Genetics of amyotrophic lateral sclerosis and frontotemporal dementia and the potential for discovering new genes and pathways underlying these neurological disorders

By Yevgeniya Abramzon

January 2022

A thesis submitted to University College London for the degree of Doctor of Philosophy

Reta Lila Weston Institute, UCL Queen Square Institute of Neurology,

University College London, London, UK

DECLARATION

I, Yevgeniya Abramzon, confirm that the work contained within this thesis is my own original work. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Despite tremendous progress in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) genetic research over the last two decades, only 20% to 40% of the underlying genetic causes have been discovered. Several genes responsible for these two neurological disorders remain to be discovered. During my Ph.D., I utilized the current state-of-the-art exome and whole-genome sequencing technologies to discover new loci underlying ALS and FTD. This genetic knowledge helps to unravel the complex etiology of these disorders and paves the way for targeted therapies.

Chapter 1 of this thesis talks about the evolution of genetic research in the past two decades. Chapter 2 introduces ALS and FTD, the two neurological disorders that are the focus of my thesis. Chapter 3 provides an overview of the genetics of ALS and FTD. Chapter 4 describes the discovery of the new gene linked to juvenile ALS, serine palmitoyltransferase long chain subunit 1 (SPTLC1), and proposes a treatment for the patients carrying mutations in this gene. Chapter 5 demonstrates that exome sequencing is a powerful technique for the analysis of rare hereditary conditions. I describe the discovery of gelsolin p.D187Y amino acid change, which was previously linked to hereditary amyloidosis type 4, in a large family in which multiple members were affected by bulbar neuropathy mimicking bulbar ALS. Chapter 6 discusses the role of the highly polymorphic locus replication factor C subunit 1 (RFC1) in ALS. Chapter 7 describes our whole-genome sequencing efforts in dementia and examines the data for the frequency of intermediate-size repeat expansions in genes previously linked to neurodegeneration.

IMPACT STATEMENT

Since March 2019, the National Institutes of Health, where I performed my graduate research, had significant occupancy restrictions due to the COVID-19 pandemic. Inability to attend the lab had a negative effect on my projects. For example, in replication factor C subunit 1 project described in chapter 6, I was unable to determine the exact size of the selected repeat expansions because I did not have access to the DNA samples that I was planning to send to another laboratory for Southern blotting. Instead, I had to rely on the data that was collected for this project before the COVID-19 closure. Unfortunately, this data did not provide enough information for accurate repeat expansion sizing. Inability to size the large repeat expansions was a significant limitation of my project.

COVID-19 restrictions also affected my intermediate alleles project described in chapter 7. In that project, using ExpansionHunter targeted I identified several sporadic ALS patients that were the carriers of ataxin-2 intermediate alleles. While ExpansionHunter is considered a precise tool for determining repeat expansion size, I was planning to confirm the results by performing repeat primed PCR on the samples carrying intermediate alleles. However, due to the laboratory restrictions, I was unable to complete this part of the project.

Finally, COVID-19 restrictions affected the serine palmitoyl transferase subunit 1 project described in chapter 4, because juvenile-onset ALS individual, who was undergoing L-serine supplementation treatment for twelve months could no longer be

observed by a neurologist to determine if there were any changes in the disease progression resulting from the treatment.

BIOGRAPHY

My path as a genetics researcher began in 2007, when I got my first internship at J. Craig Venter Genomic Institute. As a part of The Global Ocean Sampling Project, my internship focused on examining the genetic diversity of marine microorganisms to gain a better understanding of global oceanic biology.

In 2009, I joined the Laboratory of Neurogenetics at the National Institutes of Health as an Intramural Research Training Award recipient (IRTA), where I focused on unraveling the genetic etiology of amyotrophic lateral sclerosis and other related neuromuscular disorders.

In 2015, I received my bachelor's degree in Biological Sciences from the University of Maryland, USA. In 2016, I enrolled into a MPhil/PhD research program at The University College London, UK. My graduate research focused on the genetics of amyotrophic lateral sclerosis and frontotemporal dementia and the discovery of the new pathologic mutations and mechanisms that underlie these two neurological disorders.

During my graduate degree, I became the first author of three publications and coauthored seven more manuscripts published in reputable peer-reviewed scientific journals. I also had a chance to present my research at local symposiums such as The Annual Neurological Retreat at The National Institutes of Health, as well as at the other world - recognized research institutions including Georgetown University, USA and University College London, UK.

ACKNOWLEDGEMENTS

Through the five years of graduate school, I was surrounded by many wonderful people. This thesis would not have been possible without their continuous support and encouragement.

I am fully indebted to Dr. Bryan Traynor, my direct supervisor at the National Institutes of Health, for his extreme patience, guidance, enthusiasm and for pushing me farther than I would go on my own. I would also like to thank Dr. John Hardy, my chief supervisor at University College London and Pietro Fratta, my mentor at University College London for their valuable guidance.

Thanks to my colleagues and friends at the laboratory of neurogenetics at NIH: Monia Hammer, Sampath Arepalli, Sean Chong, Alan Renton, Celeste Sassi, Ramita Karra, Josh Green, Natalie Portley, and Maria Sabir for assisting with the experiments and helping me to learn new laboratory techniques. I want to especially thank Ruth Chia, a good friend, lab mate and peer mentor for her cheerful spirit and continuous help.

Thanks to my friends Maria, Inna, and Mila, who provided me with happy distractions from experiments and thesis writing. I am so lucky to have a supportive wonderful group of friends.

Thanks to my family, my mom for her unconditional love and constant encouragement, to my husband Vladimir for his support along each step of this journey. Thanks to my son Anthony for coming to this world.

Finally, thanks to the University College London and the National Institutes of Health for providing me with this amazing opportunity to learn from the leading scientists in the field of amyotrophic lateral sclerosis and frontotemporal dementia research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PUBLICATIONS SINCE PhD START

1. To Dement or Not to Dement, That Is the Question
   JAMA Neurol. (2016) doi: 10.1001/jamaneurol.2015.4984, (PMID: 26903123).
   Bryan J Traynor, **Yevgeniya A Abramzon.**

2. Exome sequencing establishes a gelsolin mutation as the cause of inherited bulbar-onset neuropathy.
   Muscle Nerve. (2017) doi: 10.1002/mus.25550, (PMID: 28039894).
   James B Caress, Janel O Johnson, **Yevgeniya A Abramzon**, Gregory A Hawkins, Raphael Gibbs, Elizabeth A Sullivan, Chamanpreet S Chahal, Bryan J Traynor.

3. NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases.
   Neurobiol Aging. (2017) doi: 10.1016/j.neurobiolaging.2017.05.009, (PMID: 28602509).
   Blauwendraat C, Faghri F, Pihlstrom L, Geiger JT, Elbaz A, Lesage S, Corvol JC, May P, Nicolas A, **Abramzon Y**, Murphy NA, Gibbs JR, Ryten M, Ferrari R, Bras J, Guerreiro R, Williams J, Sims R, Lubbe S, Hernandez DG, Mok KY, Robak L, Campbell RH, Rogaeva E, Traynor BJ, Chia R, Chung SJ; International Parkinson's Disease Genomics Consortium (IPDGC), COURAGE-PD Consortium, Hardy JA, Brice A, Wood NW, Houlden H, Shulman JM, Morris HR, Gasser T, Krüger R, Heutink P, Sharma M, Simón-Sánchez J, Nalls MA, Singleton AB, Scholz SW.

4. Genome-wide Analyses Identify KIF5A as a Novel ALS Gene
   Neuron. (2018) doi: 10.1016/j.neuron.2018.02.027, (PMID: 29566793)
   Nicolas A, Kenna KP, Renton AE, Ticozzi N, Faghri F, Chia R, Dominov JA, Kenna BJ, Nalls MA, Keagle P, Rivera AM, van Rheenen W, Murphy NA, van Vugt JJFA, Geiger JT, Van der Spek RA, Pliner HA, Shankaracharya, Smith BN, Marangi G, Topp SD, **Abramzon Y**, Gkazi AS, Eicher JD, Kenna A; ITALSGEN Consortium, Mora G, Calvo A, Mazzini L, Riva N, Mandrioli J, Caponnetto C, Battistini S, Volanti P, La Bella V, Conforti FL, Borghero G, Messina S, Simone IL, Trojsi F, Salvi F, Logullo FO, D'Alfonso S, Corrado L, Capasso M, Ferrucci L; Genomic Translation for ALS Care (GTAC) Consortium, Moreno CAM, Kamalakaran S, Goldstein DB; ALS Sequencing Consortium, Gitler AD, Harris T, Myers RM; NYGC ALS Consortium, Phatnani H, Musunuri RL, Evani US, Abhyankar A, Zody MC; Answer ALS Foundation, Kaye J, Finkbeiner S, Wyman SK, LeNail A, Lima L, Fraenkel E, Svendsen CN, Thompson LM, Van Eyk JE, Berry JD, Miller TM, Kolb SJ, Cudkowicz M, Baxi E; Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium, Benatar M, Taylor JP, Rampersaud E, Wu G, Wuu J; SLAGEN Consortium, Lauria G, Verde F, Fogh I, Tiloca C, Comi GP, Sorarù G, Cereda C; French ALS Consortium, Corcia P, Laaksovirta H, Myllykangas L, Jansson L, Valori M, Ealing J, Hamdalla H, Rollinson S, Pickering-Brown S, Orrell RW, Sidle KC, Malaspina A, Hardy J, Singleton AB, Johnson JO, Arepalli S, Sapp PC, McKenna-Yasek D, Polak M, Asress S, Al-Sarraj S, King A, Troakes C, Vance C, de Belleroche J, Baas F, Ten Asbroek ALMA, Muñoz-Blanco JL, Hernandez DG, Ding J, Gibbs JR,

Scholz SW, Floeter MK, Campbell RH, Landi F, Bowser R, Pulst SM, Ravits JM, MacGowan DJL, Kirby J, Pioro EP, Pamphlett R, Broach J, Gerhard G, Dunckley TL, Brady CB, Kowall NW, Troncoso JC, Le Ber I, Mouzat K, Lumbroso S, Heiman-Patterson TD, Kamel F, Van Den Bosch L, Baloh RH, Strom TM, Meitinger T, Shatunov A, Van Eijk KR, de Carvalho M, Kooyman M, Middelkoop B, Moisse M, McLaughlin RL, Van Es MA, Weber M, Boylan KB, Van Blitterswijk M, Rademakers R, Morrison KE, Basak AN, Mora JS, Drory VE, Shaw PJ, Turner MR, Talbot K, Hardiman O, Williams KL, Fifita JA, Nicholson GA, Blair IP, Rouleau GA, Esteban-Pérez J, García-Redondo A, Al-Chalabi A; Project MinE ALS Sequencing Consortium, Rogaeva E, Zinman L, Ostrow LW, Maragakis NJ, Rothstein JD, Simmons Z, Cooper-Knock J, Brice A, Goutman SA, Feldman EL, Gibson SB, Taroni F, Ratti A, Gellera C, Van Damme P, Robberecht W, Fratta P, Sabatelli M, Lunetta C, Ludolph AC, Andersen PM, Weishaupt JH, Camu W, Trojanowski JQ, Van Deerlin VM, Brown RH Jr, van den Berg LH, Veldink JH, Harms MB, Glass JD, Stone DJ, Tienari P, Silani V, Chiò A, Shaw CE, Traynor BJ, Landers JE.

5. Genetic analysis of neurodegenerative diseases in a pathology cohort
   Neurobiol Aging. (2019) doi: 10.1016/j.neurobiolaging.2018.11.007, (PMID: 30528841).
   Blauwendraat C, Pletnikova O, Geiger JT, Murphy NA, **Abramzon Y**, Rudow G, Mamais A, Sabir MS, Crain B, Ahmed S, Rosenthal LS, Bakker CC, Faghri F, Chia R, Ding J, Dawson TM, Pantelyat A, Albert MS, Nalls MA, Resnick SM, Ferrucci L, Cookson MR, Hillis AE, Troncoso JC, Scholz SW.

6. Mutations in the SPTLC1 gene are a cause of amyotrophic lateral sclerosis that may be amenable to serine supplementation.
   BioRxiv.(2019) doi: https://doi.org/10.1101/770339
   Janel O. Johnson[#], Ruth Chia[#], Danny E. Miller[#], Rachel Li[#], Ravindran Kumaran[#], **Yevgeniya Abramzon[#]**, Nada Alahmady[#], Faraz Faghri, Alan E. Renton, Simon D. Topp, Hannah A. Pliner, J. Raphael Gibbs, Jinhui Ding, Nathan Smith, Natalie Landeck, Michael A. Nalls, Mark R. Cookson, Olga Pletnikova, Juan Troncoso, Sonja W. Scholz, Marya S. Sabir, Sarah Ahmed, Clifton L. Dalgard, Claire Troakes, Ashley R. Jones, Aleksey Shatunov, Alfredo Iacoangeli, Ahmad Al Khleifat, Nicola Ticozzi, Vincenzo Silani, Cinzia Gellera, Ian P. Blair, Carol Dobson-Stone, John B. Kwok, Bryce K. England, Emily S. Bonkowski, The International ALS Genomics Consortium; The ITALSGEN Consortium; The FALS Sequencing Consortium; The American Genome Center; Pentti J. Tienari, David J. Stone, Karen E. Morrison, Pamela J. Shaw, Ammar Al-Chalabi, Robert H. Brown, Jr, Maura Brunetti, Andrea Calvo, Gabriele Mora, Hind Al-Saif, Marc Gotkine, Fawn Leigh, Irene J. Chang, Seth J. Perlman, Ian Glass, Christopher E. Shaw, John E. Landers, Adriano Chiò, Thomas O. Crawford, Bradley N. Smith, Bryan J. Traynor.

7. Shared polygenic risk and causal inferences in amyotrophic lateral sclerosis.
   Ann Neurol. (2019) doi: 10.1002/ana.25431, (PMID: 30723964).
   Bandres-Ciga S, Noyce AJ, Hemani G, Nicolas A, Calvo A, Mora G; ITALSGEN Consortium; **International ALS Genomics Consortium**, Tienari PJ, Stone DJ, Nalls MA, Singleton AB, Chiò A, Traynor BJ.

8. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia.
   Front Neurosci. (2020) doi: 10.3389/fnins.2020.00042, (PMID: 32116499).
   Yevgeniya A. Abramzon, Pietro Fratta, Bryan J. Traynor and Ruth Chia

9. Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis
   Neuron. (2021) doi: 10.1016/j.neuron.2020.11.005, (PMID: 33242422).
   Dewan R, Chia R, Ding J, Hickman RA, Stein TD, **Abramzon Y**, Ahmed S, Sabir MS, Portley MK, Tucci A, Ibáñez K, Shankaracharya FNU, Keagle P, Rossi G, Caroppo P, Tagliavini F, Waldo ML, Johansson PM, Nilsson CF; American Genome Center (TAGC); FALS Sequencing Consortium; Genomics England Research Consortium; International ALS/FTD Genomics Consortium (iAFGC); International FTD Genetics Consortium (IFGC); International LBD Genomics Consortium (iLBDGC); NYGC ALS Consortium; PROSPECT Consortium, Rowe JB, Benussi L, Binetti G, Ghidoni R, Jabbari E, Viollet C, Glass JD, Singleton AB, Silani V, Ross OA, Ryten M, Torkamani A, Tanaka T, Ferrucci L, Resnick SM, Pickering-Brown S, Brady CB, Kowal N, Hardy JA, Van Deerlin V, Vonsattel JP, Harms MB, Morris HR, Ferrari R, Landers JE, Chiò A, Gibbs JR, Dalgard CL, Scholz SW, Traynor BJ.

10. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture.
    Nat Genet. (2021) doi: 10.1038/s41588-021-00785-3, (PMID: 33589841)
    Chia R, Sabir MS, Bandres-Ciga S, Saez-Atienzar S, Reynolds RH, Gustavsson E, Walton RL, Ahmed S, Viollet C, Ding J, Makarious MB, Diez-Fairen M, Portley MK, Shah Z, **Abramzon Y**, Hernandez DG, Blauwendraat C, Stone DJ, Eicher J, Parkkinen L, Ansorge O, Clark L, Honig LS, Marder K, Lemstra A, St George-Hyslop P, Londos E, Morgan K, Lashley T, Warner TT, Jaunmuktane Z, Galasko D, Santana I, Tienari PJ, Myllykangas L, Oinas M, Cairns NJ, Morris JC, Halliday GM, Van Deerlin VM, Trojanowski JQ, Grassano M, Calvo A, Mora G, Canosa A, Floris G, Bohannan RC, Brett F, Gan-Or Z, Geiger JT, Moore A, May P, Krüger R, Goldstein DS, Lopez G, Tayebi N, Sidransky E; American Genome Center, Norcliffe-Kaufmann L, Palma JA, Kaufmann H, Shakkottai VG, Perkins M, Newell KL, Gasser T, Schulte C, Landi F, Salvi E, Cusi D, Masliah E, Kim RC, Caraway CA, Monuki ES, Brunetti M, Dawson TM, Rosenthal LS, Albert MS, Pletnikova O, Troncoso JC, Flanagan ME, Mao Q, Bigio EH, Rodríguez-Rodríguez E, Infante J, Lage C, González-Aramburu I, Sanchez-Juan P, Ghetti B, Keith J, Black SE, Masellis M, Rogaeva E, Duyckaerts C, Brice A, Lesage S, Xiromerisiou G, Barrett MJ, Tilley BS, Gentleman S, Logroscino G, Serrano GE, Beach TG, McKeith IG, Thomas AJ, Attems J, Morris CM, Palmer L, Love S, Troakes C, Al-Sarraj S, Hodges AK, Aarsland D, Klein G, Kaiser SM, Woltjer R, Pastor P, Bekris LM, Leverenz JB, Besser LM, Kuzma A, Renton AE, Goate A, Bennett DA, Scherzer CR, Morris HR, Ferrari R, Albani D, Pickering-Brown S, Faber K, Kukull WA, Morenas-Rodriguez E, Lleó A, Fortea J, Alcolea D, Clarimon J, Nalls MA, Ferrucci L, Resnick SM, Tanaka T, Foroud TM, Graff-Radford NR, Wszolek ZK, Ferman T, Boeve BF, Hardy JA, Topol EJ, Torkamani A, Singleton AB, Ryten M, Dickson DW, Chiò A, Ross OA, Gibbs JR, Dalgard CL, Traynor BJ, Scholz SW.

11. Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types.
    Sci Adv. (2021) doi: 10.1126/sciadv.abd9036, (PMID: 33523907)
    Saez-Atienzar S, Bandres-Ciga S, Langston RG, Kim JJ, Choi SW, Reynolds RH; International ALS Genomics Consortium; ITALSGEN, **Abramzon Y**, Dewan R, Ahmed S, Landers JE, Chia R, Ryten M, Cookson MR, Nalls MA, Chiò A, Traynor BJ.

12. Investigating RFC1 Expansions in Sporadic Amyotrophic Lateral Sclerosis
    Journal of The Neurological Sciences (Under review)
    Yevgeniya Abramzon[#], Ramita Dewan[#], Andrea Cortese, Susan Resnick, Luigi Ferrucci, Henry Houlden, Bryan J. Traynor.

ABBREVIATIONS

AD	Alzheimer disease
AFF2	AF4/FMR2 Family Member 2
ALS	Amyotrophic lateral sclerosis
AR	Androgen Receptor
ATN1	Atrophin 1
ATXN1	Ataxin 1
ATXN10	Ataxin 10
ATXN2	Ataxin 2
ATXN3	Ataxin 3
ATXN7	Ataxin 7
ATXN8OS	ATXN8 Opposite Strand LncRNA
BD	Bipolar disorder
Bis	Basophilic inclusions
C9ORF72	C9orf72-SMCR8 Complex Subunit
CACNA1A	Calcium Voltage-Gated Channel Subunit Alpha1 A
CANVAS	Cerebellar ataxia, neuropathy, and vestibular areflexia syndrome
CLN11	Ceroid lipofuscinosis, neuronal, 11
CMT2	Charcot-Marie-Tooth disease, type 2
CNBP	Cellular nucleic acid-binding protein
CNV	Copy-number variant
CSF	Cerebrospinal fluid
CSTB	Cystatin B
DIP2B	Disco Interacting Protein 2 Homolog B
DM1	Myotonic dystrophy type 1
DM2	Myotonic dystrophy type 2
DMPK	DM1 Protein Kinase
DMRV	Myopathy, rimmed vacuolar
DRPLA	Dentatorubral–pallidoluysian atrophy
EPM1	Unverricht–Lundborg disease
ET	Essential tremor

FECD	Fuchs endothelial corneal dystrophy
FMR1	FMRP Translational Regulator 1
FRA12A MR	Fragile X mental retardation
FRAXE MR	Fragile XE syndrome, mental retardation
FRDA	Friedreich's ataxia
FTD	Frontotemporal dementia
FXN	Frataxin
FXS	Fragile X syndrome
HD	Huntington disease
HDL-2	Huntington disease-like 2
HSP	Hereditary spastic paraplegia
HTT	Huntingtin
IBMPFD	Inclusion body myopathy with Paget disease and frontotemporal dementia
Indel	Insertion/deletion
JPH3	Junctophilin 3
LMN	Lower motor neurons
MND	Motor neuron disease
MSC	Mesenchymal stem cells
NADGP	Neurodegeneration, childhood onset with ataxia, dystonia & gaze palsy
NGS	Next-generation sequencing
NOP56	NOP56 Ribonucleoprotein
NTF	Neurotrophic factors
PD	Parkinson disease
PDB	Paget's disease of bone
PNFA	Progressive nonfluent aphasia
PPA	Primary progressive aphasia
PPP2R2B	Protein Phosphatase 2 Regulatory Subunit Bbeta
RFC1	Replication Factor C Subunit 1
SBMA	Spinal and bulbar muscular atrophy
SCA1	Spinocerebellar ataxia type 1
SCA10	Spinocerebellar ataxia type 10

SCA12	Spinocerebellar ataxia type 12
SCA17	Spinocerebellar ataxia type 17
SCA2	Spinocerebellar ataxia type 2
SCA3	Spinocerebellar ataxia type 3
SCA36	Spinocerebellar ataxia type 36
SCA6	Spinocerebellar ataxia type 6
SCA7	Spinocerebellar ataxia type 7
SCA8	Spinocerebellar ataxia type 8
SD	Semantic dementia
SNGP	Supranuclear gaze palsy
SNV	Single-nucleotide variants
SPG6	Spastic paraplegia 6
TBP	ATA-Box Binding Protein
TCF4	Transcription Factor 4
UMN	Upper motor neurons
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

1 INTRODUCTION

1.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

I performed the literature search and composed this chapter.

1.2 GENETIC RESEARCH IN THE PAST TWO DECADES

The last two decades have brought much excitement to the field of genetic research. Since the first draft of the human genome sequence in 2001 [1], there have been many revolutionary innovations and technical advancements.

In 2002, International HapMap Project was launched as a part of The Human Genome Project after the publication of an interim report that identified approximately 1.5 million single nucleotide polymorphisms (SNP) in the human genome.[2] The HapMap aimed to examine genetic diversity. The project had two primary outcomes: determining allelic frequencies of 1 million SNPs in eleven different populations worldwide and creating the genome-wide linkage disequilibrium map of the genes in the human population. Since then, HapMap became a valuable database for genetic association studies and a powerful source for studying the genetic factors contributing to changes in response to environmental factors.

The identification of millions of single nucleotide polymorphisms (SNPs) across the human genome and the realization that a subset of these SNPs can capture common genetic variations via linkage disequilibrium made the first genome-wide association study (GWAS) possible in 2002.[2,3] GWAS studies focus on determining associations

between SNPs and traits associated with a specific disease by comparing common genetic variants in large numbers of affected cases to those in unaffected controls. Current GWAS studies utilize microarray technology for SNP genotyping and include a meta-analysis of tens of thousands of individuals. To date, over 4,000 human GWAS studies have examined over 2,000 diseases and traits and resulted in the discovery of thousands of SNP associations.[4]

The Human Genome Project, in 2001, determined that humans have approximately 20,000 genes, which account for about 1.5% of DNA in the human genome. The other ~98% of the human genome was traditionally regarded as "junk." The ENCODE project was created in 2003 to determine the rest of the genome's role and establish a complete list of functional elements, including chromatin structure and histone modification regions, transcriptional regulatory sequences, and promoters.[5]

By 2012, this effort led to the publication of more than 30 research papers based on 1,640 datasets describing the active regions of the human genome.[6] The ENCODE was one of the most significant efforts in the history of genomics to understand the complexity and meaning of human genomic data. Upon completion in 2013, The ENCODE assigned biochemical function to approximately 80% of the human genome, providing new insights into gene organization and, most importantly, regulation mechanisms. This was the second most significant step toward understanding functional genomics since the release of the initial draft of the human genome sequence.

In 2008, the 1000 Genomes Project took advantage of the invention of the next-generation sequencing technology that significantly reduced sequencing costs. The idea was to create a detailed catalog of human genetic variation by looking at complete genome sequences of 1000 individuals in a broader context and determining the

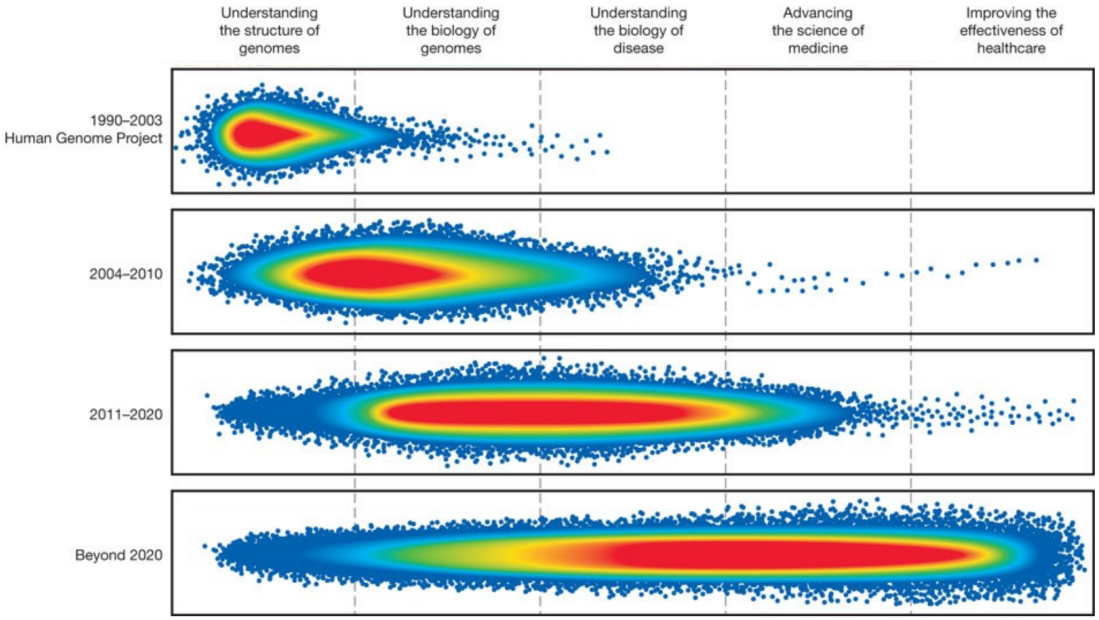
variations in coding and non-coding regions. Upon completion, the data was made available to the worldwide scientific community and became the foundation of the most extensive public catalog of human variation and genotype data.[7,8]

2009 became the rise of epigenetics studies. These studies focused on examining how methylation at the chromatin and DNA level can change genome functions. Epigenetic studies made a significant contribution to the understanding of the complexities of genetic regulation, cellular differentiation, aging, and disease and established the previously missing link between genetics, disease, and the environment.[9]

In 2010, the UK10K project compared 4,000 genomes of healthy individuals with 6,000 people living with a disease of suspected genetic cause to investigate how low-frequency and rare genetic variants contribute to human disease. The data generated by UK10K was made available to the research community during the project's life, providing access to data an order of magnitude more profound than was previously possible and enabled the discovery of novel disease-causing genes.[10]

In 2018, the 100K genomes project announced that it had completed the sequencing of 100,000 genomes from patients affected by a rare disease, an infectious disease, or cancer. The project aimed to provide data for scientific discovery, making it the first-ever research–clinical hybrid project and ultimately make genetic testing a routine technique for patient diagnostics and treatment in the UK. The data accumulated by the 1000K genomes project provided a new foundation for genetic research of rare diseases and cancer.[11]

Figure 1. Accomplishments in genomics research over the past two decades.



(Reproduced from Green et al., 2011)

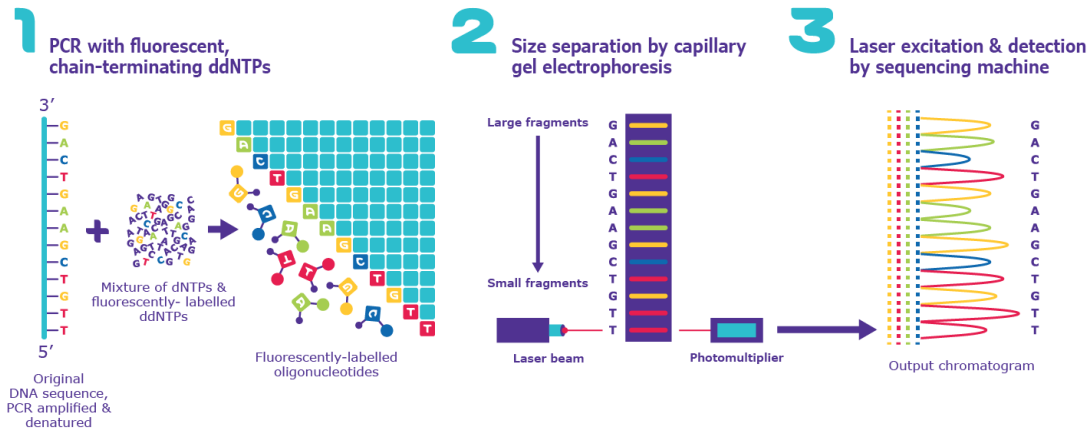
1.3 ADVANCES IN SEQUENCING TECHNOLOGIES

The last three decades have also seen significant improvements in sequencing technologies **(Figure 6)**. The first efforts at sequencing were labor-intensive and time-consuming because the process was completely manual and involved traditional test tubes and gel electrophoresis. The development of the first sequencing technology by Applied Biosystems automated dramatically sped up and refined the sequencing process. Since then, the evolution of next-generation sequencing technologies has increased the technological capabilities to find genes in tandem with decreasing costs.

1.3.1 First-generation sequencing - Sanger Sequencing

The first sequencing method, known as Sanger sequencing, was developed by Dr. Frederick Sanger in 1977. Sanger sequencing is a chain termination method that relies on the use of dideoxynucleotides (ddNTP) that has a hydrogen group (-H) at the 3' end of the molecule as opposed to the standard hydroxy functional group (-OH). The process begins with restricting enzymes to break the DNA to produce fragments, which are then replicated by polymerase chain reaction (PCR). Next, the primers anneal to the DNA region of interest. In the presence of the four nucleotides (A, T, G, C), the polymerase extends the primer by adding on the complementary nucleotides from the template to the 5' end of the DNA strand. ddNTPs added to the PCR solution randomly terminate the replicating chains by preventing more nucleotides from bonding to them. Their incorporation results in the production of DNA products of various sizes. The resulting PCR products are run on a polyacrylamide gel using an electric force to make the gel's fragments travel. Next, the gel is exposed to the laser to detect fluorescently labeled ddNTPs to determine the fragment size. In the final step, the bands are combined to form a single sequence, including all nucleotides **(Figure 2)**.

Figure 2. Three basic steps of Sanger Sequencing.



(Reproduced from www.sigmaaldrich.com)

1.3.2 Second-generation sequencing

1.3.2.1 Roche 454 Genome Sequencer FLX

Despite the low output and accuracy of the Sanger Sequencing, several new technologies were developed from the advances made by this method. Roche 454 Genome Sequencer FLX was the first next-generation sequencing technology. It took Sanger's fluorescence detection method and simplified the preparation process. The sequencer became widely available as a commercial product in 2005 and made the first mass sequencing projects possible.

Roche utilized adapter flanked fragments attached to fiber chips to hold primers, polymerize enzymes, and start the synthesis of complementary strands. The 454-sequencer utilized the emulsion PCR amplification method, which replicated the strands attaching them to capture beads. This ensured that the reaction could be detected

at specific light intensity. The samples were then loaded onto a picotiter plate, where the beads with attached DNA fragments entered individual wells. Packing beads were then added to wells to aid the spectrometer with reading the sample.

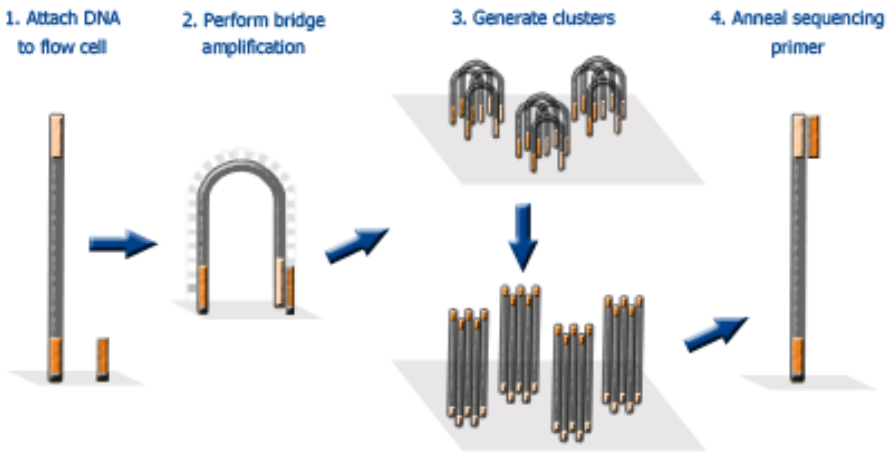
Roche 454 Genome Sequencer FLX could analyze many samples in parallel, significantly improving sequencing output compared to Sanger sequencing.

1.3.2.2 HiSeq (Illumina)

Illumina HiSeq is another platform that relies on the chain termination method. Like the other techniques, Illumina library preparation starts with DNA shearing into fragments and running the gel's fragments to separate them based on size. The 200 - 300 bp fragments are then selected for further replication through PCR amplification. Illumina uses automatic cluster generation to distribute the fragment library to the surface of a flow cell that contained attached adapters. The library fragments adhere to the flow cell by covalently binding the adapters. Next, bridge amplification generates hundreds of copies of DNA fragments of interest. This process forms library clusters. Finally, the sequencing primer is added to the 5' end of each library, and polymerase starts extending the strand adding one nucleotide at a time. When each new base is added, a camera records each cluster's location by reading the fluorescent signal emitted by the cluster. The combination of these images creates the read sequence.

Illumina HiSeq was an extremely efficient platform for its time. It would analyze more than 150 million library clusters in one run. This platform was also producing the highest output among all second-generation platforms (600 Gb). However, it had a slightly lower accuracy compared to Roche 545 FLX and SOLiD.

Figure 3. The core of HiSeq sequencing.



(Reproduced from www.eurofinsgenomics.co.in)

1.3.2.3 SOLiD (Applied Biosystems)

SOLiD stands for Sequencing by Oligonucleotide Ligation and Detection. This platform's flexibility allowed the use of this next-generation technology for different applications such as whole-genome sequencing, targeted resequencing, gene expression, small RNA analysis, and chromatin immuno-precipitation. The workflow for each of these methods differed only by sample preparation and data analysis, while the sequencing method remained the same.

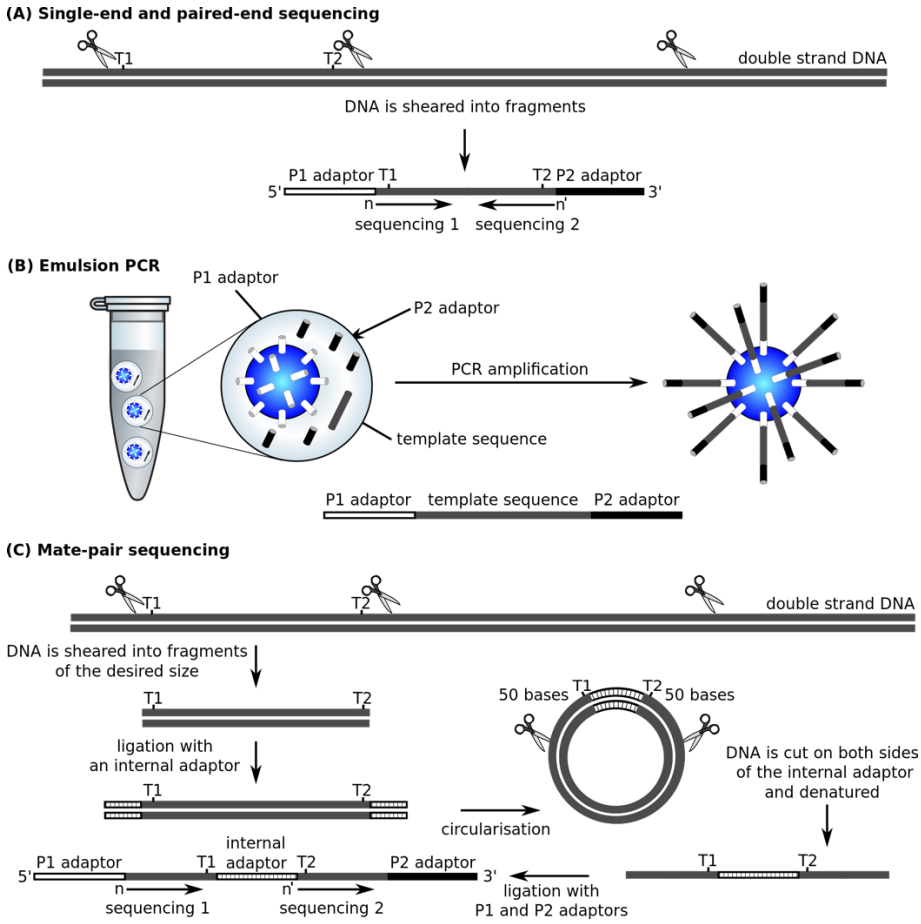
With this technology, the research community could perform a single or paired-end sequencing for the first time. In both cases, targeted DNA was sheared to a specific size, and adapters were attached to both 5' and 3' ends of the DNA to create a complete library molecule. In cases of paired-end sequencing, the internal adapter was also used to hold the two DNA fragments together.

Like Roche 454, SOLiD used emulsion PCR, but it distributed its fragments on microbeads that could vary in size and richness of slides. The users could pick the slides containing 1, 4, or 8 sections based on the targeted application.

The fluorescence signal was read out during the fragment ligation into a single strand and determined the sequence of nucleotides in a read. Data analysis was accomplished through a 2 Base Encoding system, in which a base and a color define the next base in the sequence. This system quickly detected read errors such as single-base insertions/deletions and SNP's.

SOLiD was the most accurate among its second-generation peer platforms. It made sequencing inexpensive by significantly increasing the sequencing output. The platform could sequence 265 samples and produce up to 30 Gb of data in a single run. However, this technology's main disadvantage was that it produced short reads (~50 bp), making the data unsuitable for many applications.

Figure 4. SOLiD platform sequencing.



(Reproduced from Hupe et al., 2012)

1.3.3 Third-generation sequencing

Third-generation sequencing technologies include Helicos True Single-molecule Sequencing and the Pacific Biosciences Single-molecule Real-time technology. The advancement of these technologies over the second-generation sequencing technologies is an ability to produce substantially longer reads.

1.3.3.1 Helicos tSMS

Helicos true single-molecule sequencer was the first sequencing technology that allowed for sequencing without PCR amplification, enabling shorter preparation times while decreasing the chances of error.

In the first step of the sample preparation, DNA was broken into 100 -200 bp fragments. A polyA primer was attached to the 3' end of each fragment. Next, a fluorescently labeled nucleotide was added to the polyA tail. These DNA fragments were then hybridized onto the surface of a flow cell that contained immobilized oligo-T-nucleotides complementary to the polyA primer. DNA polymerase and fluorescently labeled nucleotides were added to the flow cell, resulting in the nucleotides being added complementary to the template fragment one of a time.

Similar to Illumina HiSeq sequencing, the laser was used to the flow cell and capture the emitted fluorescent signal. However, Helicos recorded each nucleotide's addition to the DNA fragment instead of the Illumina cluster-based sequencing system. This allowed for billions of unique fragments to be independently sequenced at the same time.

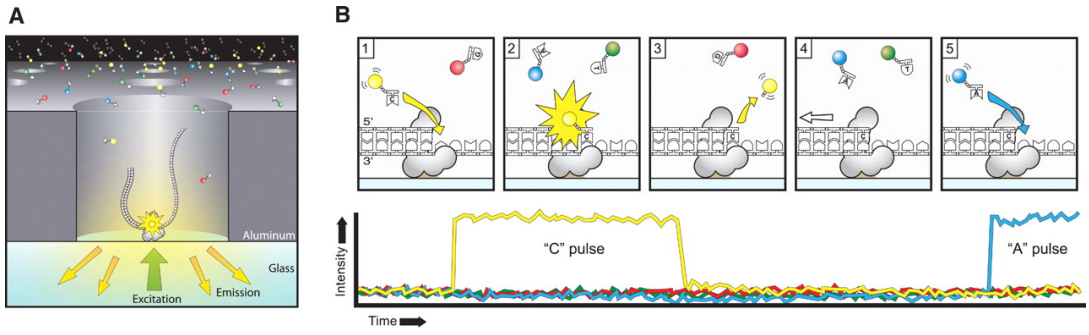
The Helicos technology has been shown to have nearly 100% accuracy because the sequencing does not involve PCR amplification. However, that was an expensive sequencing method because it relied on sequencing reagents' heavy continuous use.

1.3.3.2 Pacific Biosciences SMRT sequencing

Pacific Biosciences has developed a single-molecule real-time DNA sequencing technology (SMRT). First, fluorescence is attached to a nucleotide creating phospholinked nucleotides with a different color for each of the four nucleotides. Unlike the previous methods, the fluorescent label was attached to the terminal phosphate of a nucleotide instead of the base disclose. This allowed the DNA polymerase to cleave the fluorescent label off during the base edition. This action produced light that was captured in a nanophotonic chamber.

The SMRT sequencing used zero-mode waveguides, small pores surrounded by metal film and silicon dioxide, that enabled the section of single molecules while DNA polymerase replicates the chain inside the well. As a result, this method allowed for a simpler and faster genome assembly than comparative technologies while allowing for longer wavelengths.

Figure 5. Pacific Biosciences SMRT sequencing process.



(Reproduced from Eid et al., 2009)

1.3.4 Fourth-generation sequencing

1.3.4.1 Oxford Nanopore

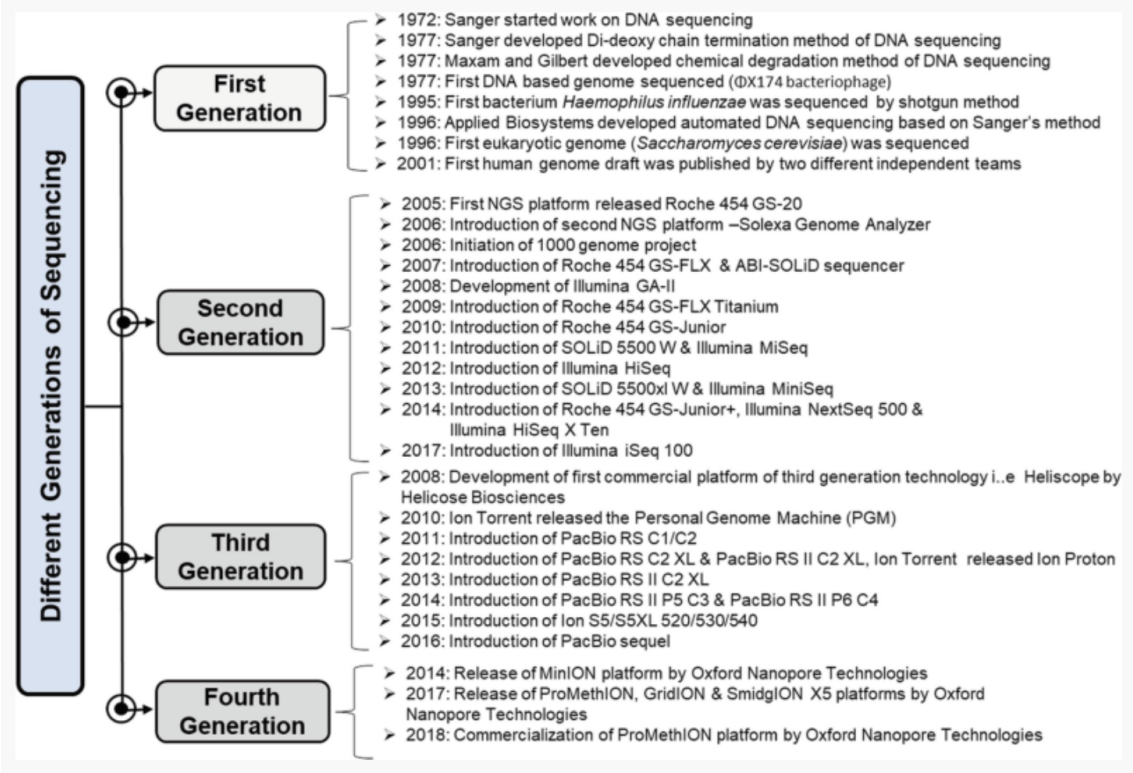
Fourth-generation sequencing technologies include four sequencing platforms developed by Oxford Nanopore Technologies: MinION, ProMethION, GridION, and SmidgION (X5). Nanopore technology is suitable for many applications such as DNA, miRNA, and protein analyses. The significant advantage of this generation of sequencing technology is that it produces reads up to 100,000bp in length, which in many cases eliminates the need for sequencing assembly.

Nanopore technology is based on a protein pore called nanopore that is attached to the synthetic polymer substrate. A transporter protein attaches itself to a DNA strand and starts "unzipping" the DNA into two complementary single strands. As DNA gets unzipped, one strand is pulled through the nanopore. Each passing nucleotide causes a disturbance in the current that passes through the nanopore. Since each of the four bases has a slightly different size and structure, they result in different current disturbances. These disturbances are recorded to produce a sequence read.

Oxford Nanopore technology is the first technology to allow real-time sequencing and analysis. The significant advantage of this technology over the previous generations is that it can produce ultra-long reads up to 100,000 bp. This, in most cases, eliminates the need for read assembly. While this technology has lower accuracy of 93% in a single read, compared to the other methods, the 100% accuracy can be achieved by running the same DNA molecule through the nanopore several times and combining the reads information to fill in the gaps.

The last two decades brought new insight into the understanding of the human genome's structure and functionality, which led to significant improvement in our understanding of many diseases' genetics. The advancements in sequencing technologies significantly reduced the cost and time need for exome and whole-genome sequencing. These advancements made sequencing a reliable and affordable method for precision diagnostics and built a strong foundation for personalized medicine development.

Figure 6. Evolution of sequencing technologies.



(Reproduced from Gupta and Verma, 2019)

1.4 SEQUENCING TECHNIQUES USED TO COLLECT DATA FOR THIS THESIS

1.4.1 Whole Exome sequencing

DNA sequencing for the diagnosis of disease continues to develop at a rapid pace. Although the exome makes up only 1.5% of the human genome, most genetic disorders are found in coding regions, and the effects of single base pair changes are still best understood in the realm of the exome. Advantages of whole-exome sequencing (WES) and analysis include faster time than whole-genome sequencing at considerably less cost. Exome sequencing is also a powerful technique to analyze hereditary diseases with ambiguous phenotypic manifestations. Furthermore, exome sequencing delivers fewer background mutations than traditional Sanger-sequencing. This is mainly an advantage with conditions that necessitate the analysis of more than one candidate gene.[12] The introduction of next-generation exome sequencing provides an unbiased analysis of all protein-coding sequences in the approximately 20,000 genes in the human genome.

Whole-exome sequencing also has its disadvantages – it only picks up about 85-90% of the sequences that encode proteins, potentially leading to missing mutations.[13] This is due to the GC-rich content of the first exon of genes. Genome sequencing overcomes this and gives more complete and even sequencing because there is no capture used.

Also, exome sequencing does not reveal the functional consequences of many sequence variants. Determining the significance of specific mutations by analyzing whether it exclusively segregates with a symptomatic phenotype is not always possible. Neurological disorders like ALS are fatal and thus reduce the pool of living individuals

that carry the disease. To overcome this problem, large-scale genomic initiatives, such as the United Kingdom's 100,000 Genomes Project, are being conducted. They will allow the scientific community to better interpret variants' significance in the general population and their pathogenic potential.[14,15]

1.4.2. Whole-genome sequencing

Whole-genome sequencing (WGS) is an increasingly used technique due to its broader coverage and decreasing cost. It is still difficult to interpret variants lying outside the protein-coding regions of the genome.[16] Although most coding variants can be detected by WES, WGS can also detect these variants. Using the most recent next-generation sequencing (NGS) technologies, Belkadi et al., in 2015 compared WES and WGS, in terms of detection rates and quality, for single-nucleotide variants (SNVs), small insertions/deletions (indels), and copy-number variants (CNVs) within the regions of the human genome covered by WES.[16] Results showed WGS was slightly but significantly more potent than WES for detecting variants in the regions covered by the exome kit, particularly for SNVs.[16] Besides, they found WGS better for detecting CNVs since it could detect variations in RNA- and protein-coding regions not covered by the exome sequencing.[16] WGS also has more reliable sequence coverage than WES, which makes this method more suitable for CNV detection.[17,18] Regions with low sequence complexity withing genome restrict the ability to design good WES capture baits, resulting in off target capture effects that produce low-coverage regions. WGS forgoes capturing and therefore is less sensitive to GC content and more likely to provide complete coverage of the entire coding region of the genome.[17]

1.4.2.1 Batch effects resulting from variation in capture methods

The rapid development of sequencing technologies in the past 20 years has led to variation in library preparation methods, sequencing reagents, flow cells, and bioinformatics tools for read alignment and variant calling. These differences created the potential for batch effects. These batch effects can occur within experiments. However, they are most marked between the experiments at the same centers or across center due to the use of different capture libraries (eg. Nimblegen or Illumina).

Batch effects are usually addressed through quality control (QC) measures. In WES, these include methods for removing batch effects in copy number variation calling and empirically derived variant filtering. To standardize the reading depth and eliminate batch effects, these algorithms rely on reference panels, principal component analysis, or singular value decomposition. Unfortunately, no standardized algorithms were developed to address the issue of batch effects in WGS. Because of this, batch effects in WGS are typically handled by adopting stringent QC measures. Another method that facilitates this batch effect is joint calling of the raw data using the same sequence alignment pipeline.

1.4.3 ALS and FTD

Various neurological diseases, including ALS, FTD, and ALS/FTD, have been shown to have a genetic basis. Mutation analysis has shown common causative factors for two major neurodegenerative disorders, ALS and FTD. Shared pathological and genetic markers and common neurological signs between these two diseases have given rise to the idea of an ALS/FTD spectrum.[19] Research has also shown that the genes responsible

for these disorders also share common pathophysiological mechanisms. The overlap among genetic factors causing the spectrum of ALS, FTD, and ALS/FTD disorders, and the coincidence of mutated alleles, including variants, have provided evidence for an oligogenic model of this disease spectrum.[19,20] Phenomena such as epistasis, oligogenic inheritance, and genetic modifications are all synonyms describing the same process - the effect of one gene/allele on the phenotypic outcome of a second gene/locus.[20] Also, genes with mutations and those with differential DNA methylation are involved in common pathways and may be critical in neurodegeneration involved in both FTD and ALS/FTD.[21] Both disorders are characterized by defects in RNA processing, protein clearance by autophagy, vesicle trafficking, mitochondrial dysfunction, and impaired protein homeostasis. Although new genes are being discovered as research progresses, more than 20 genes have now been described in the molecular mechanisms associated with both disorders – SQSTM1, TDP-43, VCP, FUS, TBK1, MATR3, GRN, CHCHD10, SOD1, and C9orf72, most of which were discovered utilizing exome or whole-genome sequencing technologies.[22]

2 NEUROLOGICAL DISORDERS EXAMINED IN THIS THESIS

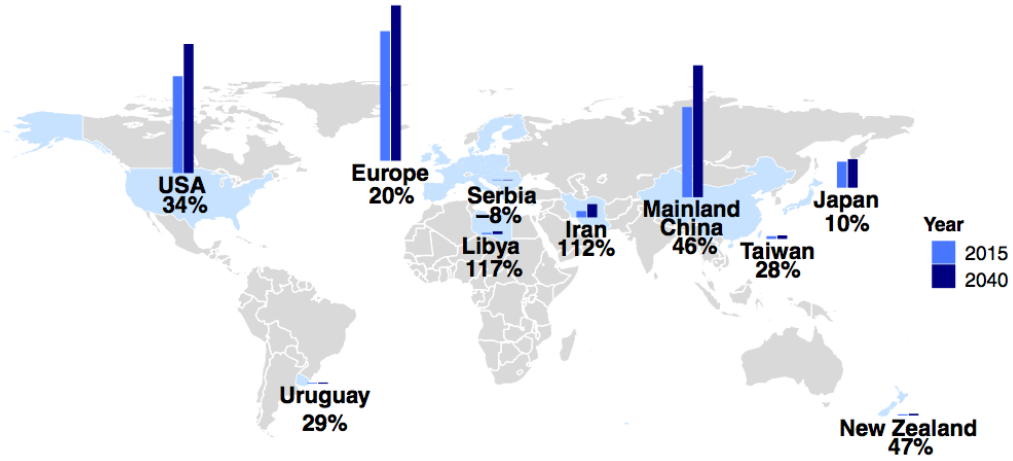
2.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

I performed the literature search and composed this chapter.

2.2 AMYOTROPHIC LATERAL SCLEROSIS

Amyotrophic lateral sclerosis (ALS), also known as motor neuron disease and Lou Gehrig's disease after New York Yankees baseball player, is an adult-onset fatal neurological disorder that results in progressive degeneration of both upper and lower motor neurons. It is a rapidly progressing neurodegenerative disorder with a poor prognosis. The majority of patients die of respiratory failure two to five years after the first onset of symptoms.[23] Less than 10% of patients survive past the five-year mark, and even fewer patients make it past ten years.[24] ALS mainly affects adults 50 years of age and older. With increased longevity due to advances in medical care plus the rising aging population, the number of ALS cases is likely to only increase in the future. Estimates show that the number of cases will rise to 400,000 by 2040, primarily due to extended average life expectancies **(Figure 7)**.[25]

Figure 7. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040.



(Reproduced from Arthur et al., 2016)

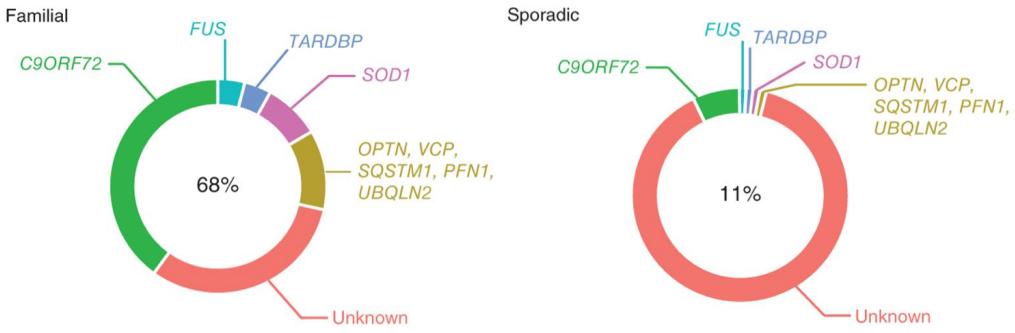
ALS results in impairment of the voluntary motor system and is associated with ubiquitin-positive intraneuronal aggregates in anterior horn cells. In contrast, FTD patients exhibit a form of neurodegenerative dementia and are associated with cortical ubiquitin-only neuropathology (FTD-U). Studies of both patient cohorts suggested a clinical and pathological overlap and shared pathogenesis between the two diseases.[26]

2.2.1 Genetics of ALS

Approximately 90% of ALS cases are sporadic, occurring with no discernable family history of the disease, and 5-10% of ALS cases are genetically linked **(Figure 8)**. Due to recent advances in genetic research, the genetic etiology of familial ALS is now known in approximately two-thirds of cases. On the other hand, genetic association studies in patients with sporadic ALS have uncovered only a small number of mutant

loci and polymorphisms, so the etiology of the majority of sporadic ALS cases remains unknown.[27,28]

Figure 8. Loci associated with sporadic and familial ALS.



(Reproduced from Renton, Chio, and Traynor, 2014)

ALS is caused by a combination of genetics and environmental factors. Approximately twenty genes have been identified and implicated in the development of ALS. Mutations in those genes increase the likelihood that a person will develop this devastating neurological disorder. Environmental factors that increase a person's risk of developing ALS include smoking, which will increase a person's risk by two times compared to that of non-smokers, heavy metal exposure, and chronic exposure to electromagnetic fields.[29]

2.2.2 Heritability of sporadic ALS

ALS is a rare neurological disorder that has a complex genetic origin. Approximately 10% of all individuals diagnosed with ALS have familial form of the disease, meaning that they have at least one other affected member in the family. The remaining 90% of

ALS cases are sporadic with no family history of the disease. A number of loci relevant to ALS pathogenesis have been identified over the years. Those explain up to two-thirds of familial cases.[30] In sporadic ALS cases however, diagnostic advancements have only helped in explaining a fraction of cases, with the etiology remaining unexplained in over 90% of patients.[28]

Over the last two decades multiple studies used twin, pedigree, and population-level genome-wide SNP data to determine the broad and narrow-sense heritability of ALS. The broad-sense heritability of ALS captures the proportion of phenotypic variation due to genetic values that may include effects due to dominance and epistasis. Twin studies comparing monozygotic and dizygotic twin pairs from overlapping clinic and population-based ALS datasets estimated the total heritability of ALS ranges between 0.38 and 0.85.[31,32] A family study based on data from a clinic-based registry estimated ALS heritability to be between 0.40 and 0.60.[33] Narrow-sense heritability of ALS, also known as SNP-based heritability, is the proportion of variance that is due to additive effects of genetic variants. GWAS study estimated SNP-based heritability of ALS to be 0.21.[34]

There are several factors that may contribute to the missing heritability in ALS. Structural variations that include deletions, insertions, inversions and microsatellites remain underexplored potential genetic modifiers of ALS and other diseases. The other factor that may contribute to the missing heritability in ALS is the potential oligogenic nature of the disease. The difficulty in uncovering the genetic determinants of ALS suggests the disease has a complex genetic architecture that may consist of combinations of gene variants differing in frequency and noxiousness.[35,36] Furthermore, many variants discovered by genome wide association studies have uncertain

significance. Low frequency variants could account for a substantial part of the missing heritability in ALS. Variants of low minor allele frequency may not be captured by current genotyping arrays and effect sizes may not be large enough to be detected by linkage analysis in families.[37]

2.2.3 Types of motor neuron disorders

There are four main subtypes of motor neuron disorders (MND): classic ALS, primary lateral sclerosis, progressive bulbar palsy, and progressive muscular atrophy. Each subtype is defined by either upper, lower, or upper and lower motor neuron involvement.

Classic ALS, characterized by the upper and lower motor neuron dysfunction, is the most common type of MND. About 70% of patients present with this form and therefore classic ALS is perhaps the most studied type of MND. Classic ALS has a strong genetic component. Disease-causing genes have been identified in 5-10% of clinical ALS.

Progressive bulbar palsy refers to a subcategory of MND patients that present with symptoms in the head or neck due to the loss of upper and lower motor neurons from those areas. Progressive bulbar palsy affects motor neurons in the brain's corticobulbar region that controls the muscles of the face, head, and neck. Pseudobulbar syndrome may present with involuntary excessive laughing and crying.[38] Approximately 25-30 % of all MND patients present with this form of the disease.

Primary lateral sclerosis (PLS) is the rarest form of MND, representing only 1-4 % of all cases.[39] PLS is due to upper motor neuron degeneration, therefore consistent with

signs and symptoms of upper motor neuron (UMN) disorder in the absence of lower motor neuron (LMN) symptoms. It is a diagnosis of exclusion, meaning individuals are diagnosed with this condition when no other medical reason can explain their progressive motor neuron dysfunction.[40,41] PLS is classified as a slowly progressive type of MND. Unlike classic ALS, the average PLS symptom duration ranges between 7.2 and 14.5 years.[42] Prognosis is the crucial differentiating factor between PLS and ALS.[43] However, it is still common for patients to be diagnosed with PLS but later re-diagnosed with ALS.

Since the initial discovery of progressive muscular atrophy (PMA) in the mid-1800s, there has been debate in the scientific community over whether it is a distinct disorder with its characteristics or subtype of MND.[44] PMA accounts for approximately 10% of all MND cases.[45] Similarly to other types of MND, PMA may start with lower motor neuron symptoms and later progress to include upper neuron symptoms. At that point the diagnosis changes from PMA to ALS.[46]

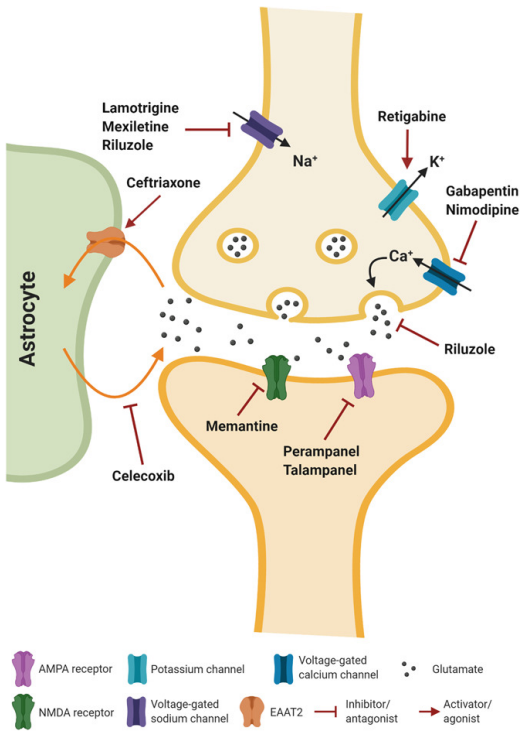
2.2.4 FDA-Approved treatments for ALS

Despite the advances in genetic research and our knowledge of the molecular and functional biology of motor neuron degeneration over the past few decades, ALS patients still have very few therapeutic options. Current FDA-approved treatments for slowing disease progression include two drugs Riluzole and Edaravone.

2.2.4.1 Riluzole

Riluzole (Rilutek) was the first FDA-approved treatment for ALS. The drug inhibits the release of neurotransmitter glutamate from the nerve cells reducing the rate of glutamate-induced neurodegeneration.[47] Riluzole mechanism of action is summarized in **Figure 9**. Riluzole is currently available only in a 50 mg tablet form. The drug is administered orally twice a day, one hour before or two hours after a meal.

Figure 9. Riluzole mechanism of action.



(Reproduced from Wobst et al., 2020)

Riluzole blocks the sodium channels in neurons and prevents increased intracellular sodium concentration, contributing to cellular death inhibition.

The effectiveness of Riluzole was first examined by Bensimon et al. in a study of 155 ALS patients in which participants were randomized to either 100 mg/day Riluzole or placebo. Riluzole recipients had a significantly longer survival of up to 12 months compared to the placebo group. The drug had an even better effect on patients with bulbar-onset ALS. In addition to prolonged survival, the group that received Riluzole demonstrated decreased muscle stiffness. However, treatment effectiveness decreased between 12 and 19 months (end of the placebo control period).[48,49]

A follow-up study by Lacomblez and colleagues focused on the dose-dependent effects of Riluzole. Nine hundred fifty-nine participants, diagnosed with clinically probable or definite ALS according to the revised El Escorial criteria and with less than five years since the symptom onset, were randomly divided into four groups to receive one of the following treatments: 50 mg, 100 mg, 200 mg of Riluzole, or placebo. The participants were followed up for 18 months.

The study showed that 50.4% of patients that received a placebo and 56.8% of those who received 100 mg/day of Riluzole were alive and did not require tracheostomy. Similarly, patients that received 200 mg/day of Riluzole had higher survival and slower disease progression than patients who received a 50 mg dose of the drug. No significant difference in performance or mortality was observed among the patients that received different doses of Riluzole. Adverse effects that included asthenia, dizziness, gastrointestinal disorders, and rises in liver enzyme activities were reported in the

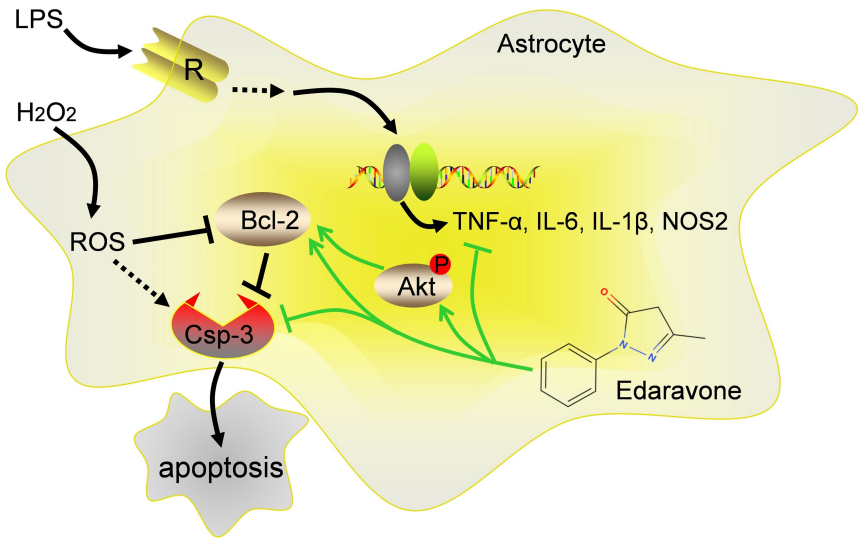
patients receiving Riluzole and were more common in patients that received the highest dose (200 mg).[50]

The analysis of randomized controlled trial data by Bensimon et al., and Lacomblez et al., suggested that Riluzole extends the survival of ALS patients by two to three months and increases the chance of surviving an additional year by approximately 9%.[48,50] However, recent data published in clinical ALS databases showed that Riluzole therapy might prolong patient survival for more than 19 months.[51]

2.2.4.2 Edaravone

In 2017, the FDA approved a new drug, Edaravone (Radicava, Radicut), to treat ALS patients in the United States. Edaravone reduces oxidative stress by neutralizing the free oxygen radicals resulting in a significant reduction in motor movement decline.[52]

Figure 10. A possible Edaravone mechanism of action in astrocytes.



(Reproduced from Guo et al., 2020)

The exact mechanism of the drug's action in ALS is not well understood. Edaravone is available in 60 mg intravenous form. The drug is administered through IV over a 1-hour infusion every day for 14 days in the initial treatment. The treatment is followed by a 14-day drug-free observation period. Successive treatments are done in cycles of daily dosing for ten days followed by 14-day drug-free periods.[31]

Phase 2 study that focused on examining therapeutic effects and safety of Edaravone involved 20 ALS patients. The participants were randomized to receive a daily dose of either 30 mg or 60 mg of Edaravone via an intravenous drip for 14 days, followed by observation for another two weeks. The four-week treatment-observation cycle was repeated six times during the six-month study period. The ALSFRS-R score change was used to determine the drug's efficacy, and the patients' CSF 3-nitrotyrosine (3NT) level, a marker for oxidative stress, was used to determine the drug pharmacokinetics. The study showed that patients receiving 60 mg of Edaravone had significantly less decline in ALSFRS-R score than before Edaravone administration. Most of the Edaravone recipients also had reduced CSF 3NT levels compared to the placebo group.[53]

A phase 3 trial was designed to confirm the efficacy and safety of Edaravone in ALS patients. Two hundred and six patients diagnosed with definite, probable, or probable laboratory-supported ALS, with less than three years since the symptom onset and with forced vital capacity (FVC) of at least 70% were recruited to the study. The participants were randomly divided into two groups to receive a placebo or a daily infusion of Edaravone for the first 14 days in cycle 1 and for 10 of the first 14 days during cycles 2 to 6. The study was conducted over 36 weeks, a 12-week assessment period followed by a 24-week treatment period. Change in ALSFRS-R score was used to measure the

primary outcomes of the study. At the end of the study, the Edaravone group had a more minor reduction of ALSFRS-R than the placebo group. However, the efficacy of Edaravone for the treatment of ALS has not been confirmed.[52]

A posthoc analysis of an open-label follow-up of 65 patients who received Edaravone treatment for 48 weeks and 58 patients who received Edaravone for 24 weeks showed a continuous benefit from receiving the drug. Patients who switched from placebo to Edaravone during the study had a 34% reduction in ALSFRS-R score compared to the patients who continued taking placebo. Furthermore, patients who received the treatment for 48 weeks showed a 38% slowdown in disease progression compared to the patients who received a placebo for the same period. The study participants reported the most common side effects were bruising at the injection site, difficulty walking, and headache.[54]

A study from Japan focused on the long-term effects of Edaravone on ALS patient survival. The study recruited 57 ALS patients between 2010 and 2016. Twenty-seven participants received Edaravone treatment administered as described in the previous studies, two weeks of daily infusion followed by a two-week drug-free observation period. Thirty study participants were in the control group and received no treatment. All patients were assessed at the beginning of the study at 6, 12, and 18 months. Changes in serum creatinine, a biomarker for ALS progression, and differences in ALSFRS-R score were used to assess the drug's effectiveness. The study showed that the patients who received Edaravone had significantly reduced ALSFRS-R scores, improved serum creatinine levels, and improved survival rate compared to the control group.[55]

Considering patients' discomfort during Edaravone IV infusion and the inconvenience of drug administration in the medical facility, an oral tablet formulation of Edaravone was developed but failed in 2011 due to poor bioavailability.[56] In 2018, a sublingual (SL) tablet formulation of Edaravone was developed. A study that aimed to determine the bioavailability of this new drug was conducted in 10 healthy volunteers that were randomized to receive one of the following treatments: (1) SL Edaravone 30mg tablet, followed by Edaravone 30mg IV infusion; or (2) Edaravone 30mg IV infusion, followed by SL Edaravone 30mg tablet. The study showed that the SL tablet's plasma concentration-time profile was similar to that with the IV infusion. The bioavailability of the SL tablet of Edaravone was 91.94% compared with IV administration. However, the authors suggested that the exposure differences can likely be addressed by increasing the SL tablet's strength.[57] Additionally, an oral suspension of Edaravone recently entered phase 3 clinical trial (ClinicalTrials.gov Identifier: NCT04165824).

2.2.5 Emerging treatments for ALS

The development of a better understanding of the pathophysiological and genetic underpinnings of ALS over the last two decades has led to the development of new therapies that target the array of pathways involved in the development of this neurological disorder, including neuroinflammation, nucleocytoplasmic transport impairment, oxidative stress, DNA damage, mitochondrial dysfunction, and RNA processing defects. Several clinical studies designed to investigate the effect of the newly developed treatments on the molecular biomarkers involved in ALS were recently approved by FDA. Five such potential ALS therapies have recently reached phase 3 clinical trials: (1) Masitinib - tyrosine kinase inhibitor, (2) Tauroursodeoxycholic acid (TUDCA) - antiapoptotic, inhibition of caspase-3, (3) Ravulizumab - an inhibitor of complement C5, and (4) NurOwn - autologous bone marrow-derived mesenchymal stem cells.

2.2.5.1 Masitinib

Masitinib is an oral tyrosine kinase inhibitor that provides a neuroprotective effect by its immunomodulatory properties, targeting microglia, macrophage, and mast cell activity, in both central and peripheral nervous systems.[58] The drug was previously studied in ALS SOD1 G93A rat models and demonstrated promising preclinical results.[59,60]

The effectiveness and safety of Masitinib were recently assessed in a phase 2/3 clinical trial (ClinicalTrials.gov Identifier: NCT02588677). The study involved 394 patients with probable, a laboratory-supported probable, or definite diagnosis of ALS according

to the revised El Escorial criteria [61]. The patients also had less than three years since the symptom onset, had FVC of at least 60%, and were on a stable dose of Riluzole for at least 30 days before joining the study. The participants were randomly divided into three groups. The first group received a daily supplement of 100 mg/day Riluzole and 3.0 mg/kg/day Masitinib. The second group received a daily dose of 100 mg/day Riluzole and 4.5 mg/kg/day Masitinib. The third group received 100 mg/day of Riluzole plus a placebo. All participants were followed up for 48 weeks (11 months).

The combination treatment of Riluzole and Masitinib showed significant benefit over placebo. Patients on the combination therapy demonstrated a 27% slowing of ALSFRS-R deterioration, a 29% lower decline in the quality of life, and 22% less respiratory decline. A time-to-event analysis demonstrated that patients on Masitinib had a 25% delay in disease progression. The best results were reported in the patients at the less severe ALS stages, suggesting that even further improvement is possible if the drug is prescribed right after the diagnosis.[58]

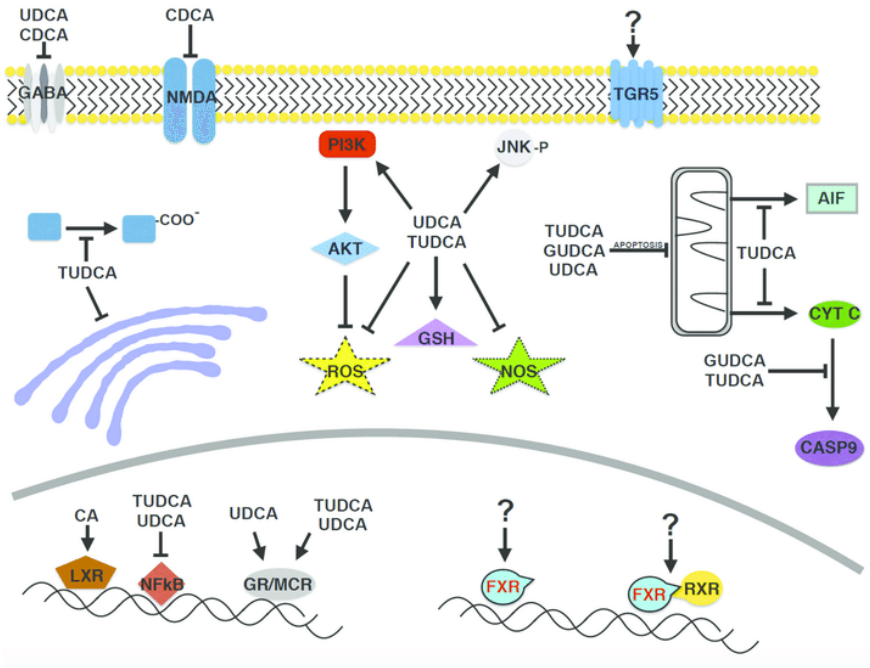
Recently, FDA approved a phase 3 randomized, double-blind, placebo-controlled trial with a planned enrollment of 495 patients (ClinicalTrials.gov Identifier: NCT03127267). This follow-up trial compares the effectiveness and safety of combinational treatment of Masitinib and Riluzole to Riluzole only.

Study participants will be randomized to a daily dose of the following: 100 mg/day Riluzole combined with Masitinib starting at 3.0 mg/kg and gradually increasing to 4.5 mg/kg after four weeks; Riluzole and 4.5mg/kg Masitinib, but after four weeks, the daily dose will be increased to 6 mg/kg; or a stable dose of 100 mg/day Riluzole and placebo. The estimated study completion date is September 2022.

2.2.5.2 Tauroursodeoxycholic acid (TUDCA)

Tauroursodeoxycholic acid (TUDCA) is a hydrophilic bile acid produced in the human liver by conjugation of ursodeoxycholic acid (UDCA) with taurine.[65] It was previously approved for the treatment of liver diseases. More recently, multiple animal models of neurodegenerative diseases demonstrated that TUDCA might have antioxidant, antiapoptotic, and neuroprotective properties in the central nervous system **(Figure 11)**.[66-69]

Figure 11. Molecular pathways implicated in the neuroprotective effects of bile acids in neurodegenerative disease models.



(Reproduced from Ackerman and Gerhard, 2016)

A phase 2 study (ClinicalTrials.gov Identifier: NCT00877604) evaluated the efficacy and tolerability of TUDCA as an add-on treatment in ALS patients taking Riluzole. Thirty-four eligible study candidates were diagnosed with limb-onset probable or definite ALS according to the revised El Escorial criteria, had symptoms for less than 18 months, had forced vital capacity of at least 75%, and were not ventilator dependent. The participants were randomized to placebo or 1 gram of TUDCA twice daily for 54 weeks following the three-month initial assessment period. The study determined that TUDCA is safe and well-tolerated and that the treatment significantly slows down ALS progression. The TUDCA-treated group had a significantly higher ALSFRS-R score than the placebo group at the end of the study.[70] There was no difference in adverse events between the treatment and placebo groups.

TUDCA was also evaluated in the sodium phenylbutyrate-taurursodiol (AMX0035) phase 2 clinical trial (ClinicalTrials.gov Identifier: NCT03127514) to assess the safety and efficacy of the drug for ALS treatment. AMX0035 is a proprietary combination of two drugs - sodium phenylbutyrate (PB) and tauroursodeoxycholic acid (TUDCA) targeting mitochondrial dysfunction and endoplasmic reticulum stress, the cellular pathways implicated in the development of ALS. The study recruited 137 qualifying ALS patients. Participants were randomized to take either 3 grams PB plus 1 gram of TUDCA twice daily or a placebo for 24 weeks.

The study showed that patients who received AMX0035 had slowed decline on the ALSFRS-R. While this new drug's benefit exceeded the benefits of Edaravone and Riluzole, no significant difference in muscle strength measures, vital lung capacity, survival, or need for tracheostomy was observed between AMX0035 and placebo groups.[71]

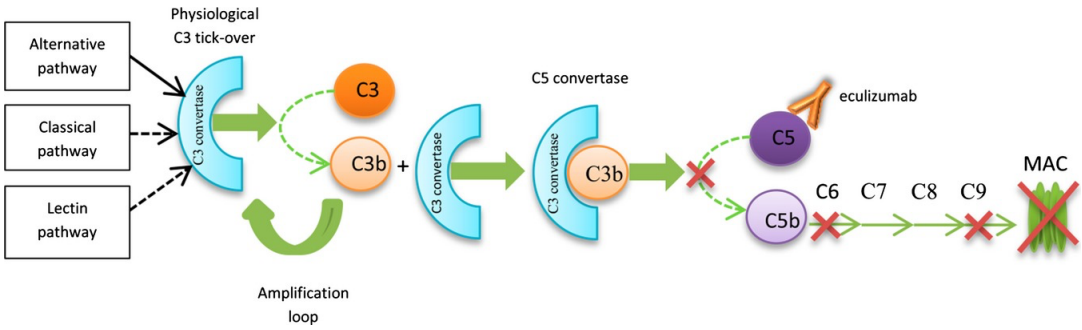
In 2019, TUDCA entered the phase 3 clinical trial (ClinicalTrials.gov Identifier: NCT03800524) to assess the drug's safety and efficacy as an add-on treatment Riluzole in patients affected by ALS. The study is enrolling 440 participants diagnosed with probable, probable laboratory-supported, or definite ALS, as defined by revised El Escorial criteria, with disease duration less than 18 months and FVC of at least 70%.

Following an initial screening period of 12 weeks, the participants will be randomized into two groups to receive TUDCA or a placebo. TUDCA (1g) will be administered orally twice daily for 18 months. The control group will receive a placebo pill on the same regimen. Additionally, all participants will be taking 100 mg Riluzole daily. The trial is expected to be complete in June 2021.

2.2.5.3 Ravulizumab

Ravulizumab (Ravulizumab-cwvz; Ultomiris) is a long-acting humanized monoclonal antibody, a relatively new type of "targeted" protein-based therapy. It is an enhanced second-generation complement C5 inhibitor created by Alexion Pharmaceuticals Incorporated. Ravulizumab mechanism of action is depicted in **Figure 12**.

Figure 12. Mechanism of action of Ravulizumab.



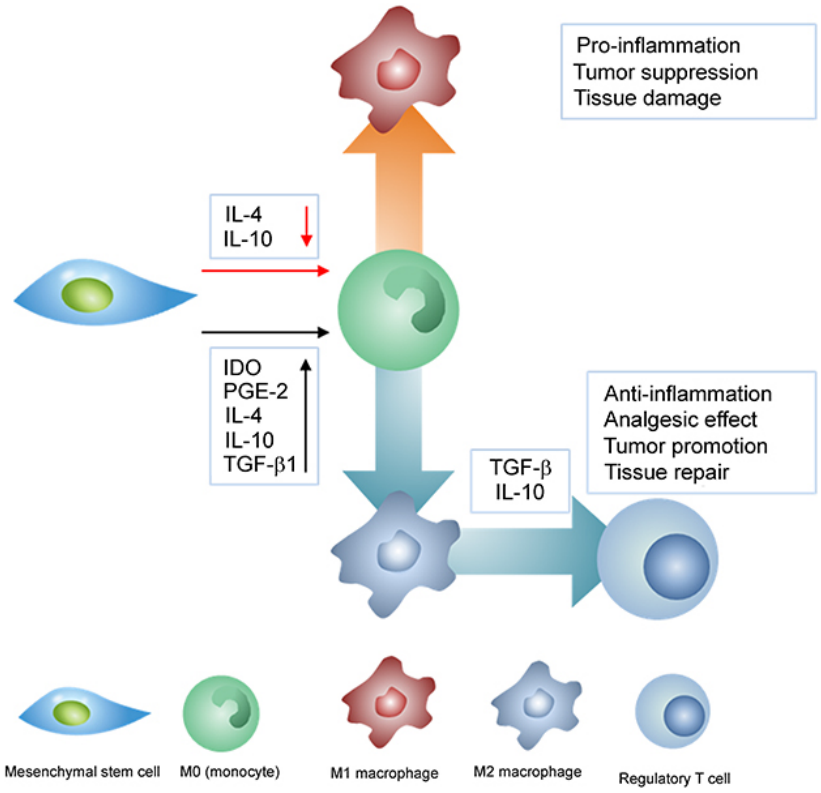
(Reproduced from Devalet et al., 2015)

Like the first-generation C5 inhibitor, eculizumab, which the same company created, Ravulizumab binds specifically to the complement protein C5, inhibiting the cleavage of C5 into C5a and C5b, thereby preventing the formation of the terminal complement complex C5b-9, that facilitates cell lysis.[72] The FDA approved Ravulizumab for the treatment of paroxysmal nocturnal hemoglobinuria and atypical hemolytic uraemic syndrome. The treatment is administered in a clinic by intravenous (IV) infusion. In March 2020, Ravulizumab entered a phase 3 clinical trial that aims to determine the efficacy and safety of this new therapy in ALS patients (ClinicalTrials.gov identifier: NCT04248465). The study plans to enroll 350 adults diagnosed with possible, laboratory-supported probable, probable, or definite ALS, with disease duration less than three years, a slow vital capacity (SVC) of > 65%, and no ventilator dependence. The participants will be randomly divided into two groups. Group 1 will receive a single dose of Ravulizumab via intravenous infusion, followed by regular maintenance dosing based on the patient's weight. Group 2 will receive a placebo. All participants will be observed for 50 weeks.[73] The study is estimated to be completed by October 2024.

2.2.5.4 NurOwn Mesenchymal stem cell - neurotrophic factor cells

NurOwn is autologous bone marrow-derived mesenchymal stem cells (MSC) induced to secrete neurotrophic factors (NTF). MSCs have been previously demonstrated to enhance neurogenesis, modulate neuroinflammation, and contribute to neuroprotection.[74] NurOwn® is delivered by a combination of intrathecal and intramuscular injections.

Figure 13. The anti-inflammatory mechanism of Mesenchymal stem cell in the central nervous system.



(Reproduced from Kwon and Yoo, 2019)

A phase 2 randomized controlled trial aimed to determine the safety and efficacy of NurOwn®. This study involved 48 patients diagnosed with possible, probable, laboratory-supported probable, or definite ALS according to El Escorial criteria and symptom duration of between 12 and 24 months. The participants were randomized into two groups to receive MSC-NTF cells transplants or placebo and were followed for six months after the treatment. Changes in ALSFRS-R scores and levels of neurotrophic factors and inflammatory biomarkers in participants' CSF were used to determine the treatment's efficacy and safety.

The majority of the MSC-NTF treated participants showed improvement in ALSFRS-R scores compared to the placebo group. This improvement was even more statistically significant in a subgroup of rapidly progressing ALS patients between 4- and 12-weeks post-treatment. Furthermore, MSC-NTF recipients had increased neuroprotective and anti-inflammatory biomarkers and significantly reduced inflammatory biomarkers' levels, suggesting a positive response to the new treatment. However, these values started declining towards the end of the follow-up period, suggesting that repeated transplantation of MSC-NTF cells is required to maintain the treatment's effectiveness.[75]

Following the phase 2 study's promising results, NurOwn entered the phase 3 clinical trial (ClinicalTrials.gov Identifier: NCT03280056) in the first quarter of 2020. The study aims to determine the safety and efficacy of 3 intrathecal injections administered at 2-month intervals. The estimated completion date for this study is December 2020.[76]

2.3 FRONTOTEMPORAL DEMENTIA

Frontotemporal dementia (FTD) is the second most common type of early-onset dementia after Alzheimer's disease.[77] It accounts for 10 to 15 percent of all dementia cases.[78] FTD is an adult-onset fatal neurological disorder that affects people between the ages of 45 and 65.[79] The average disease duration is eight years; however, it can range from 2 to 20 years.[80,81] FTD is characterized by progressive neuronal loss within the frontal and temporal cortices responsible for personality, behavioral functions, and language.

Postmortem pathology shows that FTD can result from different underlying pathologies, suggesting multiple pathogenic pathways are involved in the disease, resulting in overlapping phenotypes.[82] Genetic mutations have been found in a sizable number of clinical FTD subjects, including patients with no FTD family history. The genes and mutations will be described in Chapter 3 of this thesis.

In a recently published study, Blauwendraat et al. showed that genetic defects in various pathways could contribute to the pathogenesis of FTD, even in sporadic cases. These findings also suggest that clinical FTD results from genetic defects in several pathways due to the susceptibility of frontotemporal brain networks to insults in these pathways.[83]

Approximately 30% of FTD patients have a strong family history of the disease, and about one-fourth of patients inherited the disorder in autosomal dominant form.[84] Several genes have been linked to FTD, but MAPT protein-associated microtubules (tau), progranulin (GRN), and C9orf72 are the most common contributors to this fatal neurological disorder. The frequency of the mutations in each of these genes varies across different populations, but C9orf72 is the most common genetic cause of FTD

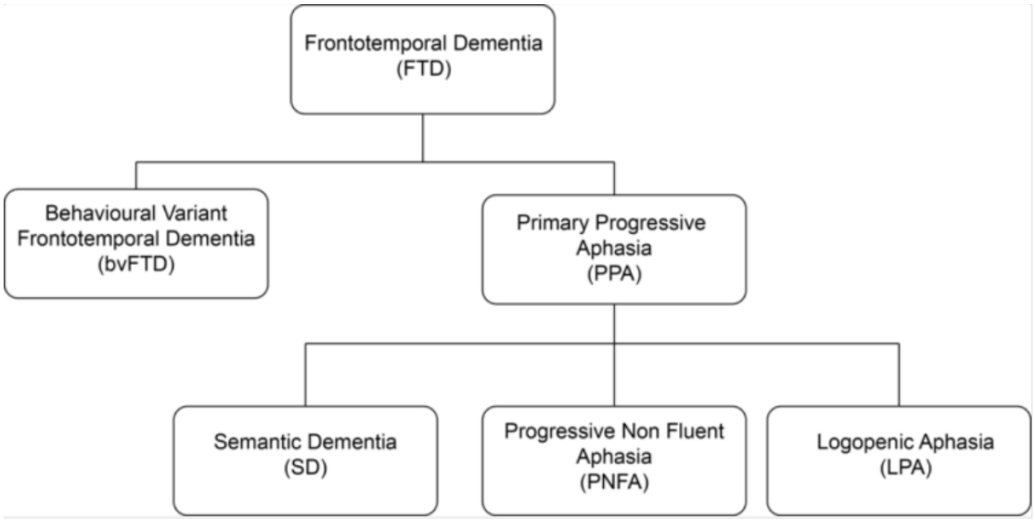
globally. Mutations on other genes are the rare cause of FTD. These genes include CHMP2B, KIF5A, CHCHD10, TBK1, TDP-43, and VCP. Other genes may not specifically cause FTD but may modify the symptoms or affect the age at symptom onset. An example of such a disease-modifying gene is transmembrane protein 106B (TMEM106B).[85]

Like ALS, there is no cure for this devastating neurological disease, but treatments such as stimulants, antipsychotics, and selective serotonin uptake inhibitors (SSRI) help alleviate symptoms.

2.3.1 Types of FTD

Two main types of FTD are recognized: behavioral type (bvFTD) and language type (PPA). PPA is further subdivided into three categories: progressive nonfluent aphasia (PNFA), semantic dementia (SD), and logopenic variant (LPA).[86] Which type of FTD a patient will develop depends on the affected areas in the frontal and temporal cortices.

Figure 14. Types of frontotemporal dementia.



(Reproduced from The Valerie Foundation)

2.3.1.1 Behavioral variant FTD

Behavioral variant (bvFTD) is the most common type of FTD. It comprises ~ 34% of all FTD cases. bvFTD affects the frontal and temporal lobes of the brain. Other affected areas include the insula and the cingulate and areas deeper in the brain, including the basal ganglia and thalamus.[87] bvFTD is characterized by executive dysfunction, changes in behavior such as disinhibition, apathy/inertia, loss of sympathy or empathy, perseverative stereotyped or ritualistic behavior, and changes in appetite and dietary preferences.[88]

2.3.1.2 Progressive nonfluent aphasia

Progressive nonfluent aphasia (PNFA) represents approximately 25% of FTD cases.[89] PNFA has involvement predominantly in the left superior temporal lobe, inferior frontal lobe, and insula. As the disease progresses, it affects other areas in the left frontal, lateral temporal, and anterior parietal lobes.[90]

PNFA is characterized by language deficits such as hesitant and effortful speech, stuttering, phonemic paraphasia, and agrammatism.[91] In later stages of the disease, speech production becomes increasingly difficult in many cases leading to depression and social withdrawal.[92] PNFA often evolves to progressive supranuclear palsy (PSP) and corticobasal syndrome (CBS). Patients develop parkinsonian symptoms such as difficulty moving, falls, and eye movement challenges.[93]

Approximately 60% of PNFA patients have abnormal accumulations of tau protein in the brain. This is especially true of patients who manifest both agrammatism and apraxia of speech. In addition, [32] PNFA tauopathies are distinct from the tau accumulates associated with Alzheimer's disease. The remaining ~40% of individuals affected by PNFA have either FTLD-TDP43 protein or an Alzheimer's pathology.

2.3.1.3 Semantic dementia

Semantic dementia (SD) is the second most common type of frontotemporal dementia. It encompasses approximately 30% of all FTD cases.[94] SD has variable age at onset, typically between 55 and 70 years.[95] The disease's duration varies from 2 to 15 years, with average patient survival being 7-8 years after the symptom onset.[96]

SD most prominently affects the anterior and inferior temporal lobes of the brain.[97] It is characterized by severe problems with understanding the meaning of words. While it is common for SD patients to be unable to name the objects, their speech usually remains unaffected.[98]

2.3.1.4 Logopenic variant PPA

Logopenic variant PPA (lvPPA) is the most recently characterized subtype of FTD.[98] Little is known about the lvPPA age of onset and survival. Like other PPA variants, the logopenic variant is considered an early-onset form of dementia. LvPPA is characterized by speech dysfunction, including spontaneous speech delays, phonological errors and paraphrasing, difficulty repeating, understanding, and finding words. However, grammar and understanding of word meaning remain relatively preserved in the early stages of the disease.[98]

MRI studies determined that lvPPA patients have subtle atrophy in the left hemisphere's core language regions, including the frontal, parietal, and temporal lobes.[99,100] As the disease progresses, these atrophies can spread to the hippocampus and affect patients' memory.[101]

Autopsies showed that Alzheimer's disease pathology most often causes lvPPA. Approximately 70% of lvPPA patients have amyloid plaques and tau neurofibrillary tangles in the brain.[102] Because of this, lvPPA is considered one of the possible early-onset presentations of Alzheimer's disease.[101] Other rare pathological discoveries in lvPPA include Lewy bodies, TDP-43 and tau.[103,104]

3 GENETICS OF AMYOTROPHIC LATERAL SCLEROSIS AND FRONTOTEMPORAL DEMENTIA

3.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

This chapter is a rewritten and an updated version of the manuscript previously published in Frontier in Neuroscience. Title: The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia.[105] I am the first author of the manuscript. I performed the literature search and wrote the manuscript.

Other authors on the original manuscript and their contributions: Pietro Fratta, Bryan Traynor and Ruth Chia revised the manuscript for important intellectual content.

3.2 BACKGROUND

Through a growing body of evidence in the literature, ALS and FTD are increasingly recognized as two entities representing a broad neurodegenerative disorder continuum. Clinical observations that both disorders may occur in the same family or person have prompted researchers to look for genetic overlap between the two diseases. Cross-sectional studies have demonstrated that roughly a half of ALS patients develop cognitive impairment, often with prominent frontal lobe features, while one-third of FTD patients develop motor dysfunction.[106,107] It is now clear that a significant genetic overlap exists between ALS and FTD and that SQSTM1, TDP-43, VCP, FUS, TBK1,

MATR3, GRN, MAPT, CHMP2B, CHCHD10, SOD1, and most significantly, C9orf72 are the main genetic players.[105] These genes and associated with them phenotypes are described in **Table 1**.

Table 1. Genes implicated in ALS and FTD.

Gene	Locus	Neurological Phenotypes	Pathway	Main localization
C9orf72	9p21.2	ALS, FTD, ALS/FTD, BD, AD, PD, Schizophrenia	Nucleocytoplasmic transport/splicing	Extracellular, nucleus, endosome, lysosome
TDP-43	1p36.22	ALS, FTD, ALS (flail arm variant), SNGP and chorea, MND	Nucleocytoplasmic transport/splicing	Nucleus
SQSTM1	5q35	ALS, FTD, AD, early-onset ALS/FTD, NADGP, PDB	Autophagy	Nucleus, cytosol, lysosome, endoplasmic reticulum, edosome
FUS	16p11.2	ALS, FTD, ALS (juvenile with BIs), ET, MND (lower), bvFTD?, PD?	Nucleocytoplasmic transport/splicing	Nucleus
SOD1	21q22.11	ALS, MND, lateral sclerosis, spastic tetraplegia and axial hypotonia	Oxidative stress	Extracellular, mitochondrion, peroxisome, nucleus , cytosol
GRN	17q21.31	FTLD, FTD with TDP-43 inclusions, NCL	Autophagy/inflammation	Plasma membrane, extracellular, endoplasmic reticulum, endosome, lysosome, golgi apparatus
VCP	9p13.3	ALS, IBMPFD and ALS, IBMPFD CMT2, HSP, DMRV, muscular dystrophy, AD? Autism?	Autophagy/mitochondreal function	Nucleus, endoplasmic reticulum, cytosol, extracellular, lysosome
MATR3	5q31.2	ALS, PLS, Distal myopathy 2, autosomal dominant distal myopathy 1	RNA metabolism and stabilization.	Nucleus
CHCHD10	22q11.23	ALS, ALS/FTD, Mitochondrial myopathy (autosomal dominant)	Mitochondrial dysfunction/ synaptic integrity	Mitochondrion, nucleus
TBK1	12q14.1	ALS, ALS/FTD, AD	Autophagy/inflammation	Nucleus, cytosol, edosome, mitochondrion
MAPT	17q21.31	FTD, PSP, PD, Pick disease of brain, supranuclear palsy (progressive atypical)	Microtubule stabilization, apoptosis	Plasma membrane, extracellular, sytoskeleton, mitochondrion, nucleus, cytosol
CHMP2B	3p11.2	ALS17, FTD3, PPA, Pick disease of brain	Autophagy	Endosome, sytosol, extracellular, nucleus, mitochondrion

(Updated from Abramzon et., al, 2020)

AD - Alzheimer disease, ALS - Amyotrophic lateral sclerosis, BD - Bipolar disorder, Bis- Basophilic inclusions, CMT2 - Charcot-Marie-Tooth disease (type 2), DMRV-Myopathy (rimmed vacuolar), ET - Essential tremor, FTD - Frontotemporal dementia, HSP- Hereditary spastic paraplegia, IBMPFD- Inclusion body myopathy with Paget disease and frontotemporal dementia, MND - Motor neuron disease, NADGP-Neurodegeneration, childhood-onset with ataxia dystonia & gaze palsy, NCL - Neuronal ceroid lipofuscinosis, PD - Parkinson disease, PDB - Paget's disease of bone, PLS – Primary lateral sclerosis, PPA – Primary progressive aphasia, PSP – Supranuclear palsy, SNGP - Supranuclear gaze palsy.

3.2.1 Chromosome 9 open reading frame 72 (C9orf72)

In 2011, a hexanucleotide repeat expansion (GGGGC)n within the C9orf72 gene on chromosome 9p21 was identified as a major genetic cause of ALS and FTD.[108,109] Since then, C9orf72 repeat expansions became the most critical cause of ALS, FTD, and ALS/FTD, accounting for approximately 11% of all ALS and 13% of all FTD cases. The discovery of this gene established that there is significantly greater genetic overlap between ALS and FTD than had been previously estimated. C9orf72 repeat expansions have also been described in other neurodegenerative disorders, including Huntington disease-like syndrome, Creutzfeldt–Jakob disease, corticobasal syndrome, supranuclear palsy, ataxia, and Parkinson's disease.[105,110-112]

Four possible mechanisms underlying C9orf72-related neurodegeneration have been described in the literature. These include (1) haploinsufficiency of C9orf72 protein [108,109,113,114], (2) RNA toxicity due to accumulation of RNA containing the GGGGCC repeat in the brain and spinal cord [108,109,115], (3) dipeptide repeat (DPR) protein toxicity arising from repeat-associated non-AUG translation occurring off the expansion [116,117]; and (4) disruption of the nucleocytoplasmic transport [118-120]. Although enough evidence exists to support all four mechanisms, it is unclear which one has the dominant role in

neurodegeneration. The possibility of multiple mechanisms acting in tandem or sequentially to cause neuronal death cannot be discounted.[105]

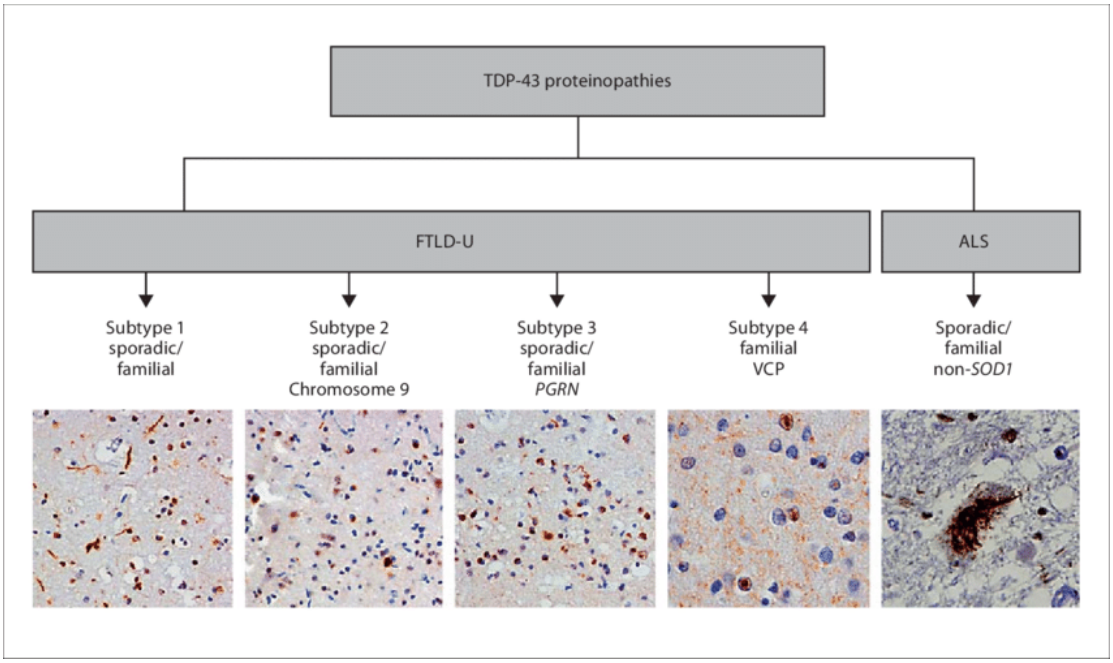
Several mouse models have been developed in an attempt to gain a better understanding of the pathological mechanism underlying C9orf72-related neurodegeneration. Though informative, these models fall short of determining the precise cause of neurodegeneration, as the available evidence supports all four models of neurodegeneration. For example, mice deficient in C9orf72 in neurons and glial cells do not display motor neuron degeneration or motor function defects associated with ALS.[121] BAC transgenic mice expressing between 100 and 1000 GGGGCC repeats formed RNA foci and dipeptide repeat proteins in the nervous system. However, there is no evidence of neurodegeneration or functional deficits.[122,123] Mice with more than 450 GGGGCC repeats exhibit moderate hippocampal neuronal loss and age-related anxiety and cognitive impairment.[105,124]

More recent mouse models demonstrated that ablation of C9orf72 in a gain-of-function C9ALS/FTD mouse model exacerbates motor behavior deficits in a dose-dependent manner.[114] Transgenic GFP-PR_{28} mice expressing arginine-rich poly(PR), the most toxic type of DPRs in neurons, partially developed neuropathological features of C9FTD/ALS.[125] Two other transgenic C9FTD/ALS mouse models demonstrated that poly(GR) affects translation and stress granule dynamics [126] and compromises mitochondrial function by binding Atp5a1 [127,105].

3.2.2 TAR DNA-binding protein 43 (TARDBP)

TDP-43 cytoplasmic and nuclear inclusions are characteristic of both ALS and FTD. Mutations in the TAR DNA-binding protein 43 gene were linked to ALS in 2008.[128] In ALS, TDP-43 cytoplasmic aggregates are typically found in neurons and glia of the primary motor cortex, brainstem motor nuclei, and spinal cord.[129,130] In FTD, TDP-43 inclusions are found in the neocortex and dentate granule cells of the hippocampus.[131,132] **Figure 15** demonstrates different types of TDP-43 proteinopathies in ALS and FTD.[105]

Figure 15. TDP-43 proteinopathies in ALS and FTD.



(Reproduced from Kwong et al., 2008)

Despite the high prevalence of TDP-43 inclusions in ALS and FTD brains, TDP-43 mutations are associated with less than 1% of all ALS cases and an even lower proportion of FTD cases.[105,133]

TDP-43 is a DNA and RNA binding protein involved in many aspects of RNA metabolism, including transcription, translation, microRNA biogenesis, and messenger RNA stabilization.[134-138] Two distinct mechanisms have been suggested to account for TDP-43-related neurodegeneration: (1) loss of function caused by sequestration of essential TDP-43 protein from the nucleus to the cytoplasm that result in the formation of cytoplasmic aggregates [139,140]; and (2) gain of function due to intrinsic toxic properties of the aggregates [141,142]. However, the toxicity of TDP-43 aggregates in ALS and FTD is still under debate. TDP-43 mutations have also been reported to negatively affect the liquid drop formation. The pathophysiological significance of this in vitro epiphenomenon, however, remains poorly understood.[105,143]

TDP-43 has a major role in repressing the inclusion of cryptic exons during splicing. Cryptic exons happen when a piece of an intron is mistakenly incorporated into the mRNA introducing frame shifts and premature termination or reduced RNA stability.[144,145]

A previous study that focused on evaluation of alternatively spliced stathmin-2 (STMN2) transcripts, as a proxy for TDP-43 pathology, demonstrated that truncated STMN2 accumulated in human induced pluripotent stem cell-derived neurons depleted of TDP-43, but not in cells with pathogenic TARDBP mutations in the absence of TDP-43 aggregation or loss of nuclear protein. The same study reported significant

associations of truncated STMN2 RNA with phosphorylated TDP-43 levels and an earlier age of disease onset in patients with FTLD-TDP.[146]

A more recent study established a direct link between variants in UNC13A, one of the strongest genetic risk factors for FTD/ALS, and loss of TDP-43 function by demonstrating that loss of TDP-43 from the nucleus in human brain, neuronal cell lines, and iPSC-derived motor neurons results in the inclusion of a cryptic exon in UNC13A mRNA and reduction in UNC13A protein expression.[147]

Another recent study that used the direct comparison of TDP-43-mediated transcription and alternative splicing in human muscle (C2C12) and neuronal (NSC34) mouse cells showed that TDP-43 displays a tissue-specific behavior targeting unique transcripts in each cell type due to cell-specific expression of RNA-binding proteins.[148]

More than fifteen mouse models have been created to unravel the pathogenic roles of TDP-43 in protein homeostasis, autophagy and clearance pathways implicated in ALS and FTD. These mouse models demonstrated that suppression of conditional TDP-43 transgene expression differentially affects early cognitive and social phenotypes in TDP-43 mice.[149] In a TDP-43$^{Q331K/Q331K}$ knock-in mouse model of ALS-FTD, TDP-43 gains function due to impaired autoregulation.[150] In TDP-43^{M337V} and TDP-43^{G298S} knock-in mice, mutant TDP-43 causes early-stage dose-dependent motor neuron degeneration.[151] Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of motor neuron degeneration.[152] Mouse models have also provided insight into how mutations in this gene may be underlying frontotemporal dementia. A recent TDP-43^{Q331K} mouse model manifested cognitive dysfunction in the absence of motor dysfunction. Pathological examination showed that normal

localization of TDP-43 within the cell, but there was evidence of perturbed regulation of TDP-43.[105,150,153]

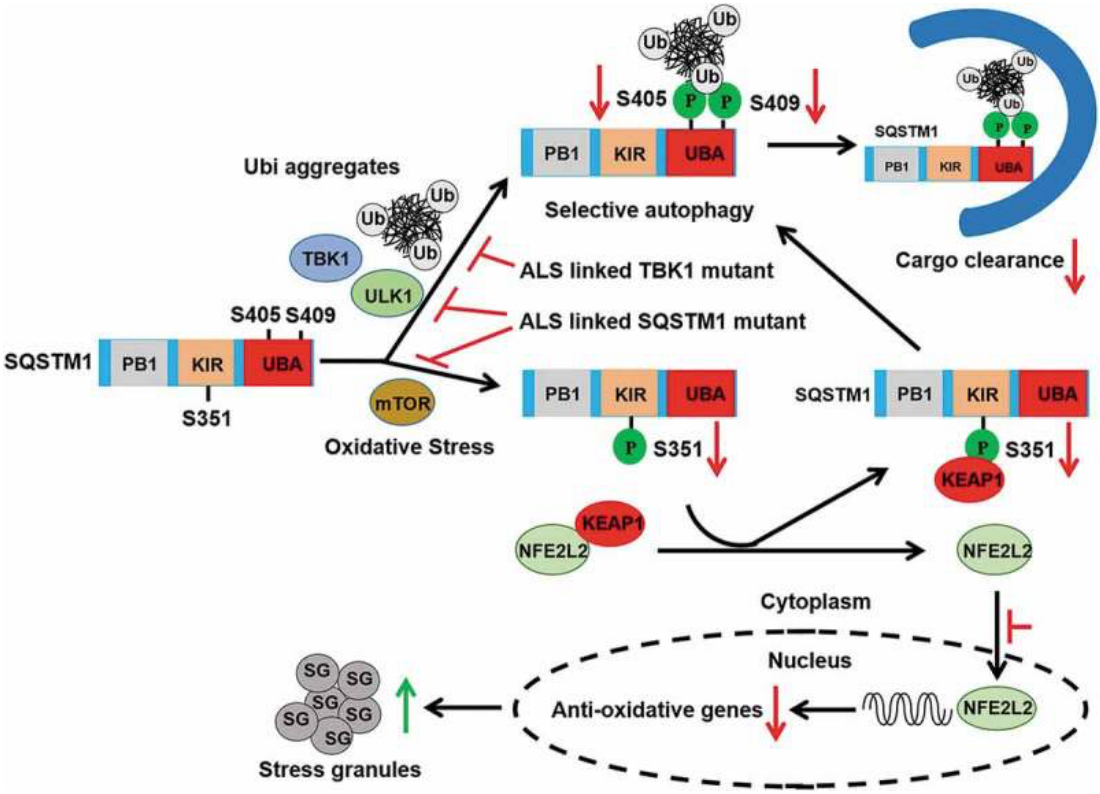
3.2.3 Sequestosome-1 (SQSTM1)

Sequestosome-1 (SQSTM1) mutations were first identified in patients with Paget's disease of bone.[154] In 2001, the gene was linked to ALS and behavioral variant FTD.[155] SQSTM1 encodes p62, a multifunctional protein involved in a wide range of cellular processes, including ubiquitin-mediated autophagy [156-158], NFKB1 signaling [159], transcription regulation [160] , and apoptosis [161]. p62 is a standard component of ubiquitin-containing inclusions in various neurological disorders, including ALS and FTD. Over 100 variants in SQSTM1 have been reported, accounting for approximately 1% of all ALS and 3% of all FTD cases. Mutated p62 protein is prone to forming aggregates. ALS patients with SQSTM1 variants have p62-positive inclusions in the motor neurons, while FTD patients have p62-positive inclusions in the hippocampus and cerebral neocortex.[105,162,163]

Accumulation of SQSTM1 comes from disturbances in the selective autophagy pathway.[164] The proposed mechanism of SQSTM1 accumulation is shown in

Figure 16.

Figure 16. SQSTM1/p62 mutations disrupt selective autophagy.



(Reproduced from Dang et al., 2019)

Similar to FUS and TDP-43, SQSTM1 undergoes liquid-liquid phase separation. Recent research indicates that cytoplasmic death-associated protein (DAXX) promotes SQSTM1/p62 phase condensation, a key step in triggering the Nrf2-mediated stress response. Polyubiquitin chain-induced p62 phase separation leads to the segregation of autophagic cargo.[105,165,166]

In the past decade, many mouse models have been created to demonstrate the connection between p62 and other ALS genes. Mitsui et al. showed SQSTM1 deficiency exacerbates disease phenotypes in $SOD1^{H46R}$ ALS mice.[167] Overexpression

of SQSTM1 leads to a significant increase in the biochemically observable insoluble SQSTM1 and polyubiquitinated protein in the spinal cord of SQSTM; $SOD1^{H46R}$ mice when compared to $SOD1^{H46R}$ mice. This finding indicates that the overexpression of p62 in $SOD1^{H46R}$ mice impairs the protein degradation pathway and thereby accelerates the onset of disease.[105,167]

From the FTD perspective, apart from developing mature-onset obesity due to impaired glucose tolerance and insulin resistance, p62 knockout mice exhibit accelerated aging phenotypes and have a significantly shorter lifespan. These mice develop cognitive dysfunction and anxiety, the symptoms characteristic of human Alzheimer's disease.[105,168]

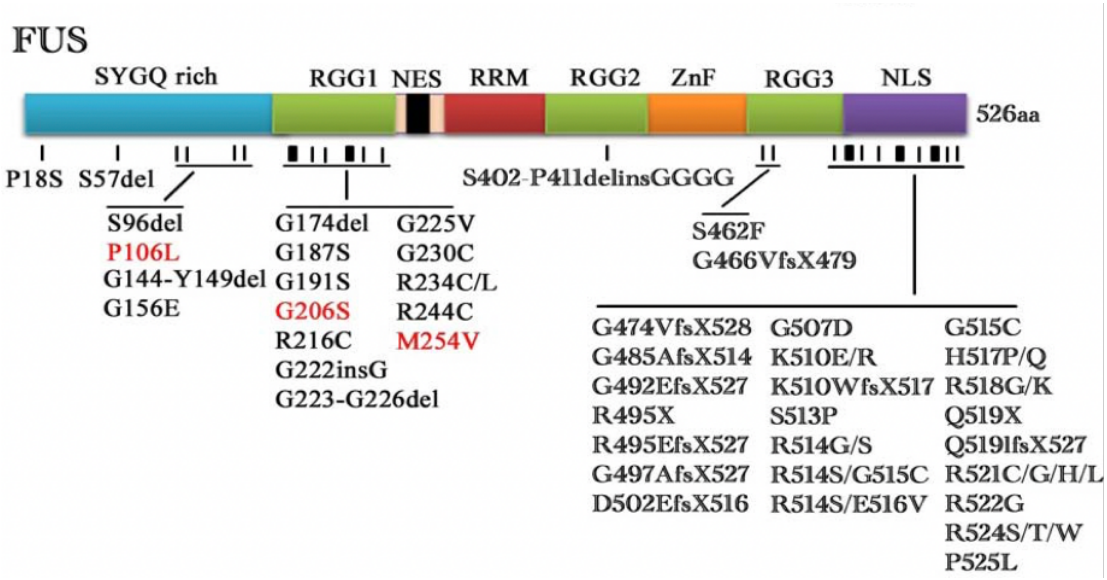
3.2.4 Fused in sarcoma (FUS)

Fused in sarcoma (FUS) is an RNA-binding protein that was associated with amyotrophic lateral sclerosis (ALS) in 2009.[169] FUS, like TDP-43, regulates various aspects of RNA metabolism, including alternative splicing, RNA translation, and transport.[169,170] FUS mutations account for approximately 1% of all ALS cases. They are also occasionally observed in FTD cases. Apart from these phenotypes, pathological FUS aggregates are found in a variety of other neurodegenerative disorders, including hereditary essential tremor, polyglutamine diseases, and Parkinson's disease.[105]

Mutations associated with ALS and FTD are located mainly in highly conserved regions of the gene and affect the nuclear localization signal (NLS) of the protein.[105] A

schematic overview of FUS protein domain structure and disease-associated mutations are shown in **Figure 17**.

Figure 17. Schematic overview of FUS protein domain structure and disease-associated mutations.



(Reproduced from Li et al., 2013)

As with TDP-43, mutations in the FUS gene are most commonly found in ALS patients. A small number of FUS mutations, including p.P106L, p.Gly174-Gly175 deletion GG, and p.M254V have been described in FTD patients without concurrent ALS.[105,26,171]

Two mechanisms have been proposed to explain how FUS mutations may induce neurodegeneration. In the loss-off-function mechanism, the depletion of nuclear FUS impairs transcription, alternative splicing, and DNA repair.[172] In a toxic gain-of-function mechanism, FUS forms cytoplasmic aggregates and spreads through neuronal tissues in a prion-like manner.[173] Both mechanisms have sufficient evidence to support

them, and different mechanisms can underpin different FUS mutations.[174,175] FUS-related neurodegeneration has recently been linked to liquid-liquid phase separation (LLPS) that is regulated by universal cellular actors such as ATP and nucleic acids through enhancement and dissolution.[176] Other recent FUS studies provided additional information on LLPS functions, mechanism, and transformation.[105,176-179]

Several mouse models have been developed through the years in an effort to determine FUS's pathogenic role in neurodegeneration. FUS knockout mice exhibit behavioral abnormalities such as hyperactivity and reduced anxiety-related behavior. However, they do not develop motor neuron dysfunction, implying that ablation of the FUS gene alone is insufficient to cause ALS.[180] Transgenic ΔNLS-FUS mice that overexpress exogenous FUS with deleted nuclear localization signal under Thy1 neuron-specific promoter develop progressive ALS phenotypes associated with the formation of ubiquitin/p62-positive FUS aggregates, neuronal loss, and gliosis. The truncation of the NLS region in $FUS^{\Delta NLS/\Delta NLS}$ mice results in the mislocalization of FUS protein from the nucleus to the cytoplasm in spinal motor neurons and cortical neurons, resulting in apoptosis [181] Furthermore, in heterozygosity, both $Fus^{\Delta NLS/+}$ mice and knock-in mice bearing another C-terminal frameshift mutation ($Fus^{\Delta 14/+}$) develop progressive motor neuron dysfunction, mimicking the early stages of ALS.[181,182] Recent FUS p.R521C knock-in mouse demonstrated that, in disease-prone neurons, stress causes mislocalization of mutant FUS into stress granules and upregulation of ubiquitin.[105,183]

3.2.5 Cu/Zn superoxide dismutase (SOD1)

SOD1 is a cytosolic protein found in the nucleus, peroxisomes, and mitochondrial intermembrane space. The SOD1 protein is encoded by the SOD1 gene, located on chromosome 21. SOD1 binds to copper and zinc ions and is one of two isozymes responsible for destroying free superoxide radicals. The SOD1 isozyme is a soluble cytoplasmic protein that acts as a homodimer to convert harmful superoxide radicals to molecular oxygen and hydrogen peroxide.[105,184]

Studies on SOD1 aggregation and toxicity demonstrated that underlying molecular mechanisms of SOD1 toxicity are different in familial and sporadic ALS. In familial ALS, SOD1 mutations gain toxicity by inducing toxic changes in protein conformation, leading to formation and accumulation of toxic SOD1 aggregates in mitochondria. This, in turn, leads to the disruption of many cellular processes, including free radical scavenging, axonal transport, mRNA splicing, mitochondrial function, and protein quality control and degradation.[185-188] Previous studies have shown that Derlin-binding region (DBR) within N-terminus of mutant SOD1 binds with high specificity to the cytosolic carboxyl-terminal region of Derilin-1, an ERAD-linked protein, subsequent ER stress–induced activation of the ASK1 pathway and ultimately apoptosis.[189] Mutant SOD1 monomers cause protein toxicity by reducing proteasome and chaperone activity and allowing aberrant protein-protein interactions.[187] Furthermore, experimental assessment of SOD1 surface hydrophobicity using a sensitive fluorescent-based assay revealed that diverse ALS-causing mutations provoke SOD1 aggregation by increasing their propensity to expose hydrophobic surfaces.[190]

In sporadic ALS the process is thought to be different. Wild-type SOD1 gains toxic function as a result of the protein's conformational changes that are induced by aberrant posttranslational mechanisms such as demetallation and overoxidation.[191-193] Oxidized wild type SOD1 was reported to inhibit axonal transport in a manner similar to mutant SOD1.[193] Misfolded "mutant-like" SOD1 protein has been observed in postmortem spinal cord samples from sporadic ALS patients.[194] Furthermore, a common conformational epitope was detected in oxidized wild type SOD1 and in familial ALS-linked mutant SOD1, but not in normal wild-type SOD1.[193]

Seeded aggregation and cell-to-cell transmission are two additional well-known toxic properties of mutant SOD1 protein. Both mutant and wild-type SOD1 proteins were shown to seed the aggregation of the wild-type protein.[195] Exosomal transmission has been demonstrated for both mutant and wild-type SOD1 aggregates, while mutated SOD1 can otherwise be released in neurosecretory vesicles through a chromogranin A- and B-mediated pathway.[195,196] Aggregates may then enter cells via micropinocytosis.[197] Furthermore, recent data indicate that SOD1 aggregates may induce disease spreading within the CNS through a prion-like transmission mechanism.[198,199]

Increased SOD1 aggregation is associated with a more aggressive disease progression and shorter survival.[200] The two most common SOD1 mutations are p.A4V in North America and p.D90A in Europe.[201] The p.A4V mutation is associated with an aggressive form of the disease with an average survival of only one year after the onset of symptoms.[202] In contrast, patients with the recessive p.D90A mutation have a slower disease progression and can survive for more than ten years [203], and tend to develop deficits in executive function at the later stages of their disease.[105,204]

Other SOD1 mutations have been associated with a behavioral FTD phenotype in various studies. Nakamura and colleagues reported a family with the p.G141X SOD1 mutation in multiple generations, in which a proband with behavioral FTD developed motor neuron disease. Pathological examination revealed motor neuron and corticospinal tract pathology associated with ALS, and neocortical and limbic system degeneration consistent with atypical FTD.[205] In another study, most affected members in a family carrying the pathogenic p.I113T SOD1 mutation died from ALS, but one member presented with behavioral FTD before developing signs of motor dysfunction.[105,206]

3.2.6 Progranulin (GRN)

FTD is familial in 25-50% of all cases and can occur as an autosomal dominant disorder with high penetrance. Genes causing familial FTD include the gene that encodes progranulin (GRN). It is unclear whether some rare FTD-related GRN variants are pathogenic and whether these mutations can also cause neurodegenerative disorders other than FTD. However, in a study by Yu et al., several variants were identified as clearly pathogenic for the causation of neurodegenerative disease, although the mechanisms by which these lead to neurodegeneration in FTD remains unclear. Pathogenic GRN variants result in reduction or loss of functional protein and are thought to result in neurodegeneration through a haploinsufficiency mechanism.[207] The two most common GRN mutations are p.R493X and p.A9D. Patients carrying the p.A9D mutation develop FTD at an earlier age (51 vs. 58 years), have shorter survival, and frequently develop Parkinson's disease.[208] Unlike other mutations in GRN, p.A9D

does not result in mRNA haploinsufficiency, suggesting the possibility of additional pathogenic mechanisms.[105,208]

Mutations in GRN were first discovered as a cause of tau-negative, ubiquitin-positive FTD in 2006.[209] GRN is a multifunctional growth factor expressed in the cerebral cortex, hippocampus, and cerebellum. On the cellular level, GRN is involved in lysosomal function and organization. More than 160 mutations had been described in GRN. These mutations are responsible for about 7.5% of all FTD cases. GRN has also been linked to Alzheimer's disease, progressive supranuclear palsy, psychiatric disorders, and corticobasal degeneration.[210] The diagnosis of FTD can be difficult because of its gradual onset and the fact that it can also be misdiagnosed as Alzheimer's disease.[105,211,212]

Progranulin deficient mice did not develop ALS-associated motor neuron deficiencies. Instead, they developed behavioral FTD-like symptoms such as depression, disinhibition, impaired learning, and social recognition deficits. Pathologically, these mice also develop an augmented age-dependent activation of microglia and astrocytes and an age-dependent increase in hippocampal ubiquitin and phosphorylated TDP-43, supporting the idea that a loss-of-function mechanism underlies GRN-related pathophysiology.[105,213]

In humans, mild cognitive impairment (MCI) is indicative of a prodromal phase of dementia, a concept that also can be used for FTD. Due to GRN null mutations, the gene is a known diagnostic biomarker for FTD. In one study, electroencephalography (EEG) rhythms were used to detect changes in brain oscillatory activity affected by GRN mutations, and MCI-FTD patients showed a significantly lower spectral power in

both alpha and theta oscillations compared to those with overt FTD. The authors concluded that EEG frequency rhythms are sensitive to different stages of FTD and can detect changes in brain oscillatory activity affected by GRN mutations. [105,214]

In 2018, Alector Inc. developed a monoclonal antibody drug AL001 that increased progranulin levels through the targeted manipulation of the sortilin-progranulin axis in patients diagnosed with FTD due to heterozygous mutations in the progranulin gene (FTD-GRN). The drug demonstrated promising outcomes in Phase II clinical trials (ClinicalTrials.gov Identifier: NCT03987295) by sustainably restoring plasma and cerebrospinal fluid progranulin levels to the normal range in all participants. In addition, the treatment decreased the plasma neurofilament light chain (NfL) levels of most symptomatic FTD-GRN participants by 14% from baseline. AL001 is currently undergoing Phase III clinical trials (ClinicalTrials.gov Identifier: NCT04374136) to evaluate the efficacy and safety of the drug. The study estimated completion date is December 30, 2023.

3.2.7 Valosin containing protein (VCP)

Mutations in the Valosin-containing protein (VCP) were first identified in 2004 as the cause of a clinical syndrome characterized by the triad of inclusion body myopathy, Paget's disease of bone, and FTD (IBMFTD).[215] VCP mutations were later linked to ALS, demonstrating how genetic mutations in a single gene can cause both ALS and FTD.[216] To date, 72 autosomal dominant mutations have been discovered in this gene. Of these, more than 30 are reported in ALS or FTD cases (including behavioral FTD, semantic dementia, and progressive non-fluent aphasia).[217-219] Many of the reported

VCP mutations are located on exon five within the N-terminal CDC48 domain, which is involved in ubiquitin-binding, meaning that mutations in this region may negatively affect the ubiquitin protein degradation pathway.[105,220,221]

A recent study by Al-Obeidi et al. showed that ~9% of patients with VCP mutations had an ALS phenotype, 4% were diagnosed with Parkinson's disease, and 2% were diagnosed with Alzheimer's disease.[219] As of today, no definite correlation between the mutation type and the incidence of clinical features associated with VCP has been established.[105,219,222]

VCP encodes a member of the AAA-ATPase enzyme family with wide-ranging functions in cell division [223], DNA repair, ubiquitin-dependent protein degradation, and apoptosis suppression.[223] Ludtmann et al. provided evidence that mutations in VCP lead to mitochondrial uncoupling due to a reduced ADP/ATP translocation by adenine nucleotide translocase.[224] Such deficiency in mitochondrial bioenergetics makes neurons especially vulnerable as they require more energy than other cell types.[105,224]

Recent mouse models of VCP showed that activation of the NLRP3 inflammasome is associated with VCP protein myopathy. Nalbandian et al. [225] reported a significant increase in the expression of NLRP3, Caspase 1, IL-1β, and IL-18 in the quadriceps of 12 and 24 months old $VCP^{R155H/+}$ heterozygous mice. Furthermore, a significant increase of $IL\text{-}1\beta^{(+)}F4/80^{(+)}Ly6C^{(+)}$ macrophages in the quadriceps and bones of the same mice were also observed and was positively correlated with high expression levels of TDP-43 and p62/SQSTM1 markers of VCP pathology and progressive muscle wasting.[105,225]

Another recent discovery showed that VCP plays a vital role in maintaining lysosomal homeostasis and TFEB activity in differentiated skeletal muscle.[226] Arhzaouy et al. showed that selective inactivation of VCP in skeletal muscles of Myl1p-cre-vcp$^{-/-}$mice results in a necrotic myopathy with increased macroautophagic/autophagic proteins and damaged lysosomes.[226] It was further demonstrated that the myofiber necrosis was preceded by the upregulation of LGALS3/Galectin-3, a marker of damaged lysosomes, and TFEB activation, suggesting early defects in the lysosomal system.[105,226]

3.2.8 Matrin 3 (MATR3)

Matrin 3 (MATR3) is an RNA/DNA binding protein that interacts with TDP-43, a disease protein linked to ALS and FTD. The protein encoded by this gene is localized in the nuclear matrix. MATR3 is a multifunctional protein involved in transcription, and it interacts with other nuclear matrix proteins to form the internal fibrogranular network.[105,227]

MATR3, previously associated with distal myopathy and bulbar dysfunction, was first linked to ALS in 2014, using exome sequencing of several Caucasian families with several members affected by ALS and dementia.[227] Fifteen variants have been described in MATR3, cumulatively accounting for less than 1% of all ALS cases; there are no reports of patients with MATR3 mutations diagnosed with FTD. A 2017 study by Marangi et al. showed that the spectrum of diseases associated with MATR3 variants includes ALS, FTD, and distal myopathy with rimmed vacuoles.[105,228]

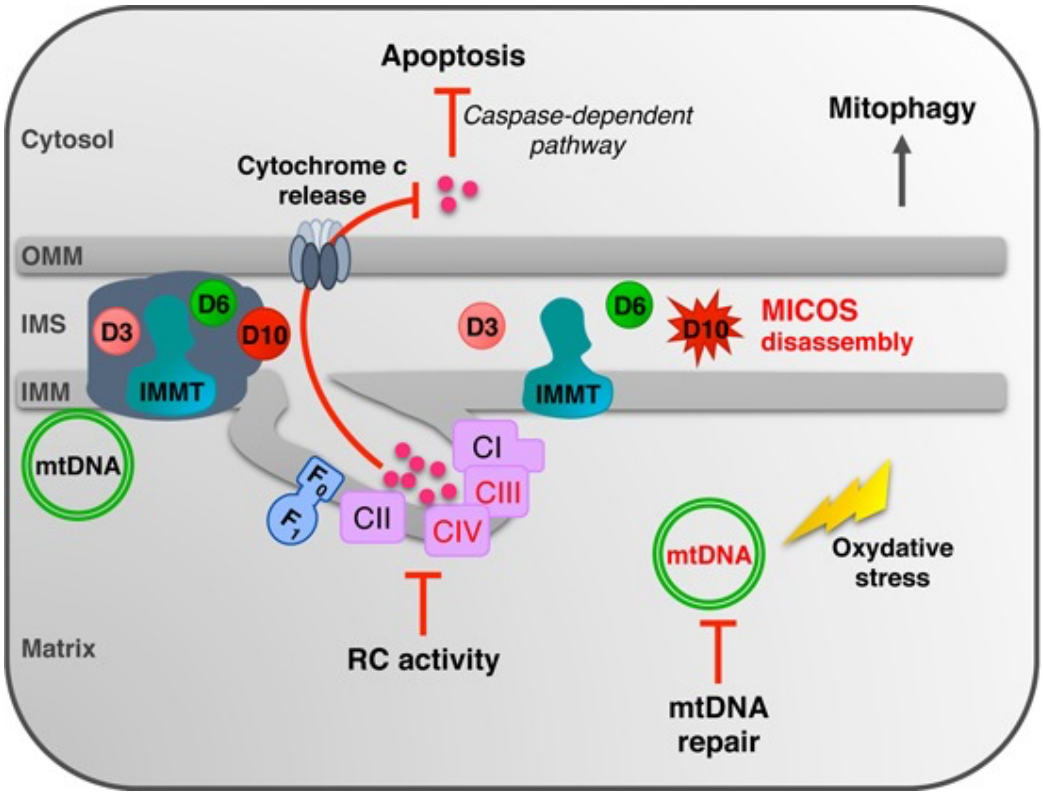
A study in mice to characterize regular MATR3 expression revealed significant variation in protein levels in different tissues. Spinal cord and muscle tissue had the lowest MATR3 levels, suggesting that these may be particularly vulnerable to protein function alterations.[229]

3.2.9 Coiled-coil-helix-coiled-coil-helix domain containing 10 (CHCHD10)

Coiled-coil-helix-coiled-coil-helix domain-containing protein 10 (CHCHD10) is a mitochondrial protein associated with ALS and FTD (including bvFTD and PPA).[230,231] It was discovered in 2014 through the exome sequencing of a large French family in which several family members were affected by autosomal dominant FTD with or without ALS, cerebellar ataxia, or mitochondrial myopathy.[232] Since then, more than 30 CHCHD10 mutations have been described in the literature. The majority of which occur in exon two of the gene encoding the non-structured N-terminus.[105,233-235]

CHCHD10 is a multifunctional protein involved in regulating mitochondrial metabolism, the synthesis of respiratory chain components, and modulation of cell apoptosis.[233] Perhaps not surprisingly, mutations in CHCHD10 lead to disassembly of the mitochondrial contact site (MICOS) complex, severe mitochondrial DNA repair deficiency after oxidative stress, disruption of oxygen consumption and ATP synthesis in cells, and disturbance of apoptotic mechanisms.[105,233] The pathogenic mechanism of CHCHD10 mutations is depicted in **Figure 18**.

Figure 18. The pathogenic mechanism of CHCHD10 mutations.



(Reproduced from Genin et al., 2016)

Recent data shows enrichment of CHCHD10 expression at the postsynaptic membrane of neuromuscular junctions.[233,236,237] Deletion of CHCHD10 in skeletal muscle of HSA-CHCHD10-/- knockout mice results in motor defects and neurotransmission impairment, indicating that muscle CHCHD10 is required for normal neurotransmission between motoneurons and skeletal muscle fibers.[236] Additionally, electron microscopy analysis of mitochondria from HSA-CHCHD10$^{-/-}$ mice revealed a large quantity of lysosome-like vesicles, indicating active mitochondria degradation and suggesting that CHCHD10 is necessary for mitochondrial structure and ATP production.[105,236,238]

Two groups independently developed CHCHD10^{S55L} knock-in mice, representative of the human CHCHD10 S59L mutation, and found that these mice developed progressive motor deficits, myopathy, cardiomyopathy, and died prematurely.[239,240] Histological examination revealed that CHCHD10 and its twin CHCHD2 form aggregates resulting in abnormal organelle morphology and function. In contrast, knock out CHCHD10 mice with a single adenine nucleotide insertion in exon two, resulting in a prematurely terminated protein, did not develop similar pathology, implying that tissue-specific toxic gain-of-function is the likely mechanism underlying CHCHD10 S59L-related neurodegeneration.[105,239]

3.2.10 TANK-binding kinase 1 (TBK1)

TANK-binding kinase 1 (TBK1) gene was discovered in 2015 through the whole-exome sequencing analysis of a large case-control cohort.[241,242] In 2016, a large genome-wide association study (GWAS) also identified the TBK1 gene on chromosome 12q14.2 as a risk locus for ALS, thus confirming the gene's association with motor neuron degeneration. TBK1 is a member of the IkB kinase family involved in autophagy, mitophagy, and innate immune signaling.[243] The protein is highly expressed in neuronal cells of the cerebral cortex, hippocampus, and lateral ventricle.[244] It also interacts with other genes implicated in ALS, such as OPTN and SQSTM1, to form the TBK1 autophagic adaptor complex.[105,245-247]

Over 90 mutations in TBK1 have been identified to date. According to a recent meta-analysis study, TBK1 loss of function and missense mutations account for 1.0% and 1.8% in ALS/FTD patients, respectively.[248] The majority of TBK1 mutations are loss

of function that results in the deletion of the C-terminal domain responsible for interaction with adaptor proteins that regulate the cellular distribution of TBK1 and activation of downstream signaling pathways.[247] Indeed, mutations tend to result in a significant decrease in TBK1 mRNA and protein expression.[105,242]

TBK1 mutations are associated with bulbar onset ALS and fast progressing behavioral FTD.[242] In ALS patients, TBK1 mutations are pathologically characterized by TDP-43 positive and p62 positive inclusions in motor neurons, as well as TDP-43 inclusions in the cortex. Similar to that observed in ALS, FTD patients harboring TBK1 mutations also carried TDP-43 inclusions in numerous brain regions and cytoplasmic p62 and ubiquitin-positive inclusions in glial cells.[105,249]

Compelling evidence exists that loss-of-function is the pathological mechanism behind TBK1-related ALS and FTD.[248,250,251] Germline deletion of TBK1 is lethal in embryonic mice suggesting that the protein plays a critical role in developmental homeostasis.[252] Recent rodent models have shown that conditional Tbk1 knockout in $Tbk1^{fl/fl}$ Nestin-Cre mice results in cognitive and motor dysfunction similar to ALS/FTD. Neuron-specific Tbk1 deletion induces morphological and biochemical alterations in neurons and glia, such as abnormal dendrites, neurofibrillary tangles, reduced dendritic spine density, as well as cortical synapse loss. Furthermore, Tbk1 knockout impairs autophagy in motor neuron-like cells, while Tbk1 over-expression extends ALS transgenic mice's survival.[105,253]

TBK1 is a central regulator of selective autophagy and inflammatory responses via IFN type I signaling.[254,255] Heterozygous deletion of the α-IFN receptor *Ifnar1* significantly prolongs the life span of $SOD1^{G93A}$ ALS mice.[256] In a 2019 study, Brenner et al. further

elucidated the connection between TBK1 and SOD1 in the mouse models.[257] The group showed that at the early stage, heterozygous Tbk1 deletion impairs autophagy in motoneurons and prepones the clinical onset and muscular denervation in $SOD1^{G93A}/Tbk1^{+/-}$ mice. In contrast, it dramatically reduces microglial neuroinflammation, slows disease progression, and prolongs mouse survival in the late stage of the disease.[105,257]

3.2.11 Microtubule-Associated Protein Tau (MAPT)

The MAPT was linked to neurodegeneration in 1998 by identifying tau mutations with neuronal and glial inclusions in 13 families diagnosed with FTDP-17.[258,259] The *MAPT* is located on chromosome 17 (17q21.31) and encodes the microtubule-associated protein tau, which has a primary role in the assembly and stability of microtubules.[260,261] It consists of 16 exons that are numbered from 0 to 14. Alternative splicing results in six tau isoforms that can be differentiated by the presence of zero, one or two N-terminal inserts, and the presence of either three (3R) or four (4R) microtubule-binding repeats in the C-terminal part of tau.[262]

To date, more than 55 MAPT pathogenic mutations have been described in the literature.[263] Most of these pathogenic mutations are clustered in exons 9-13.[264] MAPT mutations can be divided into two groups depending on the pathogenic mechanisms involved.[265,266] Group one is composed of missense mutations and deletions. These mutations modify tau protein and its function, leading to increased or decreased protein interaction with microtubules. Some of these mutations may increase protein's affinity to form tau filaments, leading to toxic aggregates. The second group of mutations is

clustered in exon 10 or its flanking regions. It interferes with the alternative splicing of exon 10, shifting from the approximately equivalent 3R:4R tau ratio to increased 4R tau, leading to an increase in filamentous inclusions and is associated with neurodegeneration.[265,267] Amongst other tau- related neurological disorders, *MAPT* mutations are responsible for 5-20% familial FTD and 0-3% sporadic FTD cases.[268-272] *These* mutations cause disease with autosomal dominant inheritance and have greater than 95% penetrance.[266]

Remarkable advances have been made in understanding the divergent physiological properties of tau in frontotemporal dementia. Comparative gene expression analysis between young and old human P301S transgenic TAU58/2 mice revealed distinct and shared pathway clusters in both young and old TAU58/2 mice, indicating that different molecular mechanisms are involved during disease progression.[273] A study by Evans et al. identified impaired ribosomal protein synthesis as a potential novel pathomechanism of tau by demonstrating decreased protein synthesis in neurons of two complementary transgenic mouse strains, K3 mice expressing K369I mutant tau and rTg4510 mice expressing P301L mutant tau.[274] Keller et al. demonstrated that CNS-targeted expression of wild type and P301L mutant tau results in robust tau hyperphosphorylation without tangle pathology, gradual development of age-progressive memory deficits. In contrast, the S320F variant, especially in combination with P301L, results in AD-type tangle pathology, focal neuroinflammation and memory impairment on an accelerated time scale. The group also showed that overexpression of wild-type tau and an FTD-associated tau variant could lead to cognitive deficits even in the absence of tangles.[275]

3.2.12 Charged Multivesicular Body Protein 2B (CHMP2B)

CHMP2B gene was linked to frontotemporal dementia in 2005 by positional cloning of a candidate gene region identified for frontotemporal dementia linked to a 15.5-Mb region on chromosome 3.[276] In 2006, mutations in this gene were identified in patients diagnosed with ALS and ALS-FTD hence, confirming the gene's link to neurodegeneration.[277] CHMP2B encodes a component of the heteromeric Endosomal Sorting Complex Required for Transport III (ESCRT-III) complex involved in the degradation and recycling of cell surface receptors.[278] The protein is expressed in all the major parts of the brain, especially in the hippocampus, cerebellum, frontal and temporal lobes.[276]

So far, 18 CHMP2B mutations have been described in the literature. Of those, 12 are linked to the ALS-FTD spectrum. Several pathogenic processes have been suggested for CHMP2B mutations: abnormal dendritic spine morphology [279], misregulation of transmembrane receptors, downregulation of a brain-specific microRNA, disruption of endo-lysosomal trafficking, and abnormal substrate degradation.[280] However, the exact mechanisms leading to disease pathogenesis and progression remain poorly understood.

Over the past decade, several mouse models have been created to better understand the diverse phenotypes and develop therapeutic approaches for the CHMP2B induced ALS and FTD. Vernay and colleagues reported that transgenic mice expressing CHMP2B^{intron5} mutation under the Thy1.2 promoter develop progressive age-dependent motor phenotype associated with behavioral changes that recapitulate ALS

and FTD.[281] A more recent study by the same team showed that neuronal overexpression of CHMP2B^{intron5} triggers an initial inflammation with disruption of lipid metabolism before the onset of motor symptoms. Furthermore, at the symptomatic stage of the disease, CHMP2B^{intron5} and SOD1^{G86R} mice share 28 common deregulated genes, known to be altered in neurodegenerative processes and particularly in ALS.[282]

Another mouse model developed by Ghazi-Noori and colleagues revealed that transgenic mice expressing CHMP2B^{Intron5} under the control of the hamster prion promoter had decreased survival and progressive neurodegenerative changes such as gliosis and age-dependent accumulation of p62- and ubiquitin-positive, but TDP-43 and FUS negative inclusions as in the FTLD-UPS pathology in patients with CHMP2B mutation. These changes were not observed in transgenic CHMP2B$^{wild\text{-}type}$ or *Chmp2b−/−* knockout mice, suggesting that CHMP2B mutations likely induce neurodegenerative changes through a gain-of-function mechanism.[283]

3.3 DISCUSSION

After three decades of research, it is evident that there is a significant genetic overlap between ALS and FTD. Mutations in C9orf72, SQSM1, VCP, CHMP2B, CHCHD10, and TBK1 mutations are most closely linked to both neurological disorders. Clinically, the ALS phenotype is most often associated with the behavioral variant FTD, while the overlap with other FTD subtypes that include language is less common. The pathophysiology underlying this observation remains poorly understood. Nevertheless, this overlap is incomplete: SOD1, TDP-43, and FUS variants are most commonly

associated with ALS and are only present in a small number of FTD patients. Similarly, MAPT and GRN are associated with FTD but not ALS.[105]

It is remarkable how the same pathways are implicated repeatedly in ALS and FTD. Both disorders are characterized by defects in RNA processing, protein clearance through autophagy, vesicle trafficking, mitochondrial dysfunction, and disrupted protein homeostasis. The genes described in this review are the key players in these pathways. TDP-43 and FUS multiple functions in mRNA processing; SQSTM1, C9orf72, VCP, CHMP2B, and TBK1 participate in autophagy and vesicle dynamics; and TDP-43, FUS, and SQSTM1 are standard components of nuclear and cytoplasmic inclusions.[284] Since there is significant genetic overlap between ALS and FTD, it is reasonable to look in FTD cases for mutations in ALS genes and vice-versa.[105]

The C9orf72 repeat expansion results in a wide variety of inter- and intra-familial phenotypes, including marked differences in the age at disease onset, the site of symptom onset, the rate and pattern of progression, the presence or absence of cognitive impairment and motor neuron degeneration, as well as disease duration. This clinical variability suggests that genetic and environmental factors play an essential role in the development and progression of the disease. Occupational exposure to toxic chemicals, heavy metals, and extremely low-frequency electromagnetic fields has previously been shown to increase the likelihood of developing neurological disorders.[285-287] Personal habits research has shown an increased risk of developing ALS in smokers and a generally worse prognosis following disease onset.[288] In comparison, alcohol intake was found to be associated with a decreased risk of developing ALS.[289] The research on head trauma and the development of neurological disorders was inconclusive.[290,291]

Recent analysis of a large genetic dataset suggested high cholesterol levels as a risk factor for ALS. The study also confirmed an association with smoking and lack of physical activity.[105,292]

Studies have shown that environmental factors can affect people's chances of developing ALS or FTD. However, the limitation of this research is that it was conducted on case cohorts that were not genetically selected. Different sets of environmental factors may have different effects on different genes. Therefore, future genetic epidemiology efforts should focus on cohorts selected based on their underlying genetic risk. Studying such population-based cohorts that have been assiduously collected and phenotyped for clinical features, genetics, epigenetics, and environmental and lifestyle exposures will be essential to these efforts.[105]

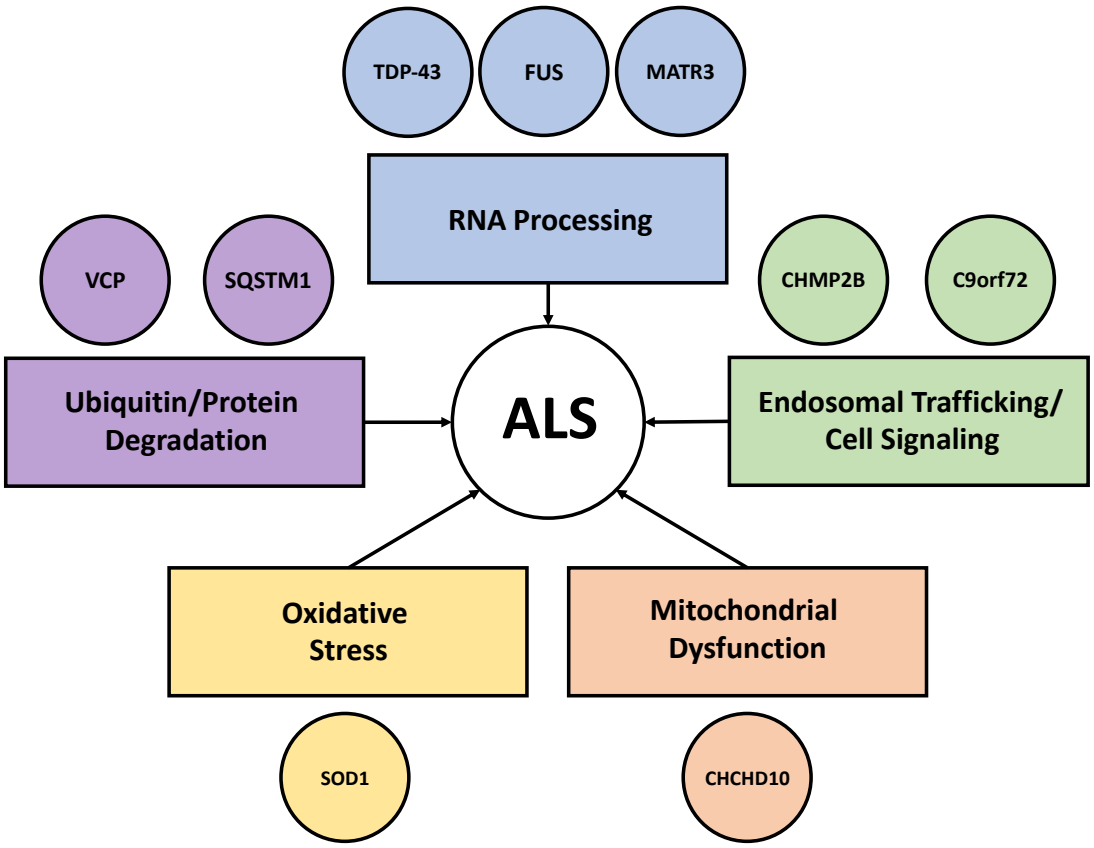
Approximately 10-15% of ALS cases can be explained by genetic mutation in roughly thirty genes linked to this neurological disorder. The etiology of the other 85-90% of cases is not well understood. These cases could be due to environmental, gene-environmental or stochastic factors. Many environmental and lifestyle factors have been implicated as risk factors for ALS. Occupational risks such as serving in the US military [293], exposures to electromagnetic radiation [294,295], lead [296-298], pesticides [299-301] and industrial chemicals [29], as well as head trauma [302] have long been thought to increase the risk of developing ALS. Population research showed that the incidence of ALS in veterans and professional athletes is two to three times higher than in the regular population.[303-305] Recent GWAS study determined that strenuous physical exercise increases the risk of developing ALS.[306] Furthermore, tobacco smoking, significantly

reduced or high body mass index (BMI)[307] were also suggested to contribute to the onset of the disease.

Al Chalabi and colleagues suggested that the manifestation of ALS is on average a six-step process, in which the predisposing variants carried by individuals interact with multiple environmental triggers.[308] Later, Chio and colleagues revised this stepwise disease development process to two step-process in ALS patients with SOD1 mutations and to four step-process in ALS patients with TDP-43 mutations.[309]

The evolution of next-generation sequencing and advances in environmental risk factors research, over the last decade, significantly transformed the field of epigenetics. The multi-omics data allowed investigation of interactions between environmental risk factors and genetic risk factors in ALS.[310] Multiple studies have demonstrated that environmental stressors can cause epigenetic alterations that can modify gene expression through DNA methylation [311-314], histone modification [315], and non-coding RNA-associated gene silencing [316-319] leading to neurodegeneration. Additionally, reactive oxygen species (ROS), a species of free radical, induced through environmental signals, have been demonstrated to cause oxidative stress and, ultimately, resulting in a range of epigenetic modifications that alter gene expression.[320-322]

Figure 19. Biological pathways enriched with ALS genes.



ALS genes described in Chapter 3 classified by common molecular pathways involved in the disease pathophysiology. Pathological variants in TDP-43, FUS and MATR3 genes lead to disruption in RNA processing; variants in CHMP2B and C9orf72 repeat expansions negatively affect endosomal trafficking and cell signaling; mutated CHCHD10 leads to mitochondrial dysfunction; mutated SOD1 leads to oxidative stress; pathological variants in VCP and SQSTM1 affect ubiquitin and protein degradation.

4 IDENTIFICATION OF MUTATIONS IN THE SPTLC1 GENE AS A CAUSE OF JUVENILE-ONSET ALS

4.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

I am a joint first author on the manuscript. I participated in the study design, performed Sanger sequencing and data analysis for the follow-up research on SPTLC1 mutations in adult-onset ALS cases. I also participated in proofreading the manuscript for critical content. This research was published in BioRxiv [323]. A modified version of this article has been recently accepted for publication in JAMA Neurology.

Other authors: Ruth Chia, Ravindran Kumaran, John Landers, Adriano Chiò, Thomas Crawford, Bradley Smith, and Bryan Traynor designed the study, wrote the report, did the literature search, and drew the figures. Olga Pletnikova, Juan Troncoso, Danny Miller, The International ALS Genomics Consortium; The ITALSGEN Consortium; the FALS Consortium, Project MinE, Ian Blair, Carol Dobson-Stone, John Kwok, Robert Brown, Andrea Calvo, Gabriele Mora, Adriano Chiò, Thomas Crawford, Christopher Shaw, Marc Gotkine, Bradley Smith and Bryan Traynor obtained samples and clinical data from patients. Janel Johnson, Ruth Chia, Ravindran Kumaran, Faraz Faghri, Alan Renton, Hannah Pliner, Simon Topp, Nada Alahmady, Raphael Gibbs, Jinhui Ding, Michael Nalls, Carol Dobson-Stone, Clifton Dalgard, Sonja Scholz, Marya Sabir, Sarah Ahmed, Ian Glass, Fawn Leigh, John Landers, Bradley Smith, Adriano Chiò, Thomas Crawford, and Bryan Traynor performed experiments and the data analysis.

4.2 BACKGROUND

Juvenile-onset ALS (JALS) is a rare pediatric motor neuron disease in which the first symptoms appear before the age of 25. Similar to adult ALS, juvenile ALS is characterized by progressive upper and lower motor neuron degeneration. Upper motor neuron signs in JALS include hyperreflexia, muscle spasticity, extensor plantar responses, and pseudobulbar syndrome (uncontrolled crying or laughing) [324]. In contrast, lower motor neuron signs include muscle weakness, hyporeflexia, cramps, and atrophy.[325] Juvenile ALS is a clinically and genetically heterogeneous disorder, with clinical, pathological, and genetic overlap with other neurological conditions.[326] The disorder is most commonly inherited in an autosomal recessive manner. However, cases with autosomal dominant inheritance have also been described in the literature.[323,327]

Unlike adult ALS, juvenile ALS cases typically have slower disease progression and prolonged survival. Most adult ALS cases have only three to five years of survival since the symptom onset, with less than 10% making it past the five-year mark and less than 5% make it past ten years. In contrast, juvenile ALS cases can live up to several decades from the time of symptom onset.[328] Less frequently JALS-plus or atypical JALS phenotypes may also exhibit additional features indicative of more extensive neurologic or multisystem involvement.[286]

Some recent studies showed that genetic factors might significantly contribute to young-onset ALS than to ALS cases overall.[164] However, the genetic profile of juvenile-onset ALS cases remains poorly understood. Exome-sequencing was used in

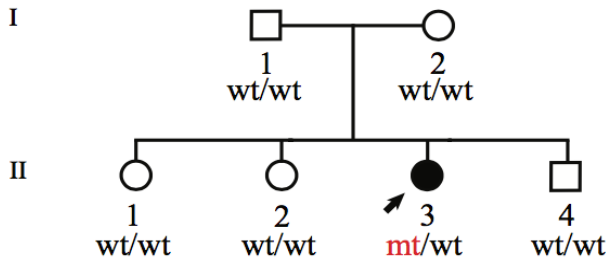
this study to compare individuals with juvenile-onset ALS to their non-affected family members and identify potential genetic lesions responsible for the disease.[286]

4.3 METHODS AND RESULTS

4.3.1 Patient description

Between March 2016 and March 2020, four unrelated juvenile-onset ALS cases were recruited and studied. All patients developed the disease in the first or second decade of life and had no previous family history of neurological disorders. Patients and their phenotypes are described below. Detailed clinical features of the four patients are summarized in **Table 2**.

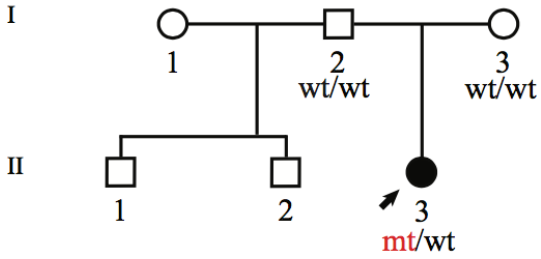
Figure 20. Pedigree of patient 1.



Patient 1 started to display growth retardation and spastic diplegia at the age of five. By the age of twenty, she developed several symptoms indicative of motor neuron degeneration: quadriplegia with significant muscle atrophy, brisk lower leg reflexes, fasciculations and weakness of the tongue, mild cognitive dysfunction, dysarthria, and

severe respiratory problems that required artificial ventilation and tracheostomy. Using the El Escorial criteria [61], the patient was diagnosed with juvenile-onset ALS. Parents and siblings of patient 1 were unaffected and not mutation carriers of known ALS genes.[286]

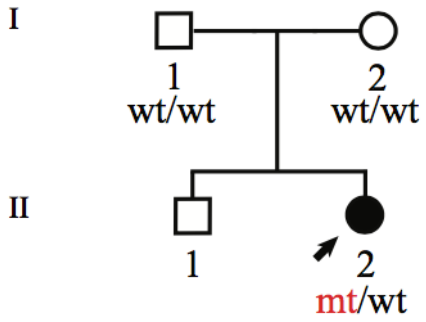
Figure 21. Pedigree of patient 2.



A second individual carrying a p.Ala20Ser amino acid change in SPTLC1 gene (patient 2) was identified through the GeneMatcher program at GeneDx, This amino acid shift arose from a mutation in an adjacent nucleotide compared to patient 1 (chr9:94874843, G>T). Patient 2 was a right-handed teenage girl with mixed African American and white race/ethnicity who presented with a six-year history of gradually progressive generalized limb and bulbar weakness. She had a long-standing history of diminished weight of unknown cause, and her school performance began to decline in her mid-teens. There was no family history of ALS or neuromuscular diseases. Her neurological examination at presentation revealed a body mass index less than the 1st percentile, exaggerated lumbar lordosis, tongue fasciculations and wasting, generalized muscle atrophy and weakness, brisk asymmetric ankle reflexes, a positive Gower's sign, and normal sensation. Neurophysiological testing revealed active and chronic denervation without evidence of sensory neuropathy. The neuropsychological evaluation showed

decreased sustained attention and impaired executive functioning. She was diagnosed with juvenile ALS based on the revised El Escorial criteria [61.286]

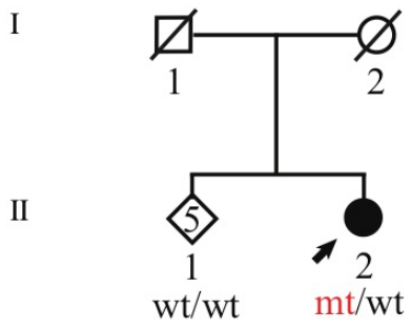
Figure 22. Pedigree of patient 3.



A third juvenile-onset individual (patient 3) who carried a p.Ser331Tyr (chr9:92047261G>T) heterozygous mutation in SPTLC1 was also identified through the GeneMatcher program at GeneDx. Patient 3 was an eleven-year-old African American girl with a history of failure to gain weight and toe-walking since the age of four. She presented at age ten with a deteriorating gait, hand weakness, right foot paresthesia, dysphagia, and increased sweating. There was no family history of ALS or neuromuscular diseases. Examination revealed marked atrophy, postural tachycardia, bilateral cataracts, a wasted, fasciculating tongue with an exaggerated jaw jerk, generalized fasciculations and weakness associated with hyperreflexia, and decreased pinprick sensation in a glove-and-stocking distribution. The patient walked abnormally due to weakness and bilateral foot drop, and she had a positive Gower's sign. Neurophysiological examination showed sensorimotor axonal neuropathy, as well as polyphasic potentials on electromyography. She was diagnosed with a juvenile ALS-

plus syndrome due to her prominent motor symptoms and modest sensory-autonomic involvement.[286]

Figure 23. Pedigree of patient 4.



Patient 4 was a thirty-four-year-old Turkish woman with a history of arm and leg weakness and atrophy since the age of fifteen. There was no family history of ALS or neuromuscular disease, and none of her five siblings had symptoms. She was diagnosed with juvenile ALS at the age of fifteen and has been on Riluzole since then. Her symptoms were slowly progressive, and there were no upper motor neuron signs on examination. During her last review at the age of thirty-four, she used a wheelchair, though she can walk short distances with assistance. She had no dysphagia, did not require oxygen, and her weight was normal. Neurophysiological examination at that time revealed denervation activity in all muscles and no evidence of multifocal motor neuropathy.[286]

Table 2. Summary of the clinical features in patients 1, 2, 3 and 4.

Clinical features	Patient 1	Patient 2	Patient 3	Patient 4
Gene change	p.Ala20Ser	p.Ala20Ser	p.Ser331Tyr	p.Leu39del
Age at onset	5 yrs	<10 yrs	4 yrs	15 yrs
Age at evaluation	20 yrs	mid-teens	11 yrs	34 yrs
BMI (Z-score)	13th pctl (-1.12)	< 1st pctl (-7.0)	< 1st pctl (-6.5)	Normal
Back deformities	Severe scoliosis	Lordosis	Normal posture	-
Foot deformities	Pes cavus	-	Pes cavus/varus	-
Walking	Non-ambulatory	Steppage	Steppage	Abnormal
Atrophy	Global, contractures	Global	Global	Global
Weakness	Generalized	Generalized	Generalized	Generalized
Reflexes	Hyporeflexia, Ach brisk	Hyporeflexia, Ach brisk	Hyperreflexia, Ach absent	Hyporeflexia
Tongue	Wasted and fasciculating	Wasted and fasciculating	Wasted and fasciculating	-
Jaw jerk	Present	-	Present	-
Respiratory	Trach. At 17 yrs	-	Dyspnea on exercise	Normal
Cognition	Executive dysfunction	Executive dysfunction	Normal	-
Sensory	Normal	Normal	Moderate pain loss in glove-stocking, painless foot ulceration	Normal
Neurophysiology				
Motor	Chronic denervation	Acute and chronic denervation	Axonal loss, Polyphasia	Denervation
Sensory	Normal	Normal	Axonal loss	Normal
Additional features	-	Sapular winging, Gower's sign	Gower's sign, vitamin D def., hyperhidrosis	Uses wheelchair

Ach - Achilles tendon; def. - deficiency; yrs - years; trach - tracheostomy; pctl – percentile.

4.3.2 Exome sequencing and alignment of juvenile-onset ALS cases

The Illumina TruSeq kit was used to enrich the DNA according to the manufacturer's protocol. A HiSeq2000 was used for sequencing of 100bp paired ends of the captured DNA. Release #19 of the reference human genome (UCSC hg 19) was used as the template for sequence alignment and variant identification. The Genome Analysis Toolkit (version 3.8, http://www.broadinstitute.org/gsa/wiki/index.php/Home_Page) was used to perform these tasks. PCR duplicates were removed using Picard Software (version 2.0.1, http://broadinstitute.github.io/picard/) before the identification of variations. In each sample, the same portion of the exome was sequenced, with a coefficient of determination of $r^2 = 0.92$ of read numbers vs. targeted sequence bait. Exome sequence data from the ALS database (n = 1,500), the control cohort (n = 4,500) and the Patient 1 family (n = 6) case were treated, sequenced and aligned in the same manner.[286]

Paternity and maternity were confirmed by identity-by-descent analysis performed using PLINK software, version 1.9.[329] Exome data was reviewed to ensure that that patient 1 and her family members are not mutation carriers in known ALS-related genes, such as SOD1, TARDBP, FUS, SETX, OPTN, UBQLN2, VCP, MATR3, VAPB, SQSTM1, CHCHD10, ALS2, CHMP2B, DCTN1, SPG11, VEGFA, PFN1, TUBA4A, TBK1, and NEK1. Furthermore, the patient's 1 family members did not carry the C9orf72 hexanucleotide repeat, as confirmed with repeat-primed PCR.[108,109] Sequence data was checked for indels or base pair variants that were only present in the affected child, but not in parents or siblings. A polymorphism was declared a true de novo mutation if present in the affected child but not present in the mother or father.

To focus on rare variants, genomic variations present in various human population databases at a frequency higher than 3.3 x 10^{-5} were removed. Subsequently, synonymous, intronic, or intragenic changes were filtered from the variant list using the ANNOVAR annotation program.[330] The sequencing databases inspected for this part of the project included the 1000 Genomes Project, the Exome Sequencing Project database, the Exome Aggregation consortium database [331], and the Haplotype Reference Consortium database [332]. To confirm the presence of the discovered sequence variations, custom primers were used for Sanger sequencing of the patient's family members.[286]

DNA from patients 2 and 3 and their families was sequenced at GeneDx, using IDT xGen Exome Research Panel (version 1.0). A HiSeq2000 was used for sequencing of 100-bp paired ends of the captured DNA. Release #19 of the reference human genome (UCSC hg 19) was used as the template for sequence alignment and identification of variants. Detailed alignment and variant calling protocols were previously described by Retterer and colleagues.[286,333]

4.3.3 Filtering process to conclude that SPTLC1 mutations cause jALS

Several points of evidence suggest that mutations on SPTLC1 gene are the likely cause of juvenile ALS. First, three unrelated patients diagnosed with juvenile ALS were identified, each of whom carried de novo mutations in SPTLC1 gene. These variants are extremely rare, as they are not present in large online databases of human polymorphisms involving ~200,000 individuals. Second, two of the patients carried the same alanine to serine amino acid shift at position 20 of the SPTLC1 protein. Although

the amino acid shift is the same, they arose from different nucleotide changes. Third, mutations in SPTLC1 are known to give rise to a neurological disease (HSAN1). Fourth, there are previous publications of a juvenile carriers of SPTLC1 mutations with similar phenotype, though they were not labeled as juvenile ALS.[334] Fifth, multiple cell-based and biochemical assays show that the Ala20Ser mutations disrupt SPTLC1 protein function. Finally, Mohassel et al, reported two other children diagnosed with juvenile ALS carrying de novo SPTLC1 Ala20Ser mutations.[335]

From a statistical perspective, considering only the two Ala20Ser mutations in this study, it is possible to quantify the probability of detecting the same protein mutation in two unrelated individuals:

$$\mathbf{(1.0\times10^{-8}) \times (1.0\times10^{-8}) \times (5.0\times10^{-6}) \times (5.0\times10^{-6}) \times (2.0\times10^{-6}) \times (2.0\times10^{-6}) = 1\times10^{-38}}$$

The following values were used for the calculation above:

- 1.0×10^{-8} (*de novo* mutation rate in the first patient)
- 1.0×10^{-8} (de novo mutation rate in the second patient)
- 5.0×10^{-6} (conservatively assuming the first Ala20Ser variant in 1 in 200,000 of the population)
- 5.0×10^{-6} (conservatively assuming the second Ala20Ser variant in 1 in 200,000 of the population)
- 2.0×10^{-6} (based on the incidence of ALS with 10% of incident cases being classified as juvenile ALS in the first patient)
- 2.0×10^{-6} (based on the incidence of ALS with 10% of incident cases being classified as juvenile ALS in the second patient)

The calculation above demonstrates that the risk of seeing two JALS individuals carrying de novo SPTLC1 Als20Ser variant by chance alone is = $1x10^{-38}$, or 1 in 100 undecillions.

4.3.4 Analysis of SPTLC1 in the follow-up adult-onset ALS cohort

A total of 5,607 adult-onset ALS cases were screened for the presence of mutations in the SPTLC1 gene. The data was generated in the following way: (i) 1,208 cases underwent exome sequencing at the Laboratory of Neurogenetics; (ii) 1,274 cases underwent genome sequencing at the Uniformed Services University on an Illumina X10 sequencer; (iii) 1,170 samples of the FALS Sequencing Consortium and 1,305 samples of the King's College London effort were sequenced similarly;[333,336] and (iv) 650 samples underwent Sanger sequencing at the Laboratory of Neurogenetics.[1] The demographic data for the follow-up adult-onset ALS cohort are summarized in **Table 3**.

Table 3. Demographic data for adult-onset ALS cases screened for SPTLC1 mutations.

	ALS cases (n = 5607)	Control subjects (n=5,710)
Age at onset/sampling (SD)	59.1 (13.0)	86.1 (6.9)
Female (%)	1,249 (41.1%)	3,370 (59.0%)
Site of onset:		
Bulbar (%)	800 (27.3%)	NA
Spinal (%)	2,095 (71.7%)	NA
Respiratory (%)	27 (1.0%)	NA
Familial disease (%)	410 (13.6%)	NA
C9orf72 repeat expansion carrier (%)	2,846 (5.3%)	NA
Race:		
White (%)	3,029 (99.6%)	5,710 (100.0%)
Black (%)	5 (0.2%)	0 (0%)
Asian (%)	1 (0.1%)	0 (0%)
Other (%)	2 (0.1%)	0 (0%)
Ethnicity:		
Non-Hispanic (%)	3,032 (99.9%)	5,710 (100.0%)
Hispanic (%)	3 (0.1%)	0 (0.0%)

A total of 5,607 ALS cases and 5,710 control subjects were used in the follow-up analysis. The average age of the cases at onset/sampling was 59.1 years old. Females represented 41.1% of cases and 59.0% of control subjects. The symptom onset distribution was as follows: 27.3 % of the ALS cohort was presented with bulbar onset, 1.0% with respiratory onset, and 71.7% with spinal onset. 5.3% of ALS cases were carriers of C9orf72 repeat expansion.

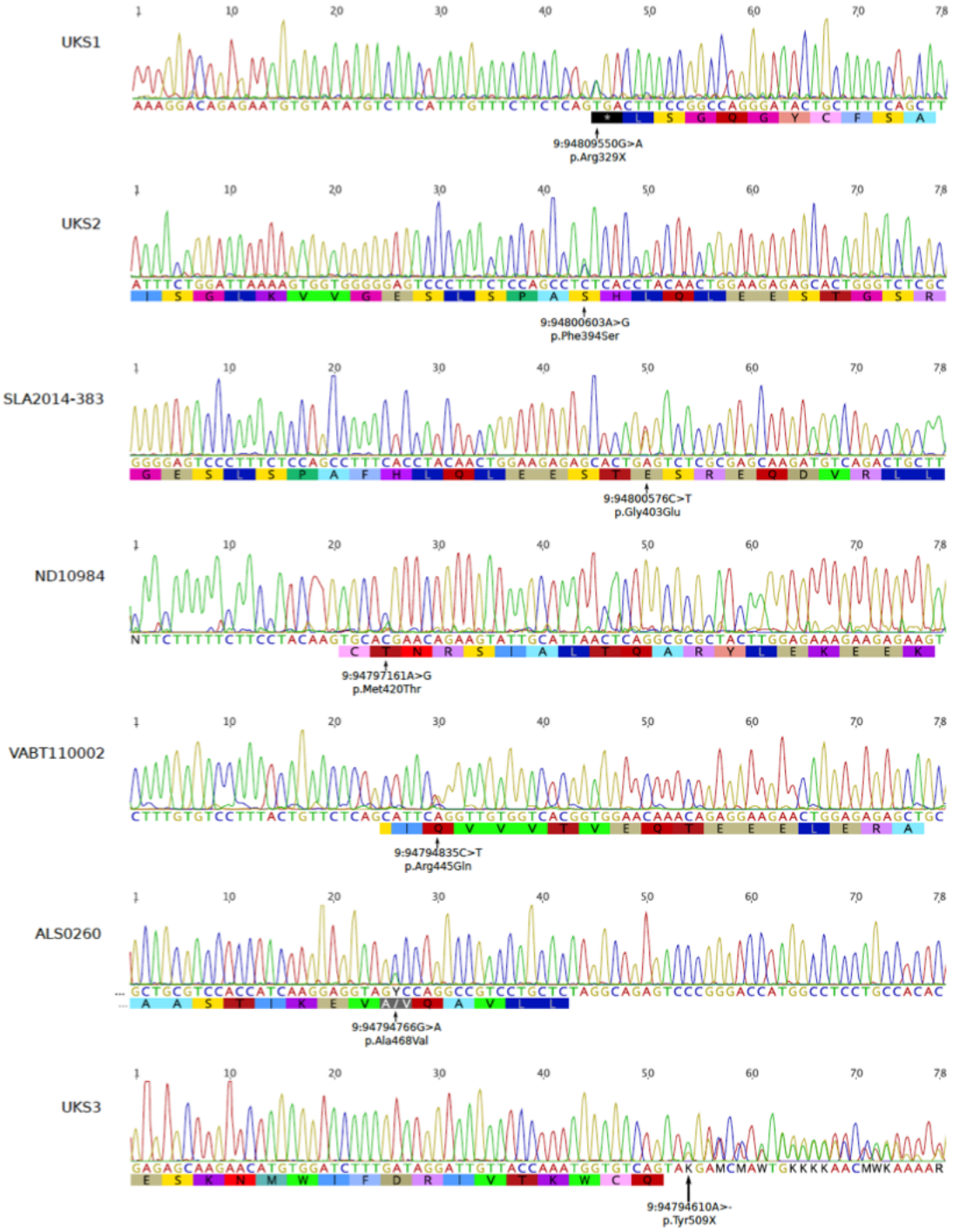
Variants in SPTLC1 were considered to be deleterious if they: (i) were not present in the 4,647 control Alzheimer's disease sequencing project (ADSP) subjects; (ii) had a frequency less than $3.3x10^{-5}$ in online databases of human polymorphisms, including the 51,592 European and 8,949 Finnish non-neurological individuals in gnomAD, and the 77,301 samples in Kaviar [337]; and (iii) were designated as 'damaging' according to four out of five prediction algorithms [338]; were identified as "stop gain" or "frameshift"; or as splice site mutations with a dbscSNV score higher than 0.6. Gene burden testing of SPTLC1 was performed using publicly available control data (gnomAD and Kaviar) as implemented in the Test Rare variants with Public Data (TRAPD) software

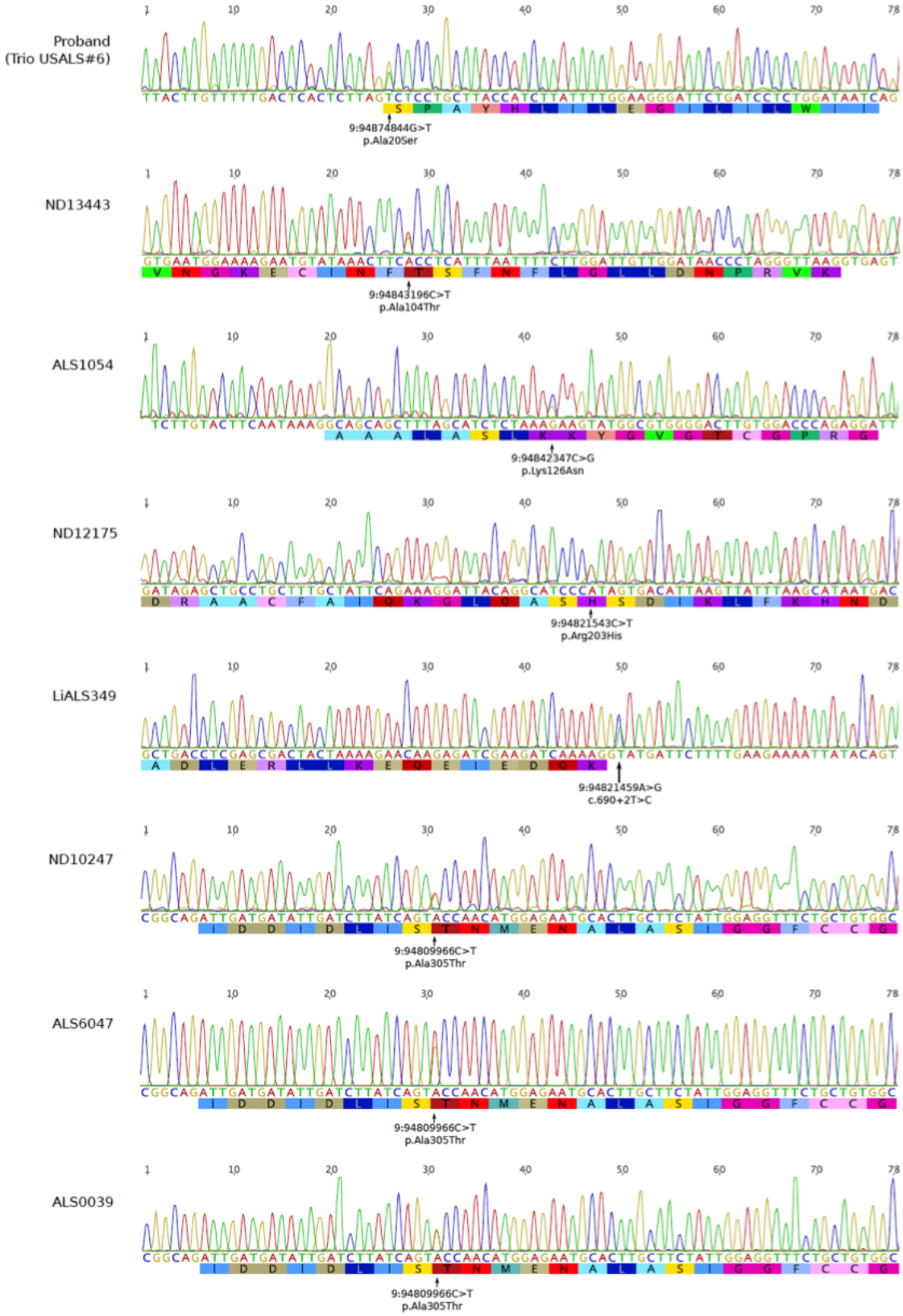
package.[339] This script performs a one-sided Fisher's exact test to determine if there is a more significant burden of qualifying variants in cases than controls for a tested gene. The threshold for statistical significance was set at $p \leq 0.05$. The presence of the variations in relevant samples was checked using the Sanger sequencing with customized primers.[323] The results are summarized in **Table 4**.

Table 4. SPTLC1 mutations identified in the follow-up cohort.

Sample and mutation		Clinical and demographic features						
Sample	*Mutation*	*Sex*	*Age*	*Onset*	*Type*	*C9orf72*	*Sensory/ autonomic*	*Country*
ND13443	p.Ala104Thr	M	55	spinal	sporadic	no	none	USA
ALS1054	p.Lys126Asn	F	66	spinal	familial	no	none	Italy
ND12175	p.Arg203His	M	38	spinal	sporadic	no	none	USA
A14LIALS68	p.Arg219X	M	49	spinal	sporadic	no	none	Finland
A14LIALS349	c.690+2T>C	M	57	spinal	familial	yes	none	Finland
B555	p.Arg236Cys	M	51	spinal	sporadic	no	NA	Italy
SLA2010-83	p.Arg240Cys	M	73	bulbar	sporadic	no	none	Italy
ND10247	p.Ala305Thr	M	76	bulbar	familial	no	none	USA
AUS145-010335	p.Ala305Thr	F	58	NA	familial	no	NA	Australia
ALS0039	p.Ala305Thr	M	64	spinal	familial	no	none	UK
SLA2011-105	p.Cys318Ser	M	58	spinal	sporadic	no	none	Italy
UKS1	p.Arg329X	M	62	spinal	sporadic	no	none	UK
UKS2	p.Phe394Ser	M	54	spinal	sporadic	no	none	UK
SLA2014-383	p.Gly403Glu	F	75	bulbar	sporadic	no	none	Italy
A356	p.Glu406Lys	F	52	spinal	sporadic	no	NA	Italy
ND10984	p.Met420Thr	M	52	spinal	sporadic	no	none	USA
VABT110002	p.Arg445Gln	M	72	spinal	sporadic	no	none	USA
ALS0260	p.Ala468Val	M	81	spinal	sporadic	no	none	Israel
UKS3	p.Tyr509X	F	56	Mixed	sporadic	no	none	UK

Figure 24. Chromatograms for SPTLC1 mutations identified in the follow-up cohort.



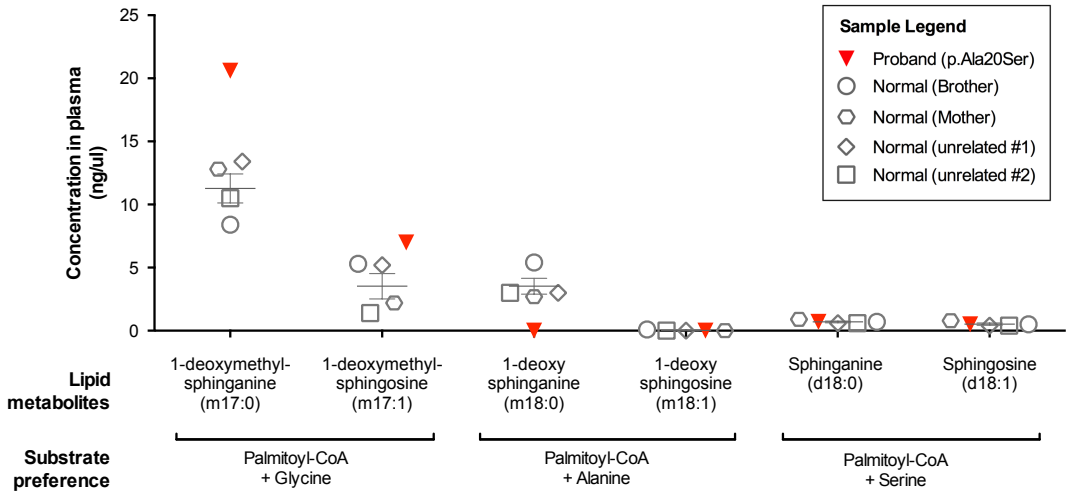

Proband
(Trio USALS#6)
9:94874844G>T
p.Ala20Ser
ND13443
9:94843196C>T
p.Ala104Thr
ALS1054
9:94842347C>G
p.Lys126Asn
ND12175
9:94821543C>T
p.Arg203His
LiALS349
9:94821459A>G
c.690+2T>C
ND10247
9:94809966C>T
p.Ala305Thr
ALS6047
9:94809966C>T
p.Ala305Thr
ALS0039
9:94809966C>T
p.Ala305Thr

4.3.5 Measurement of sphingolipid levels

The sphingolipids were measured on two separate occasions in two separate locations. This was done for convenience, to ensure data integrity, and to enable measurements before and after treatment with serine. In the first instance, the sphingolipids were extracted from the cell pellet samples using methanol extraction [340]. Chromatographic separation was performed on an Acquity Ultra Performance Liquid Chromatography (UPLC) system by injecting 5ul of the samples into an Acquity UPLC Charged Surface Hybrid C18 column. The mobile phase consisted of 0.2% v/v of formic acid in both water and acetonitrile. A gradient elution was used over 5.5 min with a flow rate of 0.4 ml/minute. The column temperature was maintained at 40°C. The sphingolipids were then detected in a Xevo TQ-S triple quadrupole-mass spectrometer using the multiple reaction monitoring method. Acquired data were analyzed using MassLynx software. Calibration equations for the sphingolipids were obtained by plotting response against concentration (ng/mL). The equation showed good linearity over the 1ng to 3,000ng range.[323]

The measurements were performed at Georgetown University Mass Spectrometry Core facility, Washington D.C. 20057. The measurement results are shown in **Figure 25**.

Figure 25. Measurement of complex sphingolipid levels in patient 1 and controls.



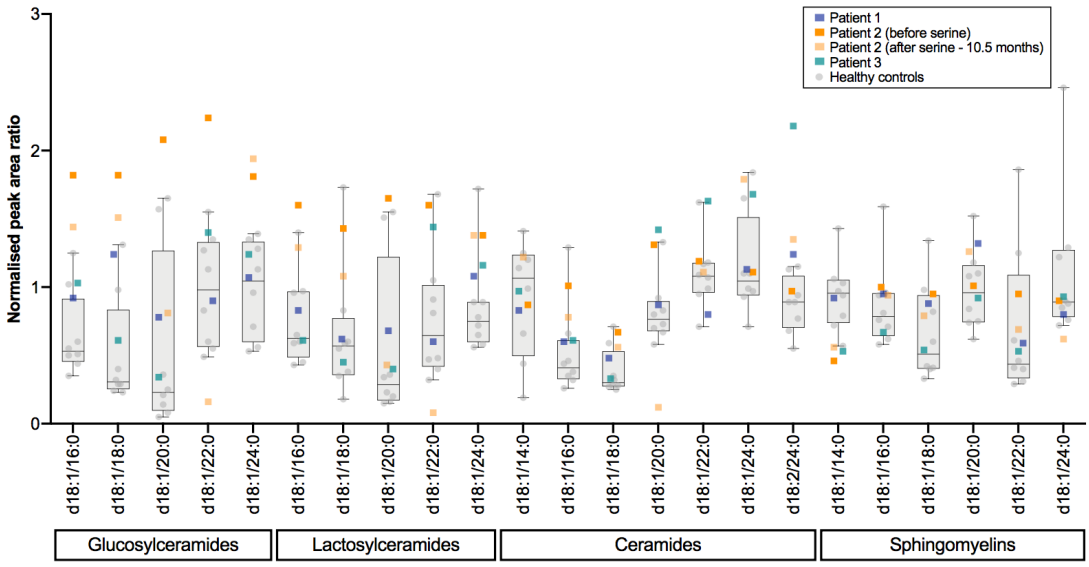
Plasma sphingolipid analysis of the patient 1 and healthy controls (unaffected mother, unaffected brother, and two unrelated healthy individuals). Patient 1 had approximately 1.8 times higher levels of the toxic sphingolipid intermediates 1-deoxymethyl-sphinganine and 1-deoxymethyl-sphingosine compared to the healthy controls.

A group led by Carsten Bönnemann that modeled the SPTLC1-associated ALS variants in induced pluripotent stem cells (iPSC) discovered an increase in de novo synthesized sphingolipids compared to the cells expressing wild type SPTLC1. Based on these discoveries, the group concluded that ALS-associated SPTLC1 variants may increase unrestrained SPT activity, therefore L-serine supplementation, might exacerbate the overproduction of complex sphingolipids and may worsen the biochemical phenotype in patients with SPTLC1-related ALS.[341] Real life data on L-serine supplementation in patient 2 did not confirm this hypothesis, though additional patients will have to be treated in the future to confirm the safety of this treatment.

To assess the L-serine supplementation's safety and metabolic effects, we performed plasma concentration measurements of four classes of complex sphingolipids

(glucosylceramides, lactosylceramides, ceramides, and sphingomyelins) formed by attachment of fatty acids with a varying number of carbons to the sphingosine backbone. The plasma concentrations of these sphingolipids were measured in three JALS carriers of SPTLC1 mutations (patients 1, 2, and 3) and eight control individuals (two unaffected family members from patients 1 and 2, and four unrelated healthy individuals). The measurement results are shown in **Figure 26**.

Figure 26. Measurement of complex sphingolipid levels in JALS cases and controls.



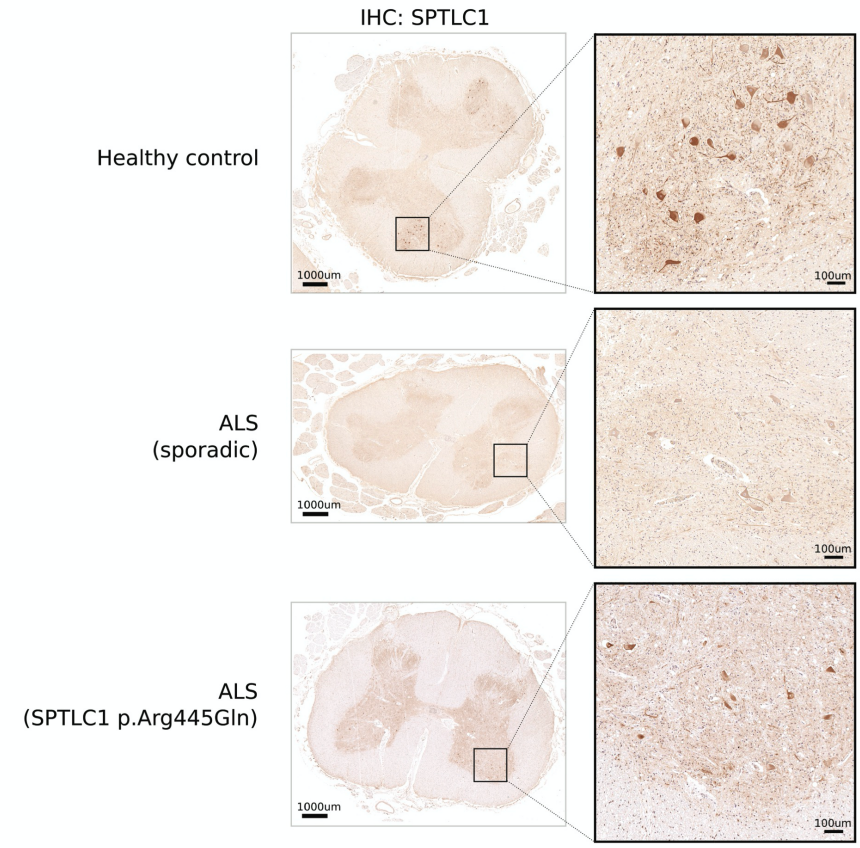
The box plots show the distribution of each complex sphingolipid for the control subjects. The median is marked by a horizontal line inside the box and whiskers representing the distribution's 5th-95th percentile. The index cases (squares) measurements were overlaid on the boxplots to show the relative levels in cases compared to controls. Patient 2 had higher glucosylceramide levels than controls before serine treatment and appeared to approach normal levels after 10.5 months of serine supplementation. No significant change of ceramides and sphingomyelins was observed in patient 2 over the same duration. The x-axis shows the different derivatives of the complex sphingolipids, and the y-axis is the normalized peak area ratio from lipid mass spectra. The measurement of complex sphingolipid levels was performed at the Biomedical Genetics Clinical Laboratory, Seattle Children's Hospital, using tandem mass spectrometry.

4.3.6 Immunohistological examination of spinal cord tissue

Formalin-fixed, paraffin-embedded spinal cord tissue blocks were sectioned to 5 um thickness and stained with a 1:50 dilution of the primary anti-SPTLC1 antibody overnight at 4°C, followed by hematoxylin counter-staining. Zeiss Wide Angle microscope with a 20x/0.8 Plan Apochromat objective was used to capture detailed images of spinal cord sections.

SPTLC1 was shown to be highly expressed in motor neurons of the anterior horn of the spinal cord in an immunohistological analysis of spinal cord tissue from a neurologically healthy control individual. In comparison, the number of motor neurons stained with SPTLC1 was reduced in autopsy tissues acquired from a patient with sporadic ALS of unknown etiology and an ALS patient carrying an SPTLC1 p.Arg445Gln mutation. The remaining motor neurons in ALS patients showed no evidence of SPTLC1 protein aggregation or mislocalization. **(Figure 27)**

Figure 27. Immunohistological examination of spinal cord tissue.



SPTLC1 is highly expressed in motor neurons of the anterior horn of the spinal cord from the healthy control individual (top panel). The number of motor neurons staining with SPTLC1 is significantly reduced in the spinal cord from a sporadic ALS patient (middle panel) and ALS patient with SPTLC1 p.Arg445Gln mutation (bottom panel) due to a decrease in the number of surviving motor neurons.

We did not perform quantification of changes in IHC staining. This decision was made due to the staining variability that is notorious with this technique. Additionally, dependent on the section of the spinal cord that was taken, staining of remaining motor neurons are not fair and accurate representation of across samples. Thus, for this purpose a decision was made to provide qualitative evidence of expression of SPTLC1

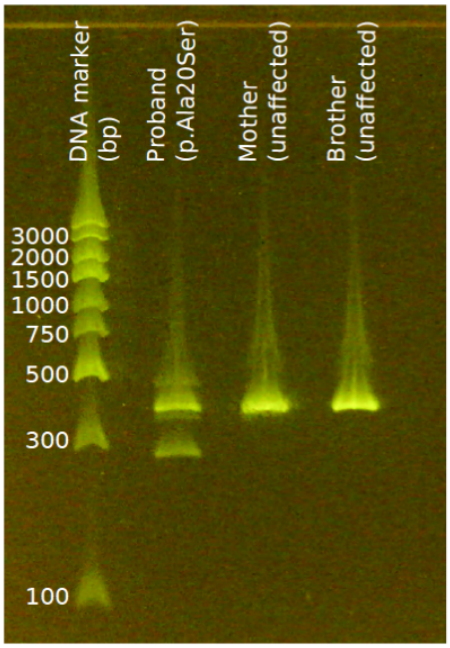
in motor neurons compared to surrounding cells and not quantitatively compare the SPTLC1 expression between samples. This is a limitation of this study.

4.3.7 SPTLC1 transcript analysis

In silico analysis using NetGene2 software predicted that the p.Ala20Ser (chr9:94874844, G>T) mutation may lead to altered splicing of SPTLC1 due to its proximity to the exon-intron border. To determine if the mutation is transcribed within the canonical transcript, TA cloning of the PCR products from patient 1 was performed, followed by Sanger sequencing of each clone. The detailed steps of transcript analysis are outlined below.

RNA was extracted from lymphoblastoid cells using Trizol (Thermo Fisher), reverse transcribed using SuperScript III First-Strand Synthesis SuperMix (Thermo Fisher), followed by PCR amplification using the primer pair CGGAGCAGTGGGTTCTGG (exon1F) and CCTCTGGGTCCACAAGTCC (exon 5R) with FastStart PCR Master Mix (Sigma Aldrich). Amplified PCR products were assessed on 2% E-Gel EX Agarose Gels (Thermo Fisher) **(Figure 28)**, cloned using the TOPO TA Cloning® Kits for Sequencing (Thermo Fisher) according to the manufacturer's protocol, and plasmids were extracted using QIAprep Spin Miniprep Kit (Qiagen). Plasmid and TA - cloned inserts were Sanger sequenced using the universal M13 forward primer.[286] Sanger sequencing results are shown in **Figure 29**.

Figure 28. Amplification of cDNA between exon 1 and exon 5.



Amplification of cDNA between exon 1 and exon 5 reverse-transcribed from RNA obtained from patient 1 carrying the p.Ala20Ser mutation and her unaffected relatives showed that the canonical transcript (~400bp) and a smaller transcript (~300bp) were present in patient 1.

Figure 29. SPTLC1 transcript analysis.

Chromatograms above show representative sequence of transcripts identified. Twenty-five clones were successfully sequenced, of which ~52% (13/25) were canonical transcripts (ten wildtype, three carrying the mutation). The remainder were alternatively spliced at exon 2 (ten sequences showed skipping of entire exon 2, two showed partial skipping of the first 28 base pairs of exon 2).

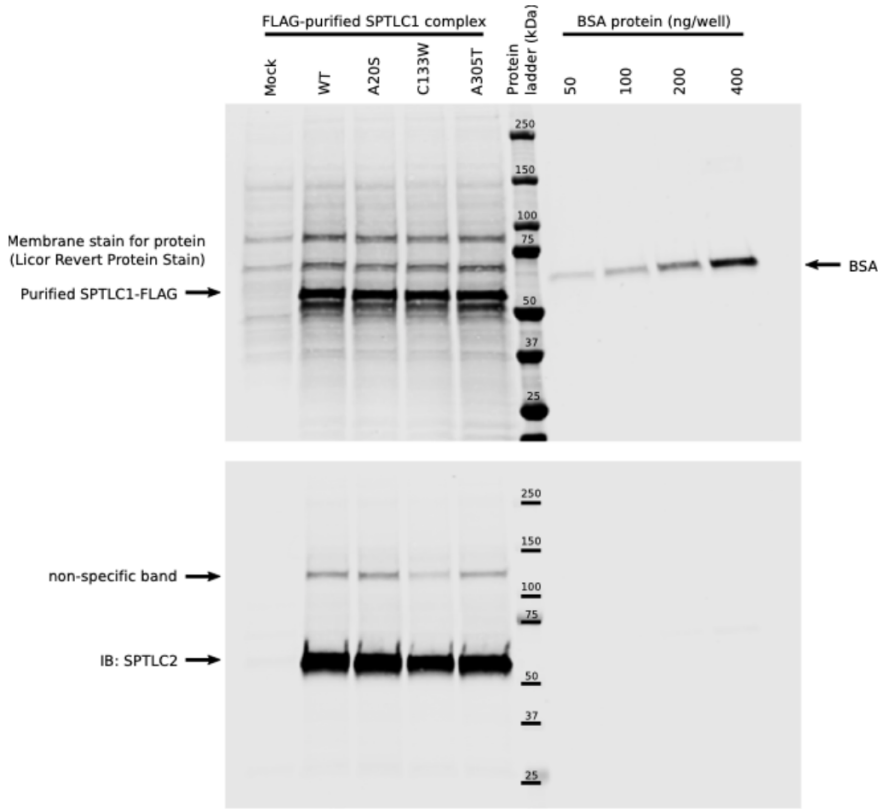
4.3.8 Western blot

Protein samples from cell lysates were prepared for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 4X sodium dodecyl sulfate sample buffer (Life Technologies) and boiled at 95 °C for ten minutes before electrophoresis on 4–20% TGX gels (Bio-Rad Laboratories Inc.). Proteins were transferred to nitrocellulose membranes using the semi-dry Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were blocked with 2X Odyssey Blocking Buffer (LI-COR Biosciences Inc.) for 30 minutes at room temperature before incubation with primary antibodies for either one hour at room temperature or overnight at 4 °C. After

primary antibody incubation, plots were washed three times for five minutes using PBS before incubation with infrared-labeled secondary antibodies (LI-COR) for one hour at room temperature. Excess secondary antibodies were removed by washing with PBS before imaging on an Odyssey CLx imaging system (LI-COR).

The following primary antibodies were used at the indicated dilutions: mouse anti-Flag 1:5,000 (Sigma-Aldrich), rabbit anti-SPTLC1 1:1,000 (Sigma-Aldrich), rabbit anti-SPTLC2 1:1000 (LifeSpan BioSciences Inc.), and rabbit anti-GAPDH (Sigma-Aldrich).[286]

Figure 30. Western blot of purified SPTLC1-FLAG proteins.



Western blot analysis was performed using peripheral blood mononuclear cells obtained from the patient 1 detected a single protein band of about 50 kD corresponding

to the canonical SPTLC1 protein. This band confirmed that the alternatively spliced variants of SPTLC1 were not translated and suggested that the full-length transcript carrying the G>T transversion translating to p.Ala20Ser mutation at the protein level was responsible for the underlying pathogenic mechanism.

4.3.9 Immunopurification of SPTLC1 protein complex

SPTLC1-FLAG proteins were immuno-purified from HEK293FT cells that were stably expressing SPTLC1. Briefly, cells from two 15 cm plates that had grown to confluence were washed twice with cold PBS, then harvested and lysed with 2 ml lysis buffer (50 mM HEPES, pH8.0, 1 mM EDTA, 0.1% (w/v) sodium monolaurate, phosphatase (Thermo Fisher) and protease inhibitor (Roche)). To aid in the solubilization of membrane proteins, the lysate was sonicated on ice for fifteen seconds at 50% power and 50% pulsation for a total of 45 seconds or until no visible clumps were observed. The lysate was clarified by centrifugation at 21,000 g for five minutes at 4 °C to remove large cellular debris followed by incubation with 40 ul EZview Red Anti-Flag M2 Affinity Gel (Sigma-Aldrich) for 2–4 hours at 4 °C with constant end-to-end mixing. Protein-gel complexes were washed three times with lysis buffer, and bound SPTLC1-FLAG protein complex was eluted with 100 ul of 120 ug/ml FLAG peptide (Sigma-Aldrich) in lysis buffer (without protease/phosphatase inhibitor) for fifteen minutes at room temperature or thirty minutes at 4 °C with constant shaking.[286]

Purified SPTLC1-FLAG was quantified by running 20 ul of the eluent on SDS-PAGE alongside 20, 40, 60, 80, 100 ng of bovine serum albumin (BSA, Thermo Fisher) protein standards per well. The gel was then transferred to nitrocellulose membrane and

stained with Revert Total Protein Stain (LI-COR), and imaged on an Odyssey CLx imaging system. The amount of FLAG-SPTLC1 was estimated based on band intensity relative to BSA's protein standard curve. Interaction of SPTLC2 with SPTLC1-FLAG was confirmed by western blot showing that the purified serine palmitoyltransferase complex was likely functional.[286]

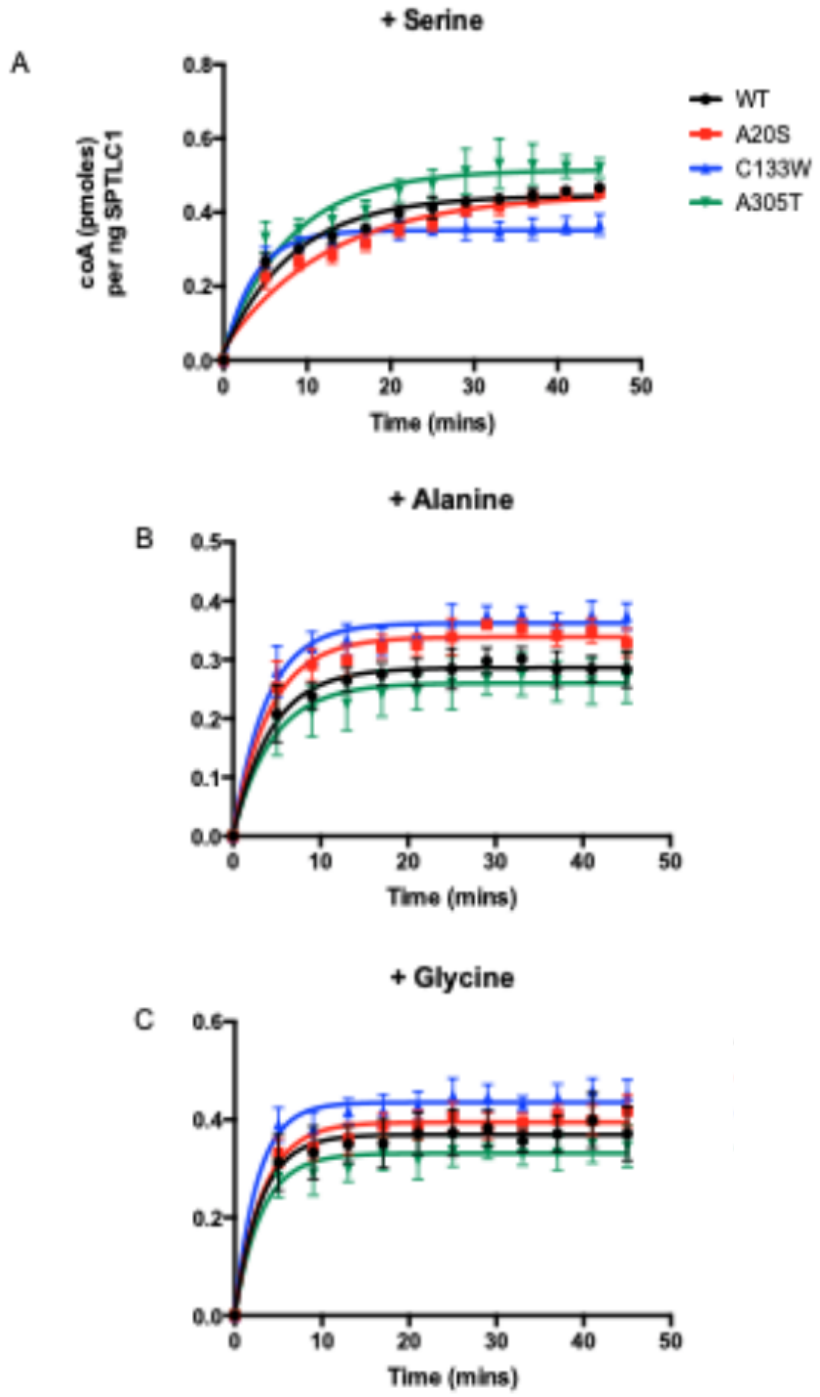
4.3.10 Photometric serine palmitoyltransferase enzymatic assay

Condensation of serine and palmitoyl-CoA by serine palmitoyltransferase enzyme produces 3-ketodihydrosphingosine and releases carbon dioxide and free coenzyme A (CoA). CoA has a free thiol/sulfhydryl (-SH) group reactive to 5,5'-dithiobis-2-nitrobenzoic acid (DTNB, Thermo Fisher). The reaction between CoA and DTNB produces mixed disulfide and 2-nitro-5-thiobenzoic acid (TNB) products that can be quantified by the absorbance of the anion (TNB^{2-}) at 412 nm. This photometric assay measured the amount of free CoA released and reflected the enzymatic activity of serine palmitoyltransferase.[286]

This assay was adapted from, and shown to be comparable to, radioactive assays measuring radiolabeled 3-ketodihydrosphingosine.[342] The assay was performed in a 96-well plate format. The diluent used for the assay was made up of 50 mM HEPES (AppliChem), pH8.0, 1 mM EDTA (KD Medical), and 0.1% (w/v) sodium monolaurate (Santa Cruz Biotechnology). Approximately 25 ng purified SPTLC1-FLAG proteins were added to 50 uM palmitoyl-CoA (Sigma-Aldrich), 5 mM L-serine (or L-alanine L-glycine, Sigma-Aldrich), 20 uM pyridoxal 5'-phosphate (Sigma-Aldrich) in a final reaction volume of 200 ul per well. Triplicate wells were assayed per wild type, per

mutant SPTLC1, and per amino acid tested. Absorbance was measured at 412 nm at the zero-time point and every two minutes for up to one hour. In place of purified SPTLC1-FLAG protein, varying CoA concentrations were included as a standard to build the calibration curve to estimate CoA released from the serine palmitoyltransferase enzymatic reaction. The estimated amount of CoA produced per ng of SPTLC1-FLAG protein was plotted over time for each amino acid.[286]

Figure 31. SPTLC1 enzymatic activity.

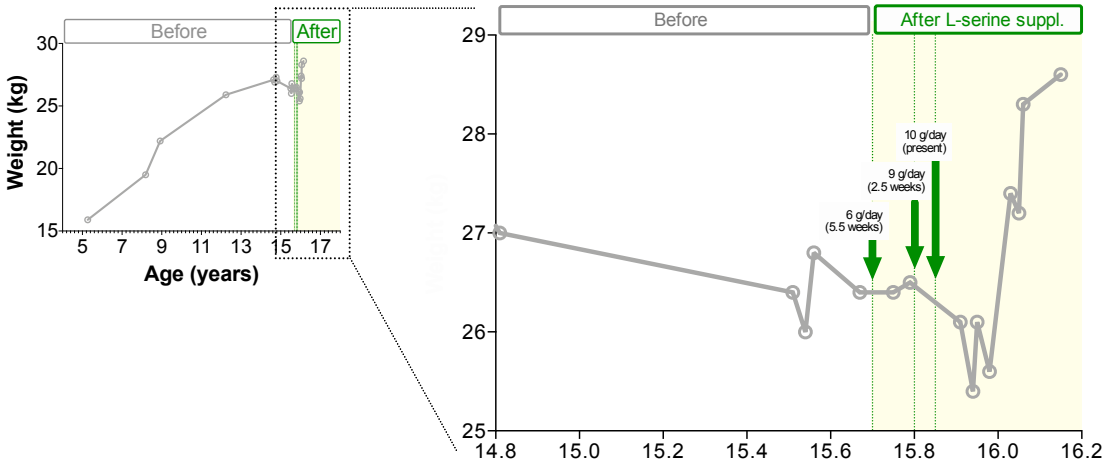


All wild-type and mutant SPTLC1 complexes (p.Ala20Ser, p.Cys133Trp, and p.Ala305Thr) are active enzymes that can utilize all three amino acids (L-alanine, L-glycine and L-serine) as substrates. The mutant SPTLC1 complexes (p.Ala20Ser and p.Cys133Trp) show preference for L-alanine or glycine over L-serine, compared to the wild-type SPTLC1 complex.

4.3.11 L-serine supplementation

Based on the previous reports that high doses of serine were well tolerated by HSAN1 patients [343], Patient 2 (the carrier of p.Ala20Ser mutation) was treated with high-dose oral serine supplementation on a compassionate basis. The treatment started with 6 grams of serine per day for five and a half weeks, then 9 grams per day for two and half weeks, and presently on 10 grams per day. While no signs of neurological improvement were observed over eleven months, the patient tolerated the treatment well and started gaining weight. Currently, she is gaining approximately one kilogram per month, which is more than she has ever gained over such a short period **(Figure 32)**.[286]

Figure 32. Patient 2's bodyweight improvement over eleven months.



4.4 SUMMARY OF THE RESULTS

To investigate a possible genetic cause of juvenile-onset ALS, exome sequencing was performed on the DNA from an individual who presented with the symptoms of motor neuron degeneration in the first decade of her life and had no family history of neurological disorders. We found a de novo mutation p.Ala20Ser in the Serine Palmitoyltransferase Long Chain Base Subunit 1 (SPTLC1) gene. Later, through the GeneDx program, we discovered two more juvenile-onset ALS individuals that carried mutations on SPTLC1. Patient 2 was the carrier of the same de novo pAla20Ser mutation as patient 1, and patient 3 had a p.Ser331Tyr mutation in SPTLC1. These mutations were not found in the 4,500 control samples from the Alzheimer's Disease database or other online databases of human population polymorphisms.

To study the role of SPTLC1 mutations in adult-onset ALS, we examined 5,607 adult-onset ALS cases. We discovered twenty new mutations in twenty-three individuals, which corresponded to 0.37% of all studied patients. These mutations were not present in our control cohort and occurred with a frequency less than $3{\cdot}3x10^{-5}$ in large databases of human polymorphisms gnomAD and Kaviar.

Finally, we treated one of the pAla20Ser carriers was treated with high-dose oral L-serine supplementation for six months. While we did not notice neurological improvements, her body mass index in the 1st percentile at the beginning of the study increased.

4.5 DISCUSSION

Several lines of evidence link the novel mutation in SPTLC1 to juvenile-onset ALS. First, the mutation was only found in patients 1 and 2 but not in their family members or in-house control dataset or the online control databases of human polymorphisms. Since the latter databases were constructed from an ethnically diverse set of individuals, our SPTLC1 mutation was likely a rare variation. Second, various SPTLC1 mutations have been previously linked to neurological diseases and clinical syndromes with symptoms similar to our patient 1. This suggests the change in SPTLC1 was indeed related to a pathological phenotype and not a background mutation. Third, SPTLC1 was known to be required for the biochemical and cellular regulation of sphingolipid metabolism, and mass spectrometry of the plasma from patient 1 revealed abnormal sphingolipid profiles. Since this perturbed pattern is consistent with disease-causing metabolic mechanisms in other neurological disorder patients [344], the SPTLC1 mutations found in our juvenile-onset ALS patients are the likely cause of the disease.

Mutations in SPTLC1 are a known cause of Hereditary sensory neuropathy type 1 (HSAN1).[345] Furthermore, a mutation resulting in a replacement of serine with phenylalanine in SPTLC1 (p.Ser331Phe; chr9:94809543; C>T) was present in a French woman displaying ALS-like symptoms similar to those shown by patient 1, including retarded growth, hypotonia, amyotrophy, and a paralyzed vocal cord together with respiratory problems.[346] SPTLC1 is a crucial unit of Serine Palmitoyltransferase, which catalyzes the reaction of L-Serine and Palmitoyl-CoA to sphinganine.[347] Several mutations in SPTLC1 change the affinity from serine towards alanine and glycine, resulting in the formation of sphingolipids lacking a hydroxy group at C1.[344] The

accumulation of deoxy-sphingolipids in plasma is a hallmark of pathogenic SPTLC1 mutations.[344] The absence of the hydroxy group impairs sphingolipid metabolism, leading to the accumulation of intermediate products within cells. In neurons, these intermediates likely affect their ability to function and ultimately result in their degradation. Indeed, in a mouse model, mutations in SPTLC1 have been found to underlie motor neuron degeneration via toxic accumulation of aberrant sphingolipid intermediates, suggesting a role of perturbed sphingolipid metabolism in the etiology of ALS. [348]

Our mass spectrometric data signified strongly elevated levels of sphingolipids in the plasma of patient 1. In contrast, sphingolipid levels were normal in each of her family members, providing evidence that perturbed sphingolipid metabolism is indeed a result of the amino acid exchange, suggesting that this perturbance could be a cause for juvenile-onset ALS.

The results of this study also suggest that SPTLC1 mutations may be a very rare genetic cause for typical adult-onset ALS. The pathology of the affected individual showed the usual signs of disrupted sphingolipid metabolism. Furthermore, a change in sphingolipid metabolism is common to several other neurological disorders with similar symptoms, such as Niemann-Pick Disease, Gaucher disease, and Alzheimer's disease. [349, 350] Importantly, aberrant sphingolipid metabolism has been linked to motor neuron degeneration in Tay-Sachs disease (also called GM2-gangliosidosis).[351] Tay-Sachs disease is caused by a recessive-autosomal mutation in hexosaminidase A and leads to the accumulation of GM2 gangliosides, a subcategory of sphingolipids, in neuronal cells. Heterozygous hexosaminidase alleles also display ALS-related

symptoms, such as motor neuron impairment, dysphagia, dysarthria, and cognitive decline, providing further evidence that a lesion in a gene responsible for sphingolipid metabolism can have drastic consequences.[352]

The results of this study suggest that early screening for SPTLC1 mutations can be used as the basis to prevent or delay the onset of ALS: dietary supplementation with serine may be able to shift the prevalence of mutant SPTLC1 use of alanine and glycine back to serine, reducing the production of deoxy-sphingolipids.[343] A diet enriched in 10% serine minimized neurotoxic sphingolipid plasma levels in Cys133Trp SPTLC1 mice and human HSAN1 patients. Indeed, a chronic daily dose of 30mg serine was previously shown to be safely tolerated and transported across the blood-brain barrier.[353] In fact, treatment of patient 2 with serine on a compassionate basis resulted in significant weight gain, which was the first time she had gained weight in several years and represented an initial proof of concept. While no signs of neurological improvement were observed, continued therapy would be required to detect such an effect.

Data collected in this study lead us to propose a new personalized medical approach. Nutritional supplementation has proven to be remarkably effective in other ALS forms: high-dose oral vitamin B2 (riboflavin) slows and even halts neurological progression in Brown-Vialetto-Van Laere cases, a rare subtype of ALS arising from mutations in the riboflavin pathway.[354-356]

5 EXOME SEQUENCING IS A POWERFUL TOOL FOR DIFFERENTIATING BETWEEN PARTIAL PHENOTYPES

5.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

This research was published in Muscle Nerve [357], and I am the third author of that manuscript. I participated in the study design, performed Sanger sequencing experiments and analysis, prepared the exome libraries and assisted in writing the manuscript. Other authors and their contributions were as follows: James Caress performed clinical and electrophysiologic examinations and wrote the manuscript. Janel Johnson and Bryan Traynor performed genetic analysis and assisted in writing the manuscript. Gregory Hawkins performed a genetic analysis and reviewed the manuscript. Raphael Gibbs performed the genetic and statistical analysis and reviewed the manuscript. Elizabeth Sullivan and Chamanpreet Chahal performed clinical examinations and reviewed the manuscript.

5.2 BACKGROUND

This study focuses on a family from the United States that presented with bulbar neuropathy. The initial diagnosis was thought to be bulbar-onset ALS, where known genes had been excluded by commercial genetic testing. For this reason, this family

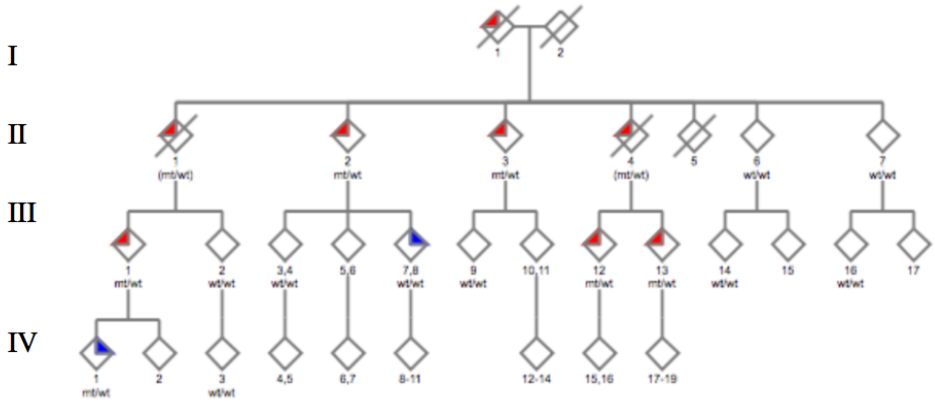
was selected for genetic research to elucidate the underlying, unknown genetic etiology. To do this, we performed exome sequencing on this family. Ultimately, the diagnosis was changed based on the genetic information revealed by these efforts, highlighting the power of modern genomic techniques to establish an accurate molecular diagnosis.

5.3 METHODS

5.3.1 Patient description

Eighteen subjects were evaluated, of which five subjects were affected by bulbar neuropathy. These patients are described in detail below, and **Figure 33** shows the pedigree of the examined family.

Figure 33. Pedigree. A large US family with European descent.



Diamonds with red indicate definitely affected family members and diamonds with blue indicate members of the family who were questionably affected.

Patient III-1. The proband was a 63-year-old woman who began experiencing facial muscle twitching at the age of 56. Within a few years, the symptoms expanded to include mild dysarthria and difficulty swallowing. She underwent blepharoplasty at age 57 to repair hanging eyelids by removing excess tissue below her eyes. Ascribed to incomplete eyelid closure, an optometrist prescribed her drops against 'dry eyes.' The dry eyes and photophobia prompted an ophthalmologist to diagnose the patient with lagophthalmos due to the blepharoplasty and inherited corneal dystrophy; the latter's significance was not explained to individual III-1. The facial muscle twitches and the tongue atrophy prompted doctors to perform an MRI, which showed deep lesions within white matter but not the brainstem. Facial muscle contractions were brief, bilateral, and non-synchronous. They were not classified as hemifacial spasms since they did not lead to eyelid closure and did not affect the corner of the mouth. The patient showed normal jaw strength and extraocular movements, and facial sensation was not affected. In the limbs, there was no hyperreflexia or distal neuropathy. The muscles of the tongue and the face – except for the masseter and sternocleidomastoid muscles - showed enlarged fibrillation and motor unit potentials with a reduction in recruitment. This was consistent with a severe and chronic neurogenic affection. Brief myokymic discharges were shown to be underlying the spontaneous muscle contractions in the face. The patient was negative for X-Linked Spinobulbar Muscular Atrophy and familial ALS genes such as SOD1, TARDBP, ANG, FIG4, and FUS based on commercial laboratory testing.

Patient II-2. This patient was 83 years old and an aunt of proband III-1 (see above). She first experienced facial twitching at 52 years of age. At 63 years, she underwent blepharoplasty to repair ptosis. Simultaneously, she had experienced bradycardia

symptoms, which led to cardiac pacemaker placement surgery. She had also attempted to treat her 'dry eyes' with bilateral corneal transplant surgery at an undisclosed age. It is unclear whether she was diagnosed with Lattice Corneal Dystrophy (LCD). She had mild problems swallowing dry food and articulating her speech. She showed frontalis muscle paralysis and bilateral facial weakness upon examination. Facial twitching was less pronounced than in the other family members. As with patient III-1, strength and reflexes in the upper and lower limbs were normal, though her feet being severely unable to sense vibrations.

Patient II-3. The 85-year-old patient II-3 - also an aunt of the proband III-1 - experienced ptosis and underwent blepharoplasty at age 76. Around age 80, she experienced problems with speech and swallowing that deteriorated to the point where food aspiration needed to be prevented using a feeding tube at age 84. In addition, the patient experienced mild sensory loss in the feet, as well as facial weakness and twitching.

Patient III-12. Patient III-12 was a 56-year-old woman who was a cousin of the proband III-1. She had experienced facial twitching and dry eyes at age 52. In contrast to other family members, facial weakness in III-12 was mild, and the tongue was normal. Similarly, sensation, reflexes, and muscle strength were normal.

Patient III-13. Patient III-13 was a 53-year-old man who was a cousin of III-1. He developed twitches in his facial muscles during his late 20s. Otherwise, the candidate was without any further symptoms until his late 40s, when subtle problems of indistinct speech started to manifest. Interestingly, despite facial twitching and weakness as well as tongue atrophy and fasciculation, he did not suffer from dysphagia. The muscle

defects are specific to the face and tongue, as sensation, reflexes, and muscle strength in the limbs were normal.

Patients III-7, **III-8,** and **IV-1**. These individuals were possibly affected by mild forms of inherited bulbar-onset neuropathy. III-7 underwent blepharoplasty at 45 years of age to treat ptosis; he also used drops to medicate dry eyes, yet there were no signs of facial muscle abnormalities or tongue fasciculation at age 64. Individual III-8 was 63 years old, exhibited subtle ptosis, and was otherwise symptom-free at the time of examination. IV-1 was a child of III-1 and showed some effects at age 37, such as fasciculation and facial asymmetry. The eyes were normal, but the patient complained of jaw cramping.

Patients I-1, **II-1**, **II-4**. These individuals were likely affected due to family reports, yet all were deceased at the time of examination, so DNA samples were not available. In addition, none of them had any tissue examination or biopsy performed that could indicate the presence of amyloid plaques and amyloidosis.

5.3.2 Exome sequencing

DNA from two affected individuals III-1 and III-12 underwent whole-exome sequencing according to the manufacturer's protocol. These two individuals were selected for the study because they are cousins and therefore the most distantly related affected individuals in the pedigree for which DNA was available. Choosing third degree relatives for identification of candidate variants is a preferred approach because

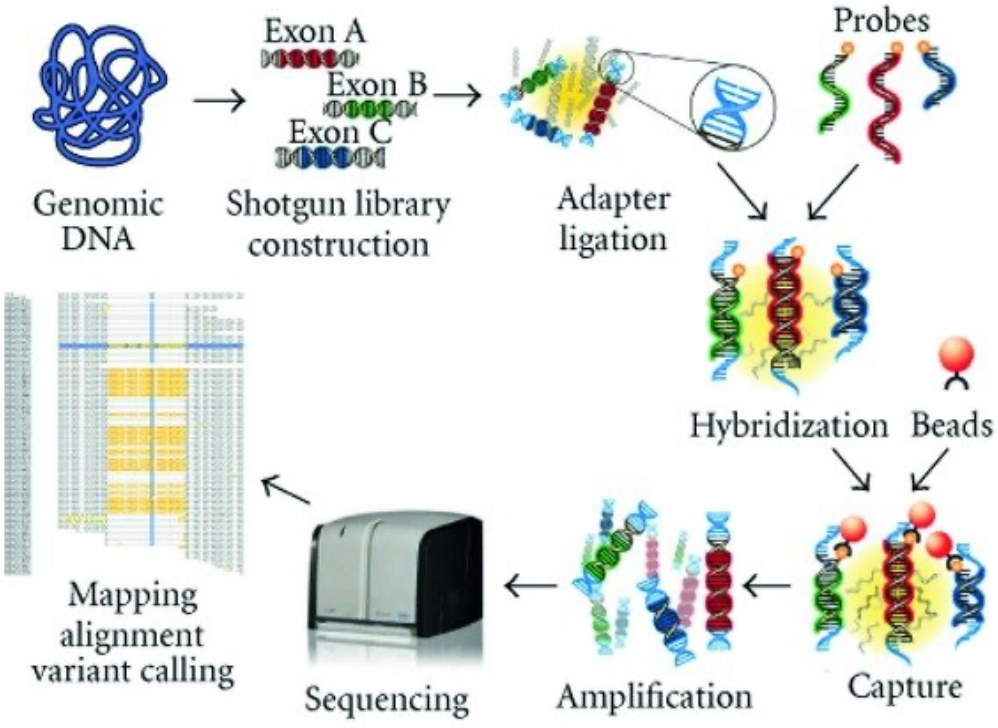
these individuals share the least amount of DNA (~12.5%), therefore reducing the chances of genetic overlap due to relatedness.

A detailed description of the protocol is provided below, divided into the various steps required to generate data.

5.3.2.1 Nimblegen SeqCap EZ Human Exome Library protocol

NimbleGen SeqCap EZ Exome capture kit was selected for this project because it had lower reagents cost and required smaller amounts of input DNA compared to NimbleGen 2.1M Human Exome Array while providing the same exome coverage. In addition, the SeqCap EZ Exome kit was previously shown to have high reproducibility and scalability. This protocol's primary benefit was that all steps could be performed at the bench without automation, and the resulting library could be sequenced on an Illumina HiSeq 2000 sequencer in our laboratory.

Figure 34. An overview of the basic steps involved in Next-Generation Sequencing.



(Reproduced from Schorderet, 2013)

1. Genomic DNA quantification

Genomic DNA (gDNA) concentration was assessed with Qubit fluorimeter 2.0 using dsDNA BR Assay pack (Invitrogen). The amount of DNA required for the following fragmentation step was 1 µg. The DNA was diluted to a total volume of 50 µL. DNA samples were vacuum concentrated or diluted as needed.

2. DNA fragmentation

The DNA sample was sheared using Covaris E210 sonicator using the following conditions: Duty Cycle: 10%, Intensity: 5, Cycles per burst: 200, Time: 6 min, Mode: Frequency sweeping, Power 23W, Temperature 6°C.

3. DNA quality check

To determine if the sheared DNA fragment length met the library requirement, that is ~150–200 bp in length, 1 µL of DNA sample was run on the Bioanalyzer 2100 (Agilent) using DNA 1000 chip.

4. DNA purification

Sheared DNA was purified with AMPure XP Magnetic Beads (Agencourt). This step was used to ensure that sheared DNA fragments are larger than 100 bp. DNA fragments >100 bp were attached to paramagnetic beads, and smaller fragments were removed using 70% ethanol wash.

The Agencourt AMPure XP beads were allowed to equilibrate at room temperature for 30 minutes and were mixed well by vortexing before the DNA cleaning procedure. Next, 216 µL of magnetic beads were added to each of the sheared DNA samples and mixed well. The mixture was incubated on the rotating platform at room temperature for 5 minutes. Next, the samples were placed on the magnetic stand, and the solution was allowed to clear for approximately two minutes. In this step, the desired DNA

fragments get attached to the magnetic beads and form a small brown-black pallet on the side of the tube. The undesired (<100bp) DNA fragments remain in the clear supernatant. The supernatant was consequently removed and discarded.

While the tubes were still on the magnet rack, the beads were washed twice (without disturbing the bead pellet) by adding 500 µL of 70% ethanol to each sample, carefully pipetting it out, and discarding after each wash. Next, the samples were let dry for five minutes. Next, the dried samples (still attached to magnetic beads) were resuspended by removing the tubes from the magnetic stand, adding 34 µL of nuclease-free water, and mixing thoroughly by vortexing. Next, the mixture was incubated for five minutes at room temperature. This step releases the bound DNA fragment from the beads back into the supernatant. Finally, the tubes were placed back on the magnetic stand, and the solution was allowed to separate. The supernatant containing the eluted DNA was transferred to a fresh 1.5-mL tube, and the beads are discarded.

Figure 35. An overview of the AMPure XP cleanup process.

(Reproduced from www.beckman.com)

5. DNA end repair

After the fragmentation and cleanup steps, each DNA fragment has a 3' overhang. These overhangs were converted into blunt ends using the End Repair Enzyme Mix in the Nimblegen End-It DNA End-Repair Kit. This was achieved by adding 3' to 5' prime exonuclease, which replaced the 3' overhang, followed by DNA polymerase, filling the remaining 5' overhang.

The end-repair reaction was prepared for each of the sheared DNA samples by mixing the following reagents: 34 µL of the purified DNA sample, 5 µL of 10× End-Repair Buffer, 5 µL of 2.5 mM dNTP mix, 5 µL 10 mM ATP, and 1 µL of End repair-enzyme mix. The reaction was incubated at room temperature for 45 minutes and placed on a heat block for ten minutes at 70°C to inactivate the enzymes. End-repaired DNA was purified using a MinElute column as follows: 250 µL of PB Buffer was added to end-repaired DNA and mixed well. Next, the mixture was loaded onto a MinElute column and filtered on the QIAvac 24 Plus vacuum. After that, each column was washed with 750 µL of PE Buffer and centrifuged at maximum speed for 1 minute. The final DNA samples were diluted with 32 µL of EB Buffer.

6. Adding an Adenine Base to End-Repaired Fragments

In this step of the library preparation, a single adenine nucleotide is added to the repaired DNA fragments through an A-tailing reaction. The adenine nucleotide forms an overhang and allows adapters that contain a single thymine overhang to base pair with the DNA fragments during the paired-end adaptor ligation step. The A-tailing

reaction was performed for each DNA sample by mixing the following reagents in a 0.2-mL PCR tube: 32 µL of end-repaired DNA, 5µL of NEB Buffer #2, 10 µL of 1 mM dATP, and 3 µL of Klenow fragment to a total reaction volume of 50 µL. The reaction was incubated on a thermocycler at 37°C for 30 minutes and then purified using a MinElute column as described in the DNA end-repair step. Next, the purified DNA was diluted with 19 µL of EB Buffer.

7. Paired-End Adaptor Ligation

The purpose of this step is to add adaptors to the DNA fragments of the previous step. Next, ligase binds the adaptor and inserts DNA fragments to form a complete library molecule. These are used to connect the sequence to the flow cell for sequencing. Adaptors contain indexes that are used to identify samples and permit multiplexing.

The paired-end adaptor ligation reaction was performed for each DNA sample by mixing the following reagents in a 1.5-mL tube: 19 µL of DNA sample from Step 6, 25 µL of 2× Rapid Ligation Buffer (LigaFast), 1 µL of PE Adaptor Oligo Mix (15 µm), 5 µL of T4 DNA ligase (3 U/µL) (LigaFast) to a total reaction volume of 50 µL. The reaction was incubated at room temperature for 15 minutes and cleaned with AMPure XP beads as described in the DNA purification step. The purified ligated DNA samples were diluted with 30 µL of nuclease-free water.

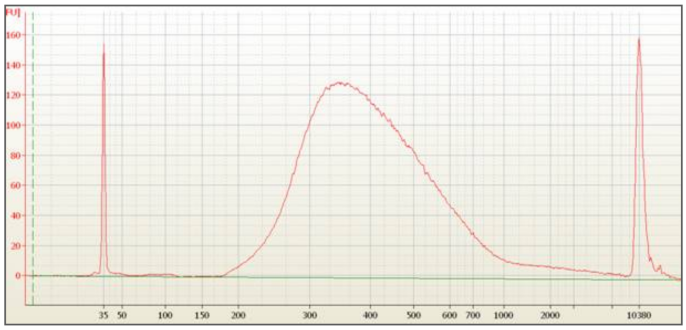
8. The Adaptor-Ligated Library Amplification

Two identical reactions were prepared for each sample in a 96 well plate. The master mix for this step was prepared as follows: 15 µL of adaptor-ligated library DNA from the previous step, 50 µL of Phusion high-fidelity PCR master mix (2×), 31 µL of nuclease-free water, 2 µL of PE-PRE1 (100 µm), and 2 µL of PE-PRE2 (100 µm) to a total reaction volume of 100 µL. The plate was placed in the thermocycler, and the libraries were amplified using the following program: 30 seconds at 98°C, 12 cycles of (10 seconds at 98°C, 30 seconds at 65°C, 30 seconds at 72°C), followed by 5 minutes at 72°C.

After amplification, two samples' reactions are combined. The amplified library was purified using one MinElute column as described in the "DNA end-repair" step and diluted with 50 µL of nuclease-free water.

The amplified pre-enrichment library's quality was checked by running 1 µL of DNA sample on a 2100 Bioanalyzer following the manufacturer's protocol. Measured library concentrations were recorded in the laboratory notebook for future reference or in case any problems arise during the following library preparation or sequencing steps.

Figure 36. An example of a good quality post ligation library.



A good quality library sample has a peak size of ~350 bp. The peaks at 25bp and 1000bp are lower and upper markers. (Reproduced from support.illumina.com)

9. Library Hybridization

4.5 µL aliquots of the SeqCap EZ Human Exome Library samples were thawed on ice. First, the hybridization mix was made by mixing the following reagents in a 1.5-mL tube: 5 µL of COT DNA (1 mg/mL), 1 µL of 1,000 µm PE-HE1, and 1 µL of 1,000 µm PE-HE2 to a total reaction volume of 9 µL. Next, each library was combined with the master mix in a separate tube, and the mixture was lyophilized in a vacuum concentrator for approximately 30 minutes at 60°C.

Following the lyophilization step, 7.5 µL of 2× hybridization buffer and 3 µL of hybridization component A were added to each sample and mixed by pipetting up and down. Then, each tube was vortexed for 10 seconds and centrifuged for 10 seconds at maximum speed. Next, the mixture was denatured in a heat block for 10 minutes at 95°C and centrifuged at maximum speed for 10 seconds at room temperature.

Denatured samples were transferred to 0.2-mL PCR tubes. 4.5 µL of the SeqCap EZ Human Exome Library was added to each denatured sample and mixed thoroughly by pipetting up and down. The tubes were incubated for 72 hours at 47°C in a thermal cycler with the heated lid set to 57°C.

10. Hybrid Capture Selection with Dynabeads M-270 Streptavidin

In this step of the library preparation, Streptavidin beads are used to purify the hybridized DNA fragments. Streptavidin beads, also known as Dynabeads, covalently bind the biotin-labeled capture probe attached to the hybridized DNA molecule while the rest of the unbound fragments were washed away. The following steps were completed before the end of the library hybridization incubation. First, for each captured library, 400 µL of 1× stringent wash buffer and 100 µL of 1× Wash buffer I were prewarmed to 47°C using heat block. At the same time, Dynabeads M-270 Streptavidin was allowed to equilibrate to room temperature for 30 minutes and mixed well by vertexing. Next, 100 µL of Dynabeads were added to each captured library, and the mix was transferred to a 1.5-mL tube. Finally, the tubes were placed on the magnetic rack, and the mixture was allowed to separate.

Next, the clear supernatant containing unbound library fragments was discarded. The Streptavidin beads with bound libraries were washed twice in the following way: tubes were removed from the magnet rack, 200 µL of 1× bead wash buffer was added to each tube. The tubes were vortexed for ten seconds and placed back on the magnetic rack. The solution was allowed to separate, and the supernatant discarded. The tubes were removed from the magnetic stand, and the beads from each library were resuspended

in 100 µL of 1× bead wash buffer. The resuspended beads were transferred to a 96-well plate (1 library = 1 well) and separated on the magnetic plate. The clear supernatant was discarded.

The hybridization mixture from the library hybridization step was then transferred to the washed beads and pipetted up and down to mix. The plate was sealed with film adhesive and incubated on the thermocycler at 47°C for 45 minutes with the thermocycler lid preset to 57°C. The samples were briefly vortexed every 15 minutes to keep the beads resuspended.

After the incubation, 100 µL of preheated to 47°C 1× wash buffer I was added to each well containing library, each reaction was transferred to a 1.5-ml tube, and the mixture vortexed for ten seconds to mix through. The beads were allowed to separate on the magnetic rack, and the clear supernatant was discarded.

The beads were washed two times with 1× stringent wash buffer as follows: 200 µL of preheated to 47°C 1× stringent wash buffer was added to the beads and mixed by pipetting up and down ten times. The mixture was incubated in the heat block for five minutes at 47°C and separated on a magnetic stand. The clear supernatant was discarded.

The beads were washed the following way: 200 µL of room-temperature 1× wash buffer was added to the beads and vortexed for 2 minutes. Tubes were placed on the magnetic rack and allowed to separate. The clear supernatant was discarded. 200 µL of room-temperature 1× wash buffer II was added to the beads and vortexed for 1 minute. Beads were separated on a magnetic stand and clear supernatant discarded. 200 µL of room-temperature 1× wash buffer III was added to the beads and vortexed for 30 seconds.

The tubes were briefly centrifuged, and the beads were allowed to separate on a magnetic rack. The supernatant was discarded when it becomes clear. The cleaned beads were resuspended in 50 µL of nuclease-free water.

11. Post-Capture Amplification

Two identical reactions are prepared for each sample in a 96 well plate. First, the following reagents were mixed to make a post-capture PCR master mix: 20 µL of bead-bound captured DNA, 50 µL of Phusion high-fidelity PCR master mix (2×), 26 µL of nuclease-free water, 2 µL of PE-POST1 (100 µm), and 2 µL of PE-POST2 (100 µm) to a total reaction volume of 100 µL. Then, the plate was placed in the thermocycler, and the libraries were amplified using the following program: 30 seconds at 98°C, 18 cycles of (10 seconds at 98°C, 30 seconds at 60°C, 30 seconds at 72°C), followed by 5 minutes at 72°C.

After amplification, two reactions for each sample were combined in a fresh 1.5-mL tube. The post-capture libraries were purified using AMPure XP beads as described in the "DNA purification" step with the following modifications: 360 µL of homogenous AMPure XP reagent were added to each reaction, and the purified DNA samples were diluted with 50 µL of QIAGEN EB buffer.

The quality of post-capture libraries was assessed by running 1 µL of each library on a high-sensitivity DNA chip using the Agilent 2100 Bioanalyzer as described in the "DNA shearing" step. The libraries showed good amplification. The peak size for an amplified library was ~350 bp. The library concentrations were recorded in the laboratory notebook for future reference.

12. Library pooling and sequencing

Library pooling was performed based on the library concentrations determined by Bioanalyzer. The final libraries were sequenced on Illumina HiSeq 2000 platform using paired-end sequencing.

5.3.3 Data analysis

Genome Analysis Toolkit (GATK) was used for exome sequence alignment and variant calling. More than 90% of the Next-Generation Sequencing reads were successfully aligned to the human reference genome. Picard Software was used to remove PCR duplicates before variant calling. Nonspecific SNPs that had been already identified in the NIH Single Nucleotide Polymorphism database (build 132) or the 1000 Genomes Project were removed based on the assumption that the mutation underlying the disease in this study was rare and thus not present in the general population. After alignment and filtering, 21 heterozygous single nucleotide variants and 2 insertion/deletions (indels) were identified that were not described in online databases of human genetic polymorphisms and were shared by both patients. Sanger sequencing of these 21 variants and 2 indels in 3 other affected individuals from the same family (patients II-2, II-3 and III-13) reduced the list to a single variant in the gelsolin gene (GSN, c.640G>T; rs121909715). The results of the Sanger sequencing examination of the 21 variants in all definitely affected family members are shown in **Table 5**. Highlighted in yellow is the gelsolin variant.

Table 5. The 21 variants shared by individuals III-1 and III-12 and the variant conformation in three other definitely affected individuals.

SNP coordinates HGRD Build 37	Function GVS	African HapMap Freq	European HapMap Freq	Asian HapMap Freq	Clinical Association	Present in dbSNP build 132	Present in 1000 Genomes	Patient III-1	Patient III-12	Patient II-2	Patient II-3	Patient III-13	Comment
chr1:152324002-152324002	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr1:152484082-152484082	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr15:68582570-68582570	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr15:70368487-70368487	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr15:74032539-74032539	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr17:38253621-38253621	missense	NA	NA	NA	unknown	no	no	het	het	wt	wt	wt	does not segregate
chr17:74093934-74093934	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	wt	does not segregate
chr18:77477922-77477922	missense	NA	NA	NA	unknown	no	no	het	het	wt	wt	het	does not segregate
chr19:18047305-18047305	missense	NA	NA	NA	unknown	no	no	het	het	wt	het	het	does not segregate
chr21:47544833-47544833	missense	NA	NA	NA	unknown	no	no	het	het	wt	wt	wt	does not segregate
chr3:13379372-13379372	missense	NA	NA	NA	unknown	no	no	het	het	het	het	wt	does not segregate
chr3:52004012-52004012	missense	NA	NA	NA	unknown	no	no	het	het	Failed	wt	wt	does not segregate
chr3:62180763-62180763	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	wt	does not segregate
chr5:143587022-143587022	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	wt	does not segregate
chr5:150920157-150920157	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	wt	does not segregate
chr8:17409321-17409321	missense	NA	NA	NA	unknown	no	no	het	het	wt	wt	wt	does not segregate
chr8:27308321-27308321	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	het	does not segregate
chr8:33246900-33246900	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	het	does not segregate
chr8:38110270-38110270	missense	NA	NA	NA	unknown	no	no	het	het	het	wt	wt	does not segregate
chr9:124073097-124073097	missense	NA	NA	NA	unknown	no	no	het	het	het	het	het	does segregate
chr9:125486365-125486365	missense	NA	NA	NA	unknown	no	no	het	het	wt	wt	wt	does not segregate

5.4 RESULTS

We utilized the advantages of whole-exome sequencing to analyze the genetic background of a large family with European ancestry in which multiple members were affected by bulbar neuropathy. In total, 18 subjects were evaluated, and five subjects were affected by bulbar-onset neuropathy. Exome sequencing was performed on two individuals from this family. More than 90% of the Next-Generation Sequencing reads were successfully aligned to the human reference genome. As described in the methods above, filtering and exome alignments yielded 21 heterozygous SNPs and 2 indels shared by both patients and not present in the NIH Single Nucleotide Polymorphism database the 1000 Genomes project databases.

Among the identified variants, we observed a variation in the Gelsolin gene (GSN, c.640G>T; rs121909715), also referred to as p.D187Y. Mutations in this gene were a known cause of hereditary amyloidosis type IV (OMIM 105120).[358] Based on these data, it rapidly became apparent that the patients did not have bulbar-onset ALS as previously thought, but rather the symptoms were consistent with hereditary amyloidosis type IV. Subsequent Sanger sequencing of those polymorphisms in three other individuals affected by bulbar-onset neuropathy confirmed that the GSN variant segregated with the disease, as it was present in all family members who showed apparent disease symptoms while absent in unaffected individuals.

5.5 DISCUSSION

Whole-exome sequencing was used in this study to identify a pathogenic GSN mutation (c.654g>t, p.Asp187Tyr) in a Caucasian-American family with a late-onset, slowly progressive bulbar neuropathy. The same mutation has been previously described as the cause of Hereditary amyloidosis type IV, also known as Hereditary gelsolin amyloidosis (HGA) and Familial amyloidosis, Finnish type. This syndrome is an autosomal dominant disorder characterized primarily by bulbar signs, corneal lattice dystrophy, progressive cranial neuropathy, and skin changes [359], symptoms that are consistent with the clinical picture observed within our family. HGA was initially described in Finland in 1969.[359] Since then, HGA patients have been increasingly reported from the other parts of the world, including Asia, North and South Americas and the Middle East.[360-363]

GSN is a ubiquitously present protein with amyloid properties.[358] Biochemically, it can bind to actin filaments and impact their capping and uncapping reactions, thus impacting cellular structures by influencing actin dynamics within the cell.[364] It is a calcium-binding protein. Upon binding of Ca^{2+} and actin, Gelsolin undergoes various conformational changes and the subsequent conformational changes that allow GSN to interact with actin molecules. Alternative splicing of Gelsolin gene produces cytoplasmic and secreted forms of the GSN protein. Previous research has demonstrated that amyloidogenic fragments are produced by aberrant processing of just the secreted form of gelsolin, also known as plasma gelsolin.[365]

The p.Asp187Tyr/Asn mutation in domain 2 of GSN gene is the most common cause of HGA. The mutation eliminates one of plasma gelsolin's four calcium-binding ligands, preventing the protein from effectively binding calcium and enabling mutant plasma gelsolin domain 2 to reach unfolded conformations. This results in the improper proteolytic processing of the protein in the Golgi apparatus and leads to the formation of 8 and 5 kDa amyloidogenic fragments that adhere to basement membranes, impairing neurons' function and physiology, cutaneous adnexa, elastic fibers, and small arterioles.[366,367,368-371] This type of amyloidotic deposits is seen as one of the most likely causes for hereditary amyloidosis type IV.[372] p.Asp187Tyr/Asn mutation has a complete penetrance in HGA patients, with only one allele being enough to cause the disease. As a result, it has been suggested that G654T leads to the development of HGA through the gain-of-function mechanism.[373]

While the underlying pathogenic mechanism of p.Asp187Asn and p.Asp187Tyr mutations is likely the same, it has been previously mentioned that the two mutations may result in different phenotypes. Danish and French carriers of p.Asp187Tyr carriers show earlier onset of bulbar and cardiac signs than Finnish p.Asp187Asn carriers. Furthermore, there have been reports of p.Asp187Tyr mutation resulting in severe phenotypes that included vocal cord paralysis and the need for mechanical ventilator support in Danish HGA patients.[374]

Hereditary amyloidosis type IV is not easy to correctly diagnose since it is rare and bulbar symptoms are not readily recognized as amyloidosis by most physicians. Interestingly, three strongly affected family members from this study first presented with tongue atrophy and fasciculation, which led doctors to consider familial ALS as a diagnosis, despite the early absence of bulbar neurological syndromes. It is very likely

that without tongue fasciculation and the possibility of bulbar-onset ALS, this disease would have gone undetected.

ALS is a complex neurological disorder. Although the main criteria for this disorder are outlined in El Escorial criteria, it is sometimes difficult to diagnose because ALS can mimic other neurological disorders such as cerebrovascular disease, cervical myelopathy, vertebral disc herniation, radiculopathy, neuropathy, and myasthenia gravis. It is currently estimated that up to 15% of ALS cases are misdiagnosed. (www.als.org) Misdiagnosis can potentially have a negative effect on the large sequencing efforts such as Project MinE that aim to identify genes associated with ALS and to provide insight in the "omics" of this neurological disorder. Considering ALS misdiagnosis rates, there is significant potential for disease misclassification in GWAS phenotype data where even small numbers of these errors can have significant impact on GWAS statistical power. Various strategies have been proposed to help address these issues through the processing of GWAS phenotype data including a mixture model for sub-phenotyping, a multivariate genome-wide association test for analyzing data on multivariate quantitative phenotypes and a joint model of multiple phenotypes.[375-378] More recent GWAS analysis frameworks such as Phenotype Latent variable Extraction of disease misdiagnosis (PheLEx) can learn and are able to correct misclassified phenotypes using structured genotype associations within a dataset.[379]

6 INVESTIGATING RFC1 REPEAT EXPANSIONS IN SPORADIC AMYOTROPHIC LATERAL SCLEROSIS

6.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

The manuscript describing this project was submitted to the Journal of the Neurological Sciences and is currently under review.

I am a joint first author on this manuscript. I participated in the study design, performed the lab work, data analysis, and drafted the manuscript. Ramita Dewan, Andrea Cortese, Henry Houlden, and Bryan Traynor participated in the study design. Ramita Dewan performed the experiments, data analysis and assisted in writing the manuscript. Andrea Cortese, Henry Houlden, Luigi Ferrucci, and Bryan Traynor reviewed the manuscript for critical content.

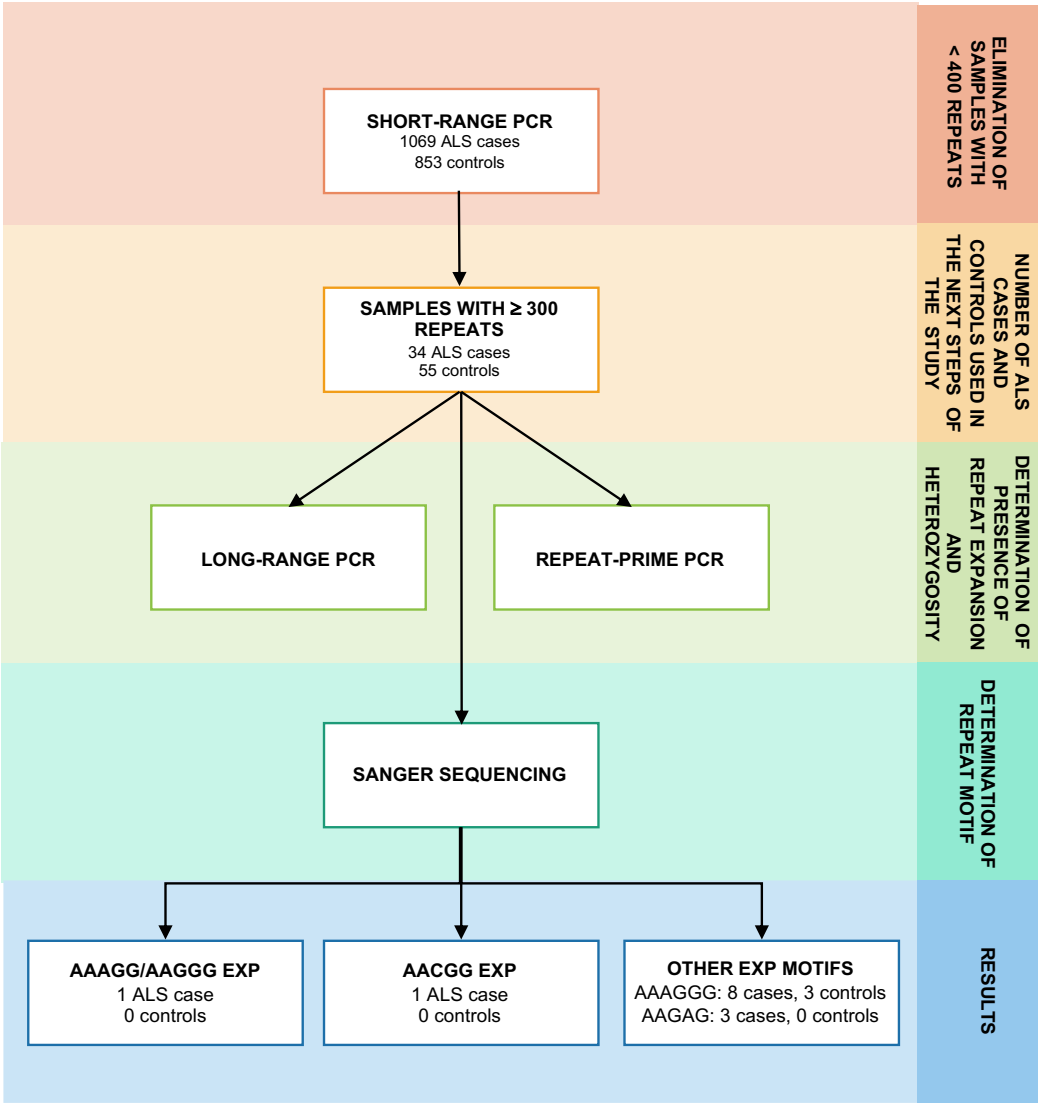
6.2 BACKGROUND

Replication factor C, subunit 1 (RFC1) encodes the large subunit of replication factor C, a five subunit, DNA polymerase accessory protein, a DNA-dependent ATPase required for DNA replication and repair. The protein acts as a DNA polymerase activator by binding to the 3' end of primers and promoting both strands' coordinated synthesis.[380,381] It may also have a role in the maintenance of telomere length and telomerase protein stability.[382] Cortese and colleagues recently described a recessive

AAGGG repeat expansion within intron 2 of this gene as a cause of cerebellar ataxia, neuropathy, vestibular areflexia syndrome (CANVAS).[383] We examined the role of RFC1 repeat expansions in patients diagnosed with sporadic ALS, based on the phenotypic overlap between the condition and CANVAS with regard to motor neuron neuropathy, the genetic pleiotropy known to occur in the disease, and the importance of other large repeat expansions, such as C9orf72, in ALS pathogenesis.

6.3 METHODS

Figure 37. Workflow of the study.



6.3.1 Patient Information

A total of 1,069 sporadic ALS patients and 853 matched neurologically healthy individuals were screened for the presence of repeat expansions in the RFC1 gene (chr4:39,287,455-39,366,380). Genomic DNA was extracted from whole blood samples obtained from the Coriell repository (www.coriell.org) for ALS patients and the (Baltimore Longitudinal Study of Aging, USA) for control subjects. All study participants were of European ancestry. ALS patients' and control subjects' demographic data are summarized in **Table 6**. All participants gave written informed consent, and the study complied with relevant ethical regulations.

Table 6. Demographic data for sporadic ALS cases screened for RFC1 repeat expansion.

	ALS (n = 1069)	Healthy controls (n = 853)
Age at onset years	57 (12.2)	N/A
Gender (male)	623 (58.3%)	467 (54.7%)
C9orf72 repeat expansion carriers	60 (5.6%)	N/A
Onset site		
Bulbar	247 (23.1%)	N/A
Limb	774 (72.4%)	N/A
Other	48 (4.5%)	N/A

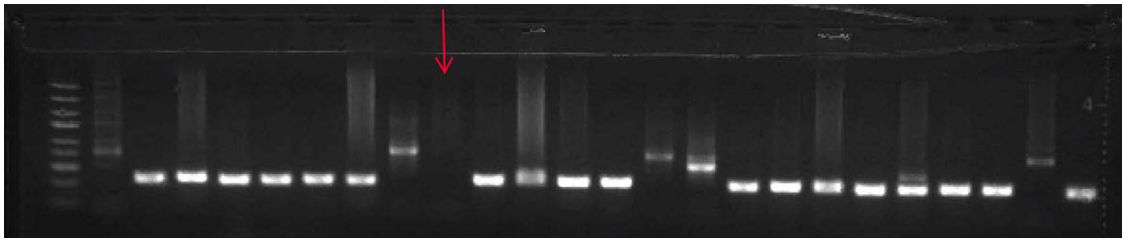
The cohorts examined in this study consisted of 1,069 USA ALS cases and 853 matched neurologically normal USA controls. Age at onset data was not available for 23 cases and was not applicable to control subjects. The site at onset data was not available for 45 cases. Family status was not available for 15 cases. C9orf72 repeat expansion carrier data was not available for 32 cases. Ethnicity data were not available for 5 cases. N/A = not applicable.

6.3.2 Standard flanking PCR

DNA quality and concentration were quantified by NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific) and Qubit 4 Fluorometer (ThermoFisher Scientific). Standard flanking PCR primers, reagents, and cycling protocol were identical to those described by Cortese et al. [383] PCR products were run on a 1% agarose gel (SeaKem® LE Agarose, 30 minutes at 115 volts and 500 microamperes) and analyzed for the presence of PCR-amplified products corresponding to the $[AAAAG]_{11}$ reference allele.

Standard flanking PCR protocol uses Roche FastStart PCR Master mix (Sigma), which can only amplify DNA segments up to 2 kilobases (kb) in length. This limitation of the standard flanking PCR was utilized to identify samples that did not amplify during this step and could potentially harbor homozygous expansions larger than 400 repeats (2 kb) in size. An example standard flanking PCR gel image is shown in **Figure 38**.

Figure 38. Standard flanking PCR.



The gel demonstrates the absence of PCR-amplifiable product during the first step of RFC1 screening (marked by red arrow).

6.3.3 Gel electrophoresis

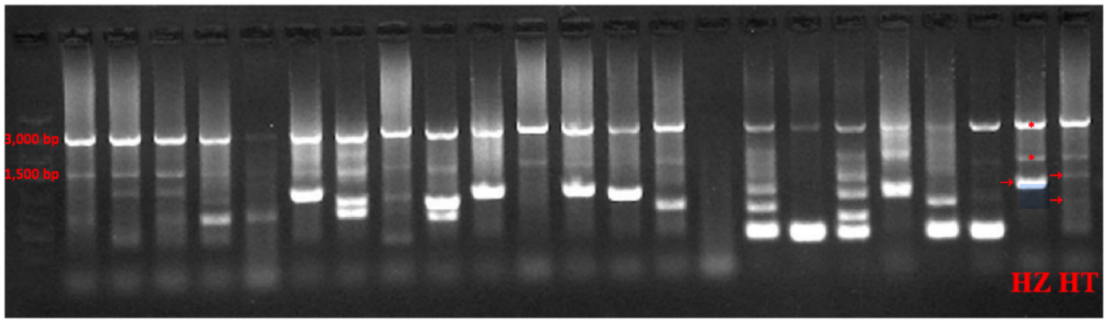
The standard flanking PCR products were run on 1% agarose gel (SeaKem® LE Agarose) for 30 minutes at 115 V and 500 mA and analyzed using the UVP ChemiDoc-It2 Imaging System (ThermoFisher Scientific) for the presence of PCR-amplified products corresponding to the [AAAAG]11 reference allele.

6.3.4 Long-range PCR

Long-range PCR was performed on all samples lacking PCR-amplifiable products on standard flanking PCR (n=89). Long-range PCR primers, reagents, and cycling protocol were identical to those used by Cortese et al. [383]

The long-range PCR protocol uses Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Beverly, MA) and longer extension times, allowing for amplification of DNA segments up to 20 kb in length. This step was used to detect the presence of a large repeat expansion and determine sample homozygosity/heterozygosity status. Like the standard flanking PCR protocol, long-range PCR products were run on a 1% agarose gel. Samples were excluded from potentially harboring homozygous mutant alleles or expansions if long-range PCR revealed products at the reference size (at least one allele with 11 pentanucleotide repeats). **(Figure 39)**

Figure 39. Long-range PCR demonstrating sample homozygosity vs. heterozygosity.



Samples with one band are homozygous, and samples with two non-artifact bands are heterozygous. Artifact bands were observed at ~1,500bp and 3,000bp. These bands were present in all samples and did not correspond to any previously described product size. Therefore, they were ignored during the analysis.

6.3.5 Verification by RPPCR

Repeat-primed PCR was performed on all samples lacking PCR-amplifiable products on standard flanking PCR (n=89). The repeat-primed PCR assay was performed for each of three pentanucleotide repeat units [AAAAG/AAAGG/AAGGG] using a previously validated repeat-primed PCR method described by Cortese et al.[383] The repeat-primed PCR assay was performed using the Eppendorf Mastercycler pro thermal cycler (Fisher Scientific) in a final reaction volume of 20µl containing Phusion Flash High-Fidelity PCR Master Mix 2X (Thermo-Fisher), DMSO 3% (Qiagen), 100ng/µl genomic DNA, 5 µM primer mix containing 0.5 µmol/L of forward primer 5′-TCAAGTGATACTCCAGCTACACCGT-3′ with a 5' 6-FAM fluorescent tag, 0.5 µmol/L anchor primer 5′- CAGGAAACAGCTATGACC -3′ (Eurofins Genomics), and

one of the reverse primer mixes for each of the RFC1 alleles. The primers are listed in **Table 7**.

Table 7. Primers used for RFC1 screening.

Short-range PCR	Fw: TCAAGTGATACTCCAGCTACACCGTTGC Rv: GTGGGAGACAGGCCAATCACTTCAG
Long-Range PCR	Fw TCAAGTGATACTCCAGCTACACCGTTGC Rv GTGGGAGACAGGCCAATCACTTCAG
Repeat-primed PCR	Fw FAM-TCAAGTGATACTCCAGCTACACCGT Anchor CAGGAAACAGCTATGACC **(AAAAG)11 allele** Rv1 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAAGAAAAGAAAAGAAAAGAAAA Rv2 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAGAAAAGAAAAGAAAAGAAAA Rv3 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAGAAAAGAAAAGAAAAGAAAA **(AAAGG)exp allele** Rv1 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAAGGAAAGGAAAGGAAAGGAAA Rv2 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAGGAAAGGAAAGGAAAGGAAA Rv3 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAGGAAAGGAAAGGAAAGGAAA **(AAGGG)exp allele** Rv1 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAAGGGAAGGGAAGGGAAGGGAA Rv2 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAAGGGAAGGGAAGGGAAGGGAA Rv3 CAGGAAACAGCTATGACCAACAGAGCAAGACTCTGT TTCAAAAGGGAAGGGAAGGGAAGGGAA

The fragment length analysis was performed on an ABI 3730xl genetic analyzer (Applied Biosystems), using a mixture of 2µl of repeat-primed PCR products, 0.5µl GeneScan 500 LIZ size standard (Thermo Fisher Scientific), and 7.5µl HiDi formamide (Applied Biosystems). The mixture was heated at 95°C for 3 minutes and then immediately cooled on ice for five minutes before loading for capillary electrophoresis. Repeat-primed PCR products were separated on an ABI3730xl DNA Analyzer (Applied Biosystems). The results were visualized using GeneMapper®v.4.0 (Applied Biosystems). Repeat-primed PCR can help determine if a sample carries a large repeat expansion and can confirm the repeat unit. However, this method cannot accurately measure repeat size for large repeat expansions due to signal drop-off after approximately 1 kb (200 repeats).

6.3.6 Sanger sequencing

Sanger sequencing of long-range PCR products was performed to confirm allelic composition in samples carrying repeat expansions using the BigDye Terminator v3.1 chemistry (ThermoFisher Scientific). Sequence analysis was performed on an ABI 3730xl DNA Analyzer (ThermoFisher Scientific), using Sequencher software (version 4.2, Gene Codes). Additionally, Sanger sequencing was used to identify unusual repeat motifs and insertions in repeat expansions.

6.3.7 Data analysis

Allele frequencies of ALS cases and control subjects were compared using chi-squared tests with Bonferroni correction, performed in R studio, version 1.1.463 as implemented in the "stats" package.

6.4 RESULTS

A total of 1,069 sporadic ALS patients and 853 neurologically healthy control individuals were screened for the presence of an expansion at the RFC1 locus. The workflow previously described by Cortese et al.[383], in which all samples were first screened by standard flanking (short-range) PCR to determine if a sample carried two non-expanded copies of the RFC1 gene was adopted for this study. Samples that did not contain two reference alleles were further analyzed using long-range PCR, repeat-primed PCR, and Sanger sequencing to determine the nature of their RFC1 repeat expansion.

Following standard flanking PCR, 89 samples in the control (n = 55) and ALS (n = 34) cohorts were selected for further analysis, based on the absence of PCR amplifiable products (6.45% and 3.18%, respectively). Combined analysis on these selected samples using long-range PCR, RP-PCR, and Sanger sequencing revealed a highly dynamic genetic locus with multiple repeat motifs and heterozygous genotypes. The summary and description of alleles observed in this study are provided in **Table 8**.

Table 8. Repeat expansion motifs in ALS cases and controls.

	ALS			Controls		
	Number of samples			Number of samples		
REPEAT	Hom	Het	Total	Hom	Het	Total
(AAAAG)exp	16	8	24	38	7	45
(AAAGG)exp	4	6	10	0	1	1
(AAGGG)exp	0	2	2	0	0	0
(AAAGGG)exp	8	0	8	2	1	3
(AAGAG)exp	2	1	3	0	0	0
(AACGG)exp	0	1	1	0	0	0
(AAAGG)n(AAAAG)exp	0	1	1	0	1	1
(AAGGG)n(AAAAG)exp	0	1	1	0	1	1
(AAGAG)n(AAAAG)exp	0	1	1	0	0	0
(AAGGG)n(AAAGG)exp	0	1	1	0	0	0
(AAAGG)n(AAAGGG)exp	0	2	2	4	3	7
(AAGGG)n(AAAGGG)exp	0	2	2	0	3	3

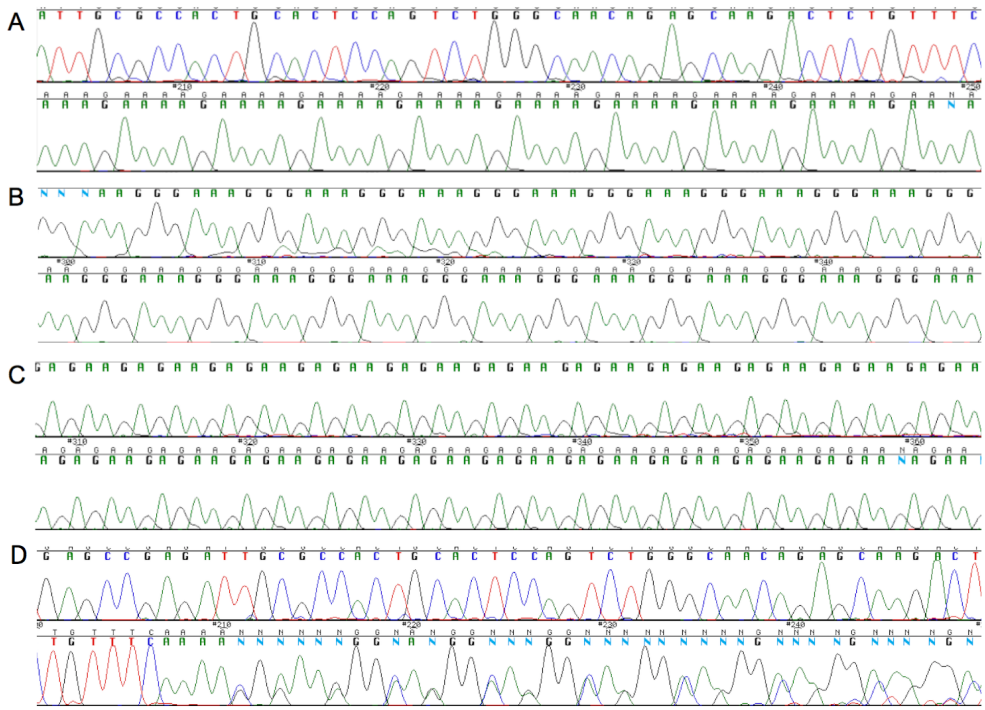
The homozygous [AAGGG]n expansion previously described as causative for CANVAS was not identified in any of our ALS cases or control subjects. Compound heterozygotes, as [AAAGG]/[AAGGG], was observed in one ALS case, but not in control subjects. This heterozygous repeat was described by Cortese et al. to be present in the normal population at the rate of 0.7%. [383].

Our cohorts revealed the presence of two previously described repeat expansion conformations [AAAGGG]exp and [AAGAG]exp. The AAAGGG expansion was present in both homozygous (n = 2) and heterozygous (n = 1) states in our control cohort. In comparison, it was only found in the homozygous state in our ALS cases (n =8). The AAGAG repeat expansion was present in both homozygous (n = 2) and heterozygous (n = 1) states in our ALS cohort but was not present in controls. The previously reported AGAGG repeat expansion was not observed in neither of our cohorts.

We discovered several ALS cases in which the repeat expansion changed its repeat unit throughout the Sanger sequencing read. These include: $[AAAGG]_n/[AAGGG]_n$ that converted to $[AAAGGG]_{exp}/[AAAGGG]_{exp}$ (ALS = 2, controls = 3) $[AAAAG]_{exp}/[AAGAG]_n$ converting to $[AAAAG]_{exp}/[AAAAG]_{exp}$ (ALS = 1, controls = 0); and $[AAAAG]_{exp}/[AAAGG]_n$ converting to $[AAAAG]_{exp}/[AAAAG]_{exp}$ (ALS = 1, controls = 1).

Additionally, a novel motif of RFC1 repeat [AACGG]n was discovered. This motif was present in a heterozygous state in a single ALS case but was not found in controls (**Figure 40**). Furthermore, we observed two types of small insertions in samples with RFC1 expansions. The first type is a single nucleotide thymine insertion after the first eleven AAAAG repeats. The second type was a 36-nucleotide insertion after the first 11 repeats. This larger insertion sequence BLATs to part of the LTR region on chromosome 4. These conformational changes were observed only in ALS patients and not in control subjects.

Figure 40. Confirmation of alternative RFC1 repeat motifs by Sanger sequencing.



A) Reference [AAAAG]11, B) alternative conformation [AAAGGG]n, C) alternative conformation [AAGAG], D) discovered in our study conformation [AACGG]n.

6.5 DISCUSSION

The recessive AAGGG expansion within the second intron of the RFC1 gene was recently reported by Cortese et al. to cause late-onset ataxia and CANVAS.[383] We performed a follow-up study to examine the role of RFC1 repeat expansion in patients diagnosed with sporadic ALS. To do this, 1,069 sporadic ALS cases and 853 US control subjects were screened following the workflow described by Cortese et al.[383] Homozygous RFC1 AAGGG repeat expansions were not observed in either ALS cases or controls, indicating that this repeat expansion is not a common cause of sporadic ALS in the United States.

While the results of this study do not support a role for the RFC1 repeat expansion in ALS, they expand on the highly dynamic nature of the RFC1 locus.[384] To date, four different repeat conformations have been observed in the general population: a wild-type sequence AAAAG (11 repeats) and longer expansions of AAAAG, AAAGG, and AAGGG sequences.[383,384] Akcimen and colleagues reported two alternative RFC1 repeat variants AAGAG and AAAGGG.[384] We confirmed the presence of RFC1 expansions containing these alternative repeat units in our cohorts. Furthermore, we discovered a novel RFC1 repeat variant [AACGG]exp that was present in 1.8% of our ALS cases that did not have an amplifiable product on short-range PCR. While the exact size of these repeat expansions is unknown due to the Southern blot's unavailability, all expansions were longer than 400 repeats because they were based on our PCR method.

This study has several limitations. First, RP-PCR, long-range PCR, and Sanger sequencing were not performed on all samples and reserved this pipeline only for the samples with two expanded alleles. Without heterozygous cases it is difficult to determine the frequency of any specific expansion in the population. Some samples may have had a single band on short-flanking PCR but may have still harbored an expanded allele. Therefore, findings reported in this study were not overall allele frequencies but showed the percentage of different RFC1 repeat expansion motifs in a subset of samples that failed to amplify during short-range PCR due to the large repeat expansion size.

Second, Cortese et al. determined the pathological cutoff for CANVAS to be 400 AAGGG repeats [383]. Consequently, our analysis focused on samples that had greater than 400 repeats. Inability to profile expansions smaller than 400 repeats is big limitation because, once again, it prevented us from establishing the allelic frequencies in our cohorts. Moreover, several repeat expansion loci are known to have different neurological phenotypes associated with large and intermediate repeat expansions. For example, full-length repeat expansions of FMR1 (greater than 200 CAG repeats) are associated with intellectual disability, autism, facial dysmorphism, and attention deficit hyperactivity disorder.[385] In comparison, premutation alleles (55-200 repeats) are related to other clinical phenotypes such as fragile X-associated tremor/ataxia syndrome, fragile X-associated primary ovarian insufficiency in females, depression, and anxiety.[385] Based on these previous findings, it would be interesting to examine the role of intermediate size RFC1 repeat expansions in ALS patients.

Third, all of our study participants were of European ancestry. It has been previously reported that the occurrence of repeat expansions at the same locus may vary across

different populations. For example, C9orf72 pathogenic repeat expansions are common among individuals with European ancestry but are rare among Asian populations [386-392,108]. It may be beneficial to examine regional variation at the RFC1 locus. Finally, this study focused on sporadic cases, and studies examining the role of RFC1 expansions in familial ALS would be of interest.

Future studies may utilize the benefits of single-molecule long-read sequencing offered by Pacific Bioscience (PacBio) and Oxford Nanopore Technologies (ONT). Recent studies show that long-read sequencing is well suited to characterizing known repeat expansions and discovering new disease-causing, disease-modifying, or risk-modifying repeat expansions that have gone undetected with conventional short-read sequencing.[393,394] Long-read sequencing studies in RFC1 expansion carriers will help determine heterogeneity and its impact on ALS onset and course.

Taken together, the results of this study do not show substantial evidence to support a role for the RFC1 repeat expansion in ALS. However, a greater understanding of the role of intermediate size RFC1 repeat expansions is needed.

6.6 Oxford Nanopore experiment

Inability to determine the size and composition of the large RFC1 repeat expansions in our ALS and control cohorts is a major limitation of this project. This information would provide us with allelic frequencies and would allow us to determine if specific allele sizes or repeat motifs could be associated with higher risk of developing ALS. Southern blot has been the gold standard for expanded allele detection and sizing for many years. Since we don't have a Southern blot in our laboratory, I was planning to

use Oxford Nanopore sequencing technology as an alternative method for repeat expansion measurement. This method has previously shown good precision in repeat expansion sizing since it produces single molecule sequencing reads often more than 100 kb in length, which in many cases eliminates the need for sequence assembly. Unfortunately, due to the COVID-19 restrictions, I was unable to complete this part of the project.

Nanopore sequencing, to determine the RFC1 allele size and composition, could be performed in three different ways: using super-long-read sequencing, CRISPR Cas9-mediated target enrichment or Read Until enrichment. Each of these methods uses a different approach to data collection.

In ultra-long-read Nanopore sequencing whole genome data is collected for each sample and the region of interest (ROI), which includes RFC1 repeat expansions, is analyzed. This is a labor-intensive method, which consists of three consecutive steps: 1) extraction of high-molecular weight (HMW) genomic DNA; 2) ultra-long DNA library construction and 3) loading of the adapter-ligated DNA onto nanopores flow cells.[395] Collected genome data is then processed using the following algorithms: Shasta - a de novo long read assembler, and MarginPolish & HELEN - a suite of nanopore assembly polishing algorithms.[396]

CRISPR-Cas9 is a method designed for targeted nanopore sequencing. In the first step of library preparation, sample DNA is dephosphorylated to prevent ligation. Next, Cas9 (CRISPR associated protein 9) cleaves the DNA at specific sites on both ends of the region of interest (ROI), exposing ligatable ends. The 3’ ends are A-tailed and

sequencing adapters are ligated only to the cleaved ends. The complete libraries are then loaded to the MinION flow cell and sequenced.

This method requires a custom panel of guide RNAs, which can be designed using the custom Alt-R™ CRISPR-Cas9 crRNA design tool. Ribonucleoprotein (RNP) complexes can be constructed by combining the guide RNA, composed of custom Alt-R CRISPR-Cas9 crRNA and tracrRNA, with Alt-R HiFi Cas9 nuclease. The resulting data would be analyzed by SAMtools [397] and nanopolish [398] data analysis tools following the previously described pipeline [399].

Read Until is another targeted Nanopore sequencing method that allows real-time selective sequencing of the ROI selected for the project. In this method, MinION reads the first four hundred base pairs of the sequence that enters the pore and if it does not match the ROI, the nanopore selectively ejects the read from the pore by reversing the polarity of the voltage across the specified pore for a short period of time. This allows the pore to proceed to the next DNA molecule faster, considerably increasing the efficiency of nanopore sequencing.[400]

In the past, Read Until sequencing libraries were prepared using the ONT Ligation Sequencing Kit (SQK-LSK109) without the DNA Control Strand (DCS) or FFPE repair in the end prep step.[400] The final sequencing data would be analyzed using UNCALLED algorithm that matches streaming nanopore current signals to a reference sequence.[400,401] Current version of Nanopore Read Until sequencing does not require sample preparation prior to sequencing. The sequencing data is analyzed in real-time with RUBRIC software that enables alignment against conventional nucleic acid references to provide the basis for sequence/reject decisions.[402]

When choosing between the methods described above, one must consider the following criteria: amount and quality of DNA needed for the experiment and time required for sample preparation, sequencing and analysis. Unrestricted ultra-long read nanopore sequencing requires a total of 5µg of DNA. It typically takes 6 hours to complete HMW DNA extraction, from 1.5 hours to 8 hours for library construction, depending on the shearing method, and up to 48 hours for DNA sequencing, and more than 28 hours for analysis.[395] However, unlike CrispR-Cas9 and Read Until, this method provides data, that can be used not only for repeat expansion sizing and motif confirmation but also to gain insight into other structural variations. CRISPR/Cas9 nanopore protocol requires approximately 3µg of genomic DNA. HMW DNA extraction process takes approximately 1.5 hours, and library preparation is approximately 2 hours. However, data analysis with samtools and nanopolish tools might take up to 48 hours. Read Until targeted sequencing method needs 5µg of total DNA. It is perhaps the most time efficient method since it does not require time for library preparation, only for DNA extraction and fragmentation, and the data is analyzed in real time.

7 EXAMINATION OF INTERMEDIATE ALLELES AS A PART OF WHOLE-GENOME SEQUENCING OF LEWY BODY AND FRONTOTEMPORAL DEMENTIAS PROJECT

7.1 STATEMENT OF CONTRIBUTION TO THIS RESEARCH

DementiaSeq was a large international project that aimed to perform whole-genome sequencing of Lewy body and frontotemporal dementias. I participated in the study design, performed the lab work, and carried out some of the data analysis focused on intermediate length repeat expansions. There are currently two publications arising from this work:

Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis.[403]

Dewan R, Chia R, Ding J, Hickman RA, Stein TD, **Abramzon Y**, Ahmed S, Sabir MS, Portley MK, Tucci A, Ibáñez K, Shankaracharya FNU, Keagle P, Rossi G, Caroppo P, Tagliavini F, Waldo ML, Johansson PM, Nilsson CF; American Genome Center (TAGC); FALS Sequencing Consortium; Genomics England Research Consortium; International ALS/FTD Genomics Consortium (iAFGC); International FTD Genetics Consortium (IFGC); International LBD Genomics Consortium (iLBDGC); NYGC ALS Consortium; PROSPECT Consortium, Rowe JB, Benussi L, Binetti G, Ghidoni R, Jabbari E, Viollet C, Glass JD, Singleton AB, Silani V, Ross OA, Ryten M,

Torkamani A, Tanaka T, Ferrucci L, Resnick SM, Pickering-Brown S, Brady CB, Kowal N, Hardy JA, Van Deerlin V, Vonsattel JP, Harms MB, Morris HR, Ferrari R, Landers JE, Chiò A, Gibbs JR, Dalgard CL, Scholz SW, Traynor BJ.

Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture.[404]

Chia R, Sabir MS, Bandres-Ciga S, Saez-Atienzar S, Reynolds RH, Gustavsson E, Walton RL, Ahmed S, Viollet C, Ding J, Makarious MB, Diez-Fairen M, Portley MK, Shah Z, **Abramzon Y**, Hernandez DG, Blauwendraat C, Stone DJ, Eicher J, Parkkinen L, Ansorge O, Clark L, Honig LS, Marder K, Lemstra A, St George-Hyslop P, Londos E, Morgan K, Lashley T, Warner TT, Jaunmuktane Z, Galasko D, Santana I, Tienari PJ, Myllykangas L, Oinas M, Cairns NJ, Morris JC, Halliday GM, Van Deerlin VM, Trojanowski JQ, Grassano M, Calvo A, Mora G, Canosa A, Floris G, Bohannan RC, Brett F, Gan-Or Z, Geiger JT, Moore A, May P, Krüger R, Goldstein DS, Lopez G, Tayebi N, Sidransky E; American Genome Center, Norcliffe-Kaufmann L, Palma JA, Kaufmann H, Shakkottai VG, Perkins M, Newell KL, Gasser T, Schulte C, Landi F, Salvi E, Cusi D, Masliah E, Kim RC, Caraway CA, Monuki ES, Brunetti M, Dawson TM, Rosenthal LS, Albert MS, Pletnikova O, Troncoso JC, Flanagan ME, Mao Q, Bigio EH, Rodríguez-Rodríguez E, Infante J, Lage C, González-Aramburu I, Sanchez-Juan P, Ghetti B, Keith J, Black SE, Masellis M, Rogaeva E, Duyckaerts C, Brice A, Lesage S, Xiromerisiou G, Barrett MJ, Tilley BS, Gentleman S, Logroscino G, Serrano GE, Beach TG, McKeith IG, Thomas AJ, Attems J, Morris CM, Palmer L, Love S, Troakes C, Al-Sarraj S, Hodges AK, Aarsland D, Klein G, Kaiser SM, Woltjer R, Pastor P, Bekris LM, Leverenz JB, Besser LM, Kuzma A, Renton AE, Goate A, Bennett

DA, Scherzer CR, Morris HR, Ferrari R, Albani D, Pickering-Brown S, Faber K, Kukull WA, Morenas-Rodriguez E, Lleó A, Fortea J, Alcolea D, Clarimon J, Nalls MA, Ferrucci L, Resnick SM, Tanaka T, Foroud TM, Graff-Radford NR, Wszolek ZK, Ferman T, Boeve BF, Hardy JA, Topol EJ, Torkamani A, Singleton AB, Ryten M, Dickson DW, Chiò A, Ross OA, Gibbs JR, Dalgard CL, Traynor BJ, Scholz SW.

In this chapter, I present a brief overview of the dementiaSeq project and then focus on the intermediate-length project which I was primarily responsible for.

7.2 BACKGROUND

7.2.1 Cohorts

Our cohort originally consisted of 3,000 individuals diagnosed with FTD, 3,000 individuals diagnosed with LBD, and 3,000 neurologically normal control individuals. Detailed phenotype data are available for each sample, including the diagnosis, age at symptom onset, gender, and clinical course. Each sample's quality and concentration have been determined using Nanodrop spectrophotometry and Quant-iT PicoGreen assay (ThermoFisher). Only DNA from individuals with European ancestry was used in this project. We did not use DNA derived from lymphoblast cell lines due to concerns about cytogenetic rearrangements.

Here, I describe the background of the dementiaSeq project including the cohort selection, sequencing methodology, and data realignment.

7.2.2 Lewy body dementia

Lewy body dementia (LBD) is the second most common form of dementia in the elderly population affecting approximately 1.3 million people in the United States.[405] The condition is characterized by progressive cognitive decline, visual hallucinations, parkinsonism, and fluctuating mental status.[406] The pathological features of LBD include Alzheimer-type pathology and widespread Lewy neurites and Lewy bodies, consisting of aggregated alpha-synuclein protein fibrils.[407] Similar to other neurodegenerative diseases, a small proportion of familial cases have been described suggesting the existence of a genetic predisposition to this condition.[408-411]

7.2.3 Frontotemporal dementia

Frontotemporal dementia (FTD) is described in detail in the "Introduction" section of this thesis. In brief, FTD is a neurodegenerative disorder of the frontal and anterior temporal lobes of the brain. FTD is one of the most common types of presenile dementias, representing 10%-20% of all dementia cases.[412] The syndrome is characterized clinically by initial behavioral and language disturbances, followed by cognitive decline leading to dementia and death within a median of seven years from symptom onset. Similar to other neurodegenerative diseases, a large proportion of FTD cases are categorized as sporadic, and the causes of this idiopathic form of the disease are largely unknown. [413,414]

7.3 Project objectives

The objective of the dementiaSeq project is to perform genome sequencing in large cohorts of patients diagnosed with LBD, FTD, as well as neurological normal control subjects; to analyze these data to identify new genetic risk loci; and to make the raw sequencing data publicly available on the dbGaP (accession number phs001963, available at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001963.v1.p1).

The genetic analyses of the whole-genome sequence data collected in this project will likely lead to the identification of new genetic loci responsible for these diseases. Each new gene linked to the etiology of LBD and FTD adds to our understanding of the cellular pathways that cause neurodegeneration. Their discovery also aids in disease modelling as well as the development and testing of targeted therapeutics.

7.4 METHODS

7.4.1.1 LBD samples

Most of our LBD samples came from autopsies and were pathologically confirmed cases according to the McKeith classification.[405] The use of a large portion of the pathologically defined LBD cohort was a significant strength of this project. It overcomes diagnostic uncertainty that arises in this group of patients due to mimic syndromes. Approximately 1,600 of these samples were from American subjects, and the remainder were from the United Kingdom.

7.4.1.2 FTD samples

FTD samples were obtained from patients diagnosed according to the Neary criteria.[415] C9orf72 status is known for all samples using a repeat primed PCR assay. We prioritized the inclusion of familial FTD cases where the causative mutation was unknown.

7.4.1.3 In-house neurologically normal samples

The control cohort consisted of 250 neurologically normal samples from the United Kingdom and 1,250 normal American samples obtained from the Baltimore Longitudinal Study of Aging (BLSA, blsa.nih.gov). Genome-wide genotype data were available for all BLSA and UK controls and used to ensure the genome sequence data's quality.

The majority of the British control genome sequence data for this project came from external sources. For this reason, the genome sequence data derived from these 250 in-house samples is helpful in the identification (and rectification) of technical artifacts.

BLSA has been active for over 50 years and is one of the most extended ongoing longitudinal studies of aging globally. It focuses on the physical and cognitive changes associated with aging, aiming to identify the genetic, physical, behavioral, and environmental factors that affect the aging process. Unprecedented demographic and clinical data have been collected for these subjects since 1958. Control subjects may ultimately develop dementia as they age. To mitigate this possibility, we only included

genome sequence European-ancestry subjects over the age of 60 and also incorporated age as a covariate in the analysis.

7.4.1 Genome sequencing

Genome sequencing was performed using the Illumina TruSeq DNA library preparation kit, and the HiSeq X10 sequencing system produces 150-base pair paired-end reads according to the manufacturer's protocol. All of the whole-genome sequencing was performed at a single location (The American Genome Center, Uniformed Services University, Bethesda, MD) to improve inter-sample variant calling. The X10 system uses patterned flow cells for increased cluster density. This method routinely achieves 35-fold coverage across 90% of the genome for each sample (a vital metric of the technical success of sequencing that allows for accurate variant calling).

7.4.2 Sequence alignment

Raw sequence data were processed from BAM/FASTQ using the current implementation of the Broad Best Practices API (application program interface) hosted by Google Genomics on the Google Cloud platform. This includes the industry-standard GATK 3.4.0 pipeline on default settings, batched re-genotyping and re-calibration, and leveraging background VCF files to harmonize datasets and reduce the likelihood of technical artifacts.

EXAMINATION OF THE INTERMEDIATE SIZE REPEAT EXPANSIONS IN ALS/FTD AND LBD COHORTS

Here, I focus the rest of the chapter on the intermediate length repeat expansion analysis.

7.5 BACKGROUND

Repeat expansion disorders are genetic disorders caused by expansions of short tandem repeats located within human genes. The first repeat expansion comprised of polyglutamine CAG repeats was discovered in Fragile X patients in 1991.[416] More than 50 repeat expansion alleles have been found since then, and more than thirty of those were linked to neurodegenerative disorders.[417]

Repeat expansions contradict Mendelian inheritance rules, which for many years were the backbone of molecular research. Mendelian inheritance implies that pathogenic mutations are static and inherited in autosomal dominant, autosomal recessive, or X-linked ways and that mutations result in the same phenotype across generations. In contrast, repeat expansions are highly unstable, often changing in size when transmitted to the subsequent generations, resulting in different phenotypes.[418]

Improvements in sequencing technologies and the development of new algorithms for repeat expansion analysis brought significant insight into the nature of the large repeat expansions. However, intermediate-sized repeat expansions that fall in the range between the reference number of repeats and just below the pathogenic threshold

remain poorly understood. This type of expanded alleles is also known as the "gray zone" because they are larger than the normal allele but not large enough to confer the same risks as the premutation allele. Intermediate alleles demonstrate a significant degree of instability and a tendency for expansion in the consequent generations.[419] Furthermore, intermediate alleles have been reported to influence the clinical phenotype and possibly modify disease onset age.[420,421]

Based on the knowledge that ALS and FTD are repeat expansion disorders and that intermediate size repeat expansions may be associated with more than one neurological phenotype, I decided to examine the frequency of the intermediate size alleles in the twenty-one genes that were previously linked to neurodegeneration in our cohorts of FTD/ALS and LBD patients.

7.6 METHODS

I performed repeat sizing of twenty-one pathogenic genetic loci previously implicated in neurodegenerative diseases (AFF2, AR, ATN1, ATXN1, ATXN2, ATXN3, ATXN7, ATXN8OS, ATXN10, C9ORF72, CACNA1A, CNBP, CSTB, DMPK, FMR1, FXN, HTT, JPH3, NOP56, PPP2R2B, TBP).[422] The genes, the associated neurological disorders, and intermediate alleles' sizes are listed in **Table 9**.

Table 9. Pathogenic genetic loci examined in this study.

Gene	Location	Repeat	Normal allele	Intermed allele	Pathogenic allele	Expansion disorder
AFF2	Xq28	CCG	5-44	45-54	>200	FRAXE MR
AR	Xq12	CAG	<34	35	>38	SBMA
ATN1	12p13.31	CAG	6-35	36-47	>48	DRPLA
ATXN1	6p22.3	CAG	6-39	36-38	39-83	SCA1
ATXN2	12q24.12	CAG	<31	32	37->200	SCA2
ATXN3	14q32.12	CAG	12-44	45-50	~60 to 87	SCA3
ATXN7	3p14.1	CAG	7-27	34-36	37-460	SCA7
ATXN10	22q13.31	ATTCT	10-32	33-280	800-4,500	SCA10
ATXN8OS	13q21.33	CTG	15-50	50-70	80-250	SCA8
DMPK	19q13.32	CTG	5-34	35–49	>50	DM1
FMR1	Xq27.3	CGG	5-44	45–54	>200	FXS
C9ORF72	9p21.2	G4C2	2-24	25-60	60->2000	ALS, FTD
CACNA1A	19p13.13	CAG	≤18	19	20-33	SCA6
CNBP	3q21.3	CCTG	<30	27-29	~75-11,000	DM2
CSTB	21q22.3	C3GC4GCG	2-3	12-17	>30	EPM1
FXN	9q21.11	GAA	5-33	34-65	66-1300	FRDA
HTT	4p16.3	CAG	<26	27-35	>36	HD
JPH3	16q24.2	CAG/CTG	6-28	29-39	>40	HDL-2
NOP56	20p13	GGCCTG	3-14	15-650	650-2500	SCA36
PPP2R2B	5q32	CAG	7-32	43–49	51–78	SCA12
TBP	6q27	CAG	25-40	41-48	49-66	SCA17

The listed intermediate allele sizes were based on the previously published literature: AFF2 [423], AR [424] , ATN1 [425], ATXN1 [426], ATXN2 [427], ATXN3 [428], ATXN7 [429], ATXN8OS [430], ATXN10 [431], C9ORF72 [432], CACNA1A [433], CNBP [434], CSTB [435], DMPK [436], FMR1 [437], FXN [438], HTT [439], JPH3 [440], NOP56 [441], PPP2R2B [442], and TBP [443].

7.6.1 Cohorts

I used whole-genome sequence data acquired in the dementia whole-genome sequencing project. The discovery cohort consisted of a total of 8,199 samples that passed the QC filtering based on Genome Analysis Tool Kit (GATK) best practices.[444] Out of those, 2,442 patients were clinically and pathologically diagnosed with frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), 2,599 patients

diagnosed with Lewy body dementia (LBD), and 3,158 were neurologically healthy controls. All individuals examined in this study were of European ancestry.

The replication case cohort consisted of 2,648 samples from the University of Massachusetts analyzed by exome sequencing. The replication control cohort had 17,703 neurologically healthy individuals from the UK 100K Genomes project, 13,670 healthy individuals from Gardiner et al. (2019)[445], and 210 neurologically healthy controls screened in our lab.

7.6.2 Data analysis

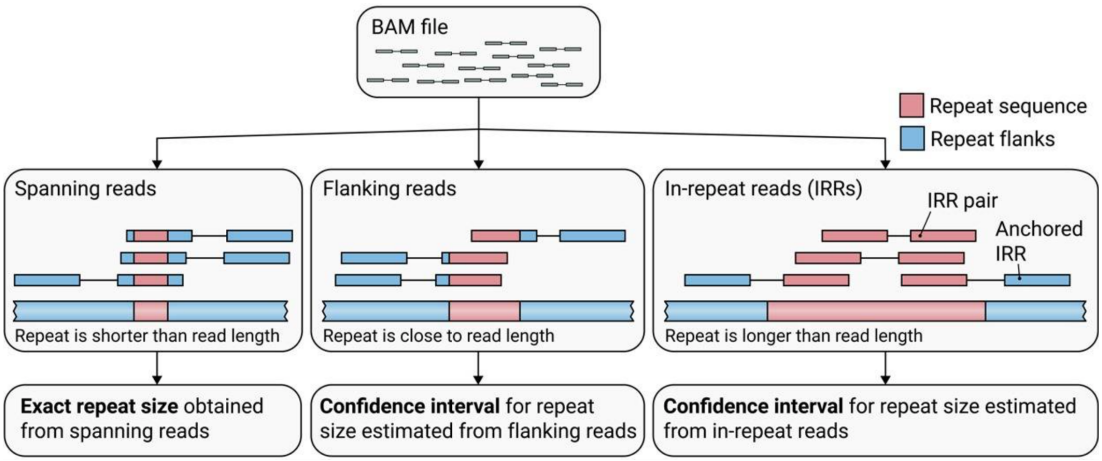
This data analysis was performed using version 3 of the ExpansionHunter targeted algorithm.[446] In the future, the data will be reanalyzed using version 4, an improved algorithm that is better at quantify loci such as PPP2R2B, and includes new loci like RFC1.

ExpansionHunter targeted version 3.0 was selected as an analytical tool for this project for several reasons. First, this software uses paired-end, PCR-free whole-genome sequencing data collected on Illumina platform to identify pathogenic repeat expansions. This is the exact method of data collection utilized in the DementiaSeq project. Second, 8 out of 21 genes selected for this project (AFF2, ATXN8OS, ATXN10, C9ORF72, CSTB, FMR1, FXN, NOP56), had the upper boundary for intermediate allele size larger than a sequencing read. The major limitation of the majority of short tandem repeats (STR) analysis software is inability to call the repeat expansions larger than the sequencing read, which is 150bp in length. In contrast,

ExpansionHunter can determine the size of repeats from just a few units in length up to large, pathogenic expansions that may be significantly longer than the read length. Third, compared to the other STR software GangSTR and STRetch, ExpansionHunter demonstrated the highest sensitivity for repeat expansion detection.[447] Finally, our group previously used the ExpansionHunter targeted in the Huntington's project [403], and it showed good accuracy in alle sizing, as confirmed by repeated primer PCR.

ExansionHunter targeted uses paired-end, PCR-free whole genome sequencing data for repeat length analysis. In the first step of the process, the algorithm separates the reads from a BAM file into three categories: 1) spanning reads - meaning the repeat spanned a total length of the read; 2) flanking reads – the reads that included the repeat and the flanking sequence on one side of the repeat; 3) in-repeat reads - meaning that repeat was entirely contained within a read. In the second step, the algorithm performed the categorized reads' analysis by assembling them back to the reference. The typical whole genome sequencing read is 150 base pairs in length. The repeat expansions contained within the read were the easiest to analyze as they did not require extra analysis. Repeats shorter than 150 bp were calculated using spanning and flanking reads. Finally, the lengths of repeats that were longer than read length were determined by counting the number of reads with "in read repeats".[446] The analysis process is summarized in **Figure 41**.

Figure 41. Analysis of repeat length by ExpansionHunter algorithm.



(Adapted from Dolzhenko et al., 2017)

7.7 RESULTS

Intermediate-size ATXN2 alleles, defined as 32 CAG repeats, were present in fifteen out of 2,442 ALS/FTD patients (0.61%) and two neurologically healthy controls (0.06%) in our discovery cohorts. Independent replication analysis on 2,438 ALS/FTD patients and discovered 12 carriers of the intermediate ATXN2 alleles (0.49%,). The ATXN2 intermediate allele results are summarized in **Table 10**. None of the other twenty genetic loci examined in this study appeared to increase the risk for developing ALS.

Table 10. Study results. ATXN2 intermediate repeat expansions in discovery and replication cohorts.

	Discovery cohort		Replication cohort	
	Expansion Carriers/ Total Screened	(%)	Expansion Carriers/ Total Screened	(%)
ALS/FTD	15/2,442	0.61	12/2,438	0.49
LBD	0/2,599	0	-	-
Controls	2/3,158	0.06	2/31,583	0.01

7.8 DISCUSSION

I examined the whole-genome sequence data collected in dementias sequencing projects that included patients diagnosed with FTD/ALS and LBD for the presence of the intermediate size repeat expansions in the twenty-one repeat expansion loci previously linked to neurodegeneration. The results of this study confirmed the previous association of intermediate size ATXN2 alleles with increased risk of developing ALS.[448,449]

Spinal cerebellar ataxia type 2 (SCA2) is a trinucleotide repeat disorder in which neurodegeneration occurs due to an expansion in the polyglutamine CAG repeat tract in the ATXN2 gene. Neurologically normal individuals typically have 22 to 23 CAG repeats, but up to 31 repeats are considered normal. Individuals with more than 34 CAG repeats develop progressive cerebellar ataxia, slow saccadic eye movement, and in some cases, parkinsonism, the clinical characteristics associated with SCA2.[450] The exact mechanism of ATXN2 pathogenicity in SCA2 remains poorly understood.

However, this gene likely causes the disease through the toxic gain-of-function mechanism similar to the other polyglutamine disorders.

In the past, SCA2 and ALS were viewed as two completely distinct neurological disorders due to the lack of overlapping symptoms. However, over the past decade, it became clear that these two neurological disorders are linked at the molecular level. ATXN2 is a cytoplasmic RNA binding protein involved in controlling RNA processing, degradation, and translation.[451] It interacts with key players involved in the pathogenesis of ALS, including TDP-43, FUS, and C9orf72.[452]

ATXN2 functions in a common pathway with TDP-43 and FUS, interacting in an RNA-dependent manner.[448] Intermediate-size expansion of ATXN2 polyglutamine tract enhances caspase-3 activated pathway, leading to TDP-43 sequestration from the nucleus to the cytoplasm where it aggregates into ATXN2-positive foci.[448] Similarly, intermediate ATXN2 expansions strengthen the interaction between ATXN2 and mutated FUS resulting in the formation of cytoplasmic FUS granules that induce endoplasmic reticulum stress, Golgi fragmentation, and apoptosis.[453] C9orf72 repeat expansions also coincide with intermediate size ATXN2 alleles in ALS-FTD patients. Co-expression of intermediate polyglutamine repeats of ATXN2 combined with C9orf72 depletion increases the aggregation of ATXN2, inducing motor neuron dysfunction and cell death.[454]

Reduction of ATXN2 levels has been previously suggested as a possible therapy for ALS since the decrease in ATXN2 has been demonstrated to significantly reduce TDP-43 aggregation and prolong survival in rodent models.[455] This therapy may be especially effective if the ATXN2 repeat expansions are discovered early in life. This

way, it may prevent cytoplasmic aggerates' formation and, consequently, significantly slow down or even stop the development of ALS and FTD. Indeed, Biogen recently developed a new antisense nucleotide drug ION541, that targets ATXN2 RNA. The drug is currently undergoing a Phase 1/2 clinical trial (ClinicalTrials.gov Identifier: NCT04494256). to assess the safety, tolerability, and pharmacokinetics in ALS patients with or without ATXN2 repeat expansions.

7.9 Uncertainty in determining the intermediate allele sizes and how they could be addressed

Intermediate alleles are defined in the literature as being larger than the typical range of CAG repeats found in the general population but smaller than premutation. However, in many cases the intermediate allele size is ambiguous. Using ATXN2 as an example, I would like to describe how one could design an experiment to determine a more precise size of intermediate alleles.

In the intermediate alleles study described in the sections 7.5-7.7, I outlined how ExpansionHunter targeted was used to identify the carriers of the intermediate ATXN2 alleles in ALS/FTD cohort. In those sections, I used 32 CAG repeats as an intermediate ATXN2 allele size. However, the size of intermediate ATXN2 alleles ranges between 24 and 34 in different publications.[448,456-458]

The possible steps of the study that aim to narrow down the intermediate ATXN2 allele size of are as follows. In the first step of the study design, the Author will select case and control cohorts. The cases should be representative of the disease, in this case ALS and the case number should have a statistical power to discover the intermediate allele

carriers. In the second step whole-genome data from the cases and controls is analyzed using ExpansionHunter targeted. Starting from the smallest allele size after the reference number of repeats, a separate repeat analysis must be performed each time increasing the allele size one repeat at a time until reaching the premutation/mutation allele size. For example, if the reference allele size for ATXN2 gene is 23 CAG the data should be analyzed for the frequency of 23+1 CAG repeats and then 24+1, 25+1 … 33 +1 until the premutation allele size is reached. The number of the ATXN2 repeat expansion carriers must be statistically evaluated in each step of the analysis to determine the biggest difference in the number of case and control carriers of this allele size. Finally, replication analysis must be performed on an independent case/control cohorts to confirm the discoveries.

8 AN OVERVIEW OF THE PROJECTS AND FUTURE DIRECTIONS

8.1 AN OVERVIEW OF THE PROJECTS

During my Ph.D., I utilized state-of-the-art exome and whole-genome sequencing technologies to discover the new loci within ALS and FTD's genetic spectrum that could potentially help unravel the complex problem of the etiology of these neurodegenerative disorders and pave the path for the development of the novel targeted therapies. Below is a brief overview of the projects presented in this thesis.

8.1.1 SPTLC1

During my graduate program, I focused on collecting and analyzing exome data. This effort led to the discovery of a new gene that is now linked to ALS: Serine palmitoyltransferase long chain subunit 1 (SPTLC1). SPTLC1 encodes a critical subunit of serine palmitoyltransferase, catalyzing the first and rate-limiting step of de novo sphingolipid synthesis. Exome sequencing revealed a novel mutation in the coding region of SPTLC1 in a juvenile-onset ALS case, who also displayed elevated levels of sphingolipid intermediates in the patient's plasma, strongly suggesting SPTLC1 has a pathological role in this disease. Furthermore, additional SPTLC1 mutations were discovered in adult ALS cases. These observations provided further support for a link between SPTLC1, ALS, and neurological pathogenesis. Data

collected in this study broadened the SPTLC1 phenotype and revealed de novo mutations behind at least a portion of sporadic instances of ALS.

8.1.2 Gelsolin

Gelsolin is a ubiquitously present protein with amyloid properties.[16] Biochemically, it can interact and bind to actin filaments and impact their capping and uncapping reactions, affecting cellular structures by influencing actin dynamics within the cell. Using exome sequencing, we also showed that the p.Asp187Tyr amino acid change is linked to an earlier onset of the disease, with the potential of a more severe progression than the p.Asp187Asn type. These discoveries suggest that exome sequencing is a powerful technique to analyze hereditary conditions that are rare and ambiguous in their phenotypic manifestations. This study shows the power of genomics to help physicians and neurologists quickly and accurately establish the correct diagnosis in patients.

8.1.3 RFC1

A homozygous pentanucleotide repeat expansion AAGGG within the Replication factor C subunit 1 (RFC1) gene was described as a common cause of cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS) syndrome. Because CANVAS patients have a variable presentation, we decided to examine a cohort of sporadic ALS patients for the presence of this repeat expansion. We did not discover any carriers of the homozygous [AAGGG] expansion in our ALS cohort. While this study results showed no substantial evidence to support a role for the RFC1 repeat

expansion in ALS, it further expanded on the highly polymorphic nature of RFC1 nature by describing a novel repeat conformation.

8.1.4 Intermediate alleles as part of dementias whole-genome sequencing project

In the past two years, I focused on learning whole-genome sequencing techniques and whole-genome sequencing analysis. I am planning to apply these techniques to the FTD whole genome sequencing project described below. The objective of the frontotemporal dementia whole-genome sequencing project was to perform genome sequencing in a large cohort of patients diagnosed with FTD and in a cohort of neurologically normal control subjects, and to make the raw data publicly available on the dbGaP (database of Genotypes and Phenotypes, accession number phs001963) web repository. This publicly available genome sequence data will act as a resource for other researchers worldwide working on FTD. The genome data generated for neurologically normal control subjects can also be used in an even more comprehensive range of neurological diseases.

The intermediate alleles project is a part of the dementia whole-genome sequencing project. This project's primary objective was to analyze the whole genome sequencing data acquired in dementia whole-genome sequencing project to determine the frequency of the intermediate size alleles in the genes that have been previously reported to cause neurodegeneration. The results of this project confirmed the previous association of Ataxin-2 intermediate alleles with an increased risk for developing ALS. Additional work will need to be done in this area as newer algorithms that enhance the

ability to call the repeat expansions become available. Access to more extensive population databases with this information will also be the key.

8.2 FUTURE DIRECTIONS

The evolution of sequencing technologies in the past two decades has revolutionized genetic research by allowing comprehensive studies of genetic variation across the entire human genome. Since the development of whole-genome sequencing, hundreds of thousands of genomes have been sequenced to build a comprehensive human genome database. The reduction in sequencing costs, data storage and processing costs have made whole genome sequencing the preferred technology for genetic research. However, research and medical limitations prevent whole-genome sequence data from reaching its full potential to bring clarity to medical diagnostics and ultimately offer personalized medicine.

A significant research-related limitation is that many whole-genome projects do not have enough statistical power. Furthermore, the current population databases are not representative of the population's racial and ethnic diversity. Genomes of people with European ancestry significantly overrepresent the majority of the whole-genome sequence databases. This strong European bias prevents the use of these databases for disease risk assessment in non-Europeans. In the future, research companies and hospitals will likely add more representatives of other populations, including rare and indigenous populations, to the currently existing and new genome databases.

The medical-related limitation of whole-genome sequence data is that although the cost of whole-genome sequencing significantly decreased over the past decade, it is still an

expensive diagnostic test typically not covered by medical insurance. Further innovations in sequencing technologies are needed to reduce the price per genome.

Another limitation of genomic research in the past two decades is the incompleteness of the human genome. The Human Genome project data released in 2003 covered 99% of the euchromatic regions of the genome; the highly repetitive heterochromatic regions were intentionally left off the genome map. These regions are mainly located in centromeres and telomeres, the parts of the chromosomes that consist of tightly packed chromatin that were previously impossible to sequence through. Additionally, these regions were omitted in the genome assembly because sequencing algorithms could not assemble the very long repeat sequences from short sequencing reads. With the invention of Oxford Nanopore technology, this task became possible. Nanopore sequencing produces single-molecule reads greater than one megabase in length. These reads completely span the repetitive regions, eliminating the need for read assembly and providing maximal precision for the large tandem repeats.

The identification of the new pathogenic variants and advances in our understanding of their pathological role in the development of neurological disorders led to the development of more effective therapies. In the past, therapies for neurological disorders were based on the patient's phenotype. Those therapies were only marginally effective because they aimed to alleviate the symptoms rather than treat the underlying pathology. Therefore, their effectiveness was decreasing with the disease progression. In contrast, gene therapies started emerging in the past decade and represent the first steps to the personalized treatments of neurological disorder patients. Gene therapy targets the underlying cause of the disease, such as mutation or perturbed mechanism. Gene therapy is based on the delivery of genetic material to cells in order to introduce

functional copies of dysfunctional genes, trophic factors and other disease-modifying genes, or silence harmful gene expression using antisense oligonucleotides (ASOs), RNA interference (RNAi), or gene-editing technology such as CRISPR.[459]

In addition to the personalized disease approach, another main advantage of gene therapy is that in the new era of genetic testing, the pathogenic mutations can be detected early in life, a long time before the first symptoms appear, therefore the use of gene therapy may prevent the cascade of the molecular events that lead to neurological damage or slow the neurodegenerative process down in individuals who already developed the disease.

Despite several decades of effort, several significant limitations prevent gene therapy from becoming a reliable form of treatment. These include effective delivery of the vector to the cell, lack of persistent gene expression in targeted cells, and immune responses to viral gene products, transgenes, or cells targeted by the vectors. Future studies should focus on these limitations to provide this life-changing treatment to patients suffering from neurological disorders.

REFERENCES

1. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. Feb 15 2001;409(6822):860-921. doi:10.1038/35057062

2. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*. Feb 15 2001;409(6822):928-33. doi:10.1038/35057149

3. Ozaki K, Ohnishi Y, Iida A, et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet*. Dec 2002;32(4):650-4. doi:10.1038/ng1047

4. EMBL-EBI. GWAS Catalog. Accessed 12.09.2020. https://www.ebi.ac.uk/gwas/downloads/summary-statistics

5. Consortium EP. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*. Oct 22 2004;306(5696):636-40. doi:10.1126/science.1105136

6. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. Sep 6 2012;489(7414):57-74. doi:10.1038/nature11247

7. Genomes Project C, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. Nov 1 2012;491(7422):56-65. doi:10.1038/nature11632

8. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. Oct 1 2015;526(7571):68-74. doi:10.1038/nature15393

9. Barros SP, Offenbacher S. Epigenetics: connecting environment and genotype to phenotype and disease. *J Dent Res*. May 2009;88(5):400-8. doi:10.1177/0022034509335868

10. Consortium UK, Walter K, Min JL, et al. The UK10K project identifies rare variants in health and disease. *Nature*. Oct 1 2015;526(7571):82-90. doi:10.1038/nature14962

11. The 100,000 Genomes Project. Accessed 12.15, 2020. https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/

12. J. Vance MT. Exome sequencing for neurological disorders. *Journal of the Neurological Sciences*. 2017;381(21):5.

13. Park JY, Clark P, Londin E, Sponziello M, Kricka LJ, Fortina P. Clinical exome performance for reporting secondary genetic findings. *Clin Chem*. Jan 2015;61(1):213-20. doi:10.1373/clinchem.2014.231456

14. Collins FS, Varmus H. A new initiative on precision medicine. *The New England journal of medicine*. Feb 26 2015;372(9):793-5. doi:10.1056/NEJMp1500523

15. Siva N. UK gears up to decode 100,000 genomes from NHS patients. *Lancet*. Jan 10 2015;385(9963):103-4. doi:10.1016/S0140-6736(14)62453-3

16. Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A*. Apr 28 2015;112(17):5473-8. doi:10.1073/pnas.1418631112

17. Meienberg J, Zerjavic K, Keller I, et al. New insights into the performance of human whole-exome capture platforms. *Nucleic Acids Res*. Jun 23 2015;43(11):e76. doi:10.1093/nar/gkv216

18. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. Jul 17 2014;511(7509):344-7. doi:10.1038/nature13394

19. Lattante S, Ciura S, Rouleau GA, Kabashi E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet*. May 2015;31(5):263-73. doi:10.1016/j.tig.2015.03.005

20. Kousi M, Katsanis N. Genetic modifiers and oligogenic inheritance. *Cold Spring Harb Perspect Med*. Jun 01 2015;5(6)doi:10.1101/cshperspect.a017145

21. Taskesen E, Mishra A, van der Sluis S, et al. Susceptible genes and disease mechanisms identified in frontotemporal dementia and frontotemporal dementia with Amyotrophic Lateral Sclerosis by DNA-methylation and GWAS. *Sci Rep*. Aug 21 2017;7(1):8899. doi:10.1038/s41598-017-09320-z

22. van Rheenen W, Shatunov A, Dekker AM, et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat Genet*. Sep 2016;48(9):1043-8. doi:10.1038/ng.3622

23. Woolley SC, York MK, Moore DH, et al. Detecting frontotemporal dysfunction in ALS: utility of the ALS Cognitive Behavioral Screen (ALS-CBS). *Amyotroph Lateral Scler*. May 3 2010;11(3):303-11. doi:10.3109/17482961003727954

24. Rowland LP, Shneider NA. Amyotrophic lateral sclerosis. Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, P.H.S.

Review. *The New England journal of medicine*. May 31 2001;344(22):1688-700. doi:10.1056/NEJM200105313442207

25. Arthur KC, Calvo A, Price TR, Geiger JT, Chio A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat Commun*. Aug 11 2016;7:12408. doi:10.1038/ncomms12408

26. Van Langenhove T, van der Zee J, Sleegers K, et al. Genetic contribution of FUS to frontotemporal lobar degeneration. *Neurology*. Feb 02 2010;74(5):366-71. doi:10.1212/WNL.0b013e3181ccc732

27. Dion PA, Daoud H, Rouleau GA. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. Research Support, Non-U.S. Gov't

Review. *Nature reviews Genetics*. Nov 2009;10(11):769-82. doi:10.1038/nrg2680

28. Renton AE, Chio A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci*. Jan 2014;17(1):17-23. doi:10.1038/nn.3584

29. Oskarsson B, Horton DK, Mitsumoto H. Potential Environmental Factors in Amyotrophic Lateral Sclerosis. *Neurol Clin*. Nov 2015;33(4):877-88. doi:10.1016/j.ncl.2015.07.009

30. Schymick JC, Talbot K, Traynor BJ. Genetics of sporadic amyotrophic lateral sclerosis. *Hum Mol Genet*. Oct 15 2007;16 Spec No. 2:R233-42. doi:10.1093/hmg/ddm215

31. Graham AJ, Macdonald AM, Hawkes CH. British motor neuron disease twin study. *J Neurol Neurosurg Psychiatry*. Jun 1997;62(6):562-9. doi:10.1136/jnnp.62.6.562

32. Al-Chalabi A, Fang F, Hanby MF, et al. An estimate of amyotrophic lateral sclerosis heritability using twin data. *J Neurol Neurosurg Psychiatry*. Dec 2010;81(12):1324-6. doi:10.1136/jnnp.2010.207464

33. Wingo TS, Cutler DJ, Yarab N, Kelly CM, Glass JD. The heritability of amyotrophic lateral sclerosis in a clinically ascertained United States research registry. *PLoS One*. 2011;6(11):e27985. doi:10.1371/journal.pone.0027985

34. Keller MF, Ferrucci L, Singleton AB, et al. Genome-wide analysis of the heritability of amyotrophic lateral sclerosis. *JAMA Neurol*. Sep 2014;71(9):1123-34. doi:10.1001/jamaneurol.2014.1184

35. Morgan S, Shatunov A, Sproviero W, et al. A comprehensive analysis of rare genetic variation in amyotrophic lateral sclerosis in the UK. *Brain*. Jun 01 2017;140(6):1611-1618. doi:10.1093/brain/awx082

36. Pang SY, Hsu JS, Teo KC, et al. Burden of rare variants in ALS genes influences survival in familial and sporadic ALS. *Neurobiol Aging*. 10 2017;58:238.e9-238.e15. doi:10.1016/j.neurobiolaging.2017.06.007

37. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. Oct 08 2009;461(7265):747-53. doi:10.1038/nature08494

38. Work SS, Colamonico JA, Bradley WG, Kaye RE. Pseudobulbar affect: an under-recognized and under-treated neurological disorder. *Adv Ther*. Jul 2011;28(7):586-601. doi:10.1007/s12325-011-0031-3

39. Singer MA, Kojan S, Barohn RJ, et al. Primary lateral sclerosis: clinical and laboratory features in 25 patients. *J Clin Neuromuscul Dis*. Sep 2005;7(1):1-9. doi:10.1097/01.cnd.0000176974.61136.45

40. Sotaniemi KA, Myllyla VV. Primary lateral sclerosis; a debated entity. *Acta Neurol Scand*. Apr 1985;71(4):334-6. doi:10.1111/j.1600-0404.1985.tb03210.x

41. Younger DS, Chou S, Hays AP, et al. Primary lateral sclerosis. A clinical diagnosis reemerges. *Arch Neurol*. Dec 1988;45(12):1304-7. doi:10.1001/archneur.1988.00520360022005

42. Singer MA, Statland JM, Wolfe GI, Barohn RJ. Primary lateral sclerosis. *Muscle Nerve*. Mar 2007;35(3):291-302. doi:10.1002/mus.20728

43. Statland JM, Barohn RJ, Dimachkie MM, Floeter MK, Mitsumoto H. Primary Lateral Sclerosis. *Neurol Clin*. Nov 2015;33(4):749-60. doi:10.1016/j.ncl.2015.07.007

44. Visser J, de Jong JM, de Visser M. The history of progressive muscular atrophy: syndrome or disease? *Neurology*. Feb 26 2008;70(9):723-7. doi:10.1212/01.wnl.0000302187.20239.93

45. Wijesekera LC, Mathers S, Talman P, et al. Natural history and clinical features of the flail arm and flail leg ALS variants. *Neurology*. Mar 24 2009;72(12):1087-94. doi:10.1212/01.wnl.0000345041.83406.a2

46. Kim WK, Liu X, Sandner J, et al. Study of 962 patients indicates progressive muscular atrophy is a form of ALS. *Neurology*. Nov 17 2009;73(20):1686-92. doi:10.1212/WNL.0b013e3181c1dea3

47. Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev*. Mar 14 2012;(3):CD001447. doi:10.1002/14651858.CD001447.pub3

48. Bensimon G, Lacomblez L, Meininger V. A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. *N Engl J Med*. Mar 3 1994;330(9):585-91. doi:10.1056/NEJM199403033300901

49. Jaiswal MK. Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. *Med Res Rev*. Mar 2019;39(2):733-748. doi:10.1002/med.21528

50. Lacomblez L, Bensimon G, Leigh PN, Guillet P, Meininger V. Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II. *Lancet*. May 25 1996;347(9013):1425-31. doi:10.1016/s0140-6736(96)91680-3

51. Hinchcliffe M, Smith A. Riluzole: real-world evidence supports significant extension of median survival times in patients with amyotrophic lateral sclerosis. *Degener Neurol Neuromuscul Dis*. 2017;7:61-70. doi:10.2147/DNND.S135748

52. Abe K, Itoyama Y, Sobue G, et al. Confirmatory double-blind, parallel-group, placebo-controlled study of efficacy and safety of edaravone (MCI-186) in amyotrophic lateral sclerosis patients. *Amyotroph Lateral Scler Frontotemporal Degener*. Dec 2014;15(7-8):610-7. doi:10.3109/21678421.2014.959024

53. Yoshino H, Kimura A. Investigation of the therapeutic effects of edaravone, a free radical scavenger, on amyotrophic lateral sclerosis (Phase II study). *Amyotroph Lateral Scler*. Dec 2006;7(4):241-5. doi:10.1080/17482960600881870

54. Shefner J, Heiman-Patterson T, Pioro EP, et al. Long-term edaravone efficacy in amyotrophic lateral sclerosis: Post-hoc analyses of Study 19 (MCI186-19). *Muscle Nerve*. Feb 2020;61(2):218-221. doi:10.1002/mus.26740

55. Okada M, Yamashita S, Ueyama H, Ishizaki M, Maeda Y, Ando Y. Long-term effects of edaravone on survival of patients with amyotrophic lateral sclerosis. *eNeurologicalSci*. Jun 2018;11:11-14. doi:10.1016/j.ensci.2018.05.001

56. Rong WT, Lu YP, Tao Q, et al. Hydroxypropyl-sulfobutyl-beta-cyclodextrin improves the oral bioavailability of edaravone by modulating drug efflux pump of enterocytes. *J Pharm Sci*. Feb 2014;103(2):730-42. doi:10.1002/jps.23807

57. Wang J, Chen X, Yuan B, et al. Bioavailability of Edaravone Sublingual Tablet Versus Intravenous Infusion in Healthy Male Volunteers. *Clin Ther*. Oct 2018;40(10):1683-1691. doi:10.1016/j.clinthera.2018.08.009

58. Mora JS, Genge A, Chio A, et al. Masitinib as an add-on therapy to riluzole in patients with amyotrophic lateral sclerosis: a randomized clinical trial. *Amyotroph*

Lateral Scler Frontotemporal Degener. Feb 2020;21(1-2):5-14. doi:10.1080/21678421.2019.1632346

59. Trias E, Ibarburu S, Barreto-Nunez R, et al. Post-paralysis tyrosine kinase inhibition with masitinib abrogates neuroinflammation and slows disease progression in inherited amyotrophic lateral sclerosis. *J Neuroinflammation*. Jul 11 2016;13(1):177. doi:10.1186/s12974-016-0620-9

60. Trias E, King PH, Si Y, et al. Mast cells and neutrophils mediate peripheral motor pathway degeneration in ALS. *JCI Insight*. Oct 4 2018;3(19)doi:10.1172/jci.insight.123249

61. Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron D. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord*. Dec 2000;1(5):293-9. doi:10.1080/146608200300079536

62. Bunton-Stasyshyn RK, Saccon RA, Fratta P, Fisher EM. SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology: New and Renascent Themes. *Neuroscientist*. Oct 2015;21(5):519-29. doi:10.1177/1073858414561795

63. Miller T, Cudkowicz M, Shaw PJ, et al. Phase 1-2 Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *N Engl J Med*. Jul 9 2020;383(2):109-119. doi:10.1056/NEJMoa2003715

64. Biogen. Update on Tofersen Clinical Program. Accessed June 23, 2021. https://www.biogen.com/en_us/als-community-update.html

65. Russell DW, Setchell KD. Bile acid biosynthesis. *Biochemistry*. May 26 1992;31(20):4737-49. doi:10.1021/bi00135a001

66. Rodrigues CM, Spellman SR, Sola S, et al. Neuroprotection by a bile acid in an acute stroke model in the rat. *J Cereb Blood Flow Metab*. Apr 2002;22(4):463-71. doi:10.1097/00004647-200204000-00010

67. Keene CD, Rodrigues CM, Eich T, Chhabra MS, Steer CJ, Low WC. Tauroursodeoxycholic acid, a bile acid, is neuroprotective in a transgenic animal model of Huntington's disease. *Proc Natl Acad Sci U S A*. Aug 6 2002;99(16):10671-6. doi:10.1073/pnas.162362299

68. Castro-Caldas M, Carvalho AN, Rodrigues E, et al. Tauroursodeoxycholic acid prevents MPTP-induced dopaminergic cell death in a mouse model of Parkinson's disease. *Mol Neurobiol*. Oct 2012;46(2):475-86. doi:10.1007/s12035-012-8295-4

69. Nunes AF, Amaral JD, Lo AC, et al. TUDCA, a bile acid, attenuates amyloid precursor protein processing and amyloid-beta deposition in APP/PS1 mice. *Mol Neurobiol*. Jun 2012;45(3):440-54. doi:10.1007/s12035-012-8256-y

70. Elia AE, Lalli S, Monsurro MR, et al. Tauroursodeoxycholic acid in the treatment of patients with amyotrophic lateral sclerosis. *Eur J Neurol*. Jan 2016;23(1):45-52. doi:10.1111/ene.12664

71. Paganoni S, Macklin EA, Hendrix S, et al. Trial of Sodium Phenylbutyrate-Taurursodiol for Amyotrophic Lateral Sclerosis. *N Engl J Med*. Sep 3 2020;383(10):919-930. doi:10.1056/NEJMoa1916945

72. Chio A, Mazzini L, Mora G. Disease-modifying therapies in amyotrophic lateral sclerosis. *Neuropharmacology*. May 1 2020;167:107986. doi:10.1016/j.neuropharm.2020.107986

73. Alexion announces planned initiation of pivotal phase 3 study of ULTOMIRIS (ravulizumab) in ALS. January 14, 2020, Accessed 11/04/2020. https://clinicaltrials.gov/ct2/show/NCT04248465

74. Uccelli A, Moretta L, Pistoia V. Mesenchymal stem cells in health and disease. *Nat Rev Immunol*. Sep 2008;8(9):726-36. doi:10.1038/nri2395

75. Berry JD, Cudkowicz M.E., Windebank A,J., et al. NurOwn, phase 2, randomized, clinical trial in patients with ALS. Safety, clinical, and biomarker results. *Neurology*. 2019;93:2294-2305.

76. Patricia I. Phase 3 Trial of NurOwn Cell Therapy Fully Enrolls 200 Patients, BrainStorm Announces. Accessed Nov 3, 2020. https://alsnewstoday.com/news-posts/2019/10/14/phase-3-nurown-trial-fully-enrolls-200-als-patients-brainstorm-says/

77. Rosso SM, Donker Kaat L, Baks T, et al. Frontotemporal dementia in The Netherlands: patient characteristics and prevalence estimates from a population-based study. *Brain*. Sep 2003;126(Pt 9):2016-22. doi:10.1093/brain/awg204

78. Onyike CU, Diehl-Schmid J. The epidemiology of frontotemporal dementia. *Int Rev Psychiatry*. Apr 2013;25(2):130-7. doi:10.3109/09540261.2013.776523

79. Snowden JS, Neary D, Mann DM. Frontotemporal dementia. *Br J Psychiatry*. Feb 2002;180:140-3. doi:10.1192/bjp.180.2.140

80. Kansal K, Mareddy M, Sloane KL, et al. Survival in Frontotemporal Dementia Phenotypes: A Meta-Analysis. *Dement Geriatr Cogn Disord*. 2016;41(1-2):109-22. doi:10.1159/000443205

81. Agarwal S, Ahmed RM, D'Mello M, et al. Predictors of survival and progression in behavioural variant frontotemporal dementia. *Eur J Neurol*. May 2019;26(5):774-779. doi:10.1111/ene.13887

82. Blauwendraat C, Wilke C, Simon-Sanchez J, et al. The wide genetic landscape of clinical frontotemporal dementia: systematic combined sequencing of 121 consecutive subjects. *Genet Med*. Feb 2018;20(2):240-249. doi:10.1038/gim.2017.102

83. Blauwendraat C, Faghri F, Pihlstrom L, et al. NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases. *Neurobiol Aging*. Sep 2017;57:247 e9-247 e13. doi:10.1016/j.neurobiolaging.2017.05.009

84. Greaves CV, Rohrer JD. An update on genetic frontotemporal dementia. *J Neurol*. Aug 2019;266(8):2075-2086. doi:10.1007/s00415-019-09363-4

85. Nicholson AM, Rademakers R. What we know about TMEM106B in neurodegeneration. *Acta Neuropathol*. Nov 2016;132(5):639-651. doi:10.1007/s00401-016-1610-9

86. Harciarek M, Kertesz A. Primary progressive aphasias and their contribution to the contemporary knowledge about the brain-language relationship. *Neuropsychol Rev*. Sep 2011;21(3):271-87. doi:10.1007/s11065-011-9175-9

87. Kleckner IR, Zhang J, Touroutoglou A, et al. Evidence for a Large-Scale Brain System Supporting Allostasis and Interoception in Humans. *Nat Hum Behav*. 2017;1doi:10.1038/s41562-017-0069

88. Johnen A, Bertoux M. Psychological and Cognitive Markers of Behavioral Variant Frontotemporal Dementia-A Clinical Neuropsychologist's View on Diagnostic Criteria and Beyond. *Front Neurol*. 2019;10:594. doi:10.3389/fneur.2019.00594

89. Johnson JK, Diehl J, Mendez MF, et al. Frontotemporal lobar degeneration: demographic characteristics of 353 patients. *Arch Neurol*. Jun 2005;62(6):925-30. doi:10.1001/archneur.62.6.925

90. Rohrer JD, Warren JD, Modat M, et al. Patterns of cortical thinning in the language variants of frontotemporal lobar degeneration. *Neurology*. May 5 2009;72(18):1562-9. doi:10.1212/WNL.0b013e3181a4124e

91. Ash S, Moore P, Vesely L, et al. Non-Fluent Speech in Frontotemporal Lobar Degeneration. *J Neurolinguistics*. Jul 1 2009;22(4):370-383. doi:10.1016/j.jneuroling.2008.12.001

92. Mesulam MM. Primary progressive aphasia. *Ann Neurol*. Apr 2001;49(4):425-32.

93. Lamb R, Rohrer JD, Lees AJ, Morris HR. Progressive Supranuclear Palsy and Corticobasal Degeneration: Pathophysiology and Treatment Options. *Curr Treat Options Neurol*. Sep 2016;18(9):42. doi:10.1007/s11940-016-0422-5

94. Coyle-Gilchrist IT, Dick KM, Patterson K, et al. Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*. May 3 2016;86(18):1736-43. doi:10.1212/WNL.0000000000002638

95. Hodges JR, Patterson K. Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurol*. Nov 2007;6(11):1004-14. doi:10.1016/S1474-4422(07)70266-1

96. Kertesz A, Jesso S, Harciarek M, Blair M, McMonagle P. What is semantic dementia?: a cohort study of diagnostic features and clinical boundaries. *Arch Neurol*. Apr 2010;67(4):483-9. doi:10.1001/archneurol.2010.55

97. Spinelli EG, Mandelli ML, Miller ZA, et al. Typical and atypical pathology in primary progressive aphasia variants. *Ann Neurol*. Mar 2017;81(3):430-443. doi:10.1002/ana.24885

98. Gorno-Tempini ML, Brambati SM, Ginex V, et al. The logopenic/phonological variant of primary progressive aphasia. *Neurology*. Oct 14 2008;71(16):1227-34. doi:10.1212/01.wnl.0000320506.79811.da

99. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology*. Mar 15 2011;76(11):1006-14. doi:10.1212/WNL.0b013e31821103e6

100. Phillips JS, Da Re F, Dratch L, et al. Neocortical origin and progression of gray matter atrophy in nonamnestic Alzheimer's disease. *Neurobiol Aging*. Mar 2018;63:75-87. doi:10.1016/j.neurobiolaging.2017.11.008

101. Mesulam M, Wicklund A, Johnson N, et al. Alzheimer and frontotemporal pathology in subsets of primary progressive aphasia. *Ann Neurol*. Jun 2008;63(6):709-19. doi:10.1002/ana.21388

102. Snowden J, Neary D, Mann D. Frontotemporal lobar degeneration: clinical and pathological relationships. *Acta Neuropathol*. Jul 2007;114(1):31-8. doi:10.1007/s00401-007-0236-3

103. Teichmann M, Migliaccio R, Kas A, Dubois B. Logopenic progressive aphasia beyond Alzheimer's--an evolution towards dementia with Lewy bodies. *J Neurol Neurosurg Psychiatry*. Jan 2013;84(1):113-4. doi:10.1136/jnnp-2012-302638

104. Chare L, Hodges JR, Leyton CE, et al. New criteria for frontotemporal dementia syndromes: clinical and pathological diagnostic implications. *J Neurol Neurosurg Psychiatry*. Aug 2014;85(8):865-70. doi:10.1136/jnnp-2013-306948

105. Abramzon YA, Fratta P, Traynor BJ, Chia R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front Neurosci*. 2020;14:42. doi:10.3389/fnins.2020.00042

106. Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*. Aug 23 2005;65(4):586-90. doi:10.1212/01.wnl.0000172911.39167.b6

107. Burrell JR, Kiernan MC, Vucic S, Hodges JR. Motor neuron dysfunction in frontotemporal dementia. *Brain*. Sep 2011;134(Pt 9):2582-94. doi:10.1093/brain/awr195

108. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. Oct 20 2011;72(2):257-68. doi:10.1016/j.neuron.2011.09.010

109. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*. Oct 20 2011;72(2):245-56. doi:10.1016/j.neuron.2011.09.011

110. Devenney EM, Ahmed RM, Halliday G, Piguet O, Kiernan MC, Hodges JR. Psychiatric disorders in C9orf72 kindreds: Study of 1,414 family members. *Neurology*. Oct 16 2018;91(16):e1498-e1507. doi:10.1212/WNL.0000000000006344

111. Hensman Moss DJ, Poulter M, Beck J, et al. C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. *Neurology*. Jan 28 2014;82(4):292-9. doi:10.1212/WNL.0000000000000061

112. Beck J, Poulter M, Hensman D, et al. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am J Hum Genet*. Mar 7 2013;92(3):345-53. doi:10.1016/j.ajhg.2013.01.011

113. Shi Y, Lin S, Staats KA, et al. Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med*. Mar 2018;24(3):313-325. doi:10.1038/nm.4490

114. Shao Q, Liang C, Chang Q, Zhang W, Yang M, Chen JF. C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. *Acta Neuropathol Commun*. Mar 4 2019;7(1):32. doi:10.1186/s40478-019-0685-7

115. Arzberger T, Schludi MH, Lehmer C, Schmid B, Edbauer D. RNA versus protein toxicity in C9orf72 ALS/FTLD. *Acta Neuropathol*. Mar 2018;135(3):475-479. doi:10.1007/s00401-018-1823-1

116. May S, Hornburg D, Schludi MH, et al. C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol*. Oct 2014;128(4):485-503. doi:10.1007/s00401-014-1329-4

117. Freibaum BD, Taylor JP. The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Front Mol Neurosci*. 2017;10:35. doi:10.3389/fnmol.2017.00035

118. Zhang K, Donnelly CJ, Haeusler AR, et al. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature*. Sep 03 2015;525(7567):56-61. doi:10.1038/nature14973

119. Jovicic A, Mertens J, Boeynaems S, et al. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci*. Sep 2015;18(9):1226-9. doi:10.1038/nn.4085

120. Freibaum BD, Lu Y, Lopez-Gonzalez R, et al. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature*. Sep 03 2015;525(7567):129-33. doi:10.1038/nature14974

121. Koppers M, Blokhuis AM, Westeneng HJ, et al. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol*. Sep 2015;78(3):426-38. doi:10.1002/ana.24453

122. Peters OM, Cabrera GT, Tran H, et al. Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron*. Dec 02 2015;88(5):902-9. doi:10.1016/j.neuron.2015.11.018

123. O'Rourke JG, Bogdanik L, Muhammad AK, et al. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron*. Dec 02 2015;88(5):892-901. doi:10.1016/j.neuron.2015.10.027

124. Jiang J, Zhu Q, Gendron TF, et al. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron*. May 04 2016;90(3):535-50. doi:10.1016/j.neuron.2016.04.006

125. Hao Z, Liu L, Tao Z, et al. Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nat Commun*. Jul 2 2019;10(1):2906. doi:10.1038/s41467-019-10956-w

126. Zhang YJ, Gendron TF, Ebbert MTW, et al. Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nat Med*. Aug 2018;24(8):1136-1142. doi:10.1038/s41591-018-0071-1

127. Choi SY, Lopez-Gonzalez R, Krishnan G, et al. C9ORF72-ALS/FTD-associated poly(GR) binds Atp5a1 and compromises mitochondrial function in vivo. *Nat Neurosci*. Jun 2019;22(6):851-862. doi:10.1038/s41593-019-0397-0

128. Sreedharan J, Blair IP, Tripathi VB, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. Mar 21 2008;319(5870):1668-72. doi:10.1126/science.1154584

129. Mackenzie IR, Rademakers R, Neumann M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol*. Oct 2010;9(10):995-1007. doi:10.1016/S1474-4422(10)70195-2

130. Bodansky A, Kim JM, Tempest L, Velagapudi A, Libby R, Ravits J. TDP-43 and ubiquitinated cytoplasmic aggregates in sporadic ALS are low frequency and

widely distributed in the lower motor neuron columns independent of disease spread. *Amyotroph Lateral Scler*. May 03 2010;11(3):321-7. doi:10.3109/17482961003602363

131. Davidson Y, Kelley T, Mackenzie IR, et al. Ubiquitinated pathological lesions in frontotemporal lobar degeneration contain the TAR DNA-binding protein, TDP-43. *Acta Neuropathol*. May 2007;113(5):521-33. doi:10.1007/s00401-006-0189-y

132. Neumann M, Sampathu DM, Kwong LK, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. Oct 06 2006;314(5796):130-3. doi:10.1126/science.1134108

133. Tan RH, Yang Y, Kim WS, et al. Distinct TDP-43 inclusion morphologies in frontotemporal lobar degeneration with and without amyotrophic lateral sclerosis. *Acta Neuropathol Commun*. Oct 27 2017;5(1):76. doi:10.1186/s40478-017-0480-2

134. Buratti E, Baralle FE. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front Biosci*. Jan 01 2008;13:867-78.

135. Strong MJ, Volkening K, Hammond R, et al. TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. *Mol Cell Neurosci*. Jun 2007;35(2):320-7. doi:10.1016/j.mcn.2007.03.007

136. Lagier-Tourenne C, Polymenidou M, Cleveland DW. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet*. Apr 15 2010;19(R1):R46-64. doi:10.1093/hmg/ddq137

137. Fiesel FC, Voigt A, Weber SS, et al. Knockdown of transactive response DNA-binding protein (TDP-43) downregulates histone deacetylase 6. *EMBO J*. Jan 06 2010;29(1):209-21. doi:10.1038/emboj.2009.324

138. Buratti E, Dork T, Zuccato E, Pagani F, Romano M, Baralle FE. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J*. Apr 02 2001;20(7):1774-84. doi:10.1093/emboj/20.7.1774

139. Roczniak-Ferguson A, Ferguson SM. Pleiotropic requirements for human TDP-43 in the regulation of cell and organelle homeostasis. *Life Sci Alliance*. Oct 2019;2(5)doi:10.26508/lsa.201900358

140. Mitra J, Guerrero EN, Hegde PM, et al. Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proc Natl Acad Sci U S A*. Mar 5 2019;116(10):4696-4705. doi:10.1073/pnas.1818415116

141. Buratti E, Baralle FE. TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem Sci*. Jun 2012;37(6):237-47. doi:10.1016/j.tibs.2012.03.003

142. Hergesheimer RC, Chami AA, de Assis DR, et al. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain*. May 1 2019;142(5):1176-1194. doi:10.1093/brain/awz078

143. Conicella AE, Zerze GH, Mittal J, Fawzi NL. ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure*. Sep 06 2016;24(9):1537-49. doi:10.1016/j.str.2016.07.007

144. Ling JP, Pletnikova O, Troncoso JC, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. Aug 7 2015;349(6248):650-5. doi:10.1126/science.aab0983

145. Jeong YH, Ling JP, Lin SZ, et al. Tdp-43 cryptic exons are highly variable between cell types. *Mol Neurodegener*. Feb 2 2017;12(1):13. doi:10.1186/s13024-016-0144-x

146. Prudencio M, Humphrey J, Pickles S, et al. Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia. *J Clin Invest*. 11 02 2020;130(11):6080-6092. doi:10.1172/JCI139741

147. Šušnjar U, Brown AL, Abbassi Y et al. Cell environment shapes TDP-43 function: implications in neuronal and muscle disease. *bioRxiv*. 2021;2021.04.20.440589doi:doi: https://doi.org/10.1101/2021.04.20.440589

148. Ma XR MP, Koike Y, Vatsavayai SC et al. TDP-43 represses cryptic exon inclusion in FTD/ALS gene UNC13A. *bioRxiv*. 2021;04.02.438213doi:https://doi.org/10.1101/2021.04.02.438213

149. Silva PR, Nieva GV, Igaz LM. Suppression of Conditional TDP-43 Transgene Expression Differentially Affects Early Cognitive and Social Phenotypes in TDP-43 Mice. *Front Genet*. 2019;10:369. doi:10.3389/fgene.2019.00369

150. White MA, Kim E, Duffy A, et al. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat Neurosci*. Apr 2018;21(4):552-563. doi:10.1038/s41593-018-0113-5

151. Ebstein SY, Yagudayeva I, Shneider NA. Mutant TDP-43 Causes Early-Stage Dose-Dependent Motor Neuron Degeneration in a TARDBP Knockin Mouse Model of ALS. *Cell Rep*. Jan 8 2019;26(2):364-373 e4. doi:10.1016/j.celrep.2018.12.045

152. Fratta P, Sivakumar P, Humphrey J, et al. Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J*. Jun 1 2018;37(11)doi:10.15252/embj.201798684

153. White MA, Kim E, Duffy A, et al. Publisher Correction: TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat Neurosci*. Aug 2018;21(8):1138. doi:10.1038/s41593-018-0160-y

154. Laurin N, Brown JP, Morissette J, Raymond V. Recurrent mutation of the gene encoding sequestosome 1 (SQSTM1/p62) in Paget disease of bone. *Am J Hum Genet*. Jun 2002;70(6):1582-8. doi:10.1086/340731

155. Fecto F, Yan J, Vemula SP, et al. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol*. Nov 2011;68(11):1440-6. doi:10.1001/archneurol.2011.250

156. Park S, Zuber C, Roth J. Selective autophagy of cytosolic protein aggregates involves ribosome-free rough endoplasmic reticulum. *Histochem Cell Biol*. Nov 12 2019;doi:10.1007/s00418-019-01829-w

157. Gao J, Perera G, Bhadbhade M, Halliday GM, Dzamko N. Autophagy activation promotes clearance of alpha-synuclein inclusions in fibril-seeded human neural cells. *J Biol Chem*. Sep 27 2019;294(39):14241-14256. doi:10.1074/jbc.RA119.008733

158. Zaffagnini G, Savova A, Danieli A, et al. Phasing out the bad-How SQSTM1/p62 sequesters ubiquitinated proteins for degradation by autophagy. *Autophagy*. 2018;14(7):1280-1282. doi:10.1080/15548627.2018.1462079

159. Foster A, Scott D, Layfield R, Rea SL. An FTLD-associated SQSTM1 variant impacts Nrf2 and NF-kappaB signalling and is associated with reduced phosphorylation of p62. *Mol Cell Neurosci*. Jul 2019;98:32-45. doi:10.1016/j.mcn.2019.04.001

160. Rea SL, Walsh JP, Layfield R, Ratajczak T, Xu J. New insights into the role of sequestosome 1/p62 mutant proteins in the pathogenesis of Paget's disease of bone. *Endocr Rev*. Aug 2013;34(4):501-24. doi:10.1210/er.2012-1034

161. Jung KT, Oh SH. Polyubiquitination of p62/SQSTM1 is a prerequisite for Fas/CD95 aggregation to promote caspase-dependent apoptosis in cadmium-exposed mouse monocyte RAW264.7 cells. *Sci Rep*. Aug 22 2019;9(1):12240. doi:10.1038/s41598-019-48684-2

162. Teyssou E, Takeda T, Lebon V, et al. Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology. *Acta Neuropathol*. Apr 2013;125(4):511-22. doi:10.1007/s00401-013-1090-0

163. Arai T, Nonaka T, Hasegawa M, et al. Neuronal and glial inclusions in frontotemporal dementia with or without motor neuron disease are immunopositive for p62. *Neurosci Lett*. May 15 2003;342(1-2):41-4.

164. Deng Z, Lim J, Wang Q, et al. ALS-FTLD-linked mutations of SQSTM1/p62 disrupt selective autophagy and NFE2L2/NRF2 anti-oxidative stress pathway. *Autophagy*. Jul 30 2019:1-15. doi:10.1080/15548627.2019.1644076

165. Herhaus L, Dikic I. Ubiquitin-induced phase separation of p62/SQSTM1. *Cell Res*. Apr 2018;28(4):389-390. doi:10.1038/s41422-018-0030-x

166. Sun D, Wu R, Zheng J, Li P, Yu L. Polyubiquitin chain-induced p62 phase separation drives autophagic cargo segregation. *Cell Res*. Apr 2018;28(4):405-415. doi:10.1038/s41422-018-0017-7

167. Mitsui S, Otomo A, Nozaki M, et al. Systemic overexpression of SQSTM1/p62 accelerates disease onset in a SOD1(H46R)-expressing ALS mouse model. *Mol Brain*. May 29 2018;11(1):30. doi:10.1186/s13041-018-0373-8

168. Kwon J, Han E, Bui CB, et al. Assurance of mitochondrial integrity and mammalian longevity by the p62-Keap1-Nrf2-Nqo1 cascade. *EMBO Rep*. Feb 01 2012;13(2):150-6. doi:10.1038/embor.2011.246

169. Kwiatkowski TJ, Jr., Bosco DA, Leclerc AL, et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. Feb 27 2009;323(5918):1205-8. doi:10.1126/science.1166066

170. Vance C, Rogelj B, Hortobagyi T, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*. Feb 27 2009;323(5918):1208-11. doi:10.1126/science.1165942

171. Huey ED, Ferrari R, Moreno JH, et al. FUS and TDP43 genetic variability in FTD and CBS. *Neurobiol Aging*. May 2012;33(5):1016 e9-17. doi:10.1016/j.neurobiolaging.2011.08.004

172. Shang Y, Huang EJ. Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis. *Brain Res*. Sep 15 2016;1647:65-78. doi:10.1016/j.brainres.2016.03.036

173. R AA. Neuronal cytoplasmic inclusions in tau, TDP-43, and FUS molecular subtypes of frontotemporal lobar degeneration share similar spatial patterns. *Folia Neuropathol*. 2017;55(3):185-192. doi:10.5114/fn.2017.70482

174. Ishigaki S, Sobue G. Importance of Functional Loss of FUS in FTLD/ALS. *Front Mol Biosci*. 2018;5:44. doi:10.3389/fmolb.2018.00044

175. An H, Skelt L, Notaro A, et al. ALS-linked FUS mutations confer loss and gain of function in the nucleus by promoting excessive formation of dysfunctional paraspeckles. *Acta Neuropathol Commun*. Jan 14 2019;7(1):7. doi:10.1186/s40478-019-0658-x

176. Kang J, Lim L, Lu Y, Song J. A unified mechanism for LLPS of ALS/FTLD-causing FUS as well as its modulation by ATP and oligonucleic acids. *PLoS Biol*. Jun 2019;17(6):e3000327. doi:10.1371/journal.pbio.3000327

177. Niaki AG, Sarkar J, Cai X, et al. Loss of Dynamic RNA Interaction and Aberrant Phase Separation Induced by Two Distinct Types of ALS/FTD-Linked FUS Mutations. *Mol Cell*. Oct 17 2019;doi:10.1016/j.molcel.2019.09.022

178. Murthy AC, Dignon GL, Kan Y, et al. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat Struct Mol Biol*. Jul 2019;26(7):637-648. doi:10.1038/s41594-019-0250-x

179. Berry J, Brangwynne CP, Haataja M. Physical principles of intracellular organization via active and passive phase transitions. *Rep Prog Phys*. Apr 2018;81(4):046601. doi:10.1088/1361-6633/aaa61e

180. Kino Y, Washizu C, Kurosawa M, et al. FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis. *Acta Neuropathol Commun*. Apr 25 2015;3:24. doi:10.1186/s40478-015-0202-6

181. Scekic-Zahirovic J, Sendscheid O, El Oussini H, et al. Toxic gain of function from mutant FUS protein is crucial to trigger cell autonomous motor neuron loss. *EMBO J*. May 17 2016;35(10):1077-97. doi:10.15252/embj.201592559

182. Devoy A, Kalmar B, Stewart M, et al. Humanized mutant FUS drives progressive motor neuron degeneration without aggregation in 'FUSDelta14' knockin mice. *Brain*. Nov 01 2017;140(11):2797-2805. doi:10.1093/brain/awx248

183. Zhang X, Wang F, Hu Y, et al. In vivo stress granule misprocessing evidenced in a FUS knock-in ALS mouse model. *Brain*. May 1 2020;143(5):1350-1367. doi:10.1093/brain/awaa076

184. Milani P, Gagliardi S, Cova E, Cereda C. SOD1 Transcriptional and Posttranscriptional Regulation and Its Potential Implications in ALS. *Neurol Res Int*. 2011;2011:458427. doi:10.1155/2011/458427

185. Liu J, Lillo C, Jonsson PA, et al. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron*. Jul 08 2004;43(1):5-17. doi:10.1016/j.neuron.2004.06.016

186. Cozzolino M, Pesaresi MG, Amori I, et al. Oligomerization of mutant SOD1 in mitochondria of motoneuronal cells drives mitochondrial damage and cell toxicity. *Antioxid Redox Signal*. Jul 2009;11(7):1547-58. doi:10.1089/ars.2009.2545

187. Kaur SJ, McKeown SR, Rashid S. Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis. *Gene*. Feb 15 2016;577(2):109-18. doi:10.1016/j.gene.2015.11.049

188. Hayashi Y, Homma K, Ichijo H. SOD1 in neurotoxicity and its controversial roles in SOD1 mutation-negative ALS. *Adv Biol Regul*. Jan 2016;60:95-104. doi:10.1016/j.jbior.2015.10.006

189. Nishitoh H, Kadowaki H, Nagai A, et al. ALS-linked mutant SOD1 induces ER stress- and ASK1-dependent motor neuron death by targeting Derlin-1. *Genes Dev*. Jun 01 2008;22(11):1451-64. doi:10.1101/gad.1640108

190. Münch C, Bertolotti A. Exposure of hydrophobic surfaces initiates aggregation of diverse ALS-causing superoxide dismutase-1 mutants. *J Mol Biol*. Jun 11 2010;399(3):512-25. doi:10.1016/j.jmb.2010.04.019

191. Guareschi S, Cova E, Cereda C, et al. An over-oxidized form of superoxide dismutase found in sporadic amyotrophic lateral sclerosis with bulbar onset shares a toxic mechanism with mutant SOD1. *Proc Natl Acad Sci U S A*. Mar 27 2012;109(13):5074-9. doi:10.1073/pnas.1115402109

192. Ezzi SA, Urushitani M, Julien JP. Wild-type superoxide dismutase acquires binding and toxic properties of ALS-linked mutant forms through oxidation. *J Neurochem*. Jul 2007;102(1):170-8. doi:10.1111/j.1471-4159.2007.04531.x

193. Bosco DA, Morfini G, Karabacak NM, et al. Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nat Neurosci*. Nov 2010;13(11):1396-403. doi:10.1038/nn.2660

194. Forsberg K, Jonsson PA, Andersen PM, et al. Novel antibodies reveal inclusions containing non-native SOD1 in sporadic ALS patients. *PLoS One*. Jul 14 2010;5(7):e11552. doi:10.1371/journal.pone.0011552

195. Grad LI, Yerbury JJ, Turner BJ, et al. Intercellular propagated misfolding of wild-type Cu/Zn superoxide dismutase occurs via exosome-dependent and -independent mechanisms. *Proc Natl Acad Sci U S A*. Mar 04 2014;111(9):3620-5. doi:10.1073/pnas.1312245111

196. Urushitani M, Sik A, Sakurai T, Nukina N, Takahashi R, Julien JP. Chromogranin-mediated secretion of mutant superoxide dismutase proteins linked to amyotrophic lateral sclerosis. *Nat Neurosci*. Jan 2006;9(1):108-18. doi:10.1038/nn1603

197. Sundaramoorthy V, Walker AK, Yerbury J, et al. Extracellular wildtype and mutant SOD1 induces ER-Golgi pathology characteristic of amyotrophic lateral sclerosis in neuronal cells. *Cell Mol Life Sci*. Nov 2013;70(21):4181-95. doi:10.1007/s00018-013-1385-2

198. McAlary L, Plotkin SS, Yerbury JJ, Cashman NR. Prion-Like Propagation of Protein Misfolding and Aggregation in Amyotrophic Lateral Sclerosis. *Front Mol Neurosci*. 2019;12:262. doi:10.3389/fnmol.2019.00262

199. McAlary L, Yerbury JJ, Cashman NR. The prion-like nature of amyotrophic lateral sclerosis. *Prog Mol Biol Transl Sci*. 2020;175:261-296. doi:10.1016/bs.pmbts.2020.07.002

200. Wang Q, Johnson JL, Agar NY, Agar JN. Protein aggregation and protein instability govern familial amyotrophic lateral sclerosis patient survival. *PLoS Biol*. Jul 29 2008;6(7):e170. doi:10.1371/journal.pbio.0060170

201. Andersen PM. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr Neurol Neurosci Rep*. Jan 2006;6(1):37-46.

202. Cudkowicz ME, McKenna-Yasek D, Sapp PE, et al. Epidemiology of mutations in superoxide dismutase in amyotrophic lateral sclerosis. *Ann Neurol*. Feb 1997;41(2):210-21. doi:10.1002/ana.410410212

203. Andersen PM, Forsgren L, Binzer M, et al. Autosomal recessive adult-onset amyotrophic lateral sclerosis associated with homozygosity for Asp90Ala CuZn-superoxide dismutase mutation. A clinical and genealogical study of 36 patients. *Brain*. Aug 1996;119 (Pt 4):1153-72.

204. Turner MR, Hammers A, Al-Chalabi A, et al. Distinct cerebral lesions in sporadic and 'D90A' SOD1 ALS: studies with [11C]flumazenil PET. *Brain*. Jun 2005;128(Pt 6):1323-9. doi:10.1093/brain/awh509

205. Nakamura M, Bieniek KF, Lin WL, et al. A truncating SOD1 mutation, p.Gly141X, is associated with clinical and pathologic heterogeneity, including frontotemporal lobar degeneration. *Acta Neuropathol*. Jul 2015;130(1):145-57. doi:10.1007/s00401-015-1431-2

206. Katz JS, Katzberg HD, Woolley SC, Marklund SL, Andersen PM. Combined fulminant frontotemporal dementia and amyotrophic lateral sclerosis associated with an I113T SOD1 mutation. *Amyotroph Lateral Scler*. Oct 2012;13(6):567-9. doi:10.3109/17482968.2012.678365

207. Yu CE, Bird TD, Bekris LM, et al. The spectrum of mutations in progranulin: a collaborative study screening 545 cases of neurodegeneration. *Arch Neurol*. Feb 2010;67(2):161-70. doi:10.1001/archneurol.2009.328

208. Chen-Plotkin AS, Martinez-Lage M, Sleiman PM, et al. Genetic and clinical features of progranulin-associated frontotemporal lobar degeneration. *Arch Neurol*. Apr 2011;68(4):488-97. doi:10.1001/archneurol.2011.53

209. Baker M, Mackenzie IR, Pickering-Brown SM, et al. Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature*. Aug 24 2006;442(7105):916-9. doi:10.1038/nature05016

210. Galimberti D, Scarpini E. Genetics and biology of Alzheimer's disease and frontotemporal lobar degeneration. *Int J Clin Exp Med*. May 15 2010;3(2):129-43.

211. Rankin KP, Baldwin E, Pace-Savitsky C, Kramer JH, Miller BL. Self awareness and personality change in dementia. *J Neurol Neurosurg Psychiatry*. May 2005;76(5):632-9. doi:10.1136/jnnp.2004.042879

212. Walker AJ, Meares S, Sachdev PS, Brodaty H. The differentiation of mild frontotemporal dementia from Alzheimer's disease and healthy aging by neuropsychological tests. *Int Psychogeriatr*. Mar 2005;17(1):57-68.

213. Yin F, Dumont M, Banerjee R, et al. Behavioral deficits and progressive neuropathology in progranulin-deficient mice: a mouse model of frontotemporal dementia. *FASEB J*. Dec 2010;24(12):4639-47. doi:10.1096/fj.10-161471

214. Moretti DV, Benussi L, Fostinelli S, Ciani M, Binetti G, Ghidoni R. Progranulin Mutations Affects Brain Oscillatory Activity in Fronto-Temporal Dementia. *Front Aging Neurosci*. 2016;8:35. doi:10.3389/fnagi.2016.00035

215. Watts GD, Wymer J, Kovach MJ, et al. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet*. Apr 2004;36(4):377-81. doi:10.1038/ng1332

216. Johnson JO, Mandrioli J, Benatar M, et al. Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron*. Dec 09 2010;68(5):857-64. doi:10.1016/j.neuron.2010.11.036

217. Saracino D, Clot F, Camuzat A, et al. Novel VCP mutations expand the mutational spectrum of frontotemporal dementia. *Neurobiol Aging*. Dec 2018;72:187 e11-187 e14. doi:10.1016/j.neurobiolaging.2018.06.037

218. Bastola P, Bilkis R, De Souza C, Minn K, Chien J. Heterozygous mutations in valosin-containing protein (VCP) and resistance to VCP inhibitors. *Sci Rep*. Jul 29 2019;9(1):11002. doi:10.1038/s41598-019-47085-9

219. Al-Obeidi E, Al-Tahan S, Surampalli A, et al. Genotype-phenotype study in patients with valosin-containing protein mutations associated with multisystem proteinopathy. *Clin Genet*. Jan 2018;93(1):119-125. doi:10.1111/cge.13095

220. Twomey EC, Ji Z, Wales TE, et al. Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science*. Aug 2 2019;365(6452)doi:10.1126/science.aax1033

221. Ganji R, Mukkavalli S, Somanji F, Raman M. The VCP-UBXN1 Complex Mediates Triage of Ubiquitylated Cytosolic Proteins Bound to the BAG6 Complex. *Mol Cell Biol*. Jul 1 2018;38(13)doi:10.1128/MCB.00154-18

222. Plewa J, Surampalli A, Wencel M, et al. A cross-sectional analysis of clinical evaluation in 35 individuals with mutations of the valosin-containing protein gene. *Neuromuscul Disord*. Sep 2018;28(9):778-786. doi:10.1016/j.nmd.2018.06.007

223. Ogura T, Wilkinson AJ. AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells*. Jul 2001;6(7):575-97.

224. Ludtmann MH, Arber C, Bartolome F, et al. Mutations in valosin-containing protein (VCP) decrease ADP/ATP translocation across the mitochondrial membrane

and impair energy metabolism in human neurons. *J Biol Chem*. Mar 30 2017;doi:10.1074/jbc.M116.762898

225. Nalbandian A, Khan AA, Srivastava R, et al. Activation of the NLRP3 Inflammasome Is Associated with Valosin-Containing Protein Myopathy. *Inflammation*. Feb 2017;40(1):21-41. doi:10.1007/s10753-016-0449-5

226. Arhzaouy K, Papadopoulos C, Schulze N, Pittman SK, Meyer H, Weihl CC. VCP maintains lysosomal homeostasis and TFEB activity in differentiated skeletal muscle. *Autophagy*. Jun 2019;15(6):1082-1099. doi:10.1080/15548627.2019.1569933

227. Johnson JO, Pioro EP, Boehringer A, et al. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci*. May 2014;17(5):664-6. doi:10.1038/nn.3688

228. Marangi G, Lattante S, Doronzio PN, et al. Matrin 3 variants are frequent in Italian ALS patients. *Neurobiol Aging*. Jan 2017;49:218 e1-218 e7. doi:10.1016/j.neurobiolaging.2016.09.023

229. Rayaprolu S, D'Alton S, Crosby K, et al. Heterogeneity of Matrin 3 in the developing and aging murine central nervous system. *J Comp Neurol*. Oct 01 2016;524(14):2740-52. doi:10.1002/cne.23986

230. Cozzolino M, Rossi S, Mirra A, Carri MT. Mitochondrial dynamism and the pathogenesis of Amyotrophic Lateral Sclerosis. *Front Cell Neurosci*. 2015;9:31. doi:10.3389/fncel.2015.00031

231. Ajroud-Driss S, Fecto F, Ajroud K, et al. Mutation in the novel nuclear-encoded mitochondrial protein CHCHD10 in a family with autosomal dominant mitochondrial myopathy. *Neurogenetics*. Jan 2015;16(1):1-9. doi:10.1007/s10048-014-0421-1

232. Chaussenot A, Le Ber I, Ait-El-Mkadem S, et al. Screening of CHCHD10 in a French cohort confirms the involvement of this gene in frontotemporal dementia with

amyotrophic lateral sclerosis patients. *Neurobiol Aging*. Dec 2014;35(12):2884 e1-4. doi:10.1016/j.neurobiolaging.2014.07.022

233. Zhou ZD, Saw WT, Tan EK. Mitochondrial CHCHD-Containing Proteins: Physiologic Functions and Link with Neurodegenerative Diseases. *Mol Neurobiol*. Sep 2017;54(7):5534-5546. doi:10.1007/s12035-016-0099-5

234. Taylor JP, Brown RH, Jr., Cleveland DW. Decoding ALS: from genes to mechanism. *Nature*. Nov 10 2016;539(7628):197-206. doi:10.1038/nature20413

235. Perrone F, Nguyen HP, Van Mossevelde S, et al. Investigating the role of ALS genes CHCHD10 and TUBA4A in Belgian FTD-ALS spectrum patients. *Neurobiol Aging*. Mar 2017;51:177 e9-177 e16. doi:10.1016/j.neurobiolaging.2016.12.008

236. Xiao Y, Zhang J, Shu X, et al. Loss of mitochondrial protein CHCHD10 in skeletal muscle causes neuromuscular junction impairment. *Hum Mol Genet*. Jul 1 2019;doi:10.1093/hmg/ddz154

237. Bannwarth S, Ait-El-Mkadem S, Chaussenot A, et al. A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain*. Aug 2014;137(Pt 8):2329-45. doi:10.1093/brain/awu138

238. Burstein SR, Valsecchi F, Kawamata H, et al. In vitro and in vivo studies of the ALS-FTLD protein CHCHD10 reveal novel mitochondrial topology and protein interactions. *Hum Mol Genet*. Jan 1 2018;27(1):160-177. doi:10.1093/hmg/ddx397

239. Anderson CJ, Bredvik K, Burstein SR, et al. ALS/FTD mutant CHCHD10 mice reveal a tissue-specific toxic gain-of-function and mitochondrial stress response. *Acta Neuropathol*. Jul 2019;138(1):103-121. doi:10.1007/s00401-019-01989-y

240. Genin EC, Madji Hounoum B, Bannwarth S, et al. Mitochondrial defect in muscle precedes neuromuscular junction degeneration and motor neuron death in

CHCHD10(S59L/+) mouse. *Acta Neuropathol*. Jul 2019;138(1):123-145. doi:10.1007/s00401-019-01988-z

241. Cirulli ET, Lasseigne BN, Petrovski S, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*. Mar 27 2015;347(6229):1436-41. doi:10.1126/science.aaa3650

242. Freischmidt A, Wieland T, Richter B, et al. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat Neurosci*. May 2015;18(5):631-6. doi:10.1038/nn.4000

243. Weidberg H, Elazar Z. TBK1 mediates crosstalk between the innate immune response and autophagy. *Sci Signal*. Aug 09 2011;4(187):pe39. doi:10.1126/scisignal.2002355

244. Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. *Science*. Jan 23 2015;347(6220):1260419. doi:10.1126/science.1260419

245. Li F, Xie X, Wang Y, et al. Structural insights into the interaction and disease mechanism of neurodegenerative disease-associated optineurin and TBK1 proteins. *Nat Commun*. Sep 13 2016;7:12708. doi:10.1038/ncomms12708

246. Morton S, Hesson L, Peggie M, Cohen P. Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. *FEBS Lett*. Mar 19 2008;582(6):997-1002. doi:10.1016/j.febslet.2008.02.047

247. Ryzhakov G, Randow F. SINTBAD, a novel component of innate antiviral immunity, shares a TBK1-binding domain with NAP1 and TANK. *EMBO J*. Jul 11 2007;26(13):3180-90. doi:10.1038/sj.emboj.7601743

248. Lamb R, Rohrer JD, Real R, et al. A novel TBK1 mutation in a family with diverse frontotemporal dementia spectrum disorders. *Cold Spring Harb Mol Case Stud*. Jun 2019;5(3)doi:10.1101/mcs.a003913

249. Van Mossevelde S, van der Zee J, Gijselinck I, et al. Clinical features of TBK1 carriers compared with C9orf72, GRN and non-mutation carriers in a Belgian cohort. *Brain*. Feb 2016;139(Pt 2):452-67. doi:10.1093/brain/awv358

250. Weinreich M, Shepheard SR, Verber N, et al. Neuropathological characterization of a novel TANK binding kinase (TBK1) gene loss of function mutation associated with amyotrophic lateral sclerosis. *Neuropathol Appl Neurobiol*. Sep 9 2019;doi:10.1111/nan.12578

251. de Majo M, Topp SD, Smith BN, et al. ALS-associated missense and nonsense TBK1 mutations can both cause loss of kinase function. *Neurobiol Aging*. Nov 2018;71:266 e1-266 e10. doi:10.1016/j.neurobiolaging.2018.06.015

252. Bonnard M, Mirtsos C, Suzuki S, et al. Deficiency of T2K leads to apoptotic liver degeneration and impaired NF-kappaB-dependent gene transcription. *EMBO J*. Sep 15 2000;19(18):4976-85. doi:10.1093/emboj/19.18.4976

253. Duan W, Guo M, Yi L, et al. Deletion of Tbk1 disrupts autophagy and reproduces behavioral and locomotor symptoms of FTD-ALS in mice. *Aging (Albany NY)*. Apr 30 2019;11(8):2457-2476. doi:10.18632/aging.101936

254. Hu YW, Zhang J, Wu XM, Cao L, Nie P, Chang MX. TANK-Binding Kinase 1 (TBK1) Isoforms Negatively Regulate Type I Interferon Induction by Inhibiting TBK1-IRF3 Interaction and IRF3 Phosphorylation. *Front Immunol*. 2018;9:84. doi:10.3389/fimmu.2018.00084

255. Perry AK, Chow EK, Goodnough JB, Yeh WC, Cheng G. Differential requirement for TANK-binding kinase-1 in type I interferon responses to toll-like

receptor activation and viral infection. *J Exp Med*. Jun 21 2004;199(12):1651-8. doi:10.1084/jem.20040528

256. Wang R, Yang B, Zhang D. Activation of interferon signaling pathways in spinal cord astrocytes from an ALS mouse model. *Glia*. Jun 2011;59(6):946-58. doi:10.1002/glia.21167

257. Brenner D, Sieverding K, Bruno C, et al. Heterozygous Tbk1 loss has opposing effects in early and late stages of ALS in mice. *J Exp Med*. Feb 4 2019;216(2):267-278. doi:10.1084/jem.20180729

258. Hutton M, Lendon CL, Rizzu P, et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*. Jun 18 1998;393(6686):702-5. doi:10.1038/31508

259. Spillantini MG, Murrell JR, Goedert M, Farlow MR, Klug A, Ghetti B. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A*. Jun 23 1998;95(13):7737-41. doi:10.1073/pnas.95.13.7737

260. Weingarten MD, Lockwood AH, Hwo SY, Kirschner MW. A protein factor essential for microtubule assembly. *Proc Natl Acad Sci U S A*. May 1975;72(5):1858-62. doi:10.1073/pnas.72.5.1858

261. Witman GB, Cleveland DW, Weingarten MD, Kirschner MW. Tubulin requires tau for growth onto microtubule initiating sites. *Proc Natl Acad Sci U S A*. Nov 1976;73(11):4070-4. doi:10.1073/pnas.73.11.4070

262. Guo T, Noble W, Hanger DP. Roles of tau protein in health and disease. *Acta Neuropathol*. May 2017;133(5):665-704. doi:10.1007/s00401-017-1707-9

263. Ghetti B, Oblak AL, Boeve BF, Johnson KA, Dickerson BC, Goedert M. Invited review: Frontotemporal dementia caused by microtubule-associated protein tau gene

(MAPT) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol Appl Neurobiol*. Feb 2015;41(1):24-46. doi:10.1111/nan.12213

264. Cruts M, Theuns J, Van Broeckhoven C. Locus-specific mutation databases for neurodegenerative brain diseases. *Hum Mutat*. Sep 2012;33(9):1340-4. doi:10.1002/humu.22117

265. Spillantini MG, Goedert M. Tau pathology and neurodegeneration. *Lancet Neurol*. Jun 2013;12(6):609-22. doi:10.1016/S1474-4422(13)70090-5

266. Takada LT. The Genetics of Monogenic Frontotemporal Dementia. *Dement Neuropsychol*. Jul-Sep 2015;9(3):219-229. doi:10.1590/1980-57642015DN93000003

267. Strang KH, Golde TE, Giasson BI. MAPT mutations, tauopathy, and mechanisms of neurodegeneration. *Lab Invest*. Jul 2019;99(7):912-928. doi:10.1038/s41374-019-0197-x

268. Houlden H, Baker M, Adamson J, et al. Frequency of tau mutations in three series of non-Alzheimer's degenerative dementia. *Ann Neurol*. Aug 1999;46(2):243-8. doi:10.1002/1531-8249(199908)46:2<243::aid-ana14>3.0.co;2-l

269. Van Langenhove T, van der Zee J, Gijselinck I, et al. Distinct clinical characteristics of C9orf72 expansion carriers compared with GRN, MAPT, and nonmutation carriers in a Flanders-Belgian FTLD cohort. *JAMA Neurol*. Mar 1 2013;70(3):365-73. doi:10.1001/2013.jamaneurol.181

270. Binetti G, Nicosia F, Benussi L, et al. Prevalence of TAU mutations in an Italian clinical series of familial frontotemporal patients. *Neurosci Lett*. Feb 20 2003;338(1):85-7. doi:10.1016/s0304-3940(02)01330-7

271. Pickering-Brown S, Hutton M. The genetics of frontotemporal dementia. *Handb Clin Neurol*. 2008;89:383-92. doi:10.1016/S0072-9752(07)01236-5

272. Kaivorinne AL, Kruger J, Kuivaniemi K, et al. Role of MAPT mutations and haplotype in frontotemporal lobar degeneration in Northern Finland. *BMC Neurol*. Dec 17 2008;8:48. doi:10.1186/1471-2377-8-48

273. Ke YD, Chan G, Stefanoska K, et al. CNS cell type-specific gene profiling of P301S tau transgenic mice identifies genes dysregulated by progressive tau accumulation. *J Biol Chem*. Sep 20 2019;294(38):14149-14162. doi:10.1074/jbc.RA118.005263

274. Evans HT, Benetatos J, van Roijen M, Bodea LG, Gotz J. Decreased synthesis of ribosomal proteins in tauopathy revealed by non-canonical amino acid labelling. *EMBO J*. Jul 1 2019;38(13):e101174. doi:10.15252/embj.2018101174

275. Koller EJ, Gonzalez De La Cruz E, Machula T, et al. Combining P301L and S320F tau variants produces a novel accelerated model of tauopathy. *Hum Mol Genet*. Oct 1 2019;28(19):3255-3269. doi:10.1093/hmg/ddz151

276. Skibinski G, Parkinson NJ, Brown JM, et al. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet*. Aug 2005;37(8):806-8. doi:10.1038/ng1609

277. Parkinson N, Ince PG, Smith MO, et al. ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology*. Sep 26 2006;67(6):1074-7. doi:10.1212/01.wnl.0000231510.89311.8b

278. van der Zee J, Urwin H, Engelborghs S, et al. CHMP2B C-truncating mutations in frontotemporal lobar degeneration are associated with an aberrant endosomal phenotype in vitro. *Hum Mol Genet*. Jan 15 2008;17(2):313-22. doi:10.1093/hmg/ddm309

279. Chassefeyre R, Martinez-Hernandez J, Bertaso F, et al. Regulation of postsynaptic function by the dementia-related ESCRT-III subunit CHMP2B. *J Neurosci*. Feb 18 2015;35(7):3155-73. doi:10.1523/JNEUROSCI.0586-14.2015

280. Gascon E, Lynch K, Ruan H, et al. Alterations in microRNA-124 and AMPA receptors contribute to social behavioral deficits in frontotemporal dementia. *Nat Med*. Dec 2014;20(12):1444-51. doi:10.1038/nm.3717

281. Vernay A, Therreau L, Blot B, et al. A transgenic mouse expressing CHMP2Bintron5 mutant in neurons develops histological and behavioural features of amyotrophic lateral sclerosis and frontotemporal dementia. *Hum Mol Genet*. Aug 1 2016;25(15):3341-3360. doi:10.1093/hmg/ddw182

282. Waegaert R, Dirrig-Grosch S, Parisot F, et al. Longitudinal transcriptomic analysis of altered pathways in a CHMP2B(intron5)-based model of ALS-FTD. *Neurobiol Dis*. Mar 2020;136:104710. doi:10.1016/j.nbd.2019.104710

283. Ghazi-Noori S, Froud KE, Mizielinska S, et al. Progressive neuronal inclusion formation and axonal degeneration in CHMP2B mutant transgenic mice. *Brain*. Mar 2012;135(Pt 3):819-32. doi:10.1093/brain/aws006

284. Weishaupt JH, Hyman T, Dikic I. Common Molecular Pathways in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Trends Mol Med*. Sep 2016;22(9):769-83. doi:10.1016/j.molmed.2016.07.005

285. Spencer PS, Lagrange E, Camu W. ALS and environment: Clues from spatial clustering? *Rev Neurol (Paris)*. Dec 2019;175(10):652-663. doi:10.1016/j.neurol.2019.04.007

286. Bozzoni V, Pansarasa O, Diamanti L, Nosari G, Cereda C, Ceroni M. Amyotrophic lateral sclerosis and environmental factors. *Funct Neurol*. Jan-Mar 2016;31(1):7-19.

287. Belbasis L, Bellou V, Evangelou E. Environmental Risk Factors and Amyotrophic Lateral Sclerosis: An Umbrella Review and Critical Assessment of Current Evidence from Systematic Reviews and Meta-Analyses of Observational Studies. *Neuroepidemiology*. 2016;46(2):96-105. doi:10.1159/000443146

288. Armon C. Smoking may be considered an established risk factor for sporadic ALS. *Neurology*. Nov 17 2009;73(20):1693-8. doi:10.1212/WNL.0b013e3181c1df48

289. E M, Yu S, Dou J, et al. Association between alcohol consumption and amyotrophic lateral sclerosis: a meta-analysis of five observational studies. *Neurol Sci*. Aug 2016;37(8):1203-8. doi:10.1007/s10072-016-2575-0

290. Turner MR, Abisgold J, Yeates DG, Talbot K, Goldacre MJ. Head and other physical trauma requiring hospitalisation is not a significant risk factor in the development of ALS. *J Neurol Sci*. Jan 15 2010;288(1-2):45-8. doi:10.1016/j.jns.2009.10.010

291. Armon C, Nelson LM. Is head trauma a risk factor for amyotrophic lateral sclerosis? An evidence based review. *Amyotroph Lateral Scler*. Jun 2012;13(4):351-6. doi:10.3109/17482968.2012.660954

292. Bandres-Ciga S, Noyce AJ, Hemani G, et al. Shared polygenic risk and causal inferences in amyotrophic lateral sclerosis. *Ann Neurol*. Apr 2019;85(4):470-481. doi:10.1002/ana.25431

293. Beard JD, Kamel F. Military service, deployments, and exposures in relation to amyotrophic lateral sclerosis etiology and survival. *Epidemiol Rev*. 2015;37:55-70. doi:10.1093/epirev/mxu001

294. Sorahan T, Mohammed N. Neurodegenerative disease and magnetic field exposure in UK electricity supply workers. *Occup Med (Lond)*. Sep 2014;64(6):454-60. doi:10.1093/occmed/kqu105

295. Huss A, Peters S, Vermeulen R. Occupational exposure to extremely low-frequency magnetic fields and the risk of ALS: A systematic review and meta-analysis. *Bioelectromagnetics*. Feb 2018;39(2):156-163. doi:10.1002/bem.22104

296. Mitsumoto H, Garofalo DC, Gilmore M, et al. Case-control study in ALS using the National ALS Registry: lead and agricultural chemicals are potential risk factors. *Amyotroph Lateral Scler Frontotemporal Degener*. Jun 17 2021:1-13. doi:10.1080/21678421.2021.1936556

297. Capozzella A, Sacco C, Chighine A, et al. Work related etiology of amyotrophic lateral sclerosis (ALS): a meta-analysis. *Ann Ig*. Sep-Oct 2014;26(5):456-72. doi:10.7416/ai.2014.2005

298. Gunnarsson LG, Bodin L. Occupational Exposures and Neurodegenerative Diseases-A Systematic Literature Review and Meta-Analyses. *Int J Environ Res Public Health*. Jan 26 2019;16(3)doi:10.3390/ijerph16030337

299. Vinceti M, Filippini T, Violi F, et al. Pesticide exposure assessed through agricultural crop proximity and risk of amyotrophic lateral sclerosis. *Environ Health*. Aug 29 2017;16(1):91. doi:10.1186/s12940-017-0297-2

300. Malek AM, Barchowsky A, Bowser R, Youk A, Talbott EO. Pesticide exposure as a risk factor for amyotrophic lateral sclerosis: a meta-analysis of epidemiological studies: pesticide exposure as a risk factor for ALS. *Environ Res*. Aug 2012;117:112-9. doi:10.1016/j.envres.2012.06.007

301. Kamel F, Umbach DM, Bedlack RS, et al. Pesticide exposure and amyotrophic lateral sclerosis. *Neurotoxicology*. Jun 2012;33(3):457-62. doi:10.1016/j.neuro.2012.04.001

302. Armon C, Albert SM. A blow to the head trauma--ALS hypothesis. *Neurology*. Apr 28 2015;84(17):1728-9. doi:10.1212/WNL.0000000000001528

303. Abel EL. Football increases the risk for Lou Gehrig's disease, amyotrophic lateral sclerosis. *Percept Mot Skills*. Jun 2007;104(3 Pt 2):1251-4. doi:10.2466/pms.104.4.1251-1254

304. Miranda ML, Alicia Overstreet Galeano M, Tassone E, Allen KD, Horner RD. Spatial analysis of the etiology of amyotrophic lateral sclerosis among 1991 Gulf War veterans. *Neurotoxicology*. Nov 2008;29(6):964-70. doi:10.1016/j.neuro.2008.05.005

305. Horner RD, Grambow SC, Coffman CJ, et al. Amyotrophic lateral sclerosis among 1991 Gulf War veterans: evidence for a time-limited outbreak. *Neuroepidemiology*. 2008;31(1):28-32. doi:10.1159/000136648

306. Julian TH, Glascow N, Barry ADF, et al. Physical exercise is a risk factor for amyotrophic lateral sclerosis: Convergent evidence from Mendelian randomisation, transcriptomics and risk genotypes. *EBioMedicine*. Jun 2021;68:103397. doi:10.1016/j.ebiom.2021.103397

307. Reich-Slotky R, Andrews J, Cheng B, et al. Body mass index (BMI) as predictor of ALSFRS-R score decline in ALS patients. *Amyotroph Lateral Scler Frontotemporal Degener*. Apr 2013;14(3):212-6. doi:10.3109/21678421.2013.770028

308. Al-Chalabi A, Calvo A, Chio A, et al. Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. *Lancet Neurol*. Nov 2014;13(11):1108-1113. doi:10.1016/S1474-4422(14)70219-4

309. Chio A, Mazzini L, D'Alfonso S, et al. The multistep hypothesis of ALS revisited: The role of genetic mutations. *Neurology*. Aug 14 2018;91(7):e635-e642. doi:10.1212/WNL.0000000000005996

310. Bradley WG, Andrew AS, Traynor BJ, Chio A, Butt TH, Stommel EW. Gene-Environment-Time Interactions in Neurodegenerative Diseases: Hypotheses and

Research Approaches. *Ann Neurosci*. Dec 2018;25(4):261-267. doi:10.1159/000495321

311. Coppede F, Stoccoro A, Mosca L, et al. Increase in DNA methylation in patients with amyotrophic lateral sclerosis carriers of not fully penetrant SOD1 mutations. *Amyotroph Lateral Scler Frontotemporal Degener*. Feb 2018;19(1-2):93-101. doi:10.1080/21678421.2017.1367401

312. Morahan JM, Yu B, Trent RJ, Pamphlett R. A genome-wide analysis of brain DNA methylation identifies new candidate genes for sporadic amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. Oct-Dec 2009;10(5-6):418-29. doi:10.3109/17482960802635397

313. Figueroa-Romero C, Hur J, Bender DE, et al. Identification of epigenetically altered genes in sporadic amyotrophic lateral sclerosis. *PLoS One*. 2012;7(12):e52672. doi:10.1371/journal.pone.0052672

314. Tremolizzo L, Messina P, Conti E, et al. Whole-blood global DNA methylation is increased in amyotrophic lateral sclerosis independently of age of onset. *Amyotroph Lateral Scler Frontotemporal Degener*. Mar 2014;15(1-2):98-105. doi:10.3109/21678421.2013.851247

315. Belzil VV, Bauer PO, Prudencio M, et al. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol*. Dec 2013;126(6):895-905. doi:10.1007/s00401-013-1199-1

316. Figueroa-Romero C, Hur J, Lunn JS, et al. Expression of microRNAs in human post-mortem amyotrophic lateral sclerosis spinal cords provides insight into disease mechanisms. *Mol Cell Neurosci*. Mar 2016;71:34-45. doi:10.1016/j.mcn.2015.12.008

317. Marcuzzo S, Bonanno S, Kapetis D, et al. Up-regulation of neural and cell cycle-related microRNAs in brain of amyotrophic lateral sclerosis mice at late disease stage. *Mol Brain*. Jan 28 2015;8:5. doi:10.1186/s13041-015-0095-0

318. Toivonen JM, Manzano R, Olivan S, Zaragoza P, Garcia-Redondo A, Osta R. MicroRNA-206: a potential circulating biomarker candidate for amyotrophic lateral sclerosis. *PLoS One*. 2014;9(2):e89065. doi:10.1371/journal.pone.0089065

319. Zhang Z, Almeida S, Lu Y, et al. Downregulation of microRNA-9 in iPSC-derived neurons of FTD/ALS patients with TDP-43 mutations. *PLoS One*. 2013;8(10):e76055. doi:10.1371/journal.pone.0076055

320. Griffiths BB, Hunter RG. Neuroepigenetics of stress. *Neuroscience*. Sep 5 2014;275:420-35. doi:10.1016/j.neuroscience.2014.06.041

321. Hunter RG, McEwen BS. Stress and anxiety across the lifespan: structural plasticity and epigenetic regulation. *Epigenomics*. Apr 2013;5(2):177-94. doi:10.2217/epi.13.8

322. Reul JM. Making memories of stressful events: a journey along epigenetic, gene transcription, and signaling pathways. *Front Psychiatry*. 2014;5:5. doi:10.3389/fpsyt.2014.00005

323. J. O. Johnson RC, D. E. Miller, R. Li, Y. Abramzon et al Mutations in the SPTLC1 gene are a cause of juvenile amyotrophic lateral sclerosis that may be amenable to serine supplementation. *bioRxiv*. 2020;doi:https://doi.org/10.1101/770339

324. Teoh HL, Carey K, Sampaio H, Mowat D, Roscioli T, Farrar M. Inherited Paediatric Motor Neuron Disorders: Beyond Spinal Muscular Atrophy. *Neural Plast*. 2017;2017:6509493. doi:10.1155/2017/6509493

325. Harms MB, Baloh RH. Clinical neurogenetics: amyotrophic lateral sclerosis. *Neurol Clin*. Nov 2013;31(4):929-50. doi:10.1016/j.ncl.2013.05.003

326. Orlacchio A, Babalini C, Borreca A, et al. SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain*. Feb 2010;133(Pt 2):591-8. doi:10.1093/brain/awp325

327. Liu ZJ, Lin HX, Liu GL, et al. The investigation of genetic and clinical features in Chinese patients with juvenile amyotrophic lateral sclerosis. *Clin Genet*. Sep 2017;92(3):267-273. doi:10.1111/cge.13015

328. Hentati A, Bejaoui K, Pericak-Vance MA, et al. Linkage of recessive familial amyotrophic lateral sclerosis to chromosome 2q33-q35. *Nat Genet*. Jul 1994;7(3):425-8. doi:10.1038/ng0794-425

329. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. Sep 2007;81(3):559-75. doi:10.1086/519795

330. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. Sep 2010;38(16):e164. doi:10.1093/nar/gkq603

331. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. Aug 18 2016;536(7616):285-91. doi:10.1038/nature19057

332. Haplotype Reference Consortium (HRC).

333. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. Jul 2016;18(7):696-704. doi:10.1038/gim.2015.148

334. Rossi F, Bruno G, Fratta M, et al. Expanding the spectrum of SPTLC1-related disorders beyond hereditary sensory and autonomic neuropathies: A novel case of the

distinct "S331 syndrome". *J Peripher Nerv Syst*. Sep 2020;25(3):308-311. doi:10.1111/jns.12394

335. Mohassel P, Donkervoort S, Lone MA, et al. Childhood amyotrophic lateral sclerosis caused by excess sphingolipid synthesis. *Nat Med*. Jul 2021;27(7):1197-1204. doi:10.1038/s41591-021-01346-1

336. Project Min EALSSC. Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. *Eur J Hum Genet*. Oct 2018;26(10):1537-1546. doi:10.1038/s41431-018-0177-4

337. Chio A, Logroscino G, Traynor BJ, et al. Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology*. 2013;41(2):118-30. doi:10.1159/000351153

338. Anderson D, Lassmann T. A phenotype centric benchmark of variant prioritisation tools. *NPJ Genom Med*. 2018;3:5. doi:10.1038/s41525-018-0044-9

339. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data. *Am J Hum Genet*. Oct 4 2018;103(4):522-534. doi:10.1016/j.ajhg.2018.08.016

340. Merrill AH, Jr., Sullards MC, Allegood JC, Kelly S, Wang E. Sphingolipidomics: high-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry. *Methods*. Jun 2005;36(2):207-24. doi:10.1016/j.ymeth.2005.01.009

341. Mohassel P, Donkervoort S, Lone MA, et al. Childhood amyotrophic lateral sclerosis caused by excess sphingolipid synthesis. *Nat Med*. Jul 2021;27(7):1197-1204. doi:10.1038/s41591-021-01346-1

342. Rutti MF, Richard S, Penno A, von Eckardstein A, Hornemann T. An improved method to determine serine palmitoyltransferase activity. *J Lipid Res*. Jun 2009;50(6):1237-44. doi:10.1194/jlr.D900001-JLR200

343. Garofalo K, Penno A, Schmidt BP, et al. Oral L-serine supplementation reduces production of neurotoxic deoxysphingolipids in mice and humans with hereditary sensory autonomic neuropathy type 1. *J Clin Invest*. Dec 2011;121(12):4735-45. doi:10.1172/JCI57549

344. Penno A, Reilly MM, Houlden H, et al. Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids. *J Biol Chem*. Apr 9 2010;285(15):11178-87. doi:10.1074/jbc.M109.092973

345. Dawkins JL, Hulme DJ, Brahmbhatt SB, Auer-Grumbach M, Nicholson GA. Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. *Nat Genet*. Mar 2001;27(3):309-12. doi:10.1038/85879

346. Rotthier A, Baets J, De Vriendt E, et al. Genes for hereditary sensory and autonomic neuropathies: a genotype-phenotype correlation. *Brain*. Oct 2009;132(Pt 10):2699-711. doi:10.1093/brain/awp198

347. Hanada K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochim Biophys Acta*. Jun 10 2003;1632(1-3):16-30.

348. Zhao L, Spassieva S, Gable K, et al. Elevation of 20-carbon long chain bases due to a mutation in serine palmitoyltransferase small subunit b results in neurodegeneration. *Proc Natl Acad Sci U S A*. Oct 20 2015;112(42):12962-7. doi:10.1073/pnas.1516733112

349. Sabourdy F, Astudillo L, Colacios C, et al. Monogenic neurological disorders of sphingolipid metabolism. *Biochim Biophys Acta*. Aug 2015;1851(8):1040-51. doi:10.1016/j.bbalip.2015.01.010

350. Haughey NJ, Bandaru VV, Bae M, Mattson MP. Roles for dysfunctional sphingolipid metabolism in Alzheimer's disease neuropathogenesis. *Biochim Biophys Acta*. Aug 2010;1801(8):878-86. doi:10.1016/j.bbalip.2010.05.003

351. Henriques A, Croixmarie V, Priestman DA, et al. Amyotrophic lateral sclerosis and denervation alter sphingolipids and up-regulate glucosylceramide synthase. *Hum Mol Genet*. Dec 20 2015;24(25):7390-405. doi:10.1093/hmg/ddv439

352. Drory VE, Birnbaum M, Peleg L, Goldman B, Korczyn AD. Hexosaminidase A deficiency is an uncommon cause of a syndrome mimicking amyotrophic lateral sclerosis. *Muscle Nerve*. Jul 2003;28(1):109-12. doi:10.1002/mus.10371

353. Levine TD, Miller RG, Bradley WG, et al. Phase I clinical trial of safety of L-serine for ALS patients. *Amyotroph Lateral Scler Frontotemporal Degener*. Feb 2017;18(1-2):107-111. doi:10.1080/21678421.2016.1221971

354. Green P, Wiseman M, Crow YJ, et al. Brown-Vialetto-Van Laere syndrome, a ponto-bulbar palsy with deafness, is caused by mutations in c20orf54. *Am J Hum Genet*. Mar 12 2010;86(3):485-9. doi:10.1016/j.ajhg.2010.02.006

355. Johnson JO, Gibbs JR, Megarbane A, et al. Exome sequencing reveals riboflavin transporter mutations as a cause of motor neuron disease. *Brain*. Sep 2012;135(Pt 9):2875-82. doi:10.1093/brain/aws161

356. Johnson JO, Gibbs JR, Van Maldergem L, Houlden H, Singleton AB. Exome sequencing in Brown-Vialetto-van Laere syndrome. *Am J Hum Genet*. Oct 08 2010;87(4):567-9; author reply 569-70. doi:10.1016/j.ajhg.2010.05.021

357. Caress JB, Johnson JO, Abramzon YA, et al. Exome sequencing establishes a gelsolin mutation as the cause of inherited bulbar-onset neuropathy. *Muscle Nerve*. Nov 2017;56(5):1001-1005. doi:10.1002/mus.25550

358. Haltia M, Prelli F, Ghiso J, et al. Amyloid protein in familial amyloidosis (Finnish type) is homologous to gelsolin, an actin-binding protein. *Biochem Biophys Res Commun*. Mar 30 1990;167(3):927-32.

359. Meretoja J. Familial systemic paramyloidosis with lattice dystrophy of the cornea, progressive cranial neuropathy, skin changes and various internal symptoms. A previously unrecognized heritable syndrome. *Ann Clin Res*. Dec 1969;1(4):314-24.

360. Asahina A, Yokoyama T, Ueda M, et al. Hereditary gelsolin amyloidosis: a new Japanese case with cutis laxa as a diagnostic clue. *Acta Derm Venereol*. Mar 2011;91(2):201-3. doi:10.2340/00015555-1011

361. Rothstein A, Auran JD, Wittpenn JR, Koester CJ, Florakis GJ. Confocal microscopy in Meretoja syndrome. *Cornea*. May 2002;21(4):364-7. doi:10.1097/00003226-200205000-00007

362. Spuler S, Emslie-Smith A, Engel AG. Amyloid myopathy: an underdiagnosed entity. *Ann Neurol*. Jun 1998;43(6):719-28. doi:10.1002/ana.410430606

363. Ardalan MR, Shoja MM, Kiuru-Enari S. Amyloidosis-related nephrotic syndrome due to a G654A gelsolin mutation: the first report from the Middle East. *Nephrol Dial Transplant*. Jan 2007;22(1):272-5. doi:10.1093/ndt/gfl548

364. McGough A, Chiu W, Way M. Determination of the gelsolin binding site on F-actin: implications for severing and capping. *Biophys J*. Feb 1998;74(2 Pt 1):764-72. doi:10.1016/S0006-3495(98)74001-9

365. Sadzynski A, Kurek K, Kononczuk T, Zendzian-Piotrowska M. [Gelsolin - variety of structure and functions]. *Postepy Hig Med Dosw (Online)*. Jun 18 2010;64:303-9. Gelsolina - roznorodnosc budowy i funkcji.

366. Kangas H, Paunio T, Kalkkinen N, Jalanko A, Peltonen L. In vitro expression analysis shows that the secretory form of gelsolin is the sole source of amyloid in gelsolin-related amyloidosis. *Hum Mol Genet*. Sep 1996;5(9):1237-43.

367. Paunio T, Kangas H, Kalkkinen N, Haltia M, Palo J, Peltonen L. Toward understanding the pathogenic mechanisms in gelsolin-related amyloidosis: in vitro expression reveals an abnormal gelsolin fragment. *Hum Mol Genet*. Dec 1994;3(12):2223-9.

368. Ratnaswamy G, Huff ME, Su AI, Rion S, Kelly JW. Destabilization of Ca2+-free gelsolin may not be responsible for proteolysis in Familial Amyloidosis of Finnish Type. *Proc Natl Acad Sci U S A*. Feb 27 2001;98(5):2334-9. doi:10.1073/pnas.041452598

369. Chen CD, Huff ME, Matteson J, et al. Furin initiates gelsolin familial amyloidosis in the Golgi through a defect in Ca(2+) stabilization. *EMBO J*. Nov 15 2001;20(22):6277-87. doi:10.1093/emboj/20.22.6277

370. Page LJ, Suk JY, Huff ME, et al. Metalloendoprotease cleavage triggers gelsolin amyloidogenesis. *EMBO J*. Dec 7 2005;24(23):4124-32. doi:10.1038/sj.emboj.7600872

371. Page LJ, Huff ME, Kelly JW, Balch WE. Ca2+ binding protects against gelsolin amyloidosis. *Biochem Biophys Res Commun*. Oct 1 2004;322(4):1105-10. doi:10.1016/j.bbrc.2004.07.125

372. Pihlamaa T, Rautio J, Kiuru-Enari S, Suominen S. Gelsolin amyloidosis as a cause of early aging and progressive bilateral facial paralysis. *Plast Reconstr Surg*. Jun 2011;127(6):2342-51. doi:10.1097/PRS.0b013e318213a0a2

373. Solomon JP, Page LJ, Balch WE, Kelly JW. Gelsolin amyloidosis: genetics, biochemistry, pathology and possible strategies for therapeutic intervention. *Crit Rev Biochem Mol Biol*. May-Jun 2012;47(3):282-96. doi:10.3109/10409238.2012.661401

374. Boysen G, Galassi G, Kamieniecka Z, Schlaeger J, Trojaborg W. Familial amyloidosis with cranial neuropathy and corneal lattice dystrophy. *J Neurol Neurosurg Psychiatry*. Nov 1979;42(11):1020-30. doi:10.1136/jnnp.42.11.1020

375. Warde-Farley D, Brudno M, Morris Q, Goldenberg A. Mixture model for sub-phenotyping in GWAS. *Pac Symp Biocomput*. 2012:363-74.

376. Yang JJ, Williams LK, Buu A. Identifying Pleiotropic Genes in Genome-Wide Association Studies for Multivariate Phenotypes with Mixed Measurement Scales. *PLoS One*. 2017;12(1):e0169893. doi:10.1371/journal.pone.0169893

377. O'Reilly PF, Hoggart CJ, Pomyen Y, et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One*. 2012;7(5):e34861. doi:10.1371/journal.pone.0034861

378. Fusi N, Lippert C, Lawrence ND, Stegle O. Warped linear mixed models for the genetic analysis of transformed phenotypes. *Nat Commun*. Sep 19 2014;5:4890. doi:10.1038/ncomms5890

379. Shafquat A, Crystal RG, Mezey JG. Identifying novel associations in GWAS by hierarchical Bayesian latent variable detection of differentially misclassified phenotypes. *BMC Bioinformatics*. May 7 2020;21(1):178. doi:10.1186/s12859-020-3387-z

380. Majka J, Burgers PM. The PCNA-RFC families of DNA clamps and clamp loaders. *Prog Nucleic Acid Res Mol Biol*. 2004;78:227-60. doi:10.1016/S0079-6603(04)78006-X

381. Tomida J, Masuda Y, Hiroaki H, et al. DNA damage-induced ubiquitylation of RFC2 subunit of replication factor C complex. *J Biol Chem*. Apr 4 2008;283(14):9071-9. doi:10.1074/jbc.M709835200

382. Dahlen M, Sunnerhagen P, Wang TS. Replication proteins influence the maintenance of telomere length and telomerase protein stability. *Mol Cell Biol*. May 2003;23(9):3031-42. doi:10.1128/mcb.23.9.3031-3042.2003

383. Cortese A, Simone R, Sullivan R, et al. Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia. *Nat Genet*. Apr 2019;51(4):649-658. doi:10.1038/s41588-019-0372-4

384. Akcimen F, Ross JP, Bourassa CV, et al. Investigation of the RFC1 Repeat Expansion in a Canadian and a Brazilian Ataxia Cohort: Identification of Novel Conformations. *Front Genet*. 2019;10:1219. doi:10.3389/fgene.2019.01219

385. Lozano R, Rosero CA, Hagerman RJ. Fragile X spectrum disorders. *Intractable Rare Dis Res*. Nov 2014;3(4):134-46. doi:10.5582/irdr.2014.01022

386. Sabatelli M, Conforti FL, Zollino M, et al. C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiol Aging*. Aug 2012;33(8):1848 e15-20. doi:10.1016/j.neurobiolaging.2012.02.011

387. Gijselinck I, Van Langenhove T, van der Zee J, 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*. Jan 2012;11(1):54-65. doi:10.1016/S1474-4422(11)70261-7

388. Chio A, Borghero G, Restagno G, et al. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. *Brain*. Mar 2012;135(Pt 3):784-93. doi:10.1093/brain/awr366

389. Cooper-Knock J, Hewitt C, Highley JR, et al. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain*. Mar 2012;135(Pt 3):751-64. doi:10.1093/brain/awr365

390. Majounie E, Renton AE, Mok K, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol*. Apr 2012;11(4):323-30. doi:10.1016/S1474-4422(12)70043-1

391. Mok KY, Koutsis G, Schottlaender LV, Polke J, Panas M, Houlden H. High frequency of the expanded C9ORF72 hexanucleotide repeat in familial and sporadic Greek ALS patients. *Neurobiol Aging*. Aug 2012;33(8):1851 e1-5. doi:10.1016/j.neurobiolaging.2012.02.021

392. Millecamps S, Boillee S, Le Ber I, et al. Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J Med Genet*. Apr 2012;49(4):258-63. doi:10.1136/jmedgenet-2011-100699

393. Ebbert MTW, Farrugia SL, Sens JP, et al. Long-read sequencing across the C9orf72 'GGGGCC' repeat expansion: implications for clinical use and genetic discovery efforts in human disease. *Mol Neurodegener*. Aug 21 2018;13(1):46. doi:10.1186/s13024-018-0274-4

394. Mitsuhashi S, Frith MC, Mizuguchi T, et al. Tandem-genotypes: robust detection of tandem repeat expansions from long DNA reads. *Genome Biol*. Mar 19 2019;20(1):58. doi:10.1186/s13059-019-1667-6

395. Gong L, Wong CH, Idol J, Ngan CY, Wei CL. Ultra-long Read Sequencing for Whole Genomic DNA Analysis. *J Vis Exp*. Mar 15 2019;(145)doi:10.3791/58954

396. Shafin K, Pesout T, Lorig-Roach R, et al. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nat Biotechnol*. Sep 2020;38(9):1044-1053. doi:10.1038/s41587-020-0503-6

397. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. Aug 15 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352

398. Loman NJ, Quick J, Simpson JT. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods*. Aug 2015;12(8):733-5. doi:10.1038/nmeth.3444

399. Gilpatrick T, Lee I, Graham JE, et al. Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nat Biotechnol*. Apr 2020;38(4):433-438. doi:10.1038/s41587-020-0407-5

400. Kovaka S, Fan Y, Ni B, Timp W, Schatz MC. Targeted nanopore sequencing by real-time mapping of raw electrical signal with UNCALLED. *Nat Biotechnol*. Apr 2021;39(4):431-441. doi:10.1038/s41587-020-0731-9

401. Payne A, Holmes N, Clarke T, Munro R, Debebe BJ, Loose M. Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nat Biotechnol*. Apr 2021;39(4):442-450. doi:10.1038/s41587-020-00746-x

402. Edwards HS, Krishnakumar R, Sinha A, Bird SW, Patel KD, Bartsch MS. Real-Time Selective Sequencing with RUBRIC: Read Until with Basecall and Reference-Informed Criteria. *Sci Rep*. Aug 7 2019;9(1):11475. doi:10.1038/s41598-019-47857-3

403. Dewan R, Chia R, Ding J, et al. Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. *Neuron*. Feb 3 2021;109(3):448-460 e4. doi:10.1016/j.neuron.2020.11.005

404. Chia R, Sabir MS, Bandres-Ciga S, et al. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. *Nat Genet*. Mar 2021;53(3):294-303. doi:10.1038/s41588-021-00785-3

405. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology*. Dec 27 2005;65(12):1863-72. doi:10.1212/01.wnl.0000187889.17253.b1

406. McKeith I. Dementia with Lewy bodies. *Dialogues Clin Neurosci*. Sep 2004;6(3):333-41.

407. Khachaturian ZS. Diagnosis of Alzheimer's disease. *Arch Neurol*. Nov 1985;42(11):1097-105. doi:10.1001/archneur.1985.04060100083029

408. Denson MA, Wszolek ZK, Pfeiffer RF, Wszolek EK, Paschall TM, McComb RD. Familial parkinsonism, dementia, and Lewy body disease: study of family G. *Ann Neurol*. Oct 1997;42(4):638-43. doi:10.1002/ana.410420415

409. Ishikawa A, Takahashi H, Tanaka H, Hayashi T, Tsuji S. Clinical features of familial diffuse Lewy body disease. *Eur Neurol*. 1997;38 Suppl 1:34-8.

410. Wakabayashi K, Hayashi S, Ishikawa A, et al. Autosomal dominant diffuse Lewy body disease. *Acta Neuropathol*. Aug 1998;96(2):207-10.

411. Ohara K, Takauchi S, Kokai M, Morimura Y, Nakajima T, Morita Y. Familial dementia with Lewy bodies (DLB). *Clin Neuropathol*. Sep-Oct 1999;18(5):232-9.

412. Knopman DS, Roberts RO. Estimating the number of persons with frontotemporal lobar degeneration in the US population. *J Mol Neurosci*. Nov 2011;45(3):330-5. doi:10.1007/s12031-011-9538-y

413. Hardy J, Rogaeva E. Motor neuron disease and frontotemporal dementia: sometimes related, sometimes not. *Exp Neurol*. Dec 2014;262 Pt B:75-83. doi:10.1016/j.expneurol.2013.11.006

414. Ratnavalli E, Brayne C, Dawson K, Hodges JR. The prevalence of frontotemporal dementia. *Neurology*. Jun 11 2002;58(11):1615-21.

415. Neary D, Snowden JS, Gustafson L, et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*. Dec 1998;51(6):1546-54.

416. Verkerk AJ, Pieretti M, Sutcliffe JS, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*. May 31 1991;65(5):905-14. doi:10.1016/0092-8674(91)90397-h

417. Depienne C, Mandel JL. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *Am J Hum Genet*. May 6 2021;108(5):764-785. doi:10.1016/j.ajhg.2021.03.011

418. Paulson H. Repeat expansion diseases. *Handb Clin Neurol*. 2018;147:105-123. doi:10.1016/B978-0-444-63233-3.00009-9

419. Semaka A, Creighton S, Warby S, Hayden MR. Predictive testing for Huntington disease: interpretation and significance of intermediate alleles. *Clin Genet*. Oct 2006;70(4):283-94. doi:10.1111/j.1399-0004.2006.00668.x

420. Rubino E, Mancini C, Boschi S, et al. ATXN2 intermediate repeat expansions influence the clinical phenotype in frontotemporal dementia. *Neurobiol Aging*. Jan 2019;73:231 e7-231 e9. doi:10.1016/j.neurobiolaging.2018.09.009

421. Entezari A, Khaniani MS, Bahrami T, Derakhshan SM, Darvish H. Screening for intermediate CGG alleles of FMR1 gene in male Iranian patients with Parkinsonism. *Neurol Sci*. Jan 2017;38(1):123-128. doi:10.1007/s10072-016-2723-6

422. Liu Q, Tong Y, Wang K. Genome-wide detection of short tandem repeat expansions by long-read sequencing. *BMC Bioinformatics*. Dec 28 2020;21(Suppl 21):542. doi:10.1186/s12859-020-03876-w

423. Jorge P, Oliveira B, Marques I, Santos R. Development and validation of a multiplex-PCR assay for X-linked intellectual disability. *BMC Med Genet*. Aug 5 2013;14:80. doi:10.1186/1471-2350-14-80

424. La Spada A. Spinal and Bulbar Muscular Atrophy. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

425. Veneziano L, Frontali M. Drpla. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

426. Opal P, Ashizawa T. Spinocerebellar Ataxia Type 1. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

427. Pulst SM. Spinocerebellar Ataxia Type 2. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

428. Paulson H, Shakkottai V. Spinocerebellar Ataxia Type 3. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

429. La Spada AR. Spinocerebellar Ataxia Type 7. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

430. Ayhan F, Ikeda Y, Dalton JC, Day JW, Ranum LPW. Spinocerebellar Ataxia Type 8. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

431. Matsuura T, Ashizawa T. Spinocerebellar Ataxia Type 10. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

432. Gossye H, Engelborghs S, Van Broeckhoven C, van der Zee J. C9orf72 Frontotemporal Dementia and/or Amyotrophic Lateral Sclerosis. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

433. Casey HL, Gomez CM. Spinocerebellar Ataxia Type 6. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

434. Schoser B. Myotonic Dystrophy Type 2. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

435. Lehesjoki AE, Kalviainen R. Progressive Myoclonic Epilepsy Type 1. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

436. Bird TD. Myotonic Dystrophy Type 1. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

437. Hunter JE, Berry-Kravis E, Hipp H, Todd PK. FMR1 Disorders. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

438. Bidichandani SI, Delatycki MB. Friedreich Ataxia. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

439. Caron NS, Wright GEB, Hayden MR. Huntington Disease. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

440. Anderson DG, Krause A, Margolis RL. Huntington Disease-Like 2. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

441. Arias M, Quintans B, Garcia-Murias M, Sobrido MJ. Spinocerebellar Ataxia Type 36 - RETIRED CHAPTER, FOR HISTORICAL REFERENCE ONLY. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

442. Fujigasaki H, Verma IC, Camuzat A, et al. SCA12 is a rare locus for autosomal dominant cerebellar ataxia: a study of an Indian family. *Ann Neurol*. Jan 2001;49(1):117-21.

443. Toyoshima Y, Onodera O, Yamada M, Tsuji S, Takahashi H. Spinocerebellar Ataxia Type 17. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.

444. Adelson RP, Renton AE, Li W, et al. Empirical design of a variant quality control pipeline for whole genome sequencing data using replicate discordance. *Sci Rep*. Nov 6 2019;9(1):16156. doi:10.1038/s41598-019-52614-7

445. Gardiner SL, Boogaard MW, Trompet S, et al. Prevalence of Carriers of Intermediate and Pathological Polyglutamine Disease-Associated Alleles Among Large Population-Based Cohorts. *JAMA Neurol*. Jun 1 2019;76(6):650-656. doi:10.1001/jamaneurol.2019.0423

446. Dolzhenko E, van Vugt J, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res*. Nov 2017;27(11):1895-1903. doi:10.1101/gr.225672.117

447. van der Sanden BPGH, Corominas J, de Groot M, et al. Systematic analysis of short tandem repeats in 38,095 exomes provides an additional diagnostic yield. *Genet Med*. 08 2021;23(8):1569-1573. doi:10.1038/s41436-021-01174-1

448. Elden AC, Kim HJ, Hart MP, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. Aug 26 2010;466(7310):1069-75. doi:10.1038/nature09320

449. Lattante S, Millecamps S, Stevanin G, et al. Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. *Neurology*. Sep 9 2014;83(11):990-5. doi:10.1212/WNL.0000000000000778

450. Dionisio PA, Amaral JD, Ribeiro MF, Lo AC, D'Hooge R, Rodrigues CM. Amyloid-beta pathology is attenuated by tauroursodeoxycholic acid treatment in

APP/PS1 mice after disease onset. *Neurobiol Aging*. Jan 2015;36(1):228-40. doi:10.1016/j.neurobiolaging.2014.08.034

451. Kozlov G, Trempe JF, Khaleghpour K, Kahvejian A, Ekiel I, Gehring K. Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proc Natl Acad Sci U S A*. Apr 10 2001;98(8):4409-13. doi:10.1073/pnas.071024998

452. Vucic S, Rothstein JD, Kiernan MC. Advances in treating amyotrophic lateral sclerosis: insights from pathophysiological studies. *Trends Neurosci*. Aug 2014;37(8):433-42. doi:10.1016/j.tins.2014.05.006

453. Farg MA, Soo KY, Warraich ST, Sundaramoorthy V, Blair IP, Atkin JD. Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. *Hum Mol Genet*. Feb 15 2013;22(4):717-28. doi:10.1093/hmg/dds479

454. Ciura S, Sellier C, Campanari ML, Charlet-Berguerand N, Kabashi E. The most prevalent genetic cause of ALS-FTD, C9orf72 synergizes the toxicity of ATXN2 intermediate polyglutamine repeats through the autophagy pathway. *Autophagy*. Aug 2 2016;12(8):1406-8. doi:10.1080/15548627.2016.1189070

455. Becker LA, Huang B, Bieri G, et al. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature*. Apr 20 2017;544(7650):367-371. doi:10.1038/nature22038

456. Shan DE, Soong BW, Sun CM, Lee SJ, Liao KK, Liu RS. Spinocerebellar ataxia type 2 presenting as familial levodopa-responsive parkinsonism. *Ann Neurol*. Dec 2001;50(6):812-5. doi:10.1002/ana.10055

457. Lee T, Li YR, Ingre C, et al. Ataxin-2 intermediate-length polyglutamine expansions in European ALS patients. *Hum Mol Genet*. May 01 2011;20(9):1697-700. doi:10.1093/hmg/ddr045

458. Naruse H, Matsukawa T, Ishiura H, et al. Association of ATXN2 intermediate-length CAG repeats with amyotrophic lateral sclerosis correlates with the distributions of normal CAG repeat alleles among individual ethnic populations. *Neurogenetics*. 05 2019;20(2):65-71. doi:10.1007/s10048-019-00570-9

459. Amado DA, Davidson BL. Gene therapy for ALS: A review. *Mol Ther*. Apr 9 2021;doi:10.1016/j.ymthe.2021.04.008