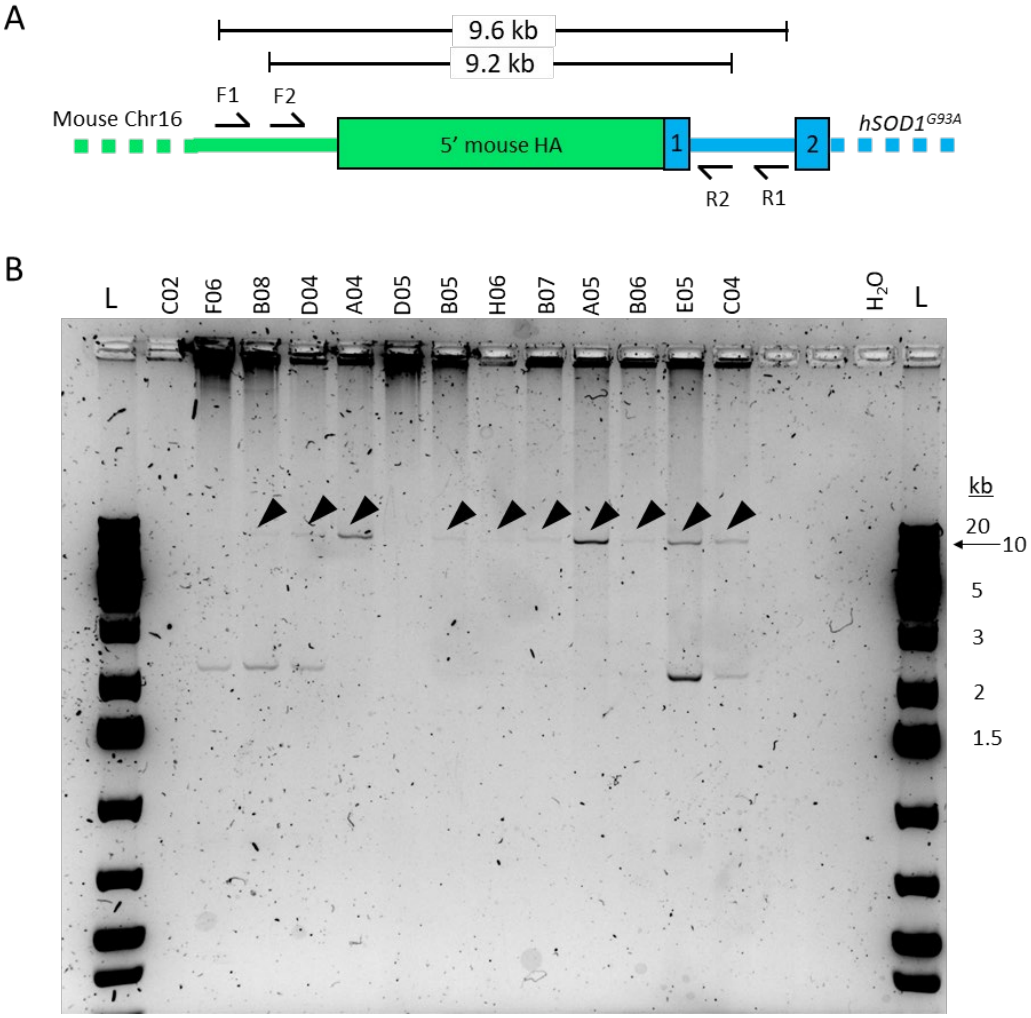



**Figure 5-2 | Nested PCR across the 5' homology arm. (A)** Schematic of the nested PCR reaction. Primers F1 and R1 were used in the first PCR reaction with a genomic DNA template. Primers F2 and R2 were used in the second PCR reaction with the amplicon from reaction 1 as a template. **(B)** Agarose gel electrophoresis image of nested PCR products from 13 ESC clones. Black arrowheads indicate clones with a correctly sized 9.2 kb amplicon. Image is representative of 3 independent PCR reactions per clone. L, Ladder.

| PCR | Primer name | Sequence (5'→ 3') | Binding site | PCR product size |
|---|---|---|---|---|
| LR-PCR across 3' junction | Sod1-3HR-F1 | CCCTTGGATGTAGTCTGAGGC | *hSOD1* 3' UTR | 9.2 kb |
| | Sod1-3HR-R1 | AGTAGCAAGATGCTGTCCCAG | Mouse genome downstream of mouse 3' HA | |
| Nested LR-PCR across 5' junction, reaction 1 | hSOD1-G93A_5'HA F1 | CAAGATGTGCCCTTTAGCCC | Mouse genome upstream of mouse 5' HA | 9.6 kb |
| | hSOD1-G93A_5'HA R1 | TGCAGCCCAATCATGACCTA | *hSOD1* intron 1 | |
| Nested LR-PCR across 5' junction, reaction 2 | hSOD1-G93A_5'HA F2 | CTTCTGATGGAGCCGCATTC | Mouse genome upstream of mouse 5' HA | 9.2 kb |
| | hSOD1-G93A_5'HA R2 | ACACAAACGGTCTCTGTTGC | *hSOD1* intron 1 | |

**Table 5-5 |PCR primers for screening of targeted ESCs**

| Reaction | Step | Temperature (°C) | Time (s) | Cycles |
|---|---|---|---|---|
| 1 | Initial denaturation | 98 | 30 | - |
| | Denaturation | 98 | 7 | 20 |
| | Annealing | 70 | 20 | |
| | Extension | 72 | 300 | |
| | Final extension | 72 | 600 | - |
| 2 | Initial denaturation | 98 | 30 | - |
| | Denaturation | 98 | 7 | 20 |
| | Annealing | 68 | 20 | |
| | Extension | 72 | 300 | |
| | Final extension | 72 | 600 | - |

**Table 5-6 | Nested PCR cycling conditions**

#### 5.3.2.2.3 Copy count ddPCR (GEMS)

To confirm single integration of the *hSOD1*$^{G93A}$ transgene the MRC Harwell Genotyping Core performed a ddPCR-based Loss-Of-Allele (LOA) copy number assay (Table 5-7). Clones with 1-2 copies of the *hSOD1*$^{G93A}$ allele were considered positive. Alongside a correctly targeted allele, it is likely the second mouse *mSod1* allele would also undergo CRISPR/Cas9 induced mutagenesis (via HDR to produce a second correctly targeted allele, or random mutagenesis via NHEJ). The presence

of an intact *mSod1* allele was not required, as subsequent breeding steps would select for offspring harbouring the *hSOD1*$^{G93A}$ allele and eliminate the mouse *mSod1* allele.

Table 5-8 contains a summary of the screening results for the *hSOD1*$^{G93A}$ ESC clones.

| Target | Forward primer 5' -> 3' | Reverse primer 5' -> 3' | 5'-FAM labelled probe 5' -> 3' |
|---|---|---|---|
| *mSod1* exon 3 | GTACCAGTGCAGGACCTCAT | AGCGTGCTGCTCACCTCT | AACATGGTGGCCCGGCGGATG |
| *hSOD1* exon 3 | GTGCAGGTCCTCACTTTAATCC | CCAGAAAGCTATCGCCATTATTACAAG | CCAAAGGATGAAGAGAGGTAACAAGATGC |

**Table 5-7 | Primers/probes used in ddPCR assays to screen for integration of the *hSOD1*$^{G93A}$ transgene**

| | | | ddPCR copy number assay | |
|---|---|---|---|---|
| ***hSOD1*$^{G93A}$ ESC clone** | **Results from LR-PCR across 3' junction** | **Results from LR-PCR across 5' junction** | ***mSod1* allele copy number** | ***hSOD1*$^{G93A}$ allele copy number** |
| A04 | Positive | Positive | 0.97 | 0.96 |
| A05 | Positive | Positive | 0.98 | 3.17 |
| B05 | Positive | Positive | 0 | 2.15 |
| B06 | Positive | Positive | 0 | 0.99 |
| B07 | Positive | Negative | 0.96 | 5.34 |
| B08 | Positive | Negative | 0 | 3.52 |
| C02 | Positive | Negative | 1.92 | 0.93 |
| C04 | Positive | Negative | 1.02 | 3.22 |
| D04 | Positive | Positive | 0 | 34.9 |
| D05 | Positive | Negative | 0 | 0.9 |
| E05 | Positive | Positive | 0 | 1.02 |
| F06 | Positive | Negative | 0 | 2.2 |
| H06 | Positive | Positive | 0 | 17 |
| Control | | | 0.96 | 0.99 |

**Table 5-8 | Summary of screening results for the *hSOD1*$^{G93A}$ ESC clones**

#### 5.3.2.3 *ESC microinjection and breeding of hSOD1$^{G93A}$ animals*

*hSOD1$^{G93A}$* mice were generated at The Wellcome Centre for Human Genetics by microinjection of modified ESCs into C57BL/6N albino donor blastocysts. The ESC clones chosen for microinjection were A04, B05, B06 and E05. Microinjection of clones A04 and B06 yielded chimeric male offspring which were then crossed to C57BL/6N females, with the offspring genotyped to screen for germline transmission (GLT) of the humanised *hSOD1$^{G93A}$* allele. ESC line B06 was the only line to result in germ-line transmission of the *hSOD1$^{G93A}$* allele. The neomycin selection cassette was removed by crossing *hSOD1$^{G93A}$* mice to C57BL/6N Flpo-expressing mice.

Sperm from heterozygous *hSOD1$^{G93A}$* neo-negative males was used for IVF with C57BL/6NTac females to import the line into the Mary Lyon Centre at MRC Harwell. The *SOD1$^{G93A}$* allele is conditional and can be converted from *SOD1$^{WT}$* to *SOD1$^{G93A}$* following Cre-Lox recombination. Half of the IVF-derived 1-cell embryos were microinjected with *Cre* mRNA to delete the *loxP* flanked exons and enable expression of *hSOD1$^{G93A}$*, whereas the remaining embryos did not undergo Cre recombination and therefore express *hSOD1$^{WT}$*. The *hSOD1$^{WT}$* and *hSOD1$^{G93A}$* offspring were crossed together to generate *hSOD1$^{WT}$*/*hSOD1$^{G93A}$* compound heterozygotes, which were then intercrossed to produce an experimental cohort. This experimental cohort will undergo motor and behavioural phenotyping in future experiments.

### 5.3.3 *hSOD1$^{G93A}$* allele quality control

To assess the integrity of the humanised *SOD1$^{G93A}$* locus, targeted locus amplification (TLA) and short-read Illumina sequencing were performed by Cergentis on splenocytes isolated from a heterozygous *mSod1*/*hSOD1$^{G93A}$* male animal which had not undergone cre recombination. For enrichment, two primer pairs were designed by Cergentis against the *hSOD1$^{G93A}$* locus (Table 5-9). This provided a ~250kb island of sequencing coverage centred around the SOD1 locus. Sequencing read depth across the humanised sequence ranged from 400x to >5000x with two exceptions:

1. Regions close to the detection primers had much higher sequencing read depth, ranging from 100,000x to >300,000x.
2. A 1.8 kb region surrounding the 5' breakpoint had very low sequencing coverage, ranging from 0x to 15x.

To confirm a correct targeting event, sequencing reads were aligned to the mouse (mm10) reference genome, the human (hg38) reference genome, the sequence of the hSOD1-G93A-pJAZZ targeting construct, and the expected sequence of the correctly targeted allele (Figure 5-3 A-D). Our targeting was performed In C57BL/6N ESCs. However, we aligned to the C57BL/6J reference genome rather than C57BL/6N because the C57BL/6J genome is the most complete and well-annotated mouse genome available, and because there are no known differences between C57BL/6J and C57BL/6N strains at the *Sod1* locus.

Alignment to the mouse reference genome showed the loss of the mouse *Sod1* gene (Figure 5-3 A). A lack of sequencing coverage elsewhere in the genome confirmed that transgene integration occurred exclusively at the mouse *Sod1* locus in chr16 (Figure 5-3 F).

Alignment of reads to the human reference genome showed insertion of the human *SOD1* gene, indicated by a region of high sequencing coverage (Figure 5-3 B). This alignment also confirmed the presence of the SOD1-G93A mutation. The short-read sequencing approach could not distinguish between the duplicated exons; however, Sanger sequencing of the hSOD1-G93A-pJAZZ vector confirmed the G93A mutation to be in exon 4, not exon 4' (Figure 5-1 C).

A 1.8 kb region of low sequencing coverage surrounding the 5' breakpoint was evident in alignments to both the mouse and human genome (red bar in Figure 5-3 A, B). A 180bp region of the humanised *SOD1* intron 1 and two short regions of 5' mouse sequence had no sequencing coverage. However, 3 reads spanned the 5' breakpoint between the mouse *Sod1* 5' UTR and human *SOD1* exon 1, confirming that this sequence was integrated into the mouse genome; although lack of surrounding

sequence information makes it impossible to confirm correct integration of this sequence at the humanised *SOD1* locus.

The abrupt drop-off in sequencing coverage around the 5' breakpoint could be cause for concern. However, examination of reads which mapped to this region revealed an assortment of sequences past the drop-off point, with no sequence consensus (Figure 5-3 E). This confirms that the drop-off in sequencing coverage was due to a technical bias in the TLA process [422], rather than a genuine deletion or rearrangement where we would expect to see a consensus in sequence from another part of the genome. As part of the TLA process, fragmentation of DNA was performed using a restriction enzyme with the recognition sequence GATC, resulting in a series of GATC-flanked fragments which are religated, decrosslinked, amplified and sequenced. The region of low sequencing coverage corresponded to a GATC-flanked fragment containing a stretch of ~450 nucleotides with 80% GC content, which likely compromised the amplification and sequencing of the entire fragment, resulting in the observed sharp drop-off in sequencing coverage. Cergentis regularly observes similar sequencing biases in TLA projects.

Alignment of reads to the hSOD1-G93A-pJAZZ vector sequence revealed a sharp drop-off in sequencing at the ends of the mouse homology arms (Figure 5-3 C). No reads mapping to the vector backbone were detected, indicating that only the region between the homology arms was integrated into the mouse genome, likely via HDR.

Finally, alignment of reads to the expected sequence of the humanised *SOD1* locus indicated seamless humanisation of the mouse *Sod1* gene, with no unexpected sequence alterations detected (Figure 5-3 D). Alignment to the expected locus also confirmed deletion of the Neomycin selection cassette, evidenced by a complete lack of reads mapped to the sequence between the FRT sites.

Together, the four alignments described above strongly indicate a correct, HDR-mediated targeting event in the *hSOD1$^{G93A}$* mouse line.

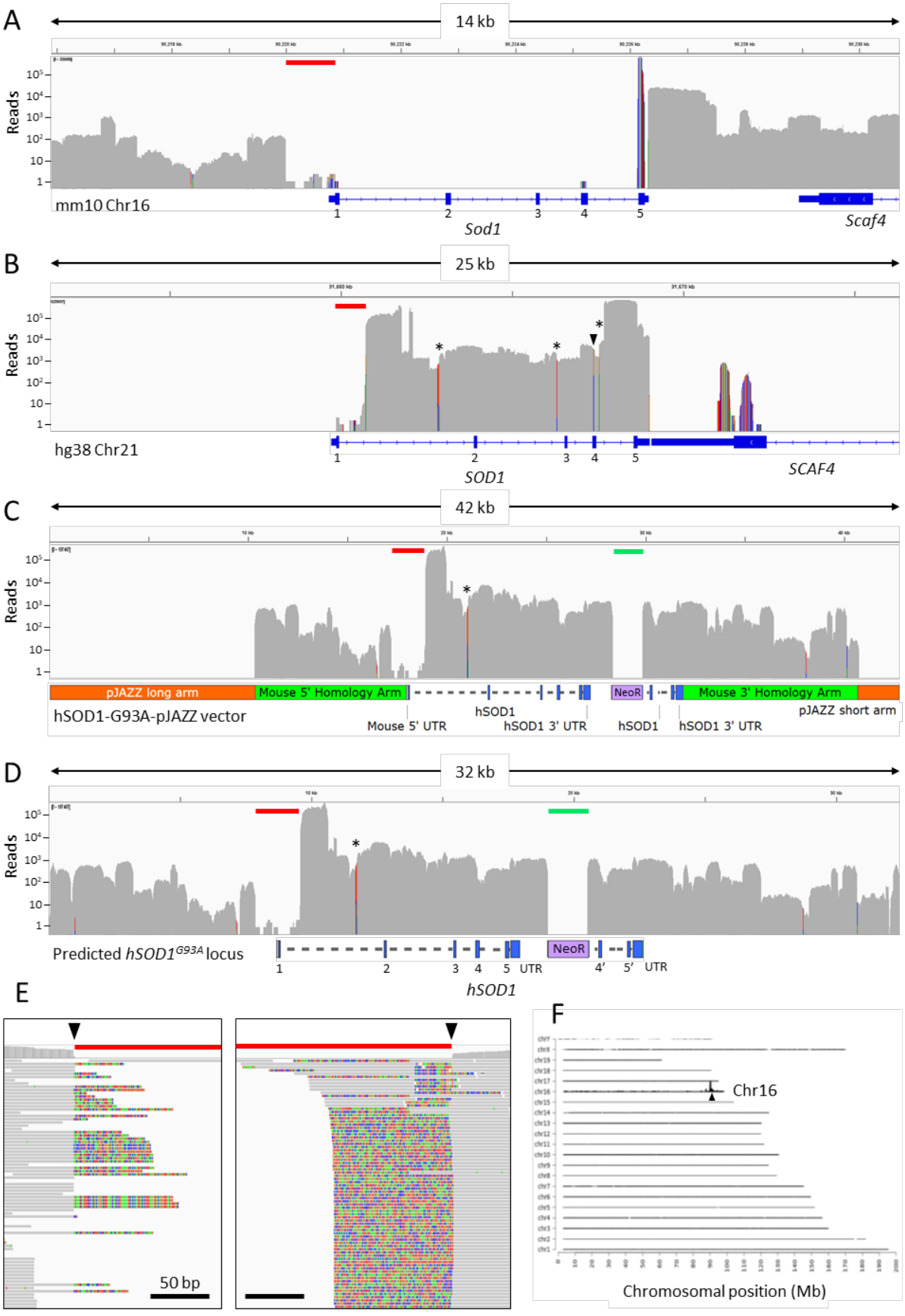

A
14 kb
Reads
mm10 Chr16
Sod1
Scaf4
B
25 kb
Reads
hg38 Chr21
SOD1
SCAF4
C
42 kb
Reads
pJAZZ long arm
Mouse 5' Homology Arm
NeoR
Mouse 3' Homology Arm
hSOD1-G93A-pJAZZ vector
hSOD1
Mouse 5' UTR
hSOD1 3' UTR
hSOD1
hSOD1 3' UTR
pJAZZ short arm
D
32 kb
Reads
Predicted hSOD1G93A locus
NeoR
UTR
UTR
hSOD1
E
50 bp
F
Chr16
Chromosomal position (Mb)

**Figure 5-3 | Targeted locus amplification (TLA) and short-read sequencing of *hSOD1$^{G93A}$* locus.** Alignment of TLA sequencing reads to **(A)** the mouse (mm10) reference genome **(B)** the human (hg38) reference genome, **(C)** the hSOD1-G93A-pJAZZ vector sequence and **(D)** the predicted *hSOD1$^{G93A}$* locus. Arrowhead indicates G93A mutation. Asterisks indicate flagged nucleotides that were checked manually. Red bar indicates area of poor sequencing coverage. Green bar indicates deletion of the neomycin resistance selection cassette by Flp-FRT recombination. **(E)** Expanded view of region of poor sequencing coverage around the 5' breakpoint showing extended reads. Mismatched bases compared to the predicted *hSOD1$^{G93A}$* locus are coloured; green = A, red = T, blue = C, orange = G. Arrowheads indicate drop-off of sequencing coverage. **(F)** Sequencing coverage across the mouse genome using *hSOD1$^{G93A}$* enrichment primer set 1. The chromosomes are indicated on the y-axis, the chromosomal position on the x-axis. Arrowhead indicates identified integration site at mouse *Sod1* locus on chromosome 16. No other integration sites were identified.

After accounting for engineered features in the hSOD1-G93A-pJAZZ targeting construct there was one misalignment to the human reference genome in a GC-rich region of intron 1 of the humanised *SOD1* gene (leftmost asterisk in Figure 5-3 B-D). The base was incorrectly called as G, rather than A, in 64% of reads. However, the majority of sequencing reads which miscalled this base systematically misidentified two further bases ~30 bp away – although these bases were called correctly in the majority of reads – suggesting the errors are likely due to a sequencing artefact associated with a subset of reads. One possible cause of these sequencing errors is high GC content of the surrounding sequence; GC content was >80% in a 20bp window surrounding each miscalled base. Another possibility is that these reads originated elsewhere in the humanised *SOD1* locus or the untargeted mouse *Sod1* locus and were incorrectly mapped to this sequence.

Several other nucleotides were flagged because they had an increased error rate in our sequencing (Figure 5-3 B); these occurred exclusively in *SOD1* introns and in each case the expected base was called in the majority of reads.

| Assay | Primer | Sequence (5'→ 3') | Binding site |
|---|---|---|---|
| *hSOD1*$^{G93A}$ enrichment primer set 1 | forward primer | TGGCCTATTCCGTGTTTAAA | *hSOD1* intron 1 |
| | reverse primer | GAGGCGTCTTAAGAATAGCA | *hSOD1* intron 1 |
| *hSOD1*$^{G93A}$ enrichment primer set 2 | forward primer | TTTGGCTTGTGGTGTAATTG | *hSOD1* intron 4 (and intron 4') |
| | reverse primer | AAGTGAGAAACCCAATCCTG | *hSOD1* exons 5 (and exon 5') |

**Table 5-9 | Detection primers for Xdrop enrichment of the *hSOD1*$^{G93A}$ locus.**

## 5.4 Discussion

The high-copy transgenic *tgSOD1*$^{G93A}$ mouse is the most widely used animal model of ALS, and has yielded many important insights into the disease. However, for reasons described extensively in section 1.5, novel mouse models of ALS are needed which more accurately model early-stage disease and exhibit greater physiological relevance. This chapter describes our efforts to develop a humanised *hSOD1*$^{G93A}$ mouse, which we hope will provide useful insights into the mechanisms driving the development and early progression of SOD1-ALS.

The cloning of the *hSOD1*$^{G93A}$ targeting vector involved manipulation of a large BAC construct containing a region of duplicated sequence. Large DNA constructs such as BACs typically lack unique (i.e. single or low frequency cutting cutting) restriction sites needed for standard restriction enzyme-based cloning. Using CRISPR/Cas9 offers a useful alternative to restriction enzymes as custom guides can be designed to target desired sites with high specificity. Our method of using CRISPR/Cas9 in lieu of suitable restriction enzymes is fairly novel, yet we have found it highly effective in seamless cloning applications. The duplicated region prevented us from using a recombineering-based approach, as the self-complementary sequence of the duplicated region would likely recombine and excise the internal sequence from the construct. For these reasons a commercial cloning company also failed to generate the *hSOD1*$^{G93A}$ targeting construct.

Generation of the *hSOD1*$^{G93A}$ targeting construct required a different edit to be introduced in each copy of the duplicated sequence in the hSOD1-L126delTT BAC. A multitude of methods exist for site-directed mutagenesis of DNA [536–538]; however, the large size of the construct and the duplicated region prevented the use of traditional PCR mutagenesis methods. Instead, a custom approach was developed which combined PCR mutagenesis with CRISPR/Cas9, Gibson assembly and the pJAZZ linear vector system. This strategy was highly efficient and yielded the correct final construct in >60% of clones (data not shown). This method is sequence nonspecific and only requires suitable guides flanking the target mutation(s), making it easily adaptable to other constructs. The use-case for this strategy is fairly narrow, as simpler mutagenesis methods exist and are compatible with most cloning projects; however, we predict this method could be useful for a subset of complicated cloning projects.

Quality control of the final hSOD1-G93A-pJAZZ targeting construct included Sanger sequencing of modified sequences and PacBio long-read sequencing of the entire vector, from which we determined the vector sequence was correct with a high degree of certainty.

Quality control of the humanised *hSOD1*$^{G93A}$ allele included LR-PCR across both homology arms and ddPCR-based copy number assays in edited ESCs, prior to microinjection. Quality control also included TLA and short-read sequencing of the *hSOD1*$^{G93A}$ locus in cells from a *hSOD1*$^{G93A}$ mouse. This sequencing confirmed there was single transgene integration site in the *hSOD1*$^{G93A}$ mouse genome, situated at the *Sod1* locus, and that the resultant allele harboured a seamless humanisation of the mouse *Sod1* gene from exon 1 to the 3' UTR, an engineered duplication and the SOD1-G93A mutation. No unexpected sequencing results were observed.

Phenotyping and molecular characterisation of *hSOD1*$^{G93A}$ mice is yet to begin; breeding of an experimental cohort is underway. Planned studies include:

1. Analysis of the expression and dismutase activity of the SOD1$^{G93A}$ protein in various tissues.

2. Longitudinal motor and behavioural phenotyping of a cohort of *hSOD1*$^{G93A}$ animals to establish disease severity and progression.
3. An analysis of cellular phenotypes related to neurodegeneration in ALS, including examination of motor neuron loss and neuromuscular junction denervation.

Following the above characterisation experiments, we plan to disseminate the *hSOD1*$^{G93A}$ mouse line to the wider scientific community. Our initial characterisation studies will focus on constitutive expression of mutant *SOD1*$^{G93A}$. However, the conditional aspect of this strain – the allele can be converted from *SOD1*$^{WT}$ to *SOD1*$^{G93A}$ following Cre-Lox recombination – will enable rescue experiments and studies involving tissue-specific expression of mutant SOD1.

# Chapter 6: Progress towards the creation of a humanised mouse model of C9orf72-ALS/FTD

## 6.1 Introduction

A hexanucleotide GGGGCC repeat expansion in intron 1 of the *C9orf72* gene is the most common cause of familial ALS and FTD in European populations [237]. Although intermediate expansions in the order of 20-30 repeats have been associated with ALS [539], most C9orf72-ALS/FTD patients carry expansions of hundreds, or even thousands, of repeats [20]. Pathology in C9orf72-ALS/FTD is believed to arise through three main mechanisms: loss of function of the C9orf72 protein, toxicity from *C9orf72* repeat RNA, and toxicity from dipeptide repeat proteins (DPRs) produced as a result of non-ATG translation of the repeat [238].

Numerous attempts have been made to model C9orf72-ALS/FTD *in vitro* and *in vivo*. *In vitro* models of C9orf72-ALS/FTD consist of patient-derived iPSCs with varying repeat sizes, or engineered cell lines that typically overexpress relatively short repeats [238]. A wide range of *in vivo* models exist, including *C. elegans*, *Drosophila,* Zebrafish, and mouse [540]. However, with the exception of mouse, most *in vivo* models of C9orf72-ALS/FTD express short repeats, with the majority of models harbouring <100 GGGGCC repeats [271,285,286,541].

Existing mouse models have been created through somatic transgenesis by introducing short *C9orf72* repeats into the CNS via AAV vectors, or via traditional transgenesis using synthetic repeat constructs [298,393] or C9orf72-ALS/FTD patient-derived bacterial artificial chromosomes (BACs) (reviewed in reference [238]). To date, four different *C9orf72*-repeat-BAC transgenic mouse lines have been published, with repeat sizes ranging from 100 to 1000 repeats [272–274,390]. ALS-associated phenotypes in these lines are highly variable; whilst all studies reported *C9orf72* repeat RNA foci and DPRs, only one model – C9-500 – has been found to exhibit motor phenotypes and TDP-43 pathology characteristic of ALS [274,395]. However, some controversy exists around the phenotypes

present in C9-500 mice, with a separate multi-laboratory study finding no evidence of neurodegeneration [396].

The unexpected phenotypic differences between BAC transgenic strains may result from differences in repeat length, differences in the genetic backgrounds used, and/or differences in the transgene integration site, which in turn may impact repeat stability and DPR/RNA-foci expression level. These differences are compounded by repeat length heterogeneity that was observed within each strain. Furthermore, none of the BAC transgenic strains display *C9orf72* downregulation, as seen in patients, due to the presence of intact mouse *C9orf72*.

To circumvent these issues, we are creating a humanised mouse model of C9orf72-ALS/FTD. One key step towards this goal was to stabilise and manipulate a large *C9orf72* GGGGCC repeat in bacterial culture, so that further cloning could produce a targeting construct for humanisation of the mouse *C9orf72* locus in ESCs. Numerous studies have cloned repeat DNA of various sizes and nucleotide sequences [542–545]; however, large GGGGCC repeats remained refractory to cloning due to their 100% GC content and instability in bacterial culture [272,286,546].

In this chapter I first describe the work undertaken to accurately size and stabilise BAC-derived GGGGCC repeat DNA, and the challenges which prevented cloning of these BAC-derived repeats into mouse targeting constructs. Secondly, I describe our alternative strategy of *de novo* synthesis of large GGGGCC repeats. Thirdly, I describe the use of large synthesised GGGGCC repeats in the generation of targeting constructs for genetic engineering of mouse ESCs and human iPSCs. Finally, I describe the generation of a humanised *C9orf72* mouse with a large GGGGCC repeat.

The complexity of the cloning described in this chapter necessitated multiple strategies to be run in parallel to maximise the chance of success (Figure 6-1). I worked closely with Dr Remya Nair on this project; some strategies we each pursued independently, whereas others were a collaborative

effort. I have highlighted occasions where work was performed solely by Dr Nair, included to provide context.

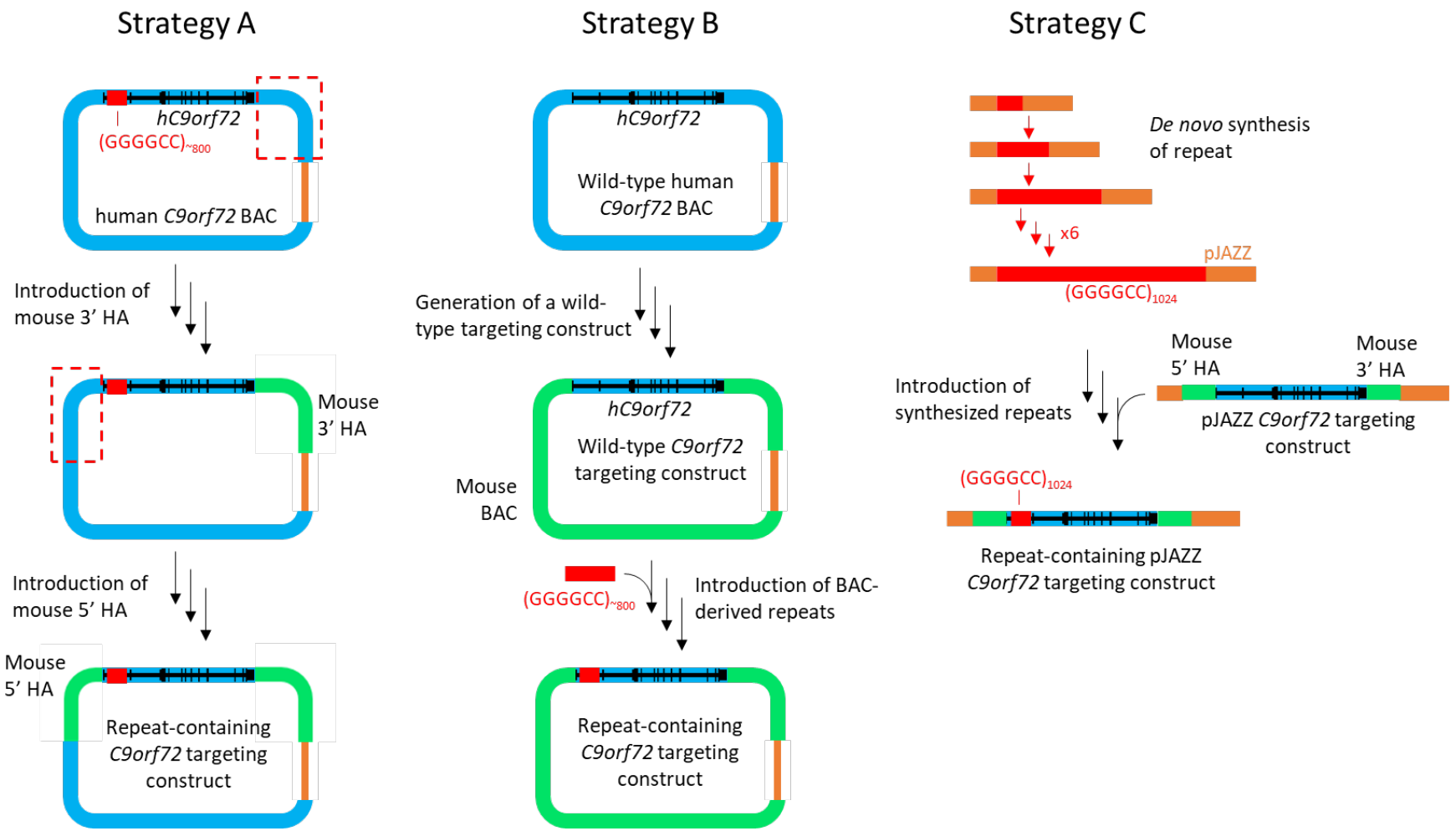



**Figure 6-1 | Strategies for generation of a targeting construct to humanize mouse *C9orf72* and introduce a large hexanucleotide repeat. (A)** Strategy for introduction of mouse homology arms into a human *C9orf72* repeat-carrying BAC. **(B)** Strategy involving transfer of wild-type human *C9orf72* into a mouse BAC, followed by introduction of a patient BAC-derived GGGGCC repeat**. (C)** Strategy to synthesise a large GGGGCC repeat *de novo*, then introduce the repeat into a linear targeting construct containing wild-type human *C9orf72* and mouse homology arms. Mouse sequence is depicted in green, human sequence in blue, and vector sequence in orange. Diagram not drawn to scale.

## 6.2 Aims

1. To generate a targeting construct to humanise the mouse *C9orf72* gene and introduce a large C9orf72 GGGGCC repeat expansion.
2. To generate a targeting construct to introduce a large *C9orf72* repeat expansion into human iPSCs.
3. To generate a humanised *C9orf72* mouse with a large GGGGCC repeat.

## 6.3 Results

### 6.3.1 Growing and sizing a large hexanucleotide repeat in the *C9orf72* gene within a BAC construct

The starting material for our *C9orf72* cloning was a 168.5 kb BAC vector derived from a *C9orf72*-ALS patient, containing the human *C9orf72* gene and ~800 GGGGCC hexanucleotide repeats, flanked by 26.4 kb 5' and 119.5 kb 3' sequence from the human *C9orf72* locus (BAC clone RP11-27J8) [272].

#### *6.3.1.1 Development of a novel method to size C9orf72 repeats using CRISPR/Cas9*

The primary existing method for sizing *C9orf72* repeats was Southern blotting. However, Southern blotting is low-throughput and labour intensive. Therefore, to facilitate high-throughput screening of sub-clones for un-retracted repeats, we developed a CRISPR/Cas9-based screening method [429]. Dr Nair designed 3 sgRNA guides against the human *C9orf72* gene (Table 6-1); sgRNA 7 (1 kb upstream of the repeat), sgRNA 6 (840 bp downstream of the repeat), and sgRNA 5 (5.4 kb downstream of the repeat) (Figure 6-2 A). Cas9 digestion of un-retracted *C9orf72* BAC DNA using guides 6 + 7 or guides 5 + 7 released a 6.7 kb band or a 11.2 kb band respectively. This band was then visualised using pulsed-field agarose electrophoresis, allowing rapid and reliable identification of clones of interest (Figure 6-2 B-C).

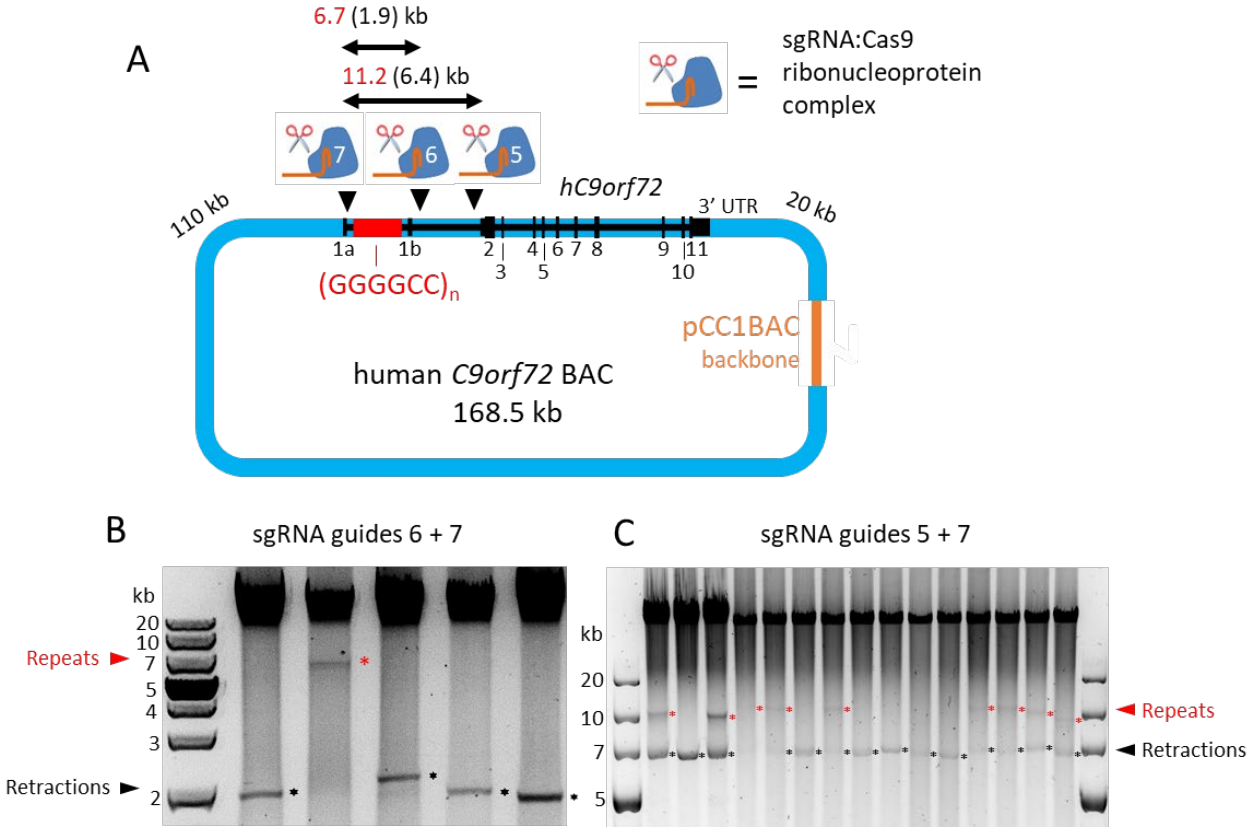



**Figure 6-2 | CRISPR/Cas9-based method to size *C9orf72* repeats. (A)** Schematic of the *C9orf72*-ALS patient BAC vector, which harbours the human *C9orf72* gene, a GGGGCC hexanucleotide repeat expansion and ~130kb of surrounding genomic sequence. The target sites of sgRNA 5, 6 and 7 are indicated, along with the expected size of the released fragment in red (assuming 800 GGGGCC repeats). The expected size of the released fragment from a fully retracted BAC is indicated in brackets. **(B-C)** Agarose gel electrophoresis images of DNA from *C9orf72*-BAC sub-clones digested with CRISPR-Cas9 and sgRNA guides **(B)** 6 + 7 or **(C)** 5 + 7. Red asterisks indicate repeat-containing bands, black asterisks indicate large retractions. Figure adapted from Nair et al., 2020 [429].

| **sgRNA name** | **Guide sequence (5'→ 3')** | ***In vitro* synthesis template sequence (5'→ 3')** | **Binding site** |
|---|---|---|---|
| sgRNA 5 | CATAACCAGAG AGTTCACTG | CGTAATACGACTCACTATAGG**CATAACCAGAGAGTTCACTG** GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCC GTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT | ~5.3 kb 3' of repeat |
| sgRNA 6 | TTTCTGAATACA AAGCCTGG | CGTAATACGACTCACTATAGG**TTTCTGAATACAAAGCCTGG** GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCC GTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT | 840 bp 3' of repeat |
| sgRNA 7 | AACGTTTTAATC ATTCACCG | CGTAATACGACTCACTATAGG**AACGTTTTAATCATTCACCG** GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT | ~1 kb 5' of repeat |

**Table 6-1 | Cas9 sgRNAs for sizing the *C9orf72* repeat.** The *in vitro* synthesis templates consist of a T7 promotor sequence, the sgRNA guide sequence (highlighted in bold), and a tracrRNA sequence.

#### 6.3.1.2 *Optimising growth conditions of a BAC construct containing a large C9orf72 hexanucleotide repeat*

Large expanded repeats are inherently unstable in bacterial culture [547]. While generating their C9orf72-BAC-transgenic line, O'Rourke and colleagues reported major retractions in repeat length in the majority of sub-clones [272]. Similarly, we observed retractions in repeat length in the vast majority of sub-clones of human *C9orf72* BAC (Figure 6-3). Even sub-clones containing full-length repeats were heterogeneous in nature; the vast majority of sub-clones showed a bimodal distribution of ~50 and ~800 repeats, where full-length ~800 repeat DNA made up a small proportion of the total BAC DNA isolated. Many studies have reported that incubation conditions designed to reduce the bacterial growth-rate lead to improved retention of expanded repeats [548,549], including GGGGCC repeats [274,546]. Hence, we sought to improve the yield of full-length repeats through manipulation of the bacterial culture conditions.

Incubation temperature has a significant impact on the stability of repeat DNA in bacterial culture [549]. To assess the impact of incubation temperature on the stability of the repeat, we grew *C9orf72* BAC sub-clones at a range of temperatures: 30°C, 28°C and room-temperature (~21°C) and examined the proportion of un-retracted sub-clones (Figure 6-3 A-C). We found that although almost all clones had evidence of retracted repeats, clones with bimodal repeat length – where at least some proportion of the DNA contained full-length repeat – were more common at incubation temperatures ≤28°.

The stability of repeat DNA is improved in recombination-deficient bacteria [550,551]. A derivative of the recombination-deficient *E. coli* strain DH10B, which carries a mutant recA recombinase (*rec*A1), was used for the majority of our *C9orf72* BAC culture. However, we also tested stability of the repeat in Stbl4 cells (Invitrogen), another recombination-deficient *E. coli* strain which has mutations in the methylation-dependent restriction system, in addition to the *rec*A1 genotype to further stabilise

direct repeats [550]. However, we did not observe any significant improvement in full-length repeat retention in Stbl4 cells compared to DH10B cells (Figure 6-3 D).

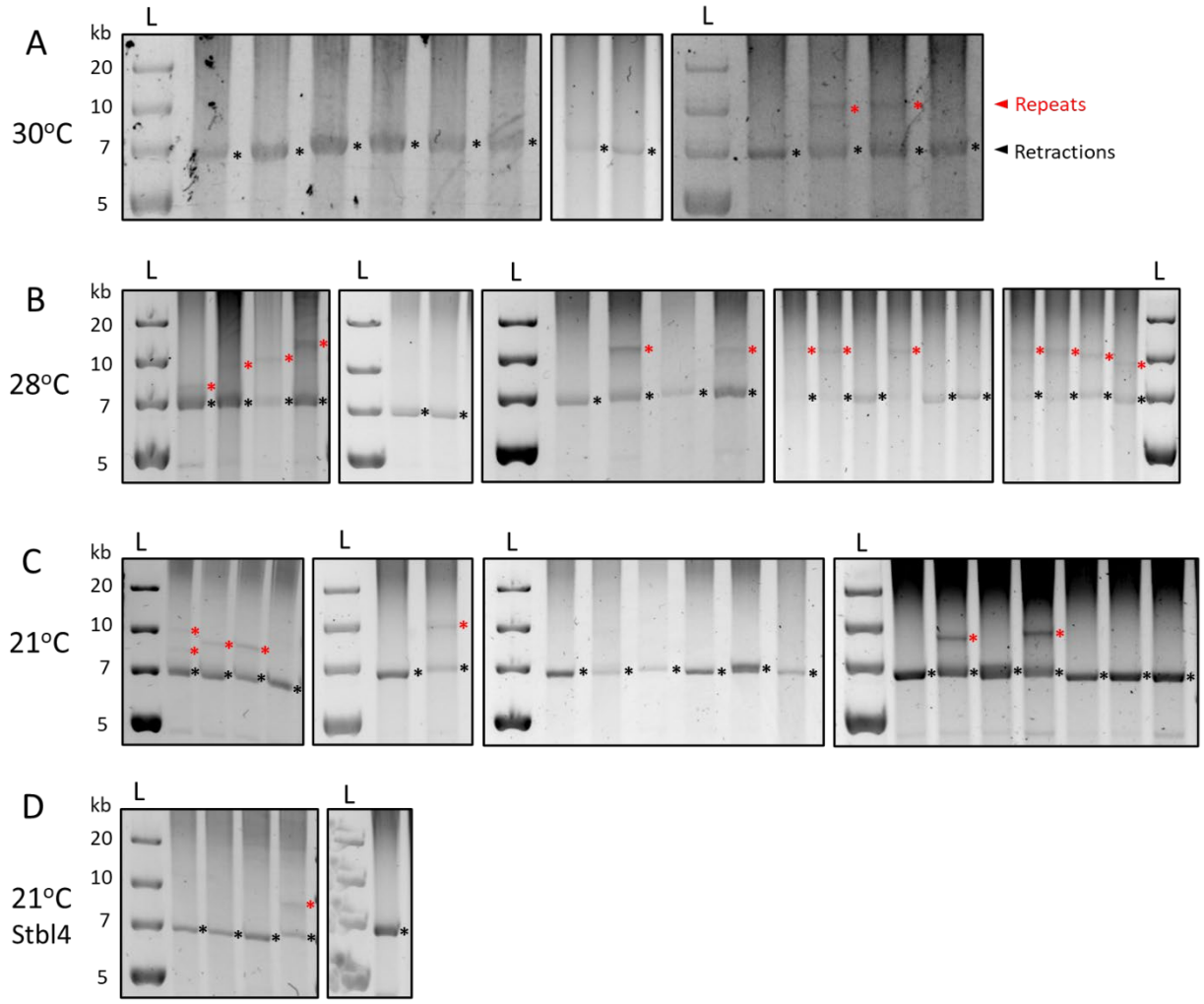



**Figure 6-3 | Optimisation of culture conditions for a repeat-containing human *C9orf72* BAC.** Representative agarose gel electrophoresis images of *C9orf72* repeat BAC DNA isolated from *E. coli* DH10B cultures grown at **(A)** 30°C, **(B)** 28°C and **(C)** 21°C, or **(D)** *E. coli* Stbl4 cultures grown at 21°C. BAC DNA was digested with CRISPR/Cas9 and sgRNA guides 5 + 7 to visualise the repeat. The expected size of the released fragment from a BAC containing 800 GGGGCC repeats is 11.2 kb, and the expected fragment size from a fully retracted BAC is 6.4 kb. Red asterisks indicate repeat-containing bands, black asterisks indicate large retractions. L, ladder.

From the above observations, incubation of DH10B bacteria at ≤28°C was chosen for optimal retention of the full-length repeat. However, regardless of these improvements, retention was still poor. Clones harbouring retracted repeats appeared to have a strong growth advantage upon

subculturing, especially in rapid growth conditions, effectively obscuring the full-length repeat in the majority of clones.

#### *6.3.1.3 Long-read sequencing of a C9orf72 repeat expansion within a BAC construct*

We sent a sample of BAC DNA from a full-size repeat-harbouring clone for PacBio long-read sequencing. Although coverage across the repeat was poor, with only a single read spanning the entire repeat, this sequencing confirmed the presence of a large GGGGCC repeat, 728 repeat units in length, with 100% GC content and ~95% of the sequence reading GGGGCC. Small interruptions consisting of loss or gain of single G or C nucleotides were observed throughout the repeat (see Appendix 8.5 for full sequence). It is likely that these interruptions were due to sequencing errors in the single PacBio sequencing read, rather than genuine interruptions in the repeat DNA sequence; this was supported by sanger sequencing of ~350 bp into the 5' and 3' ends of the repeat, which did not uncover any interruptions to the GGGGCC sequence, including at two sites where interruptions were detected by PacBio sequencing (Figure 6-4**,** Table 6-2). However, genuine interruptions further into the repeat sequence cannot be ruled out.

Both PacBio and Sanger sequencing revealed an insertion/ deletion event immediately 3' to the repeat (-GTGGTC +CGGGCCCG) (Figure 6-4). Similar 3' sequence variants have been reported in a number of C9orf72-ALS/FTD patients [552].

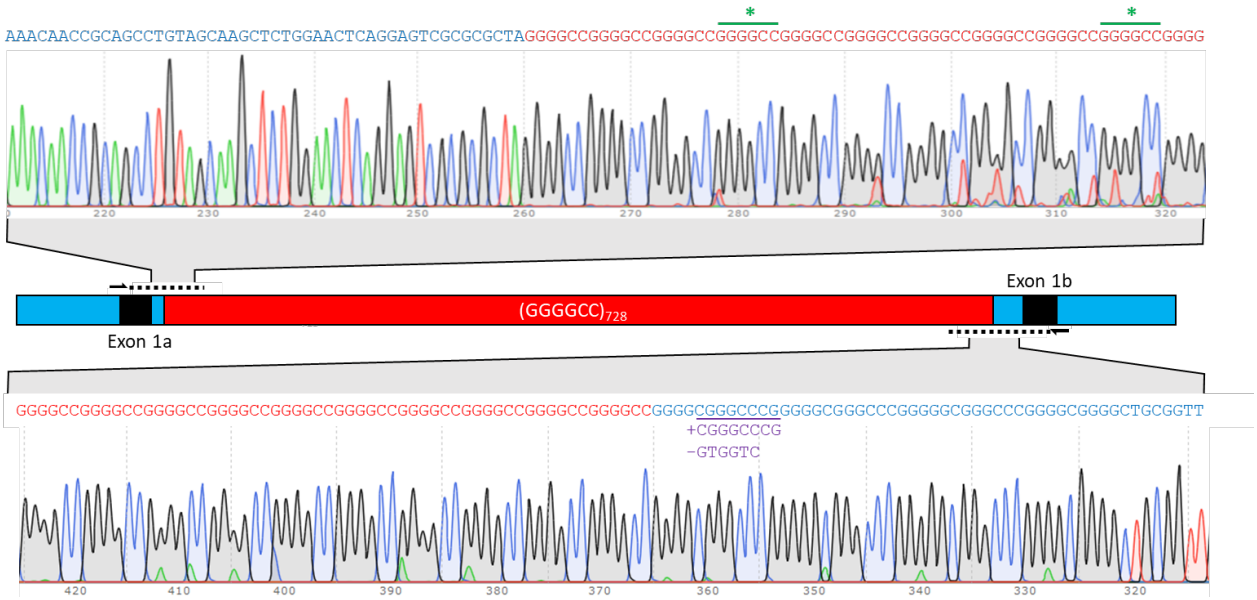



**Figure 6-4 | Sanger sequencing into 5' and 3' ends of a patient-derived *C9orf72* repeat.** Repeat sequence is denoted in red, non-repeat sequence is denoted in blue. Green asterisks highlight the sites of two PacBio sequencing errors (that read as GGGCC) which were sequenced as the expected GGGGCC by sanger sequencing. Purple text highlights an insertion/deletion event (GTGGTC -> CGGGCCCG). Each sequencing trace shown is representative of sequencing traces obtained from 3 separate sequencing primers at each end of the repeat. Figure adapted from Nair et al., 2020 [429].

| Primer name | Sequence (5'→ 3') | Binding site |
|---|---|---|
| hC9_Rep_Seq_Fwd3 | CACCCTCTCTCCCCACTACT | 5' of repeat |
| hC9_Rep_Seq_Fwd4 | GCGTCAAACAGCGACAAGTT | 5' of repeat |
| hC9_Rep_Seq_Fwd5 | GCCCACGTAAAAGATGACGC | 5' of repeat |
| hC9_Rep_Seq_Rev1 | CAAGGAAGAGGCCAGATCCC | 3' of repeat |
| hC9_Rep_Seq_Rev4 | ATGCAGGCAATTCCACCAGT | 3' of repeat |
| hC9_Rep_Seq_Rev5 | AAGGAGACAGCTCGGGTACT | 3' of repeat |

**Table 6-2 | Sequencing primers for sanger sequencing of the *C9orf72* repeat**

### 6.3.2 Creation of a mouse *C9orf72* targeting construct

The purpose of this cloning was to generate a targeting construct containing the full human C9orf72 gene from the start of intron 1 through to the TAA stop codon in exon 11 (mouse chr4:35,225,928-35,192,881, replaced with human chr9:27,573,708-27,548,235), including a large GGGGCC repeat expansion in intron 1, flanked by mouse homology arms (HAs) of sufficient length to allow homologous recombination with the mouse C9orf72 locus in ESCs or zygotes.

The starting materials for this cloning were the human *C9orf72* repeat-containing BAC vector, described in section 6.3.1, and a 161.7 kb mouse BAC containing the mouse C9orf72 gene and the surrounding genomic locus (BAC clone B6Ng01-92L5).

#### *6.3.2.1 Strategy A: introduction of mouse homology arms into a human C9orf72 repeat-carrying BAC*

The initial cloning strategy for this project involved introducing large mouse HAs flanking the human C9orf72 gene in the human C9orf72 BAC. Our strategy for introducing the 3' mouse HA is depicted in Figure 6-5. If this editing was successful we planned to use the same methodology to introduce a 5' mouse HA.

The strategy utilised Red/ET recombination, an efficient way to seamlessly engineer large DNA fragments using short (~50 bp) homologies [553–555], also termed recombineering. High sequence homology between the human and mouse C9orf72 loci – particularly towards the 3' end of the C9orf72 gene – precluded the use of recombineering to directly transfer the human C9orf72 gene into the mouse BAC (datum not shown). Instead, we attempted to introduce mouse HAs upstream and downstream of the human C9orf72 gene individually using a strategy similar to 'Assisted large fragment insertion by Red/ET-recombination' (ALFIRE), which is a 2-step method that utilises an *RpsL*-counter-selection system to improve recombination efficiency between large DNA fragments [556]. *RpsL* confers dominant sensitivity to streptomycin [557].

We first amplified an *Rpsl*-Amp$^R$ resistance cassette with extended primers (Table 6-3) to introduce (50bp) human and mouse HAs flanking *Rpsl*-Amp$^R$ in the resulting PCR product (*Rpsl*-Amp$^R$-HA) (Figure 6-5 B, Step 1). This PCR product was then introduced into the human C9orf72 BAC downstream of the TAA stop codon using recombineering (Figure 6-5 B, Step 2). Recombinants were plated on LB plates with chloramphenicol (12.5 µg/mL), kanamycin (15 µg/mL) and ampicillin (100 µg/mL). Putative positive clones were screened by colony PCR (Table 6-3) which confirmed introduction of the *Rpsl*-Amp$^R$-HA cassette into the human C9orf72 BAC (hC9-BAC-*Rpsl*-HA). In

parallel, a 36.4 kb DNA fragment was isolated from the mouse C9orf72 BAC by SacII and PacI digestion followed by gel purification, containing 11.6 kb of the mouse C9orf72 gene and 3'UTR and 25.2 kb of downstream sequence, including an FRT-flanked neo/kan$^R$ cassette (mC9-HA). Recombination was then attempted between the hC9-BAC-HA and mC9-HA, to generate a recombinant BAC containing the human C9orf72 gene, up to the TAA stop codon, followed by the mouse c9orf72 3'UTR and 25.2 kb 3' mouse HA (Figure 6-5 B, Step 3). Recombinants were plated on LB plates with chloramphenicol (12.5 μg/mL), kanamycin (15 μg/mL) and streptomycin (50 or 15 μg/mL); carryover of undigested hC9-BAC-HA was prevented using *RpsI* counter-selection. However, we did not obtain any colonies from 3 independent recombineering attempts, suggesting that this recombineering was highly inefficient.

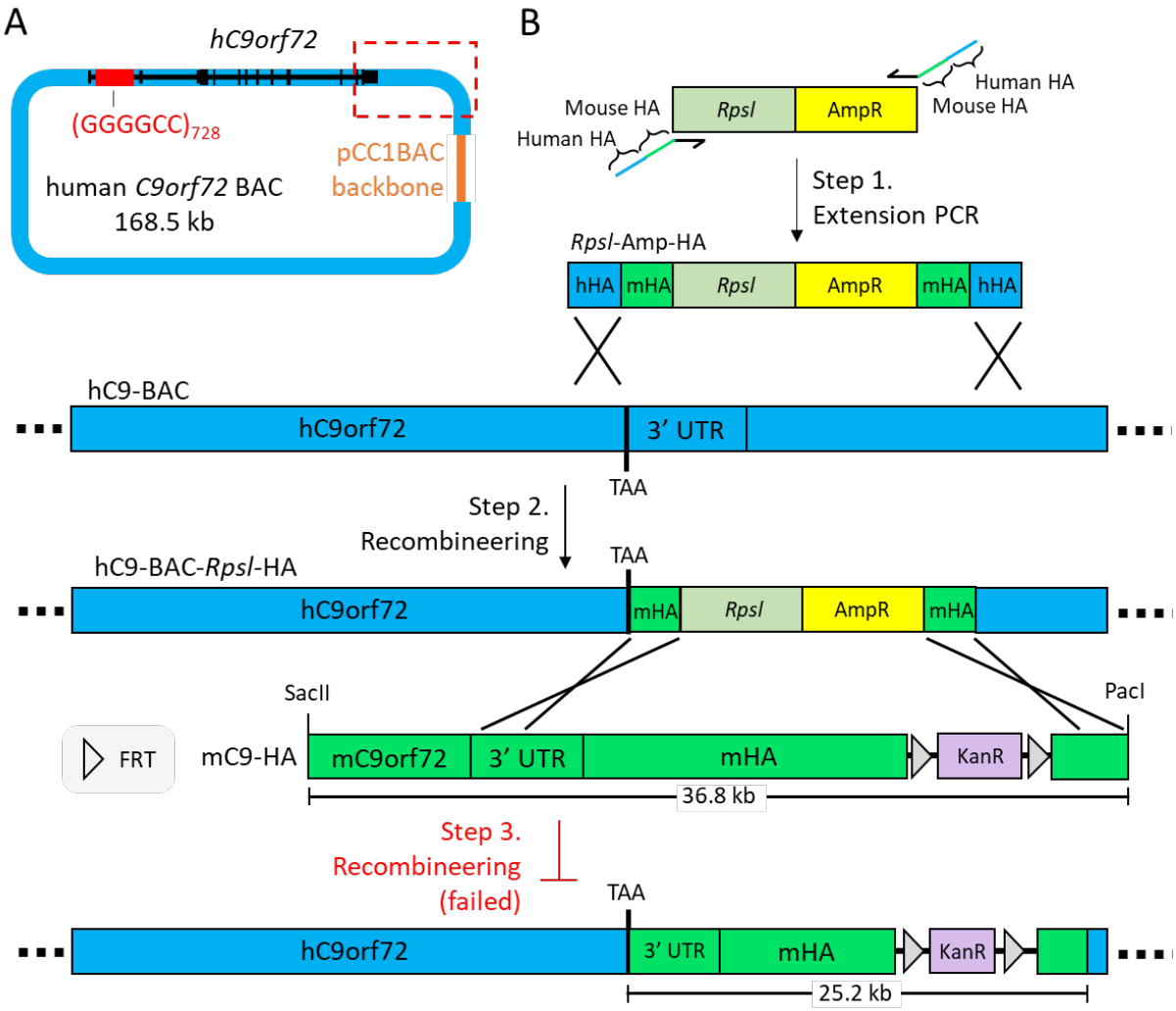



**Figure 6-5 | Strategy for introduction of a mouse 3' homology arm (HA) into a patient-derived *C9orf72* BAC. (A)** Schematic of the human *C9orf72* BAC. Red box indicates the 3' HA. **(B)** 3-step strategy to introduce a ~25 kb mouse 3' HA downstream of the human *C9orf72* gene in a patient-derived BAC. Step 1 involved extension PCR to generate an *Rpsl*-AmpR selection/counter-selection cassette flanked by short human and mouse HAs. Step 2 involved introduction of this selection/counter-selection cassette into the human BAC directly 3' of the TAA stop codon by recombineering. Step 3 involved recombineering to replace the selection/counter-selection cassette with a 25.2 kb mouse HA; however, no colonies were obtained from step 3. If this strategy was successful, a similar approach would have been used to introduce a mouse 5' HA upstream of the human *C9orf72* gene, to generate a final mouse targeting construct. Mouse sequence is depicted in green, human sequence in blue. Diagram not drawn to scale. mHA, mouse homology arm; hHA, human homology arm; AmpR, Ampicillin resistance cassette; KanR, Kanamycin resistance cassette.

| PCR | Primer name | Sequence (5'→ 3') | Binding site | PCR product size |
|---|---|---|---|---|
| Extension PCR of *Rpsl*-Amp$^R$ | ALFIRE-RHA-fwd | GACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAATGACTTTTTAACCGTGTGGTTTGCTGTGTCTGTCTCTTCACAGTCACACCTGCTGTTACAGTAGGGATAACAGGGTAATGGCCTGGTGATGATGGCGGG | 5' end of the Rpsl-Amp cassette | 1690 bp |
| | ALFIRE-RHA-rev | TAATAAATTTGCGGCCGCTAATACGACTCACTATAGGGAGAGGATCCGCGCTCAAAGACAAAAACCACATGATCATCTCGTTGGATGCAGAAAAAGCATTTTACCCTGTTATCCCTATTACCAATGCTTAATCAGTG | 3' end of the Rpsl-Amp cassette | |
| Colony PCR of hC9-BAC-*Rpsl*-HA 5' junction | VecA_hBAC4 | TGTTTAGTGTGTATTGTCTGCCA | 3' end of human C9orf72 | 560 bp |
| | VecA_hBAC5 | ACAGACACAGCAAACCACAC | Mouse C9orf72 3' UTR (mouse HA) | |
| Colony PCR of hC9-BAC-*Rpsl*-HA 3' junction | Amp_F | GAGGCGGATAAAGTTGCAGG | Amp$^R$ gene | 560 bp |
| | BB_R | AGCGGGGCACATTTCATTAC | BAC backbone | |

**Table 6-3 | Primers used in the creation of hC9-BAC-Rpsl-HA.** The human homology arms are highlighted in blue, and the mouse homology arms are highlighted in green.

Furthermore, the recombineering of constructs containing a large hexanucleotide repeat was potentially problematic; very short regions of homology are sufficient for recombination to occur [555], hence recombineering would likely cause the repeat sequence, which was highly homologous to itself, to recombine internally and be removed from the BAC. Considering this, we next devised alternative strategies that did not rely on recombineering the repeats.

#### *6.3.2.2 Strategy B: Creation of a wild-type C9orf72 targeting construct, followed by introduction of patient-BAC-derived repeats*

In parallel with Strategy A, Dr Nair used recombineering to produce a BAC targeting construct containing the wild-type human *C9orf72* gene, with no repeat, and large mouse HAs (32.9 kb 5' and 35 kb 3') (hC9orf72-WT-mHA BAC). This construct also included a hygromycin resistance cassette

flanked by piggyBac transposon-specific inverted terminal repeats (ITRs) [558], inserted into an TTAA site 69 bp 5' of the repeats to allow selection of bacterial/ESC clones followed by removal of the selection cassette in a scarless manner. This construct increased our flexibility when cloning the large repeat, as we could now clone a smaller region surrounding the repeat with the aim of seamlessly integrating it into the hC9orf72-WT-mHA BAC.

The availability of full-size repeat DNA was the limiting factor for our cloning. As described in section 6.3.1.2, the full-size repeat in the hC9-BAC was highly unstable in bacterial culture. Hence, we sought to stabilise the repeat by cloning it into an intermediate vector – a modified pSMART BAC (Lucigen) – designed to stabilise unstable sequences. Cloning with the full-size repeat proved very challenging; therefore, the below Cpf1 and Cas9 cloning strategies were attempted in parallel.

#### 6.3.2.2.1 Cpf1 cloning

The large size of the human BAC and the requirement for precise junctions between the human and mouse HAs precluded the use of restriction enzymes for traditional restriction-ligation cloning. Instead, we used CRISPR technology, which enables highly sequence-specific nuclease activity restrained only by the location of PAM sites [559]. Cas12a (Cpf1) is a programmable nuclease that generates a staggered double strand break following cleavage of DNA, leaving a 4 or 5 nt 5'-overhang that is amenable to 'sticky end' cloning [560]. Cpf1 has previously been applied to 'sticky end' cloning of large constructs [561]; hence, we sought to apply this method to clone the repeats into pSMART.

First, we introduced a 2.2 kb 'landing pad' into the pSMART BAC (Figure 6-6 A, step 1). The landing pad was designed containing target sequences for several predicted crRNAs (Table 6-4) flanking an *RpsI* counter-selection cassette. pSMART BAC pre-linearized with BamHI was purchased from Lucigen. The landing pad was synthesised as a gBlock (IDT), digested with BamHI-HF, ligated into the linearized pSMART BAC using T4 ligase, and chemically transformed into *E. coli* strain DH5α. Transformants were plated on LB plates with chloramphenicol (12.5 µg/mL) and ampicillin (100

μg/mL). The resulting colonies were screened for the pSMART-cpf1 plasmid using colony PCR (Table 6-5), with putative positive clones confirmed by sanger sequencing (Table 6-6).

We next screened 6 crRNAs – one 5' (crRNA A) and five 3' (crRNAs B-F) of the repeat (Table 6-4) – for cutting efficiency *in vitro* on a synthetic DNA template (gBlock; IDT)(Figure 6-6 B). crRNA A showed good cutting efficiency *in vitro*. Of the crRNAs which cut 3' of the repeat, crRNAs B and F showed the highest cutting efficiency *in vitro*.

| **crRNA name** | **Guide sequence (5'→ 3')** | **In vitro synthesis template sequence (5'→ 3')** | **Binding site** |
|---|---|---|---|
| crRNA A | CCACCCTCTCTCCCCACTA | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**CCACCCTCTCTCCCCACTA** | 124 bp 5' of repeat |
| crRNA B | CCTCTCATTTCTCTGACCG | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**CCTCTCATTTCTCTGACCG** | 200 bp 3' of repeat |
| crRNA C | GCCTCTAGCGACTGGTGGA | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**GCCTCTAGCGACTGGTGGA** | 247 bp 3' of repeat |
| crRNA D | AGCAAGTCTGTGTCATCTC | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**AGCAAGTCTGTGTCATCTC** | 701 bp 3' of repeat |
| crRNA E | AACATCAGGATCGGGCTTC | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**AACATCAGGATCGGGCTTC** | ~5.3 kb 3' of repeat |
| crRNA F | GCCCCTAACCCATAGCAAC | GCGTAATACGACTCACTATAGGAATGGATAATTTCTACTAAGTGTAGAT**GCCCCTAACCCATAGCAAC** | ~6 kb 3' of repeat |

**Table 6-4 | Cpf1 crRNA sequences**

We then digested the pSMART-cpf1 plasmid with Cpf1 using crRNA guide pairs A and B or A and F to excise the 600bp *Rpsl* cassette (Figure 6-6 A, step 2). The large fragment containing the pSMART BAC backbone was identified by size selection on agarose gel electrophoresis and isolated by gel purification.

| **Primer name** | **Sequence (5'→ 3')** | **Binding site** | **PCR product size** |
|---|---|---|---|
| VecA_hBAC1 | ACAAGACTGCAGAGATCCGC | Mouse C9orf72 exon 1a | 1221 bp |
| Amp_F2 | GTATCATTGCAGCACTGGGG | $Amp^R$ gene | |

**Table 6-5 | pSMART-cpf1 colony PCR primers**

| Primer name | Sequence (5'→ 3') | Binding site |
|---|---|---|
| SL1 | CCAGTCCAGTTACGCTGGAGT | pSMART backbone 3' end |
| SR4 | AAATCATACCAACATGGTCAA | pSMART backbone 5' end |

**Table 6-6 | pSMART-cpf1 sequencing primers**

We then attempted to cut the human C9orf72 BAC with the same crRNA guides to excise the repeat (Figure 6-6 A, step 3). We screened three pairs of cpf1 crRNAs for efficiency of cutting the human C9orf72 BAC; crRNA A combined with crRNA B, D or F. However, none of the crRNA pairs tested produced a repeat-containing fragment visible by PFGE (data not shown), suggesting that cpf1 cut the human BAC with low efficiency. This approach was ultimately abandoned due to poor availability of repeat BAC DNA, and the success of other cloning strategies.

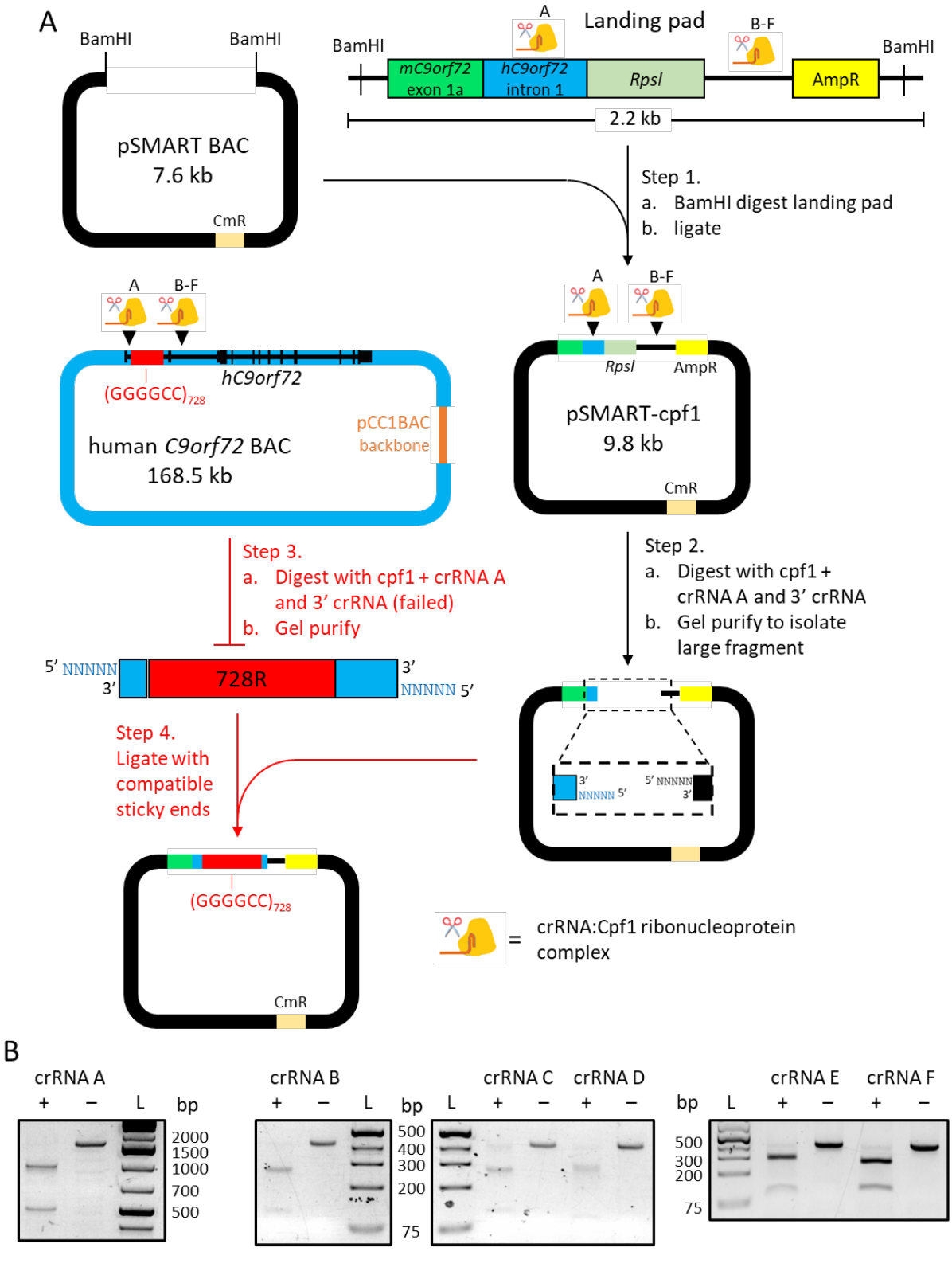



**Figure 6-6 | CRISPR/Cpf1-based strategy to move a large *C9orf72* repeat from a patient-derived BAC into an intermediate vector. (A)** Diagram of our strategy to stabilise a large GGGGCC repeat by moving it from a human BAC into a modified pSMART BAC, which is designed for cloning unstable sequences. Step 1 involved introduction of a 'landing pad' into the pSMART BAC by restriction-ligation cloning. Step 2 involved digestion of pSMART-cpf1 BAC with Cpf1 and a pair of crRNA guides. Cpf1

digestion leaves 4-5 nt sticky ends. Step 3 involved Cpf1 digestion of the repeat BAC with the same crRNA guides as step 2 to isolate the repeat in a fragment containing sticky ends compatible with those in the digested pSMART-cpf1 BAC. This step failed due to inefficient Cpf1 digestion. Step 4, which was not attempted due to failure of step 3, involved sticky-end ligation of the isolated repeat into the digested pSMART-cpf1 BAC. Mouse sequence is depicted in green, human sequence in blue and vector sequence in black or orange. **(B)** Agarose gel electrophoresis images showing Cpf1 cutting efficiency with 6 different crRNAs on gBlock DNA. L, ladder; bp, base pairs.

#### 6.3.2.2.2 Cas9 cloning

We also attempted to use a 'Cas9-Assisted Targeting of CHromosome segments' (CATCH) approach [562] to transfer the repeats into the pSMART BAC. This method combines the programmable nuclease activity of Cas9 with overlap extension PCR and Gibson assembly to seamlessly clone large DNA constructs into cloning vectors. This cloning strategy is outlined in Figure 6-7.

We used a modified pSMART BAC generated in previous cloning attempts (Vector A), which included an insert with a zeocin resistance ($zeo^R$) cassette. We first designed extended primers to amplify Vector A and incorporate 25 bp of overlap with sequence flanking the sgRNA sites 6 and 7 in the human C9orf72 BAC (Table 6-7). We performed long-range PCR to generate an ~8.5 kb amplicon which contained the pSMART BAC backbone, the $zeo^R$ cassette and 25 bp overlaps (Figure 6-7, Step 1). The PCR reaction was digested with DpnI, which cleaves only methylated DNA, to remove residual Vector A. The correct ~8.5 kb amplicon was identified by size selection on PFGE and isolated by gel purification.

Human C9orf72 BAC with repeats was digested with Cas9 using sgRNAs 6 and 7 (Table 6-1). The ~6.7 kb DNA fragment containing the repeat was identified by size selection on PFGE and isolated by gel purification (Figure 6-7, Step 2). The repeat DNA obtained from multiple Cas9 digests were pooled to obtain enough DNA for Gibson assembly. The gel-purified repeat DNA and pSMART BAC amplicon were combined in a Gibson assembly reaction, followed by transformation into Stbl4 cells (Figure 6-7, Step 3). Transformants were plated on LB plates with Chloramphenicol (10 µg/mL). The resulting colonies were screened using colony PCR (Table 6-8); however, carryover of Vector A resulted in colonies, none of which were positive for the repeat insert, in three transformation attempts from

two independent Gibson assembly reactions. Further use of this cloning strategy was limited by the availability of repeat DNA.

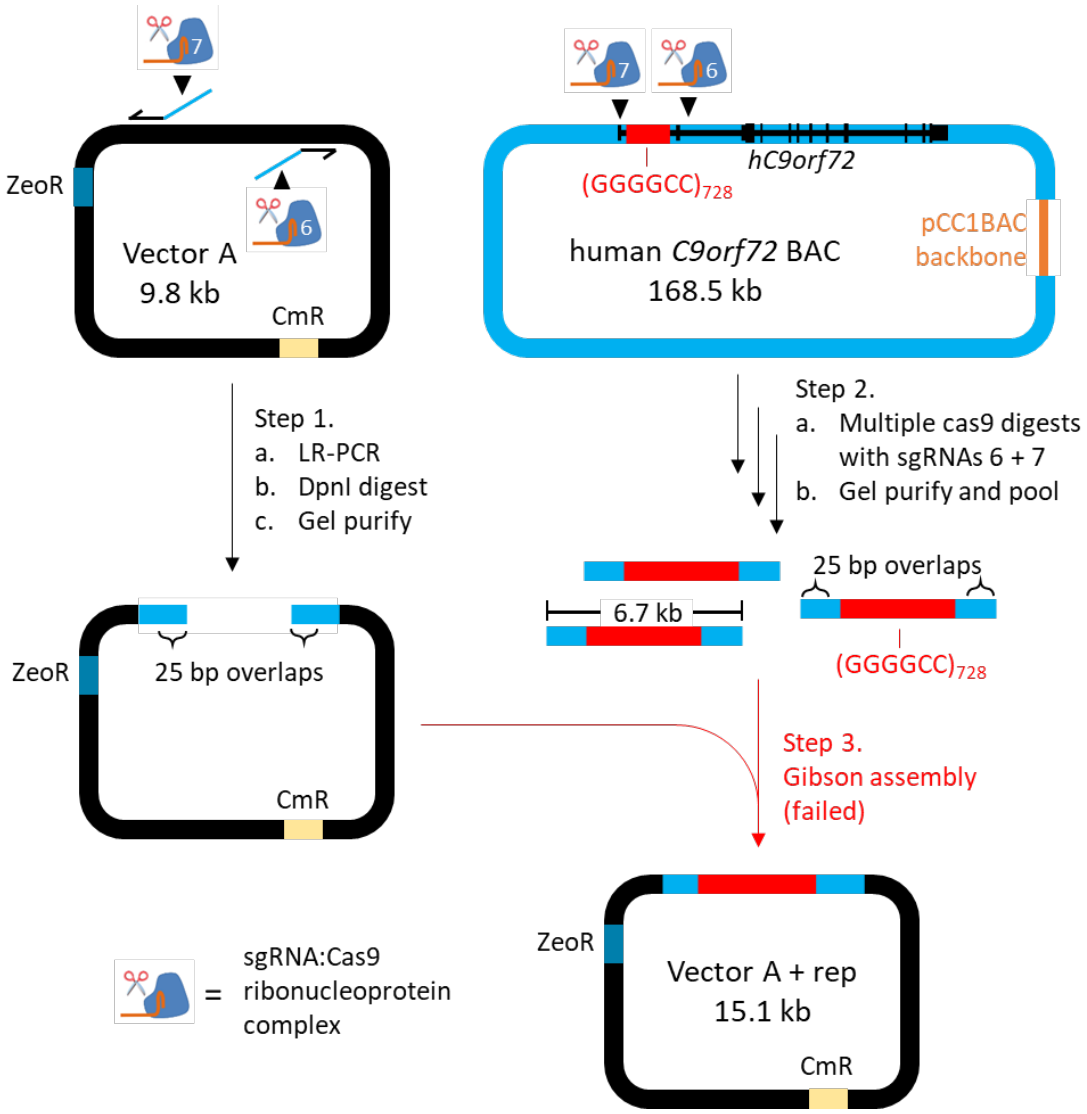



**Figure 6-7 | CRISPR/Cas9-based strategy to move a large *C9orf72* repeat from a patient-derived BAC into an intermediate vector.** This figure describes our 3 step Cas9-based strategy to stabilise a large GGGGCC repeat by moving it from a human BAC into a modified pSMART BAC (Vector A). Step 1 involved long-range PCR with extended primers to incorporate target sequences for sgRNAs 6 and 7 and short regions of flanking human sequence. Step 2 involved Cas9 digestion of human *C9orf72* BAC with sgRNAs 6 + 7 to isolate the 6.7 kb repeat-containing fragment. Multiple Cas9 digests were pooled to obtain sufficient quantities of DNA. Step 3 involved Gibson assembly to clone the isolated repeat-containing fragment into Vector A. Step 3 failed due to carryover of Vector A, resulting in colonies not positive for the repeat. Mouse sequence is depicted in green, human sequence in blue and vector sequence in black or orange. LR-PCR, long-range PCR.

| Primer name | Sequence (5'→ 3') | Binding site | PCR product size |
|---|---|---|---|
| pSmart-Zeo sgRNA6 fwd | ATCTGGCCGCGTTGAACACCACCAGGCTTTGT ATTCAGAAACCCCTCTGTGTTCTGTGGTA | Vector A | ~8.5 kb |
| pSmart-Zeo sgRNA7 rev | ATCACACAGAACGTTTTAATCATTCACCGAGG CTGTCATGGGCTTTGTAG | Vector A | |

**Table 6-7 | Primers for LR-PCR of Vector A.** Sequences homologous to the human C9orf72 gene are shown in blue.

| Primer name | Sequence (5'→ 3') | Binding site | PCR product size |
|---|---|---|---|
| Cpf1 pSmart 5' Fwd | ACATTTGCGGCCTAACTGTG | Vector A backbone | 1192 bp |
| hC9internal_1F | TCAGTACCCGAGCTGTCTCC | human c9orf72 3' of repeat | |
| C9orf72 REP 3' R | CCAGCTTCGGTCAGAGAAAT | human c9orf72 3' of repeat | 360 bp (without repeat). Unable to amplify large repeat. |
| C9orf72 REP 5' F | CTCTCCCCACTACTTGCTCT | human c9orf72 5' of repeat | |

**Table 6-8 | Colony PCR primers to screen for 'Vector A + rep' construct**

#### *6.3.2.3 Strategy C: Creation of a C9orf72 targeting construct using synthesised repeats*

The challenges involved in cloning the large patient-derived BAC hexanucleotide repeat led us to pursue an alternative strategy – de novo-synthesis of the repeat. We employed recursive directional ligation (RDL), a technique which allows cloning of highly repetitive DNA sequences from smaller units. RDL typically utilises type IIs or type III restriction enzymes, which cut asymmetrically outside of their recognition sequence, to build repetitive sequence in an iterative process [543,563]. RDL has previously been employed to build vectors containing up to 103 GGGGCC repeats for expression in drosophila [286]. However, repeats larger than 103 GGGGCC units were unstable and prone to retractions in the pBluescript vector backbone the authors used.

For this cloning we utilised a linear vector – pJAZZ (Lucigen) – based on the coliphage N15 genome, which offers a number of advantages over the pBluescript vector and pSMART BAC [547]. Firstly, the pJAZZ vector does not supercoil, thereby reducing torsional stress and disfavouring secondary structures otherwise induced in circular plasmid vectors. Secondly, the linear nature of the vector

allows more flexibility for cloning as each arm of the vector can be isolated independently following a single cut. Finally, pJAZZ has been shown to stabilise CGG triplet repeats [547], which are highly unstable in bacterial culture, similar to GGGGCC repeats.

#### 6.3.2.3.1 Removal of an existing AarI site from the pJAZZ vector

Our repeat-building strategy is outlined in Figure 6-8. The enzymes used in our RDL strategy were AarI (type IIs restriction enzyme) and XmaI. The pJAZZ backbone already contained an AarI recognition site within the coliphage N15 protelomerase gene, so we first introduced a silent mutation into this site to destroy the AarI site without affecting the peptide sequence of the protelomerase gene. The pJAZZ linear plasmid was digested with PpuMI and SacI in separate single-digest reactions and fragments were isolated by PFGE followed by gel purification (Figure 6-8, Step 1). Next, a 395 bp gBlock was designed which was homologous with the region between the PpuMI and SacI sites, except for a T->A point mutation in the AarI recognition site. The gBlock had ~75 bp overlaps with the pJAZZ fragment ends which were amenable to Gibson assembly. The two pJAZZ fragments and the gBlock were assembled in a Gibson assembly reaction and transformed into DH10B cells (Figure 6-8, Step 2). Transformants were plated on LB plates with 30 µg/mL kanamycin, IPGT (1 mM) and x-gal (20 µg/mL) for blue-white screening. DNA from the resulting blue clones was isolated, screened by AarI digest and sequenced by sanger sequencing (Table 6-9).

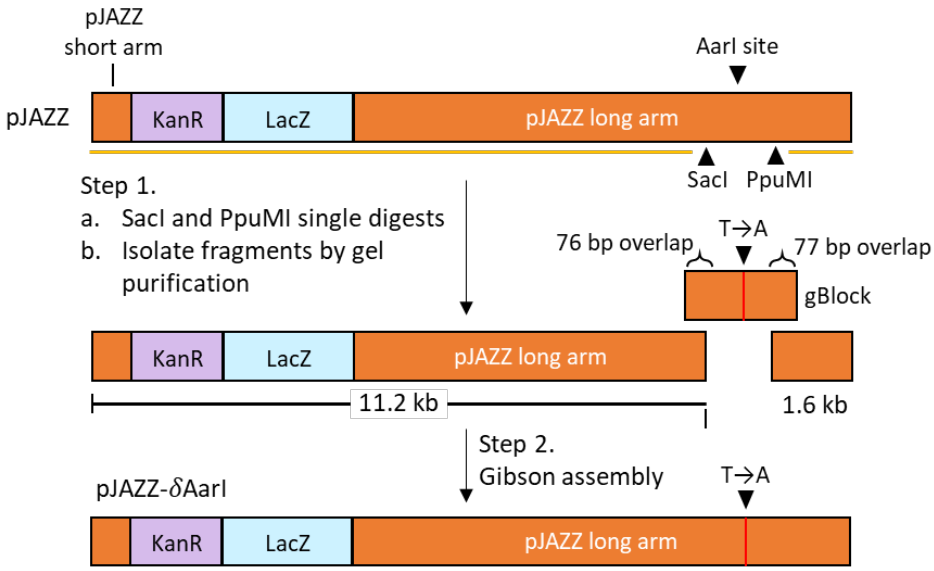



**Figure 6-8 | Strategy for removal of an AarI site from the pJAZZ vector.** Step 1 involved restriction digests followed by gel purification to excise the region of the pJAZZ construct containing the AarI site. Step 2 involved a Gibson assembly reaction between the pJAZZ fragments and a synthetic gBlock homologous to the region surrounding the AarI site, but with a T→A point mutation abolishing the AarI recognition sequence. KanR, Kanamycin resistance cassette.

| Primer name | Sequence (5'→ 3') | Binding site |
|---|---|---|
| pJAZZ AarI seq Fwd 1 | TAGCCGGTACCTGGAGTTTG | pJAZZ-δAarI 5' of AarI site |
| pJAZZ AarI seq Rev 1 | CTCTTCGGTTGGTTCATCGC | pJAZZ-δAarI 3' of AarI site |

**Table 6-9 | Primers for sequencing pJAZZ-δAarI**

#### 6.3.2.3.2 Repeat building up to 1024 GGGGCC repeats by Dr Nair

pJAZZ with the AarI site destroyed (pJAZZ-δAarI) was used by Dr Remya Nair for RDL to build a large repeat (Figure 6-9). In brief, an insert containing 4 GGGGCC repeats (Repeat unit A) was purchased as two imperfectly aligned single stranded oligonucleotides from IDT which were then annealed, giving rise to a 4 nt 5' overhang compatible with the 5' overhang generated by XmaI. This insert was cloned into the pJAZZ vector using XmaI (pJAZZ-δAarI-repA). (Figure 6-9 A, Step 1). Repeat Unit A was designed with recognition sites for AarI (5') and XmaI (3'), which produce compatible 4 nt 5' overhangs when digested. pJAZZ-δAarI-repA was digested with AarI and XmaI in separate reactions, following which the DNA fragment from each reaction that contained the repeat– the pJAZZ long arm in the case of AarI and the pJAZZ short arm in the case of XmaI - was identified by size selection

on agarose gel electrophoresis and isolated by gel purification (Figure 6-9 A, B, Step 2). The two fragments were then seamlessly ligated together, resulting in a construct with 8 uninterrupted GGGGCC repeats (pJAZZ-$\delta$AarI-repB), with the restriction recognition sites only maintained at the ends of the repeat (Figure 6-9 A, B, Step 3). Subsequent rounds of repeat building proceeded in the same manner, using the product from the previous round as the starting material, with the length of the repeat doubling each round. Dr Nair carried out 8 rounds of RDL, resulting in a construct with 1024 repeats. Repeats up to 1024 GGGGCC units in length were found to be stable in the pJAZZ vector, with few retraction events observed.

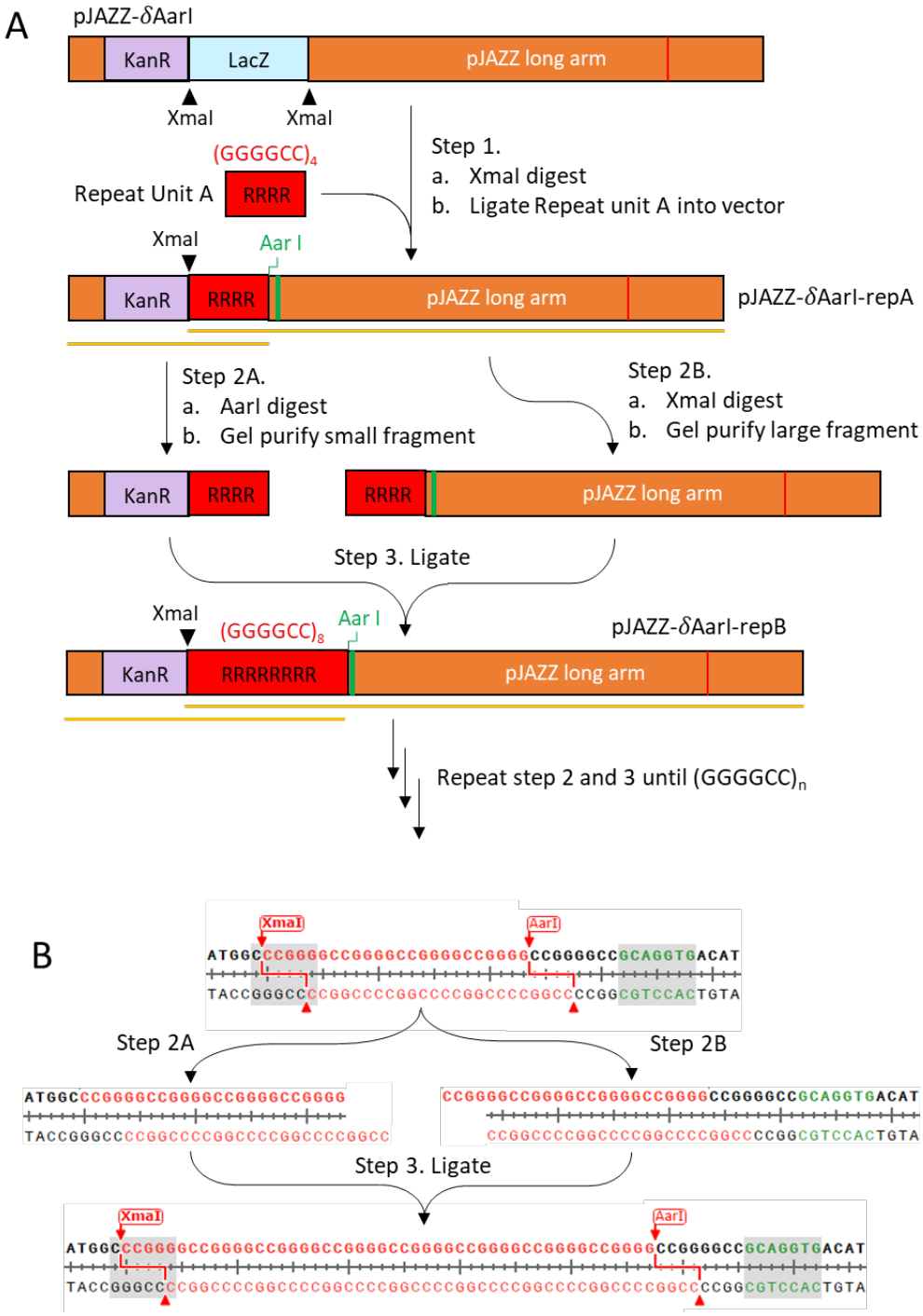



**Figure 6-9 | *De novo* GGGGCC repeat synthesis strategy. (A)** This figure describes our RDL-based strategy for building a large GGGGCC repeat *de novo*. Step 1 involved restriction-ligation cloning to introduce a synthetic DNA fragment containing 4 tandem GGGGCC repeats into the pJAZZ vector. Step 2A and 2B, which were performed in parallel, involve restriction enzyme digests of the pJAZZ-repeat construct, followed by gel purification of the repeat-containing fragment from each reaction. Step 3 involved ligation of the isolated fragments to seamlessly generate a GGGGCC repeat which was double the length of the input repeat. Steps 2 and 3 were repeated 8 times to obtain a construct with 1024 repeats. **(B)** Diagram of the DNA sequence edited in one round of RDL. Repeat DNA is shown in red and the restriction enzyme recognition sequences are highlighted in grey. KanR, Kanamycin resistance cassette; R, single GGGGCC repeat unit.

#### 6.3.2.3.3 Repeat building up to 2048 GGGGCC repeats

I then continued to expand the repeat using the same RDL protocol, generating a construct harbouring 2048 repeats. To generate a construct with 2048 repeats, the pJAZZ construct containing 1024 repeats was digested with AarI and XmaI in separate reactions. The fragment from each digest containing the repeat – 8.4 kb and 16.5 kb for AarI and XmaI respectively – was identified by size selection on agarose gel electrophoresis and isolated by gel purification. The fragments were ligated in a 1:1 ratio using electroligase to form a 24.9 kb construct with 2048 uninterrupted repeats. This construct was transformed into Bigeasy TSA cells and transformants were plated on LB plates with Kanamycin (30 µg/mL). DNA was isolated and screened for repeat size by XbaI digest (Figure 6-10).

The above restriction digests revealed retractions in a large proportion of clones. Unlike retractions in the human C9orf72 repeat-carrying BAC, which were frequently bimodal in nature, retractions in the pJAZZ 2048 repeat construct appeared as a smear of DNA, presumably from a mixed population of partial retractions. In an attempt to increase the proportion of DNA containing ~2048 repeats, we further subcloned this construct twice. In each round of subcloning we screened clones by XbaI digest and chose clones with the largest repeat size to grow further. Ultimately, we succeeded in obtaining 4 clones with ~1800-2048 repeats – as determined by PFGE (Figure 6-10, bottom panel).

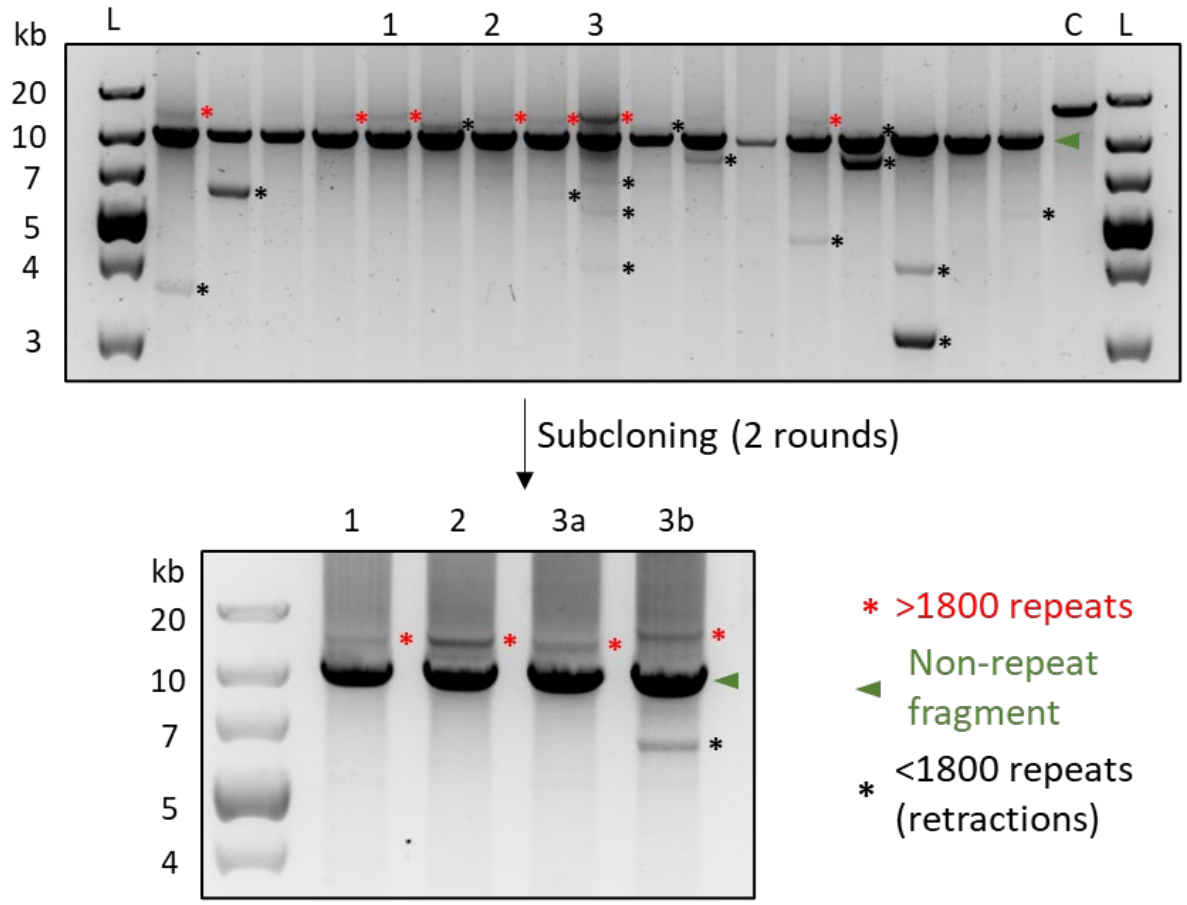



**Figure 6-10 | Screening of large GGGGCC repeats.** Agarose gel electrophoresis images of XbaI digests of pJAZZ constructs containing ~2000 GGGGCC repeats. Numbered putative positive clones in the top panel were subcloned twice to obtain the clones in the bottom panel. The expected fragment size for a clone carrying 2000 repeats is 14.7 kb. The expected fragment size for a fully retracted repeat is 2.7 kb (no fragments of this size were observed). An additional non-repeat-containing fragment appears at 10.2 kb. Red asterisks indicate fragments containing >1800 repeats. Black asterisks indicate fragments containing <1800 repeats. Green arrowhead indicates non-repeat-containing fragment. C, undigested control; L, ladder.

#### 6.3.2.3.4 Generation of a mouse targeting construct by Dr Nair

Considering the instability of constructs containing >1024 repeats, we decided that the final mouse targeting construct should contain 1024 repeats. The final mouse targeting construct was generated by Dr Remya Nair in a complex cloning process (Figure 6-11). The cloning strategy is described in brief below; a detailed description of this cloning is beyond the scope of this thesis.

Firstly, a 6.3 kb segment of the hC9orf72-WT-mHA BAC was transferred into the pJAZZ vector backbone; this segment consisted of mouse *C9orf72* exon 1A, 3.6 kb of wild-type human *C9orf72*

intron 1 (including the repeat locus), and a Hygromycin resistance cassette flanked by PiggyBac ITRs (Figure 6-11, Step 1). Next, a gBlock landing pad was introduced into the wild-type human *C9orf72* gene using Cas9 cloning and Gibson assembly (Figure 6-11, Step 2).The landing pad contained AarI sites which, when cleaved, resulted in 4nt sticky ends (5'-CCGG-3') compatible with the repeat. The 1024x repeat was then seamlessly cloned into this construct using restriction-ligation cloning with AarI (Figure 6-11, Step 3).

Separately, the full length wild-type human *C9orf72* gene with ~5 kb mouse HAs (37.6 kb in total) was cloned from the hC9orf72-WT-mHA BAC into pJAZZ using Cas9 cloning (hC9orf72-WT-mHA-pJAZZ; (Figure 6-11, Step 6).

To incorporate the wider human C9orf72 gene and 5kb mouse HAs into the intermediate repeat-carrying pJAZZ construct, each arm of the repeat-carrying construct was replaced with the equivalent arm from hC9orf72-WT-mHA-pJAZZ (Figure 6-11, Steps 5 and 7-9).

The above cloning resulted in a final mouse targeting construct (hC9orf72-rep-mHA-pJAZZ) consisting of mouse *C9orf72* 5'UTR and noncoding exon 1A, followed by human *C9orf72* sequence from the start of intron 1 through to the TAA stop codon in exon 11, followed by the mouse noncoding region of exon 11 and downstream mouse sequences. The construct harboured 1024 GGGGCC repeats within intron 1 of the humanised *C9orf72* gene and a PiggyBac ITR-flanked Neomycin selection cassette, also within intron 1. The humanised region was flanked by ~5 kb mouse HAs for homologous recombination with the mouse *C9orf72* locus. The correct sequence of the final hC9orf72-rep-mHA-pJAZZ mouse targeting construct was confirmed using next generation Illumina sequencing, performed by Source Bioscience.

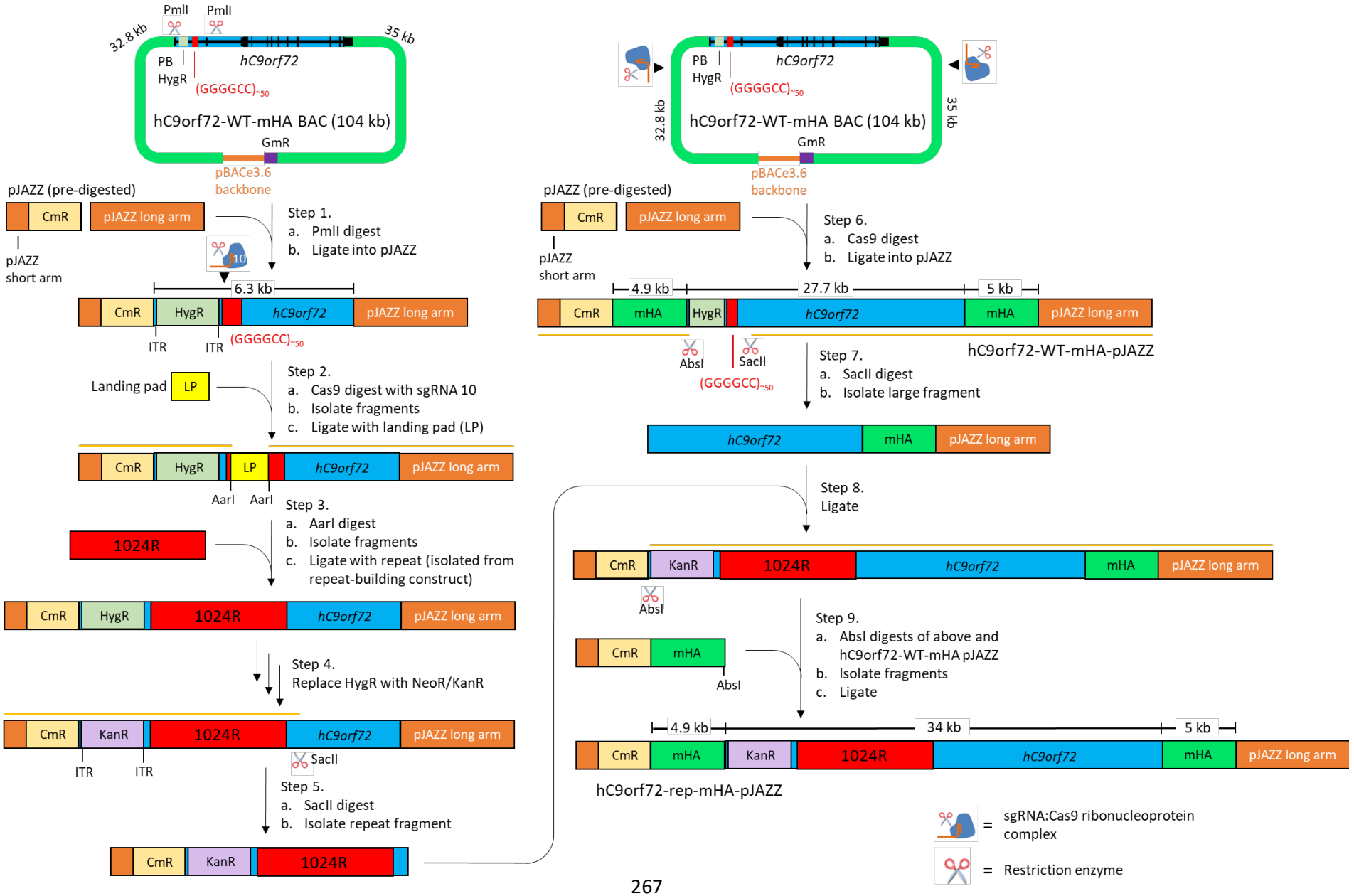

PmlI
PmlI
32.8 kb
35 kb
PB
HygR
hC9orf72
(GGGGCC)~50
hC9orf72-WT-mHA BAC (104 kb)
GmR
pBACe3.6
backbone
pJAZZ (pre-digested)
CmR
pJAZZ long arm
pJAZZ
short arm
Step 1.
a. PmlI digest
b. Ligate into pJAZZ
10
6.3 kb
CmR
HygR
hC9orf72
pJAZZ long arm
ITR
ITR
(GGGGCC)~50
Landing pad
LP
Step 2.
a. Cas9 digest with sgRNA 10
b. Isolate fragments
c. Ligate with landing pad (LP)
CmR
HygR
LP
hC9orf72
pJAZZ long arm
AarI
AarI
1024R
Step 3.
a. AarI digest
b. Isolate fragments
c. Ligate with repeat (isolated from repeat-building construct)
CmR
HygR
1024R
hC9orf72
pJAZZ long arm
Step 4.
Replace HygR with NeoR/KanR
CmR
KanR
1024R
hC9orf72
pJAZZ long arm
ITR
ITR
SacII
Step 5.
a. SacII digest
b. Isolate repeat fragment
CmR
KanR
1024R
PB
HygR
hC9orf72
(GGGGCC)~50
hC9orf72-WT-mHA BAC (104 kb)
GmR
32.8 kb
35 kb
pBACe3.6
backbone
pJAZZ (pre-digested)
CmR
pJAZZ long arm
pJAZZ
short arm
Step 6.
a. Cas9 digest
b. Ligate into pJAZZ
4.9 kb
27.7 kb
5 kb
CmR
mHA
HygR
hC9orf72
mHA
pJAZZ long arm
hC9orf72-WT-mHA-pJAZZ
AbsI
SacII
(GGGGCC)~50
Step 7.
a. SacII digest
b. Isolate large fragment
hC9orf72
mHA
pJAZZ long arm
Step 8.
Ligate
CmR
KanR
1024R
hC9orf72
mHA
pJAZZ long arm
AbsI
CmR
mHA
AbsI
Step 9.
a. AbsI digests of above and hC9orf72-WT-mHA pJAZZ
b. Isolate fragments
c. Ligate
4.9 kb
34 kb
5 kb
CmR
mHA
KanR
1024R
hC9orf72
mHA
pJAZZ long arm
hC9orf72-rep-mHA-pJAZZ
= sgRNA:Cas9 ribonucleoprotein complex
= Restriction enzyme

**Figure 6-11 | Strategy for the generation of a mouse *C9orf72* targeting construct, performed by Dr Nair.** Mouse sequence is depicted in green, human sequence in blue. Diagram not drawn to scale. 1024R, 1024 tandem GGGGCC repeats; LP, landing pad; mHA, mouse homology arm; CmR, chloramphenicol resistance cassette; HygR, hygromycin resistance cassette; KanR, kanamycin resistance cassette; GmR, gentamycin resistance cassette; ITR, PiggyBac inverted terminal repeat.

### 6.3.3 Generation of a humanised *C9orf72* mouse by Dr Ben Davies

The hC9orf72-rep-mHA-pJAZZ mouse targeting construct was sent to Dr Ben Davies and colleagues at the Wellcome Centre for Human Genetics, Oxford to perform the ESC targeting and microinjection. Editing of naïve ESCs to introduce the *hC9orf72* allele was achieved by CRISPR/Cas9 assisted homology directed repair (HDR) using the hC9orf72-rep-mHA-pJAZZ construct as a DNA donor (Figure 6-12).

Briefly, hC9orf72-rep-mHA-pJAZZ and CRISPR-Cas9 RNPs were introduced into the C57Bl/6NTac mouse ESC line JM8.F6 by electroporation. Putative positive transformants were selected for neomycin resistance by growth in ESC media supplemented with neomycin. Repeat size was estimated by Southern blotting, performed by Dr Nair (data not shown). *hC9orf72* mice were then generated by microinjection of modified ESCs into C57BL/6N albino donor blastocysts. Chimeric male offspring which were then crossed to C57BL/6N females, with the offspring genotyped to screen for germline transmission (GLT) of the humanised *hC9orf72* allele. Once GLT was obtained, sperm from heterozygous *hC9orf72* males was used for IVF with C57BL/6NTac females to import the line into the Mary Lyon Centre at MRC Harwell. The offspring are currently being crossed with mice expressing PiggyBac transposase to excise the neomycin selection cassette.

Quality control of the *hC9orf72* mice is ongoing at the time of writing and includes targeted locus amplification and short-read sequencing, Nanopore long-read sequencing, and southern blotting to determine the size of the repeat.

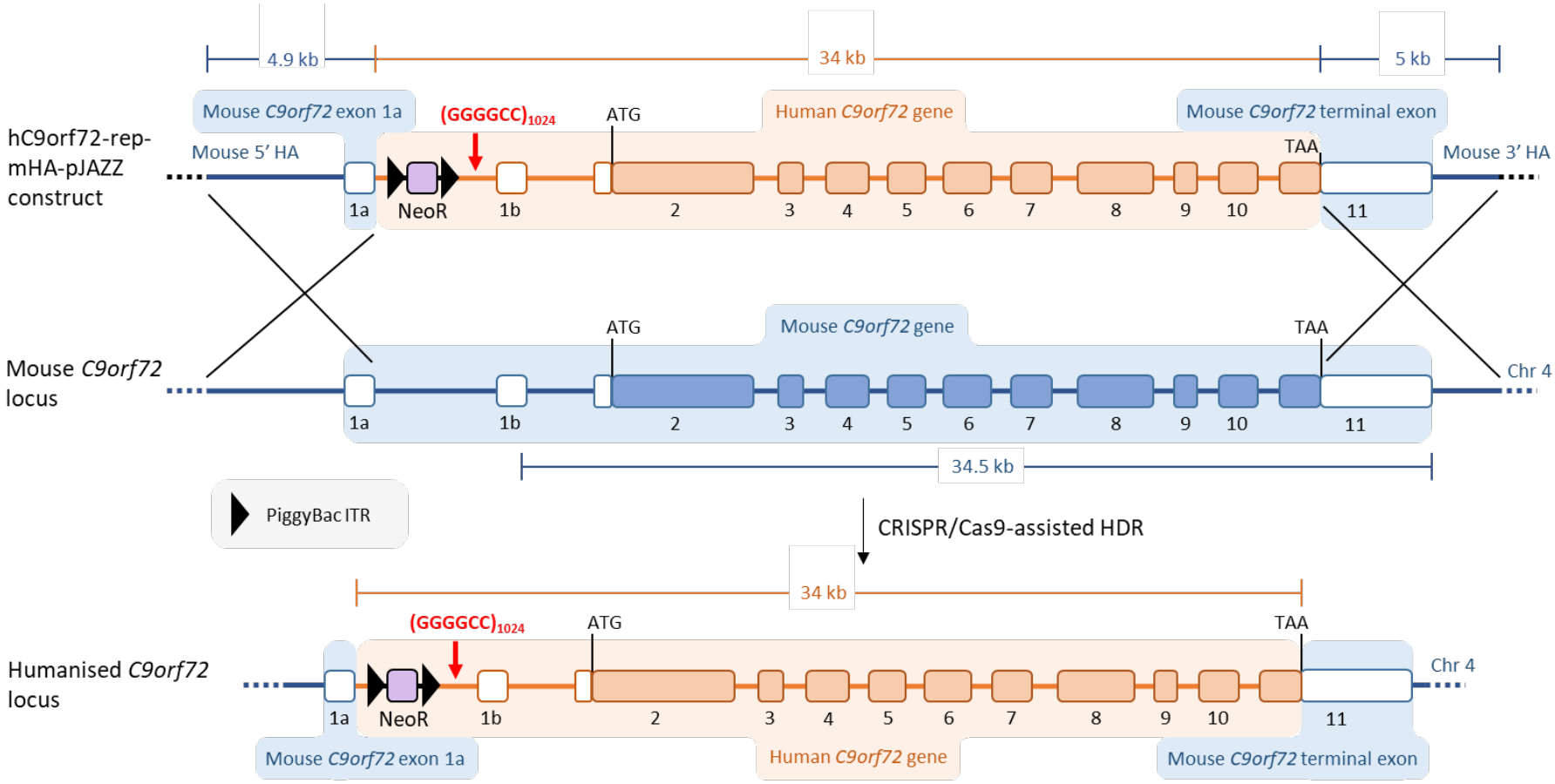



**Figure 6-12 | Humanisation of the mouse *C9orf72* locus by CRISPR/Cas9-assisted homology-directed repair (HDR).** The mouse *C9orf72* locus was humanised from the start of intron 1 through to the TAA stop codon in exon 11. 34.5 kb of mouse sequence was replaced by 34 kb of human sequence (including 6.1 kb of GGGGCC repeat sequence). All coding exons were humanised. White rectangles indicate non-coding exons; filled rectangles indicate coding exons. A PiggyBac ITR-flanked neomycin selection cassette within human intron 1 was used for selection of transformed ESC clones, and will be removed by crossing to mice expressing PiggyBac transposase.

### 6.3.4 Creation of a targeting construct to introduce a large *C9orf72* repeat expansion into iPSCs

The purpose of this cloning was to modify our existing constructs to generate a targeting construct for creating C9orf72-repeat knock-in iPSCs. This construct was designed to contain the 1024 GGGGCC repeats flanked by short human HAs (1 kb 5' and 2.9 kb 3').

A DNA fragment containing 1.9 kb of the human C9orf72 gene, without repeats, was generated by Dr Remya Nair through Cas9 digestion of the human C9orf72 BAC (2kb hC9). I ligated this fragment into the pJAZZ vector using clonesmart DNA ligase, and transformed the ligation product into bigeasy TSA cells (Figure 6-13 A, Step 1). Transformants were plated on LB plates with 30 µg/mL kanamycin, IPGT (1mM) and x-gal (20µg/mL) for blue-white screening. White colonies were selected and screened for the correct insert by XbaI digest, the resulting vector was named 'pJAZZ+2kb hC9'.

Next, pJAZZ+2kb hC9 was digested with Cas9 using sgRNA 10, which cuts 2 bp into the 5' end of the repeat. The small 3.3 kb fragment, containing the 1 kb 5' HA, was then identified by size selection on PFGE and isolated by gel purification (Figure 6-13 A, Step 2).

A second construct – pJAZZ+1024rep Hygro – was generated by Dr Remya Nair as an intermediate step in the creation the mouse targeting construct, containing 1024 GGGGCC repeats and a 2.9 kb 3' human HA, but no 5' HA. pJAZZ+1024rep Hygro was digested with Cas9 using sgRNA 10, resulting in a small 5 kb fragment and a large 19.4 kb repeat-containing fragment (Figure 6-13 A, Step 3). The Cas9 digestion mix was heat inactivated.

The small 'pJAZZ+2kb hC9' fragment and the DNA products of the 'pJAZZ+1024rep Hygro' Cas9 digestion were ligated using clonesmart DNA ligase and transformed into bigeasy TSA cells (iPSC-1024rep) (Figure 6-13 A, Step 4). Transformants were plated on LB plates with 30 µg/mL kanamycin. DNA was isolated from clones and screened by XbaI, KflI, and BamHI-HF digests. DNA from a putative positive clone was sequenced. Non-repeat regions of the construct were sequenced by Sanger

sequencing, including the break point at the 5’ end of the repeat (Figure 6-13 B). Sanger sequencing quality in the 3’ HA was poor due to high GC content. The correct sequence was confirmed by Nanopore sequencing of the entire construct performed by Dr Matthew Parker at the University of Sheffield (data not shown). Repeat length was confirmed by Nanopore sequencing and restriction digestion (Figure 6-13 C, D). In addition to the full-length repeat, Nanopore sequencing detected a retracted repeat (~300 repeat units) in approximately 1/3 of reads (Figure 6-13 C). However, Nanopore sequencing is known to sequence shorter repeats with higher efficiency and hence retracted DNA likely makes up a smaller population than estimated by sequencing. This observation is supported by restriction digests where retracted DNA was not detectable (Figure 6-13 D).

This iPSC targeting construct has been sent to our collaborators at The Jackson Laboratory and KCL, who will perform the iPSC targeting.

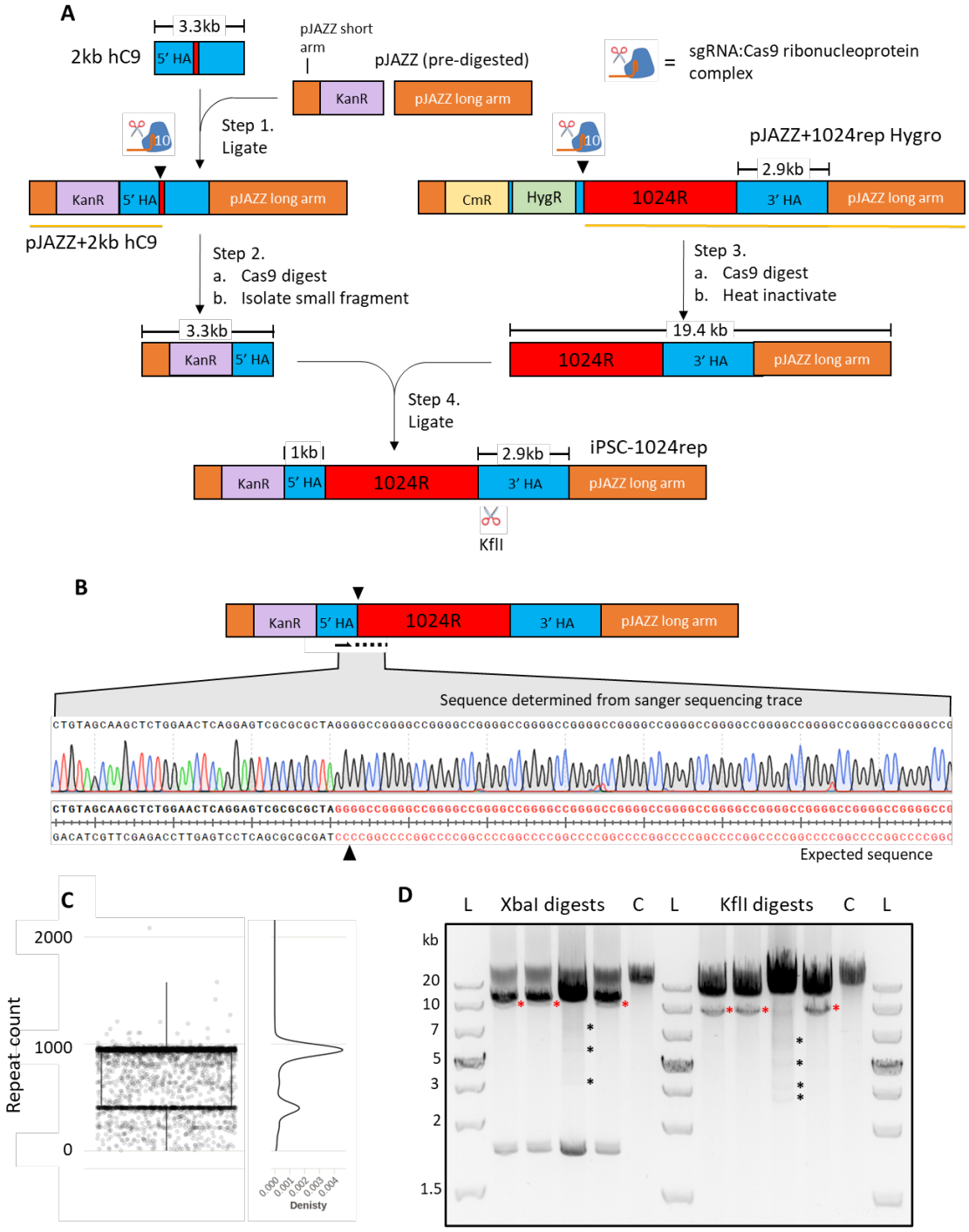



**Figure 6-13 | Generation of a human iPSC *C9orf72* targeting construct. (A)** Strategy for the generation of a human iPSC *C9orf72* targeting construct. CmR, chloramphenicol resistance cassette; HygR, hygromycin resistance cassette; KanR, neomycin/kanamycin resistance cassette; HA, homology arm; R, GGGGCC repeat unit. **(B)** Representative sanger sequencing trace from across the break point (2bp into the 5′ end of the repeat). Arrow indicates break point resulting from Cas9 digestion. GGGGCC repeat sequence is shown in red. **(C)** Graph of repeat lengths obtained from Nanopore sequencing of the construct. A retraction from 1000 to ~300 repeat units is evident in a subset of DNA. **(D)** Agarose gel electrophoresis images of DNA from iPSC-1024rep clones digested with XbaI and KflI. The expected size of the released fragment from a construct containing 1024 GGGGCC repeats is 10034 bp for XbaI digests and 9810 bp for KflI digests. Red asterisks indicate bands from constructs harbouring 1024 repeats, black asterisks indicate large retractions. C, undigested control; L, ladder.

## 6.4 Discussion

### 6.4.1 Cloning

The work in this chapter describes the creation of targeting constructs to introduce large C9orf72 GGGGCC repeats into mouse ESCs and human iPSCs. This cloning was a high-risk multi-year project which required multiple strategies to be run in parallel to maximise the chance of success.

Ultimately, our RDL-based strategy was successful. Two previous studies have utilised RDL for building GGGGCC repeats [286,546]. Both these studies used BspQI and EcoO109I as the restriction enzymes for RDL, with a maximum of 156 repeats cloned in a stable manner. Importantly, the constructs generated were cloning and expression vectors which contained the repeat but lacked any surrounding genomic sequence. Our work represents a significant advance over these existing studies for two reasons: 1) We have grown and maintained far larger GGGGCC repeats (up to ~2000 repeats). 2) We have created constructs containing the repeat flanked by sequence from the endogenous genomic locus, allowing the constructs to be used as templates for homology-directed-repair (HDR) to generate knock-in models. To achieve this we completely redesigned the RDL protocol, changing the vectors and enzymes to allow the large repeat to be grown, stabilised and inserted into the wild-type human C9orf72 gene in a scarless manner.

Our findings emphasise the utility of the pJAZZ vector for cloning tandem repeats. Another group has previously cloned tandem FMR1 CGG repeats, up to 660 bp in size, and non-repetitive inserts up to 30 kb in size using the pJAZZ vector [547]. The 2000 GGGGCC repeats cloned in the present work constitute 12 kb of 100% GC sequence, an 18-fold increase in repeat length compared to the original study. Constructs harbouring ~2000 repeats were highly unstable, even in the pJAZZ vector, and multiple rounds of sub-cloning were required to obtain clones carrying the full-length repeat. Hence, 12 kb of GGGGCC repeats likely represents the upper limit of the pJAZZ vector, although the retention of other tandem repeat sequences is likely to vary. Regardless, the majority of repeat

expansion disorders are caused by repeats <12 kb in length [564]; therefore, this size limit is unlikely to be issue for modelling other repeat expansion disorders using pJAZZ.

Our repeat building strategy can be applied to other unstable tandem-repeat DNA sequences, although the choice of enzymes impose some restrictions. Our laboratory is already utilising the pJAZZ construct and RDL to build large FMR1 CGG repeats for the generation of novel mouse models of Fragile X syndrome (FXS).

### 6.4.2 Further use of constructs

#### *6.4.2.1 Humanised C9orf72-ALS/FTD mouse model*

Our mouse targeting construct carries the entire human *C9orf72* gene from the start of intron 1 through to the TAA stop codon, including 1024 GGGGCC repeats, enabling humanisation of the mouse *C9orf72* gene and introduction of the repeat in a single ESC targeting step. The rationale for using genomic humanisation to study ALS is detailed in section 1.5.2.2.3. In the case of *C9orf72*, genomic humanisation is particularly desirable because loss-of-function mechanisms may contribute to C9orf72-ALS/FTD disease pathology [238], something that is difficult to model in existing BAC transgenic C9orf72-ALS/FTD mouse models.

We have generated a fully humanised *hC9orf72* mouse model, with quality control ongoing at the time of writing. Following confirmation of correct targeting of the *hC9orf72* locus the next step will be to phenotype *hC9orf72* mice.

#### *6.4.2.2 C9orf72-ALS/FTD iPSC lines*

To date, studies investigating C9orf72-ALS/FTD in human iPSCs have utilised cells derived from C9orf72-ALS/FTD patients [269,565,566]. Large phenotypic variability exists between C9orf72-ALS/FTD patients [567,568] and, therefore, it is unsurprising that large phenotypic variability exists between different patient-derived iPSC lines. iPSC lines from patients vary widely in their genomic background and capacity for differentiation, making it challenging to determine if a biological effect is driven by

ALS-relevant disease processes or other genomic differences. This variability necessitates the use of stringent controls; the control iPSC lines used in studies are sometimes derived from age/sex-matched healthy control participants [269,270,569], however the modern 'gold standard' is to generate "isogenic" control lines through deletion of the expanded repeat via CRISPR-Cas9 gene editing [570,571].

In theory, these isogenic iPSC pairs should circumvent the issue of variability by allowing the comparison of expanded vs unexpanded GGGGCC repeats in the same genetic context. In practice, off-target CRISPR-Cas9 editing events and chromosomal abnormalities resulting from the expansion of edited clones can go undetected and cause erroneous results [572].

In light of these challenges, a number of groups are now pursuing an alternative approach; the generation of multiple mutant iPSC lines from well-characterised reference lines [572,573]. This approach allows multiple mutations to be compared in the same genetic context to uncover common disease pathways, while reducing the probability of off-target editing events through standardised quality control assays [572].

Disease-causing point mutations in other familial ALS genes (e.g. SOD1, FUS, TDP-43) are relatively straightforward to engineer using CRISPR/Cas9; however, introduction of a large C9orf72 repeat is far more challenging due to the large size and instability of the repeat-carrying HDR template. Through stabilising the repeat, our human iPSC targeting construct makes introduction of a large C9orf72 repeat feasible, and thus provides a new utility for the study of C9orf72-ALS/FTD. This construct has been shared with our collaborators – Professor Bill Skarnes and Dr Michael Ward at the iNDI project, and Dr Sarah Mizielinska and Dr Marc-David Ruepp at King's College London – who are performing the iPSC targeting.

# Chapter 7: Discussion

This thesis describes the creation and characterisation of novel humanised mouse models of SOD1-ALS and C9orf72-ALS/FTD. The rationale for humanisation is based on the need for mouse models with increased physiological relevance to help improve the translation of preclinical research into successful treatments for disease. The central research questions of this project are:

1. What are the physiological effects of humanisation of SOD1 in mice?
2. What phenotypes do humanised mouse models of SOD1-ALS develop?
3. What phenotypes do humanised mouse models of C9orf72-ALS/FTD develop?

The answer to these questions will guide the use of humanised ALS mouse models and enable their use as an effective tool for ALS researchers. The work presented in this thesis goes some way towards addressing these questions. Question 1 has been addressed by the study of phenotypes in *hSOD1*$^{WT}$ mice. Question 2 has been partially addressed by the phenotypic characterisation of *hSOD1*$^{A4V}$ mice and the creation of *hSOD1*$^{G93A}$ mice, and future work will examine the phenotypes in *hSOD1*$^{G93A}$ mice. Question 3 has begun to be addressed through the creation of a *hC9orf72* mouse model, with future work to involve quality control and characterisation of this novel mouse strain.

## 7.1 Summary of key findings

Through the characterisation of a humanised mouse model of *hSOD1*$^{WT}$ in chapter 3, this project demonstrated that humanisation of *SOD1* is well-tolerated in mice. Despite retention of the mouse promotor in *hSOD1*$^{WT}$ mice, we unexpectedly observed decreased expression of human *SOD1* in skeletal muscle and liver, but not in CNS tissues. Lower bodyweight was the only phenotype observed in *hSOD1*$^{WT}$ mice, which may be indicative of SOD1 LOF in peripheral tissues in these animals. *hSOD1*$^{WT}$ mice did not display any of the behavioural deficits or indicators of peripheral axonopathy observed in *Sod1*$^{-/-}$ null mice [48].

Characterisation of the humanised SOD1-ALS model *hSOD1*$^{A4V}$ revealed mild metabolic phenotypes without overt neurodegeneration. There was a lack of behavioural phenotypes, although detection of phenotypes in aged animals was hampered by low n numbers. Very few transcriptional changes were observed in aged *hSOD1*$^{A4V}$ animals. We observed an inconsistent NMJ denervation phenotype in TA muscle which may result from reduced NMJ functionality in aged *hSOD1*$^{A4V}$ mice. SOD1 activity was markedly decreased in all tissues as a result of the A4V mutation; hence, the observed metabolic and NMJ denervation phenotypes may be a result of SOD1 LOF.

In chapter 5, I described the creation of a *hSOD1*$^{G93A}$ mouse model, a genomically humanised counterpart to the widely used transgenic *tgSOD1*$^{G93A}$ mouse [338]. The construct used for gene targeting was created through editing of an existing construct. Quality control of the humanised *hSOD1*$^{G93A}$ mouse strain demonstrated the expected seamless humanisation of the mouse *Sod1* locus and introduction of the SOD1-G93A mutation.

In chapter 6, I described our efforts to clone a large C9orf72 GGGGCC repeat to be used in the generation of a humanised *hC9orf72* mouse line and novel C9orf72-ALS/FTD iPSC lines. We used RDL to generate constructs containing up to 2048 GGGGCC repeats. Constructs containing 1024 repeats were used to create targeting constructs for mouse ESCs and human iPSCs. Quality control of the *hC9orf72* mice is ongoing at the time of writing. Targeting of human iPSCs is ongoing at the time of writing.

## 7.2 Knock-in mouse models are not sufficient to model SOD1 GOF pathology

The assumption when beginning this project was that humanised models of SOD1-ALS would have mild, but detectable phenotypes, and that these would be relevant for human disease due to the physiological nature of humanised models.

This assumption was based on findings from the physiological (but not humanised) ENU mutant *Sod1*$^{D83G}$ strain, which displayed a number of ALS-relevant phenotypes, including MN loss and

behavioural phenotypes [354]. Similarly, knock-in mice with mutations in *Tardbp* [355] and partially humanised truncated *Fus* [361] display pathological phenotypes. We also have unpublished data from our own group showing mild phenotypes in humanised *FUS*$^{P525L}$ and *FUS*$^{Q519IFS}$ mice. These studies show that mutations in ALS genes can have pathogenic effects even at physiological expression levels.

The existence of an existing *tgSOD1*$^{A4V}$ transgenic model [71,338] was not considered detrimental to our hypothesis, as very little work had been done on this strain, and it had not been probed with sensitive modern molecular biology techniques. We believed that *hSOD1*$^{A4V}$ mice would recapitulate similar phenotypes to those noted in *Sod1*$^{D83G}$ mice, but with the added dimensions of humanisation and the A4V mutation – which causes particularly aggressive disease in humans. However this was not the case, and the phenotypes in *hSOD1*$^{A4V}$ mice were exceedingly mild, with no clear SOD1 GOF phenotypes. The discrepancy between phenotypes in *Sod1*$^{D83G}$ and *hSOD1*$^{A4V}$ mice likely result from differences in biochemistry between mouse and human SOD1, and from differences between the effects of the D83G and A4V mutations on SOD1 toxicity and function.

SOD1 LOF seems particularly important for the phenotypes noted in knock-in models of SOD1-ALS. This is supported by observations in four knock-in SOD1-ALS strains created by Dr Robert Brown and colleagues (Table 1-3). Of these four new strains, only *Sod1*$^{G85R}$ [372] and *Sod1*$^{H71Y}$ (unpublished) mice show phenotypes, whereas *Sod1*$^{A4V}$ and *Sod1*$^{G93A}$ mice do not (unpublished). A key difference between these strains is that the G85R and H71Y mutations completely abolish Sod1 enzymatic function, whereas the A4V and G93A mutations do not. The phenotypes observed in *Sod1*$^{G85R}$ mice are limited to homozygous animals, and have been primarily attributed to Sod1 LOF [372]. Similarly to *hSOD1*$^{A4V}$ mice, none of these four strains show significant SOD1 GOF phenotypes.

Sod1$^{D83G}$ also exhibits complete loss of enzymatic function, and Sod1 LOF phenotypes were observed in *Sod1*$^{D83G}$ mice [354]. However, Joyce et al. also identified some phenotypes they attributed to Sod1 GOF – such as MN loss – with GOF phenotypes being defined as phenotypes which do not appear in

*Sod1*$^{-/-}$ mice. The finding of MN loss in *Sod1*$^{D83G}$ mice is an outlier compared to other knock-in models; 23% loss of LMNs was noted at 15 weeks of age, however this remained stable until at least 52 weeks of age (the oldest age tested) suggesting an atypical pattern of MN degeneration without the progression seen in other SOD1-ALS mouse models.

Our data, and the findings of Dr Robert Brown and colleagues, suggest that knock-in mouse models are not sufficient to model SOD1 GOF pathology. This should inform further studies of humanised SOD1-ALS mice moving forward. However, we cannot be certain this statement is true for all SOD1 mutations, and questions remain regarding Sod1 GOF phenotypes in *Sod1*$^{D83G}$ mice.

## 7.3 Why isn't the *SOD1*$^{A4V}$ mutation as severe in mice as in humans?

Considering the importance of SOD1 LOF to the phenotype of SOD1 knock-in mice, and the fact that *hSOD1*$^{A4V}$ mice retain 15-25% SOD1 activity in the CNS, it is perhaps unsurprising that phenotypes in *hSOD1*$^{A4V}$ mice were exceedingly mild. However, this still does not explain *why* no SOD1 GOF phenotypes were observed in *hSOD1*$^{A4V}$ mice, especially considering the severity of the SOD1-A4V mutation in humans [452–454]. There are 3 possible factors that could explain differing severity of *SOD1*$^{A4V}$ in humans and mice: 1. Age, 2. species-specific biochemistry and 3. anatomy.

### 7.3.1 Age

Age is the primary risk factor for ALS. It is possible that knock-in mice don't live long enough to develop severe SOD1 GOF phenotypes. As discussed in section 1.5.1, ALS is a late onset disease and symptom onset is often preceded by years or decades of presymptomatic pathogenic changes [335,336]. Considering the average age of disease onset in SOD1-A4V patients is 47 years old [454], pathological GOF effects would have to manifest some 24 times faster to see disease phenotypes in mice by 2 years of age, and that would only represent the start of symptomatic disease.

If we consider age as the main factor preventing *hSOD1*$^{A4V}$ mice from exhibiting overt GOF phenotypes, then it is possible that these animals are in a pre-symptomatic prodromal state. This is

supported by our findings of mild metabolic disturbances in *hSOD1$^{A4V}$* mice, which although are more likely to result from SOD1 LOF, may have some aspect of SOD1 GOF pathology. We have seen similar indications of a prodromal state in humanised *hFUS$^{P525L}$* mice, which show some mild motor symptoms without loss of motor neurons (unpublished observations).

It is important to note that SOD1-ALS patients carry SOD1 mutations from birth and do not exhibit symptoms until mid-late life. However, at the point of diagnosis, SOD1-ALS disease progression is typically rapid, the pathology has reached a critical point and runaway mechanisms are in play. It is clear that, for whatever reason, *hSOD1$^{A4V}$* mice never reach the point of runaway disease. This is also the case in *tgSOD1$^{A4V}$* mice (founder line A1073), which carry 4.7 ± 0.4 transgene copies [71,338], indicating that substantially higher expression of *SOD1$^{A4V}$* would be required to accelerate the development SOD1 GOF phenotypes such that phenotypes are detectable within the lifetime of the mouse.

### 7.3.2 Species-specific biochemistry

It is possible that the A4V mutation has differential effects on the SOD1 interactome in mice vs in humans. It has been shown that different SOD1 mutations substantially alter the interactions between SOD1 and other proteins [574]. Mutant SOD1 has a number of abnormal interaction partners, and it is certainly conceivable that these abnormal interaction partners may be different in mice and humans due to differences in the proteome. On the other hand, some abnormal interaction partners of mutant SOD1 are clearly conserved between mice and humans; for example, the abnormal interaction of mutant SOD1 with Derlin-1 is observed in both human cell lines and in mice [171,575].

Transgenic SOD1-ALS mouse models with low copy number are rare, making it difficult to determine whether SOD1-A4V leads to particularly mild phenotypes compared to other SOD1 mutants. However, low copy substrains of *tgSOD1$^{G93A}$* and *tgSOD1$^{L126delTT}$* mice have similar transgene copy numbers to *tgSOD1$^{A4V}$* mice, allowing for comparison. Mice carrying 4 copies of the *SOD1$^{G93A}$* transgene (VLE *tgSOD1$^{G93A}$*) show prolonged survival and variable motor phenotypes which are only

~80% penetrant by 24 months old. Approximately 18% of VLE *tgSOD1*$^{G93A}$ mice reached end stage disease [345]. Transgenic *tgSOD1*$^{L126delTT}$ mice carrying 4 or 5 copies of the *SOD1*$^{L126delTT}$ transgene show survival deficits – average survival was 372.4 and 480.2 days for 4 and 5 transgene copies, respectively – and severe motor symptoms including hindlimb paralysis, whereas mice carrying 1 or 3 transgene copies show no symptoms [370].

The phenotypes in *tgSOD1*$^{G93A}$ and *tgSOD1*$^{L126delTT}$ mice carrying 4 transgene copies are more severe than *tgSOD1*$^{A4V}$ mice, which do not show any survival deficits or overt motor phenotypes [71], suggesting that the A4V mutation is less severe than G93A or L126delTT in mice. This is curious as the A4V mutation is one of the most severe *SOD1* mutations in humans [452–454]. Taken together, these observations suggest that some undetermined features of mouse physiology cause SOD1$^{A4V}$ to be less severe than some other SOD1 mutants.

A key caveat to this analysis is that it is based on transgene copy number which, due to epigenetic factors, does not always correlate well with gene expression. SOD1 protein expression has been reported in these strains, but interpretation is complicated by the differing stability of the different SOD1 mutants. For accurate comparison, measurement of *SOD1* mRNA expression relative to mouse *Sod1* is required, however this has not been reported for *tgSOD1*$^{A4V}$, VLE *tgSOD1*$^{G93A}$ or *tgSOD1*$^{L126delTT}$ mice.

### 7.3.3 Anatomy

There are key differences in MN anatomy between humans and mice which may influence their susceptibility to degeneration. Beyond the clear differences in length (sciatic MNs can be longer than 1 meter in humans, but only extend a couple of centimetres in mice), MN soma volume and dendritic tree size are estimated to be up to 10 times larger in humans than mice [576–579]. A recent transcriptomics study has linked large MN size and the expression of cytoskeletal genes to their selective vulnerability in ALS, suggesting size is a key factor for degeneration [579]. Mouse MNs are also much more excitable than their human counterparts [578]. Additionally, human NMJs have been

shown to be morphologically distinct from mouse NMJs – human NMJs are smaller and more fragmented – and significant differences in the synaptic proteome are observed between our species [580].

MNs have substantial energy requirements, and yet have limited energy stores, requiring constant provision of nutrients for ATP production [581]. This high metabolic demand leaves MNs vulnerable to both energetic stress, and oxidative stress due to high ROS production [140,581,582]. These issues are likely exacerbated in human MNs due to their extremely large size, and may contribute to the relatively lower pathogenicity of *SOD1*$^{A4V}$ in mice. We know that differences in MN anatomy do not completely prevent mice from getting ALS – this has been demonstrated in several transgenic SOD1-ALS mouse models – however, it may make them less susceptible to the disease.

## 7.4 The suitability of mice as a model of SOD1 loss-of-function (LOF)

*SOD1* is an essential gene in human cell lines [583] and iPSCs [584], and patients homozygous for the *SOD1* truncating mutation p.C112Wfs*11 (SOD1$^{C112X\text{-}Hom}$) present with an infantile-onset motor neuron disease distinct from ALS [221–223], indicating an essential role for SOD1 in the development and maintenance of the motor system. Conversely, *Sod1* is nonessential in mice, being one of only 19 human genes essential *in vitro* that are nonessential in mouse [585], and *Sod1*$^{-/-}$ mice exhibit relatively mild phenotypes [48]. It is difficult to make a direct comparison between the severity of phenotypes in SOD1$^{C112X\text{-}Hom}$ patients and *Sod1*$^{-/-}$ mice, however a broad examination of the phenotypes suggests that SOD1 LOF is more severe in humans.

The largest study of 8 SOD1$^{C112X\text{-}Hom}$ patients found UMN and LMN dysfunction; cerebellar, brain stem and fronto-temporal atrophy; a haematological phenotype consisting of anaemia and reduced red blood cell (RBC) survival; and developmental phenotypes, including low set ears and overlapping toes [223]. The predominant presentation of CNS phenotypes from around 8 months of age indicates a particular vulnerability of the human nervous system to SOD1 LOF. Some evidence also suggests a

survival deficit in SOD1$^{C112X\text{-}Hom}$ patients; all SOD1$^{C112X\text{-}Hom}$ patients identified to date are less than 14 years old, and the sibling of one patient presenting with a similar disease course died at age 6 [223].

*Sod1*$^{-/-}$ mice also exhibit predominantly CNS symptoms, with a phenotype consisting of peripheral axonopathy without motor neuron loss [445,446,451,586]. This is accompanied by motor phenotypes and tremor starting at around 13 months of age [439,446]. *Sod1*$^{-/-}$ mice also have an accelerated decline in cognition, evident at the earliest tested time point of 13 months old, which has been attributed to an accelerated aging phenotype [587]. Anaemia and reduced RBC survival have also been reported in *Sod1*$^{-/-}$ mice [529]. *Sod1*$^{-/-}$ mice have approximately 30% shorter lifespan, a difference attributed to increased incidence of hepatocellular carcinoma, rather than CNS phenotypes [588,589].

Although the symptoms of SOD1 LOF affect similar physiological systems in humans and mice, they occur relatively earlier and are markedly more severe in humans. In light of this, the question arises: why is SOD1 LOF more severe in humans?

The phenotypes in SOD1 SOD1$^{C112X\text{-}Hom}$ patients may result from cellular damage caused by excess ROS, which would be consistent with the established role of SOD1 as a key antioxidant enzyme [32]. In support of this hypothesis, Park et al. observed lower levels of GSH in RBCs, accompanied by reduced activity of Gpx1, which may indicate increased oxidative stress in these cells [223]. Additionally, fibroblasts from a SOD1$^{C112X\text{-}Hom}$ patient could not be grown under standard culture conditions due to extreme oxygen sensitivity, and required low-oxygen growth conditions and supplementation with ascorbate [221].

Ascorbate has high reactivity with superoxide [590,591], and is a key antioxidant with an established protective role against oxidative stress in the CNS [493,592,593]. Hence, ascorbate may represent an antioxidant defence mechanism of particular importance in the absence of SOD1; indeed, this was proposed following the study of a SOD1$^{C112X\text{-}Hom}$ patient [221]. One difference between the mouse and human antioxidant defence systems is that mice can synthesise ascorbate, predominantly in the liver

[594], whereas humans cannot. This is due to loss of the terminal enzyme in the ascorbate biosynthetic pathway (gulonolactone oxidase) in humans, making ascorbate an essential dietary component [595,596]. Ascorbate is found at especially high concentrations in neurons [494], and neurons appear to be particularly vulnerable to ascorbate deficiency [592], potentially because of their high ROS production and sensitivity to oxidative stress [140,581,582].

It may be possible that, in mice, endogenous ascorbate synthesis compensates for the lack of SOD1, thereby making the effects of SOD1 LOF less severe. We found blood ascorbate levels to be unchanged in *hSOD1$^{A4V}$* mice, however this is somewhat unsurprising as ascorbate levels in cells and biofluids are subject to tight homeostatic regulation [597] and increased ascorbate synthesis may be balanced by increased uptake into tissues. Blood ascorbate levels were similarly unchanged in SOD1$^{C112X\text{-}Hom}$ patients [223].

The influence of ascorbate metabolism is just one possible reason for the discrepancy in SOD1 LOF phenotypes between humans and mice. It is also possible that other antioxidant systems, such as glutathione, play a differing role between our two species. Glutathione metabolism was found to be altered in RBCs from SOD1$^{C112X\text{-}Hom}$ patients [223] and in the liver of *Sod1$^{-/-}$* mice [495], indicating that SOD1 LOF has an effect on this system. We also found glutathione metabolism to be perturbed in *hSOD1$^{A4V}$* mice; however, we noted a change in the blood GSH/GSSG ratio indicative of lower oxidative stress. This result was the opposite of what was expected; *hSOD1$^{A4V}$* mice have lower levels of a key antioxidant enzyme, and yet appear to have *less* oxidative stress. However, blood ascorbate and GSH/GSSG do not necessarily reflect the redox state in all tissues, and so we have planned a tissue metabolomics experiment to examine metabolism and redox state in individual tissues, including the SC and liver.

It is also possible that other differences between human and mouse biochemistry or anatomy are responsible for the differing severity of SOD1 LOF phenotypes. For example, the larger size and high

metabolic requirements of human MNs may render them more susceptible to oxidative damage resulting from SOD1 LOF.

Regardless of the exact mechanism, the severe phenotype of SOD1$^{C112X\text{-}Hom}$ patients suggests that SOD1 LOF plays a more important role in humans than SOD1 mouse models suggest, and warrants further study. Accordingly, oxidative stress may have a greater impact on pathology in human SOD1-ALS than mouse models suggest. This influence of ascorbate metabolism could be investigated further by crossing *hSOD1$^{A4V}$* (or *Sod1$^{-/-}$*) mice to *gulo$^{-/-}$* mice, which lack gulonolactone oxidase and hence are unable to synthesise ascorbate [598]. Crossing humanised SOD1-ALS mice to *gulo$^{-/-}$* mice may help to bring the severity of murine SOD1 LOF phenotypes more in line with those in humans.

## 7.5 Metabolic phenotypes precede neurodegeneration in *hSOD1$^{A4V}$* mice

Numerous studies have implicated metabolic dysfunction in the aetiology of ALS. Lower body mass index (BMI) is associated with faster disease progression and decreased survival in ALS patients [599–602] and ALS risk is reduced in high BMI individuals [603,604]. Hypermetabolism is observed in approximately 50% of ALS patients [605,606] with metabolic changes being observed early in disease and even prior to symptom onset [477,601].

Our results support the hypothesis that metabolic changes are a very early phenomenon in SOD1-ALS, as we observed mild metabolic changes in *hSOD1$^{A4V}$* mice despite no overt neurodegeneration. The changes we observed included lower bodyweight and fat mass in female *hSOD1$^{A4V}$* mice, disturbed glucose metabolism in mice of both sexes, a number of altered metabolites in plasma from mice of both sexes, and mild transcriptional changes related to metabolism in the SC of aged male mice (females not examined). None of these changes are individually compelling, but together describe a mild metabolic phenotype in *hSOD1$^{A4V}$* mice. This metabolic phenotype requires further investigation to identify changes in specific metabolic pathways, and to determine whether these changes are a result of SOD1 LOF or GOF effects.

One caveat to our findings was that metabolic phenotypes were only detected in *hSOD1*$^{A4V}$ mice on a C57BL/6J background and not a mixed C57BL/6J x 129S1/SvImJ background, a difference which remains unexplained.

We did observe one potential neurodegenerative phenotype, which was loss of functionality of TA NMJs in aged *hSOD1*$^{A4V}$ mice, but further experiments are needed to confirm this result.

## 7.6 New methods for cloning repeat expansions and unstable DNA

Cloning of a large GGGGCC repeat expansion was a difficult but not intractable challenge. We ultimately succeeded in cloning GGGGCC DNA up to 2000 repeat units in length using RDL and CRISPR cloning in combination with the pJAZZ linear vector.

Our approach is broadly applicable to cloning of other repeat DNA sequences, although the limited number of type IIs and type III restriction enzymes impose some restrictions on the sequences that can be cloned. The development of this approach has resulted in a number of collaborations from groups interested in our C9orf72 targeting constructs, as well as groups interested in applying our repeat cloning method to other repeat expansions.

## 7.7 Future research into the phenotypes of humanised *hSOD1* and *hC9orf72* mice

There are a number of avenues for further research to address questions raised in this study and complete the study aims. Below are some key areas.

### 7.7.1 *hSOD1*$^{A4V}$ mice

#### *7.7.1.1 Metabolic phenotypes*

Our findings indicate a mild metabolic phenotype in *hSOD1*$^{A4V}$ mice. We performed metabolomics analysis on plasma, however this approach does not reveal the metabolic changes occurring in individual tissues, which will be important for understanding the contribution of different tissues to disease pathogenesis. To investigate this further we have planned a tissue metabolomics study using SC, TA muscle and liver from aged *hSOD1*$^{A4V}$ mice.

Elevation of xylulose in plasma from *hSOD1*$^{A4V}$/*hSOD1*$^{A4V}$ mice was among the strongest phenotypes seen in this study, but is difficult to explain. Changes in ROS and the balance of redox couples NAD+/NADH and NADP+/NADPH can affect oxidoreductase activity and synthesis of xylulose [607]. Considering the emerging role of SOD1 in modulating redox homeostasis [39,608] then an investigation of the redox balance in *hSOD1*$^{A4V}$ mice, particularly in the liver, may be an avenue to explain the overproduction of xylulose.

#### *7.7.1.2 Molecular phenotypes*

We examined SOD1 aggregation in the lumbar SC of aged *hSOD1*$^{A4V}$ mice by immunohistochemistry, which lacks the sensitivity to detect low levels of misfolded or aggregated SOD1. We also observed background staining which impacted detection. To investigate SOD1 misfolding/aggregation in greater detail, we plan to use a more sensitive ELISA assay to detect misfolded/aggregated SOD1 in tissues from 18-month-old *hSOD1*$^{A4V}$ mice.

We performed bulk RNA-seq on the lumbar SC of *hSOD1*$^{A4V}$ mice. Future studies could probe neuron-specific transcriptional changes using single-cell RNA-seq, or bulk RNA-seq of MNs isolated by laser capture microdissection. Additional insights could also be gained by re-analysing the existing bulk RNA-seq data using two-way ANOVA with age and genotype as factors (possible in the EdgeR software package) rather than pairwise analysis. This would allow *hSOD1*$^{A4V}$ mice of different ages to be analysed together, and hence increase the power to detect subtle changes in gene expression. For gene ontology analysis, a tissue-specific list of background genes could be used, rather than all mouse genes, which could potentially increase the power to detect significantly enriched GO terms.

#### *7.7.1.3 Neurodegenerative phenotypes*

We observed an inconsistent NMJ phenotype which may be due to a reduction in TA NMJ functionality in aged *hSOD1*$^{A4V}$ mice. The first step to investigate this further would be to replicate our existing findings using internalised labelled $H_CT$ as an axonal marker. If our findings of NMJ denervation in aged animals are replicated, then TA NMJ functionality could be probed further by

double labelling with antibodies – to visualise the axons and presynaptic termini – and $H_CT$, to determine if there are structurally intact NMJs that fail to internalise $H_CT$ to a reduction in function.

Further studies could also attempt to look at changes in NMJ structure in aged *hSOD1*$^{A4V}$ mice using NMJ-analyser [518]; however, resolving NMJs to the level of detail required for structural analysis may be challenging due to high background staining in tissues from older mice.

Although we did not detect any motor phenotypes in *hSOD1*$^{A4V}$ mice, our longitudinal study allows for the examination of interrelationships between phenotypes. For example, it would be possible to investigate if there is a correlation between metabolic phenotypes (e.g. bodyweight, fat mass, IPGTT results) and motor phenotypes (e.g. grip strength, rotarod performance), and determine if these correlations vary between the genotypes. Such analyses are complex but could uncover novel associations between phenotypes, and genotype-specific differences that are not possible to detect when each test is evaluated separately.

### 7.7.2 *hSOD1*$^{G93A}$ mice

Phenotypic characterisation of *hSOD1*$^{G93A}$ mice is planned. Initial characterisation will consist of examination of SOD1 levels and activity, and basic behavioural and motor phenotyping. We expect that *hSOD1*$^{G93A}$ animals will display a lack of SOD1 GOF phenotypes and, considering that SOD1$^{G93A}$ retains enzyme activity, it is also likely that LOF effects will be mild. Future studies with *hSOD1*$^{G93A}$ animals should focus on molecular and metabolic phenotypes, as it is very unlikely that disease-relevant behavioural phenotypes will manifest.

### 7.7.3 *hC9orf72* mice

We have planned TLA and short-read sequencing, southern blotting (to determine repeat size) and nanopore sequencing of the *hC9orf72* locus. Assuming that these final quality control measures confirm correct targeting of the *hC9orf72* locus and the presence of a reasonable size repeat, the next step will be to phenotype *hC9orf72* mice. Considering the lack of neurodegenerative

phenotypes in the majority of C9-BAC models (Table 1-4), it is likely that the phenotype of *hC9orf72* mice will also be mild, and hence phenotyping efforts should focus on molecular and metabolic phenotypes.

## 7.8 Potential applications of humanised *SOD1* mice

Despite the lack of strong phenotypes in *hSOD1*$^{A4V}$ mice, and presumably also in *hSOD1*$^{G93A}$ mice, there are still several applications for humanised SOD1 mouse models that could be pursued. Potential uses fall into 2 groups:

1. Exploring ways to make the phenotype more aggressive.
2. Using the models in their current form to study early-stage disease processes and the pharmacological/genetic manipulation of human SOD1 *in vivo*.

### 7.8.1 Approaches to increase the severity of phenotypes in humanised *SOD1* mice

#### *7.8.1.1 Stacking mutations*

One possible approach to exacerbate the phenotype of humanised *SOD1* mice is to combine multiple *SOD1* mutations together in a single humanised gene. A similar method has been used in amyloid precursor protein (APP) knock-in mice to model Alzheimer's disease, whereby the addition of more mutations into *APP* worsens the phenotype. However, these mutations do not affect the sequence of the pathological amyloid β peptide (Aβ), but rather increase the expression of Aβ and the ratio of $A\beta_{42}$ to $A\beta_{40}$ [609]. However, in the case of *SOD1* these mutations would affect the coding sequence. This approach may increase the probability of severe phenotypes developing within the lifetime of the mouse, depending on the toxicity of the *SOD1* mutant created; however, it would also reduce the relevance to human disease by creating novel *SOD1* mutants not observed in human ALS patients.

#### 7.8.1.2 Crosses to *Gulo*$^{-/-}$ knockout mice

As described in section 7.4, mice can synthesise the antioxidant free radical scavenger ascorbate, which may responsible for the reduced impact of SOD1 LOF in mice vs humans. Hence, one way to exacerbate SOD1 LOF phenotypes in humanised *SOD1* mice may be to inhibit their ability to synthesise ascorbate by crossing to *Gulo*$^{-/-}$ knockout mice, which lack L-gulono-γ-lactone oxidase, the rate-limiting enzyme in the biosynthesis of ascorbate [598].

#### 7.8.1.3 Manipulation of genetic background

It is well established that genetic background modulates the phenotypes of SOD1-ALS mouse models [501–504]. Studies in *tgSOD1*$^{G93A}$ mice suggest that a C57BL/6J background is among the most protective against degeneration, whereas ALR, NOD.Rag1KO, FVB/NJ are the most susceptible backgrounds [502]. It is possible that crossing humanised *SOD1* mice onto one of these three backgrounds may exacerbate the phenotypes. On the other hand, crossing *hSOD1*$^{A4V}$ mice onto a mixed C57BL/6J x 129S1/SvImJ background – which we believed would be more susceptible than C57BL/6J – did not exacerbate the phenotype (although we did not phenotype animals beyond 12 months old), and instead improved some of the metabolic phenotypes.

#### 7.8.1.4 Manipulation of environmental factors

The development of ALS is believed to be a multistep process, whereby multiple 'hits' are required to overwhelm cells' capacity to cope with pathological changes and cause disease [335,336]. Potential hits include genetic risk factors, environmental risk factors, and time.

The approaches described thus far relate to altering genetic risk factors, however future studies with *hSOD1*$^{A4V}$ mice could also explore modulating environmental factors to introduce additional hits and exacerbate any phenotypes. Environmental factors that could be modulated include: diet – caloric restriction has been shown to reduce the lifespan of *tgSOD1*$^{G93A}$ mice [610,611], whereas high-fat and high-energy diets are protective [612–614]; sleep stress – circadian rhythm dysfunction has been found

to worsen phenotypes in *tgSOD1*$^{G93A}$ mice [615]; physical activity – higher lifetime physical activity has been shown in some studies to be risk factor for ALS [335,616,617].

### 7.8.2 Potential uses of humanised *SOD1* mice in their current form

#### *7.8.2.1 Testing of antisense oligonucleotides (ASOs)*

ASOs targeting SOD1 represent a promising avenue for the treatment of SOD1-ALS. The SOD1 ASO Tofersen was recently granted accelerated approval by the FDA [224], has been given a positive recommendation by the EMA [225], and is under review by NICE [226]. Current SOD1 ASOs, including Tofersen, reduce overall SOD1 levels and are not selective for the mutant protein. More marked silencing may increase the efficacy of SOD1 ASOs, but runs the risk of increasing adverse effects due to SOD1 LOF. Hence, the development of ASOs that are specific for the mutant protein is highly desirable.

Heterozygous *hSOD1*$^{WT}$/*hSOD1*$^{A4V}$ mice represent an ideal model to test the efficiency and specificity of *SOD1*$^{A4V}$-specific ASOs, as these mice mimic the genotype of SOD1$^{A4V}$ patients. We have recently sent *hSOD1*$^{A4V}$ animals to a collaborator for this purpose. *hSOD1*$^{WT}$ mice are also a useful model to investigate the ASO silencing of total SOD1, as the efficiency of silencing can now be investigated in a *in vivo* model which expresses human SOD1 at physiologically relevant levels.

#### *7.8.2.2 Investigation of prion-like mechanisms in SOD1-ALS*

As described in section 1.3.2.2, prion-like mechanisms may play a role in the spreading of SOD1 pathology. Previous *in vivo* investigations into the seeding of SOD1 pathology have utilised transgenic SOD1 mouse models, particularly *tgSOD1*$^{G85R}$ [77] and *tgSOD1*$^{G85R\text{-}YFP}$ mice [69]. It is unclear how the level of transgene expression and the structural properties of the different SOD1 mutants impact templated misfolding of mutant SOD1 *in vivo*. The *hSOD1*$^{A4V}$ and *hSOD1*$^{G93A}$ mice described in this thesis may prove useful models to investigate the templated misfolding of SOD1$^{A4V}$ and SOD1$^{G93A}$, especially as they express mutant SOD1 at physiologically relevant levels.

In addition to the templated misfolding of mutant SOD1, WT SOD1 may undergo templated misfolding and contribute to disease. A previous study in our group attempted to investigate prion-like templated misfolding of WT human SOD1 using transgenic *tgSOD1$^{WT}$* mice; however, the overexpression of SOD1 in these animals lead to high levels of SOD1 misfolding at baseline, and additional misfolding induced by seeding was impossible to detect. *hSOD1$^{WT}$* mice may be a better model system to investigate this area, as the physiological level of human SOD1 expression should reduce background misfolding and enable detection of seeded misfolding.

#### *7.8.2.3 Conditional expression of SOD1$^{G93A}$*

In *hSOD1$^{G93A}$* mice, the expression of *SOD1$^{G93A}$* is conditional – the allele can be converted from *SOD1$^{WT}$* to *SOD1$^{G93A}$* following Cre-Lox recombination. In theory, this will enable rescue experiments and studies involving tissue-specific expression of mutant SOD1$^{G93A}$. In practice, the utility of this conditional strain will depend on whether *hSOD1$^{G93A}$* mice develop overt phenotypes which can then be rescued.

## 7.9 Conclusions

This thesis has presented the creation and characterisation of two novel humanised SOD1 mouse models, *hSOD1$^{WT}$* and *hSOD1$^{A4V}$*, and the creation of a further two humanised ALS mouse models *hSOD1$^{G93A}$* and *hC9orf72*, which are yet to undergo characterisation.

Experiments with *hSOD1$^{WT}$* mice demonstrated that humanisation of SOD1 is well-tolerated. I have presented an exploratory analysis of *hSOD1$^{A4V}$* mice, including behavioural, molecular and metabolic phenotyping. We observed mild metabolic disturbances in *hSOD1$^{A4V}$* mice which may be of relevance to ALS, although further work is needed to explore the metabolic changes in more detail.

This work has highlighted the challenges involved in the modelling of SOD1-ALS using mice. Physiological levels of expression of mutant SOD1 are not sufficient to see overt SOD1 GOF phenotypes within the lifetime of the mouse. Yet overexpression of mutant SOD1 at the level

required to see SOD1 GOF phenotypes likely also gives rise to artefactual phenotypes. This is a particularly intractable problem which highlights the importance of using a range of mouse models to study disease.

Mice may not be the best model organism to study SOD1 LOF, as the recent discovery of patients with no SOD1 activity demonstrate that SOD1 LOF results in a severe neuromotor disorder, lending strength to the existing argument that SOD1 LOF contributes to pathology in SOD1-ALS [48]. The impact of SOD LOF is markedly more severe in humans than in mice. The mild metabolic phenotypes in *hSOD1*$^{A4V}$ mice, and lower bodyweight in *hSOD1*$^{WT}$ mice, may be attributed to SOD1 LOF, although more work is required to investigate the relative impact of SOD1 LOF and GOF mechanisms in humanised *SOD1* mice.

Humanised mouse models are not intended to replace transgenic models, as each class of model has different advantages and drawbacks. They will instead form part of the toolbox available to researchers; for use in answering specific questions. The phenotype of humanised *hSOD1*$^{A4V}$ mice was extremely mild, and there is little reason to believe that humanised *hSOD1*$^{G93A}$ mice will be significantly more severe; hence the questions answerable with these models relate to early-stage disease processes, and the study or pharmacological/genetic manipulation of human SOD1 *in vivo*.

# References


1. Chiò, A. *et al.* Global Epidemiology of Amyotrophic Lateral Sclerosis: A Systematic Review of the Published Literature. *Neuroepidemiology* **41**, 118–130 (2013).

2. Magnussen, M. J. & Glass, J. D. Chapter 2 - Natural History of Amyotrophic Lateral Sclerosis. in *Molecular and Cellular Therapies for Motor Neuron Diseases* (eds. Boulis, N., O'Connor, D. & Donsante, A.) 25–41 (Academic Press, 2017). doi:10.1016/B978-0-12-802257-3.00002-X.

3. Picher-Martel, V., Brunet, F., Dupré, N. & Chrestian, N. The Occurrence of FUS Mutations in Pediatric Amyotrophic Lateral Sclerosis: A Case Report and Review of the Literature. *J. Child Neurol.* **35**, 556–562 (2020).

4. Lehky, T. & Grunseich, C. Juvenile Amyotrophic Lateral Sclerosis: A Review. *Genes* **12**, 1935 (2021).

5. Talbot, K. Motor neuron disease: THE BARE ESSENTIALS. *Pract. Neurol.* **9**, 303–309 (2009).

6. Fujimura-Kiyono, C. *et al.* Onset and spreading patterns of lower motor neuron involvements predict survival in sporadic amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **82**, 1244–1249 (2011).

7. Bang, J., Spina, S. & Miller, B. L. Frontotemporal dementia. *The Lancet* **386**, 1672–1682 (2015).

8. Ringholz, G. M. *et al.* Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **65**, 586–590 (2005).

9. Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in als and FTD: Disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).

10. Abramzon, Y. A., Fratta, P., Traynor, B. J. & Chia, R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front. Neurosci.* **14**, (2020).

11. Saeed, M. *et al.* Age and founder effect of SOD1 A4V mutation causing ALS. *Neurology* **72**, 1634–1639 (2009).

12. Wang, H., Guan, L. & Deng, M. Recent progress of the genetics of amyotrophic lateral sclerosis and challenges of gene therapy. *Front. Neurosci.* **17**, (2023).

13. Al-Chalabi, A. *et al.* An estimate of amyotrophic lateral sclerosis heritability using twin data. *J. Neurol. Neurosurg. Psychiatry* **81**, 1324–1326 (2010).

14. Mejzini, R. *et al.* ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Front. Neurosci.* **13**, (2019).

15. Zou, Z.-Y. *et al.* Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J. Neurol. Neurosurg. Psychiatry* **88**, 540–549 (2017).

16. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* **17**, 17–23 (2014).

17. Hutton, M. *et al.* Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**, 702–705 (1998).

18. Baker, M. *et al.* Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature* **442**, 916–919 (2006).

19. Cruts, M. *et al.* Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature* **442**, 920–924 (2006).

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

21. Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72**, 257–268 (2011).

22. Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).

23. Kwiatkowski, T. J. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).

24. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).

25. Rutherford, N. J. *et al.* Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet.* **4**, e1000193 (2008).

26. Sreedharan, J. *et al.* TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science* **319**, 1668–1672 (2008).

27. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).

28. Hergesheimer, R. C. *et al.* The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain* **142**, 1176–1194 (2019).

29. Strange, R. W. *et al.* The Structure of Holo and Metal-deficient Wild-type Human Cu, Zn Superoxide Dismutase and its Relevance to Familial Amyotrophic Lateral Sclerosis. *J. Mol. Biol.* **328**, 877–891 (2003).

30. Arnesano, F. *et al.* The Unusually Stable Quaternary Structure of Human Cu,Zn-Superoxide Dismutase 1 Is Controlled by Both Metal Occupancy and Disulfide Status*. *J. Biol. Chem.* **279**, 47998–48003 (2004).

31. McCord, J. M. & Fridovich, I. Superoxide Dismutase: AN ENZYMIC FUNCTION FOR ERYTHROCUPREIN (HEMOCUPREIN). *J. Biol. Chem.* **244**, 6049–6055 (1969).

32. Fridovich, I. Superoxide radical and superoxide dismutases. *Annu. Rev. Biochem.* **64**, 97–112 (1995).

33. Corson, L. B., Strain, J. J., Culotta, V. C. & Cleveland, D. W. Chaperone-facilitated copper binding is a property common to several classes of familial amyotrophic lateral sclerosis-linked superoxide dismutase mutants. *Proc. Natl. Acad. Sci.* **95**, 6361–6366 (1998).

34. Tokuda, E. & Furukawa, Y. Copper Homeostasis as a Therapeutic Target in Amyotrophic Lateral Sclerosis with SOD1 Mutations. *Int. J. Mol. Sci.* **17**, 636 (2016).

35. Wei, J.-P. J., Srinivasan, C., Han, H., Valentine, J. S. & Gralla, E. B. Evidence for a Novel Role of Copper-Zinc Superoxide Dismutase in Zinc Metabolism. *J. Biol. Chem.* **276**, 44798–44803 (2001).

36. Agbas, A. *et al.* Activation of brain calcineurin (Cn) by Cu-Zn superoxide dismutase (SOD1) depends on direct SOD1–Cn protein interactions occurring in vitro and in vivo. *Biochem. J.* **405**, 51–59 (2007).

37. Wang, X., Culotta, V. C. & Klee, C. B. Superoxide dismutase protects calcineurin from inactivation. *Nature* **383**, 434–437 (1996).

38. Reddi, A. R. & Culotta, V. C. SOD1 Integrates Signals from Oxygen and Glucose to Repress Respiration. *Cell* **152**, 224–235 (2013).

39. Tsang, C. K., Liu, Y., Thomas, J., Zhang, Y. & Zheng, X. F. S. Superoxide dismutase 1 acts as a nuclear transcription factor to regulate oxidative stress resistance. *Nat. Commun.* **5**, 3446 (2014).

40. Li, X. *et al.* A new function of copper zinc superoxide dismutase: as a regulatory DNA-binding protein in gene expression in response to intracellular hydrogen peroxide. *Nucleic Acids Res.* **47**, 5074–5085 (2019).

41. Bunton-Stasyshyn, R. K. A., Saccon, R. A., Fratta, P. & Fisher, E. M. C. SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology: New and Renascent Themes. *The Neuroscientist* (2014) doi:10.1177/1073858414561795.

42. Wang, J., Xu, G. & Borchelt, D. R. Mapping superoxide dismutase 1 domains of non-native interaction: roles of intra- and intermolecular disulfide bonding in aggregation. *J. Neurochem.* **96**, 1277–1288 (2006).

43. Seetharaman, S. V., Taylor, A. B., Holloway, S. & Hart, P. J. Structures of Mouse SOD1 and Human/Mouse SOD1 Chimeras. *Arch. Biochem. Biophys.* **503**, 183–190 (2010).

44. The Human Protein Atlas. https://v15.proteinatlas.org/.

45. Marklund, S. L. Extracellular superoxide dismutase in human tissues and human cell lines. *J. Clin. Invest.* **74**, 1398–1403 (1984).

46. Lindenau, J., Noack, H., Possel, H., Asayama, K. & Wolf, G. Cellular distribution of superoxide dismutases in the rat CNS. *Glia* **29**, 25–34 (2000).

47. Crapo, J. D., Oury, T., Rabouille, C., Slot, J. W. & Chang, L. Y. Copper,zinc superoxide dismutase is primarily a cytosolic protein in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10405–10409 (1992).

48. Saccon, R. A., Bunton-Stasyshyn, R. K. A., Fisher, E. M. C. & Fratta, P. Is SOD1 loss of function involved in amyotrophic lateral sclerosis? *Brain* **136**, 2342–2358 (2013).

49. ALSoD. https://alsod.ac.uk/.

50. Shaw, B. F. & Valentine, J. S. How do ALS-associated mutations in superoxide dismutase 1 promote aggregation of the protein? *Trends Biochem. Sci.* **32**, 78–85 (2007).

51. Prudencio, M. & Borchelt, D. R. Superoxide dismutase 1 encoding mutations linked to ALS adopts a spectrum of misfolded states. *Mol. Neurodegener.* **6**, 77 (2011).

52. Lindberg, M. J., Byström, R., Boknäs, N., Andersen, P. M. & Oliveberg, M. Systematically perturbed folding patterns of amyotrophic lateral sclerosis (ALS)-associated SOD1 mutants. *Proc. Natl. Acad. Sci.* **102**, 9754–9759 (2005).

53. Sato, T. *et al.* Rapid disease progression correlates with instability of mutant SOD1 in familial ALS. *Neurology* **65**, 1954–1957 (2005).

54. Wang, Q., Johnson, J. L., Agar, N. Y. R. & Agar, J. N. Protein Aggregation and Protein Instability Govern Familial Amyotrophic Lateral Sclerosis Patient Survival. *PLOS Biol.* **6**, e170 (2008).

55. Prudencio, M., Hart, P. J., Borchelt, D. R. & Andersen, P. M. Variation in aggregation propensities among ALS-associated variants of SOD1: Correlation to human disease. *Hum. Mol. Genet.* **18**, 3217–3226 (2009).

56. Vassall, K. A. *et al.* Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2210–2215 (2011).

57. McAlary, L., Aquilina, J. A. & Yerbury, J. J. Susceptibility of Mutant SOD1 to Form a Destabilized Monomer Predicts Cellular Aggregation and Toxicity but Not In vitro Aggregation Propensity. *Front. Neurosci.* **10**, (2016).

58. Shi, Y., Acerson, M. J., Abdolvahabi, A., Mowery, R. A. & Shaw, B. F. Gibbs Energy of Superoxide Dismutase Heterodimerization Accounts for Variable Survival in Amyotrophic Lateral Sclerosis. *J. Am. Chem. Soc.* **138**, 5351–5362 (2016).

59. Berdyński, M. *et al.* SOD1 mutations associated with amyotrophic lateral sclerosis analysis of variant severity. *Sci. Rep.* **12**, 103 (2022).

60. Rothstein, J. D. Current hypotheses for the underlying biology of amyotrophic lateral sclerosis. *Ann. Neurol.* **65**, S3–S9 (2009).

61. Le Gall, L. *et al.* Molecular and Cellular Mechanisms Affected in ALS. *J. Pers. Med.* **10**, 101 (2020).

62. Shibata, N. *et al.* Intense Superoxide Dismutase-1 Immunoreactivity in Intracytoplasmic Hyaline Inclusions of Familial Amyotrophic Lateral Sclerosis with Posterior Column Involvement. *J. Neuropathol. Exp. Neurol.* **55**, 481–490 (1996).

63. Shibata, N., Asayama, K., Hirano, A. & Kobayashi, M. Immunohistochemical study on superoxide dismutases in spinal cords from autopsied patients with amyotrophic lateral sclerosis. *Dev. Neurosci.* **18**, 492–498 (1996).

64. Kato, S. *et al.* New consensus research on neuropathological aspects of familial amyotrophic lateral sclerosis with superoxide dismutase 1 (SOD1) gene mutations: inclusions containing SOD1 in neurons and astrocytes. *Amyotroph. Lateral Scler. Mot. Neuron Disord. Off. Publ. World Fed. Neurol. Res. Group Mot. Neuron Dis.* **1**, 163–184 (2000).

65. Bruijn, L. I. *et al.* Aggregation and Motor Neuron Toxicity of an ALS-Linked SOD1 Mutant Independent from Wild-Type SOD1. *Science* **281**, 1851–1854 (1998).

66. Johnston, J. A., Dalton, M. J., Gurney, M. E. & Kopito, R. R. Formation of high molecular weight complexes of mutant Cu,Zn-superoxide dismutase in a mouse model for familial amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12571–12576 (2000).

67. Wang, J., Xu, G. & Borchelt, D. R. High molecular weight complexes of mutant superoxide dismutase 1: age-dependent and tissue-specific accumulation. *Neurobiol. Dis.* **9**, 139–148 (2002).

68. Wang, J. *et al.* Copper-binding-site-null SOD1 causes ALS in transgenic mice: aggregates of non-native SOD1 delineate a common feature. *Hum. Mol. Genet.* **12**, 2753–2764 (2003).

69. Wang, J. *et al.* Progressive aggregation despite chaperone associations of a mutant SOD1-YFP in transgenic mice that develop ALS. *Proc. Natl. Acad. Sci.* **106**, 1392–1397 (2009).

70. Cheroni, C. *et al.* Accumulation of human SOD1 and ubiquitinated deposits in the spinal cord of SOD1G93A mice during motor neuron disease progression correlates with a decrease of proteasome. *Neurobiol. Dis.* **18**, 509–522 (2005).

71. Deng, H.-X. *et al.* Conversion to the amyotrophic lateral sclerosis phenotype is associated with intermolecular linked insoluble aggregates of SOD1 in mitochondria. *Proc. Natl. Acad. Sci.* **103**, 7142–7147 (2006).

72. Kato, S. Amyotrophic lateral sclerosis models and human neuropathology: similarities and differences. *Acta Neuropathol. (Berl.)* **115**, 97–114 (2008).

73. Jaarsma, D. *et al.* Human Cu/Zn Superoxide Dismutase (SOD1) Overexpression in Mice Causes Mitochondrial Vacuolization, Axonal Degeneration, and Premature Motoneuron Death and Accelerates Motoneuron Disease in Mice Expressing a Familial Amyotrophic Lateral Sclerosis Mutant SOD1. *Neurobiol. Dis.* **7**, 623–643 (2000).

74. Wang, L. *et al.* Wild-type SOD1 overexpression accelerates disease onset of a G85R SOD1 mouse. *Hum. Mol. Genet.* **18**, 1642–1651 (2009).

75. Xu, G. *et al.* Direct and indirect mechanisms for wild-type SOD1 to enhance the toxicity of mutant SOD1 in bigenic transgenic mice. *Hum. Mol. Genet.* **24**, 1019–1035 (2015).

76. Gill, C. *et al.* SOD1-positive aggregate accumulation in the CNS predicts slower disease progression and increased longevity in a mutant SOD1 mouse model of ALS. *Sci. Rep.* **9**, 6724 (2019).

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

78. Karch, C. M., Prudencio, M., Winkler, D. D., Hart, P. J. & Borchelt, D. R. Role of mutant SOD1 disulfide oxidation and aggregation in the pathogenesis of familial ALS. *Proc. Natl. Acad. Sci.* **106**, 7774–7779 (2009).

79. Jonsson, P. A. *et al.* Minute quantities of misfolded mutant superoxide dismutase-1 cause amyotrophic lateral sclerosis. *Brain* **127**, 73–88 (2004).

80. Zetterström, P. *et al.* Soluble misfolded subfractions of mutant superoxide dismutase-1s are enriched in spinal cords throughout life in murine ALS models. *Proc. Natl. Acad. Sci.* **104**, 14157–14162 (2007).

81. Brotherton, T. E. *et al.* Localization of a toxic form of superoxide dismutase 1 protein to pathologically affected tissues in familial ALS. *Proc. Natl. Acad. Sci.* **109**, 5505–5510 (2012).

82. Redler, R. L., Fee, L., Fay, J. M., Caplow, M. & Dokholyan, N. V. Non-native Soluble Oligomers of Cu/Zn Superoxide Dismutase (SOD1) Contain a Conformational Epitope Linked to Cytotoxicity in Amyotrophic Lateral Sclerosis (ALS). *Biochemistry* **53**, 2423–2432 (2014).

83. Zhu, C., Beck, M. V., Griffith, J. D., Deshmukh, M. & Dokholyan, N. V. Large SOD1 aggregates, unlike trimeric SOD1, do not impact cell viability in a model of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **115**, 4661–4665 (2018).

84. Silveira, J. R. *et al.* The most infectious prion protein particles. *Nature* **437**, 257–261 (2005).

85. Reiss, A. B., Arain, H. A., Stecker, M. M., Siegart, N. M. & Kasselman, L. J. Amyloid toxicity in Alzheimer's disease. *Rev. Neurosci.* **29**, 613–627 (2018).

86. Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **216**, 136–144 (1982).

87. Imran, M. & Mahmood, S. An overview of human prion diseases. *Virol. J.* **8**, 1–9 (2011).

88. Prusiner, S. B. Neurodegenerative Diseases and Prions. *N. Engl. J. Med.* **344**, 1516–1526 (2001).

89. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).

90. McAlary, L., Plotkin, S. S., Yerbury, J. J. & Cashman, N. R. Prion-Like Propagation of Protein Misfolding and Aggregation in Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.* **12**, (2019).

91. Collinge, J. & Clarke, A. R. A General Model of Prion Strains and Their Pathogenicity. *Science* **318**, 930–936 (2007).

92. Greenwald, J. & Riek, R. Biology of Amyloid: Structure, Function, and Regulation. *Structure* **18**, 1244–1260 (2010).

93. Chattopadhyay, M. *et al.* Initiation and elongation in fibrillation of ALS-linked superoxide dismutase. *Proc. Natl. Acad. Sci.* **105**, 18663–18668 (2008).

94. Chattopadhyay, M. *et al.* The Disulfide Bond, but Not Zinc or Dimerization, Controls Initiation and Seeded Growth in Amyotrophic Lateral Sclerosis-linked Cu,Zn Superoxide Dismutase (SOD1) Fibrillation*. *J. Biol. Chem.* **290**, 30624–30636 (2015).

95. Münch, C., O'Brien, J. & Bertolotti, A. Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proc. Natl. Acad. Sci.* **108**, 3548–3553 (2011).

96. Ivanova, M. I. *et al.* Aggregation-triggering segments of SOD1 fibril formation support a common pathway for familial and sporadic ALS. *Proc. Natl. Acad. Sci.* **111**, 197–201 (2014).

97. Ayers, J. I. *et al.* Experimental transmissibility of mutant SOD1 motor neuron disease. *Acta Neuropathol. (Berl.)* **128**, 791–803 (2014).

98. Ayers, J. I., Fromholt, S. E., O'Neal, V. M., Diamond, J. H. & Borchelt, D. R. Prion-like propagation of mutant SOD1 misfolding and motor neuron disease spread along neuroanatomical pathways. *Acta Neuropathol. (Berl.)* **131**, 103–114 (2016).

99. Bergh, J. *et al.* Structural and kinetic analysis of protein-aggregate strains in vivo using binary epitope mapping. *Proc. Natl. Acad. Sci.* **112**, 4489–4494 (2015).

100. Bidhendi, E. E. *et al.* Two superoxide dismutase prion strains transmit amyotrophic lateral sclerosis–like disease. *J. Clin. Invest.* **126**, 2249–2253 (2016).

101. Wang, L.-Q. *et al.* Cryo-EM structure of an amyloid fibril formed by full-length human SOD1 reveals its conformational conversion. *Nat. Commun.* **13**, 3491 (2022).

102. Zeineddine, R. *et al.* SOD1 protein aggregates stimulate macropinocytosis in neurons to facilitate their propagation. *Mol. Neurodegener.* **10**, 57 (2015).

103. Grad, L. I. *et al.* Intercellular propagated misfolding of wild-type Cu/Zn superoxide dismutase occurs via exosome-dependent and -independent mechanisms. *Proc. Natl. Acad. Sci.* **111**, 3620–3625 (2014).

104. Crown, A. *et al.* Tryptophan residue 32 in human Cu-Zn superoxide dismutase modulates prion-like propagation and strain selection. *PLOS ONE* **15**, e0227655 (2020).

105. Pokrishevsky, E., Hong, R. H., Mackenzie, I. R. & Cashman, N. R. Spinal cord homogenates from SOD1 familial amyotrophic lateral sclerosis induce SOD1 aggregation in living cells. *PLOS ONE* **12**, e0184384 (2017).

106. Ekhtiari Bidhendi, E. *et al.* Mutant superoxide dismutase aggregates from human spinal cord transmit amyotrophic lateral sclerosis. *Acta Neuropathol. (Berl.)* **136**, 939–953 (2018).

107. Keskin, I. *et al.* Peripheral administration of SOD1 aggregates does not transmit pathogenic aggregation to the CNS of SOD1 transgenic mice. *Acta Neuropathol. Commun.* **9**, 111 (2021).

108. Johnson, B. S. *et al.* TDP-43 Is Intrinsically Aggregation-prone, and Amyotrophic Lateral Sclerosis-linked Mutations Accelerate Aggregation and Increase Toxicity*. *J. Biol. Chem.* **284**, 20329–20339 (2009).

109. Nomura, T. *et al.* Intranuclear Aggregation of Mutant FUS/TLS as a Molecular Pathomechanism of Amyotrophic Lateral Sclerosis*. *J. Biol. Chem.* **289**, 1192–1202 (2014).

110. Arseni, D. *et al.* Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature* **601**, 139–143 (2022).

111. Arseni, D. *et al.* TDP-43 forms amyloid filaments with a distinct fold in type A FTLD-TDP. *Nature* **620**, 898–903 (2023).

112. Sun, Y. *et al.* Molecular structure of an amyloid fibril formed by FUS low-complexity domain. *iScience* **25**, 103701 (2022).

113. Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K. & Nukina, N. A Seeding Reaction Recapitulates Intracellular Formation of Sarkosyl-insoluble Transactivation Response Element (TAR) DNA-binding Protein-43 Inclusions*♦. *J. Biol. Chem.* **286**, 18664–18672 (2011).

114. Nonaka, T. *et al.* Prion-like Properties of Pathological TDP-43 Aggregates from Diseased Brains. *Cell Rep.* **4**, 124–134 (2013).

115. Porta, S. *et al.* Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat. Commun.* **9**, 1–15 (2018).

116. Ding, X. *et al.* Spreading of TDP-43 pathology via pyramidal tract induces ALS-like phenotypes in TDP-43 transgenic mice. *Acta Neuropathol. Commun.* **9**, 15 (2021).

117. Kumar, S. T. *et al.* Seeding the aggregation of TDP-43 requires post-fibrillization proteolytic cleavage. *Nat. Neurosci.* **26**, 983–996 (2023).

118. Kong, J. & Xu, Z. Massive mitochondrial degeneration in motor neurons triggers the onset of amyotrophic lateral sclerosis in mice expressing a mutant SOD1. *J. Neurosci. Off. J. Soc. Neurosci.* **18**, 3241–3250 (1998).

119. Jaarsma, D. *et al.* CuZn superoxide dismutase (SOD1) accumulates in vacuolated mitochondria in transgenic mice expressing amyotrophic lateral sclerosis-linked SOD1 mutations. *Acta Neuropathol. (Berl.)* **102**, 293–305 (2001).

120. Higgins, C. M., Jung, C. & Xu, Z. ALS-associated mutant SOD1G93A causes mitochondrial vacuolation by expansion of the intermembrane space and by involvement of SOD1 aggregation and peroxisomes. *BMC Neurosci.* **4**, 16 (2003).

121. Liu, J. *et al.* Toxicity of Familial ALS-Linked SOD1 Mutants from Selective Recruitment to Spinal Mitochondria. *Neuron* **43**, 5–17 (2004).

122. Sasaki, S., Warita, H., Murakami, T., Abe, K. & Iwata, M. Ultrastructural study of mitochondria in the spinal cord of transgenic mice with a G93A mutant SOD1 gene. *Acta Neuropathol. (Berl.)* **107**, 461–474 (2004).

123. Afifi, A. K., Aleu, F. P., Goodgold, J. & MacKay, B. Ultrastructure of atrophic muscle in amyotrophic lateral sclerosis. *Neurology* **16**, 475–475 (1966).

124. Atsumi, T. The ultrastructure of intramuscular nerves in amyotrophic lateral sclerosis. *Acta Neuropathol. (Berl.)* **55**, 193–198 (1981).

125. Nakano, Y., Hirayama, K. & Terao, K. Hepatic Ultrastructural Changes and Liver Dysfunction in Amyotrophic Lateral Sclerosis. *Arch. Neurol.* **44**, 103–106 (1987).

126. Sasaki, S. & Iwata, M. Ultrastructural study of synapses in the anterior horn neurons of patients with amyotrophic lateral sclerosis. *Neurosci. Lett.* **204**, 53–56 (1996).

127. Sasaki, S. & Iwata, M. Ultrastructural change of synapses of Betz cells in patients with amyotrophic lateral sclerosis. *Neurosci. Lett.* **268**, 29–32 (1999).

128. Tafuri, F., Ronchi, D., Magri, F., Comi, G. P. & Corti, S. SOD1 misplacing and mitochondrial dysfunction in amyotrophic lateral sclerosis pathogenesis. *Front. Cell. Neurosci.* **9**, (2015).

129. Higgins, C. M. J., Jung, C., Ding, H. & Xu, Z. Mutant Cu, Zn Superoxide Dismutase that Causes Motoneuron Degeneration Is Present in Mitochondria in the CNS. *J. Neurosci.* **22**, RC215–RC215 (2002).

130. Field, L. S., Furukawa, Y., O'Halloran, T. V. & Culotta, V. C. Factors Controlling the Uptake of Yeast Copper/Zinc Superoxide Dismutase into Mitochondria*. *J. Biol. Chem.* **278**, 28052–28059 (2003).

131. Ahtoniemi, T., Jaronen, M., Keksa-Goldsteine, V., Goldsteins, G. & Koistinaho, J. Mutant SOD1 from spinal cord of G93A rats is destabilized and binds to inner mitochondrial membrane. *Neurobiol. Dis.* **32**, 479–485 (2008).

132. Cozzolino, M. *et al.* Oligomerization of Mutant SOD1 in Mitochondria of Motoneuronal Cells Drives Mitochondrial Damage and Cell Toxicity. *Antioxid. Redox Signal.* **11**, 1547–1558 (2009).

133. Magrané, J. *et al.* Mutant SOD1 in neuronal mitochondria causes toxicity and mitochondrial dynamics abnormalities. *Hum. Mol. Genet.* **18**, 4552–4564 (2009).

134. Igoudjil, A. *et al.* In Vivo Pathogenic Role of Mutant SOD1 Localized in the Mitochondrial Intermembrane Space. *J. Neurosci.* **31**, 15826–15837 (2011).

135. Pasinelli, P. *et al.* Amyotrophic Lateral Sclerosis-Associated SOD1 Mutant Proteins Bind and Aggregate with Bcl-2 in Spinal Cord Mitochondria. *Neuron* **43**, 19–30 (2004).

136. Vande Velde, C., Miller, T. M., Cashman, N. R. & Cleveland, D. W. Selective association of misfolded ALS-linked mutant SOD1 with the cytoplasmic face of mitochondria. *Proc. Natl. Acad. Sci.* **105**, 4022–4027 (2008).

137. Israelson, A. *et al.* Misfolded Mutant SOD1 Directly Inhibits VDAC1 Conductance in a Mouse Model of Inherited ALS. *Neuron* **67**, 575–587 (2010).

138. Pedrini, S. *et al.* ALS-linked mutant SOD1 damages mitochondria by promoting conformational changes in Bcl-2. *Hum. Mol. Genet.* **19**, 2974–2986 (2010).

139. Pickles, S. *et al.* ALS-linked misfolded SOD1 species have divergent impacts on mitochondria. *Acta Neuropathol. Commun.* **4**, 43 (2016).

140. Hemerková, P. & Vališ, M. Role of Oxidative Stress in the Pathogenesis of Amyotrophic Lateral Sclerosis: Antioxidant Metalloenzymes and Therapeutic Strategies. *Biomolecules* **11**, 437 (2021).

141. Fukai, T. & Ushio-Fukai, M. Superoxide Dismutases: Role in Redox Signaling, Vascular Function, and Diseases. *Antioxid. Redox Signal.* **15**, 1583–1606 (2011).

142. Cunha-Oliveira, T. *et al.* Oxidative Stress in Amyotrophic Lateral Sclerosis: Pathophysiology and Opportunities for Pharmacological Intervention. *Oxid. Med. Cell. Longev.* **2020**, 5021694 (2020).

143. Hosaka, T., Tsuji, H. & Tamaoka, A. Biomolecular Modifications Linked to Oxidative Stress in Amyotrophic Lateral Sclerosis: Determining Promising Biomarkers Related to Oxidative Stress. *Processes* **9**, 1667 (2021).

144. Bogdanov, M. *et al.* Increased oxidative damage to DNA in ALS patients. *Free Radic. Biol. Med.* **29**, 652–658 (2000).

145. Simpson, E. P., Henry, Y. K., Henkel, J. S., Smith, R. G. & Appel, S. H. Increased lipid peroxidation in sera of ALS patients: A potential biomarker of disease burden. *Neurology* **62**, 1758–1765 (2004).

146. Sohmiya, M. *et al.* An increase of oxidized coenzyme Q-10 occurs in the plasma of sporadic ALS patients. *J. Neurol. Sci.* **228**, 49–53 (2005).

147. Babu, G. N. *et al.* Oxidant–antioxidant imbalance in the erythrocytes of sporadic amyotrophic lateral sclerosis patients correlates with the progression of disease. *Neurochem. Int.* **52**, 1284–1289 (2008).

148. Murata, T., Ohtsuka, C. & Terayama, Y. Increased mitochondrial oxidative damage in patients with sporadic amyotrophic lateral sclerosis. *J. Neurol. Sci.* **267**, 66–69 (2008).

149. Cova, E. *et al.* Time course of oxidant markers and antioxidant defenses in subgroups of amyotrophic lateral sclerosis patients. *Neurochem. Int.* **56**, 687–693 (2010).

150. Ikawa, M. *et al.* Increased oxidative stress is related to disease severity in the ALS motor cortex: A PET study. *Neurology* **84**, 2033–2039 (2015).

151. Bruijn, L. I. *et al.* Elevated free nitrotyrosine levels, but not protein-bound nitrotyrosine or hydroxyl radicals, throughout amyotrophic lateral sclerosis (ALS)-like disease implicate tyrosine nitration as an aberrant in vivo property of one familial ALS-linked superoxide dismutase 1 mutant. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7606–7611 (1997).

152. Ferrante, R. J. *et al.* Increased 3-nitrotyrosine and oxidative damage in mice with a human copper/zinc superoxide dismutase mutation. *Ann. Neurol.* **42**, 326–334 (1997).

153. Andrus, P. K., Fleck, T. J., Gurney, M. E. & Hall, E. D. Protein oxidative damage in a transgenic mouse model of familial amyotrophic lateral sclerosis. *J. Neurochem.* **71**, 2041–2048 (1998).

154. Mahoney, D. J., Kaczor, J. J., Bourgeois, J., Yasuda, N. & Tarnopolsky, M. A. Oxidative stress and antioxidant enzyme upregulation in SOD1-G93A mouse skeletal muscle. *Muscle Nerve* **33**, 809–816 (2006).

155. Szeliga, M. Peroxiredoxins in Neurodegenerative Diseases. *Antioxidants* **9**, 1203 (2020).

156. Kim, K. Glutathione in the Nervous System as a Potential Therapeutic Target to Control the Development and Progression of Amyotrophic Lateral Sclerosis. *Antioxidants* **10**, 1011 (2021).

157. Ezzi, S. A., Urushitani, M. & Julien, J.-P. Wild-type superoxide dismutase acquires binding and toxic properties of ALS-linked mutant forms through oxidation. *J. Neurochem.* **102**, 170–178 (2007).

158. Mulligan, V. K. *et al.* Early Steps in Oxidation-Induced SOD1 Misfolding: Implications for Non-Amyloid Protein Aggregation in Familial ALS. *J. Mol. Biol.* **421**, 631–652 (2012).

159. Martins, D. & English, A. M. SOD1 oxidation and formation of soluble aggregates in yeast: relevance to sporadic ALS development. *Redox Biol.* **2**, 632–639 (2014).

160. Liu, K. X. *et al.* Neuron-specific antioxidant OXR1 extends survival of a mouse model of amyotrophic lateral sclerosis. *Brain* **138**, 1167–1181 (2015).

161. Lin, J. H., Walter, P. & Yen, T. S. B. Endoplasmic Reticulum Stress in Disease Pathogenesis. *Annu. Rev. Pathol.* **3**, 399–425 (2008).

162. Jeon, Y.-M., Kwon, Y., Lee, S. & Kim, H.-J. Potential roles of the endoplasmic reticulum stress pathway in amyotrophic lateral sclerosis. *Front. Aging Neurosci.* **15**, 1047897 (2023).

163. Ilieva, E. V. *et al.* Oxidative and endoplasmic reticulum stress interplay in sporadic amyotrophic lateral sclerosis. *Brain* **130**, 3111–3123 (2007).

164. Hetz, C. *et al.* XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. *Genes Dev.* **23**, 2294–2306 (2009).

165. Ito, Y. *et al.* Involvement of CHOP, an ER-stress apoptotic mediator, in both human sporadic ALS and ALS model mice. *Neurobiol. Dis.* **36**, 470–476 (2009).

166. Medinas, D. B. *et al.* Endoplasmic reticulum stress leads to accumulation of wild-type SOD1 aggregates associated with sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **115**, 8209–8214 (2018).

167. Tobisawa, S. *et al.* Mutant SOD1 linked to familial amyotrophic lateral sclerosis, but not wild-type SOD1, induces ER stress in COS7 cells and transgenic mice. *Biochem. Biophys. Res. Commun.* **303**, 496–503 (2003).

168. Kikuchi, H. *et al.* Spinal cord endoplasmic reticulum stress associated with a microsomal accumulation of mutant superoxide dismutase-1 in an ALS model. *Proc. Natl. Acad. Sci.* **103**, 6025–6030 (2006).

169. Nagata, T. *et al.* Increased ER stress during motor neuron degeneration in a transgenic mouse model of amyotrophic lateral sclerosis. *Neurol. Res.* **29**, 767–771 (2007).

170. Saxena, S., Cabuy, E. & Caroni, P. A role for motoneuron subtype–selective ER stress in disease manifestations of FALS mice. *Nat. Neurosci.* **12**, 627–636 (2009).

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

172. Atkin, J. D. *et al.* Mutant SOD1 inhibits ER-Golgi transport in amyotrophic lateral sclerosis. *J. Neurochem.* **129**, 190–204 (2014).

173. Ligon, L. A. *et al.* Mutant superoxide dismutase disrupts cytoplasmic dynein in motor neurons. *Neuroreport* **16**, 533–536 (2005).

174. Bilsland, L. G. *et al.* Deficits in axonal transport precede ALS symptoms in vivo. *Proc. Natl. Acad. Sci.* **107**, 20523–20528 (2010).

175. Magrané, J., Cortez, C., Gan, W.-B. & Manfredi, G. Abnormal mitochondrial transport and morphology are common pathological denominators in SOD1 and TDP43 ALS mouse models. *Hum. Mol. Genet.* **23**, 1413–1424 (2014).

176. Tosolini, A. P. *et al.* BDNF-dependent modulation of axonal transport is selectively impaired in ALS. *Acta Neuropathol. Commun.* **10**, 121 (2022).

177. Kieran, D. *et al.* A mutation in dynein rescues axonal transport defects and extends the life span of ALS mice. *J. Cell Biol.* **169**, 561–567 (2005).

178. Xie, Y. *et al.* Endolysosomal Deficits Augment Mitochondria Pathology in Spinal Motor Neurons of Asymptomatic fALS Mice. *Neuron* **87**, 355–370 (2015).

179. Otomo, A. *et al.* High-throughput quantitative analysis of axonal transport in cultured neurons from SOD1H46R ALS mice by using a microfluidic device. *Neurosci. Res.* **174**, 46–52 (2022).

180. Rhymes, E. R. *et al.* Bimodal regulation of axonal transport by the GDNF-RET signalling axis in healthy and diseased motor neurons. *Cell Death Dis.* **13**, 584 (2022).

181. Zhang, F. *et al.* Interaction between Familial Amyotrophic Lateral Sclerosis (ALS)-linked SOD1 Mutants and the Dynein Complex*. *J. Biol. Chem.* **282**, 16691–16699 (2007).

182. Ström, A.-L. *et al.* Interaction of Amyotrophic Lateral Sclerosis (ALS)-related Mutant Copper-Zinc Superoxide Dismutase with the Dynein-Dynactin Complex Contributes to Inclusion Formation*. *J. Biol. Chem.* **283**, 22795–22805 (2008).

183. Bosco, D. A. *et al.* Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nat. Neurosci.* **13**, 1396–1403 (2010).

184. Vande Velde, C. *et al.* Misfolded SOD1 Associated with Motor Neuron Mitochondria Alters Mitochondrial Shape and Distribution Prior to Clinical Onset. *PLoS ONE* **6**, e22031 (2011).

185. De Vos, K. J. *et al.* Familial amyotrophic lateral sclerosis-linked SOD1 mutants perturb fast axonal transport to reduce axonal mitochondria content. *Hum. Mol. Genet.* **16**, 2720–2728 (2007).

186. Moller, A., Bauer, C. S., Cohen, R. N., Webster, C. P. & De Vos, K. J. Amyotrophic lateral sclerosis-associated mutant SOD1 inhibits anterograde axonal transport of mitochondria by reducing Miro1 levels. *Hum. Mol. Genet.* **26**, 4668–4679 (2017).

187. Fransson, S., Ruusala, A. & Aspenström, P. The atypical Rho GTPases Miro-1 and Miro-2 have essential roles in mitochondrial trafficking. *Biochem. Biophys. Res. Commun.* **344**, 500–510 (2006).

188. MacAskill, A. F. *et al.* Miro1 Is a Calcium Sensor for Glutamate Receptor-Dependent Localization of Mitochondria at Synapses. *Neuron* **61**, 541–555 (2009).

189. Zhang, F. *et al.* Miro1 deficiency in amyotrophic lateral sclerosis. *Front. Aging Neurosci.* **7**, 100 (2015).

190. De Vos, K. J. & Hafezparast, M. Neurobiology of axonal transport defects in motor neuron diseases: Opportunities for translational research? *Neurobiol. Dis.* **105**, 283–299 (2017).

191. Marinković, P. *et al.* Axonal transport deficits and degeneration can evolve independently in mouse models of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **109**, 4296–4301 (2012).

192. Van Den Bosch, L., Van Damme, P., Bogaert, E. & Robberecht, W. The role of excitotoxicity in the pathogenesis of amyotrophic lateral sclerosis. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1762**, 1068–1082 (2006).

193. King, A. E., Woodhouse, A., Kirkcaldie, M. T. K. & Vickers, J. C. Excitotoxicity in ALS: Overstimulation, or overreaction? *Exp. Neurol.* **275**, 162–171 (2016).

194. Ikonomidou, C., Qin Qin, Y., Labruyere, J. & Olney, J. W. Motor neuron degeneration induced by excitotoxin agonists has features in common with those seen in the SOD-1 transgenic mouse model of amyotrophic lateral sclerosis. *J. Neuropathol. Exp. Neurol.* **55**, 211–224 (1996).

195. Vandenberghe, W., Robberecht, W. & Brorson, J. R. AMPA Receptor Calcium Permeability, GluR2 Expression, and Selective Motoneuron Vulnerability. *J. Neurosci.* **20**, 123–132 (2000).

196. Van Den Bosch, L., Vandenberghe, W., Klaassen, H., Van Houtte, E. & Robberecht, W. Ca(2+)-permeable AMPA receptors and selective vulnerability of motor neurons. *J. Neurol. Sci.* **180**, 29–34 (2000).

197. Rothstein, J. D., Martin, L. J. & Kuncl, R. W. Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis. *N. Engl. J. Med.* **326**, 1464–1468 (1992).

198. Rothstein, J. D., Van Kammen, M., Levey, A. I., Martin, L. J. & Kuncl, R. W. Selective loss of glial glutamate transporter GLT-1 in amyotrophic lateral sclerosis. *Ann. Neurol.* **38**, 73–84 (1995).

199. Fray, A. E. *et al.* The expression of the glial glutamate transporter protein EAAT2 in motor neuron disease: an immunohistochemical study. *Eur. J. Neurosci.* **10**, 2481–2489 (1998).

200. Sasaki, S., Komori, T. & Iwata, M. Excitatory amino acid transporter 1 and 2 immunoreactivity in the spinal cord in amyotrophic lateral sclerosis. *Acta Neuropathol. (Berl.)* **100**, 138–144 (2000).

201. Canton, T., Pratt, J., Stutzmann, J. M., Imperato, A. & Boireau, A. Glutamate uptake is decreased tardively in the spinal cord of FALS mice. *Neuroreport* **9**, 775–778 (1998).

202. Bendotti, C. *et al.* Transgenic SOD1 G93A mice develop reduced GLT-1 in spinal cord without alterations in cerebrospinal fluid glutamate levels. *J. Neurochem.* **79**, 737–746 (2001).

203. Warita, H. *et al.* Tardive decrease of astrocytic glutamate transporter protein in transgenic mice with ALS-linked mutant SOD1. *Neurol. Res.* **24**, 577–581 (2002).

204. Deitch, J. S., Alexander, G. M., Del Valle, L. & Heiman-Patterson, T. D. GLT-1 glutamate transporter levels are unchanged in mice expressing G93A human mutant SOD1. *J. Neurol. Sci.* **193**, 117–126 (2002).

205. Doble, A. The pharmacology and mechanism of action of riluzole. *Neurology* **47**, S233-241 (1996).

206. Albo, F., Pieri, M. & Zona, C. Modulation of AMPA receptors in spinal motor neurons by the neuroprotective agent riluzole. *J. Neurosci. Res.* **78**, 200–207 (2004).

207. Pramatarova, A., Laganière, J., Roussel, J., Brisebois, K. & Rouleau, G. A. Neuron-Specific Expression of Mutant Superoxide Dismutase 1 in Transgenic Mice Does Not Lead to Motor Impairment. *J. Neurosci.* **21**, 3369–3374 (2001).

208. Jaarsma, D., Teuling, E., Haasdijk, E. D., Zeeuw, C. I. D. & Hoogenraad, C. C. Neuron-Specific Expression of Mutant Superoxide Dismutase Is Sufficient to Induce Amyotrophic Lateral Sclerosis in Transgenic Mice. *J. Neurosci.* **28**, 2075–2088 (2008).

209. Wang, L. *et al.* Restricted expression of mutant SOD1 in spinal motor neurons and interneurons induces motor neuron pathology. *Neurobiol. Dis.* **29**, 400–408 (2008).

210. Gong, Y. H., Parsadanian, A. S., Andreeva, A., Snider, W. D. & Elliott, J. L. Restricted Expression of G86R Cu/Zn Superoxide Dismutase in Astrocytes Results in Astrocytosis But Does Not Cause Motoneuron Degeneration. *J. Neurosci.* **20**, 660–665 (2000).

211. Dobrowolny, G. *et al.* Skeletal Muscle Is a Primary Target of SOD1G93A-Mediated Toxicity. *Cell Metab.* **8**, 425–436 (2008).

212. Wong, M. & Martin, L. J. Skeletal muscle-restricted expression of human SOD1 causes motor neuron degeneration in transgenic mice. *Hum. Mol. Genet.* **19**, 2284–2302 (2010).

213. Martin, L. J. & Wong, M. Skeletal Muscle-Restricted Expression of Human SOD1 in Transgenic Mice Causes a Fatal ALS-Like Syndrome. *Front. Neurol.* **11**, (2020).

214. Beers, D. R. *et al.* Wild-type microglia extend survival in PU.1 knockout mice with familial amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **103**, 16021–16026 (2006).

215. Boillée, S. *et al.* Onset and Progression in Inherited ALS Determined by Motor Neurons and Microglia. *Science* **312**, 1389–1392 (2006).

216. Ilieva, H., Polymenidou, M. & Cleveland, D. W. Non–cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. *J. Cell Biol.* **187**, 761–772 (2009).

217. Philips, T. & Rothstein, J. Glial cells in Amyotrophic Lateral Sclerosis. *Exp. Neurol.* **262PB**, 111–120 (2014).

218. Vahsen, B. F. *et al.* Non-neuronal cells in amyotrophic lateral sclerosis — from pathogenesis to biomarkers. *Nat. Rev. Neurol.* **17**, 333–348 (2021).

219. Peggion, C. *et al.* SOD1 in ALS: Taking Stock in Pathogenic Mechanisms and the Role of Glial and Muscle Cells. *Antioxidants* **11**, 614 (2022).

220. Kaur, S. J., McKeown, S. R. & Rashid, S. Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis. *Gene* **577**, 109–118 (2016).

221. Andersen, P. M. *et al.* Phenotype in an Infant with SOD1 Homozygous Truncating Mutation. *N. Engl. J. Med.* **381**, 486–488 (2019).

222. Park, J. H. *et al.* SOD1 deficiency: a novel syndrome distinct from amyotrophic lateral sclerosis. *Brain* **142**, 2230–2237 (2019).

223. Park, J. H. *et al.* The motor system is exceptionally vulnerable to absence of the ubiquitously expressed superoxide dismutase-1. *Brain Commun.* **5**, fcad017 (2023).

224. Blair, H. A. Tofersen: First Approval. *Drugs* **83**, 1039–1043 (2023).

225. European Medicines Agency. New treatment for rare motor neurone disease recommended for approval. https://www.ema.europa.eu/en/news/new-treatment-rare-motor-neurone-disease-recommended-approval.

226. National Institute for Health and Care Excellence. Tofersen for treating amyotrophic lateral sclerosis caused by SOD1 gene mutations. https://www.nice.org.uk/guidance/awaiting-development/gid-ta11353.

227. Miller, T. M. *et al.* Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *N. Engl. J. Med.* **387**, 1099–1110 (2022).

228. Miller, T. *et al.* Phase 1–2 Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *N. Engl. J. Med.* **383**, 109–119 (2020).

229. Ince, P. G., Tomkins, J., Slade, J. Y., Thatcher, N. M. & Shaw, P. J. Amyotrophic lateral sclerosis associated with genetic abnormalities in the gene encoding Cu/Zn superoxide dismutase: molecular pathology of five new cases, and comparison with previous reports and 73 sporadic cases of ALS. *J. Neuropathol. Exp. Neurol.* **57**, 895–904 (1998).

230. Mackenzie, I. R. A. *et al.* Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Ann. Neurol.* **61**, 427–434 (2007).

231. Forsberg, K. *et al.* Novel Antibodies Reveal Inclusions Containing Non-Native SOD1 in Sporadic ALS Patients. *PLOS ONE* **5**, e11552 (2010).

232. Forsberg, K. *et al.* Misfolded SOD1 inclusions in patients with mutations in C9orf72 and other ALS/FTD-associated genes. *J. Neurol. Neurosurg. Psychiatry* **90**, 861–869 (2019).

233. Tokuda, E. *et al.* Wild-type Cu/Zn-superoxide dismutase is misfolded in cerebrospinal fluid of sporadic amyotrophic lateral sclerosis. *Mol. Neurodegener.* **14**, 42 (2019).

234. Jeon, G. S. *et al.* Pathological Modification of TDP-43 in Amyotrophic Lateral Sclerosis with SOD1 Mutations. *Mol. Neurobiol.* **56**, 2007–2021 (2019).

235. Shan, X., Vocadlo, D. & Krieger, C. Mislocalization of TDP-43 in the G93A mutant SOD1 transgenic mouse model of ALS. *Neurosci. Lett.* **458**, 70–74 (2009).

236. Okamoto, Y. *et al.* An autopsy case of SOD1-related ALS with TDP-43 positive inclusions. *Neurology* **77**, 1993–1995 (2011).

237. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: A cross-sectional study. *Lancet Neurol.* **11**, 323–330 (2012).

238. Balendra, R. & Isaacs, A. M. C9orf72 -mediated ALS and FTD: multiple pathways to disease. *Nat. Rev. Neurol.* **14**, 544–558 (2018).

239. Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H. & Hayes, M. J. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics* **29**, 499–503 (2013).

240. Farg, M. A. *et al.* C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23**, 3579–3595 (2014).

241. Sellier, C. *et al.* Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* **35**, 1276–1297 (2016).

242. Sullivan, P. M. *et al.* The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* **4**, 51 (2016).

243. Yang, M. *et al.* A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci. Adv.* **2**, e1601167 (2016).

244. Webster, C. P., Smith, E. F., Grierson, A. J. & Vos, K. J. D. C9orf72 plays a central role in Rab GTPase-dependent regulation of autophagy. *Small GTPases* **9**, 399–408 (2018).

245. Cason, S. E. & Holzbaur, E. L. F. Axonal transport of autophagosomes is regulated by dynein activators JIP3/JIP4 and ARF/RAB GTPases. *J. Cell Biol.* **222**, e202301084 (2023).

246. Deinhardt, K. *et al.* Rab5 and Rab7 Control Endocytic Sorting along the Axonal Retrograde Transport Pathway. *Neuron* **52**, 293–305 (2006).

247. Webster, C. P. *et al.* The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* **35**, 1656–1676 (2016).

248. Chitiprolu, M. *et al.* A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nat. Commun.* **9**, 1–18 (2018).

249. Wang, M. *et al.* C9orf72 associates with inactive Rag GTPases and regulates mTORC1-mediated autophagosomal and lysosomal biogenesis. *Aging Cell* **19**, e13126 (2020).

250. Ugolino, J. *et al.* Loss of C9orf72 Enhances Autophagic Activity via Deregulated mTOR and TFEB Signaling. *PLOS Genet.* **12**, e1006443 (2016).

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

252. O'Rourke, J. G. *et al.* C9orf72 is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).

253. Shi, Y. *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).

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

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

256. Frick, P. *et al.* Novel antibodies reveal presynaptic localization of C9orf72 protein and reduced protein levels in C9orf72 mutation carriers. *Acta Neuropathol. Commun.* **6**, 72 (2018).

257. Waite, A. J. *et al.* Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5-1779.e13 (2014).

258. McGoldrick, P., Lau, A., You, Z., Durcan, T. M. & Robertson, J. Loss of C9orf72 perturbs the Ran-GTPase gradient and nucleocytoplasmic transport, generating compositionally diverse Importin β-1 granules. *Cell Rep.* **42**, 112134 (2023).

259. Koppers, M. *et al.* C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann. Neurol.* **78**, 426–438 (2015).

260. Shao, Q. *et al.* C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. *Acta Neuropathol. Commun.* **7**, 32 (2019).

261. Zhu, Q. *et al.* Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. *Nat. Neurosci.* **23**, 615–624 (2020).

262. Pettersson, O. J., Aagaard, L., Jensen, T. G. & Damgaard, C. K. Molecular mechanisms in DM1 — a focus on foci. *Nucleic Acids Res.* **43**, 2433–2441 (2015).

263. Galloway, J. N. & Nelson, D. L. Evidence for RNA-mediated toxicity in the fragile X-associated tremor/ataxia syndrome. *Future Neurol.* **4**, 785–798 (2009).

264. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80**, 415–428 (2013).

265. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci.* **110**, E4968–E4977 (2013).

266. Gendron, T. F. *et al.* Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol. (Berl.)* **126**, 829–844 (2013).

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

268. Mehta, A. R. *et al.* Improved detection of RNA foci in C9orf72 amyotrophic lateral sclerosis post-mortem tissue using BaseScope™ shows a lack of association with cognitive dysfunction. *Brain Commun.* **2**, (2020).

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

270. Sareen, D. *et al.* Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. *Sci. Transl. Med.* **5**, 208ra149-208ra149 (2013).

271. Chew, J. *et al.* C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* **348**, 1151–1154 (2015).

272. O'Rourke, J. G. *et al.* C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* **88**, 892–901 (2015).

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

274. Liu, Y. *et al.* C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron* **90**, 521–534 (2016).

275. Herranz-Martin, S. *et al.* Viral delivery of C9orf72 hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits. *Dis. Model. Mech.* **10**, 859–868 (2017).

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

277. DeJesus-Hernandez, M. *et al.* In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers. *Acta Neuropathol. (Berl.)* **134**, 255–269 (2017).

278. Vatsavayai, S. C., Nana, A. L., Yokoyama, J. S. & Seeley, W. W. C9orf72-FTD/ALS pathogenesis: evidence from human neuropathological studies. *Acta Neuropathol. (Berl.)* **137**, 1–26 (2019).

279. Božič, T., Zalar, M., Rogelj, B., Plavec, J. & Šket, P. Structural Diversity of Sense and Antisense RNA Hexanucleotide Repeats Associated with ALS and FTLD. *Molecules* **25**, 525 (2020).

280. Dodd, D. W., Tomchick, D. R., Corey, D. R. & Gagnon, K. T. Pathogenic C9ORF72 Antisense Repeat RNA Forms a Double Helix with Tandem C:C Mismatches. *Biochemistry* **55**, 1283–1286 (2016).

281. Fratta, P. *et al.* C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2**, (2012).

282. Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195–200 (2014).

283. McEachin, Z. T., Parameswaren, J., Raj, N., Bassell, G. J. & Jiang, J. RNA-mediated toxicity in C9orf72 ALS and FTD. *Neurobiol. Dis.* **145**, 105055 (2020).

284. Xu, Z. *et al.* Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7778–7783 (2013).

285. Swinnen, B. *et al.* A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol. (Berl.)* **135**, 427–443 (2018).

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

287. Moens, T. G. *et al.* Sense and antisense RNA are not toxic in Drosophila models of C9orf72-associated ALS/FTD. *Acta Neuropathol. (Berl.)* **135**, 445–457 (2018).

288. Conlon, E. G. *et al.* The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *eLife* **5**, e17820 (2016).

289. Wang, Q., Conlon, E. G., Manley, J. L. & Rio, D. C. Widespread intron retention impairs protein homeostasis in C9orf72 ALS brains. *Genome Res.* **30**, 1705–1715 (2020).

290. Ishiguro, A., Lu, J., Ozawa, D., Nagai, Y. & Ishihama, A. ALS-linked FUS mutations dysregulate G-quadruplex-dependent liquid–liquid phase separation and liquid-to-solid transition. *J. Biol. Chem.* **297**, 101284 (2021).

291. Burguete, A. S. *et al.* GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife* **4**, e08881 (2015).

292. Chong, Z. Z., Menkes, D. L. & Souayah, N. Pathogenesis underlying hexanucleotide repeat expansions in C9orf72 gene in amyotrophic lateral sclerosis. *Rev. Neurosci.* (2023) doi:10.1515/revneuro-2023-0060.

293. Schmitz, A., Pinheiro Marques, J., Oertig, I., Maharjan, N. & Saxena, S. Emerging Perspectives on Dipeptide Repeat Proteins in C9ORF72 ALS/FTD. *Front. Cell. Neurosci.* **15**, (2021).

294. Lee, K.-H. *et al.* C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **167**, 774-788.e17 (2016).

295. Wen, X. *et al.* Antisense Proline-Arginine RAN Dipeptides Linked to C9ORF72-ALS/FTD Form Toxic Nuclear Aggregates that Initiate In Vitro and In Vivo Neuronal Death. *Neuron* **84**, 1213–1225 (2014).

296. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129–133 (2015).

297. Tao, Z. *et al.* Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum. Mol. Genet.* **24**, 2426–2441 (2015).

298. Verdone, B. M. *et al.* A mouse model with widespread expression of the C9orf72-linked glycine–arginine dipeptide displays non-lethal ALS/FTD-like phenotypes. *Sci. Rep.* **12**, 5644 (2022).

299. Murakami, T. *et al.* ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **88**, 678–690 (2015).

300. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).

301. Boeynaems, S. *et al.* Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. *Mol. Cell* **65**, 1044-1055.e5 (2017).

302. White, M. R. *et al.* C9orf72 Poly(PR) Dipeptide Repeats Disturb Biomolecular Phase Separation and Disrupt Nucleolar Function. *Mol. Cell* **74**, 713-728.e6 (2019).

303. Frottin, F. *et al.* The nucleolus functions as a phase-separated protein quality control compartment. *Science* **365**, 342–347 (2019).

304. Andrade, N. S. *et al.* Dipeptide repeat proteins inhibit homology-directed DNA double strand break repair in C9ORF72 ALS/FTD. *Mol. Neurodegener.* **15**, 13 (2020).

305. Fumagalli, L. *et al.* C9orf72-derived arginine-containing dipeptide repeats associate with axonal transport machinery and impede microtubule-based motility. *Sci. Adv.* **7**, eabg3013 (2021).

306. Kwon, I. *et al.* Poly-dipeptides encoded by the *C9orf72* repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).

307. Mori, K. *et al.* The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science* **339**, 1335–1338 (2013).

308. May, S. *et al.* C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol. (Berl.)* **128**, 485–503 (2014).

309. Zhang, Y.-J. *et al.* Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol. (Berl.)* **128**, 505–524 (2014).

310. Mackenzie, I. R. A. *et al.* Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta Neuropathol. (Berl.)* **130**, 845–861 (2015).

311. Schludi, M. H. *et al.* Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathol. (Berl.)* **130**, 537–555 (2015).

312. Chang, Y.-J., Jeng, U.-S., Chiang, Y.-L., Hwang, I.-S. & Chen, Y.-R. The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties*. *J. Biol. Chem.* **291**, 4903–4911 (2016).

313. Zhang, Y.-J. *et al.* C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* **19**, 668–677 (2016).

314. Schludi, M. H. *et al.* Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol. (Berl.)* **134**, 241–254 (2017).

315. Khosravi, B. *et al.* Cell-to-cell transmission of C9orf72 poly-(Gly-Ala) triggers key features of ALS/FTD. *EMBO J.* **39**, e102811 (2020).

316. Nihei, Y. *et al.* Poly-glycine–alanine exacerbates C9orf72 repeat expansion-mediated DNA damage via sequestration of phosphorylated ATM and loss of nuclear hnRNPA3. *Acta Neuropathol. (Berl.)* **139**, 99–118 (2020).

317. Freibaum, B. D. & Taylor, J. P. The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Front. Mol. Neurosci.* **10**, 1–9 (2017).

318. Mackenzie, I. R. *et al.* Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol. (Berl.)* **126**, 859–879 (2013).

319. Davidson, Y. S. *et al.* Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol. Commun.* **2**, 70 (2014).

320. Schipper, L. J. *et al.* Prevalence of brain and spinal cord inclusions, including dipeptide repeat proteins, in patients with the C9ORF72 hexanucleotide repeat expansion: a systematic neuropathological review. *Neuropathol. Appl. Neurobiol.* **42**, 547–560 (2016).

321. Saberi, S. *et al.* Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathol. (Berl.)* **135**, 459–474 (2018).

322. Sakae, N. *et al.* Poly-GR dipeptide repeat polymers correlate with neurodegeneration and Clinicopathological subtypes in C9ORF72-related brain disease. *Acta Neuropathol. Commun.* **6**, 63 (2018).

323. Baborie, A. *et al.* Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathol. Appl. Neurobiol.* **41**, 601–612 (2015).

324. Vatsavayai, S. C. *et al.* Timing and significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. *Brain* **139**, 3202–3216 (2016).

325. Fisher, E. M. C. & Bannerman, D. M. Mouse models of neurodegeneration: Know your question, know your mouse. *Sci. Transl. Med.* **11**, 1–16 (2019).

326. Todd, T. W. & Petrucelli, L. Modelling amyotrophic lateral sclerosis in rodents. *Nat. Rev. Neurosci.* 1–21 (2022) doi:10.1038/s41583-022-00564-x.

327. Chinwalla, A. T. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).

328. Cozzolino, M. *et al.* Cysteine 111 Affects Aggregation and Cytotoxicity of Mutant Cu,Zn-superoxide Dismutase Associated with Familial Amyotrophic Lateral Sclerosis. *J. Biol. Chem.* **283**, 866 (2008).

329. Xu, W.-C. *et al.* Pathological hydrogen peroxide triggers the fibrillization of wild-type SOD1 via sulfenic acid modification of Cys-111. *Cell Death Dis.* **9**, 1–17 (2018).

330. DuVal, M. G. *et al.* Tryptophan 32 mediates SOD1 toxicity in a in vivo motor neuron model of ALS and is a promising target for small molecule therapeutics. *Neurobiol. Dis.* **124**, 297–310 (2019).

331. Lin, S. *et al.* Comparison of the transcriptional landscapes between human and mouse tissues. *Proc. Natl. Acad. Sci.* **111**, 17224–17229 (2014).

332. Breschi, A., Gingeras, T. R. & Guigó, R. Comparative transcriptomics in human and mouse. *Nat. Rev. Genet.* **18**, 425–440 (2017).

333. Hernández, F., Merchán-Rubira, J., Vallés-Saiz, L., Rodríguez-Matellán, A. & Avila, J. Differences Between Human and Murine Tau at the N-terminal End. *Front. Aging Neurosci.* **12**, (2020).

334. Schoch, K. M. *et al.* Increased 4R-Tau Induces Pathological Changes in a Human-Tau Mouse Model. *Neuron* **90**, 941–947 (2016).

335. Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat. Rev. Neurol.* **9**, 617–628 (2013).

336. Al-Chalabi, A. *et al.* Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. *Lancet Neurol.* **13**, 1108–1113 (2014).

337. Lo, C. W. Localization of low abundance dna sequences in tissue sections by in situ hybridization. *J. Cell Sci.* **81**, 143–162 (1986).

338. Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science* **264**, 1772–1775 (1994).

339. Alexander, G. M. *et al.* Effect of transgene copy number on survival in the G93A SOD1 transgenic mouse model of ALS. *Mol. Brain Res.* **130**, 7–15 (2004).

340. Acevedo-Arozena, A. *et al.* A comprehensive assessment of the SOD1G93A low-copy transgenic mouse, which models human amyotrophic lateral sclerosis. *Dis. Model. Mech.* **4**, 686–700 (2011).

341. Molnar-Kasza, A. *et al.* Evaluation of Neuropathological Features in the SOD1-G93A Low Copy Number Transgenic Mouse Model of Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.* **14**, (2021).

342. Zwiegers, P., Lee, G. & Shaw, C. A. Reduction in hSOD1 copy number significantly impacts ALS phenotype presentation in G37R (line 29) mice: Implications for the assessment of putative therapeutic agents. *J. Negat. Results Biomed.* **13**, 1–9 (2014).

343. Graffmo, K. S. *et al.* Expression of wild-type human superoxide dismutase-1 in mice causes amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **22**, 51–60 (2013).

344. Afshar, P. *et al.* Overexpression of human SOD1 leads to discrete defects in the cerebellar architecture in the mouse. *Front. Neuroanat.* **11**, 1–13 (2017).

345. Deitch, J. S. *et al.* Phenotype of Transgenic Mice Carrying a Very Low Copy Number of the Mutant Human G93A Superoxide Dismutase-1 Gene Associated with Amyotrophic Lateral Sclerosis. *PLOS ONE* **9**, e99879 (2014).

346. Chandler, K. J. *et al.* Relevance of BAC transgene copy number in mice: transgene copy number variation across multiple transgenic lines and correlations with transgene integrity and expression. *Mamm. Genome* **18**, 693–708 (2007).

347. De Giorgio, F., Maduro, C., Fisher, E. M. C. & Acevedo-Arozena, A. Transgenic and physiological mouse models give insights into different aspects of amyotrophic lateral sclerosis. *DMM Dis. Models Mech.* **12**, (2019).

348. Guo, Y. *et al.* HO-1 induction in motor cortex and intestinal dysfunction in TDP-43 A315T transgenic mice. *Brain Res.* **1460**, 88–95 (2012).

349. Hatzipetros, T. *et al.* C57BL/6J congenic Prp-TDP43A315T mice develop progressive neurodegeneration in the myenteric plexus of the colon without exhibiting key features of ALS. *Brain Res.* **1584**, 59–72 (2014).

350. Goodwin, L. O. *et al.* Large-scale discovery of mouse transgenic integration sites reveals frequent structural variation and insertional mutagenesis. *Genome Res.* **29**, 494–505 (2019).

351. Achilli, F. *et al.* The SOD1 transgene in the G93A mouse model of amyotrophic lateral sclerosis lies on distal mouse chromosome 12. *Amyotroph. Lateral Scler.* **6**, 111–114 (2005).

352. Srivastava, A. *et al.* Discovery of transgene insertion sites by high throughput sequencing of mate pair libraries. *BMC Genomics* **15**, 367 (2014).

353. Acevedo-Arozena, A. *et al.* ENU Mutagenesis, a Way Forward to Understand Gene Function. *Annu. Rev. Genomics Hum. Genet.* **9**, 49–69 (2008).

354. Joyce, P. I. *et al.* A novel SOD1-ALS mutation separates central and peripheral effects of mutant SOD1 toxicity. *Hum. Mol. Genet.* **24**, 1883–1897 (2015).

355. Fratta, P. *et al.* Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J.* **37**, e98684 (2018).

356. White, M. A. *et al.* TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat. Neurosci.* **21**, 552–563 (2018).

357. Casola, S. Mouse Models for miRNA Expression: The ROSA26 Locus. in *MicroRNAs and the Immune System: Methods and Protocols* (ed. Monticelli, S.) 145–163 (Humana Press, Totowa, NJ, 2010). doi:10.1007/978-1-60761-811-9_10.

358. Gordon, D. *et al.* Single-copy expression of an amyotrophic lateral sclerosis-linked TDP-43 mutation (M337V) in BAC transgenic mice leads to altered stress granule dynamics and progressive motor dysfunction. *Neurobiol. Dis.* **121**, 148–162 (2019).

359. Nair, R. R. *et al.* Uses for humanised mouse models in precision medicine for neurodegenerative disease. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **30**, 173–191 (2019).

360. Zhu, F., Nair, R. R., Fisher, E. M. C. & Cunningham, T. J. Humanising the mouse genome piece by piece. *Nat. Commun.* **10**, 1845 (2019).

361. Devoy, A. *et al.* Humanized mutant FUS drives progressive motor neuron degeneration without aggregation in 'FUSDelta14' knockin mice. *Brain* **140**, 2797–2805 (2017).

362. Devoy, A. *et al.* Generation and analysis of innovative genomically humanized knockin SOD1, TARDBP (TDP-43), and FUS mouse models. *iScience* **24**, 103463 (2021).

363. Wong, P. C. *et al.* An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria. *Neuron* **14**, 1105–1116 (1995).

364. Chang-Hong, R. *et al.* Neuroprotective effect of oxidized galectin-1 in a transgenic mouse model of amyotrophic lateral sclerosis. *Exp. Neurol.* **194**, 203–211 (2005).

365. Wang, J. *et al.* Fibrillar Inclusions and Motor Neuron Degeneration in Transgenic Mice Expressing Superoxide Dismutase 1 with a Disrupted Copper-Binding Site. *Neurobiol. Dis.* **10**, 128–138 (2002).

366. Lino, M. M., Schneider, C. & Caroni, P. Accumulation of SOD1 Mutants in Postnatal Motoneurons Does Not Cause Motoneuron Pathology or Motoneuron Disease. *J. Neurosci.* **22**, 4825–4832 (2002).

367. Ripps, M. E., Huntley, G. W., Hof, P. R., Morrison, J. H. & Gordon, J. W. Transgenic mice expressing an altered murine superoxide dismutase gene provide an animal model of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **92**, 689–693 (1995).

368. Jonsson, P. A. *et al.* Disulphide-reduced superoxide dismutase-1 in CNS of transgenic amyotrophic lateral sclerosis models. *Brain* **129**, 451–464 (2006).

369. Wang, J. *et al.* Somatodendritic accumulation of misfolded SOD1-L126Z in motor neurons mediates degeneration: αB-crystallin modulates aggregation. *Hum. Mol. Genet.* **14**, 2335–2347 (2005).

370. Watanabe, Y. *et al.* Mouse motor neuron disease caused by truncated SOD1 with or without C-terminal modification. *Mol. Brain Res.* **135**, 12–20 (2005).

371. Han-Xiang, D. *et al.* Molecular dissection of ALS-associated toxicity of SOD1 in transgenic mice using an exon-fusion approach. *Hum. Mol. Genet.* **17**, 2310–2319 (2008).

372. Dominov, J. A. *et al.* Up-regulation of cholesterol synthesis pathways and limited neurodegeneration in a knock-in Sod1 mutant mouse model of ALS. *bioRxiv* 2023.05.05.539444 (2023) doi:10.1101/2023.05.05.539444.

373. Chiu, A. Y. *et al.* Age-Dependent Penetrance of Disease in a Transgenic Mouse Model of Familial Amyotrophic Lateral Sclerosis. *Mol. Cell. Neurosci.* **6**, 349–362 (1995).

374. Amendola, J., Verrier, B., Roubertoux, P. & Durand, J. Altered sensorimotor development in a transgenic mouse model of amyotrophic lateral sclerosis. *Eur. J. Neurosci.* **20**, 2822–2826 (2004).

375. Zundert, B. van *et al.* Neonatal Neuronal Circuitry Shows Hyperexcitable Disturbance in a Mouse Model of the Adult-Onset Neurodegenerative Disease Amyotrophic Lateral Sclerosis. *J. Neurosci.* **28**, 10864–10874 (2008).

376. Vinsant, S. *et al.* Characterization of early pathogenesis in the SOD1G93A mouse model of ALS: part II, results and discussion. *Brain Behav.* **3**, 431–457 (2013).

377. Alves, C. J. *et al.* Early motor and electrophysiological changes in transgenic mouse model of amyotrophic lateral sclerosis and gender differences on clinical outcome. *Brain Res.* **1394**, 90–104 (2011).

378. Mancuso, R., Oliván, S., Osta, R. & Navarro, X. Evolution of gait abnormalities in SOD1G93A transgenic mice. *Brain Res.* **1406**, 65–73 (2011).

379. Azzouz, M. *et al.* Progressive motor neuron impairment in an animal model of familial amyotrophic lateral sclerosis. *Muscle Nerve* **20**, 45–51 (1997).

380. Fischer, L. R. *et al.* Amyotrophic lateral sclerosis is a distal axonopathy: evidence in mice and man. *Exp. Neurol.* **185**, 232–240 (2004).

381. Gerber, Y. N., Sabourin, J.-C., Rabano, M., Vivanco, M. d M. & Perrin, F. E. Early Functional Deficit and Microglial Disturbances in a Mouse Model of Amyotrophic Lateral Sclerosis. *PLoS ONE* **7**, e36000 (2012).

382. Mead, R. J. *et al.* Optimised and Rapid Pre-clinical Screening in the SOD1G93A Transgenic Mouse Model of Amyotrophic Lateral Sclerosis (ALS). *PLoS ONE* **6**, e23244 (2011).

383. Barnéoud, P., Lolivier, J., Sanger, D. J., Scatton, B. & Moser, P. Quantitative motor assessment in FALS mice: a longitudinal study. *NeuroReport* **8**, 2861 (1997).

384. Hegedus, J., Putman, C. T. & Gordon, T. Time course of preferential motor unit loss in the SOD1G93A mouse model of amyotrophic lateral sclerosis. *Neurobiol. Dis.* **28**, 154–164 (2007).

385. Hegedus, J., Putman, C. T., Tyreman, N. & Gordon, T. Preferential motor unit loss in the SOD1G93A transgenic mouse model of amyotrophic lateral sclerosis. *J. Physiol.* **586**, 3337–3351 (2008).

386. Gal, J., Ström, A.-L., Kilty, R., Zhang, F. & Zhu, H. p62 accumulates and enhances aggregate formation in model systems of familial amyotrophic lateral sclerosis. *J. Biol. Chem.* **282**, 11068–11077 (2007).

387. Stieber, A., Gonatas, J. O. & Gonatas, N. K. Aggregation of ubiquitin and a mutant ALS-linked SOD1 protein correlate with disease progression and fragmentation of the Golgi apparatus. *J. Neurol. Sci.* **173**, 53–62 (2000).

388. Tu, P. H. *et al.* Transgenic mice carrying a human mutant superoxide dismutase transgene develop neuronal cytoskeletal pathology resembling human amyotrophic lateral sclerosis lesions. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 3155–3160 (1996).

389. Gould, T. W. *et al.* Complete Dissociation of Motor Neuron Death from Motor Dysfunction by Bax Deletion in a Mouse Model of ALS. *J. Neurosci.* **26**, 8774–8786 (2006).

390. Jiang, J. *et al.* Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* **90**, 535–550 (2016).

391. Riemslagh, F. W. *et al.* Inducible expression of human C9ORF72 36× G4C2 hexanucleotide repeats is sufficient to cause RAN translation and rapid muscular atrophy in mice. *Dis. Model. Mech.* **14**, dmm044842 (2021).

392. Zhang, Y.-J. *et al.* Poly(GR) impairs protein translation and stress granule dynamics in C9orf72 -associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nat. Med.* **24**, 1136–1142 (2018).

393. Choi, S. Y. *et al.* C9ORF72-ALS/FTD-associated poly(GR) binds Atp5a1 and compromises mitochondrial function in vivo. *Nat. Neurosci.* **22**, 851–862 (2019).

394. Milioto, C. *et al.* PolyGR and polyPR knock-in mice reveal a conserved neuroprotective extracellular matrix signature in C9orf72 ALS/FTD neurons. *Nat. Neurosci.* **27**, 643–655 (2024).

395. Nguyen, L. *et al.* Survival and Motor Phenotypes in FVB C9-500 ALS/FTD BAC Transgenic Mice Reproduced by Multiple Labs. *Neuron* **108**, 784-796.e3 (2020).

396. Mordes, D. A. *et al.* Absence of Survival and Motor Deficits in 500 Repeat C9ORF72 BAC Mice. *Neuron* **108**, 775-783.e4 (2020).

397. 029099 - C9-500 Strain Details. https://www.jax.org/strain/029099.

398. Crabbe, J. C., Wahlsten, D. & Dudek, B. C. Genetics of Mouse Behavior: Interactions with Laboratory Environment. *Science* **284**, 1670–1672 (1999).

399. Richter, S. H., Garner, J. P. & Würbel, H. Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nat. Methods* **6**, 257–261 (2009).

400. Alegre, M.-L. Mouse microbiomes: overlooked culprits of experimental variability. *Genome Biol.* **20**, 108 (2019).

401. Blacher, E. *et al.* Potential roles of gut microbiome and metabolites in modulating ALS in mice. *Nature* **572**, 474–480 (2019).

402. Kohnken, R. A. & Schwahn, D. J. Lack of Chronic Histologic Lesions Supportive of Sublethal Spontaneous Seizures in FVB/N Mice. *Comp. Med.* **66**, 105–111 (2016).

403. Nguyen, L. *et al.* Antibody Therapy Targeting RAN Proteins Rescues C9 ALS/FTD Phenotypes in C9orf72 Mouse Model. *Neuron* **105**, 645-662.e11 (2020).

404. Pattamatta, A. *et al.* Repeat length increases disease penetrance and severity in C9orf72 ALS/FTD BAC transgenic mice. *Hum. Mol. Genet.* **29**, 3900–3918 (2020).

405. Hendricks, E. *et al.* The C9ORF72 repeat expansion alters neurodevelopment. *Cell Rep.* **42**, 112983 (2023).

406. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

407. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

408. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).

409. Rogers, D. C. *et al.* SHIRPA, a protocol for behavioral assessment: validation for longitudinal study of neurological dysfunction in mice. *Neurosci. Lett.* **306**, 89–92 (2001).

410. Carter, R. J. *et al.* Characterization of Progressive Motor Deficits in Mice Transgenic for the Human Huntington's Disease Mutation. *J. Neurosci.* **19**, 3248–3257 (1999).

411. Mendes, C. S. *et al.* Quantification of gait parameters in freely walking rodents. *BMC Biol.* **13**, (2015).

412. Mandillo, S. *et al.* Early motor deficits in mouse disease models are reliably uncovered using an automated home-cage wheel-running system: a cross-laboratory validation. *Dis. Model. Mech.* **7**, 397–407 (2014).

413. Austin, A. *et al.* Sectioning and Counting of Motor Neurons in the L3 to L6 Region of the Adult Mouse Spinal Cord. *Curr. Protoc.* **2**, e428 (2022).

414. Sleigh, J. N., Burgess, R. W., Gillingwater, T. H. & Cader, M. Z. Morphological analysis of neuromuscular junction development and degeneration in rodent lumbrical muscles. *J. Neurosci. Methods* **227**, 159–165 (2014).

415. Bolatto, C., Olivera-Bravo, S. & Cerri, S. Dissection of Single Skeletal Muscle Fibers for Immunofluorescent and Morphometric Analyses of Whole-Mount Neuromuscular Junctions. *J. Vis. Exp. JoVE* (2021) doi:10.3791/62620.

416. Zetterström, P., Andersen, P. M., Brännström, T. & Marklund, S. L. Misfolded superoxide dismutase-1 in CSF from amyotrophic lateral sclerosis patients. *J. Neurochem.* **117**, 91–99 (2011).

417. Lehmann, M. *et al.* Aggregate-selective antibody attenuates seeded aggregation but not spontaneously evolving disease in SOD1 ALS model mice. *Acta Neuropathol. Commun.* **8**, 161 (2020).

418. Weydert, C. J. & Cullen, J. J. Measurement of superoxide dismutase, catalase and glutathione peroxidase in cultured cells and tissue. *Nat. Protoc.* **5**, 51–66 (2010).

419. Peskin, A. V. & Winterbourn, C. C. Assay of superoxide dismutase activity in a plate assay using WST-1. *Free Radic. Biol. Med.* **103**, 188–191 (2017).

420. Madsen, E. B., Höijer, I., Kvist, T., Ameur, A. & Mikkelsen, M. J. Xdrop: Targeted sequencing of long DNA molecules from low input samples using droplet sorting. *Hum. Mutat.* **41**, 1671–1679 (2020).

421. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

422. de Vree, P. J. P. *et al.* Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. *Nat. Biotechnol.* **32**, 1019–1025 (2014).

423. Pang, Z. *et al.* MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **49**, W388–W396 (2021).

424. Bevilacqua, M. & Bro, R. Can We Trust Score Plots? *Metabolites* **10**, 278 (2020).

425. Tinevez, J.-Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).

426. Mielcarek, M. *et al.* HDAC4-Myogenin Axis As an Important Marker of HD-Related Skeletal Muscle Atrophy. *PLOS Genet.* **11**, e1005021 (2015).

427. Untergasser, A. *et al.* Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).

428. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).

429. Nair, R. R. *et al.* Sizing, stabilising, and cloning repeat-expansions for gene targeting constructs. *Methods San Diego Calif* (2020) doi:10.1016/j.ymeth.2020.07.007.

430. Concordet, J.-P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).

431. Perez, A. R. *et al.* GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* **35**, 347–349 (2017).

432. Hofemeister, H. *et al.* Recombineering, transfection, Western, IP and ChIP methods for protein tagging via gene targeting or BAC transgenesis. *Methods* **53**, 437–452 (2011).

433. Elroy-Stein, O., Bernstein, Y. & Groner, Y. Overproduction of human Cu/Zn-superoxide dismutase in transfected cells: extenuation of paraquat-mediated cytotoxicity and enhancement of lipid peroxidation. *EMBO J.* **5**, 615–622 (1986).

434. Karch, C. M. & Borchelt, D. R. AGGREGATION MODULATING ELEMENTS IN MUTANT HUMAN SUPEROXIDE DISMUTASE 1†. *Arch. Biochem. Biophys.* **503**, 175–182 (2010).

435. Grad, L. I. *et al.* Intermolecular transmission of superoxide dismutase 1 misfolding in living cells. *Proc. Natl. Acad. Sci.* **108**, 16398–16403 (2011).

436. Cozzolino, M. *et al.* Cysteine 111 affects aggregation and cytotoxicity of mutant Cu,Zn-superoxide dismutase associated with familial amyotrophic lateral sclerosis. *J. Biol. Chem.* **283**, 866–874 (2008).

437. Karch, C. M. & Borchelt, D. R. A Limited Role for Disulfide Cross-linking in the Aggregation of Mutant SOD1 Linked to Familial Amyotrophic Lateral Sclerosis*. *J. Biol. Chem.* **283**, 13528–13537 (2008).

438. Delahaye, C. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. *PLOS ONE* **16**, e0257521 (2021).

439. Muller, F. L. *et al.* Absence of CuZn superoxide dismutase leads to elevated oxidative stress and acceleration of age-dependent skeletal muscle atrophy. *Free Radic. Biol. Med.* **40**, 1993–2004 (2006).

440. Nardo, G. *et al.* Counteracting roles of MHCI and CD8+ T cells in the peripheral and central nervous system of ALS SOD1G93A mice. *Mol. Neurodegener.* **13**, 42 (2018).

441. Zhang, S. *et al.* The superoxide dismutase 1 (SOD1) 3'UTR maintains high expression of the SOD1 gene in cancer cells: the involvement of the RNA binding protein AUF-1. *Free Radic. Biol. Med.* **85**, 33–44 (2015).

442. Wang, Q. *et al.* MicroRNA-377 is up-regulated and can lead to increased fibronectin production in diabetic nephropathy. *FASEB J.* **22**, 4126–4135 (2008).

443. Wang, L., Xu, J., Liu, H., Li, J. & Hao, H. PM2.5 inhibits SOD1 expression by up-regulating microRNA-206 and promotes ROS accumulation and disease progression in asthmatic mice. *Int. Immunopharmacol.* **76**, 105871 (2019).

444. Wang, X. *et al.* Knockouts of SOD1 and GPX1 Exert Different Impacts on Murine Islet Function and Pancreatic Integrity. *Antioxid. Redox Signal.* **14**, 391–401 (2011).

445. Larkin, L. M. *et al.* Skeletal muscle weakness due to deficiency of CuZn-superoxide dismutase is associated with loss of functional innervation. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **301**, R1400–R1407 (2011).

446. Flood, D. G. *et al.* Hindlimb Motor Neurons Require Cu/Zn Superoxide Dismutase for Maintenance of Neuromuscular Junctions. *Am. J. Pathol.* **155**, 663–672 (1999).

447. Sakellariou, G. K. *et al.* Neuron-specific expression of CuZnSOD prevents the loss of muscle mass and function that occurs in homozygous CuZnSOD-knockout mice. *FASEB J.* **28**, 1666–1681 (2014).

448. Zhang, Y. *et al.* CuZnSOD gene deletion targeted to skeletal muscle leads to loss of contractile force but does not cause muscle atrophy in adult mice. *FASEB J.* **27**, 3536–3548 (2013).

449. Sakellariou, G. K. *et al.* Comparison of Whole Body SOD1 Knockout with Muscle-Specific SOD1 Knockout Mice Reveals a Role for Nerve Redox Signaling in Regulation of Degenerative Pathways in Skeletal Muscle. *Antioxid. Redox Signal.* **28**, 275–295 (2018).

450. Fischer, L. R. *et al.* SOD1 targeted to the mitochondrial intermembrane space prevents motor neuropathy in the Sod1 knockout mouse. *Brain* **134**, 196–209 (2011).

451. Fischer, L. R., Li, Y., Asress, S. A., Jones, D. P. & Glass, J. D. Absence of SOD1 leads to oxidative stress in peripheral nerve and causes a progressive distal motor axonopathy. *Exp. Neurol.* **233**, 163–171 (2012).

452. Rosen, D. R. *et al.* A frequent ala 4 to val superoxide dismutase-1 mutation is associated with a rapidly progressive familial amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **3**, 981–987 (1994).

453. Cudkowicz, M. E. *et al.* Epidemiology of mutations in superoxide dismutase in amyotrophic lateal sclerosis. *Ann. Neurol.* **41**, 210–221 (1997).

454. Juneja, T., Pericak-Vance, M. A., Laing, N. G., Dave, S. & Siddique, T. Prognosis in Familial Amyotrophic Lateral Sclerosis: Progression and Survival in Patients with Glu100gly and Ala4val Mutations in Cu,Zn Superoxide Dismutase. *Neurology* **48**, 55–57 (1997).

455. Hough, M. A. *et al.* Dimer destabilization in superoxide dismutase may result in disease-causing properties: Structures of motor neuron disease mutants. *Proc. Natl. Acad. Sci.* **101**, 5976–5981 (2004).

456. Rakhit, R. *et al.* Monomeric Cu,Zn-superoxide Dismutase Is a Common Misfolding Intermediate in the Oxidation Models of Sporadic and Familial Amyotrophic Lateral Sclerosis*[boxs]. *J. Biol. Chem.* **279**, 15499–15504 (2004).

457. Borchelt, D. R. *et al.* Superoxide dismutase 1 with mutations linked to familial amyotrophic lateral sclerosis possesses significant activity. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 8292–8296 (1994).

458. Bowling, A. C., Schulz, J. B., Brown Jr., R. H. & Beal, M. F. Superoxide Dismutase Activity, Oxidative Damage, and Mitochondrial Energy Metabolism in Familial and Sporadic Amyotrophic Lateral Sclerosis. *J. Neurochem.* **61**, 2322–2325 (1993).

459. Ratovitski, T. *et al.* Variation in the Biochemical/Biophysical Properties of Mutant Superoxide Dismutase 1 Enzymes and the Rate of Disease Progression in Familial Amyotrophic Lateral Sclerosis Kindreds. *Hum. Mol. Genet.* **8**, 1451–1460 (1999).

460. Tiwari, A. & Hayward, L. J. Familial Amyotrophic Lateral Sclerosis Mutants of Copper/Zinc Superoxide Dismutase Are Susceptible to Disulfide Reduction*. *J. Biol. Chem.* **278**, 5984–5992 (2003).

461. Farrawell, N. E. & Yerbury, J. J. Mutant Cu/Zn Superoxide Dismutase (A4V) Turnover Is Altered in Cells Containing Inclusions. *Front. Mol. Neurosci.* **14**, (2021).

462. Kiskinis, E. *et al.* Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell* **14**, 781–795 (2014).

463. Fernández-Beltrán, L. C. *et al.* A Transcriptomic Meta-Analysis Shows Lipid Metabolism Dysregulation as an Early Pathological Mechanism in the Spinal Cord of SOD1 Mice. *Int. J. Mol. Sci.* **22**, 9553 (2021).

464. Wolf, N. I., van Spaendonk, R. M., Hobson, G. M. & Kamholz, J. PLP1 Disorders. in *GeneReviews®* (eds. Adam, M. P. et al.) (University of Washington, Seattle, Seattle (WA), 1993).

465. Muffat, J. & Walker, D. W. Apolipoprotein D. *Cell Cycle Georget. Tex* **9**, 269–273 (2010).

466. Jones, S. F. & Infante, J. R. Molecular Pathways: Fatty Acid Synthase. *Clin. Cancer Res.* **21**, 5434–5438 (2015).

467. Li, G. *et al.* Adipose tissue–specific ablation of Ces1d causes metabolic dysregulation in mice. *Life Sci. Alliance* **5**, (2022).

468. Deguise, M. *et al.* Abnormal fatty acid metabolism is a core component of spinal muscular atrophy. *Ann. Clin. Transl. Neurol.* **6**, 1519–1532 (2019).

469. Andrés-Benito, P. *et al.* Lipid alterations in human frontal cortex in ALS-FTLD-TDP43 proteinopathy spectrum are partly related to peroxisome impairment. *Neuropathol. Appl. Neurobiol.* **47**, 544–563 (2021).

470. Perez, K. *et al.* Single nuclei profiling identifies cell specific markers of skeletal muscle aging, frailty, and senescence. *Aging* **14**, 9393–9422 (2022).

471. Pan, L. INTEGRATIVE OMICS DATA ANALYSIS TO DISCOVER NOVEL SIGNATURES IN COMPLEX DISEASES. (Karolinska Institutet, 2023).

472. Bernardini, C. *et al.* Mitochondrial Network Genes in the Skeletal Muscle of Amyotrophic Lateral Sclerosis Patients. *PLoS ONE* **8**, e57739 (2013).

473. Capitanio, D. *et al.* Molecular Signatures of Amyotrophic Lateral Sclerosis Disease Progression in Hind and Forelimb Muscles of an SOD1G93A Mouse Model. *Antioxid. Redox Signal.* **17**, 1333–1350 (2012).

474. Tefera, T. W., Steyn, F. J., Ngo, S. T. & Borges, K. CNS glucose metabolism in Amyotrophic Lateral Sclerosis: a therapeutic target? *Cell Biosci.* **11**, 14 (2021).

475. Hop, P. J. *et al.* Genome-wide study of DNA methylation shows alterations in metabolic, inflammatory, and cholesterol pathways in ALS. *Sci. Transl. Med.* **14**, eabj0264 (2022).

476. Maksimovic, K., Youssef, M., You, J., Sung, H.-K. & Park, J. Evidence of Metabolic Dysfunction in Amyotrophic Lateral Sclerosis (ALS) Patients and Animal Models. *Biomolecules* **13**, 863 (2023).

477. Xia, K. *et al.* Mutation-specific metabolic profiles in presymptomatic amyotrophic lateral sclerosis. *Eur. J. Neurol.* **30**, 87–95 (2023).

478. Niessen, H. G. *et al.* Metabolic progression markers of neurodegeneration in the transgenic G93A-SOD1 mouse model of amyotrophic lateral sclerosis. *Eur. J. Neurosci.* **25**, 1669–1677 (2007).

479. Bame, M., Grier, R. E., Needleman, R. & Brusilow, W. S. A. Amino Acids as biomarkers in the SOD1G93A mouse model of ALS. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1842**, 79–87 (2014).

480. Henriques, A. *et al.* Amyotrophic lateral sclerosis and denervation alter sphingolipids and up-regulate glucosylceramide synthase. *Hum. Mol. Genet.* **24**, 7390–7405 (2015).

481. Lee, H. *et al.* Multi-omic analysis of selectively vulnerable motor neuron subtypes implicates altered lipid metabolism in ALS. *Nat. Neurosci.* **24**, 1673–1685 (2021).

482. Vaz, F. M. & Wanders, R. J. A. Carnitine biosynthesis in mammals. *Biochem. J.* **361**, 417–429 (2002).

483. Sharma, S. & Black, S. M. CARNITINE HOMEOSTASIS, MITOCHONDRIAL FUNCTION, AND CARDIOVASCULAR DISEASE. *Drug Discov. Today Dis. Mech.* **6**, e31–e39 (2009).

484. Dodge, J. C. *et al.* Neutral Lipid Cacostasis Contributes to Disease Pathogenesis in Amyotrophic Lateral Sclerosis. *J. Neurosci.* **40**, 9137–9147 (2020).

485. Blasco, H. *et al.* Lipidomics Reveals Cerebrospinal-Fluid Signatures of ALS. *Sci. Rep.* **7**, 17652 (2017).

486. González De Aguilar, J.-L. Lipid Biomarkers for Amyotrophic Lateral Sclerosis. *Front. Neurol.* **10**, (2019).

487. Goutman, S. A. *et al.* Metabolomics identifies shared lipid pathways in independent amyotrophic lateral sclerosis cohorts. *Brain* **145**, 4425–4439 (2022).

488. Hartmann, H., Ho, W. Y., Chang, J.-C. & Ling, S.-C. Cholesterol dyshomeostasis in amyotrophic lateral sclerosis: cause, consequence, or epiphenomenon? *FEBS J.* **289**, 7688–7709 (2022).

489. Jia, R. *et al.* Characteristics of serum metabolites in sporadic amyotrophic lateral sclerosis patients based on gas chromatography-mass spectrometry. *Sci. Rep.* **11**, 20786 (2021).

490. Lanznaster, D. *et al.* Metabolic Profile and Pathological Alterations in the Muscle of Patients with Early-Stage Amyotrophic Lateral Sclerosis. *Biomedicines* **10**, 1307 (2022).

491. Worley, B. & Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **1**, 92–107 (2013).

492. Owen, J. B. & Butterfield, D. A. Measurement of Oxidized/Reduced Glutathione Ratio. in *Protein Misfolding and Cellular Stress in Disease and Aging: Concepts and Protocols* (eds. Bross, P. & Gregersen, N.) 269–277 (Humana Press, Totowa, NJ, 2010). doi:10.1007/978-1-60761-756-3_18.

493. Rice, M. E. Ascorbate regulation and its neuroprotective role in the brain. *Trends Neurosci.* **23**, 209–216 (2000).

494. Rice, M. E. & Russo-Menna, I. Differential compartmentalization of brain ascorbate and glutathione between neurons and glia. *Neuroscience* **82**, 1213–1223 (1998).

495. Wang, S. K., Weaver, J. D., Zhang, S. & Lei, X. G. Knockout of SOD1 promotes conversion of selenocysteine to dehydroalanine in murine hepatic GPX1 protein. *Free Radic. Biol. Med.* **51**, 197–204 (2011).

496. Blasco, H. *et al.* Panel of Oxidative Stress and Inflammatory Biomarkers in ALS: A Pilot Study. *Can. J. Neurol. Sci. J. Can. Sci. Neurol.* **44**, 90–95 (2017).

497. Tohgi, H. *et al.* Increase in oxidized NO products and reduction in oxidized glutathione in cerebrospinal fluid from patients with sporadic form of amyotrophic lateral sclerosis. *Neurosci. Lett.* **260**, 204–206 (1999).

498. Muyderman, H. & Chen, T. Mitochondrial dysfunction in amyotrophic lateral sclerosis – a valid pharmacological target? *Br. J. Pharmacol.* **171**, 2191–2205 (2014).

499. Tafuri, F., Ronchi, D., Magri, F., Comi, G. P. & Corti, S. SOD1 misplacing and mitochondrial dysfunction in amyotrophic lateral sclerosis pathogenesis. *Front. Cell. Neurosci.* **9**, 1–12 (2015).

500. Doetschman, T. Influence of Genetic Background on Genetically Engineered Mouse Phenotypes. *Methods Mol. Biol. Clifton NJ* **530**, 423–433 (2009).

501. Heiman-Patterson, T. D. *et al.* Background and gender effects on survival in the TgN(SOD1-G93A)1Gur mouse model of ALS. *J. Neurol. Sci.* **236**, 1–7 (2005).

502. Heiman-Patterson, T. D. *et al.* Effect of genetic background on phenotype variability in transgenic mouse models of amyotrophic lateral sclerosis: A window of opportunity in the search for genetic modifiers. *Amyotroph. Lateral Scler.* **12**, 79–86 (2011).

503. Mancuso, R. *et al.* Effect of genetic background on onset and disease progression in the SOD1-G93A model of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **13**, 302–310 (2012).

504. Nardo, G. *et al.* New insights on the mechanisms of disease course variability in ALS from mutant SOD1 mouse models. *Brain Pathol.* **26**, 237–247 (2016).

505. Simpson, E. M. *et al.* Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat. Genet.* **16**, 19–27 (1997).

506. Threadgill, D. W., Yee, D., Matin, A., Nadeau, J. H. & Magnuson, T. Genealogy of the 129 inbred strains: 129/SvJ is a contaminated inbred strain. *Mamm. Genome* **8**, 390–393 (1997).

507. Pizzasegola, C. *et al.* Treatment with lithium carbonate does not improve disease progression in two different strains of SOD1 mutant mice. *Amyotroph. Lateral Scler.* **10**, 221–228 (2009).

508. Nardo, G. *et al.* Transcriptomic indices of fast and slow disease progression in two mouse models of amyotrophic lateral sclerosis. *Brain* **136**, 3305–3332 (2013).

509. Marino, M. *et al.* Differences in protein quality control correlate with phenotype variability in 2 mouse models of familial amyotrophic lateral sclerosis. *Neurobiol. Aging* **36**, 492–504 (2015).

510. Verma, S. *et al.* Neuromuscular Junction Dysfunction in Amyotrophic Lateral Sclerosis. *Mol. Neurobiol.* **59**, 1502–1527 (2022).

511. Frey, D. *et al.* Early and Selective Loss of Neuromuscular Synapse Subtypes with Low Sprouting Competence in Motoneuron Diseases. *J. Neurosci.* **20**, 2534–2542 (2000).

512. Comley, L. H. *et al.* Motor unit recovery following Smn restoration in mouse models of spinal muscular atrophy. *Hum. Mol. Genet.* **31**, 3107–3119 (2022).

513. Filali, M., Lalonde, R. & Rivest, S. Sensorimotor and cognitive functions in a SOD1 G37R transgenic mouse model of amyotrophic lateral sclerosis. *Behav. Brain Res.* **225**, 215–221 (2011).

514. Hadzipasic, M. *et al.* Selective degeneration of a physiological subtype of spinal motor neuron in mice with SOD1-linked ALS. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16883–16888 (2014).

515. OLIVÁN, S. *et al.* Comparative study of behavioural tests in the SOD1G93A mouse model of amyotrophic lateral sclerosis. *Exp. Anim.* **64**, 147–153 (2015).

516. Dadon-Nachum, M., Melamed, E. & Offen, D. The "Dying-Back" Phenomenon of Motor Neurons in ALS. *J. Mol. Neurosci.* **43**, 470–477 (2011).

517. Surana, S. *et al.* The travel diaries of tetanus and botulinum neurotoxins. *Toxicon* **147**, 58–67 (2018).

518. Mejia Maza, A. *et al.* NMJ-Analyser identifies subtle early changes in mouse models of neuromuscular disease. *Sci. Rep.* **11**, 12251 (2021).

519. Alhindi, A., Boehm, I. & Chaytow, H. Small junction, big problems: Neuromuscular junction pathology in mouse models of amyotrophic lateral sclerosis (ALS). *J. Anat.* **241**, 1089–1107 (2022).

520. Korobeynikov, V. A., Lyashchenko, A. K., Blanco-Redondo, B., Jafar-Nejad, P. & Shneider, N. A. Antisense oligonucleotide silencing of FUS expression as a therapeutic approach in amyotrophic lateral sclerosis. *Nat. Med.* **28**, 104–116 (2022).

521. Kurahashi, T. *et al.* A malfunction in triglyceride transfer from the intracellular lipid pool to apoB in enterocytes of SOD1-deficient mice. *FEBS Lett.* **586**, 4289–4295 (2012).

522. Lee, J., Homma, T., Kobayashi, S., Ishii, N. & Fujii, J. Unveiling systemic organ disorders associated with impaired lipid catabolism in fasted SOD1-deficient mice. *Arch. Biochem. Biophys.* **654**, 163–171 (2018).

523. Muscogiuri, G. *et al.* Genetic Disruption of SOD1 Gene Causes Glucose Intolerance and Impairs β-Cell Function. *Diabetes* **62**, 4201–4207 (2013).

524. Wang, L., Jiang, Z. & Lei, X. G. Knockout of SOD1 altered murine hepatic glycolysis, gluconeogenesis, and lipogenesis. *Free Radic. Biol. Med.* **53**, 1689–1696 (2012).

525. Uchiyama, S., Shimizu, T. & Shirasawa, T. CuZn-SOD Deficiency Causes ApoB Degradation and Induces Hepatic Lipid Accumulation by Impaired Lipoprotein Secretion in Mice*. *J. Biol. Chem.* **281**, 31713–31719 (2006).

526. Kurahashi, T. *et al.* An SOD1 deficiency enhances lipid droplet accumulation in the fasted mouse liver by aborting lipophagy. *Biochem. Biophys. Res. Commun.* **467**, 866–871 (2015).

527. Zhu, Y., Carvey, P. M. & Ling, Z. Age-related changes in glutathione and glutathione-related enzymes in rat brain. *Brain Res.* **1090**, 35–44 (2006).

528. Levy, M. A., Tsai, Y.-H., Reaume, A. & Bray, T. M. Cellular response of antioxidant metalloproteins in Cu/Zn SOD transgenic mice exposed to hyperoxia. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **281**, L172–L182 (2001).

529. Iuchi, Y. *et al.* Elevated oxidative stress in erythrocytes due to a SOD1 deficiency causes anaemia and triggers autoantibody production. *Biochem. J.* **402**, 219–227 (2007).

530. Wiese, S. *et al.* Isolation and enrichment of embryonic mouse motoneurons from the lumbar spinal cord of individual mouse embryos. *Nat. Protoc.* **5**, 31–38 (2010).

531. Kim, R. B., Irvin, C. W., Tilva, K. R. & Mitchell, C. S. State of the field: An informatics-based systematic review of the SOD1-G93A amyotrophic lateral sclerosis transgenic mouse model. *Amyotroph. Lateral Scler. Front. Degener.* **17**, 1–14 (2016).

532. Bendotti, C. *et al.* Early vacuolization and mitochondrial damage in motor neurons of FALS mice are not associated with apoptosis or with changes in cytochrome oxidase histochemical reactivity. *J. Neurol. Sci.* **191**, 25–33 (2001).

533. Mourelatos, Z., Gonatas, N. K., Stieber, A., Gurney, M. E. & Dal Canto, M. C. The Golgi apparatus of spinal cord motor neurons in transgenic mice expressing mutant Cu,Zn superoxide dismutase becomes fragmented in early, preclinical stages of the disease. *Proc. Natl. Acad. Sci.* **93**, 5472–5477 (1996).

534. Petrov, D., Mansfield, C., Moussy, A. & Hermine, O. ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment? *Front. Aging Neurosci.* **9**, 68 (2017).

535. Watanabe, Y., Kato, S., Adachi, Y. & Nakashima, K. Frameshift, nonsense and non amino acid altering mutations in SOD1 in familial ALS: report of a Japanese pedigree and literature review. *Amyotroph. Lateral Scler. Mot. Neuron Disord. Off. Publ. World Fed. Neurol. Res. Group Mot. Neuron Dis.* **1**, 251–258 (2000).

536. Edelheit, O., Hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9**, 61 (2009).

537. Reikofski, J. & Tao, B. Y. Polymerase chain reaction (PCR) techniques for site-directed mutagenesis. *Biotechnol. Adv.* **10**, 535–547 (1992).

538. Zhang, K. *et al.* A high-efficiency method for site-directed mutagenesis of large plasmids based on large DNA fragment amplification and recombinational ligation. *Sci. Rep.* **11**, 10454 (2021).

539. Iacoangeli, A. *et al.* C9orf72 intermediate expansions of 24–30 repeats are associated with ALS. *Acta Neuropathol. Commun.* **7**, 115 (2019).

540. Gois, A. M., Mendonça, D. M. F., Freire, M. A. M. & Santos, J. R. IN VITRO AND IN VIVO MODELS OF AMYOTROPHIC LATERAL SCLEROSIS: AN UPDATED OVERVIEW. *Brain Res. Bull.* **159**, 32–43 (2020).

541. Wang, X. *et al.* Forward Genetic Screen in Caenorhabditis elegans Suggests F57A10.2 and acp-4 As Suppressors of C9ORF72 Related Phenotypes. *Front. Mol. Neurosci.* **9**, 113 (2016).

542. Jiang, S.-W., Trujillo, M. A. & Eberhardt, N. L. An Efficient Method for Generation and Subcloning of Tandemly Repeated DNA Sequences with Defined Length, Orientation and Spacing. *Nucleic Acids Res.* **24**, 3278–3279 (1996).

543. Meyer, D. E. & Chilkoti, A. Genetically Encoded Synthesis of Protein-Based Polymers with Precisely Specified Molecular Weight and Sequence by Recursive Directional Ligation: Examples from the Elastin-like Polypeptide System. *Biomacromolecules* **3**, 357–367 (2002).

544. Ordway, J. M. & Detloff, P. J. In vitro synthesis and cloning of long CAG repeats. *BioTechniques* **21**, 609–610, 612 (1996).

545. Yu, Z. *et al.* Androgen-dependent pathology demonstrates myopathic contribution to the Kennedy disease phenotype in a mouse knock-in model. *J. Clin. Invest.* **116**, 2663–2672 (2006).

546. Rohilla, K. J. *et al.* Systematic microsatellite repeat expansion cloning and validation. *Hum. Genet.* **139**, 1233–1246 (2020).

547. Godiska, R. *et al.* Linear plasmid vector for cloning of repetitive or unstable sequences in Escherichia coli. *Nucleic Acids Res.* **38**, e88 (2010).

548. Bowater, R. P., Rosche, W. A., Jaworski, A., Sinden, R. R. & Wells, R. D. Relationship between Escherichia coli growth and deletions of CTG.CAG triplet repeats in plasmids. *J. Mol. Biol.* **264**, 82–96 (1996).

549. Sarkar, P. S., Chang, H. C., Boudi, F. B. & Reddy, S. CTG repeats show bimodal amplification in E. coli. *Cell* **95**, 531–540 (1998).

550. Trinh, T., Jessee, J., Bloom, F. & Hirsch, V. AN ESCHERICHIA COLI STRAIN FOR THE STABLE PROPAGATION OF RETROVIRAL CLONES AND DIRECT REPEAT SEQUENCES. *Focus* **16**, 78–80 (1994).

551. Jakupciak, J. P. & Wells, R. D. Genetic Instabilities of Triplet Repeat Sequences by Recombination. *IUBMB Life* **50**, 355–359 (2000).

552. Nordin, A. *et al.* Sequence variations in C9orf72 downstream of the hexanucleotide repeat region and its effect on repeat-primed PCR interpretation: a large multinational screening study. *Amyotroph. Lateral Scler. Front. Degener.* **18**, 256–264 (2017).

553. Zhang, Y., Buchholz, F., Muyrers, J. P. P. & Stewart, A. F. A new logic for DNA engineering using recombination in Escherichia coli. *Nat. Genet.* **20**, 123–128 (1998).

554. Yu, D. *et al.* An efficient recombination system for chromosome engineering in Escherichia coli. *Proc. Natl. Acad. Sci.* **97**, 5978–5983 (2000).

555. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009).

556. Rivero-Müller, A., Lajić, S. & Huhtaniemi, I. Assisted large fragment insertion by Red/ET-recombination (ALFIRE)—an alternative and enhanced method for large fragment recombineering. *Nucleic Acids Res.* **35**, e78 (2007).

557. Hosted, T. J. & Baltz, R. H. Use of rpsL for dominance selection and gene replacement in Streptomyces roseosporus. *J. Bacteriol.* **179**, 180–186 (1997).

558. Zhao, S. *et al.* PiggyBac transposon vectors: the tools of the human gene encoding. *Transl. Lung Cancer Res.* **5**, 120–125 (2016).

559. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

560. Bin Moon, S. *et al.* Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3'-overhang. *Nat. Commun.* **9**, (2018).

561. Lei, C. *et al.* The CCTL (Cpf1-assisted Cutting and Taq DNA ligase-assisted Ligation) method for efficient editing of large DNA constructs in vitro. *Nucleic Acids Res.* **45**, 1–7 (2017).

562. Jiang, W. *et al.* Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* **6**, 8101 (2015).

563. McDaniel, J. R., MacKay, J. A., Quiroz, F. G. & Chilkoti, A. Recursive directional ligation by plasmid reconstruction allows rapid and seamless cloning of oligomeric genes. *Biomacromolecules* **11**, 944–952 (2010).

564. Paulson, H. Repeat expansion diseases. *Handb. Clin. Neurol.* **147**, 105–123 (2018).

565. Selvaraj, B. T., Livesey, M. R. & Chandran, S. Modeling the C9ORF72 repeat expansion mutation using human induced pluripotent stem cells. *Brain Pathol.* **27**, 518–524 (2017).

566. Lines, G., Casey, J. M., Preza, E. & Wray, S. Modelling frontotemporal dementia using patient-derived induced pluripotent stem cells. *Mol. Cell. Neurosci.* **109**, 103553 (2020).

567. Foxe, D. *et al.* Intrafamilial Phenotypic Variability in the C9orf72 Gene Expansion: 2 Case Studies. *Front. Psychol.* **9**, (2018).

568. Ende, E. L. van der *et al.* Unravelling the clinical spectrum and the role of repeat length in C9ORF72 repeat expansions. *J. Neurol. Neurosurg. Psychiatry* **92**, 502–509 (2021).

569. Phatnani, H. *et al.* An integrated multi-omic analysis of iPSC-derived motor neurons from C9ORF72 ALS patients. *iScience* **24**, 103221 (2021).

570. Ababneh, N. A. *et al.* Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Hum. Mol. Genet.* **29**, 2200–2217 (2020).

571. Beckers, J. *et al.* A toxic gain-of-function mechanism in C9orf72 ALS impairs the autophagy-lysosome pathway in neurons. *Acta Neuropathol. Commun.* **11**, 151 (2023).

572. Ramos, D. M., Skarnes, W. C., Singleton, A. B., Cookson, M. R. & Ward, M. E. Tackling neurodegenerative diseases with genomic engineering: A new stem cell initiative from the NIH. *Neuron* **109**, 1080–1083 (2021).

573. Kwart, D. *et al.* A Large Panel of Isogenic APP and PSEN1 Mutant Human iPSC Neurons Reveals Shared Endosomal Abnormalities Mediated by APP β-CTFs, Not Aβ. *Neuron* **104**, 256-270.e5 (2019).

574. Huai, J. & Zhang, Z. Structural Properties and Interaction Partners of Familial ALS-Associated SOD1 Mutants. *Front. Neurol.* **10**, (2019).

575. Tsuburaya, N. *et al.* A small-molecule inhibitor of SOD1-Derlin-1 interaction ameliorates pathology in an ALS mouse model. *Nat. Commun.* **9**, 2668 (2018).

576. Kawamura, Y. & Dyck, P. J. Lumbar motoneurons of man: III. The number and diameter distribution of large- and intermediate- diameter cytons by nuclear columns. *J. Neuropathol. Exp. Neurol.* **36**, 956–963 (1977).

577. McHanwell, S. & Biscoe, T. J. The sizes of motoneurons supplying hindlimb muscles in the mouse. *Proc. R. Soc. Lond. B Biol. Sci.* **213**, 201–216 (1981).

578. Manuel, M., Chardon, M., Tysseling, V. & Heckman, C. J. Scaling of Motor Output, From Mouse to Humans. *Physiology* **34**, 5–13 (2019).

579. Yadav, A. *et al.* A cellular taxonomy of the adult human spinal cord. *Neuron* **111**, 328-344.e7 (2023).

580. Jones, R. A. *et al.* Cellular and Molecular Anatomy of the Human Neuromuscular Junction. *Cell Rep.* **21**, 2348–2356 (2017).

581. Vandoorne, T., De Bock, K. & Van Den Bosch, L. Energy metabolism in ALS: an underappreciated opportunity? *Acta Neuropathol. (Berl.)* **135**, 489–509 (2018).

582. Barber, S. C. & Shaw, P. J. Oxidative stress in ALS: Key role in motor neuron injury and therapeutic target. *Free Radic. Biol. Med.* **48**, 629–641 (2010).

583. Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015).

584. Tian, R. *et al.* CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron* **104**, 239-255.e12 (2019).

585. Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).

586. Jang, Y. C. *et al.* Increased superoxide in vivo accelerates age-associated muscle atrophy through mitochondrial dysfunction and neuromuscular junction degeneration. *FASEB J.* **24**, 1376–1390 (2010).

587. Logan, S. *et al.* Accelerated decline in cognition in a mouse model of increased oxidative stress. *GeroScience* **41**, 591–607 (2019).

588. Elchuri, S. *et al.* CuZnSOD deficiency leads to persistent and widespread oxidative damage and hepatocarcinogenesis later in life. *Oncogene* **24**, 367–380 (2005).

589. Zhang, Y. *et al.* Dietary restriction attenuates the accelerated aging phenotype of Sod1−/− mice. *Free Radic. Biol. Med.* **60**, 300–306 (2013).

590. Halliwell, B. & Foyer, C. H. Ascorbic acid, metal ions and the superoxide radical. *Biochem. J.* **155**, 697–700 (1976).

591. Shen, J. *et al.* Ascorbate oxidation by iron, copper and reactive oxygen species: review, model development, and derivation of key rate constants. *Sci. Rep.* **11**, 7417 (2021).

592. Harrison, F. E. & May, J. M. Vitamin C function in the brain: vital role of the ascorbate transporter SVCT2. *Free Radic. Biol. Med.* **46**, 719–730 (2009).

593. Figueroa-Méndez, R. & Rivas-Arancibia, S. Vitamin C in Health and Disease: Its Role in the Metabolism of Cells and Redox State in the Brain. *Front. Physiol.* **6**, (2015).

594. Chatterjee, I. B., Majumder, A. K., Nandi, B. K. & Subramanian, N. Synthesis and Some Major Functions of Vitamin C in Animals*. *Ann. N. Y. Acad. Sci.* **258**, 24–47 (1975).

595. Nishikimi, M., Fukuyama, R., Minoshima, S., Shimizu, N. & Yagi, K. Cloning and chromosomal mapping of the human nonfunctional gene for L-gulono-gamma-lactone oxidase, the enzyme for L-ascorbic acid biosynthesis missing in man. *J. Biol. Chem.* **269**, 13685–13688 (1994).

596. Tsao, C. S. An overview of ascorbic acid chemistry and biochemistry. *Vitam. C Health Dis.* 25–58 (1997).

597. Przybyło, M. & Langner, M. On the physiological and cellular homeostasis of ascorbate. *Cell. Mol. Biol. Lett.* **25**, 32 (2020).

598. Maeda, N. *et al.* Aortic wall damage in mice unable to synthesize ascorbic acid. *Proc. Natl. Acad. Sci.* **97**, 841–846 (2000).

599. Jawaid, A. *et al.* A decrease in body mass index is associated with faster progression of motor symptoms and shorter survival in ALS. *Amyotroph. Lateral Scler.* **11**, 542–548 (2010).

600. Marin, B. *et al.* Alteration of nutritional status at diagnosis is a prognostic factor for survival of amyotrophic lateral sclerosis patients. *J. Neurol. Neurosurg. Psychiatry* **82**, 628–634 (2011).

601. Peter, R. S. *et al.* Life course body mass index and risk and prognosis of amyotrophic lateral sclerosis: results from the ALS registry Swabia. *Eur. J. Epidemiol.* **32**, 901–908 (2017).

602. Dardiotis, E. *et al.* Body mass index and survival from amyotrophic lateral sclerosis. *Neurol. Clin. Pract.* **8**, 437–444 (2018).

603. O'Reilly, É. J. *et al.* Premorbid body mass index and risk of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Front. Degener.* **14**, 205–211 (2013).

604. Nakken, O., Meyer, H. E., Stigum, H. & Holmøy, T. High BMI is associated with low ALS risk: A population-based study. *Neurology* **93**, e424–e432 (2019).

605. Bouteloup, C. *et al.* Hypermetabolism in ALS patients: an early and persistent phenomenon. *J. Neurol.* **256**, 1236–1242 (2009).

606. Fayemendy, P. *et al.* Hypermetabolism is a reality in amyotrophic lateral sclerosis compared to healthy subjects. *J. Neurol. Sci.* **420**, (2021).

607. Sellés Vidal, L., Kelly, C. L., Mordaka, P. M. & Heap, J. T. Review of NAD(P)H-dependent oxidoreductases: Properties, engineering and application. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1866**, 327–347 (2018).

608. Montllor-Albalate, C. *et al.* Sod1 integrates oxygen availability to redox regulate NADPH production and the thiol redoxome. *Proc. Natl. Acad. Sci.* **119**, e2023328119 (2022).

609. Saito, T. *et al.* Single App knock-in mouse models of Alzheimer's disease. *Nat. Neurosci.* **17**, 661–663 (2014).

610. Patel, B. P., Safdar, A., Raha, S., Tarnopolsky, M. A. & Hamadeh, M. J. Caloric Restriction Shortens Lifespan through an Increase in Lipid Peroxidation, Inflammation and Apoptosis in the G93A Mouse, an Animal Model of ALS. *PLOS ONE* **5**, e9386 (2010).

611. Hamadeh, M. J., Rodriguez, M. C., Kaczor, J. J. & Tarnopolsky, M. A. Caloric restriction transiently improves motor performance but hastens clinical onset of disease in the Cu/Zn-superoxide dismutase mutant G93A mouse. *Muscle Nerve* **31**, 214–220 (2005).

612. Coughlan, K. S., Halang, L., Woods, I. & Prehn, J. H. M. A high-fat jelly diet restores bioenergetic balance and extends lifespan in the presence of motor dysfunction and lumbar spinal cord motor neuron loss in TDP-43A315T mutant C57BL6/J mice. *Dis. Model. Mech.* **9**, 1029–1037 (2016).

613. Dupuis, L., Oudart, H., René, F., de Aguilar, J.-L. G. & Loeffler, J.-P. Evidence for defective energy homeostasis in amyotrophic lateral sclerosis: Benefit of a high-energy diet in a transgenic mouse model. *Proc. Natl. Acad. Sci.* **101**, 11159–11164 (2004).

614. Fergani, A. *et al.* Increased peripheral lipid clearance in an animal model of amyotrophic lateral sclerosis. *J. Lipid Res.* **48**, 1571–1580 (2007).

615. Huang, Z. *et al.* Circadian Rhythm Dysfunction Accelerates Disease Progression in a Mouse Model With Amyotrophic Lateral Sclerosis. *Front. Neurol.* **9**, 218 (2018).

616. Scarmeas, N., Shih, T., Stern, Y., Ottman, R. & Rowland, L. P. Premorbid weight, body mass, and varsity athletics in ALS. *Neurology* **59**, 773–775 (2002).

617. Huisman, M. H. B. *et al.* Lifetime physical activity and the risk of amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **84**, 976–981 (2013).

# Appendices

## 8.1 Full lists of differentially expressed genes (DEGs) in *hSOD1$^{A4V}$* mice

| Gene Name | baseMean | log2Fold Change | SE of log2Fold Change | pvalue | padj | Significant comparison |
|---|---|---|---|---|---|---|
| **6-month-old *hSOD1$^{A4V}$* spinal cord** | | | | | | |
| *Hbb-bs* | 1954.43 | 0.74 | 0.10 | 7.28E-14 | 1.67E-09 | WT vs Hom |
| *Apod* | 1908.22 | -0.58 | 0.09 | 5.78E-11 | 5.34E-07 | WT vs Hom |
| *Plp1* | 23050.22 | -0.49 | 0.08 | 6.97E-11 | 5.34E-07 | WT vs Hom |
| *Cacng2* | 754.99 | -0.67 | 0.12 | 5.97E-08 | 3.25E-04 | WT vs Hom |
| *Rpl17* | 643.12 | -0.58 | 0.11 | 7.07E-08 | 3.25E-04 | WT vs Hom |
| *Hbb-bt* | 405.06 | 0.71 | 0.13 | 1.01E-07 | 3.86E-04 | WT vs Hom |
| *Phip* | 176.19 | -0.90 | 0.17 | 2.07E-07 | 6.79E-04 | WT vs Hom |
| *Hba-a1* | 855.17 | 0.58 | 0.12 | 7.63E-07 | 2.19E-03 | WT vs Hom |
| *Syp* | 752.97 | -0.48 | 0.10 | 4.35E-06 | 9.36E-03 | WT vs Hom |
| *Hba-a2* | 607.56 | 0.54 | 0.12 | 4.35E-06 | 9.36E-03 | WT vs Hom |
| *Cfl1* | 886.46 | -0.47 | 0.10 | 4.48E-06 | 9.36E-03 | WT vs Hom |
| *Olig1* | 898.13 | 0.41 | 0.09 | 5.52E-06 | 1.06E-02 | WT vs Hom |
| *Syne1* | 861.13 | 0.40 | 0.09 | 8.14E-06 | 1.44E-02 | WT vs Hom |
| *Map1b* | 6888.16 | -0.32 | 0.07 | 1.52E-05 | 2.49E-02 | WT vs Hom |
| *ND4* | 56406.47 | -0.36 | 0.08 | 1.88E-05 | 2.88E-02 | WT vs Hom |
| *Rpl26* | 864.54 | -0.45 | 0.11 | 2.69E-05 | 3.86E-02 | WT vs Hom |
| *Tmsb4x* | 3207.73 | -0.33 | 0.08 | 3.06E-05 | 3.90E-02 | WT vs Hom |
| *Ap3b2* | 614.06 | -0.47 | 0.11 | 3.06E-05 | 3.90E-02 | WT vs Hom |
| *LOC118568383* | 381.56 | 0.49 | 0.12 | 3.93E-05 | 4.75E-02 | WT vs Hom |
| *Rpl4* | 1224.09 | -0.43 | 0.11 | 4.50E-05 | 5.17E-02 | WT vs Hom |
| *Rps24* | 1684.97 | -0.39 | 0.10 | 4.95E-05 | 5.41E-02 | WT vs Hom |
| *2900060B14Rik* | 255.09 | 0.66 | 0.17 | 5.98E-05 | 6.24E-02 | WT vs Hom |
| *Ptpro* | 24.02 | 1.85 | 0.47 | 8.08E-05 | 7.79E-02 | WT vs Hom |
| *Klf9* | 1670.74 | 0.28 | 0.07 | 8.15E-05 | 7.79E-02 | WT vs Hom |
| *Cox7c* | 1906.48 | -0.29 | 0.07 | 9.34E-05 | 8.45E-02 | WT vs Hom |
| *Cnih2* | 55.42 | -1.10 | 0.28 | 1.03E-04 | 8.45E-02 | WT vs Hom |
| *Gm16287* | 14.73 | 2.52 | 0.65 | 1.04E-04 | 8.45E-02 | WT vs Hom |
| *Ndn* | 612.00 | -0.38 | 0.10 | 1.06E-04 | 8.45E-02 | WT vs Hom |
| *Gm6579* | 4.47 | -5.59 | 1.44 | 1.07E-04 | 8.45E-02 | WT vs Hom |
| *Vamp1* | 2831.94 | -0.34 | 0.09 | 1.21E-04 | 9.16E-02 | WT vs Hom |
| *Gnas* | 5553.10 | -0.27 | 0.07 | 1.24E-04 | 9.16E-02 | WT vs Hom |
| *Apod* | 2021.51 | -0.49 | 0.08 | 2.23E-09 | 5.51E-06 | WT vs Het |
| *mt-Rnr2* | 155301.43 | 0.25 | 0.05 | 3.74E-07 | 4.61E-04 | WT vs Het |
| *Pcsk1n* | 1546.50 | -0.37 | 0.08 | 1.16E-06 | 9.54E-04 | WT vs Het |
| *Vamp1* | 3015.35 | -0.25 | 0.06 | 1.75E-05 | 1.08E-02 | WT vs Het |

| | | | | | | |
|---|---|---|---|---|---|---|
| *mt-Rnr1* | 63020.04 | 0.24 | 0.06 | 2.65E-05 | 1.31E-02 | WT vs Het |
| *Pip5k1a* | 277.57 | 0.53 | 0.13 | 8.04E-05 | 3.31E-02 | WT vs Het |
| *Rhob* | 413.17 | -0.44 | 0.11 | 9.90E-05 | 3.49E-02 | WT vs Het |
| *Mobp* | 5351.61 | -0.23 | 0.06 | 1.15E-04 | 3.55E-02 | WT vs Het |
| *Usf2* | 483.89 | -0.44 | 0.11 | 1.38E-04 | 3.79E-02 | WT vs Het |
| *Tmsb4x* | 3369.27 | -0.28 | 0.07 | 1.61E-04 | 3.97E-02 | WT vs Het |
| *Ccnl1* | 248.53 | 0.55 | 0.15 | 1.89E-04 | 4.25E-02 | WT vs Het |
| *Tmem35a* | 420.78 | 0.52 | 0.14 | 3.18E-04 | 6.14E-02 | WT vs Het |
| *Rpl4* | 1283.19 | -0.38 | 0.11 | 3.23E-04 | 6.14E-02 | WT vs Het |
| *Nisch* | 599.16 | -0.37 | 0.11 | 4.59E-04 | 8.09E-02 | WT vs Het |
| *Syp* | 803.01 | -0.37 | 0.11 | 5.04E-04 | 8.30E-02 | WT vs Het |
| *Slc25a5* | 725.02 | -0.31 | 0.09 | 5.43E-04 | 8.38E-02 | WT vs Het |
| *Eid1* | 545.54 | 0.45 | 0.13 | 6.25E-04 | 8.88E-02 | WT vs Het |
| *Gm2a* | 398.45 | -0.50 | 0.15 | 6.47E-04 | 8.88E-02 | WT vs Het |
| *Hba-a1* | 790.26 | 0.91 | 0.11 | 1.21E-16 | 2.81E-12 | Het vs Hom |
| *Hbb-bs* | 1961.87 | 0.75 | 0.10 | 8.17E-13 | 9.48E-09 | Het vs Hom |
| *Hba-a2* | 575.49 | 0.76 | 0.11 | 9.73E-12 | 7.53E-08 | Het vs Hom |
| *Hbb-bt* | 391.35 | 0.88 | 0.14 | 1.20E-09 | 6.96E-06 | Het vs Hom |
| *Plp1* | 21691.94 | -0.32 | 0.06 | 2.29E-07 | 1.06E-03 | Het vs Hom |
| *Nfkbia* | 386.34 | 0.62 | 0.13 | 8.16E-07 | 3.16E-03 | Het vs Hom |
| *Phip* | 166.81 | -0.76 | 0.16 | 4.14E-06 | 1.37E-02 | Het vs Hom |
| *Arf5* | 63.12 | -1.76 | 0.41 | 2.11E-05 | 6.11E-02 | Het vs Hom |
| *Klf9* | 1655.91 | 0.33 | 0.08 | 2.96E-05 | 7.63E-02 | Het vs Hom |
| *Pou3f1* | 111.39 | 1.14 | 0.27 | 3.35E-05 | 7.79E-02 | Het vs Hom |
| **12-month-old *hSOD1$^{A4V}$* spinal cord** | | | | | | |
| *Hbb-bs* | 859.17 | 0.68 | 0.08 | 1.02E-17 | 2.44E-13 | WT vs Hom |
| *Hbb-bt* | 185.82 | 0.95 | 0.12 | 2.85E-15 | 3.40E-11 | WT vs Hom |
| *Pisd-ps3* | 317.29 | -0.68 | 0.11 | 1.67E-10 | 1.33E-06 | WT vs Hom |
| *Hba-a1* | 99.73 | 0.91 | 0.16 | 1.16E-08 | 6.91E-05 | WT vs Hom |
| *Hba-a2* | 72.98 | 0.86 | 0.19 | 4.33E-06 | 2.07E-02 | WT vs Hom |
| *Hba-a1* | 90.15 | 1.17 | 0.20 | 8.99E-09 | 2.14E-04 | Het vs Hom |
| *Hbb-bt* | 179.74 | 0.91 | 0.16 | 2.86E-08 | 3.40E-04 | Het vs Hom |
| *Hbb-bs* | 789.80 | 0.84 | 0.17 | 4.26E-07 | 3.38E-03 | Het vs Hom |
| *Chchd2* | 1116.73 | -0.40 | 0.08 | 2.41E-06 | 1.43E-02 | Het vs Hom |
| **21-month-old *hSOD1$^{A4V}$* spinal cord** | | | | | | |
| *Fasn* | 696.27 | -1.53 | 0.29 | 1.52E-07 | 2.92E-03 | WT vs Het |
| *Ces1d* | 118.87 | -2.36 | 0.46 | 2.78E-07 | 2.92E-03 | WT vs Het |

**Table a-1 | Full list of differentially expressed genes (DEGs) in the spinal cord of male *hSOD1$^{A4V}$* mice.** SE, standard error; WT, *hSOD1$^{WT}$*/*hSOD1$^{WT}$*; Het, *hSOD1$^{WT}$*/*hSOD1$^{A4V}$*; Hom, *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$*.

| Gene Name | baseMean | log2Fold Change | SE of log2Fold Change | pvalue | padj | Significant comparison |
|---|---|---|---|---|---|---|
| **6-month-old *hSOD1$^{A4V}$* TA muscle** | | | | | | |
| *Des* | 5936.21 | 0.58 | 0.09 | 3.91E-10 | 6.77E-07 | WT vs Het |
| *Hspb8* | 882.00 | 0.78 | 0.13 | 4.99E-10 | 6.77E-07 | WT vs Het |
| *Ddit4l* | 868.17 | -0.86 | 0.16 | 5.24E-08 | 7.79E-05 | Het vs Hom |
| *Asb2* | 1208.56 | -0.84 | 0.16 | 7.62E-08 | 7.79E-05 | Het vs Hom |
| *Ampd1* | 1422.60 | -0.65 | 0.13 | 1.86E-07 | 1.27E-04 | Het vs Hom |
| *Eif4ebp2* | 238.73 | -1.01 | 0.20 | 2.53E-07 | 2.29E-04 | WT vs Het |
| *Trdn* | 10903.95 | 0.78 | 0.16 | 5.05E-07 | 2.95E-04 | WT vs Het |
| *mt-Rnr2* | 252345.53 | -0.74 | 0.15 | 5.44E-07 | 2.95E-04 | WT vs Het |
| *Rpl13a* | 995.89 | 0.67 | 0.13 | 6.81E-07 | 3.08E-04 | WT vs Het |
| *Nceh1* | 154.08 | -1.13 | 0.23 | 1.16E-06 | 4.15E-04 | WT vs Het |
| *Aqp4* | 81.95 | -1.27 | 0.26 | 1.23E-06 | 1.43E-02 | WT vs Hom |
| *Rps27a* | 261.11 | 0.79 | 0.16 | 1.39E-06 | 4.15E-04 | WT vs Het |
| *Rps24* | 1217.05 | 0.79 | 0.16 | 1.44E-06 | 4.15E-04 | WT vs Het |
| *ND2* | 206386.28 | -0.55 | 0.12 | 1.53E-06 | 4.15E-04 | WT vs Het |
| *Ptrhd1* | 131.69 | 1.18 | 0.25 | 2.06E-06 | 5.08E-04 | WT vs Het |
| *1700042G15Rik* | 181.17 | 6.28 | 1.34 | 2.76E-06 | 1.43E-02 | WT vs Hom |
| *Rbm25* | 516.29 | 0.76 | 0.16 | 3.14E-06 | 7.09E-04 | WT vs Het |
| *Sh3rf2* | 58.67 | -1.66 | 0.36 | 3.58E-06 | 1.43E-02 | WT vs Hom |
| *Lmcd1* | 917.81 | 0.86 | 0.19 | 4.79E-06 | 2.45E-03 | Het vs Hom |
| *Ndufa5* | 1953.11 | 0.64 | 0.14 | 4.92E-06 | 9.93E-04 | WT vs Het |
| *Chchd2* | 3590.21 | 0.77 | 0.17 | 5.56E-06 | 9.93E-04 | WT vs Het |
| *Rps26* | 1189.45 | 0.54 | 0.12 | 5.56E-06 | 9.93E-04 | WT vs Het |
| *Cyb5r1* | 215.08 | 0.89 | 0.20 | 5.86E-06 | 9.93E-04 | WT vs Het |
| *Myh1* | 8083.13 | -0.57 | 0.13 | 6.28E-06 | 1.00E-03 | WT vs Het |
| *Ddit4l* | 776.74 | -0.80 | 0.18 | 6.31E-06 | 1.89E-02 | WT vs Hom |
| *Mustn1* | 621.30 | 0.94 | 0.21 | 1.11E-05 | 2.65E-02 | WT vs Hom |
| *Gm41498* | 10.15 | 4.67 | 1.08 | 1.43E-05 | 2.85E-02 | WT vs Hom |
| *Aste1* | 6.02 | 6.00 | 1.41 | 1.97E-05 | 3.37E-02 | WT vs Hom |
| *Ndufb3* | 1130.01 | 0.51 | 0.12 | 2.30E-05 | 3.47E-03 | WT vs Het |
| *Cep85l* | 130.16 | -0.89 | 0.21 | 2.46E-05 | 3.51E-03 | WT vs Het |
| *Nedd4* | 1120.49 | -0.67 | 0.16 | 2.70E-05 | 3.66E-03 | WT vs Het |
| *Gm52631* | 27.88 | 2.74 | 0.66 | 3.48E-05 | 5.19E-02 | WT vs Hom |
| *Rnf115* | 376.02 | 1.00 | 0.24 | 3.48E-05 | 1.42E-02 | Het vs Hom |
| *Rps23* | 354.73 | 0.72 | 0.18 | 4.41E-05 | 5.68E-03 | WT vs Het |
| *Calm3* | 385.74 | 0.68 | 0.17 | 4.61E-05 | 5.68E-03 | WT vs Het |
| *Gsk3b* | 640.19 | 0.66 | 0.16 | 5.32E-05 | 1.81E-02 | Het vs Hom |
| *Eef1d* | 296.03 | 0.63 | 0.16 | 5.41E-05 | 6.24E-03 | WT vs Het |
| *Zc3h15* | 403.14 | 0.57 | 0.14 | 5.56E-05 | 6.24E-03 | WT vs Het |
| *Ampd1* | 1433.88 | 0.49 | 0.12 | 5.76E-05 | 6.24E-03 | WT vs Het |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Aamdc* | 306.27 | 0.65 | 0.16 | 7.03E-05 | 7.33E-03 | WT vs Het |
| *Zfp827* | 101.30 | 0.97 | 0.25 | 7.41E-05 | 9.84E-02 | WT vs Hom |
| *Lrpprc* | 409.84 | -0.66 | 0.17 | 7.60E-05 | 2.03E-02 | Het vs Hom |
| *Nexn* | 5703.52 | 0.42 | 0.11 | 7.74E-05 | 7.77E-03 | WT vs Het |
| *Zbtb20* | 2996.76 | 0.47 | 0.12 | 7.95E-05 | 2.03E-02 | Het vs Hom |
| *Enah* | 742.71 | 0.76 | 0.19 | 8.74E-05 | 9.85E-02 | WT vs Hom |
| *Rnf115* | 360.46 | 0.83 | 0.21 | 9.07E-05 | 9.85E-02 | WT vs Hom |
| *Ndufb3* | 1163.63 | -0.54 | 0.14 | 9.89E-05 | 2.25E-02 | Het vs Hom |
| *Ddx6* | 313.16 | -0.63 | 0.16 | 1.11E-04 | 1.07E-02 | WT vs Het |
| *Cul5* | 627.59 | -0.50 | 0.13 | 1.67E-04 | 1.56E-02 | WT vs Het |
| *Rpl35* | 372.64 | 0.60 | 0.16 | 1.88E-04 | 1.70E-02 | WT vs Het |
| *Rhou* | 136.66 | -0.94 | 0.25 | 2.06E-04 | 1.80E-02 | WT vs Het |
| *Txnip* | 1673.18 | -0.65 | 0.18 | 2.50E-04 | 2.12E-02 | WT vs Het |
| *Csnk1e* | 260.04 | 0.65 | 0.18 | 2.54E-04 | 4.51E-02 | Het vs Hom |
| *Rpl26* | 624.62 | 0.57 | 0.16 | 2.64E-04 | 2.17E-02 | WT vs Het |
| *Fzd4* | 193.69 | -0.79 | 0.22 | 2.80E-04 | 2.18E-02 | WT vs Het |
| *Atp6v1g1* | 345.67 | 0.60 | 0.17 | 2.82E-04 | 2.18E-02 | WT vs Het |
| *Snrk* | 203.40 | 0.78 | 0.21 | 2.93E-04 | 4.51E-02 | Het vs Hom |
| *Sf3b1* | 491.05 | 0.48 | 0.13 | 3.00E-04 | 4.51E-02 | Het vs Hom |
| *Rpl23* | 999.56 | -0.45 | 0.12 | 3.07E-04 | 4.51E-02 | Het vs Hom |
| *Klf9* | 2713.11 | -0.35 | 0.10 | 3.23E-04 | 2.39E-02 | WT vs Het |
| *Pabpn1* | 233.87 | 0.83 | 0.23 | 3.26E-04 | 2.39E-02 | WT vs Het |
| *Luc7l3* | 399.20 | 0.62 | 0.17 | 3.35E-04 | 2.39E-02 | WT vs Het |
| *F830016B08Rik* | 391.93 | -0.84 | 0.23 | 3.37E-04 | 4.51E-02 | Het vs Hom |
| *Nfkbia* | 423.40 | 0.68 | 0.19 | 3.37E-04 | 4.51E-02 | Het vs Hom |
| *F830016B08Rik* | 399.12 | 0.62 | 0.17 | 3.62E-04 | 2.52E-02 | WT vs Het |
| *Tacc2* | 1773.45 | 0.50 | 0.14 | 3.63E-04 | 4.51E-02 | Het vs Hom |
| *Rps23* | 370.22 | -0.70 | 0.20 | 3.75E-04 | 4.51E-02 | Het vs Hom |
| *Csnk1e* | 243.19 | -0.59 | 0.17 | 3.79E-04 | 2.56E-02 | WT vs Het |
| *Tcea3* | 1382.07 | 0.47 | 0.13 | 3.87E-04 | 2.56E-02 | WT vs Het |
| *Lrpprc* | 399.88 | 0.60 | 0.17 | 4.05E-04 | 2.61E-02 | WT vs Het |
| *Ank1* | 1246.27 | 0.36 | 0.10 | 4.35E-04 | 2.74E-02 | WT vs Het |
| *Usf2* | 481.98 | 0.46 | 0.13 | 4.51E-04 | 2.78E-02 | WT vs Het |
| *Ndufa5* | 2136.78 | -0.46 | 0.13 | 4.60E-04 | 4.87E-02 | Het vs Hom |
| *Rps25* | 775.05 | -0.48 | 0.14 | 4.69E-04 | 4.87E-02 | Het vs Hom |
| *Tcea3* | 1418.97 | -0.52 | 0.15 | 4.77E-04 | 4.87E-02 | Het vs Hom |
| *Dvl1* | 150.89 | 0.73 | 0.21 | 4.89E-04 | 2.89E-02 | WT vs Het |
| *Gm26917* | 474952.57 | -0.41 | 0.12 | 4.96E-04 | 2.89E-02 | WT vs Het |
| *Rala* | 255.03 | 0.58 | 0.17 | 5.01E-04 | 2.89E-02 | WT vs Het |
| *Pfkfb1* | 135.91 | -0.93 | 0.27 | 5.34E-04 | 2.93E-02 | WT vs Het |
| *Fhl3* | 634.70 | -0.59 | 0.17 | 5.37E-04 | 2.93E-02 | WT vs Het |
| *Ndufa1* | 418.14 | 0.56 | 0.16 | 5.40E-04 | 2.93E-02 | WT vs Het |
| *Gmpr* | 475.03 | -0.56 | 0.16 | 5.44E-04 | 5.15E-02 | Het vs Hom |
| *Diaph1* | 589.78 | 0.57 | 0.16 | 5.55E-04 | 5.15E-02 | Het vs Hom |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Snrk* | 190.51 | -0.71 | 0.21 | 5.82E-04 | 3.08E-02 | WT vs Het |
| *Mrpl33* | 444.01 | 0.50 | 0.14 | 5.91E-04 | 3.08E-02 | WT vs Het |
| *Gm54301* | 386.84 | -0.67 | 0.20 | 6.12E-04 | 3.13E-02 | WT vs Het |
| *Csde1* | 4540.75 | -0.30 | 0.09 | 6.38E-04 | 3.20E-02 | WT vs Het |
| *Map2k6* | 258.21 | -0.60 | 0.17 | 6.39E-04 | 5.68E-02 | Het vs Hom |
| *Klf3* | 125.95 | -0.82 | 0.24 | 6.51E-04 | 3.21E-02 | WT vs Het |
| *Gsk3b* | 594.11 | -0.58 | 0.17 | 7.17E-04 | 3.43E-02 | WT vs Het |
| *Rpl17* | 389.24 | 0.59 | 0.17 | 7.22E-04 | 3.43E-02 | WT vs Het |
| *Anapc13* | 290.62 | -0.56 | 0.17 | 7.45E-04 | 6.33E-02 | Het vs Hom |
| *Rplp2* | 1873.88 | 0.42 | 0.12 | 7.51E-04 | 3.51E-02 | WT vs Het |
| *Tnnc2* | 69815.22 | 0.36 | 0.11 | 7.74E-04 | 3.56E-02 | WT vs Het |
| *Fbxo40* | 682.33 | 0.52 | 0.15 | 8.04E-04 | 6.33E-02 | Het vs Hom |
| *Rps27a* | 283.64 | -0.62 | 0.19 | 8.05E-04 | 6.33E-02 | Het vs Hom |
| *Asb2* | 1277.11 | 0.49 | 0.15 | 8.46E-04 | 3.77E-02 | WT vs Het |
| *Pdhx* | 297.26 | 0.51 | 0.15 | 8.49E-04 | 3.77E-02 | WT vs Het |
| *Myo18a* | 780.17 | 0.50 | 0.15 | 8.67E-04 | 6.46E-02 | Het vs Hom |
| *Atp5md* | 2520.91 | 0.38 | 0.11 | 8.94E-04 | 3.87E-02 | WT vs Het |
| *Cacna2d1* | 1155.95 | -0.38 | 0.11 | 9.06E-04 | 3.87E-02 | WT vs Het |
| *Rpl21* | 708.80 | 0.46 | 0.14 | 9.14E-04 | 3.87E-02 | WT vs Het |
| *Aopep* | 202.23 | 0.83 | 0.25 | 9.14E-04 | 6.46E-02 | Het vs Hom |
| *Fyco1* | 721.15 | -0.49 | 0.15 | 9.29E-04 | 3.88E-02 | WT vs Het |
| *Cd93* | 295.80 | 0.55 | 0.17 | 9.32E-04 | 6.46E-02 | Het vs Hom |
| *Spag9* | 524.29 | -0.43 | 0.13 | 9.46E-04 | 3.88E-02 | WT vs Het |
| *Rbm38* | 454.44 | 0.59 | 0.18 | 9.74E-04 | 6.46E-02 | Het vs Hom |
| *Des* | 6517.78 | -0.39 | 0.12 | 9.79E-04 | 6.46E-02 | Het vs Hom |
| *Rpl23* | 972.50 | 0.41 | 0.13 | 1.06E-03 | 4.27E-02 | WT vs Het |
| *Trdn* | 12369.87 | -0.46 | 0.14 | 1.10E-03 | 7.01E-02 | Het vs Hom |
| *Rpl35* | 393.93 | -0.54 | 0.17 | 1.14E-03 | 7.03E-02 | Het vs Hom |
| *Rpl3l* | 1660.44 | 0.48 | 0.15 | 1.14E-03 | 4.52E-02 | WT vs Het |
| *Srp68* | 236.06 | 0.53 | 0.16 | 1.15E-03 | 4.52E-02 | WT vs Het |
| *Rps24* | 1372.19 | -0.49 | 0.15 | 1.17E-03 | 7.03E-02 | Het vs Hom |
| *Myh8* | 165.96 | -0.69 | 0.21 | 1.21E-03 | 4.68E-02 | WT vs Het |
| *Fbxo40* | 624.99 | -0.39 | 0.12 | 1.29E-03 | 4.88E-02 | WT vs Het |
| *Rps7* | 553.08 | 0.49 | 0.15 | 1.30E-03 | 4.88E-02 | WT vs Het |
| *Ift20* | 153.66 | 0.59 | 0.18 | 1.34E-03 | 4.96E-02 | WT vs Het |
| *Pabpc4* | 2282.74 | 0.37 | 0.12 | 1.38E-03 | 4.96E-02 | WT vs Het |
| *Svil* | 684.74 | -0.40 | 0.12 | 1.40E-03 | 4.96E-02 | WT vs Het |
| *Pygm* | 11300.31 | -0.32 | 0.10 | 1.40E-03 | 4.96E-02 | WT vs Het |
| *Nfatc3* | 128.21 | -0.76 | 0.24 | 1.41E-03 | 4.96E-02 | WT vs Het |
| *Macf1* | 221.52 | 0.56 | 0.17 | 1.45E-03 | 8.49E-02 | Het vs Hom |
| *Mpz* | 183.56 | -0.74 | 0.23 | 1.53E-03 | 8.66E-02 | Het vs Hom |
| *Sh3glb1* | 1838.81 | -0.31 | 0.10 | 1.62E-03 | 5.60E-02 | WT vs Het |
| *Mrpl51* | 472.77 | 0.56 | 0.18 | 1.63E-03 | 5.60E-02 | WT vs Het |
| *Mif* | 501.42 | 0.44 | 0.14 | 1.64E-03 | 9.08E-02 | Het vs Hom |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Nudt9* | 162.32 | 0.60 | 0.19 | 1.65E-03 | 5.60E-02 | WT vs Het |
| *Rps12* | 1003.69 | 0.42 | 0.13 | 1.71E-03 | 5.71E-02 | WT vs Het |
| *Rpl31* | 1095.61 | 0.40 | 0.13 | 1.74E-03 | 5.72E-02 | WT vs Het |
| *Calm1* | 5455.36 | 0.32 | 0.10 | 1.75E-03 | 5.72E-02 | WT vs Het |
| *Atp5j* | 1568.65 | 0.39 | 0.12 | 1.78E-03 | 5.73E-02 | WT vs Het |
| *Rpl21* | 740.54 | -0.44 | 0.14 | 1.79E-03 | 9.65E-02 | Het vs Hom |
| *Pa2g4* | 204.72 | -0.60 | 0.19 | 1.81E-03 | 5.75E-02 | WT vs Het |
| *Aurkaip1* | 207.59 | 0.61 | 0.20 | 1.82E-03 | 5.75E-02 | WT vs Het |
| *Gm26917* | 502105.77 | 0.45 | 0.14 | 1.92E-03 | 9.78E-02 | Het vs Hom |
| *Dot1l* | 304.41 | 0.66 | 0.21 | 1.96E-03 | 9.78E-02 | Het vs Hom |
| *Acot2* | 197.16 | 0.89 | 0.29 | 1.98E-03 | 9.78E-02 | Het vs Hom |
| *Nrap* | 280.54 | -0.76 | 0.25 | 1.99E-03 | 6.21E-02 | WT vs Het |
| *Atp1b1* | 1548.24 | 0.59 | 0.19 | 2.01E-03 | 9.78E-02 | Het vs Hom |
| *Ppp2r5c* | 274.52 | -0.51 | 0.17 | 2.02E-03 | 6.21E-02 | WT vs Het |
| *Rpl41* | 603.37 | 0.40 | 0.13 | 2.04E-03 | 6.21E-02 | WT vs Het |
| *Calm3* | 415.45 | -0.55 | 0.18 | 2.10E-03 | 1.00E-01 | Het vs Hom |
| *Mrpl12* | 525.91 | 0.42 | 0.14 | 2.19E-03 | 6.58E-02 | WT vs Het |
| *Celf1* | 427.14 | -0.43 | 0.14 | 2.24E-03 | 6.62E-02 | WT vs Het |
| *Cox7b* | 3115.25 | 0.36 | 0.12 | 2.25E-03 | 6.62E-02 | WT vs Het |
| *Gps2* | 270.70 | 0.51 | 0.17 | 2.33E-03 | 6.80E-02 | WT vs Het |
| *Sub1* | 674.17 | -0.42 | 0.14 | 2.36E-03 | 6.81E-02 | WT vs Het |
| *Arhgef12* | 243.45 | -0.49 | 0.16 | 2.51E-03 | 7.04E-02 | WT vs Het |
| *Myo18a* | 733.33 | -0.44 | 0.15 | 2.51E-03 | 7.04E-02 | WT vs Het |
| *Ttn* | 74332.81 | -0.35 | 0.11 | 2.56E-03 | 7.04E-02 | WT vs Het |
| *Ecsit* | 183.84 | 0.59 | 0.20 | 2.57E-03 | 7.04E-02 | WT vs Het |
| *Sgk1* | 305.59 | -0.54 | 0.18 | 2.57E-03 | 7.04E-02 | WT vs Het |
| *Gas5* | 134.04 | 0.82 | 0.27 | 2.66E-03 | 7.22E-02 | WT vs Het |
| *Lsm12* | 148.47 | -0.62 | 0.21 | 2.76E-03 | 7.31E-02 | WT vs Het |
| *ND4* | 213575.61 | 0.45 | 0.15 | 2.77E-03 | 7.31E-02 | WT vs Het |
| *Jup* | 173.13 | -0.61 | 0.20 | 2.81E-03 | 7.31E-02 | WT vs Het |
| *Tuba4a* | 1101.21 | 0.39 | 0.13 | 2.81E-03 | 7.31E-02 | WT vs Het |
| *Msi2* | 879.98 | -0.44 | 0.15 | 2.83E-03 | 7.31E-02 | WT vs Het |
| *Ndufs5* | 421.85 | 0.47 | 0.16 | 2.95E-03 | 7.51E-02 | WT vs Het |
| *Sparcl1* | 892.04 | 0.44 | 0.15 | 2.97E-03 | 7.51E-02 | WT vs Het |
| *Cab39* | 487.27 | -0.44 | 0.15 | 3.18E-03 | 7.94E-02 | WT vs Het |
| *Ywhae* | 1599.59 | 0.34 | 0.12 | 3.19E-03 | 7.94E-02 | WT vs Het |
| *Psmc4* | 186.50 | 0.54 | 0.18 | 3.25E-03 | 8.00E-02 | WT vs Het |
| *Upf2* | 257.32 | 0.50 | 0.17 | 3.30E-03 | 8.06E-02 | WT vs Het |
| *Actb* | 1531.44 | -0.42 | 0.14 | 3.35E-03 | 8.10E-02 | WT vs Het |
| *Prkar2a* | 1263.46 | -0.36 | 0.12 | 3.42E-03 | 8.16E-02 | WT vs Het |
| *Usp2* | 994.06 | -0.42 | 0.14 | 3.43E-03 | 8.16E-02 | WT vs Het |
| *Mrps36* | 126.67 | 0.61 | 0.21 | 3.50E-03 | 8.25E-02 | WT vs Het |
| *Cep85* | 167.09 | -0.57 | 0.19 | 3.65E-03 | 8.53E-02 | WT vs Het |
| *Idh3b* | 1300.48 | 0.38 | 0.13 | 3.91E-03 | 8.99E-02 | WT vs Het |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Rps25* | 761.79 | 0.41 | 0.14 | 3.91E-03 | 8.99E-02 | WT vs Het |
| *Mrfap1* | 729.65 | 0.39 | 0.14 | 4.07E-03 | 9.26E-02 | WT vs Het |
| *Hsp90ab1* | 5115.62 | 0.36 | 0.13 | 4.13E-03 | 9.34E-02 | WT vs Het |
| *Nkiras1* | 200.49 | -0.48 | 0.17 | 4.25E-03 | 9.53E-02 | WT vs Het |
| *Smarca5* | 162.22 | -0.53 | 0.19 | 4.30E-03 | 9.55E-02 | WT vs Het |
| *Ddx17* | 147.21 | 0.68 | 0.24 | 4.36E-03 | 9.61E-02 | WT vs Het |
| *Cops9* | 507.84 | 0.39 | 0.14 | 4.55E-03 | 9.95E-02 | WT vs Het |
| **12-month-old *hSOD1$^{A4V}$* TA muscle** | | | | | | |
| *Hbb-bs* | 1730.53 | 1.26 | 0.12 | 2.76E-24 | 5.48E-20 | Het vs Hom |
| *Hbb-bt* | 359.04 | 1.20 | 0.12 | 1.08E-23 | 1.08E-19 | Het vs Hom |
| *Hba-a2* | 204.76 | 1.15 | 0.18 | 9.74E-11 | 6.46E-07 | Het vs Hom |
| *Eda2r* | 33.77 | 1.37 | 0.29 | 2.60E-06 | 1.29E-02 | Het vs Hom |
| *Klf2* | 82.71 | 1.06 | 0.25 | 1.65E-05 | 6.55E-02 | Het vs Hom |
| **21-month-old *hSOD1$^{A4V}$* TA muscle** | | | | | | |
| *Myl3* | 413.64 | -0.78 | 0.15 | 2.96E-07 | 5.59E-03 | Het vs Hom |

**Table a-2 | Full list of differentially expressed genes (DEGs) in the TA muscle of male *hSOD1$^{A4V}$* mice.** SE, standard error; WT, *hSOD1$^{WT}$*/*hSOD1$^{WT}$*; Het, *hSOD1$^{WT}$*/*hSOD1$^{A4V}$*; Hom, *hSOD1$^{A4V}$*/*hSOD1$^{A4V}$*.

## 8.2 Full list of changed metabolites in plasma from *hSOD1$^{A4V}$* mice

| Metabolite name | Log2 Fold Change Het vs WT | Log2 Fold Change Hom vs WT | F Value | P Value | Adjusted P Value |
|---|---|---|---|---|---|
| **4-month-old animals** | | | | | |
| Biliverdin | -0.382 | 1.131 | 6.642 | 2.00E-03 | 0.423 |
| Niacinamide | -0.797 | 0.540 | 4.827 | 1.01E-02 | 0.742 |
| 2,3,4,5,6,7-Hexahydroxyheptanoic acid | -1.253 | -0.425 | 4.371 | 1.53E-02 | 0.742 |
| Hexose | -1.135 | -0.242 | 3.960 | 2.23E-02 | 0.742 |
| Histamine | -1.200 | -0.626 | 3.539 | 3.29E-02 | 0.742 |
| L-2-Hydroxyglutaric acid / 3-Hydroxyglutaric acid | -0.362 | -1.197 | 3.481 | 3.48E-02 | 0.742 |
| LysoPC(18:0/0:0) | -1.134 | -0.813 | 3.106 | 4.94E-02 | 0.742 |
| **6-month-old animals** | | | | | |
| LysoPC(0:0/22:4) | 1.460 | 1.122 | 5.252 | 5.24E-03 | 0.466 |
| Inosine | -0.528 | 0.900 | 5.184 | 5.61E-03 | 0.466 |
| Guanosine | -0.914 | 0.433 | 4.977 | 6.89E-03 | 0.466 |
| Alpha-ketoisovaleric acid | 1.317 | 0.557 | 4.338 | 1.31E-02 | 0.466 |
| LysoPC(0:0/20:4) | 1.221 | 1.139 | 4.024 | 1.79E-02 | 0.466 |
| Ketoleucine | 1.270 | 0.550 | 4.017 | 1.80E-02 | 0.466 |
| L-Phenylalanine | 1.290 | 0.842 | 3.989 | 1.85E-02 | 0.466 |
| Hypoxanthine | -0.228 | 0.991 | 3.976 | 1.88E-02 | 0.466 |
| L-leucine / L-Isoleucine | 1.171 | 0.325 | 3.749 | 2.35E-02 | 0.466 |
| Xanthosine | -1.226 | -0.653 | 3.627 | 2.66E-02 | 0.466 |

| | | | | | |
|---|---|---|---|---|---|
| Xanthine | -0.420 | 0.762 | 3.553 | 2.86E-02 | 0.466 |
| LysoPC(22:4(7Z,10Z,13Z,16Z)/0:0) | 1.211 | 0.684 | 3.518 | 2.97E-02 | 0.466 |
| LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0) | 1.193 | 0.903 | 3.495 | 3.04E-02 | 0.466 |
| Icosa-8,11,14-trienoylcarnitine | 1.062 | 1.133 | 3.395 | 3.35E-02 | 0.466 |
| LysoPC(0:0/22:6) | 1.152 | 0.828 | 3.217 | 4.01E-02 | 0.466 |
| LysoPE(0:0/16:0) | 1.144 | 0.675 | 3.135 | 4.35E-02 | 0.466 |
| 3-Hydroxycapric acid | 1.140 | 0.780 | 3.127 | 4.38E-02 | 0.466 |
| Allantoin | -1.127 | -0.609 | 3.058 | 4.70E-02 | 0.466 |
| **9-month-old animals** | | | | | |
| LysoPC(22:4(7Z,10Z,13Z,16Z)/0:0) | 0.761 | 1.445 | 6.209 | 4.23E-03 | 0.544 |
| Azelaic acid | -0.936 | -1.195 | 5.961 | 5.13E-03 | 0.544 |
| 2-Methylglutaric acid / Methylglutaric acid | -0.875 | -1.325 | 5.359 | 8.28E-03 | 0.575 |
| LysoPC(20:4(5Z,8Z,11Z,14Z)/0:0) | 0.470 | 1.318 | 5.025 | 1.09E-02 | 0.575 |
| Pimelic acid / 3-Methyladipic acid | -0.528 | -1.212 | 4.694 | 1.42E-02 | 0.603 |
| Suberic acid | -0.793 | -1.183 | 4.049 | 2.44E-02 | 0.683 |
| LysoPC(0:0/20:4) | 0.408 | 1.192 | 3.921 | 2.71E-02 | 0.683 |
| N-Formyl-L-methionine | -0.712 | -1.071 | 3.855 | 2.87E-02 | 0.683 |
| LysoPC(18:2(9Z,12Z)/0:0) | -0.123 | -0.923 | 3.843 | 2.90E-02 | 0.683 |
| Glutarylcarnitine | -0.969 | -0.316 | 3.546 | 3.74E-02 | 0.747 |
| 4-Acetamidobutanoic acid | -0.148 | -0.962 | 3.315 | 4.56E-02 | 0.747 |
| L-Palmitoylcarnitine | 0.531 | -0.454 | 3.278 | 4.72E-02 | 0.747 |
| L-2-Hydroxyglutaric acid / 3-Hydroxyglutaric acid | -0.295 | -1.063 | 3.265 | 4.77E-02 | 0.747 |
| Isocitric acid | -0.905 | -0.962 | 3.227 | 4.93E-02 | 0.747 |
| **13-month-old animals** | | | | | |
| LysoPC(18:3(9Z,12Z,15Z)/0:0) | -1.133 | -1.628 | 5.397 | 4.53E-03 | 0.269 |
| Palmitoleic acid | 0.454 | -0.955 | 5.145 | 5.83E-03 | 0.269 |
| Methylguanosine | 0.124 | -1.211 | 5.119 | 5.98E-03 | 0.269 |
| LysoPC(0:0/18:3) | -0.726 | -1.546 | 4.802 | 8.21E-03 | 0.269 |
| 4-Acetamidobutanoic acid | -0.201 | -1.353 | 4.705 | 9.05E-03 | 0.269 |
| LysoPC(0:0/18:2(9Z,12Z)) | -0.589 | -1.503 | 4.700 | 9.10E-03 | 0.269 |
| Creatinine | 0.146 | -1.130 | 4.615 | 9.90E-03 | 0.269 |
| Palmitic acid | -0.779 | -1.519 | 4.589 | 1.02E-02 | 0.269 |
| 4-Hydroxybenzoic acid | 1.120 | -0.021 | 4.335 | 1.31E-02 | 0.270 |
| Pyroglutamic acid | 1.280 | 1.266 | 4.242 | 1.44E-02 | 0.270 |
| LysoPC(18:2(9Z,12Z)/0:0) | -0.438 | -1.388 | 4.222 | 1.47E-02 | 0.270 |
| Oleic acid / Elaidic acid | 0.097 | -1.094 | 4.102 | 1.65E-02 | 0.270 |
| 2-Ketobutyric acid | 0.014 | -1.149 | 4.100 | 1.66E-02 | 0.270 |
| LysoPC(14:0/0:0) | -0.375 | -1.313 | 3.871 | 2.08E-02 | 0.315 |
| Hippuric acid | -0.043 | -1.115 | 3.629 | 2.66E-02 | 0.347 |
| alpha-Linolenic acid | 0.225 | -0.930 | 3.620 | 2.68E-02 | 0.347 |
| LysoPC(0:0/14:0) | -0.334 | -1.248 | 3.553 | 2.86E-02 | 0.347 |
| N-Formyl-L-methionine | -0.342 | -1.240 | 3.481 | 3.08E-02 | 0.347 |
| 9-Hexadecenoylcarnitine | -0.141 | -1.143 | 3.451 | 3.17E-02 | 0.347 |

| | | | | | |
|---|---|---|---|---|---|
| Benzoic acid | -1.193 | -1.065 | 3.419 | 3.27E-02 | 0.347 |
| Kynurenic acid | 0.082 | -0.956 | 3.121 | 4.41E-02 | 0.417 |
| 9,12-Hexadecadienoylcarnitine | -0.047 | -1.036 | 3.111 | 4.45E-02 | 0.417 |
| 3-Hydroxybutyric acid | 0.267 | -0.796 | 2.996 | 5.00E-02 | 0.417 |
| **20-month-old animals** | | | | | |
| 2-Ketobutyric acid | -0.878 | -1.089 | 5.421 | 6.31E-03 | 0.693 |
| (S)-3-Hydroxyisobutyric acid | -0.637 | -1.009 | 4.054 | 2.13E-02 | 0.693 |
| LysoPC(20:3(8Z,11Z,14Z)/0:0) | -0.084 | 0.823 | 3.917 | 2.41E-02 | 0.693 |
| Stearoylcarnitine | -0.090 | -0.886 | 3.664 | 3.03E-02 | 0.693 |
| LysoPE(0:0/18:2(9Z,12Z)) | 0.577 | -0.326 | 3.450 | 3.68E-02 | 0.693 |
| alpha-Dimorphecolic acid | 0.666 | -0.214 | 3.374 | 3.95E-02 | 0.693 |
| Hexose | -0.880 | -0.304 | 3.307 | 4.20E-02 | 0.693 |
| LysoPC(22:5(4Z,7Z,10Z,13Z,16Z)/0:0) | -0.875 | -0.331 | 3.303 | 4.21E-02 | 0.693 |
| 2,3,4,5,6,7-Hexahydroxyheptanoic acid | -0.883 | -0.359 | 3.269 | 4.35E-02 | 0.693 |
| 3-Hydroxytetradecanoic acid | 0.597 | -0.265 | 3.222 | 4.54E-02 | 0.693 |
| Itaconic acid | 0.789 | 0.055 | 3.168 | 4.77E-02 | 0.693 |
| 3-Hydroxyisovalerylcarnitine | 0.825 | 0.157 | 3.162 | 4.80E-02 | 0.693 |
| Palmitoleic acid | -0.609 | -0.890 | 3.146 | 4.87E-02 | 0.693 |

**Table a-3 | Full list of significantly changed metabolites from the LC-MS analysis of plasma from *hSOD1$^{A4V}$* mice.**

| Metabolite name | Log2 Fold Change Het vs WT | Log2 Fold Change Hom vs WT | F Value | P Value | Adjusted P Value |
|---|---|---|---|---|---|
| **4-month-old animals** | | | | | |
| 3-Hydroxymethylglutaric acid | -0.772 | -1.345 | 5.446 | 6.98E-03 | 0.341 |
| 1,5-Anhydrosorbitol (1,5-AG) | -0.507 | -1.367 | 5.399 | 7.27E-03 | 0.341 |
| 3,4-Dihydroxybutyric acid | -0.448 | -1.065 | 4.140 | 2.12E-02 | 0.452 |
| beta-Hydroxybutyric acid (3-HBA) | -0.814 | -1.090 | 3.455 | 3.87E-02 | 0.452 |
| MG(18:2(9Z,12Z)/0:0/0:0) | -0.542 | -1.152 | 3.396 | 4.07E-02 | 0.452 |
| **6-month-old animals** | | | | | |
| Alanine | 1.086 | 0.915 | 3.441 | 3.20E-02 | 0.552 |
| MG(18:1(9Z)/0:0/0:0) | 0.960 | 0.209 | 3.004 | 4.96E-02 | 0.552 |
| Elaidic acid | 0.991 | 0.300 | 3.001 | 4.97E-02 | 0.552 |
| **9-month-old animals** | | | | | |
| Glutamine | -0.407 | 0.943 | 6.588 | 3.29E-03 | 0.281 |
| Ornithine | -0.720 | 0.468 | 4.886 | 1.25E-02 | 0.281 |
| 1,11-Undecanedicarboxylic acid | -0.350 | 0.837 | 4.734 | 1.41E-02 | 0.281 |
| Pyroglutamic acid | -0.426 | 0.709 | 4.566 | 1.62E-02 | 0.281 |
| scyllo-Inositol | -1.006 | -1.052 | 4.490 | 1.72E-02 | 0.281 |
| Adipic acid | -1.081 | -0.310 | 4.441 | 1.79E-02 | 0.281 |
| Aminomalonic acid | -1.175 | -0.779 | 4.224 | 2.14E-02 | 0.288 |

| | | | | | |
|---|---|---|---|---|---|
| Arachidonic acid | 0.193 | 1.042 | 3.865 | 2.89E-02 | 0.340 |
| Undecanedioic acid | -0.948 | -0.180 | 3.249 | 4.90E-02 | 0.408 |
| **13-month-old animals** | | | | | |
| Glycolic acid | 0.501 | -0.977 | 5.992 | 2.84E-03 | 0.267 |
| Glutaric acid | -1.296 | -1.492 | 4.413 | 1.30E-02 | 0.516 |
| Succinic acid | -1.225 | -1.358 | 3.845 | 2.26E-02 | 0.516 |
| scyllo-Inositol | -0.796 | -1.436 | 3.661 | 2.70E-02 | 0.516 |
| myo-Inositol | -0.722 | -1.428 | 3.624 | 2.80E-02 | 0.516 |
| Isocitric acid | -0.100 | -1.103 | 3.263 | 3.98E-02 | 0.516 |
| 3-Hydroxymethylglutaric acid | 0.298 | -0.790 | 3.229 | 4.11E-02 | 0.516 |
| **20-month-old animals** | | | | | |
| Xylulose | 0.139 | 1.734 | 23.206 | 4.43E-08 | 0.000 |
| Proline | 1.085 | 0.323 | 5.948 | 4.53E-03 | 0.141 |
| Glutamine | -0.238 | 0.863 | 5.645 | 5.83E-03 | 0.141 |
| Urea | 0.786 | -0.300 | 5.444 | 6.89E-03 | 0.141 |
| Cysteine | -0.550 | 0.470 | 4.998 | 1.00E-02 | 0.165 |
| Cholesterol | -0.532 | -1.049 | 4.369 | 1.72E-02 | 0.188 |
| Glycine | 0.692 | 0.928 | 4.341 | 1.76E-02 | 0.188 |
| Methionine | -0.198 | 0.764 | 4.298 | 1.83E-02 | 0.188 |
| Glyceric acid | -0.535 | 0.421 | 3.811 | 2.80E-02 | 0.250 |
| Glucose | -0.656 | 0.243 | 3.716 | 3.05E-02 | 0.250 |
| Phenylalanine | -0.322 | 0.593 | 3.547 | 3.54E-02 | 0.264 |
| Pyruvic acid | 0.418 | 0.926 | 3.369 | 4.15E-02 | 0.272 |
| MG(18:0/0:0/0:0) | -0.828 | -0.156 | 3.251 | 4.61E-02 | 0.272 |
| Sarcosine | 0.842 | 0.599 | 3.244 | 4.64E-02 | 0.272 |

**Table a-4 | Full list of significantly changed metabolites from the GC-MS analysis of plasma from *hSOD1$^{A4V}$* mice.**

## 8.3 Representative immunofluorescence images from *hSOD1$^{A4V}$* embryonic spinal cord primary cultures

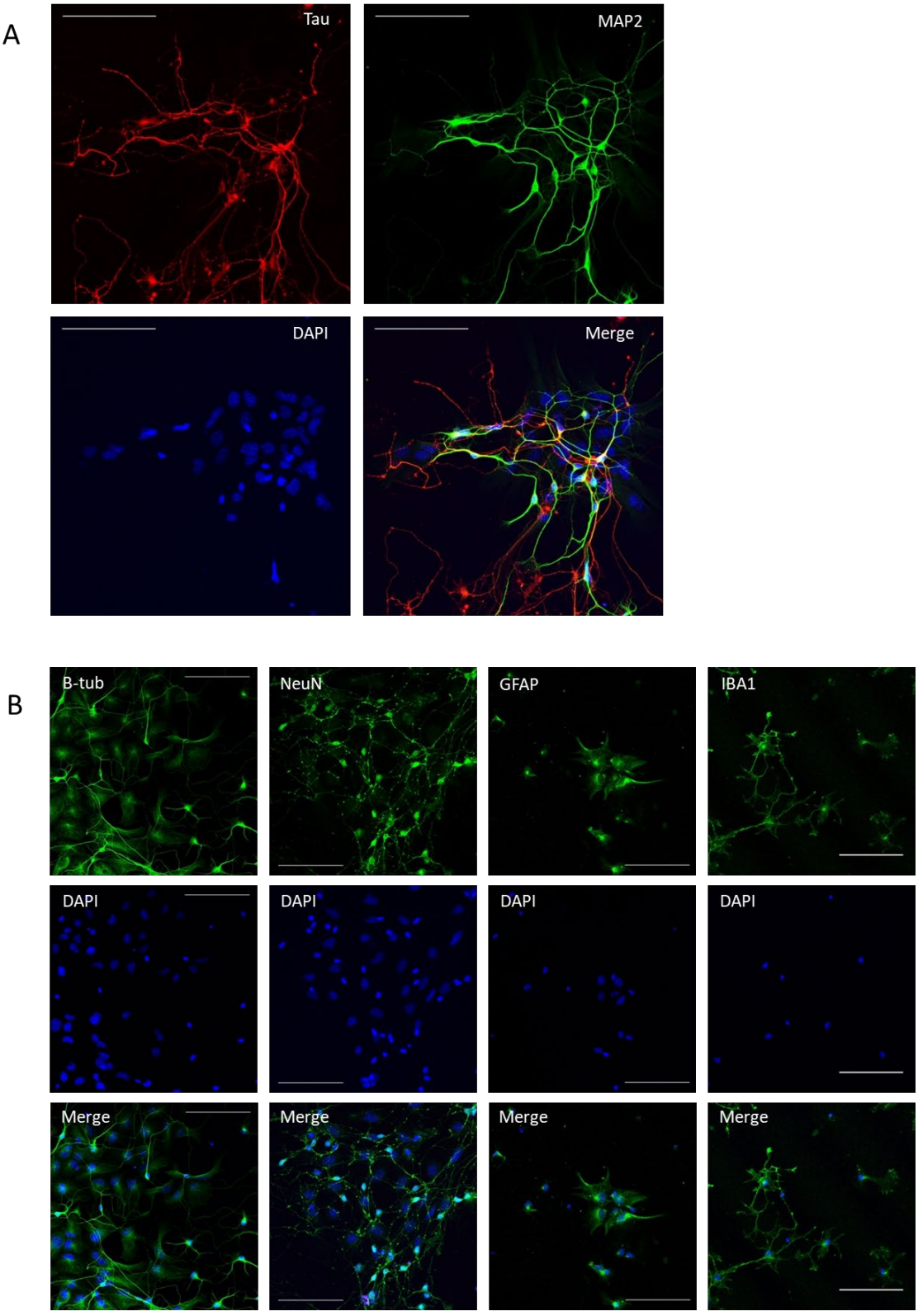


**Figure a-1 | Immunofluorescence images of neuronal/glial markers in spinal cord primary cultures from E13.5 *hSOD1$^{A4V}$* mice. (A)** Staining for Tau and MAP2 shows neuronal morphology. **(B)** Staining for β-tubulin (B-tub), NeuN, GFAP and IBA1 shows the presence of neurons, astrocytes and microglia. Images taken at 5 DIV. Scale bar = 100 µm.

## 8.4 Representative images of neuromuscular junction (NMJ) staining in aged *hSOD1$^{A4V}$* mice

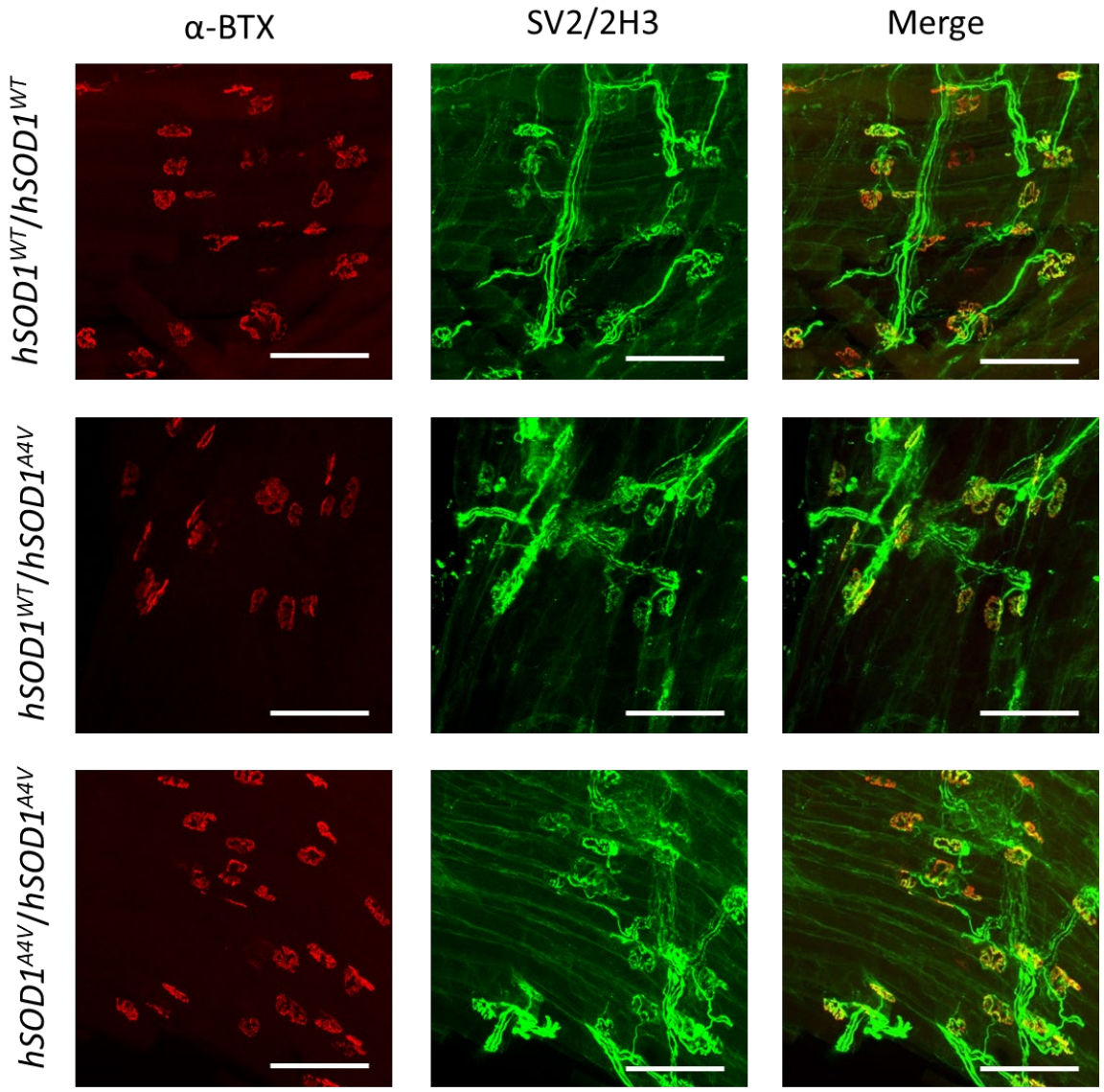



**Figure a-2 | Representative images of neuromuscular junction (NMJ) staining in lumbrical muscles from 24-month-old *hSOD1$^{A4V}$* mice.** α-bungarotoxin (α-BTX), red; synaptophysin/neurofilament (SV2/2H3), green. Scale bar 100 µm.

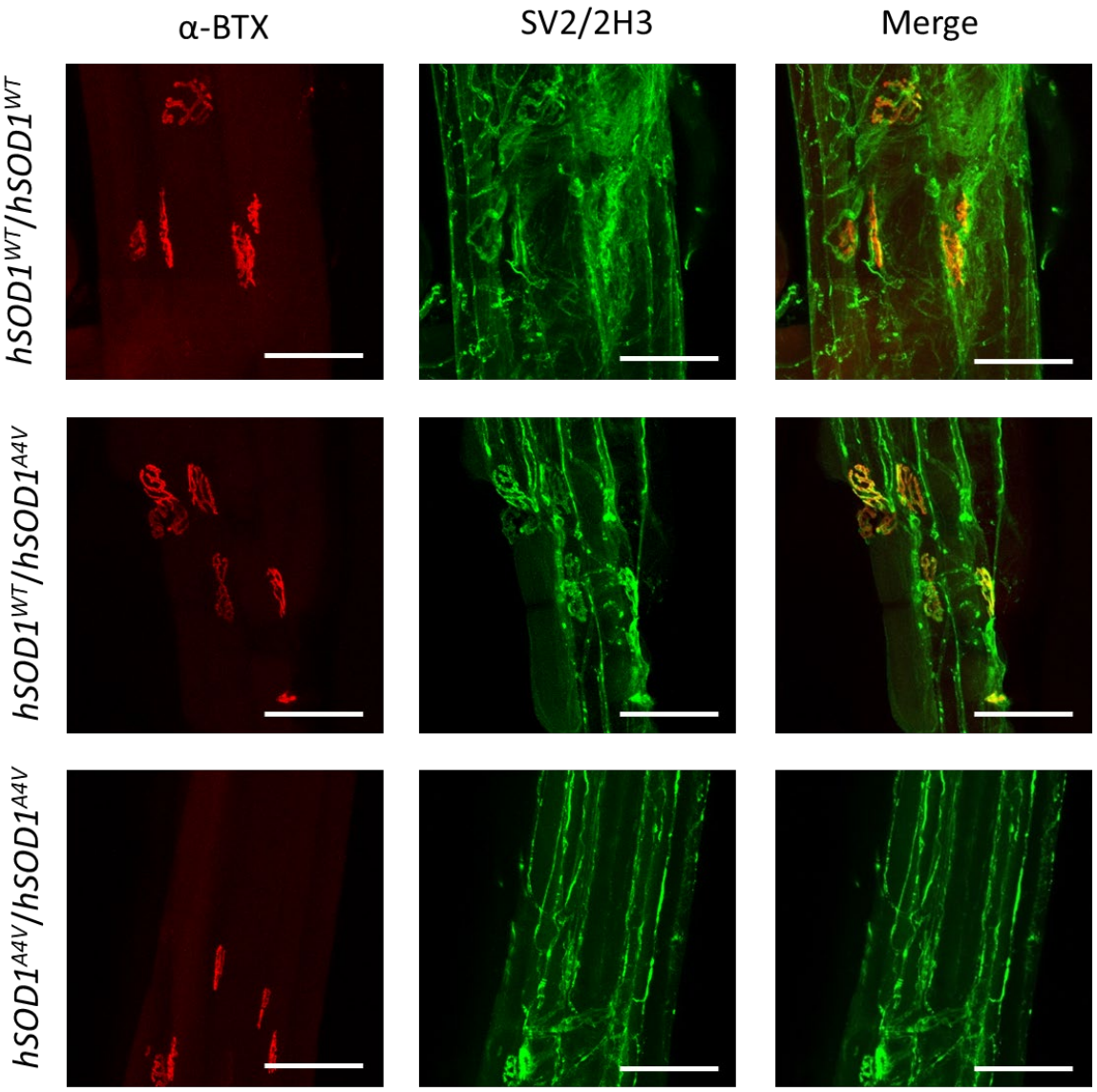



**Figure a-3 | Representative images of neuromuscular junction (NMJ) staining in tibialis anterior (TA) muscles from 24-month-old *hSOD1$^{A4V}$* mice.** α-bungarotoxin (α-BTX), red; synaptophysin/neurofilament (SV2/2H3), green. Scale bar 100 µm.

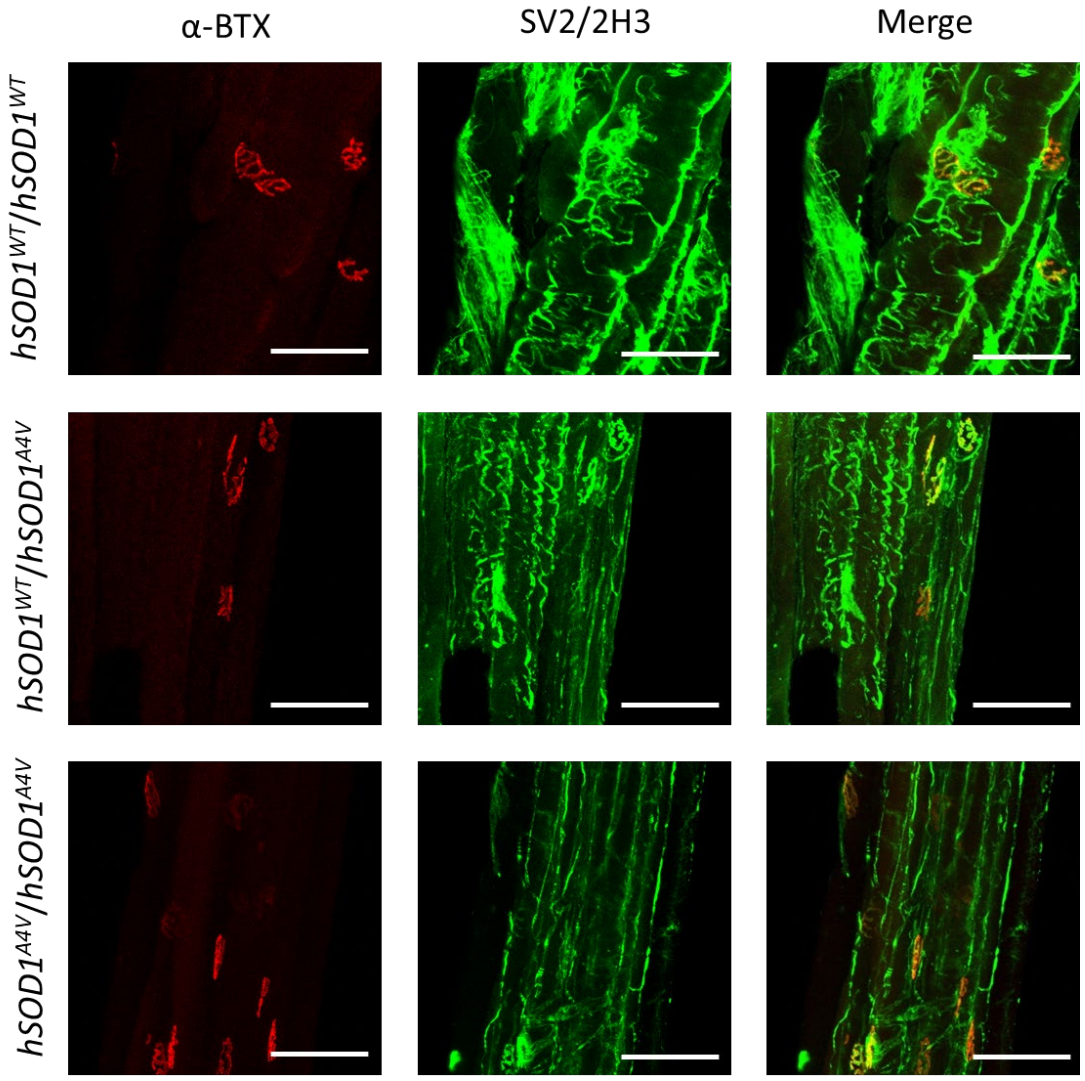



**Figure a-4 | Representative images of neuromuscular junction (NMJ) staining in soleus muscles from 24-month-old *hSOD1$^{A4V}$* mice.** α-bungarotoxin (α-BTX), red; synaptophysin/neurofilament (SV2/2H3), green. Scale bar 100 µm.

## 8.5 Full sequence of a C9orf72 repeat expansion within the RP11-27J8 BAC construct

Sequence of the *C9orf72* GGGGCC repeat (red) within BAC clone RP11-27J8 (100bp of sequence flanking the repeat shown in black) derived from PacBio long-read sequencing. Interruptions to the expected GGGGCC sequence are shown in green. Two interruptions to the GGGGCC sequence (highlighted in yellow) were checked by Sanger sequencing and found to follow the expected GGGGCC sequence, indicating these interruptions are due to errors in the PacBio sequencing read

rather than genuine interruptions to the repeat sequence. An insertion/deletion event (GTGGTC -> CGGGCCCG) present in BAC clone RP11-27J8 is shown in purple.

GGTGAACAAGAAAAGACCTGATAAAGATTAACCAGAAGAAAACAAGGAGGGAAACAACCGCAGCCTGTAGCAAGCTCTGGAAC
TCAGGAGTCGCGCGCTAGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGCCGGGGCCGGGGCCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
CCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCCGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGG
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGGCCGGGGCC

## 8.6 Full Western blot images

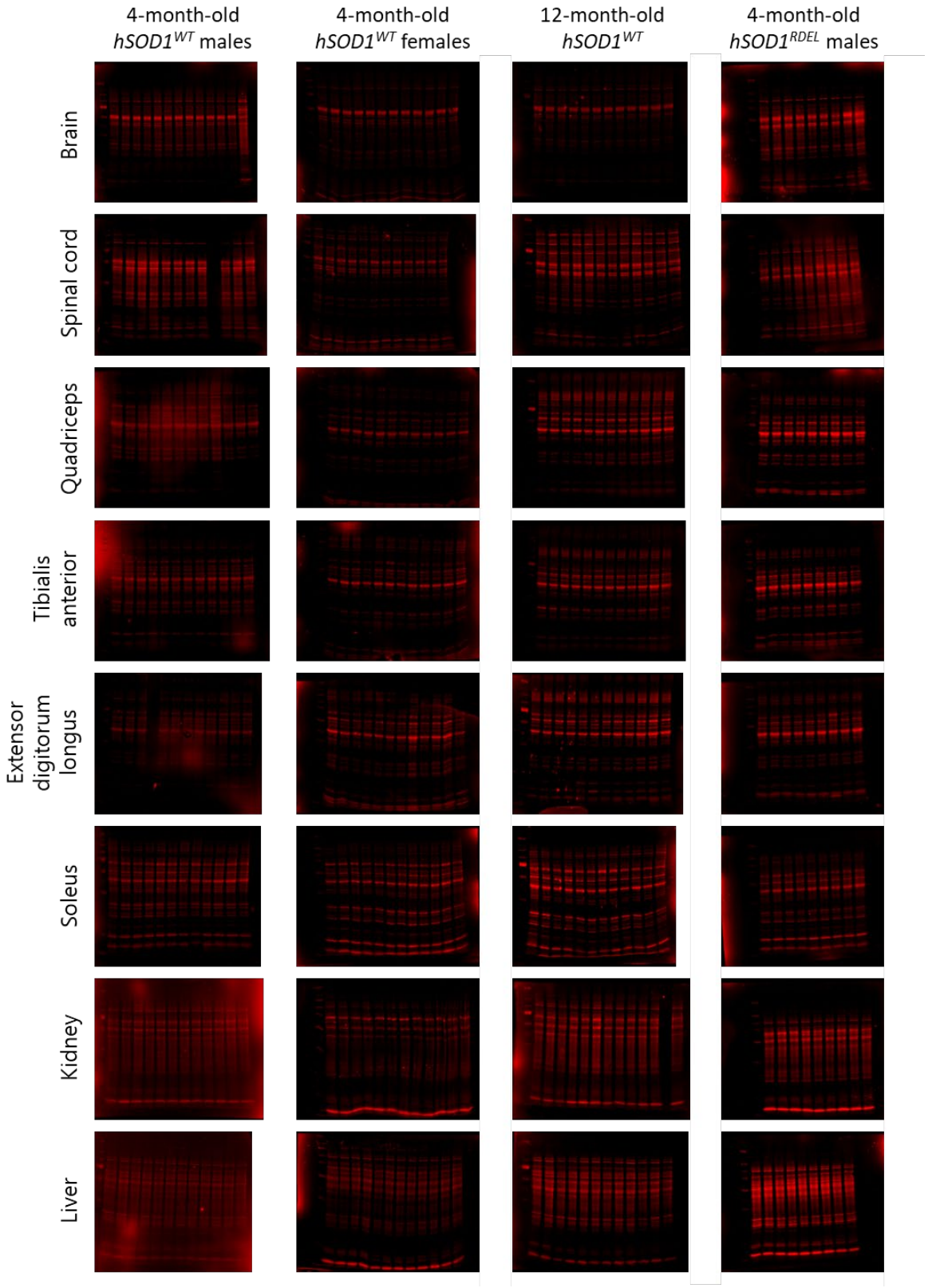



**Figure a-5 | Total protein staining images of Western blots of tissues from *hSOD1$^{WT}$* and *hSOD1$^{RDEL}$* mice.** Blots from figure 3-6 and 3-7.

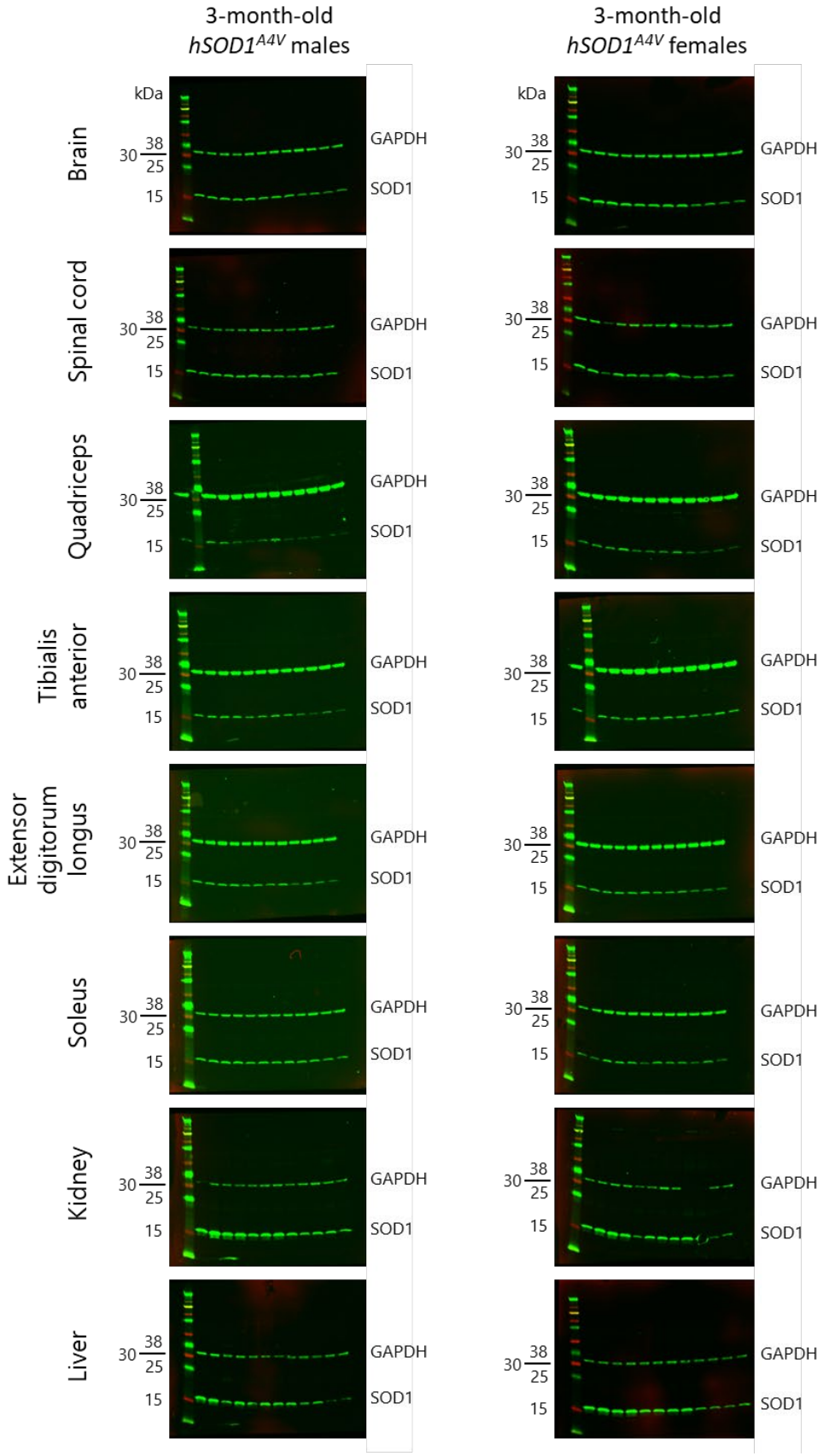



**Figure a-6 | Full images of Western blots of tissues from *hSOD1$^{A4V}$* mice.** Blots from figure 4-4.