# GENETIC CHARACTERIZATION OF AMYOTROPHIC LATERAL SCLEROSIS (ALS): INSIGHTS INTO DISEASE ARCHITECTURE, PHENOTYPIC PRESENTATION, AND THERAPEUTIC IMPLICATIONS

Università di Torino
**PhD in Neuroscience**

Candidate:
Maurizio Grassano

Supervisors
Prof. Andrea Calvo
Prof. Adriano Chiò

# Co-Authorship Statement

**For all manuscripts included in this Dissertation, I am listed as the first author or co-author. For all data chapters, I contributed to conceptualization, methodology, data curation and cleaning, formal analyses, statistical analysis, interpretation of the results, writing of the original draft of the manuscript, and performing manuscript revisions. Aside from Chaptes 6 and 7, which were collaborative effors, I wrote more than 90% of each manuscript.**

Index

# 1. Introduction

**Amyotrophic lateral sclerosis (ALS) is the most common motor neuron disease characterized by the progressive degeneration of motor neurons in the cerebral cortex, brain stem, and spinal cord [1,2,3]. This leads to progressive paralysis, muscle wasting, and ultimately results in death within a few years of symptom onset. While the majority of ALS cases are sporadic (sALS) and occur without a family history, approximately 5-10% have a familial background (fALS) [4]. Although the remaining 90% of cases are apparently sporadic, pathogenic mutations identified in familial ALS were also described in many sporadic cases [5,6]. Additionally, even in sporadic cases, there is evidence to suggest that genetic factors play a role, although their specific contribution is not yet fully understood. In some cases, the absence of a positive family history of ALS could be due to factors such as incomplete penetrance or inheritance involving multiple genes.**

**Indeed, familial aggregation studies, twin studies, and heritability studies based on genome-wide association studies (GWAS) confirmed a substantial genetic component (61% by twin studies and 21% by GWAS) in ALS [7,8,9]. These findings suggest that there are likely shared underlying pathways contributing to the development of ALS in both familial and sporadic cases.**

**Pathogenic variants in the most common ALS-associated genes, superoxide dismutase 1 (SOD1), TAR DNA-binding protein (TARDBP), fused in sarcoma (FUS), and chromosome 9 open reading frame 72 (C9orf72), account for approximately 60% of familial cases and about 10% of sporadic ALS [10]. More recently, whole-exome sequencing and whole-genome sequencing have identified several other ALS-causing genes to the point that we now know the underlying genetic mutation in about three-quarters of familial and one-fifth of sporadic ALS cases in populations of European ancestries [10].**

## 1.1 The genetic basis of Amyotrophic Lateral Sclerosis

### 1.1.1 The four major ALS genes

#### SOD1

In 1993, linkage analysis in neurological disease identified a chromosome 21q locus segregating with ALS in 23 families. Subsequent mutation screening and segregation analysis revealed pathogenic variants in the SOD1 gene as the cause of ALS [11,12]. SOD1 mutations account for approximately 12% of familial and 1-2% of sporadic ALS cases in populations of European ancestries [13]. Most disease-causing SOD1 variants are missense mutations, and they can be inherited in an autosomal dominant or recessive manner [13]. The mutant SOD1 protein has a toxic gain of function and leads to neuronal cell death through various mechanisms [14]. This discovery paved the way for molecular therapy development targeting the misfolded SOD1 protein [15,16,17].

### TARDBP

The discovery of nuclear depletion and cytoplasmic accumulation of TAR DNA-binding protein (TDP43) in motor neurons of ALS patients in 2006 represented a watershed in our understanding of motor neuron degeneration [18]. Indeed, abnormal cytoplasmic TDP43 inclusions are a hallmark of ALS and are also found in other neurological conditions [19]. Subsequent studies identified mutations in the TARDBP gene as a cause of autosomal dominant ALS [20,21]. TARDBP mutations are less common than SOD1 mutations, accounting for 4% of familial and 1% of sporadic cases [10]. TDP43's role as an RNA-binding protein and its involvement in RNA processing suggest its importance in ALS pathogenesis [19,22].

### FUS

A year after the discovery of TARDBP, pathogenic variants were found in the FUS gene, which encodes a similar RNA-binding protein [23,24]. FUS mutations account for 4% of familial and 1% of sporadic ALS cases [10]. Similar to TARDBP, FUS mutations lead to nuclear to cytoplasmic mislocalization and the formation of cytoplasmic inclusions. The disruption of RNA metabolism is implicated in ALS pathogenesis [25].

### C9orf72

Finally, in 2011, a hexanucleotide repeat expansion in the C9orf72 gene was discovered as the most common genetic cause of ALS and FTD [26,27]. The expansion accounts for about 40% of familial and 7% of sporadic ALS cases [28]. Expanded repeats in C9orf72 disrupt RNA metabolism and lead to the production of dipeptide repeats [29,30]. The C9orf72 repeat expansion is associated with various mechanisms of neurodegeneration and has also been reported in other neurological disorders [31,32].

### 1.1.2 Recent mendelian genes discoveries in ALS

#### ANXA11

**The whole-exome investigation of multiple fALS cases revealed the role of ANXA11 mutations in ALS pathogenesis. ANXA11 encodes Annexin A11, a calcium-dependent phospholipid-binding protein that is expressed globally in the body [33]. Mutant ANXA11 tends to form neurotoxic cytoplasmic aggregates in neurons.**

#### C21orf2

**Genome-wide association approaches found an association between C21orf2 and an increased risk of ALS [34]. NEK1 and C21orf2 have been found to interact with each other and are involved in various cellular processes, including microtubule assembly, DNA damage response and repair, and mitochondrial function [35].**

#### CHCHD10

**Variants in CHCHD10 have also been associated with various central and peripheral neurodegenerative disorders, including ALS, frontotemporal dementia, frontotemporal lobar degeneration, parkinsonism, Alzheimer's disease, autosomal dominant mitochondrial myopathy, adult-onset spinal muscular atrophy, and Charcot-Marie-Tooth type 2 [36]. Mutations in CHCHD10 are relatively uncommon as a cause of ALS, but they may be more prevalent among patients diagnosed with frontotemporal dementia [37]. CHCHD10 is a nuclear-encoded mitochondrial protein that is localized in the mitochondrial intermembrane space and plays a crucial role in maintaining mitochondrial dynamics and cellular bioenergetics. Additionally, CHCHD10 also interacts with TDP43, promoting its retention in the nucleus. CHCHD10 mutations disrupt this localization, leading to the accumulation of TDP43 in the cytoplasm and synaptic damage [38].**

#### DNAJC7

**Rare variants in DNAJC7 gene resulting in protein truncation or damaging missense mutations in the were found to be more enriched in ALS cases [33], and functional evidence supports the involvement of the gene in ALS. DNAJC7 belongs to the Hsp40 heat shock protein family: these molecular chaperones are responsible for the folding and clearance of misfolded proteins, including the elimination of misfolded TDP43 [34].**

#### KIF5A

**KIF5A gene has been recently linked to ALS. [35,36] Other than ALS, mutations in KIF5A can lead to other distinct phenotypes depending on their location within the gene, especially hereditary spastic paraplegia (SPG10) or Charcot-Marie-Tooth type 2 hereditary neuropathy (CMT2) [37]. ALS causing variants in KIF5A are located in the cargo-binding domain of the C-terminal region and are predicted to affect the splicing of exon 27. These mutations lead to a toxic gain of function involving abolished auto-inhibition of the kinesin activity, resulting in hyperactive axonal transport [38].**

### MATR3

**MATR3 is a nuclear protein that contains RNA and DNA binding domains and primarily regulates gene expression. It forms a complex with two other ALS-associated RNA-binding proteins, TDP43 and FUS [39]. Several variants in MATR3 have been identified in patients with sporadic ALS since 2014 [40,41]. Surprisingly, unlike TDP43 and FUS, where mutations cause the mutant protein to relocate from the nucleus to the cytoplasm, studies have shown that the subcellular localization of mutant MATR3 is generally unaffected [42]. Overall, the contribution of MATR3 to the development of ALS or ALS with frontotemporal dementia is relatively rare [41].**

### NEK1

**Loss-of-function mutations in the NEK1 gene have been associated with sporadic ALS, indicating its involvement in the disease [43]. Multiple case-control studies have provided additional evidence supporting the association of NEK1 with ALS [44], estimating that NEK1 mutations may account for approximately 2% of all ALS cases. NEK1 interacts with proteins involved in lipid trafficking within the endosome and endoplasmic reticulum, but recent data suggested that NEK1 mutants could exhibit compromised DNA damage repair mechanisms that lead to increased cell death of motoneurons. [45]. Available clinical data suggests that individuals with NEK1 mutations typically exhibit typical ALS symptoms without dementia.**

### SPTLC1

**Recently, de novo mutations in the SPTLC1 gene have been identified as a cause of juvenile ALS [46.47]. Mutations in SPTLC1 disrupt an essential enzyme complex in the sphingolipid synthesis pathway, which is also implicated in autosomal dominant hereditary sensory autonomic neuropathy, type 1A (HSAN1A) [48]. This broadens the connection between sphingolipid metabolism disruption and motor neuron disease: functional assays showed that SPTLC1 mutations alter the enzyme's function, leading to abnormal utilization of certain substrates [46,47]. Intriguingly, nutritional supplementation with SPTLC1 substrates, as serine, could be a viable therapeutic strategy for treating SPTLC1-related ALS.**

#### TBK1

**Mutations in the TBK1 gene are found in approximately 1% of familial ALS patients and 1% of sporadic ALS patients [49]. TBK1 is a multidomain protein involved in various crucial cellular processes associated with ALS, including neuroinflammation, ubiquitin-proteasome systems, and autophagy [50]. Pathogenic variants in TBK1 are typically located in the kinase and coiled-coil domains, suggesting that these mutations may disrupt the normal functioning of downstream regulatory pathways [50].**

### 1.1.3 Genetic risk factors

**It has been suggested that approximately 8.5% of the heritability of ALS can be attributed to common SNPs. However, currently, only 0.2% of the variation in liability can be explained by six common susceptibility loci located in or near the genes UNC13A, SARM1, MOBP and SCFD1, [51]. Additionally, it has been now extensively demonstrated that Ataxin 2 (ATXN2) polyglutamine repeats int the range of 29 to 33 are associated with an increased risk of ALS [52]. Both ATXN2 and UNC13A are also putative prognostic factors associated with a faster disease progression and a shorter survival [53]. However the clinical relevance of genetic testing for (complex) genetic risk factors in ALS has been unclear until recently.**

## 1.2 Objectives of this dissertation

**To unravel the heterogeneous nature of ALS, gain insights into pathogenic mechanisms, and identify therapeutic strategies and targets, a deeper comprehension of ALS genetics is essential. A fundamental step in this direction involves understanding which genetic factors are commonly implicated in ALS and their key clinical and epidemiological characteristics.**

**However, providing a detailed description of the genetic background of ALS in a clinical setting is challenging, particularly because genomic studies are often conducted on large, heterogeneous, multi-centric cohorts. Moreover, multi-cohort studies may not adequately account for known epidemiological biases and environmental factors.**

**In this dissertation, I analyzed a population consisting of ALS cases from the Piemonte and Valle d'Aosta Register for ALS (PARALS), an epidemiological prospective register covering two regions in Italy since 1995. By utilizing genetic data from an epidemiological cohort, we gained advantages**

**through the availability of a large amount and various types of phenotypic information for comparison. ALS is a complex disease influenced by genetic and environmental risk factors, and leveraging epidemiological data alongside whole-genome analysis from a population-based registry helped minimize the impact of environmental effects and enhanced the power to detect the genetic factors associated with the disease. The PARALS registry also provided the opportunity to analyze a large patient cohort in a disease characterized by significant interindividual variability while maintaining relative genetic homogeneity.**

**The objectives of this study were as follows: First, I performed single-variant association testing and gene-wide burden analysis to evaluate the genetic architecture of ALS in our cohort and prioritize candidate genes (Chapter 2). Next, I implemented a comprehensive approach that incorporated gene-specific information and combined multiple in silico predictions to distinguish ALS-causing rare variants. This framework was adapted to whole-genome sequencing data to select variants meeting the criteria for clinical significance, providing an updated evaluation of the prevalence of disease-causing mutations in ALS (Chapter 3). Once the efficacy of this variant classification framework was demonstrated in identifying ALS-associated genes and variants, a similar approach was adopted to explore whether the genetic pleiotropy of neurodegenerative diseases could also have clinical implications (Chapter 4,5). Rare variants in genes associated with neurodegenerative diseases were analyzed, evaluating their contribution to ALS risk and phenotypic presentation. To validate the results, the analysis was replicated using data from a large-scale genomic ALS consortium, project MinE (https://www.projectmine.com/).**

**Subsequently, the impact of genes and genotypes on ALS phenotype was evaluated (Chapter 6 and Chapter 7). By demonstrating that the most common genetic risk factor for ALS also had a specific effect on disease phenotype, the work presented in these sections laid the groundwork for identifying other genetic factors that could jointly contribute to disease risk and phenotypic presentation. Finally, this approach was applied to demonstrate that a common genetic factor related to other neurodegenerative disorders (the HTT gene) is also associated with the pathogenesis and progression of motor neuron disorders (Chapter 8).**

## 1.3 References


1. **Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., ... & Amyotrophic Lateral Sclerosis. The Lancet 377(9769), 942-955 (2011).**

2. **Logroscino, G., Traynor, B.J., Hardiman, O., Chiò, A., Mitchell, D., Swingler, R.J., ... & Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. Neuroepidemiology 41(2), 118-130 (2013).**

3. **van Es, M.A., Hardiman, O., Chio, A., Al-Chalabi, A., Pasterkamp, R.J., & Veldink, J.H. Amyotrophic lateral sclerosis. The Lancet 390(10107), 2084-2098 (2017).**

4. **Brown, R.H., & Al-Chalabi, A. Amyotrophic lateral sclerosis. New England Journal of Medicine 377(2), 162-172 (2017).**

5. **Renton, A.E., Chio, A., & Traynor, B.J. State of play in amyotrophic lateral sclerosis genetics. Nature Neuroscience 17(1), 17-23 (2014).**

6. **Al-Chalabi, A., & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. Nature Reviews Neurology 9(11), 617-628 (2013).**

7. **Keller, M.F. et al. Genome-wide analysis of the heritability of amyotrophic lateral sclerosis. JAMA Neurology 71, 1123-1134 (2014).**

8. **Al-Chalabi, A. et al. An estimate of amyotrophic lateral sclerosis heritability using twin data. Journal of Neurology, Neurosurgery, and Psychiatry 81, 1324-1326 (2010).**

9. **Fang, F. et al. Familial aggregation of amyotrophic lateral sclerosis. Annals of Neurology 66, 94-99 (2009).**

10. **Chia, R., Chio, A., & Traynor, B.J. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. The Lancet Neurology 17, 94-102 (2018).**

11. **Siddique, T. et al. Linkage of a gene causing familial amyotrophic lateral sclerosis to chromosome 21 and evidence of genetic-locus heterogeneity. New England Journal of Medicine 324, 1381-1384 (1991).**

12. **Rosen, D.R. et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362, 59-62 (1993).**

13. **McCann, E.P. et al. Evidence for polygenic and oligogenic basis of Australian sporadic amyotrophic lateral sclerosis. Journal of Medical Genetics 58, 87-95 (2021).**

14. Hayashi, Y., Homma, K., & Ichijo, H. SOD1 in neurotoxicity and its controversial roles in SOD1 mutation-negative ALS. Advances in Biological Regulation 60, 95-104 (2016).

15. Raoul, C. et al. Lentiviral-mediated silencing of SOD1 through RNA interference retards disease onset and progression in a mouse model of ALS. Nat. Med. 11, 423–428 (2005).

16. Ralph, G. S. et al. Silencing mutant SOD1 using RNAi protects against neurodegeneration and extends survival in an ALS model. Nat. Med. 11, 429–433 (2005).

17. Smith, R. A. et al. Antisense oligonucleotide therapy for neurodegenerative disease. J. Clin. Invest. 116, 2290–2296 (2006).

18. Neumann, M. et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314, 130–133 (2006).

19. Besser, L. M., Teylan, M. A. & Nelson, P. T. Limbic predominant age-related TDP-43 encephalopathy (LATE): clinical and neuropathological associations. J. Neuropathol. Exp. Neurol. 79, 305–313 (2020).

20. Kabashi, E. et al. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat. Genet. 40, 572–574 (2008).

21. Sreedharan, J. et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science 319, 1668–1672 (2008).

22. Suk, T. R. & Rousseaux, M. W. C. The role of TDP-43 mislocalization in amyotrophic lateral sclerosis. Mol. Neurodegener. 15, 45 (2020).

23. Kwiatkowski, T. J. Jr et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science 323, 1205–1208 (2009).

24. Vance, C. et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 323, 1208–1211 (2009).

25. Bosco, D. A. et al. Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. Hum. Mol. Genet. 19, 4160–4175 (2010).

26. Renton, A. E. et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72, 257–268 (2011).

27. DeJesus-Hernandez, M. et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245–256 (2011).

28. Majounie, E. et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol. 11, 323–330 (2012).

29. Ash, P. E. et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 77, 639–646 (2013).

30. Balendra, R. & Isaacs, A. M. C9orf72-mediated ALS and FTD: multiple pathways to disease. Nat. Rev. Neurol. 14, 544–558 (2018).

31. Abramzon, Y. A., Fratta, P., Traynor, B. J. & Chia, R. The overlapping genetics of amyotrophic lateral sclerosis and frontotemporal dementia. Front. Neurosci. 14, 42 (2020).

32. Majounie, E. et al. Repeat expansion in C9ORF72 in Alzheimer's disease. N. Engl. J. Med. 366, 283–284 (2012).

33. Farhan, S. M. K. et al. Exome sequencing in amyotrophic lateral sclerosis implicates a novel gene, DNAJC7, encoding a heat-shock protein. Nat. Neurosci. 22, 1966–1974 (2019).

34. Jinwal, U. K. et al. Cdc37/Hsp90 protein complex disruption triggers an autophagic clearance cascade for TDP-43 protein. J. Biol. Chem. 287, 24814–24820 (2012).

35. Brenner D, Yilmaz R, Muller K, € et al. Hot-spot KIF5A mutations cause familial ALS. Brain 2018; 141:688–697.

36. Nicolas A, Kenna KP, Renton AE, et al. Genome-wide analyses identify KIF5A as a novel ALS gene. Neuron 97, 1268.e6–1283.e6 (2018).

37. Liu YT, Laura´ M, Hersheson J, et al. Extended phenotypic spectrum of KIF5A mutations: from spastic paraplegia to axonal neuropathy. Neurology 83, 612–619 (2014).

38. Baron, D. M. et al. ALS-associated KIF5A mutations abolish autoinhibition resulting in a toxic gain of function. Cell Rep. 39, 110598 (2022).

39. Coelho MB, Attig J, Bellora N, et al. Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. EMBO J. 34, 653–668 (2015).

40. Johnson JO, Pioro EP, Boehringer A, et al. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. Nat Neurosci. 17, 664–666 (2014).

41. Leblond CS, Gan-Or Z, Spiegelman D, et al. Replication study of MATR3 in familial and sporadic amyotrophic lateral sclerosis. Neurobiol Aging. 37, 209.e17–209.e21 (2016).

42. Gallego-Iradi MC, Clare AM, Brown HH, Janus C, Lewis J, Borchelt DR. Subcellular localization of Matrin 3 containing mutations associated with ALS and distal myopathy. PLoS One. 10, e0142144 (2015).

43. Brenner D, Müller K, Wieland T, et al. NEK1 mutations in familial amyotrophic lateral sclerosis. Brain J Neurol. 139, e28 (2016).

44. Kenna KP, van Doormaal PTC, Dekker AM, et al. NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. Nat Genet. 48, 1037–1042 (2016).

45. Higelin J, et al. NEK1 loss-of-function mutation induces DNA damage accumulation in ALS patient-derived motoneurons. Stem Cell Res. 30, 150

46. Mohassel, P. et al. Childhood amyotrophic lateral sclerosis caused by excess sphingolipid synthesis. Nat. Med. 27, 1197–1204 (2021).

47. Johnson, J. O. et al. Association of variants in the SPTLC1 gene with juvenile amyotrophic lateral sclerosis. JAMA Neurol. 78, 1236–1248 (2021).

48. Penno, A. et al. Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids. J. Biol. Chem. 285, 11178–11187 (2010).

49. Freischmidt A, Wieland T, Richter B, et al. Haploinsufficiency of TBK1 causes familial ALS and frontotemporal dementia. Nat Neurosci. 18, 631–636 (2015).

50. Oakes JA, Davies MC, Collins MO. TBK1: a new player in ALS linking autophagy and neuroinflammation. Mol Brain. 10, 5 (2017).

51. van Rheenen W, et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet. 2016;48(9):1043–1048. doi: 10.1038/ng.3622

52. Elden AC, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature. 2010;466(7310):1069–1075. doi: 10.1038/nature09320

53. Chio, A., Moglia, C., Canosa, A., Manera, U., Grassano, M., Vasta, R., ... & Calvo, A. (2023). Association of Copresence of Pathogenic Variants Related to Amyotrophic Lateral Sclerosis and Prognosis. Neurology.

# 2. Mutational analysis of known ALS genes in a large Italian population-based cohort

As published in ***Grassano, M., Calvo, A., Moglia, C., Brunetti, M., Barberis, M., Sbaiz, L., ... & Chiò, A. (2021). Mutational analysis of known ALS genes in an Italian population-based cohort. Neurology, 96(4), e600-e609.***

**AUTHORS**

**Maurizio Grassano1, Andrea Calvo1, Cristina Moglia1, Maura Brunetti2, Marco Barberis2, Luca Sbaiz2, Antonio Canosa1, Umberto Manera1; International ALS Genomics Consortium; The American Genomic Center; Luigi Ferrucci3, Rosario Vasta1, Lucia Corrado 4, Sandra D'Alfonso 4, Letizia Mazzini 5, Sonja W. Scholz 6,7, Clifton Dalgard 6,8,9, Jinhui Ding 10, Raphael J. Gibbs 6,10, Bryan J. Traynor 6,7, 11, Adriano Chiò 1,12**

**Author Affiliations**

**[1] ''Rita Levi Montalcini'' Department of Neuroscience, University of Turin, Turin, Italy**

**[2] Laboratory of Genetics, Department of Clinical Pathology, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy**

**[3] Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD, 21224, USA**

**[4] Department of Health Sciences Interdisciplinary Research Center of Autoimmune Diseases, D "Amedeo Avogadro" University of Eastern Piedmont, Novara, Italy;**

**[5] ALS Center, Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy**

**[6] Neurodegenerative Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA**

[7] Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA

[8] Department of Anatomy, Physiology & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA

[9] The American Genome Center, Collaborative Health Initiative Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA

[10] Biocomputational Group, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA

[11] Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA

[12] Institute of Cognitive Sciences and Technologies, National Council of Research, Rome, Italy

## 2.1 Abstract

### Importance

**Amyotrophic Lateral Sclerosis is a disease with relevant genetic component. However, the genetic causes of ALS remain undefined in most of the patients.**

### Objective

**To assess the burden of rare genetic variants and estimate the contribution of known ALS genes in an Italian population cohort.**

### Design

**We performed a cohort-wide genome sequencing in a population based cohort. A panel of 45 ALS genes was analyzed to identify potential disease-causing genetic variants and to evaluate the gene-wide burden of rare variants among our population.**

### Settings

**This cohort-based genome-sequencing was conducted on patients included in the Piedmont and Valle d'Aosta Registry for ALS (PARALS).**

### Participants

**A cohort of 959 ALS patients was compared with 677 healthy controls.**

**Results**

**Gene-wide association tests demonstrated a strong association for SOD1, whose rare variants are the second most common cause of disease after C9orf72 expansion. A lower signal was observed for TARDBP, proving that its effect on our cohort is driven by few known causal variants. We detected rare variants in other known ALS genes who did not surpassed statistical significance in gene-wise tests, thus highlighting that their contribution to disease risk in our cohort is limited.**

**Conclusions**

**We identified potential disease-causing variants in 11.9% of our patients. We identified the genes most frequently involved in our cohort and confirmed the contribution of rare variants in disease risk. Our results provide further insight into the pathological mechanism of the disease and demonstrated the importance of genome wide sequencing as a diagnostic utility.**

## 2.2 Introduction

**Amyotrophic lateral sclerosis (ALS, OMIM #105400) is a devastating neurological disorder characterized by progressive degeneration of upper and lower motor neurons resulting in progressive paralysis and ultimately death from respiratory failure. Approximately 10% of ALS cases have a family history, whereas the remainder occurs randomly.1 Since the discovery of pathogenic mutations in SOD1 in 19932, researchers have made great strides in delineating the genetic basis underlying ALS. Today more than forty genes have been reported in association with the disease, and the genetic etiology of approximately two-thirds of familial cases is known.1**

**With few exceptions, mutational screening has been performed on samples collected from patients attending ALS clinics. Such clinic-based cohorts represent a self-selected subset of the ALS population, making it challenging to determine true mutational rates. In contrast, the Piedmont and Valle d'Aosta Registry for ALS (PARALS) was established in 1995 to study the epidemiological characteristics of the disease in Northern Italy.4 Using this registry offers an opportunity to study the frequency of mutations in a population-based cohort that has been in operation for over twenty years and has near-complete case ascertainment.**

**Here, we analyze the frequency and burden of mutations in known ALS genes within a population-based cohort of nearly 1,000 Italian cases that underwent genome sequencing. Our efforts provide insight into the precise frequency of mutations in ALS genes in Northern Italy.**

## 2.3 Methods

### Study participants

**Our cohort consisted of 959 patients included in PARALS. 5 Supplementary Table T1 lists the clinical and demographic details of this cohort.**

### Mutation screening

**Genome sequencing of the cohort was performed using paired-end sequencing on an Illumina HiSeq X10 platform after library preparation using the TruSeq according to manufacturer's protocol. We extracted variant information from the genome sequence data for genes with robust evidence supporting their association with ALS (see Supplementary Table T2). All subjects were screened for the pathological repeat expansion in C9orf72.3**

### Variant filtering

**Variants were annotated based on allele frequencies in Non-Finnish European derived from the gnomAD (https://gnomad.broadinstitute.org/) and TOPMed (https://www.nhlbiwgs.org/) databases. As standard in gene-burden analysis, we set the minor allele frequency (MAF) threshold at 5% and 1%.6 A lower threshold (MAF < 0.0033%) was also selected based on the epidemiology of the studied condition.5**

### Statistical Analysis

**To identify association signals, we applied several single-variant and gene-burden tests to our dataset. Supplementary Table T3 details these statistical methods. As none of these tests offer a clear advantage in terms of performance, we evaluated the concordance across the different tests to assess the importance of each gene. Supplementary Material provides a detailed description of the bioinformatics and statistical analysis methods used in this paper.**

## 2.4 Results

### Frequency of mutations in ALS genes

**We analyzed the sequences of forty-four ALS-related genes in a population-based cohort of 959 Italian ALS cases and 657 unaffected controls. Of these, 71 patients carried the C9orf72 pathogenic repeat expansions, representing 7.4% of the entire cohort and 42.1% of the familial cases. An additional 34 cases carried causative mutations, representing a further 3.5% of the overall cohort and 13.1% of the familial cases.**

**Figure 1 shows the mutational frequency of the genes tested in our cohort, and Supplementary Table T4 describes the clinical features of the mutated cases.**

### Single-variant analysis

**After adjustment for multiple testing, only a single variant was associated with the disease. This was a c.2009T>A change resulting in a valine to glutamic acid change at residue 670 of the neurofilament heavy (NEFH) gene (p-value=3.7x10-5, FDR-adjusted p-value=0.047). This variant (rs190692435) decreased the risk of ALS (OR 0.28). Variants that are known to cause ALS did not surpass the significance threshold in our cohort, likely reflecting the fact that our sample size was insufficiently powered to detect them (Table 1).**

### Gene-burden analysis results

**NEFH was significantly associated with the risk of ALS under the SKAT-O model at a minor allele threshold of 5% (p-value=7.7x10-5, FDR p-value=0.0035, see Figure 2). The gene-level test remained significant after removing the top variants in NEFH from the analysis.**

**The only other significant gene was SOD1 (p-value under Madson-Browning analysis = 8.4x10-4, FDR p-value=0.036, see Figure 1). Based on Sanger sequence data, we had previously reported that variant to be the most prevalent in our cohort after the C9orf72 expansion.6**

**Repeating the gene-burden analysis at a more stringent minor allele threshold of 1% found that SOD1 remained significantly associated with the risk of ALS (p-value under the KBAC model = 7.0x10-4, FDR p-value=0.032). TARDBP was the second most enriched gene, though it did not reach statistical significance (p-value under the KBAC model = 1.9x10-3, FDR p-value=0.084).**

**Investigating ultra-rare variants (MAF < 0.0033%), we observed the persistence of the association of SOD1 (p-value under the SKATt-O model = 1.9x10-4, FDR p-value=0.025). In contrast, we obtained a less significant p-value for TARDBP reflecting the exclusion of a common founder mutation at this threshold (see Supplementary Figure S1 for details).**

## 2.5 Discussion

**We sequenced the genomes of 959 patients diagnosed with ALS and 677 healthy control obtained from an Italian population-based registry and estimated the frequency of mutations of major ALS genes in this prospectively ascertained, population-based epidemiologic series.13 This constitutes an opportunity to understand the contribution of variants and genes to the genetic and phenotypic spectrum of ALS. We found that a significant portion of our cohort (n=105 cases, 11.0%) carried mutations that are known to be pathogenic.**

**Our analysis demonstrates the utility of genome sequencing to assess the role of multiple genes in a disease rapidly. Undertaking this scope of analysis with traditional Sanger sequencing would be costly and time-consuming. Even if we applied restrictive criteria in the definition of causal genetic variants, we demonstrated that genome sequencing yields a broader insight into the genetic architecture of ALS than targeted sequencing or even exome sequencing. Whole-genome sequencing provides a complete catalog of rare variants present in cases, allowing us to establish the role of noncoding and intergenic genetic variation.**

**In gene-based rare-variant association tests, SOD1 was the only gene to reach a significance threshold for all of the frequency thresholds tested. Interestingly, we found low-frequency variants in NEFH that lowered the risk of disease. There are few previous reports in the literature concerning protective variants in ALS. However, neurofilaments are currently regarded as a promising biomarker in ALS, possibly reflecting an involvement in neuronal degeneration. 14**

**The relatively small size of our cohort limited the power to detect a significant association. Furthermore, we did not consider other types of genetic mutations, such as intronic variants and structural variants that are known to be relevant to ALS. Thus, the mutational frequencies presented here are considered to be floor estimates. Nevertheless, our data do allow us to conclude that coding mutations in these genes are not a common cause of sporadic disease.**

**In summary, we performed a cohort-wide genome sequencing that can capture genetic variation across the entire disease spectrum in our population. Our study provides evidence and an effective way to evaluate the role of low-frequency and rare variants underpinning the complex nature of ALS.**

## 2.6 References


**1. Renton AE, Chiò A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. Nat Neurosci. 2014;17(1):17-23.**

**2. Laaksovirta H, Peuralinna T, Schymick JC, et al. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol. 2010;9(10):978-985.**

**3. Borghero G, Pugliatti M, Marrosu F, et al. Genetic architecture of ALS in Sardinia. Neurobiol Aging. 2014;35(12):2882.e7-2882.e12.**

**4. Chiò A, Mora G, Calvo A, et al. Epidemiology of ALS in Italy: a 10-year prospective population-based study. Neurology. 2009;72(8):725-731.**

**5. Chiò A, Mora G, Moglia C, et al. Secular Trends of Amyotrophic Lateral Sclerosis: The Piemonte and Valle d'Aosta Register. JAMA Neurol. 2017;74(9):1097-1104.**

**6. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011;89(1):82-93.**

**7. Lee S, Emond MJ, Bamshad MJ, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet. 2012;91(2):224-237.**

**8. Liu DJ, Leal SM. A novel adaptive method for the analysis of next-generation sequencing data to detect complex trait associations with rare variants due to gene main effects and interactions. PLoS Genet. 2010;6(10):e1001156.**

**9. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet. 2008;83(3):311-321.**

**10. Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genet Epidemiol. 2010;34(2):188-193.**

**11. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009;5(2):e1000384.**

**12. Lin D-Y, Tang Z-Z. A General Framework for Detecting Disease Associations with Rare Variants in Sequencing Studies. Am J Hum Genet. 2011;89(3).**

**13. Chiò A, et al. Extensive genetics of ALS: a population-based study in Italy. Neurology. 2012. 79(2): 1983-1989.**

**14. Steinacker P, Feneberg E, Weishaupt J, et al. Neurofilaments in the diagnosis of motoneuron diseases: a prospective study on 455 patients. Journal of Neurology, Neurosurgery & Psychiatry 2016;87:12-20.**

## 2.7 Tables And Figures

**Figure 1. Charts showing the frequency of major ALS genes in our cohort and their distribution among ALS cases with a family history of the disease.**

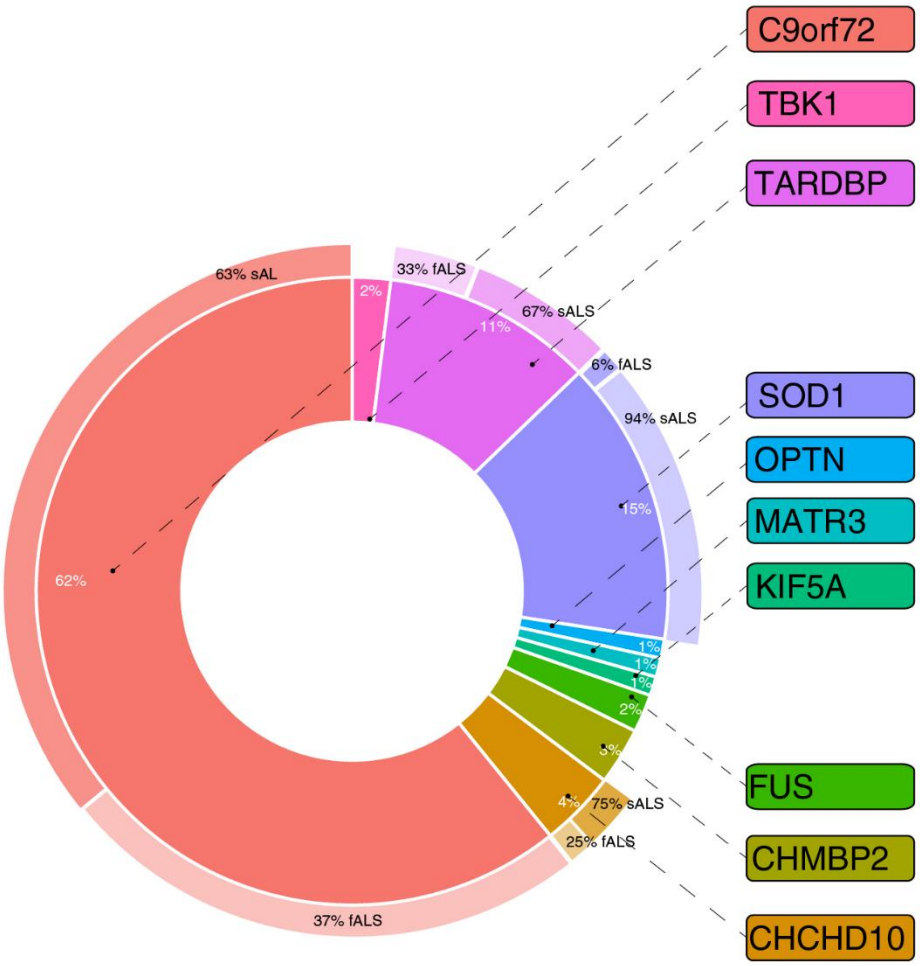



**Figure 2. Result of rare-variant association testing with MAF < 5% (A), < 1% (B) and < 0.0033% (C). Solid line: p-value 0.05. Dashed line: Bonferroni-adjusted p-value 0.001.**

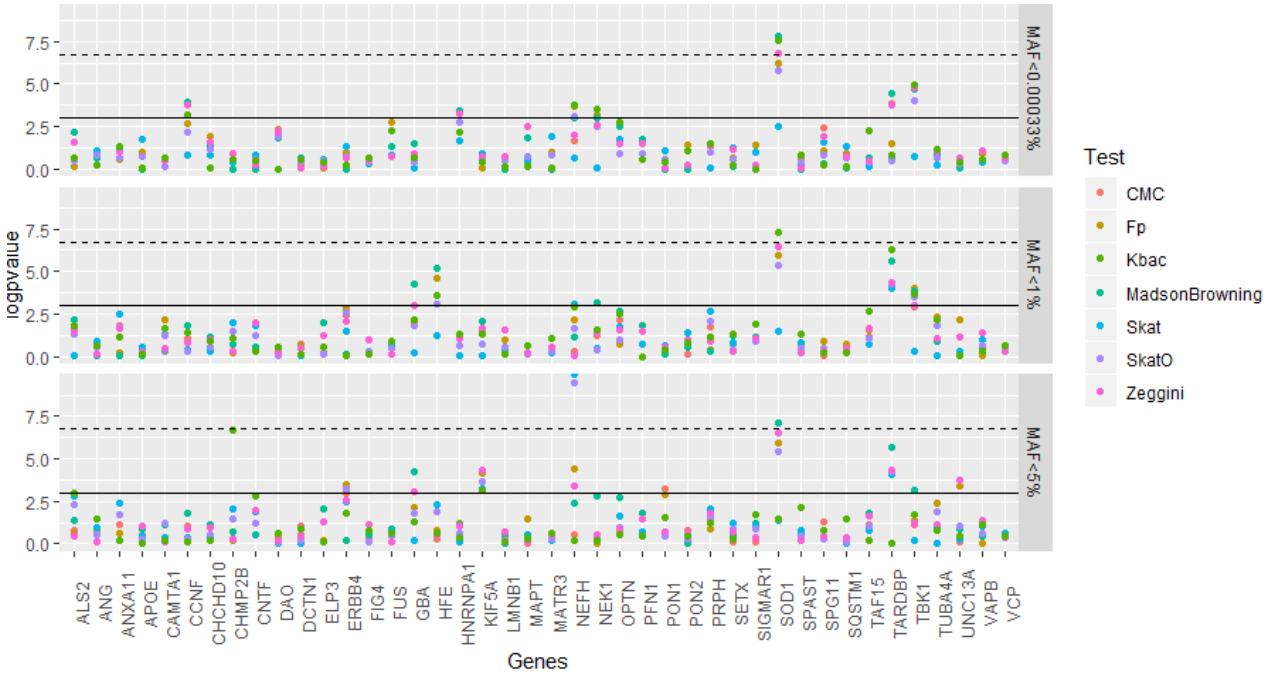



**Table 1.** **Summary of the results from single-variant test.**

**Nucleotide, nucleotide change. AA, amino acid change. AF, allele frequency in our cohort. OR, odds ratio from Firth regression, p-valueS, p-value from Single Score test. p-valueF, p-value from Firth regression (see also Supplementary Methods). Values in bold withstands correction for Bonferroni-adjusted significance threshold or False Discovery Ratio (FDR).**

| Gene | Exon | Nucleotide | AA | AF | OR | p-valueS | FDR | p-valueF |
|---|---|---|---|---|---|---|---|---|
| NEFH | 4 | T2009A | V670E | 0.0138 | 0.28 | 0.000037 | 0.0489 | 0.000106 |
| NEFH | 4 | C2015A | A672E | 0.0156 | 0.31 | 0.000054 | 0.0713 | 0.000127 |
| NEFH | 4 | 1952_1975del | E658_K665 del | 0.0171 | 0.37 | 0.000393 | 0.5152 | 0.000588 |
| TARDBP | 6 | G1144A | A382T | 0.0025 | 5.54 | 0.017639 | 1 | 0.057657 |
| SOD1 | 5 | G435C | L145F | 0.0012 | 5.52 | 0.093306 | 1 | 0.243403 |

# 3. A systematic evaluation of genetic mutations in ALS: a population-based study

**As published in *Grassano, M., Calvo, A., Moglia, C., Sbaiz, L., Brunetti, M., Barberis, M., ... & Chiò, A. (2022). Systematic evaluation of genetic mutations in ALS: a population-based study. Journal of Neurology, Neurosurgery & Psychiatry, 93(11), 1190-1193.***

**AUTHORS**

**Maurizio Grassano MD, Andrea Calvo MD, Cristina Moglia MD, PhD, PhD, Luca Sbaiz BSc, Maura Brunetti BSc, Marco Barberis BSc, Federico Casale PhD, Umberto Manera MD, Rosario Vasta MD, Antonio Canosa MD, PhD, Sandra D'Alfonso PhD, Lucia Corrado PhD, Letizia Mazzini MD, Clifton Dalgard PhD, The American Genome Center, Ramita Karra PhD , Ruth Chia PhD, Bryan J. Traynor° MD, PhD, Adriano Chiò° MD, PhD, FAAN**

°These authors contributed equally to this work

**Author Affiliations**

**ALS Center, ''Rita Levi Montalcini'' Department of Neuroscience, University of Turin, Turin, Italy (Maurizio Grassano, Andrea Calvo, Cristina Moglia, Federico Casale, Umberto Manera, Rosario Vasta, Antonio Canosa, Adriano Chiò);**

**S.C. Neurologia 1U, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy (Andrea Calvo, Cristina Moglia, Umberto Manera, Antonio Canosa, Adriano Chiò);**

**Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA (Maurizio Grassano, Ruth Chia, Ramita Karra, Bryan J. Traynor);**

**Laboratory of Genetics, Department of Clinical Pathology, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy (Maura Brunetti, Marco Barberis, Luca Sbaiz);**

**Department of Health Sciences, Interdisciplinary Research Center of Autoimmune Diseases, "Amedeo Avogadro" University of Eastern Piedmont, Novara, Italy (Lucia Corrado, Sandra D'Alfonso);**

**ALS Center, Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy (Letizia Mazzini);**

**Department of Anatomy, Physiology & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA (Clifton Dalgard);**

**The American Genome Center, Collaborative Health Initiative Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA (Clifton Dalgard);**

**Reta Lila Weston Institute, UCL Queen Square Institute of Neurology, University College London, London WC1N 1PJ, UK (Bryan J. Traynor);**

**Department of Neurology, Johns Hopkins University Medical Center, Baltimore, MD 21287, USA (Bryan J. Traynor);**

**National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD, USA (Bryan J. Traynor);**

**ASO Rapid Development Laboratory, Therapeutics Development Branch, National Center for Advancing Translational Sciences, NIH, Rockville, MD, USA (Bryan J. Traynor);**

**Institute of Cognitive Sciences and Technologies, National Council of Research, Rome, Italy (Adriano Chiò).**

## 3.1 Abstract

### Importance

**A genetic diagnosis in ALS can inform genetic counselling, prognosis and, in the light of incoming gene-targeted therapy, management and treatment. However, conventional genetic testing strategies for ALS are often costly and time-consuming.**

### Objective

**To evaluate the analytical and clinical validity of whole-genome sequencing (WGS) as a standard diagnostic genetic test for ALS and to provide a comprehensive analysis of the frequency and spectrum of known genetic factors in ALS.**

### Design, Setting, and Participants

**In this population-based cohort study, 1043 ALS patients recruited in the Piemonte and Valle d'Aosta Register for ALS (PARALS) from 2007 to 2016 were included. A control cohort of 755 healthy individuals was used for population-specific mutation filtering.**

### Exposures

**All samples were screened by WGS for variants in 42 ALS-related genes. Variants were classified by likelihood of pathogenicity according to a pipeline that accounted for allele frequency, absence in the controls and pathogenicity predictions. Repeated expansion lengths of the *C9orf72* and *ATXN2* genes were also estimated from WGS data.**

### Main Outcomes and Measures

**The primary study outcome was the diagnostic yield of WGS in ALS cases in a population-based setting.**

### Results

**A total of 279 ALS cases (26.9%) received a genetic diagnosis. Disease-related mutations were observed in 75.2% of patients with a family history of ALS and 21.5% of sporadic ALS cases. The mutation rate among early-onset ALS patients was 43.9%, compared to 19.7% of late-onset patients. An additional 14.6% of the cohort had a genetic factor that increase disease risk and worsen prognosis. 1.3% of ALS cases carried > 1 genetic variant.**

### Conclusions and Relevance

**This study suggests that WGS has high validity for genetic diagnosis in the clinical setting. Additionally, our results provide a detailed picture of the genetic basis of ALS in the general population. WGS should be considered as a standard genetic test for ALS since it is time and cost-efficient and can be adapted to include different genes and types of mutations as our understanding of the genetic mechanism underlying ALS evolves.**

## 3.2 Introduction

**Considerable progress has been made in unravelling the complex molecular mechanisms underlying ALS[1]. Nevertheless, it is not easy to gauge the impact of these genetic advances across the ALS space. Estimates of mutation carrier rate among ALS patients seen in specialist centers may be skewed due to referral bias[2,3]. This knowledge gap hinders efforts to develop therapies and to counsel patients effectively. For this reason, we systematically analyzed a population-based cohort of ALS patients using whole-genome sequencing. Our comprehensive approach provides a snapshot of what we currently know about the genetic architecture of ALS. In the process. The current study builds on our previous work, where we showed that whole-genome sequencing is a reliable tool to assess gene burden effects[4].**

## 3.3 Methods

### Patients and controls cohort

**We analyzed ALS patients who had been enrolled in Piemonte and Valle d'Aosta Register for ALS (PARALS). All samples have been previously analyzed for mutations in the *SOD1, TARDBP*, *FUS* and *C9orf72* genes (eMaterials). An additional 755 healthy individuals underwent whole genome sequencing and were used as control data for mutation filtering.**

### Sequencing and bioinformatics analysis

**Whole-genome sequencing methodology and quality control filters have been detailed elsewhere[4–6] and are described in eMaterials. We extracted variant information for 46 genes previously implicated in ALS pathogenesis (Figure 1 and eMaterials). We estimated the repeat lengths of the *C9orf72* and *ATXN2* repeat expansions using ExpansionHunter – Targeted (version 0.3)[7]. For variants classification, we employed a framework based on the 2015 ACMG-AMP guidelines[8]. Loss-of-function and previously reported ALS variants were considered pathogenic unless present in the control cohort. The remaining rare variants (defined as minor allele frequency less than 0.0001 in the non-Finnish European population) were then classified based on computational prediction and expert review (eMaterials) and reported if deemed to be potentially pathogenic (Figure 1).**

### Statistical analysis

**A two-tailed Fisher exact test was used to evaluate the genetic association between *ATXN2* CAG-repeat sizes and ALS. The burden of multiple variant carriers was assessed by a binomial test[9]. The analyses were performed in R (version 3.6.0).**

## 3.4 Results

**We analyzed a population-based cohort of 1,043 patients who had been diagnosed with ALS and enrolled in Piemonte and Valle d'Aosta Register for ALS (PARALS). The 1,043 patients represented 71.3% of incident cases during the 2007–2016 period. We found 96 mutations known to cause ALS among 203 cases, representing 19.5% of our cohort. An additional 76 patients (7.4%) carried potentially pathogenic variants based on our classification pipeline, bringing the total number of ALS cases with disease-causing variants to approximately a quarter of ALS cases (26.9%, n=279, Figure 2A and eTable 1). The distribution of variants across ALS genes is reported in Figure 2B. The most common mutation type was C9orf72 (present in 7.7% of ALS cases), followed by SOD1 (2.0%), NEK1 (1.8%), TARDBP (1.4%), and KIF5A (0.8%). All mutations were detected by both traditional Sanger sequencing and whole-genome sequencing screening, confirming the ability of whole-genome sequencing to identify relevant mutations correctly.**

### Diagnostic yield in familial and early-onset ALS cases

**Subdividing our cohort, we detected disease-causing mutations in 88 (75.2%) familial ALS cases. We also observed a high mutation rate among early-onset ALS patients (age at onset<50 years, n=43, 43.9%) compared to late-onset patients (age at onset>75 years, n=41, 19.7%). Nearly half of the mutations observed in the elderly cohort were due to the *C9orf72* repeat expansion, illustrating the reduced penetrance observed with this mutation[10]. Furthermore, 21.5% of apparently sporadic ALS patients carried a disease-causing variant (Figure 2A).**

### Risk factor and prognostic variants

**Intermediate-length *ATXN2* CAG expansion (30-33 repeats) was the only high-risk genetic factor (defined as odds ratio≥2.0) identified in our cohort (odds ratio=2.84, 95% confidence interval=1.45-5.57, p=0.0023). The ATXN2 expansions were present in 41 (3.9%) patients. The prognostic *UNC13A* rs12608932 CC genotype was observed in 10.7% (n=112) of our ALS cases.**

### Oligogenic cases

**Our cohort's rate of oligogenic ALS cases was 1.3% (n=13) (eTable 2A). However, the proportion of oligogenic patients was not higher than expected (binomial p=0.98) based on the frequency of monogenic (25.6%) and non-mutated cases (73.1%) (eTable 2B).**

## 3.5 Discussion

**We detected a high rate of patients carrying pathogenic mutations (26.9%) in our population-based ALS cohort. This represents the largest percentage explained by genetic causes for any cohort reported to date[2,11,12]. As we sequenced an ALS cohort from a population-based registry, our data also represent a detailed audit of what is currently known about the genetic architecture of this fatal neurodegenerative condition in the general population. Our findings reflect both our increasing knowledge of the genetic architecture of ALS and the power of whole-genome sequencing to identify these variants. For example, the C9orf72 repeat expansion accounted for nearly one-third of all cases where the mutation was known.**

**Our results support the use of whole-genome sequencing in ALS patients at the time of their diagnosis, irrespective of family history, age at onset or clinical phenotype[13]. Whole-genome sequencing is a flexible tool that does not rely on a pre-determined set of genes and a variant prioritization process but can be adapted to different types of mutations as the collection of ALS-related genes evolves. The continuous improvement in variant calling and interpretation, along with our improving understanding of ALS genetics, will facilitate those endeavors. Ideally, periodic reevaluations of the whole-genome sequence data will become a feed-forward loop that enhances our knowledge with each iteration.**

**Our data do not support the oligogenicity theory of ALS, as it occurs very infrequently within the Italian population. The future discovery of more genes involved in ALS pathogenesis may reveal true oligogenicity. However, the existing data do not currently uphold a role for this disease mechanism.**

**Our primary reason for championing whole-genome sequencing as a routine test lies in the therapeutic implications for ALS patients. For example, we detected *ATXN2* intermediate trinucleotide expansion in 41 ALS subjects (3.9%) and confirmed that it increased ALS risk. This result is clinically relevant since antisense oligonucleotide therapy targeting *ATXN2* is in human trials at present[14]. Considering *ATXN2, C9orf72*, *SOD1 and FUS*, all genes currently under study for the treatment with ASOs[15], at least 142 patients (13.6% of the cohort) could be candidates for gene-based therapy (Figure 3D). Whole-genome sequencing allows those variants to be screened in a low-cost, rapid manner compared to traditional methodologies.**

**Our study is not without limitations. Above all, the lack of definite criteria for variant interpretation increases the risk of misclassifying ALS patients as wild-type or mutated. While our internal framework was reasonably robust, we do not maintain that it is definitive. Over time, the availability of updated mutation databases based on ever-larger cohort sizes and improved interpretation algorithms will enhance the analysis of these variants. Ambitious research programs are collecting whole-genome sequencing data from millions of human genomes to pair with phenotypical and long-term clinical data.**

**To summarize, our population-based evaluation sheds light on the complex genetics of ALS and provides a valuable benchmark of where the field currently stands. We show that whole-genome sequencing increases diagnostic yield and facilitates the assessment of the pathogenic role of involved genes. This information will be crucial for clinical care as precision therapeutics emerge as effective treatments. In anticipation of that, we consider whole-genome sequencing as the first-tier genetic test on all ALS patients.**

## 3.6 References


**1. Chia R, Chiò A, Traynor BJ. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *The Lancet Neurology*. 2018;17(1):94-102. doi:10.1016/S1474-4422(17)30401-5**

**2. Shepheard SR, Parker MD, Cooper-Knock J, et al. Value of systematic genetic screening of patients with amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. Published online February 14, 2021:jnnp-2020-325014. doi:10.1136/jnnp-2020-325014**

**3. Logroscino G, Marin B, Piccininni M, et al. Referral bias in ALS epidemiological studies. *PLoS One*. 2018;13(4):e0195821. doi:10.1371/journal.pone.0195821**

**4. Grassano M, Calvo A, Moglia C, et al. Mutational analysis of known ALS genes in an Italian population-based cohort. *Neurology*. Published online November 2020. doi:10.1212/WNL.0000000000011209**

**5. 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*. 2021;53(3):294-303. doi:10.1038/s41588-021-00785-3**

**6. Dewan R, Chia R, Ding J, et al. Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. *Neuron*. 2021;109(3):448-460.e4. doi:10.1016/j.neuron.2020.11.005**

**7. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35(22):4754-4756. doi:10.1093/bioinformatics/btz431**

**8. Richards S, Aziz N, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-423. doi:10.1038/gim.2015.30**

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

**10. Murphy NA, Arthur KC, Tienari PJ, Houlden H, Chiò A, Traynor BJ. Age-related penetrance of the C9orf72 repeat expansion. *Sci Rep*. 2017;7(1):2116. doi:10.1038/s41598-017-02364-1**

**11. Cady J, Allred P, Bali T, et al. Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes: Rare Variants in ALS Genes. *Ann Neurol*. 2015;77(1):100-113. doi:10.1002/ana.24306**

**12. McCann EP, Henden L, Fifita JA, et al. Evidence for polygenic and oligogenic basis of Australian sporadic amyotrophic lateral sclerosis. *J Med Genet*. 2021;58(2):87-95. doi:10.1136/jmedgenet-2020-106866**

**13. Pecoraro V, Mandrioli J, Carone C, Chiò A, Traynor BJ, Trenti T. The NGS technology for the identification of genes associated with the ALS. A systematic review. *Eur J Clin Invest*. 2020;50(5). doi:10.1111/eci.13228**

**14. Scoles DR, Meera P, Schneider MD, et al. Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. *Nature*. 2017;544(7650):362-366. doi:10.1038/nature22044**

**15. Cappella M, Pradat PF, Querin G, Biferi MG. Beyond the Traditional Clinical Trials for Amyotrophic Lateral Sclerosis and The Future Impact of Gene Therapy. *JND*. 2021;8(1):25-38. doi:10.3233/JND-200531**

## 3.7 Tables And Figures

**Figure 1. Simplified flowchart describing the steps taken to filter and prioritize variants in ALS genes.MAF, minor allele frequency; NFE, non-Finnish Europeans; ACMG, American College of Medical Genetic.**

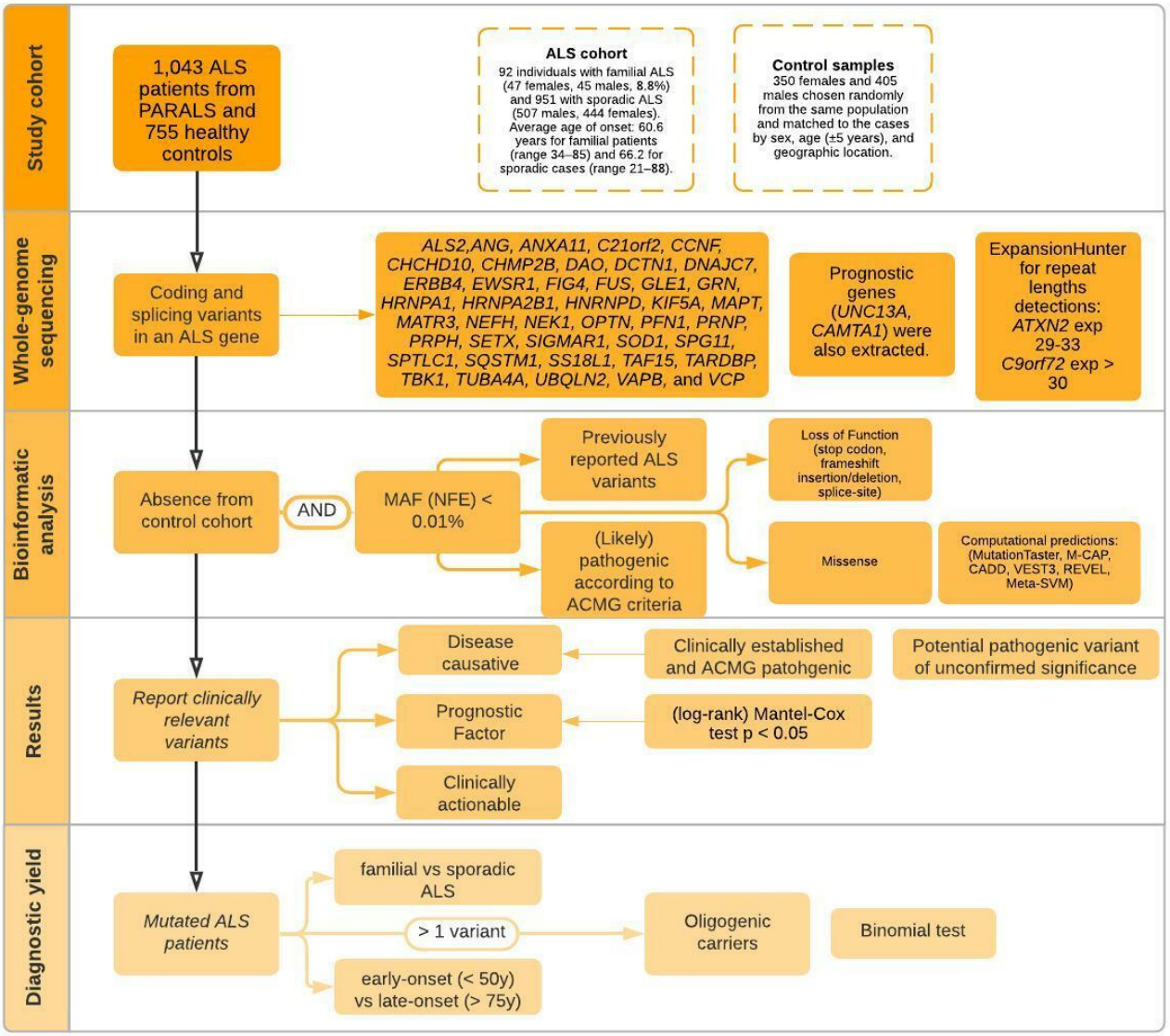
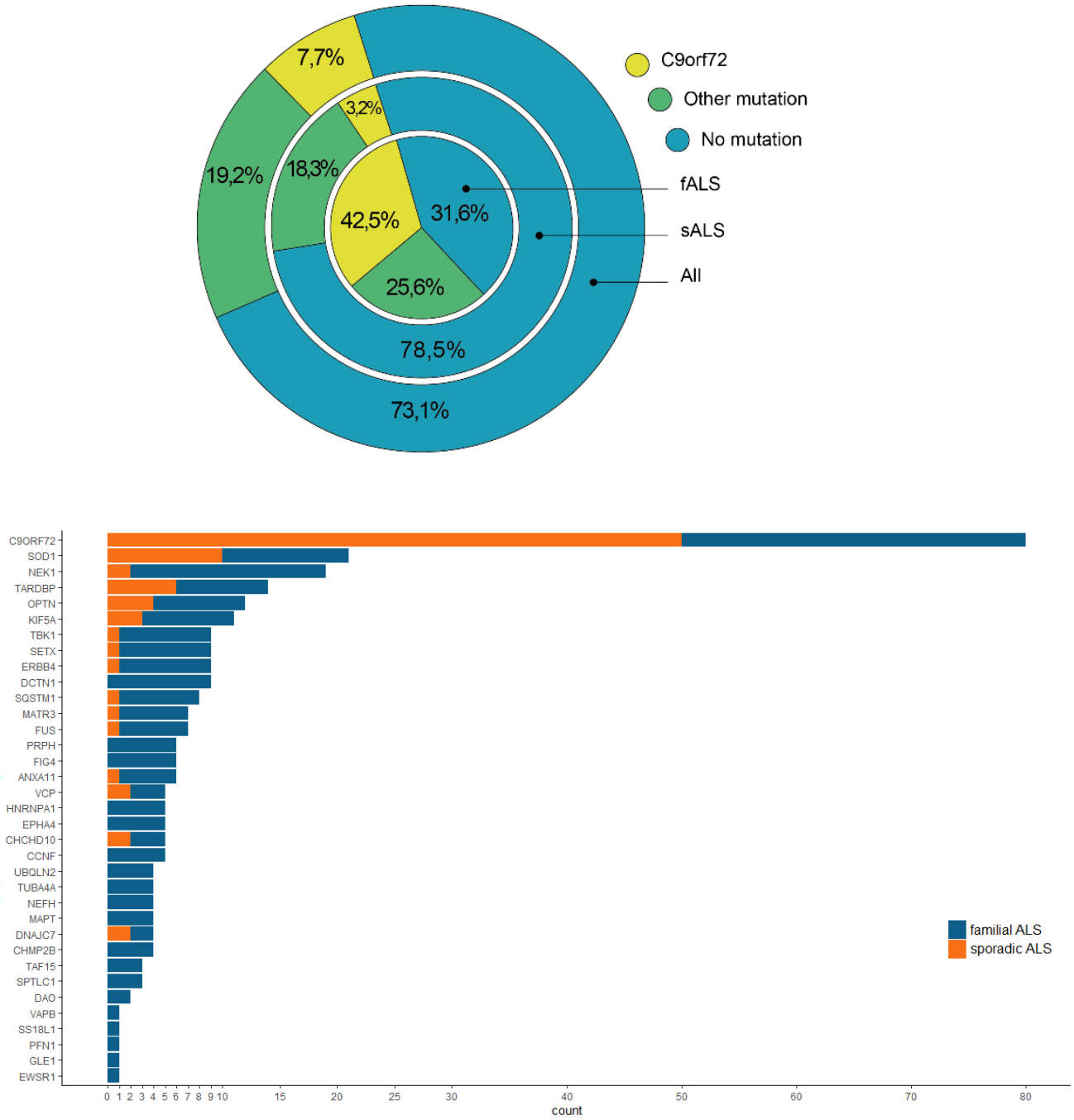


**Figure 2.** **Distribution of mutated genes across our cohort**

**The frequency of mutations in all ALS cases (out band), sporadic cases (middle band), and familial cases (inner circle). (B) The number of potentially pathogenic variants identified per gene**



## 3.8 Supplementary

### Supplementary Materials

**PATIENTS AND CONTROLS CHARACTERISTIC**

**Patients were included in PARALS if they met definite, probable, or probable laboratory-supported ALS diagnosis according to the revised El Escorial criteria1. Control samples were composed of 350 females and 405 males. The 755 controls were chosen randomly from the same population and matched to the cases by sex, age (±5 years), and geographic location. Control subjects were excluded from the study if they had a first-degree relative with a neurological or psychiatric disorder.**

**STANDARD PROTOCOL APPROVALS AND PATIENT CONSENTS**

**All subjects participating in the study gave written informed consent. Ethical approval was obtained from the medical ethical review board of the A.O.U Citta della Salute e della Scienza di Torino, Italy.**

**WHOLE-GENOME SEQUENCING**

**Whole-genome sequencing was performed at The American Genome Center at the Uniformed Services University on the Walter Reed National Military Medical Center campus in Bethesda, MD, USA. Briefly, libraries were prepared using TruSeq DNA PCR-Free High Throughput Library Prep Kit (Illumina Inc.) as per the manufacturer's instructions. Sequencing was performed on an Illumina HiSeq X10 sequencer using pairedend 150 base pair reads, and the data were processed according to Genome Analysis Toolkit's (GATK) best practices (https://software.broadinstitute.org/gatk/best-practices/). Variant quality control was performed using the GATK variant quality score method with default filters using Genome Reference Consortium Human Build 38 as the reference.**

**ALS-RELATED GENES**

**We extracted variant information from the data for the following genes: ALS2, ANXA11, ATXN2, C21orf2, C9orf72, CCNF, CHCHD10, CHMP2B, DAO, DCTN1, DNAJC7, ERBB4, EWSR1, FIG4, FUS, GLE1, GRN, HRNPA1, HRNPA2B1, HNRNPD, KIF5A, MAPT, MATR3, NEFH, NEK1, OPTN, PFN1, PRNP, PRPH, SETX, SIGMAR1, SOD1, SPG11, SPTLC1, SQSTM1, SS18L1, TAF15, TARDBP, TBK1, TUBA4A, UBQLN2, VAPB, and VCP. Prognostic genes (UNC13A, CAMTA1) were also extracted.**

**VARIANT ANNOTATION**

**Annotation was then performed using ANNOVAR v2020-06-07 (https://annovar.openbioinformatics.org) and KGGseq v1.0 (http://pmglab.top/kggseq/) using the gnomAD database (version 2.1.1) to determine minor allele frequency (MAF) in the European-**

**derived population. The current study was focused on coding variants due to the limitations in interpreting non-coding variants.**

**EXPANSION SCREENING**

**The samples were screened for C9orf72 expansions and the microsatellite repeat in ATXN2 using repeat-primed PCR methodology as previously described.2,3 A cut-off of 30 repeated expansion and the characteristic sawtooth pattern was considered pathogenic for C9orf722. ATXN2 CAG expansions were deemed intermediate if they were within the range of 30-33 repeats3. ExpansionHunter - Targeted software (version 0.3) was used to estimate repeat lengths of known, disease-causing expansions in samples undergoing whole-genome sequencing. This algorithm has been validated using experimentally confirmed samples carrying expansions4. The performance of whole-genome sequencing in terms of sensitivity for repeated expansion has been validated in the same ALS cohort 5,6.**

**VARIANT INTERPRETATION**

**We set a conservative MAF frequency threshold of <0.01% based on the epidemiology of ALS We defined Loss of Function (LoF) variants when the sequence changes were predicted to be a premature stop codon, a frameshift causing insertion/deletion (indel), or a splice-site disrupting variant located in the canonical splice sites (+1 and +2, –1 and –2) that cause the premature termination codon falling < 50–55 nucleotides upstream of the 3′ most exon–exon junctions. Loss of function variants was considered deleterious unless they exceeded the 0.01% MAF threshold. For all other type of variants, the combination of different sets of algorithms was considered using recommended treshold [MutationTaster, Mendelian Clinically Applicable Pathogenicity (MCAP), CADD, Variant Effect Scoring Tool (VEST3), Rare Exome Variant Ensemble Learner (REVEL), Metaanalytic Support Vector Machine (Meta-SVM)]7. hose variants underwent an independent review by clinical and genetic ALS experts and were confirmed to be clinically reportable if agreed by consensus. When available, gene-specific information data was used in the classification**

**STATISTICAL ANALYSIS**

**A 2-tailed Fisher exact test was used to evaluate the genetic association between ATXN2 CAG repeat sizes and ALS (significance set at P < 0.05). The burden of multiple variant carriers was assessed by a binomial test8. Age at disease onset and disease progression were also assessed across oligogenic and monogenic with analysis of variance (ANOVA). The analyses were performed in R (version 3.6.0).**

**DATA AVAILABILITY STATEMENT**

**The individual-level sequence data are available on dbGaP (accession number: phs001963.v1.p1) upon motivated request by interested researchers.**

**Supplementary Figure 1.** **Summary of identified variants.**

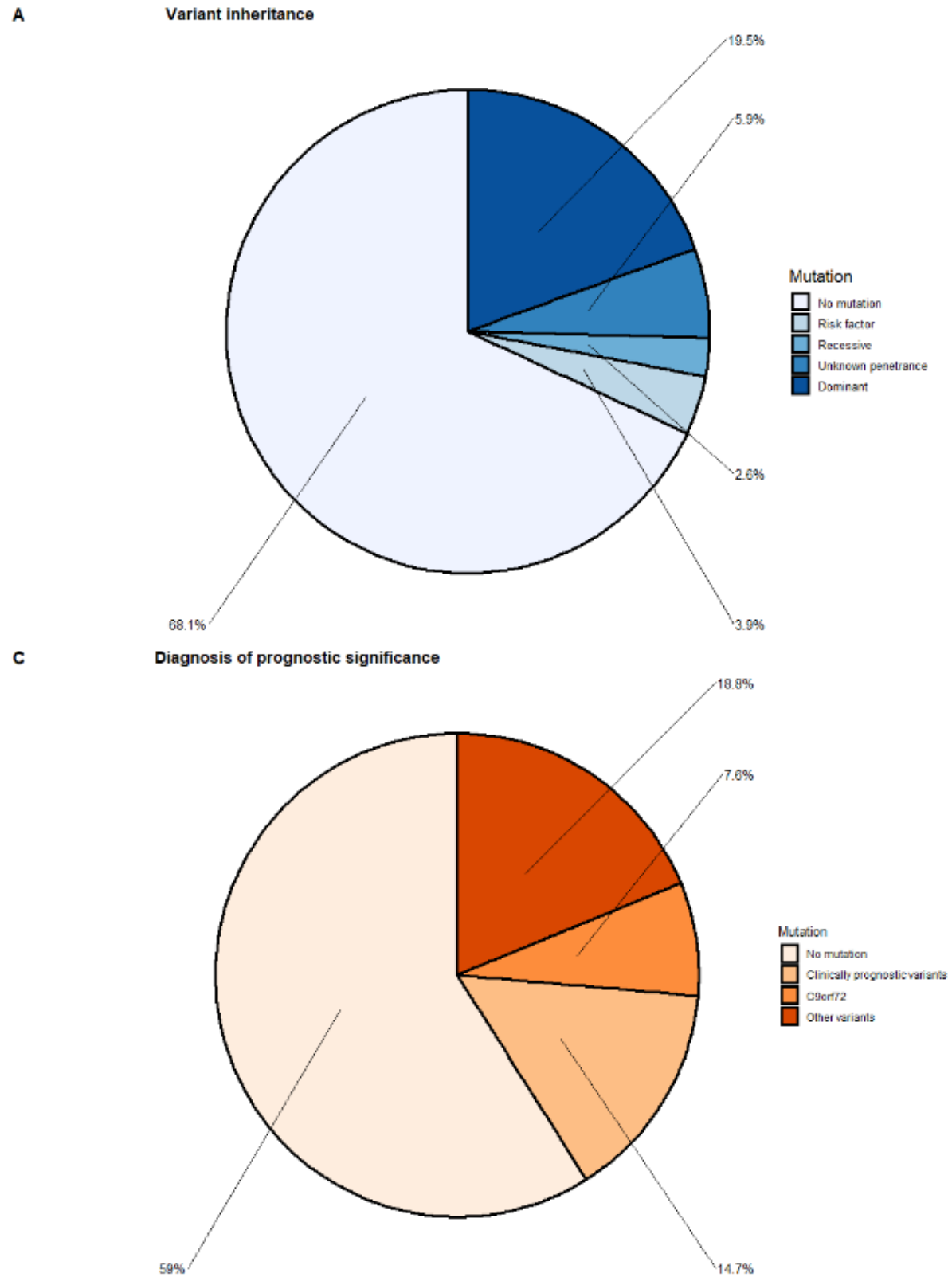
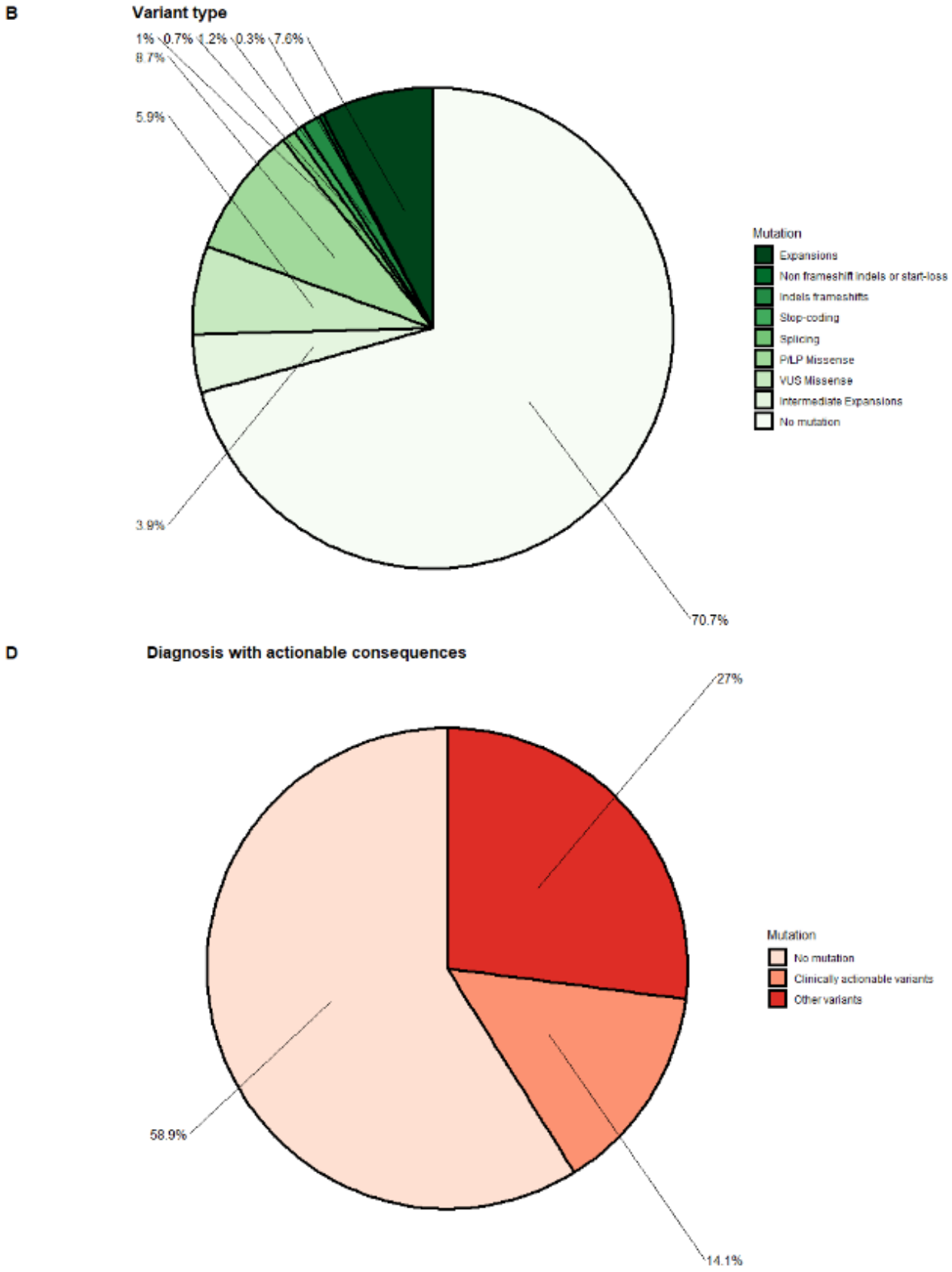


**Figure. 3. Summary of identified variants (continued).**



**Supplementary Table 1. Variants detected in our cohort previously reported in ALS cases.**

| Gene | Transcript | Nucleotide change | Amino acid change | Exon | Function |
|---|---|---|---|---|---|
| *ALS2* | NM_020919 | c.4867C>T | p.L1623F | exon33 | missense |
| *ALS2* | NM_020919 | c.3836+1G>A | . | exon24 | splicing |
| *ALS2* | NM_020919 | c.3797C>G | p.P1266R | exon24 | missense |
| *ALS2* | NM_020919 | c.3462G>T | p.Q1154H | exon21 | missense |
| *ALS2* | NM_020919 | c.3394C>A | p.R1132S | exon21 | missense |
| *ALS2* | NM_020919 | c.3206G>A | p.G1069E | exon19 | missense |
| *ALS2* | NM_020919 | c.2979+1G>A | . | exon17 | splicing |
| *ALS2* | NM_020919 | c.2479A>T | p.T827S | exon13 | missense |
| *ALS2* | NM_020919 | c.1804C>T | p.R602C | exon8 | missense |
| *ALS2* | NM_020919 | c.1649C>T | p.P550L | exon7 | missense |
| *ALS2* | NM_020919 | c.1628A>G | p.D543G | exon6 | missense |
| *ALS2* | NM_020919 | c.1550C>G | p.A517G | exon6 | missense |
| *ALS2* | NM_020919 | c.1472-1G>C | . | exon6 | splicing |
| *ALS2* | NM_020919 | c.37G>A | p.G13R | exon3 | missense |
| *ANXA11* | NM_001157 | c.922C>T | p.R308X | exon8 | stopgain |
| *ANXA11* | NM_001157 | c.102G>A | p.M34I | exon3 | startloss |
| *ANXA11* | NM_001157 | c.744+2C>T | . | exon7 | splicing |
| *C21ORF2* | NM_001271441 | c.418_420del | p.G140del | exon5 | non-frameshift deletion |
| *C21ORF2* | NM_001271441. | c.735_736insCGT.. | p.P245_V246insR | exon6 | non-frameshift insertion |
| *C21ORF2* | NM_001271441. | c.701_702insGC.. | p.V235Pfs*38 | exon6 | frameshift insertion |
| *CCNF* | NM_001761 | c.139C>T | p.L47F | exon2 | missense |
| *CCNF* | NM_001761 | c.316C>G | p.L106V | exon4 | missense |
| *CCNF* | NM_001761 | c.1043C>T | p.P348L | exon9 | missense |
| *CCNF* | NM_001761 | c.1132G>A | p.V378I | exon11 | missense |
| *CHCHD10* | NM_2137200 | c.239C>T | p.P80L | exon2 | missense |
| *CHMP2B* | NM_014043 | c.74A>G | p.Q25R | exon2 | missense |
| *CHMP2B* | NM_014043 | c.85A>G | p.I29V | exon2 | missense |
| *DAO* | NM_001917 | c.34G>A | p.G12R | exon2 | missense |
| *DAO* | NM_001917 | c.992G>A | p.G331E | exon11 | missense |
| *DCTN1* | NM_001190836 | c.3575C>T | p.A1192V | exon28 | missense |
| *DCTN1* | NM_001190836 | c.3574G>A | p.A1192T | exon28 | missense |
| *DCTN1* | NM_001190836 | c.2588T>C | p.M863T | exon20 | missense |
| *DCTN1* | NM_001190836 | c.2522A>G | p.Y841C | exon20 | missense |
| *DCTN1* | NM_001190836 | c.2321C>G | p.P774R | exon18 | missense |

| | | | | | |
|---|---|---|---|---|---|
| *DCTN1* | NM_001190836 | c.2147C>T | p.T716M | exon17 | missense |
| *DCTN1* | NM_001190836 | c.1484G>C | p.R495P | exon12 | missense |
| *DCTN1* | NM_001190836 | c.964C>T | p.Q322* | exon8 | stopgain |
| *DCTN1* | NM_001190836 | c.884T>C | p.V295A | exon7 | missense |
| *DNAJC7* | NM_003315 | c.611G>A | p.R204Q | exon7 | missense |
| *DNAJC7* | NM_003315 | c.203G>A | p.R68Q | exon3 | missense |
| *EPHA4* | NM_001304537 | c.2060C>G | p.S687C | exon12 | missense |
| *EPHA4* | NM_001304537 | c.1702T>C | p.S568P | exon9 | missense |
| *EPHA4* | NM_001304537 | c.1058T>C | p.L353P | exon4 | missense |
| *ERBB4* | NM_005235 | c.3446G>T | p.G1149V | exon27 | missense |
| *ERBB4* | NM_005235 | c.2525G>A | p.R842Q | exon21 | missense |
| *ERBB4* | NM_005235 | c.1913G>A | p.W638* | exon16 | stopgain |
| *ERBB4* | NM_005235 | c.1912T>C | p.W638R | exon16 | missense |
| *ERBB4* | NM_005235 | c.1898G>C | p.C633S | exon16 | missense |
| *ERBB4* | NM_005235 | c.1772A>G | p.E591G | exon15 | missense |
| *ERBB4* | NM_005235 | c.586C>T | p.R196C | exon5 | missense |
| *ERBB4* | NM_005235 | c.532G>C | p.V178L | exon4 | missense |
| *ERBB4* | NM_005235 | c.308G>A | p.R103H | exon3 | missense |
| *EWSR1* | NM_001163285 | c.1447A>G | p.M483V | exon14 | missense |
| *EWSR1* | NM_001163285 | c.1775_1776insT | p.Q593Ffs*20 | exon17 | frameshift |
| *FIG4* | NM_014845 | c.30C>G | p.S10R | exon1 | missense |
| *FIG4* | NM_014845 | c.52T>C | p.Y18H | exon1 | missense |
| *FIG4* | NM_014845 | c.2200G>A | p.E734K | exon20 | missense |
| *FIG4* | NM_014845 | c.2467C>T | p.Q823* | exon22 | stopgain |
| *FIG4* | NM_014845 | c.2639T>A | p.I880N | exon23 | missense |
| *FIG4* | NM_014845 | c.2650C>T | p.Q884* | exon23 | stopgain |
| *FUS* | NM_001170634 | c.785A>G | p.N262S | exon7 | missense |
| *FUS* | NM_001170634 | c.1480C>T | p.R494* | exon14 | stopgain |
| *FUS* | NM_001170634 | c.1539G>C | p.R513S | exon15 | missense |
| *FUS* | NM_001170634 | c.1552C>G | p.Q518E | exon15 | missense |
| *GLE1* | NM_001003722 | c.1771C>T | p.Q591* | exon12 | stopgain |
| *GRN* | NM_002087 | c.763C>A | p.L255M | exon8 | missense |
| *GRN* | NM_002087 | c.1562G>A | p.C521Y | exon12 | missense |
| *GRN* | NM_002087 | c.1595C>A | p.T532N | exon12 | missense |
| *HNRNPA1* | NM_031157 | c.380A>G | p.Q127R | exon4 | missense |
| *HNRNPA1* | NM_031157 | c.666C>G | p.F222L | exon6 | missense |
| *HNRNPA1* | NM_031157 | c.824G>T | p.G275V | exon8 | missense |
| *HNRNPA1* | NM_031157 | c.876C>G | p.N292K | exon8 | missense |
| *HNRNPA1* | NM_031157 | c.883G>A | p.G295R | exon8 | missense |
| *KIF5A* | NM_004984 | c.340C>T | p.R114* | exon4 | stopgain |

| | | | | | |
|---|---|---|---|---|---|
| *KIF5A* | NM_004984 | c.1463T>G | p.L488R | exon14 | missense |
| *KIF5A* | NM_004984 | c.2152C>T | p.R718W | exon19 | missense |
| *KIF5A* | NM_004984 | c.2263G>A | p.E755K | exon20 | missense |
| *KIF5A* | NM_004984 | c.2757del | p.K920Nfs*128 | exon25 | frameshift |
| *MAPT* | NM_016834 | c.149T>C | p.I50T | exon3 | missense |
| *MAPT* | NM_016834 | c.736G>A | p.G246S | exon7 | missense |
| *MAPT* | NM_016834 | c.880T>C | p.S294P | exon9 | missense |
| *MAPT* | NM_016834 | c.913G>A | p.V305I | exon9 | missense |
| *MATR3* | NM_018834 | c.368A>G | p.D123G | exon2 | missense |
| *MATR3* | NM_018834 | c.2148+2T>+GG | . | exon14 | splicing |
| *MATR3* | NM_018834 | c.2371+2T>+AT | . | exon14 | splicing |
| *NEFH* | NM_021076 | c.1036C>T | p.R346C | exon2 | missense |
| *NEFH* | NM_021076 | c.1723C>T | p.P575S | exon4 | missense |
| *NEFH* | NM_021076 | c.1783C>T | p.P595S | exon4 | missense |
| *NEFH* | NM_021076 | c.2461A>G | p.K821E | exon4 | missense |
| *NEK1* | NM_001199399 | c.3502C>T | p.H1168Y | exon32 | missense |
| *NEK1* | NM_001199399 | c.3451A>C | p.I1151L | exon32 | missense |
| *NEK1* | NM_001199399 | c.3419_3422del | p.I1140Rfs*17 | exon31 | frameshift |
| *NEK1* | NM_001199399 | c.3373G>A | p.E1125K | exon31 | missense |
| *NEK1* | NM_001199399 | c.3214delA | p.T1072Lfs*20 | exon30 | frameshift |
| *NEK1* | NM_001199399 | c.2816C>G | p.S939* | exon28 | stopgain |
| *NEK1* | NM_001199399 | c.2698G>T | p.D900Y | exon28 | missense |
| *NEK1* | NM_001199399 | c.2523_2526del | p.N841Kfs*53 | exon26 | frameshift |
| *NEK1* | NM_001199399 | c.1226G>A | p.W409* | exon15 | stopgain |
| *NEK1* | NM_001199399 | c.1129_1132del | p.Q377Rfs*7 | exon13 | frameshift |
| *NEK1* | NM_001199399 | c.781C>T | p.R261C | exon10 | missense |
| *NEK1* | NM_001199399 | c.577C>T | p.L193F | exon9 | missense |
| *NEK1* | NM_001199399 | c.464+1G>AG | . | exon6 | splicing |
| *NEK1* | NM_001199399 | c.449C>G | p.A150G | exon7 | missense |
| *NEK1* | NM_001199399 | c.380G>A | p.R127Q | exon6 | missense |
| *OPTN* | NM_001008212 | c.247C>T | p.R83C | exon4 | missense |
| *OPTN* | NM_001008212 | c.265C>T | p.Q89* | exon4 | stopgain |
| *OPTN* | NM_001008212 | c.332T>G | p.L111R | exon4 | missense |
| *OPTN* | NM_001008212 | c.403G>T | p.E135* | exon5 | stopgain |
| *OPTN* | NM_001008212 | c.644G>A | p.R215K | exon7 | |
| *OPTN* | NM_001008212 | c.917_921del | p.T307Sfs*3 | exon9 | frameshift |
| *OPTN* | NM_001008212 | c.1499T>C | p.L500P | exon13 | missense |
| *OPTN* | NM_001008212 | c.1634G>A | p.R545Q | exon15 | missense |
| *OPTN* | NM_001008212 | c.1639C>T | p.Q547* | exon15 | stopgain |
| *OPTN* | NM_001008212 | c.1719G>A | p.M573I | exon15 | missense |

| | | | | | |
|---|---|---|---|---|---|
| *PFN1* | NM_005022 | c.67G>A | p.V23M | exon1 | missense |
| *PRPH* | NM_006262 | c.487G>T | p.D163Y | exon1 | missense |
| *PRPH* | NM_006262 | c.611T>A | p.V204E | exon3 | missense |
| *PRPH* | NM_006262 | c.996+1G>A | . | exon6 | splicing |
| *PRPH* | NM_006262 | c.1024G>A | p.E342K | exon6 | missense |
| *PRPH* | NM_006262.4 | c.1222_1223insG | p.S408Cfs*19 | exon7 | frameshift |
| *PRPH* | NM_006262 | c.1409A>C p.Y470S | | exon9 | missense |
| *SETX* | NM_015046 | c.7240C>T | p.R2414* | exon25 | stopgain |
| *SETX* | NM_015046 | c.6848_6851del | p.T2283Kfs*32 | exon22 | frameshift |
| *SETX* | NM_015046 | c.6685A>G | p.M2229V | exon21 | missense |
| *SETX* | NM_015046 | c.1427A>G | p.H476R | exon10 | missense |
| *SETX* | NM_015046 | c.1427A>C | p.H476P | exon10 | missense |
| *SETX* | NM_015046 | c.1178T>C | p.L393P | exon10 | missense |
| *SETX* | NM_015046 | c.638C>T | p.S213F | exon6 | missense |
| *SETX* | NM_015046 | c.62A>T | p.Y21F | exon3 | missense |
| *SETX* | NM_015046 | c.4A>G | p.S2G | exon3 | missense |
| *SOD1* | NM_000454 | c.16G>A | p.V6M | exon1 | missense |
| *SOD1* | NM_000454 | c.59A>G | p.N20S | exon1 | missense |
| *SOD1* | NM_000454 | c.115C>G | p.L39V | exon2 | missense |
| *SOD1* | NM_000454 | c.197A>G | p.N66S | exon3 | missense |
| *SOD1* | NM_000454 | c.217G>A | p.G73S | exon3 | missense |
| *SOD1* | NM_000454 | c.271G>A | p.D91N | exon4 | missense |
| *SOD1* | NM_000454 | c.281G>A | p.G94D | exon4 | missense |
| *SOD1* | NM_000454 | c.357+1G>+T | . | exon5 | splicing |
| *SOD1* | NM_000454 | c.409A>T | p.K137* | exon5 | stopgain |
| *SOD1* | NM_000454 | c.435G>C | p.L145F | exon5 | missense |
| *SOD1* | NM_000454 | c.435G>T | p.L145F | exon5 | missense |
| *SOD1* | NM_000454 | c.442G>A | p.G148S | exon5 | missense |
| *SPAST* | NM_001363875 | c.455T>C | p.I152T | exon2 | missense |
| *SPAST* | NM_001363875 | c.806C>T | p.S269F | exon5 | missense |
| *SPAST* | NM_001363875 | c.1526A>G | p.D509G | exon14 | missense |
| *SPG11* | NM_025137 | c.6877C>T | p.R2293W | exon38 | missense |
| *SPG11* | NM_025137 | c.6857G>C | p.R2286P | exon38 | missense |
| *SPG11* | NM_025137 | c.6094T>C | p.C2032R | exon32 | missense |
| *SPG11* | NM_025137 | c.6010T>G | p.L2004V | exon32 | missense |
| *SPG11* | NM_025137 | c.5414G>A | p.R1805H | exon30 | missense |
| *SPG11* | NM_025137 | c.3095C>T | p.P1032L | exon17 | missense |
| *SPG11* | NM_025137 | c.148C>T | p.Q50* | exon1 | stopgain |
| *SQSTM1* | NM_003900 | c.447C>G | p.D149E | exon3 | missense |
| *SQSTM1* | NM_003900 | c.775G>A | p.V259M | exon6 | missense |

| | | | | | |
|---|---|---|---|---|---|
| *SQSTM1* | NM_003900 | c.775G>C | p.V259L | exon6 | missense |
| *SQSTM1* | NM_003900 | c.1043C>T | p.P348L | exon7 | missense |
| *SQSTM1* | NM_003900 | c.1084G>A | p.E362K | exon7 | missense |
| *SQSTM1* | NM_003900 | c.1142C>T | p.A381V | exon7 | missense |
| *SS18L1* | NM_001301778 | c.253A>C | p.M85L | exon7 | missense |
| *TAF15* | NM_139215 | c.116G>A | p.G39E | exon4 | missense |
| *TAF15* | NM_139215 | c.329A>G | p.Y110C | exon6 | missense |
| *TAF15* | NM_139215 | c.1088+1G>A | . | exon14 | splicing |
| *TARDBP* | NM_007375 | c.800A>G | p.N267S | exon6 | missense |
| *TARDBP* | NM_007375 | c.1144G>A | p.A382T | exon6 | missense |
| *TARDBP* | NM_007375 | c.1169A>G | p.N390S | exon6 | missense |
| *TARDBP* | NM_007375 | c.1178C>T | p.S393L | exon6 | missense |
| *TBK1* | NM_013254 | c.254T>C | p.I85T | exon4 | missense |
| *TBK1* | NM_013254 | c.454G>C | p.V152L | exon5 | missense |
| *TBK1* | NM_013254 | c.521A>G | p.Y174C | exon5 | missense |
| *TBK1* | NM_013254 | c.992+1G>A | . | exon9 | splicing |
| *TBK1* | NM_013254 | c.1343_1346del | p.I450Kfs*15 | exon12 | frameshift |
| *TIA1* | NM_022173 | c.1154C>T | p.T385I | exon13 | missense |
| *TIA1* | NM_022173 | c.746A>C | p.K249T | exon10 | missense |
| *TIA1* | NM_022173 | c.95C>A | p.P32H | exon2 | missense |
| *TIA1* | NM_022173 | c.94C>A | p.P32T | exon2 | missense |
| *TUBA4A* | NM_001278552 | c.1312G>A | p.D438N | exon4 | missense |
| *TUBA4A* | NM_001278552 | c.1184_1190del | p.G395Afs*67 | exon4 | frameshift deletion |
| *TUBA4A* | NM_001278552 | c.1045A>T | p.T349S | exon4 | missense |
| *TUBA4A* | NM_001278552 | c.148delG | p.A50Qfs*89 | exon2 | frameshift deletion |
| *UBQLN2* | NM_013444 | c.401C>T | p.T134I | exon1 | missense |
| *UBQLN2* | NM_013444 | c.1172A>G | p.Y391C | exon1 | missense |
| *UBQLN2* | NM_013444 | c.1505G>A | p.G502D | exon1 | missense |
| *VAPB* | NM_004738 | c.332C>T | p.P111L | exon4 | missense |
| *VCP* | NM_001354927 | c.2086C>T | p.R696C | exon16 | missense |
| *VCP* | NM_001354927 | c.2002A>G | p.R668G | exon15 | missense |
| *VCP* | NM_001354927 | c.265T>C | p.Y89H | exon4 | missense |

## Supplementary Table 2A. Oligogenic combinations in our cohort

| Mutation 1 | Mutation 2 |
|---|---|
| *C9orf72* GGGGCC exp | *CHCHD10*:c.239C>T:p.P80L |
| *C9orf72* GGGGCC exp | *KIF5A*:c.T1463G:p.L488R |
| *C9orf72* GGGGCC exp | *C21orf2*:c.700_701ins:p.V246 Wfs*35 |
| C9orf72 GGGGCC exp | *NEK1*:c.G3373A:p.E1125K |
| *C9orf72* GGGGCC exp | *OPTN*:c.G644A:p.R215K |
| *SOD1*:c.59A>G:p.N20S | *FUS*:c.1542G>C:p.R514S |
| *SOD1*:c.115C>G:p.L39V | *MATR3*:c.2148+2->GGAC |
| *SOD1*:c.271G>A:p.D91N | *CHCHD10*:c.239C>T:p.P80L |
| *SOD1*:c.435G>C:p.L145F | *OPTN*:c.403G>T:p.E135* |
| *SOD1*:c.435G>C:p.L145F | *ERBB4*:c.308G>A:p.R103H |
| *TARDBP*:c.1144G>A:p.A382T | *UBQLN2*:c.1612G>C:p.V538L |
| *DCTN1*:c.C673T:p.Q225* | *MATR3*:c.2148+2->GGAC |
| *ANXA11*:c.C922T:p.R308X | *ERBB4*:c.A1772G:p.E591G |

## Supplementary Table 2B. Analysis of oligogenic frequency and effect of multiple variants on age at disease onset in our cohort

| | Oligogenic cases | Monogenic cases | Binomial p-value | ANOVA p-value |
|---|---|---|---|---|
| **Cases (% cohort)** | 13 (1.3%) | 266 (25.6%) | 0.98 | - |
| **Age at onset (SD)** | 59.9 (11.9) | 63.9 (11.2) | - | 0.1310 |

## Supplementary References


1. **Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*. 2000;1(5). doi:10.1080/146608200300079536**

2. **Renton AE, Majounie E, Waite A, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*. 2011;72(2). doi:10.1016/j.neuron.2011.09.010**

3. **Van Damme P, Veldink JH, van Blitterswijk M, et al. Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology*. 2011;76(24):2066-2072. doi:10.1212/WNL.0b013e31821f445b**

4. **Dolzhenko E, van Vugt JJFA, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Research*. 2017;27(11). doi:10.1101/gr.225672.117**

5. **Grassano M, Calvo A, Moglia C, et al. Mutational analysis of known ALS genes in an Italian population-based cohort. *Neurology*. Published online November 2020. doi:10.1212/WNL.0000000000011209**

6. **Dewan R, Chia R, Ding J, et al. Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. *Neuron*. 2021;109(3):448-460.e4. doi:10.1016/j.neuron.2020.11.005**

7. **Anderson D, Lassmann T. A phenotype centric benchmark of variant prioritisation tools. *npj Genom Med*. 2018;3(1):5. doi:10.1038/s41525-018-0044-9**

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

# 4. GBA variants influence cognitive status in amyotrophic lateral sclerosis

**As published in *Canosa, A., Grassano, M., Moglia, C., Iazzolino, B., Peotta, L., Gallone, S., ... & Chio, A. (2022). GBA variants influence cognitive status in amyotrophic lateral sclerosis. Journal of Neurology, Neurosurgery & Psychiatry, 93(4), 453-455.***

**AUTHORS**

**Antonio Canosa # [1 2 3], Maurizio Grassano # [1 4], Cristina Moglia [1 2], Barbara Iazzolino [1], Laura Peotta [1], Salvatore Gallone [2], Maura Brunetti [1], Marco Barberis [5], Luca Sbaiz [5], Francesca Palumbo [1], Sara Cabras [1], Umberto Manera [1], Rosario Vasta [1], Bryan Traynor [4 6], Lucia Corrado [7], Sandra D'Alfonso [7], Letizia Mazzini [8], Andrea Calvo # [1 2], Adriano Chio # [2 3 9].**

**Author Affiliations**

**1"Rita Levi Montalcini" Department of Neuroscience, ALS Centre, University of Turin, Turin, Italy.**

**2SC Neurologia 1U, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy.**

**3Institute of Cognitive Sciences and Technologies, National Council of Research, Rome, Italy.**

**4Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD, USA.**

**5Laboratory of Genetics, Department of Clinical Pathology, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy.**

**6Department of Neurology, Johns Hopkins University Medical Center, Baltimore, MD, USA.**

**7Department of Health Sciences Interdisciplinary Research Center of Autoimmune Diseases, University of Eastern Piedmont Amedeo Avogadro School of Medicine, Novara, Piemonte, Italy.**

**8ALS Center, Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy.**

**9"Rita Levi Montalcini" Department of Neuroscience, ALS Centre, University of Turin, Turin, Italy achio@usa.net.**

**#Contributed equally.**

## 4.1 Introduction

**Amyotrophic lateral sclerosis (ALS) is a relentlessly progressive degenerative disease of upper and lower motor neurons. Approximately 15% of patients display clinical features consistent with frontotemporal dementia (FTD) and 35% display milder degrees of cognitive and behavioural impairment at some stage during their illness.1 Several genes have been reported to cause both ALS and FTD. Nevertheless, it remains unclear why some patients with ALS develop cognitive impairment, while other cases, often within the same family, remain unaffected.**

**The GBA gene (OMIM *606463) encodes glucocerebrosidase (GCase), a lysosomal enzyme that converts glucocerebroside into glucose and ceramide. Heterozygous GBA mutations increase the risk of Parkinson's disease (PD) and the risk of cognitive impairment in patients with PD.2**

**It is increasingly recognised that variants in genes causing Mendelian neurodegenerative diseases may exhibit pleiotropic effects and impact the phenotypic heterogeneity of those disorders. Moreover, lysosomal dysfunction has recently been associated with both Dementia with Lewy Bodies and FTD spectrum (online supplemental table 1). Based on this, we postulated that GBA variants may influence the cognitive status of patients with ALS.**

## 4.2 Methods

**We examined the GBA variants' association with the risk of cognitive impairment in 751 patients with ALS from the population-based Piemonte and Valle d'Aosta Register for ALS that had undergone both a detailed neuropsychological evaluation and a whole-genome sequencing screening.3 Patients were classified as ALS with normal cognitive function (ALS-CN), ALS-FTD and ALS with intermediate cognitive deficits. The characteristics of the study population and a detailed description of neuropsychological testing and genetic screening are reported in online supplementarl materials in the Methods section.**

**A mutational screening of GBA exonic variants was performed and their frequencies were compared with an internal control cohort (see online supplemental methods, Subjects). To assess whether**

**pathogenic rare variants (minor allele frequency<1%) in GBA contribute to cognitive decline risk in ALS, a gene-based rare variants association test was performed as previously described.3 In the following step of the analysis, only variants known to be a risk factor for cognitive decline in PD were considered. First, a binomial test was used to assess the prevalence of GBA mutations across cognitive groups. Then, a linear mixed-effects model was used to test for associations between GBA genotype and cognitive functioning while including the following covariates: sex, age, site of disease onset, bulbar signs at diagnosis, rate of ALS Functional Rating Scale-Revised (ALS-FRS-R) decline and C9orf72 status. Further details on the statistical analysis are reported in online supplementary materials. All statistical analyses were performed in R V.3.6.0 (https://www.r-project.org/).**

## 4.3 Results

**The gene-based rare variant association test identified an enrichment of rare GBA variants in patients with ALS with intermediate cognitive dysfunction (p SKAT-O=0.000005), but not in ALS-FTD cases (p SKAT-O=0.184) (online supplemental table 2).**

**We identified one GBA mutation (p.N409S), known to cause Gaucher Disease in homozygous carriers, one likely pathogenic variant (p.R209H) and two GBA polymorphisms that are known to increase the risk of dementia in patients with PD (p.E365K and p.T408M). The remaining identified coding GBA variants are reported in online supplemental table 3. The frequency of GBA variants was not increased in our cohort as compared with healthy controls (online supplemental table 4). Thirteen out of 18 (72.2%) of patients with ALS carrying GBA variants displayed cognitive impairment in the form of FTD or intermediate cognitive phenotypes. In contrast, cognitive impairment was observed among 47.1% (298 out of 733) of patients with ALS not carrying GBA variants (binomial test p value=0.0357). We repeated the analysis excluding C9orf72 expansion carriers and the difference remained significant (p=0.0486). To confirm the effect of GBA variants on cognitive phenotype, we modelled the association between the GBA variants and cognitive impairment using a linear mixed-effects model that controlled for relevant covariates (figure 1). In the mixed-effects model, GBA mutation status was associated with the clinical diagnosis of cognitive impairment (OR=3.74, 95% CI 1.25 to 12.72, p=0.023). This effect was not seen when considering either only the ALS-FTD phenotype or the intermediate phenotype.**

## 4.4 Discussion

**We observed that the burden of rare variants in the GBA gene was associated with cognitive impairment in patients with ALS. Patients carrying known pathogenic GBA variants (p.E365K, p.T408M, p.N409S) were three times more likely to develop cognitive impairment compared with non-carriers, independently of age, sex, site of onset, bulbar involvement, rate of ALS-FRS-R decline and C9orf72 status.**

**In multivariate analysis, we identified an effect of GBA variants only when ALS-CN cases were compared with patients with FTD and intermediate deficits combined. However, the results of the burden test suggest that this finding is primarily driven by patients with intermediate deficits. The course of cognitive deterioration among patients with ALS may partially explain our findings: recent studies have shown that cognitive impairment may worsen over time and that it is correlated to more severe motor deficits.4 The neuropsychological assessment at diagnosis might have captured an early phase of the trajectory of cognitive deterioration over time. Nonetheless, also the small number of GBA variant carriers may have conditioned such findings.**

**A possible role of GBA in the neurodegenerative process underlying ALS is suggested by increasing evidence of the involvement of endolysosomal dysfunction in ALS pathogenesis. Several genes causing ALS and FTD, including C9orf72, TBK1, OPTN, SQSTM1 and VCP, are related to lysosomal function and protein degradation. This research field deserves further attention as several therapeutic agents targeting lysosomal pathways have been proposed for neurological diseases.5 Our results expand the spectrum of neurodegenerative diseases for which heterozygous GBA variants represent a detrimental factor. It is possible that such variants are kept in populations because they provide some biological advantage.**

**As a limitation of our study, we acknowledge that we could not evaluate whether different variants had a variable impact on the risk of cognitive impairment, on the pattern of cognitive deficits and on other clinical characteristics, mostly due to the relatively small number of GBA variant carriers (online supplemental tables 5 and 6).**

**In conclusion, we found that variants of the GBA gene are associated with an increased risk of cognitive impairment in patients with ALS. Our results broaden the spectrum of genetic factors that modulate the vulnerability of patients with ALS to cognitive dysfunction and strengthen the role of lysosomal impairment in the neurodegenerative process underlying ALS, highlighting that genes can modify not only the risk of ALS but also modulate different aspects of its phenotype. Addressing the**

**gap in our understanding of the role that genetic modifiers play in ALS is essential for diagnosis, prognosis, and therapy development.**

## 4.5 References

**1. Phukan J, Elamin M, Bede P, et al.. The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study. J Neurol Neurosurg Psychiatry 2012;83:102–8. 10.1136/jnnp-2011-300188**

**2. Sidransky E, Nalls MA, Aasly JO, et al.. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med 2009;361:1651–61. 10.1056/NEJMoa0901281 [PMC free article]**

**3. Grassano M, Calvo A, Moglia C. Mutational analysis of known ALS genes in an Italian population-based cohort. Neurology 2020;96. 10.1212/WNL.0000000000011209 [PMC free article]**

**4. Beeldman E, Govaarts R, de Visser M, et al.. Progression of cognitive and behavioural impairment in early amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2020;91:779–80. 10.1136/jnnp-2020-322992 [PMC free article]**

**5. Peng W, Minakaki G, Nguyen M, et al.. Correction to: preserving lysosomal function in the aging brain: insights from neurodegeneration. Neurotherapeutics 2019;16:920–1. 10.1007/s13311-019-00756-x [PMC free article]**

## 4.6 Tables And Figures

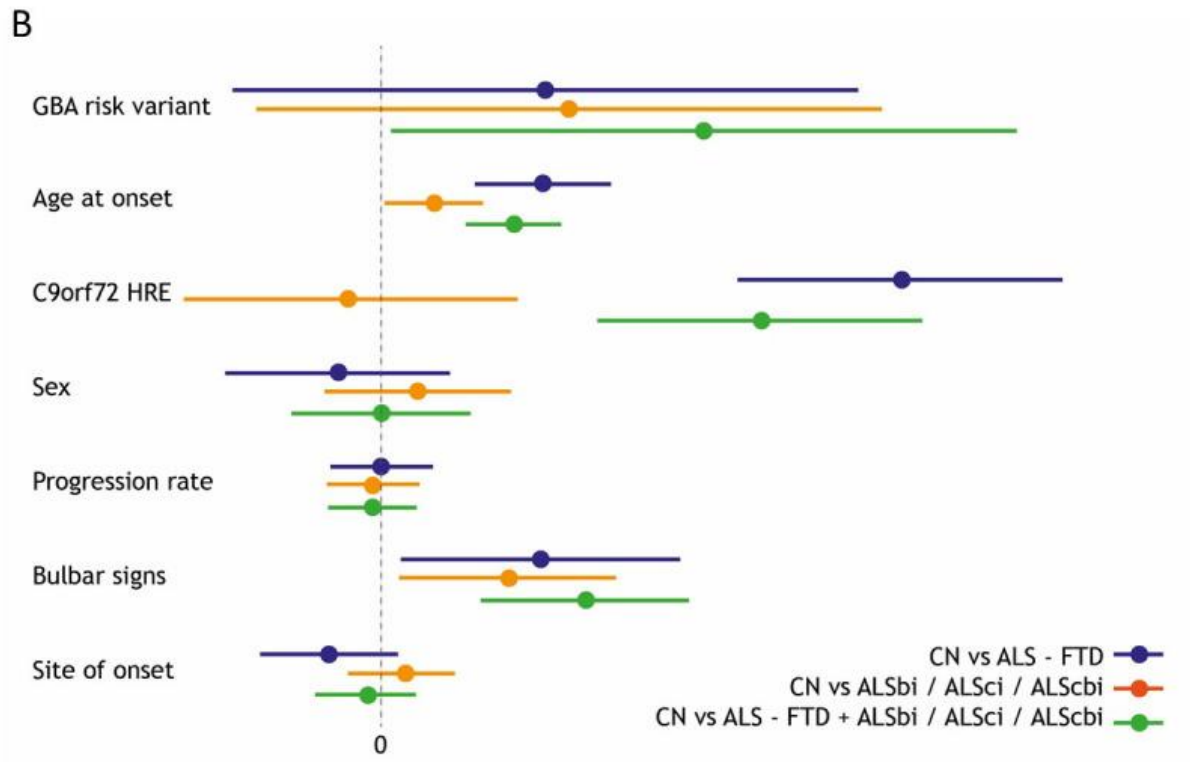
### Figure 1

**(A) Cognitive phenotype frequencies in GBA risk variant carriers. Given the strong influence of C9orf72 on cognitive status, here we performed the binomial test without C9orf72 expansion carriers to rule out its impact on the results. (B) Results of the linear mixed-effects model, coefficient and 95% CI. Model estimates are expressed as eOR. Colours correspond to the results of the three different models (1) ALS-CN versus ALS-FTD (blue); (2) ALS-CN versus ALSbi+ALSci+ALScbi (red); (3) ALS-CN versus ALS-FTD+ALSbi+ALSci+ALScbi (green). See online supplemental file 1 for further information about cognitive classification. ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia.**

A

| Phenotype | *GBA* risk variants carriers | *GBA* risk variants non carriers | Binomial test P-value (vs Cn) |
|---|---|---|---|
| ALS-FTD | 4 (3.7%) | 105 (96.3%) | 0.1197 |
| ALSbi/Ci/Cbi | 8 (4.0%) | 194 (96.0%) | 0.0749 |
| ALS-FTD + ALSbi/Ci/Cbi | 12 (3.9%) | 299 (96.1%) | 0.0486 |
| CN | 5 (1.4%) | 365 (98.6%) | |
| Total | 17 | 664 | |

B



## 4.7 Supplementary

### Supplementary Methods

#### SUBJECTS

**The study population consisted of incident patients identified through the Piemonte and Valle d'Aosta Register for ALS (PARALS), diagnosed with definite, probable, and probable laboratory-supported ALS according to the El Escorial revised diagnostic criteria[1] between**

**January 1st 2007 and December 31st 2015. The PARALS is a prospective epidemiologic registry established in 1995 in Piemonte and Valle d'Aosta regions in Northern Italy.[2] Patients were followed at the ALS Centers in Turin and Novara. Patients with a history of other neurologic disorders affecting cognition (neurodegenerative diseases other than FTD, major stroke, severe head injuries, mental retardation), alcohol and drug dependence, severe mental illness, and use of high-dose psychoactive medications were excluded from the analysis.[3] Non-Italian speaking incident patients were also excluded. None of the study participants showed oxygen saturation <92% based on pulse oximetry at the time of their neuropsychological assessment. Six hundred seventy-seven healthy controls, matched by age, sex, and ancestry (based on the origin of both parents), were recruited from the patients' general practitioners. Out of 1313 incident cases in the period 2007-2015, 751 ALS patients from the participant ALS Centers (57.2%) underwent whole-genome sequencing and full neuropsychological assessment.**

#### NEUROPSYCHOLOGICAL EVALUATION

**An extensive list of neuropsychological tests was performed on each patient.[4] These included: Mini-Mental State Examination (MMSE); Letter and category fluency tests; Frontal Assessment Battery (FAB); Digit Span Forward and Backward; Trail-Making Test (TMT) A and B; Rey Auditory Verbal Learning Test (RAVLT), immediate and delayed recall; Babcock Story Recall Test (BSRT), immediate and delayed recall; Rey-Osterrieth Complex Figure (ROCF), copy and delayed recall; and Raven's Colored Progressive Matrices (CPM47). Neurobehavioral dysfunction was determined by patient history, direct observation by the neuropsychologist, and the family form of the Frontal Systems Behavior Scale completed by a close relative or caregiver. The Edinburgh Cognitive and Behavioral ALS Screen (ECAS) was not included in the battery since the Italian validation was published in 2018.[5] Anxiety and depression were assessed using the Hospital Anxiety and Depression Scale (HADS); the item "I feel slowed down" was discussed with patients in order to have them not to refer to physical disability.[3] The raw data of the neuropsychological tests were adjusted for age and years of**

**education according to the Italian normative. The battery was administered following the same sequence to avoid differential interference. All study participants were tested at diagnosis or during the first follow-up visit (2 months later).**

## COGNITIVE CATEGORIZATION

**Patients' cognitive status was classified according to the revised ALS-FTD Consensus Criteria[6] into five categories:**

- **ALS with normal cognition (ALS-CN);**
- **ALS with behavioral impairment (ALSbi);**
- **ALS with cognitive impairment (ALSci);**
- **ALS with cognitive and behavioral impairment (ALScbi)**
- **ALS with FTD (ALS-FTD).**

**Patients were retrospectively classified blindly by two neuropsychologists with expertise in ALS (BI, LP). The concordance rate was over 90% for all diagnoses. When there was disagreement, the case was discussed until a consensus diagnosis was established.[4] In the analyses the ALSbi, ALSci and ALScbi group were collapsed into a single group, i.e. intermediate cognitive phenotype.**

## WHOLE-GENOME SEQUENCING

**All eligible patients and healthy subjects were screened for *GBA* variants through whole-genome sequencing (WGS). Blood samples and genomic DNA have been processed according to standard protocols: library preparation and read sequencing have been performed as per manufacturer protocol; read alignment, variant calling and quality control have been performed according to standard protocols. WGS methods have been fully described elsewhere[7]. Annotation was performed with KGGseq v1.0 (pgmlab.top/kggseq). A repeat-primed PCR determined the presence of the *C9orf72* hexanucleotide expansion. A threshold of ≥30 repeats with the typical sawtooth pattern was considered pathological.[8]**

## STATISTICAL ANALYSIS

**Descriptive statistics were calculated for baseline demographics and cognitive testing. Survival was calculated from onset to death or censoring date (December 31st, 2019) using the KaplanMeier method. Patients with tracheostomy were coded as deceased on the date of the procedure.**

**Comparisons were performed using the log-rank test. In the first step of the analyses, only variants known to be a risk factor for cognitive decline in PD were considered. A single-variant association test was used to compare the frequency of *GBA* variants between ALS cases and healthy subjects.[9] A binomial test was used to assess the prevalence of *GBA* mutations across cognitive groups. Given the strong influence of *C9orf72* on cognitive status, we performed the binomial test with and without *C9orf72* expansion carriers to rule out its impact on the results. Linear mixed-effects models were used to test for associations between *GBA* genotype and cognitive functioning. The covariates included in this model were age at onset, sex, site of onset (classified as bulbar or spinal), the presence of bulbar signs at diagnosis, the rate of ALS Functional Rating Scale-Revised (ALS-FRS-R) decline at the time of diagnosis, and *C9orf72* status. ALS-FRS-R decline was calculated as follows: (48 - ALS-FRS-R at diagnosis) / months from disease onset to diagnosis. All statistical analyses were performed in R v.3.6.0 (http://www.r-project. org). R scripts are available on GitHub (http://github.com). To assess whether pathogenic rare variants in *GBA* contribute to cognitive decline risk in ALS, a genebased rare variants association test was also performed. This analysis provided an independent approach from the assumption that only *GBA* variants associated with PD may influence cognitive status. Only rare variants were included, and minor allele frequency threshold was set at < 5%. The analysis was performed confronting ALS patients with cognitive decline (i.e. ALSFTD and ALSCi/Cbi/Bi) vs ALS patients with normal cognitive function. The rare variant burden was assessed using the Sequence Kernel Association Test (SKAT) and the Sequence Kernel Association Test - Optimized (SKAT-O) as implemented in RVtests (http://zhanxw.github.io/rvtests/).**

## Supplementary Tables

**Supplementary Table 1. Genes and Risk Factors associated with Parkinson Disease (PD), Dementia with Lewy Bodies (DLB), frontotemporal dementia (FTD) and Amyotrophic Lateral Sclerosis (ALS) that have been implicated in Lysosomal function.**

| Gene | Disease | Functions involving lysosomes |
|---|---|---|
| *SNCA* | Early-onset PD, DLB | Recruitment of proteins to lysosomal damage sites |
| *PARK2* | Early-onset PD | Mitophagy |
| *PINK1* | Early-onset PD | Mitophagy |
| *LRRK2* | Late-onset PD | Regulattion of lysosomal pH and homesostasis, and lysosome-Golgi trafficking |
| *GBA* | Late-onset PD, DLB | Accumulation of cholesterol in lysosomes |
| *VPS35* | Late-onset PD | Endosome–Golgi trafficking; autophagy |
| *C9ORF72* | ALS/FTD | Autophagy induction and stress granule autophagy |
| *GRN* | FTD | Regulation of lysosomal pH |
| *TBK1* | ALS/FTD | Clearance of damaged mitochondria, recruitment to lysosomes by α-synuclein |
| *OPTN* | ALS/FTD | Clearance of damaged mitochondria, recruitment to lysosomes by α-synuclein |
| *CHMP2B* | ALS/FTD | Regulation of lysosomal trafficking |
| *CHCHD2* | PD | Mitochondrial quality control |
| *CHCHD10* | ALS/FTD | Mitochondrial quality control |
| *VCP* | ALS/FTD | Autophagosome maturation and clearance of damaged lysosomes |
| *SQSTM1* | ALS/FTD | It targets protein aggregates for lysosomal degradation |

**Supplementary Table 2. Results from the gene-based rare variant association test**

| Burden test | FTD | ALS-Bi / ALS-Ci / ALS-Cbi |
|---|---|---|
| **SKAT** | 0.196 | 0.0125 |
| **SKAT-O** | 0.184 | 0.00005 |

**Supplementary Table 3. List of *GBA* variants detected in our ALS series. We identified three common *GBA* polymorphisms (p.E365K, p.T408M, p.N409S), which are known risk factors for DLB**

**and cognitive impairment in PD[10–13]. Six of the other variants have not been reported to be associated with PD.**

| Nucleotide | AA | Exon | SNP id | gnomAD NFE | CADD score | ClinVar | Disease |
|---|---|---|---|---|---|---|---|
| c.C1319T | p.P440L | 10 | rs74598136 | . | 26.2 | Pathogenic | Gaucher's disease |
| c.G1279A | p.E427K | 10 | rs149171124 | 0.0003 | 23.2 | . | . |
| c.A1226G | p.N409S | 10 | rs76763715 | 0.002 | 22.7 | Pathogenic | Gaucher's disease, PD, DLB |
| c.C1223T | p.T408M | 9 | rs75548401 | 0.0091 | 22.2 | Likely benign | PD |
| c.G1093A | p.E365K | 9 | rs2230288 | 0.0121 | 17.33 | Pathogenic | PD |
| c.A928G | p.S310G | 8 | rs1057942 | 0.0000 | 14.81 | . | . |
| c.C740G | p.T247S | 7 | . | . | 22.0 | . | . |
| c.T634A | p.S212T | 7 | rs398123533 | . | 1.94 | VUS | |
| c.G626A | p.R209H | 7 | rs749416070 | 0.00004 | 17.53 | . * | Reported in PD[14] |
| c.A38G | p.K13R | 3 | rs150466109 | 0.0002 | 0.003 | Benign | |

**AA: amino acid; SNP: Single Nucleotide Polymorphism; gnomAD NFE: exome frequency in Non-Finnish European.**

*** Classified as likely pathogenic according to the following American College of Medical Genetic (ACMG) criteria[15]: PM1 moderate (Hot-spot of length 17 amino-acids has 7 non-VUS missense/in-frame variants, 6 pathogenic and 1 benign), PM2 moderate (GnomAD exomes homozygous allele count = 0 is less than 3 for recessive gene GBA), PM5 moderate (Alternative variants Arg209Pro and Arg209Cys are classified as Pathogenic), PP2 supporting (269 out of 282 non-VUS missense variants in *GBA* gene are pathogenic).**

**Supplementary Table 4. Comparison of *GBA* risk variants frequency between ALS patients and healthy controls. These variants were found in 18 ALS patients (2.40% of our cohort) and 15 healthy controls (2.22%). The variant frequencies observed in our Italian population are consistent with those reported in gnomAD. The single-variant analysis confirmed that *GBA* variants are not a risk factor for developing ALS. We did not observe homozygote cases or compound heterozygous cases for *GBA* variants. Furthermore, we did not observe mutations in ALS-related genes (*SOD1*, *TARDBP***

**and *FUS*) in *GBA* risk variants carriers. We identified one patient who carried both a *GBA* risk variant (p.N409S) and the *C9orf72* repeat expansion.**

| *GBA* variant | n. Heterozygous ALS cases (n=751) | n. Heterozygous HC (n=677) | p-value | gnomAD NFE |
|---|---|---|---|---|
| c.1093G>A (p.E365K) | 3 | 5 | 0.49 | 0.0121 |
| c.1223C>T (p.T408M) | 8 | 6 | 0.61 | 0.0091 |
| c.1226A>G (p.N409S) | 7 | 4 | 0.76 | 0.002 |

**Supplementary Table 5. Clinical and demographic characteristics of ALS patients grouped by GBA genotype. The GBA risk variants did not influence survival or the motor phenotype. Only two of the ALS patients carrying a GBA risk variant exhibited extrapyramidal signs. Both patients had slowly progressive disease, and they developed** extrapyramidal signs years **after the initial ALS diagnosis. None of the GBA risk variants carriers had a family history of Parkinson's disease or other neurodegenerative diseases. The effect of GBA mutations on cognitive but not motor phenotype could be explained by a lower susceptibility of motor neurons: it should be noted that while extrapyramidal involvement is established in GBA homozygous carriers, motor neuron involvement has only been recently reported in few cases with Gaucher's Disease and one subject with a heterozygous GBA mutation. 16,17**

| | ***GBA* carriers (n=18)** | ***GBA* non carriers (n=733)** | **P** |
|---|---|---|---|
| N (% male) | 5 (27.8%) | 404 (55.1%) | 0.04 |
| Age onset (yrs, | 65.2 (10.5) | 65.7 (9.9) | 0.84 |
| Bulbar onset, n | 4 (22.2%) | 244 (33.3%) | 0.46 |
| Bulbar signs at | 9 (50.0%) | 369 (49.2%) | 1 |
| *C9orf72* | 1 (5.6%) | 66 (9.0%) | 0.93 |
| Median survival | 2.65 (1.68-5.26) | 3.26 (1.68-5.89) | 0.43* |
| Mean ALSFRS-R | 1.31 (1.21) | 0.99 (1.74) | 0.33 |

*** Survival was calculated with Kaplan-Meier curves and significance with Log-rank test Supplementary Table 6. Cognitive phenotypes of *GBA* risk variant carriers. Due to the relatively small number of *GBA* variant carriers, we could not properly draw any conclusion on whether *GBA* variants were predominantly associated with cognitive or behavioural** impairment or with a distinct pattern of cognitive deficits. However, it should be noted that the only variant causing Gaucher's Disease of this group (p.N409S) was associated with cognitive

impairment in 7 out of the 8 carriers (87.5%): it could be postulated that more disruptive variants have a larger influence on cognitive function. Longitudinal neuropsychological studies and the application of neuroimaging techniques **might allow further clarification on this issue. As recently reported in cross-sectional study using $^{18}$F-2-fluoro-2-deoxy-D-glucose-PET,[18] the extent of metabolic brain changes in ALS patients reflects the degree of cognitive impairment, paralleling brain metabolic alterations observed in FTD over time.**[19]

## REFERENCES TO SUPPLEMENTARY MATERIALS


1. **Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*. 2000;1(5). doi:10.1080/146608200300079536**

2. **Chiò A, Mora G, Moglia C, et al. Secular Trends of Amyotrophic Lateral Sclerosis. *JAMA Neurology*. 2017;74(9). doi:10.1001/jamaneurol.2017.1387**

3. **Montuschi A, Iazzolino B, Calvo A, et al. Cognitive correlates in amyotrophic lateral sclerosis: a population-based study in Italy. *Journal of Neurology, Neurosurgery & Psychiatry*. 2015;86(2):168-173. doi:10.1136/jnnp-2013-307223**

4. **Iazzolino B, Pain D, Peotta L, et al. Validation of the revised classification of cognitive and behavioural impairment in ALS. *J Neurol Neurosurg Psychiatry*. 2019;90(7):734-739. doi:10.1136/jnnp-2018-319696**

5. **Poletti B, Solca F, Carelli L, et al. Cognitive-behavioral longitudinal assessment in ALS: the Italian Edinburgh Cognitive and Behavioral ALS screen (ECAS). *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2018;19(5-6):387-395. doi:10.1080/21678421.2018.1473443**

6. **Strong MJ, Abrahams S, Goldstein LH, et al. Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2017;18(3-4). doi:10.1080/21678421.2016.1267768**

7. **Grassano M, Calvo A, Moglia C, et al. Mutational analysis of known ALS genes in an Italian population-based cohort. *Neurology*. Published online November 2020. doi:10.1212/WNL.0000000000011209**

8. **Renton AE, Majounie E, Waite A, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*. 2011;72(2). doi:10.1016/j.neuron.2011.09.010**

9. **Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics*. 2014;95(1):5-23. doi:10.1016/j.ajhg.2014.06.009**

10. **Mata IF, Leverenz JB, Weintraub D, et al. *GBA* Variants are associated with a distinct pattern of cognitive deficits in Parkinson's disease: Cognitive Profile in *GBA* -Related PD. *Mov Disord*. 2016;31(1):95-102. doi:10.1002/mds.26359**

11. **Greuel A, Trezzi J, Glaab E, et al. *GBA* Variants in Parkinson's Disease: Clinical, Metabolomic, and Multimodal Neuroimaging Phenotypes. *Mov Disord*. 2020;35(12):2201-2210. doi:10.1002/mds.28225**

12. **Davis MY, Johnson CO, Leverenz JB, et al. Association of *GBA* Mutations and the E326K Polymorphism With Motor and Cognitive Progression in Parkinson Disease. *JAMA Neurol*. 2016;73(10):1217. doi:10.1001/jamaneurol.2016.2245**

13. **Cilia R, Tunesi S, Marotta G, et al. Survival and dementia in GBA -associated Parkinson's disease: The mutation matters. *Ann Neurol*. 2016;80(5):662-673. doi:10.1002/ana.24777**

14. **Heijer JM, Cullen VC, Quadri M, et al. A LARGE-SCALE Full *GBA1* Gene Screening in Parkinson's Disease in the Netherlands. *Mov Disord*. 2020;35(9):1667-1674. doi:10.1002/mds.28112**

15. **Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-423. doi:10.1038/gim.2015.30**

16. **Pozzilli V, Giona F, Ceccanti M, et al. A case of motor neuron involvement in Gaucher disease. *Molecular Genetics and Metabolism Reports*. 2019;21:100540. doi:10.1016/j.ymgmr.2019.100540**

17. **Oliveira LM, Rastin T, Nimmo GAM, et al. Occurrence of Amyotrophic Lateral Sclerosis in Type 1 Gaucher Disease. *Neurol Genet*. 2021;7(4):e600. doi:10.1212/NXG.0000000000000600**

18. **Canosa A, Moglia C, Manera U, et al. Metabolic brain changes across different levels of cognitive impairment in ALS: a $^{18}$ F-FDG-PET study. *J Neurol Neurosurg Psychiatry*. Published online November 23, 2020:jnnp-2020-323876. doi:10.1136/jnnp-2020-323876**

19. **Diehl-Schmid J, Grimmer T, Drzezga A, et al. Decline of cerebral glucose metabolism in frontotemporal dementia: a longitudinal 18F-FDG-PET-study. *Neurobiology of Aging*. 2007;28(1):42-50. doi:10.1016/j.neurobiolaging.2005.11.002**

# 5. Neurodegeneration-associated rare variants contribute to ALS risk and clinical phenotype

## 5.1 Abstract

**Amyotrophic Lateral Sclerosis (ALS) is a devastating neurodegenerative disease with a complex genetic architecture. While significant progress has been made in understanding its genetic underpinnings, much remains to be discovered. Rare variants have been implicated in contributing to the missing heritability and phenotypic heterogeneitY of ALS, yet their detection poses challenges due to their individual rarity and the need for large sample sizes. In this study, we employed whole-genome sequencing and rare variant association testing in a population-based cohort of ALS patients to investigate the role of rare variants in genes associated with various neurodegenerative diseases. By collectively analyzing these variants, we aimed to uncover their contribution to ALS risk and phenotypic presentation. Our findings revealed that individuals harboring rare high-impact variants in neurodegeneration-associated genes exhibited an increased risk of ALS. Furthermore, the presence of novel high-impact variants significantly heightened ALS risk, highlighting their potential as important contributors to disease susceptibility. Importantly, we observed an overlap of rare genetic variants between ALS and other neurodegenerative diseases, shedding light on shared underlying mechanisms. Additionally, the impact of these variants on ALS phenotypes was explored, revealing associations with specific motor phenotypes and an increased risk of developing frontotemporal dementia in ALS patients. This study underscores the pleiotropic effects of rare variants in neurodegenerative diseases, expanding our understanding of ALS genetics and offering potential avenues for improved diagnostic and therapeutic approaches. The identification of rare variants in genes associated with multiple neurodegenerative diseases emphasizes the importance of comprehensive genetic evaluation in ALS patients, paving the way for more personalized and precise genetic counseling and patient classification.**

## 5.2 Introduction

**Despite notable advancements in revealing the genetic underpinnings of Amyotrophic Lateral Sclerosis (ALS), there is still much to uncover about its genetic architecture [1].**

**Although rare variants are known to contribute to the missing heritability and phenotypic heterogeneity of ALS, their detection remains challenging. Rare variants with substantial impact on protein function are individually rare and require large sample sizes for reliable detection, which poses a challenge even for the largest genome-wide association studies (GWAS). However, group-**

**based approaches for rare variant association analysis offer promising avenues for identifying novel genetic associations and understanding additional disease risk [2] even with more modest sample sizes.**

**Neurodegenerative diseases are a broad spectrum of progressive neurological disorders whose etiology remains not yet completely understood. Although traditionally considered distinct entities, there has been increasing evidence suggesting clinical, pathological and genetic overlap between them [3]. It has been especially observed that many genetic loci could influence multiple neurodegenerative diseases [4], a phenomenon referred to as pleiotropy. Despite the increasing interest in leveraging pleiotropy for genomic discoveries, the clinical implications of genetic pleiotropy have not been fully explored. Although pleiotropic variants offer new possibilities for diagnosis, therapeutics, and intervention, their interpretation poses significant challenges from a clinical perspective**

**In this study, we employ whole-genome sequencing (WGS) and rare variant association testing to investigate the role of rare variants in genes associated with neurodegenerative diseases in a population-based cohort of ALS patients from the Piemonte and Valle d'Aosta Register for ALS (PARALS). By analyzing these variants collectively, our aim is to examine whether rare variants in genes associated with various neurodegenerative diseases also contribute to the risk and phenotypic presentation of ALS. Additionally, by studying variants linked to different phenotypes, we can gain insights into the associations across diverse neurodegenerative diseases and explore potential shared features often observed among them.**

**By elucidating the contribution of rare variants and exploring the pleiotropic nature of neurodegenerative diseases, this study has the potential to advance our understanding of ALS genetics and pave the way for improved diagnostic and therapeutic approaches.**

## 5.3 Methods

### Study cohort

**The study population consists of 1324 ALS patients from PARALS Register and 746 healthy controls of identical ancestry. The characteristics of PARALS have been extensively described previously [5]. To minimize the risk of patients misdiagnosis, we only included in the analysis the 948 ALS patients who met the El Escorial revised diagnostic criteria for definite or probable laboratory-supported ALS [6].**

**Clinical features (age at diagnosis, sex, site of disease onset, cognitive impairment, date of death/tracheostomy, ALSFRS-R score) were retrieved from registry data. Briefly. ALSFRS-R score mean monthly decline (ΔALSFRS-R) was calculated using the following formula: (48 – ALSFRS-R score at diagnosis)/(months from onset to diagnosis). A total of 684 patients (75.5%) underwent an extensive cognitive battery at the time of diagnosis and were classified according to the Consensus Criteria for diagnosing frontotemporal cognitive and behavioral syndromes in ALS. [7] Characteristics of cognitive testing are extensively described elsewhere. [8] All patients underwent screening for C9orf72 expansion and pathogenic mutations in SOD1, FUS, or TARDBP genes, as previously described. [9]**

### Whole-genome sequencing and variant filtering

**DNA was extracted from blood samples following established protocols. Screening for the C9orf72 intronic expansion was performed with repeat-primed PCR: pathogenic expansions were identified by repeat lengths of ≥30 units with the characteristic sawtooth pattern. Whole-genome sequencing was conducted at The American Genome Center, located at the Uniformed Services University on the Walter Reed National Military Medical Center campus in Bethesda, Maryland. Library preparation was performed using the TruSeq DNA PCR-Free High Throughput Library Prep Kit from Illumina, following the manufacturer's instructions. The sequencing was carried out on an Illumina HiSeqX10 sequencer, generating paired-end 150 base pair reads. Sequence reads were processed using the Genome Analysis Toolkit (GATK) best practices, as outlined by the Broad Institute. Variant quality control was performed using the GATK variant quality score method with default filters. The Genome Reference Consortium Human Build 38 served as the reference genome for the analysis. Variants were annotated with KGGseq version 1.0 (http://pmglab.top/kggseq/doc10/UserManual.html). Only coding nonsynonymous, and loss-of-function (frameshift insertion-deletions, splice-variants, premature stop codon, stop-loss) were retained for analysis. WGS procedures and variant filtering are further detailed elsewhere [10]**

**We defined as high-impact variants the ultra rare (Minor Allele Frequency, MAF < 0.01% in gnomAD; https://gnomad.broadinstitute.org/) and novel (absent from gnomAD) variants that either had loss-of-function consequences or were missense variants with (a) a ClinVar score of 4 or 5 or (b) a Combined Annotation-Dependent Depletion (CADD) score >20. The analysis included high-impact variants in genes associated with a clinical neurodegenerative phenotype in Ensembl (https://www.ensembl.org/).**

### Statistical analysis

**The burden of high-impact rare variants in case-control analysis was assessed with a logistic regressions model. Linear and logistic regression models were applied to test whether carriers of neurodegenerative variants manifested different phenotypic features.**

**We also repeated the analysis by grouping variants according to their respective neurodegenerative disease family: Alzheimer's Disease (AD), Frontotemporal Dementia (FTD), Other Dementia, Charcot-Marie-Tooth and motor neuropathy (CMT), Hereditary Spastic Paresis (HSP), Parkinson's Disease (PD) amd Spinal Muscolar Atrophy (SMA) (Table 1).**

**All statistical tests were two-sided and the level of statistical significance was set at 0.05. All analyses were carried out using R version 4.1.2.**

### Replication cohort and analysis

**For replication, we downloaded from Project MinE [11] (4366 ALS cases and 1832 controls) the summary data for ultra-rare (MAF < 0.01%), high-impact (loss-of-functions variants or missense variants with CADD > 20) variants in the same neurodegeneration-associated genes identified in our cohort. Burden testing on the publicly available replication cohort was performed as implemented in the Test Rare Variants With Public Data (TRAPD) software package version 1.0.14 [12]. The statistical significance threshold was set at a 1-tailed P value less than .05. Project MinE data are available at http://databrowser.projectmine.com/. TRAPD is available at (https://github.com/mhguo1/TRAPD).**

## 5.4 Results

### Study cohort description

**A total of 1694 participants were included in the analysis, comprising 948 individuals diagnosed with ALS and 746 healthy controls. Among the ALS patients, 496 (52.3%) were male, with a median age at diagnosis of 67.9 years (IQR: 60.3 – 74.4). The distribution of phenotypes among ALS patients was as follows: 308 (32.5%) had a bulbar phenotype, 301 (31.8%) had a classic phenotype, 58 (6.1%) had a flail arm phenotype, 180 (19%) had a flail leg phenotype, 85 (9.0%) had a pyramidal phenotype, and 16 (1.7%) had a respiratory phenotype. Cognitive impairment was observed in 548 (57.8%) patients, with 124 (13.1%) of them presenting frontotemporal dementia (FTD). The baseline demographic and disease characteristics of the included ALS patients are summarized in Table 2.**

### Rare variants in neurodegeneration genes and ALS Risk

**Overall, individuals harboring at least one rare high-impact variant in genes associated with neurodegenerative diseases exhibited a higher risk of ALS, although borderline for statistical significance (OR = 1.28, 95%CI = 0.99 - 1.66, p = 0.059). However, when considering only novel high-impact variants, a significantly increased risk was observed in individuals with at least one such mutation compared to those without (OR = 1.79, 95%CI = 1.28 - 2.53, p = 0.001). The analysis performed after excluding the ALS patients who carried a known high-penetrance mutations in ALS-associated genes confirmed that the presence of high-impact ultra-rare or novel variants related to neurodegenerative mechanisms significantly increased the risk of ALS (OR = 1.31, 95%CI = 1.01 - 1.71, p = 0.046, and OR = 1.75, 95%CI = 1.24 - 2.50, p = 0.002, respectively). We also assessed variants in Project MinE and observed a significant enrichment of high impact ultra-rare (OR = 1.31, 95%CI = 1.01 - 1.71, p = 0.046) and novel (OR = 1.31, 95%CI = 1.01 - 1.71, p = 0.046) variants in neurodegeneration genes in ALS cases compared to healthy controls.**

### Rare genetic variants overlap between ALS and other neurodegenerative diseases

**The presence of a novel mutation in Parkinson-associated genes more than tripled the risk of ALS (OR = 3.58, 95%CI = 1.43 - 10.85, p = 0.011). Similarly, carrying a novel mutation in CMT-associated genes was associated with a more than threefold increased risk of ALS (OR = 3.28, 95%CI = 1.62 - 7.36, p = 0.002). Additional associations between rare variants in disease-associated genes are reported in Table 3.**

### Effect of rare neurodegenerative variants on ALS phenotype

**No significant associations were found between the presence of at least one mutation or novel mutation in neurodegeneration-associated genes and age of disease onset, survival, diagnosis delay, initial diagnosis score, and monthly loss in ALSFRS-R score until diagnosis (Table 4). Interestingly, a significant increase in the risk of a flail arm motor phenotype was observed in patients with at least one mutation across all considered genes (OR = 2.06, 95%CI = 1.09 - 3.90, p = 0.026), in Parkinson-associated genes (OR = 3.28, 95%CI = 1.28 - 8.38, p = 0.013), and in HSP-associated genes (OR = 2.86, 95%CI = 1.24). Additionally, harboring variants in one of these neurodegeneration genes appeared to increase the increase of developing an FTD-ALS phenotype (OR: 1.82, 95%CI: 1.02 - 3.27, p = 0.044). Curiously, however, this effect was not significant if only novel variants were considered (OR: 1.54, 95%CI: 0.83 - 2.87) (Table 5).**

## 5.5 Discussion

**The emerging evidence on the pleiotropic effects of rare variants in neurodegenerative diseases highlights their role as modifiers of disease risk and phenotype, including in ALS [2,13].**

**Our study adds to this understanding by demonstrating that rare variants in neurodegeneration-associated genes significantly increase the risk of ALS. These findings suggest that the missing heritability observed in ALS cases may be attributed, at least in part, to these rare variants. Notably, our results emphasize the importance of evaluating rare variants in the genetic background, even in sporadic ALS cases without established causative mutations.**

**When we analyzed mutations grouped by corresponding neurodegenerative diseases, we observed an elevated risk in individuals carrying at least one variant in genes associated with Parkinson's disease (PD) or motor neuropathy. This observation aligns with previous studies that have reported an enrichment of PD-related variants in ALS-FTD cohorts [14].**

**Furthermore, our investigation into the impact of mutational load in neurodegeneration-associated genes on ALS phenotype revealed that rare variants in these genes increase the risk of developing full-blown frontotemporal dementia (FTD) in ALS patients. Similar findings were observed in the context of Parkinson's disease, where a higher burden of rare variants in genes associated with cognitive impairment was observed in PD patients with dementia compared to those with motor-only symptoms [15]. One possible explanation for this effect is that by increasing the overall 'neurodegenerative burden', the presence of rare variants could tip the balance toward a more complex neurodegenerative phenotype.**

**A particularly intriguing finding was the association between a flail arm phenotype and a higher burden of rare variants in neurodegenerative disease genes, specifically those associated with PD and hereditary spastic paraplegia (HSP). The flail-arm phenotype predominantly involves lower motor neuron degeneration, making it highly unlikely that misclassification or misdiagnosis influenced our results, as neither PD nor HSP present with lower motor neuron symptoms. This also highlights the genetic heterogeneity underlying ALS phenotypes: further elucidating the genotypic correlates of clinical phenotypes could greatly enhance patient classification and our understanding of the biological basis of the disease.**

**While our analyses provide valuable insights into the complex genetic architecture of motor neuron degeneration, it is important to note that we cannot conclude that all observed rare variants are directly driving ALS pathogenesis, nor can we estimate the disease risk associated with each specific variant. However, based on our data we could reasonably assume that these rare variants influence the risk of developing ALS. Therefore, if future studies replicate our findings, it would be prudent for genetic counseling to encompass not only the analysis of ALS-specific genes but also a broader evaluation of a spectrum of neurodegenerative disease genes.**

**In summary, our study sheds light on the intricate genetic landscape of motor neuron degeneration. By uncovering the impact of rare variants and highlighting their role as phenotypic modifiers in ALS, we provide a foundation for considering a more comprehensive evaluation of neurodegenerative disease genes in genetic counseling. Such an approach has the potential to improve disease risk assessment, patient classification, and our overall understanding of the underlying biology of ALS.**

## 5.6 References


1. Goutman SA, Hardiman O, Al-Chalabi A, Chió A, Savelieff MG, Kiernan MC, Feldman EL. Emerging insights into the complex genetics and pathophysiology of amyotrophic lateral sclerosis. Lancet Neurol. 2022 May;21(5):465-479. doi: 10.1016/S1474-4422(21)00414-2. Epub 2022 Mar 22. PMID: 35334234; PMCID: PMC9513754

2. van Rheenen W, van der Spek RAA, Bakker MK, et al. Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. Nat Genet. 2021 Dec;53(12):1636-1648. doi: 10.1038/s41588-021-00973-1. Epub 2021 Dec 6. Erratum in: Nat Genet. 2022 Mar;54(3):361. PMID: 34873335; PMCID: PMC8648564

3. Ahmed RM, Devenney EM, Irish M, et al. Neuronal network disintegration: common pathways linking neurodegenerative diseases. J Neurol Neurosurg Psychiatry. 2016 Nov;87(11):1234-1241. doi: 10.1136/jnnp-2014-308350. Epub 2016 May 12. PMID: 27172939; PMCID: PMC5099318.

4. Qiao J, Wang T, Shao Z, Zhu Y, Zhang M, Huang S, Zeng P. Genetic correlation and gene-based pleiotropy analysis for four major neurodegenerative diseases with summary statistics. Neurobiol Aging. 2023 Apr;124:117-128. doi: 10.1016/j.neurobiolaging.2022.12.012. Epub 2022 Dec 31. PMID: 36740554.

5. Chiò A, Mora G, Moglia C, Manera U, Canosa A, Cammarosano S, Ilardi A, Bertuzzo D, Bersano E, Cugnasco P, Grassano M, Pisano F, Mazzini L, Calvo A; Piemonte and Valle d'Aosta Register for ALS (PARALS). Secular Trends of Amyotrophic Lateral Sclerosis: The Piemonte and Valle d'Aosta Register. JAMA Neurol. 2017 Sep 1;74(9):1097-1104. doi: 10.1001/jamaneurol.2017.1387. PMID: 28692730; PMCID: PMC5710181.

6. Brooks BR, Miller RG, Swash M, Munsat TL; World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord. 2000 Dec;1(5):293-9. doi: 10.1080/146608200300079536. PMID: 11464847.

7. Strong MJ, Abrahams S, Goldstein LH, et al. Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. Amyotroph Lateral Scler Frontotemporal Degener. 2017 May;18(3-4):153-174. doi: 10.1080/21678421.2016.1267768. Epub 2017 Jan 5. PMID: 28054827; PMCID: PMC7409990.

8. Chiò A, Moglia C, Canosa A, Manera U, Vasta R, Brunetti M, Barberis M, Corrado L, D'Alfonso S, Bersano E, Sarnelli MF, Solara V, Zucchetti JP, Peotta L, Iazzolino B, Mazzini L, Mora G, Calvo A. Cognitive impairment across ALS clinical stages in a population-based cohort. Neurology. 2019 Sep 3;93(10):e984-e994. doi: 10.1212/WNL.0000000000008063. Epub 2019 Aug 13. PMID: 31409738; PMCID: PMC6745732.

9. Grassano M, Calvo A, Moglia C, et al. Mutational Analysis of Known ALS Genes in an Italian Population-Based Cohort. Neurology. 2021 Jan 26;96(4):e600-e609. doi: 10.1212/WNL.0000000000011209. Epub 2020 Nov 18. PMID: 33208543; PMCID: PMC7905787.

10. Grassano M, Calvo A, Moglia C, et al. Systematic evaluation of genetic mutations in ALS: a population-based study. J Neurol Neurosurg Psychiatry. 2022 Jul 27;93(11):1190–3. doi: 10.1136/jnnp-2022-328931. Epub ahead of print. PMID: 35896380; PMCID: PMC9606529.

11. Project MinE ALS Sequencing Consortium. Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. Eur J Hum Genet. 2018 Oct;26(10):1537-1546. doi: 10.1038/s41431-018-0177-4. Epub 2018 Jun 28. PMID: 29955173; PMCID: PMC6138692.

12. 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. 2018 Oct 4;103(4):522-534. doi: 10.1016/j.ajhg.2018.08.016. Epub 2018 Sep 27. PMID: 30269813; PMCID: PMC6174288.

13. Canosa A, Grassano M, Moglia C, et al. GBA variants influence cognitive status in amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry. 2022 Apr;93(4):453-455. doi: 10.1136/jnnp-2021-327426. Epub 2021 Sep 28. PMID: 34583942; PMCID: PMC8921570.

14. Dilliott AA, Abdelhady A, Sunderland KM, et al. Contribution of rare variant associations to neurodegenerative disease presentation. NPJ Genom Med. 2021 Sep 28;6(1):80. doi: 10.1038/s41525-021-00243-3. PMID: 34584092; PMCID: PMC8478934.

15. Schulte EC, Fukumori A, Mollenhauer B,et al. Rare variants in β-Amyloid precursor protein (APP) and Parkinson's disease. Eur J Hum Genet. 2015 Oct;23(10):1328-33. doi: 10.1038/ejhg.2014.300. Epub 2015 Jan 21. PMID: 25604855; PMCID: PMC4592093.

## 5.7 Tables And Figures

### Table 1

**List of neurodegeneration-related genes where variants were identified**

| |
|---|
| **ALZHEIMER DISEASE (AD)** |
| APP, APOE, PRX, PSEN1 |
| **ATAXIA** |
| AFG3L2, ATP1A3, DNMT1, FBXO7, FLRT1, MAPT, MORC2, PMP22, PNKP, RTN2, SERPINI1, UBAP1, UCHL1 |
| **CMT (NEUROPATHY)** |
| AFG3L2, AIFM1, ATL1, BSCL2, DDHD2, DNAJC6, DSTYK, HACE1, KIDINS220, KIF1A, MAPT, MARS2, MFN2, PLP1, PMP22, PNPLA6, POLG, PSEN1, SACS, SYNJ1, TECPR2, TFG, TRIP4 |
| **OTHER DEMENTIA** |
| APOE, GBA2, MAG, POLG, PSEN1, RTN2, SLC33A1, SNCA, |
| **FRONTOTEMPORAL DEMENTIA (FTD)** |
| CPT1C, PRKAR1B |
| **HEREDITARY SPASTIC PARAPARESIS (HSP)** |
| ALDH18A1, AMPD2, AP4B1, AP4E1, AP4M1, AP4S1, AP5Z1, ARL6IP1, ATL1, ATP13A2, B4GALNT1, BSCL2, C19orf12, CAPN1, CHCHD2, CPT1C, CYLD, CYP7B1, DDHD1, DDHD2, DNMT1, DSTYK, EGR2, ENTPD1, ERLIN1, ERLIN2, FA2H, FBXO7, GBA2, HACE1, HK1, HSPD1, IGHMBP2, KIF1A, KIF1C, KLC2, LRRK2, NAIP, NIPA1, NT5C2, PLA2G6, PLEKHG5, PLP1, PNPLA6, PSEN2, REEP1, SH3TC2, SLC33A1, SNCB, SPART, TECPR2, TFG, UBA1, UBAP1, VPS35, VPS37A, WASHC5, ZFYVE26 |

| PARKINSON'S DISEASE (PD) |
|---|
| ATP1A3, ATP6AP2, DNAJB2, DNAJC13, DNAJC6, FARS2, L1CAM, LAS1L, LRRK2, MAG, MAPT, PCYT2, PINK1, SNCA, SPG7, SYNJ1, UCHL1, VAMP1, VPS13C |
| **SPINAL MUSCULAR ATROPHY (SMA)** |
| AR, ASAH1, ASCC1, ATP7A, DDHD2, IBA57, IGHMBP2, KLC2, MPZ, MTPAP, NAIP, PLA2G6, TREM2, TRIP4, TRPV4 |

**Table 2.**

**ALS sample characteristics. FTD = Frontotemporal dementia**

| | **Overall** |
|---|---|
| **Sex:** Male | 496 (52.3%) |
| **Age at onset (years)** | |
| Median [Q1,Q3] | 66.9 [59.2,73.4] |
| **Age at diagnosis (years)** | |
| Median [Q1,Q3] | 67.9 [60.3,74.4] |
| **Diagnosis delay (months)** | |
| Median [Q1,Q3] | 9.04 [5.10,14.0] |
| **ALSFRS-R at diagnosis** | |
| Median [Q1,Q3] | 43.0 [38.0,45.0] |
| **Monthly ΔALSFRS-R until diagnosis** | |
| Median [Q1,Q3] | 0.591 [0.286,1.22] |
| **Survival (months)** | |
| Median [Q1,Q3] | 22.0 [11.0,40.0] |
| **Motor phenotype** | |
| Bulbar | 308 (32.5%) |
| Classical | 301 (31.8%) |
| Flail Arm | 58 (6.1%) |
| Flail Leg | 180 (19.0%) |

| | |
|---|---|
| Pyramidal | 85 (9.0%) |
| Respiratory | 16 (1.7%) |
| **Cognitive phenotype** | |
| Normal | 276 (29.1%) |
| Cognitive impairment | 548 (57.8%) |
| FTD | 124 (13.1%) |

**Table 3.**

**Risk of ALS deriving from carrying at least one variant in neurodegenerative genes. Univariate logistic regression models' effects are reported as Odds Ratios, 95% Confidence Intervals (CI), and corresponding p-values (p). CMT = Charcot-Marie-Tooth, HSP = Hereditary Spastic Paraplegia, FTD = Frontotemporal Dementia.**

| | | | ALS | | |
|---|---|---|---|---|---|
| *Sample* | *Disease* | *Variant type* | *Odds Ratio* | *95%CI* | *p* |
| All subjects | | All | 1.28 | 0.99 – 1.66 | 0.059 |
| | | Novel | **1.79** | 1.28 – 2.53 | **0.001** |
| Excluding patients with | | All | **1.31** | 1.01 – 1.71 | **0.046** |
| | | Novel | **1.75** | 1.24 – 2.50 | **0.002** |
| | Parkinson | All | **1.76** | 1.04 – 3.07 | **0.041** |
| | | Novel | **3.58** | 1.43 – 10.85 | **0.011** |
| | CMT | All | **2.34** | 1.41 – 4.00 | **0.001** |
| | | Novel | **3.28** | 1.62 – 7.36 | **0.002** |
| | HSP | All | 1.12 | 0.75 – 1.69 | 0.587 |
| | | Novel | 1.14 | 0.68 – 1.93 | 0.624 |
| | Dementia | All | 0.33 | 0.09 – 0.98 | 0.061 |
| | | Novel | 0.62 | 0.08 – 3.73 | 0.597 |
| | Alzheimer | All | 0.15 | 0.01 – 0.90 | 0.083 |
| | | Novel | | | NC |
| | FTD | All | 0.46 | 0.06 – 2.37 | 0.373 |
| | | Novel | 0.93 | 0.04 – 23.46 | 0.957 |

Table 4.

Phenotype analysis of ALS patients carrying ≥ 1 variant in genes related to neurodegeneration: results from linear regression analysis

| | ≥ 1 **Variant** | | ≥ 1 **Novel** | | **HSP** | |
|---|---|---|---|---|---|---|
| *Phenotypes* | *Coefficie* | *p* | *Coeffic* | *p* | *Coeffic* | *p* |
| Age of disease | -1.11 | 0.253 | -0.60 | 0.609 | 0.43 | 0.784 |
| Survival, mos | -1.36 | 0.533 | -3.26 | 0.214 | -1.90 | 0.587 |
| Diagnosis | -2.34 | 0.505 | -1.92 | 0.099 | **-3.67** | **0.018** |
| Score at | -0.02 | 0.968 | -0.20 | 0.764 | 0.01 | 0.991 |
| Monthly | 0.23 | 0.158 | 0.33 | 0.096 | 0.03 | 0.904 |

## Table 5.

**Phenotype analysis of ALS patients carrying ≥ 1 variant in genes related to neurodegeneration: results from logistic regression analysis**

| | **All** | | **Novel** | | **Parkinson** | | **HSP** | | **CMT** | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Phenotypes* | *Odds Ratio* | *p* | *Odds Ratio* | *p* | *Odds Ratio* | *p* | *Odds Ratio* | *p* | *Odds Ratio* | *p* |
| Motor (reference : Classic) | | | | | | | | | | |
| Bulbar | 0.83 (0.54 - 1.28) | 0.390 | 0.83 (0.49 - | 0.495 | 0.89 (0.40 - 1.98) | 0.767 | 0.75 (0.37 - | 0.416 | 0.58 (0.28 - | 0.132 |
| Flail Arm | **2.06** (1.09 - 3.90) | **0.026** | **2.54** (1.26 - | **0.009** | **3.28** (1.28 - 8.38) | **0.013** | **2.86** (1.24 - | **0.013** | 1.15 (0.41 - | 0.787 |
| Flail Leg | 0.67 (0.39 - 1.14) | 0.138 | 0.59 (0.30 - | 0.119 | 0.49 (0.16 - 1.52) | 0.215 | 0.49 (0.19 - | 0.140 | 0.6 (0.26 - | 0.229 |
| Pyramidal | 0.61 (0.30 - 1.25) | 0.176 | 0.64 (0.27 - | 0.315 | 0.24 (0.03 - 1.89) | 0.177 | 0.33 (0.07 - | 0.141 | 0.77 (0.28 - | 0.608 |
| Cognitive (reference :normal) | | | | | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Altered | 1.33 (0.88 - 2.00) | 0.174 | 1.17 (0.73 - | 0.516 | 1.32 (0.63 - 2.79) | 0.459 | 1.26 (0.68 - | 0.464 | 1.21 (0.62 - | 0.56 9 |
| FTD | **1.82** (1.02 - 3.27) | **0.044** | 1.54 (0.83 - | 0.174 | 0.66 (0.18 - 2.44) | 0.533 | 1.36 (0.58 - | 0.479 | 2.17 (0.96 - | 0.06 3 |

# 6.1 Clinical and Metabolic Signature of UNC13A rs12608932 Variant in Amyotrophic Lateral Sclerosis

**As published in *Calvo, A., Canosa, A., Moglia, C., Manera, U., Grassano, M., Vasta, R., ... & Chio, A. (2022). Clinical and Metabolic Signature of UNC13A rs12608932 Variant in Amyotrophic Lateral Sclerosis. Neurology Genetics, 8(6).***

**AUTHORS**

**Andrea Calvo[1,2], MD, PhD,* Antonio Canosa[1,2], MD, PhD,* Cristina Moglia[1,2], MD, PhD,* Umberto Manera[1,2], MD, Maurizio Grassano[1], MD, Rosario Vasta[1], MD, Francesca Palumbo[1], MD, Paolo Cugnasco[1], Salvatore Gallone[,2], MD, Maura Brunetti[,1], BSc, Fabiola De Marchi[4], MD, PhD, Vincenzo Arena[5], MD, Marco Pagani[6], MD, Clifton Dalgard[7,8], PhD, Sonja W. Scholz[9,10], MD, PhD, Ruth Chia[11], PhD, Lucia Corrado[12], PhD, Sandra Dalfonso[12], PhD, Letizia Mazzini[4], MD,* Bryan J. Traynor[11,13], MD, PhD,* and Adriano Chio[1,2,14], MD, PhD, FAAN ***

## Author Affiliations

**1 ''Rita Levi Montalcini'' Department of Neuroscience, University of Turin, Turin, Italy**

**2 Neurologia 1, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy**

**3 Laboratory of Genetics, Department of Clinical Pathology, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy**

**4 ALS Center, Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy**

**5 Positron Emission Tomography Centre AFFIDEA-IRMET S.p.A**

**6 Department of Medical Radiation Physics and Nuclear Medicine (M.P.), Karolinska University Hospital, Stockholm, Sweden**

**7 The American Genome Center, Collaborative Health Initiative Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA**

**8 Department of Anatomy, Physiology & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA**

**9 Neurodegenerative Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA**

**10 Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA**

**11 Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, MD 20892, USA**

**12 Department of Health Sciences Interdisciplinary Research Center of Autoimmune Diseases, D "Amedeo Avogadro" University of Eastern Piedmont, Novara, Italy;**

**13 Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD, 21224, USA**

**14 Institute of Cognitive Sciences and Technologies, National Council of Research, Rome, Italy**

## 6.1 Abstract

### Background and Objectives

**To characterize the clinical and cognitive behavioral phenotype and brain 18F-2-fluoro-2-deoxy-d-glucose-PET (18F-FDG-PET) metabolism of patients with amyotrophic lateral sclerosis (ALS) carrying the rs12608932 variant of the UNC13A gene.**

### Methods

**The study population included 1,409 patients with ALS without C9orf72, SOD1, TARDBP, and FUS mutations identified through a prospective epidemiologic ALS register. Control participants included 1,012 geographically matched, age-matched, and sex-matched participants. Clinical and cognitive differences between patients carrying the C/C rs12608932 genotype and those carrying the A/A + A/C genotype were assessed. A subset of patients underwent 18F-FDG-PET.**

### Results

**The C/C genotype was associated with an increased risk of ALS (odds ratio: 1.54, 95% confidence interval 1.18–2.01, p = 0.001). Patients with the C/C genotype were older, had more frequent bulbar onset, and manifested a higher rate of weight loss. In addition, they showed significantly reduced performance in the letter fluency test, fluency domain of Edinburgh Cognitive and Behavioural ALS**

**Screen (ECAS) and story-based empathy task (reflecting social cognition). Patients with the C/C genotype had a shorter survival (median survival time, C/C 2.25 years, interquartile range [IQR] 1.33–3.92; A/A + C/C: 2.90 years, IQR 1.74–5.41; p = 0.0001). In Cox multivariable analysis, C/C genotype resulted to be an independent prognostic factor. Finally, patients with a C/C genotype had a specific pattern of hypometabolism on brain 18F-FDG-PET extending to frontal and precentral areas of the right hemisphere.**

**Discussion**

**C/C rs12608932 genotype of UNC13A is associated with a specific motor and cognitive/behavioral phenotype, which reflects on 18F-FDG-PET findings. Our observations highlight the importance of adding the rs12608932 variant in UNC13A to the ALS genetic panel to refine the individual prognostic prediction and reduce heterogeneity in clinical trials.**

## 6.2 Introduction

**Amyotrophic lateral sclerosis (ALS) is a degenerative disorder of the CNS, characterized by the progressive involvement of several neuronal systems, namely the lower and upper motor neurons and neurons of the frontotemporal cortices. Approximately 10% of people with ALS have a familial history of ALS or frontotemporal dementia (FTD),1 while the disease appears to be sporadic in the remaining participants. Moreover, at least 25% of patients with ALS carry a genetic mutation, which is the likely cause of the disease.2**

**ALS is a clinically heterogeneous disease with diverse clinical presentations and outcomes.1 The modifiers of the ALS phenotype are only partially known; they include age, sex, and genetics3 and likely environmental and lifestyle factors, such as exercise and elevated low density lipoproteins cholesterol.4-6 The common variant rs12608932, located within an intron of UNC13A gene on chromosome 19p13.3, may influence susceptibility to disease and survival among patients with ALS7-9 and may modulate response to pharmacologic treatment.10**

**While the biological mechanisms underlying the rs12608932 variant effect remain unclear, the phenotypic and MRI characteristics of patients carrying this variant have been recently reported in a cohort of patients with ALS of Dutch ancestry.11 In this study, we used a large population-based cohort of Italian patients to characterize the clinical and behavioral/cognitive phenotype of patients**

**with ALS carrying the rs12608932 variant of the UNC13A gene. We also assessed their metabolic correlates using brain 18F-2-fluoro-2-deoxy-d-glucose-PET (18F-FDG-PET).**

## 6.3 Methods

**The rs12608932 variant of the UNC13 gene was assessed in 1,584 patients with ALS identified through the Piemonte and Valle d'Aosta Register for ALS, a prospective population-based register active in Northern Italy since 1995. The characteristics of the register have been reported elsewhere.12 For this study, we included cases with ALS diagnosed between 2007 and 2018. We excluded 175 patients with C9orf72 (92), SOD1 (30), TARDBP (20), and FUS (12) mutations. Patients met the revised El Escorial diagnostic criteria for definite, probable, and probable laboratory-supported ALS.13 A total of 265 cases have already been reported in a previous study.8**

**Amyotrophic lateral sclerosis functional rating scale - revised (ALSFRS-R) mean monthly decline (ΔALSFRS-R) was calculated using the following formula: (48 − ALSFRS-R score at diagnosis)/(months from onset to diagnosis).14 Similarly, weight mean monthly decline (ΔWeight) was calculated as (Weight at diagnosis − healthy body weight)/months from onset to diagnosis.**

**A total of 565 patients underwent an extensive cognitive battery administered within 2 months of their diagnosis, encompassing executive function, memory, visuospatial function, language, and social cognition domains.15 The battery included the following neuropsychological tests: Letter Fluency Test (FAS); Category Fluency Test; Frontal Assessment Battery; Trail Making Test A, B, and B-A; Rey-Osterrieth Complex Figure Test: immediate recall (IR) and differed recall (DR); Rey Auditory Verbal Learning Test: IR and DR; Babcock Story Recall Test: IR and DR; Digit Span Forward and Backward; Raven Colored Progressive Matrices (47); Mini-Mental State Examination; and the Story-Based Empathy Task (SET). The raw scores of all tests were age-corrected, sex-corrected, and education-corrected using the recent Italian normative.16 Patients were classified into 5 categories according to the revised ALS-FTD Consensus Criteria.17**

**Neurobehavioral dysfunction was determined with the Frontal Systems Behavior Scale (FrSBe), using the family form evaluated by a close relative/caregiver (scores: normal ≤59, borderline 60–64; and pathologic ≥65). For this study, we considered the change in points for each of the 3 domains of FrSBe (apathy, disinhibition, and executive) between the premorbid and postillness scores. If patients had scores reflecting a frontal systems abnormality both in the premorbid and postillness forms, they were considered pathologic only if there was an increase of ≥10 points between the premorbid and the disease scores.16 Anxiety and depression were assessed with the Hospital**

**Anxiety and Depression Scale. The test item "I feel slowed down" was discussed with patients to have them not refer to physical disability.18**

### Controls

**Control participants consisted of 1,012 geographically matched, age-matched, and sex-matched Italian participants, identified through patients' general practitioners. Of them, 129 were tested with the cognitive battery.**

### Genetic Analysis

**All patients included in the study were tested for SOD1 (all exons), TARDBP (exon 6 only), FUS (exons 14 and 15), and the C9orf72 repeat expansion using the method described elsewhere.19**

**The UNC13A rs12608932 variant was genotyped by whole-genome sequencing on an Illumina HiSeqX10 sequencer in 1,319 cases and 771 controls. The remaining 265 cases and 241 controls were genotyped on Infinium HumanHap550 bead chips (Illumina, Inc., San Diego, CA) as per the manufacturer's specification. Sequencing methods have been detailed elsewhere.2,8 These data are publicly available on the dbGAP repository (ncbi.nlm.nih.gov/gap; phs001963.v1.p1 and phs000101.v3.p1).**

### PET Acquisition

**Before brain 18F-FDG-PET, patients fasted for at least 6 hours, and blood glucose was less than 7.2 mmol/L in all cases. After a 20-minute rest, approximately 185 MBq of 18F-FDG was injected. Data acquisition started 60 minutes after the injection. PET/CT scans were performed on a Discovery ST-E System (General Electric). Brain CT (slice thickness of 3.75 mm, 140 kV, 60–80 mAs) and PET scan were sequentially acquired, and the former was used for the attenuation correction of PET data. The PET images were reconstructed with 4 iterations and 28 subsets with an initial voxel size of 2.34 × 2.34 × 2.00 mm, and data were collected in 128 × 128 matrices.**

### PET Statistical Analysis

**SPM12 implemented in Matlab R2018b (MathWorks, Natick, MA) was used for image spatial normalization to a customized brain 18F-FDG-PET template.20 Normalization with a subcortical reference region was not performed because all brain regions have been demonstrated to be potentially affected in ALS. Intensity normalization was performed using the 0.8 default Statistical Parametric Mapping value of gray matter threshold, and images were smoothed with a 10-mm filter and submitted to statistical analysis. Patients carrying the A/A or the A/C genotypes were merged and compared with the group with the C/C genotype, using the 2-sample t test model of SPM12. To**

**achieve a better characterization of patients' brain metabolic state, we compared each patient group (AA + AC and CC) with 40 healthy controls. We considered eligible as controls participants referred to the PET Center for suspected lung cancer, but for whom oncologic diseases were excluded, and brain PET scan, medical history for neurologic disorders, and neurologic examination resulted negative. Because the sample size of the 2 patient groups was unbalanced (AA + AC = 277, CC = 40), for the comparison with controls, 40 of 277 participants were randomly selected from the AA + AC group and then submitted to the analysis. In the comparison of each patient group with healthy controls (HCs) age at PET and sex were used as covariates. The height threshold was set at $p < 0.001$ ($p < 0.05$ family-wise error (FWE)-corrected at cluster level).**

**Age at PET, sex, site of onset (spinal/bulbar), and King stage at PET were used as covariates to compare the 2 groups of patients, and the height threshold was set at $p < 0.001$ ($p < 0.05$ FWE-corrected at cluster level). Only clusters containing >125 contiguous voxels were considered significant. Brodmann areas were identified at a 0–2 mm range from the Talairach coordinates of the SPM output isocenters corrected by Talairach Client (talairach.org/index.html).**

### Statistical Methods

**UNC13A rs12608932 was tested for association with the risk of developing disease under the following genetic models: additive (A/A vs A/C vs C/C genotypes), dominant (A/A vs A/C + CC), and recessive (A/A + A/C vs C/C). A p value < 0.05 was considered statistically significant. The genetic model underlying the association of this variant with disease risk was assessed using the max-statistic, which selects the largest test statistic from the dominant, recessive, and additive models. In addition, we calculated the Bayesian information criteria (BIC) and performed a likelihood ratio test on the residual deviance to compare the goodness of fit of the 3 genetic models. These analyses were performed using R (version 4.0). The Mann-Whitney U test was used for comparisons of continuous variables. Survival was calculated with Kaplan-Meier curves and compared using the log-rank test, setting the onset date as day 0 and the date of death or tracheostomy as the primary end point. The last day of follow-up for censored cases was December 31, 2020. None of the patients were lost to follow-up. Multivariable analysis for survival was performed with the Cox proportional hazards model (stepwise backward) with a retention criterion of $p < 0.1$. A p value < 0.05 was considered significant. The following variables were included in the model: age (continuous), diagnostic delay (continuous), sex (male vs female), site of onset (bulbar vs spinal), King staging, ΔALSFRS-R (continuous), ΔFVC (forced vital capacity)% (continuous), and ∆Weight (continuous). The cognitive tests were analyzed using the Mann-Whitney U test on age-matched, sex-matched, and**

**education-corrected scores. A p value < 0.05 (2-tailed) was considered significant. These analyses were conducted using the SPSS 26.0 statistical package (SPSS, Chicago, IL).**

### Standard Protocol Approvals, Registrations, and Patient Consents

**The study was approved by the Ethical Committees of the 2 regional ALS Expert Centers (Comitato Etico Azienda Ospedaliero-Universitaria Città della Salute e della Scienza, Torino, and Comitato Etico Azienda Ospedaliero-Universitaria Maggiore della Carità, Novara). Patients and controls provided written informed consent before enrollment. The databases were anonymized according to Italian law for the protection of privacy.**

## 6.4 Results

**The study population included 1,409 patients ([780, 55.4%] males, median onset age at onset of 67.8, interquartile range [IQR] 59.8–73.9) and 1,012 controls ([521, 51.5%] males, median age of 65.0, IQR 57–72). The patient and control cohorts did not differ for the main demographic variables (eTables 1 and 2, links.lww.com/NXG/A551).**

**Of the 1,409 patients included in the study, 662 (47.0%) carried the A/A genotype, 563 (40.0%) the A/C genotype, and 184 (13.1%) the C/C genotype. Allele frequencies were 0.67 (A) and 0.33 (C). The corresponding frequencies among controls were 509 for the A/A genotype (50.3%), 413 for the A/C genotype (40.8%), and 90 for the C/C genotype (8.9%). Allele frequencies in controls were 0.71 (A) and 0.29 (C). Allele frequency was significantly different among cases and controls (p = 0.001). Allele frequency did not deviate from the Hardy-Weinberg equilibrium among controls (p = 0.92), while the slight deviation observed among cases (p = 0.051) likely reflects the increased risk associated with the C allele.**

**Patients with the C/C genotype had an increased risk of ALS (odds ratio [OR] = 1.54, 95% confidence interval [CI] = 1.18–2.01, p = 0.001). Similarly, the C allele also increased the risk (OR = 1.19, 95% CI = 1.05–1.35, p = 0.006). UNC13A rs12608932 was associated with an increased risk of developing ALS in additive (p = 7.8 × 10−6), dominant (p = 0.0071), and recessive (p = 3.8 × 10−6) models. Post hoc analysis demonstrated that the increased disease risk was indeed most prominent among C/C carriers (OR 1.76, 97.5% CI 1.40–2.21, p = 1.5 × 10–6) rather than A/C carriers (OR 1.10, 97.5% CI 0.95–1.26, p = 0.196): in this regard, recessive inheritance was selected as the model that maximized**

test statistics of the genetic case-control study. Model comparison also revealed that the recessive model was the best fit based on a lower BIC and a significant reduction in residual deviance ($p = 0.01562$) when confronted with the additive and dominant models. According to those results, we adopted the recessive model for further analyses.

## Clinical Features

Clinical and demographic characteristics of patients with A/A + A/C vs C/C genotypes are summarized in Table 1. Patients with the C/C genotype had a higher age at onset ($p = 0.04$), higher rate of bulbar-onset disease ($p = 0.001$), and a higher ΔWeight ($p = 0.003$) compared with those with the A/A + A/C phenotype. The ΔALSFRS-R was also higher in the C/C genotype but did not reach statistical significance. The demographics of the cohorts are summarized in eTable 3, links.lww.com/NXG/A551.

Comorbid FTD was diagnosed in 43 (8.4%) patients with A/A + A/C genotypes compared with 9 (10.8%) patients with the C/C genotype ($p = 0.30$). A comparison of cognitive test age-adjusted and education-adjusted scores in non–ALS-FTD ALS patients A/A + A/C genotypes and C/C genotype with controls is listed in eTable 4, links.lww.com/NXG/A551. Overall, ALS patients had worse scores in several tests, exploring frontal function than population-based controls. In addition, we found significantly lower scores in patients with the C/C genotype in the FAS ($p = 0.01$), SET–Emotion Attribution (SET-EA) ($p = 0.01$), and SET–Global Score ($p = 0.04$), as well as in the Fluency domain of Edinburgh Cognitive and Behavioural ALS Screen (ECAS) ($p = 0.014$) (Tables 2 and 3). By contrast, no differences were found in the frequency of neurobehavioral symptoms and anxiety and depression scores.

## Survival

Patients with the C/C genotype had a shorter survival than those with AA + AC genotypes (C/C: median survival time = 2.25 years, IQR = 1.33–3.92; A/A + C/C: survival = 2.90, IQR = 1.74–5.41; $p = 0.0001$) (Figure 1). Furthermore, the C/C genotype remained an independent predictor of survival in Cox multivariable analysis (eTable 5, links.lww.com/NXG/A551).

Survival from onset according to a recessive model in patients with the C/C genotype (green Line) vs A/A + A/C genotypes (blue line). Ticks indicate censored patients. $p < 0.0001$. ALS = amyotrophic lateral sclerosis.

## PET Findings

**The AA + AC participants showed a relative hypometabolism when compared with HCs in the left temporal and occipital cortex (Figure 2A, eTable 6, links.lww.com/NXG/A551). In comparison with HCs, the C/C group showed extensive clusters of relative hypometabolism encompassing bilateral frontal and temporal cortices, left occipital cortex, left precuneus, right caudate head and putamen, and right parietal cortex (Figure 2B; eTable 6, links.lww.com/NXG/A551).**

**Compared with the A/A + A/C genotypes, patients with the C/C genotype showed a significant (p = 0.000001) relative hypometabolism in the right frontal cortex (middle and inferior frontal gyri and precentral gyrus) (Figure 2, eTable 7, links.lww.com/NXG/A551). No clusters of relative hypermetabolism were observed (Figure 3).**

## 6.5 Discussion

**We have assessed a large population-based cohort of patents with ALS to identify the phenotypic characteristics of patients carrying the C/C genotype of the common variant rs12608932 of the UNC13A gene. Based on a recessive model, patients carrying the C/C genotype were older, had more frequent bulbar-onset disease and manifested a higher rate of weight loss than those with the A/A + A/C genotype. We did not find an excess of patients with ALS-FTD in those carrying the C/C phenotype, but we observed a reduced performance in the FAS and the Fluency domain of ECAS and in social cognition (SET test). Patients with the C/C genotype had significantly shorter survival, and the presence of this genotype was independent of known outcome modifiers in Cox multivariable analysis. Finally, patients with the C/C genotype had a specific pattern of hypometabolism on brain 18F-FDG-PET extending to frontal and precentral areas of the right hemisphere.**

**To date, the only study exploring the phenotype of UNC13A genotypes found that C/C carriers had an older age at symptom onset, displayed higher rates of bulbar-onset disease, had a higher incidence of ALS-FTD, and had a lower forced vital capacity at diagnosis.11 We did find that patients with C/C variant had lower FVC% values; however, in an exploratory analysis, we found that this difference disappeared when stratifying FVC% according to the type of onset (bulbar onset, median values, C/C 79.5% (IQR 61–96), A/A + A/C 84% (IQR 62–100), p = 0.41; spinal onset, median values, C/C 96% (IQR 83–112), A/A + A/C 97% (IQR 81–108), p = 0.79).**

**A relevant finding of our study is the higher rate of weight loss (ΔWeight) in patients carrying the C/C variant, a difference that persisted when stratifying patients according to the presence of bulbar**

**symptoms during diagnosis (bulbar symptoms present, C/C genotype 0.48 [IQR 0.10–1.21] kg/mo, A/A + A/C 0.43 [IQR 0–1.01] kg/mo, p = 0.04; bulbar symptoms not present, median values, C/C 0.26 [IQR 0–0.85] kg/mo, A/A + A/C 0.07 [IQR 0–0.57] kg/mo, p = 0.01). Therefore, more severe weight loss appears to be a specific feature of UNC13A, independent from the reduction of nutritional intake due to dysphagia, but likely related to other factors, such as increased energy expenditure (i.e., the so-called hypermetabolic state)21 and central mechanisms (hypothalamic atrophy).22**

**In contrast with a previous study,11 we did not observe a significantly higher incidence of ALS-FTD in the C/C genotype group. This difference is likely because we classified patients' cognitive impairment based on an extensive battery of cognitive/behavioral tests, as suggested by the revised ALS-FTD Consensus Criteria,17 while another study11 used only on ECAS. Of interest when moving to a single-test level, the C/C genotype group showed reduced performance in FAS and the Fluency domain of ECAS, i.e., 2 tests examining frontal functions. The only 2 other studies reporting cognitive tests showed impaired performance in C/C genotype patients in the language domain of ECAS11 or in Digit Span backward test.23**

**Patients with ALS in our cohort have worse scores than controls in SET, a test exploring the theory of mind (ToM). Of interest patients with the C/C genotype had a significantly more severe impairment in this function than those with the A/A + A/C genotype, particularly the SET-EA subtest, indicating their reduced ability to attribute emotion to others in social situations.24 Conversely, the 2 groups did not differ in cognitive ToM (SET-IA subtest). Concerning behavioral function in our cohort, patients with the C/C genotype showed a nonsignificant higher prevalence of disinhibition.**

**The negative influence of the UNC13A C/C genotype on ALS survival has been reported in populations with European and Chinese ancestry.7-9,25 Three phenotypic characteristics of patients with the C/C phenotype observed in our cohort, i.e., older age, bulbar onset, and rate of weight loss, are well-known negative prognostic factors of ALS.26,27 However, in our series, the C/C polymorphism remained an independent prognostic factor in Cox multivariable analysis, indicating that its effect on survival cannot be entirely explained by its phenotypic characteristics and is based on other still undisclosed mechanisms.**

**This study is the first to have assessed brain 18F-FDG-PET in relation to UNC13A genotypes in ALS. We found that patients with the C/C genotype compared with those with A/A + A/C genotypes had a cluster of relative hypometabolism in a large area of the right frontal cortex, particularly the middle and inferior frontal gyri and the precentral gyrus. Of interest a brain 18F-FDG-PET study reported that in ALS, impairment in ToM was correlated to bilateral dorsomedial and dorsolateral prefrontal cortices hypometabolism.28 In addition, in a previous PET study, the same areas were**

**found to distinguish ALS patients with and without cognitive impairment, although bilaterally located,29 indicating that selective cognitive impairment is a relevant feature of patients with ALS with the UNC13A C/C genotype. Studies assessing regional gray matter density with voxel-based morphometry and fMRI showed that affective ToM impairment was correlated to changes in right frontoinsular and anterior cingulate cortices.30,31 Studies with MRI assessing cortical thickness in relation to UNC13A genotypes reported cortical thinning in bilateral frontal and temporal cortices in patients with the C/C genotype compared with those with the A/A + A/C genotype.11,22 The difference between out PET findings and those related to MRI cortical thickness is probably due to the different characteristics of PET, which is a functional methodology, and MRI, which is mainly structural.**

**UNC13A protein has different functions, including regulating neurotransmitter and presynaptic vesicle release. However, it is unknown whether the intronic rs12608932 variant changes the function of the UNC13A protein or acts through different mechanisms or genetic variants that are in linkage disequilibrium with this single-nucleotide polymorphism (SNP). A recent study showed that TDP-43 depletion induces robust inclusion of a cryptic exon (CE) within UNC13A, resulting in nonsense-mediated decay and protein loss.32 In particular, the rs12608932 risk allele modifies the binding of TDP-43 to the UNC13A pre-mRNA, probably making UNC13A CE more susceptible to even the partial TDP-43 loss, which occurs early in degenerating neurons.32**

**This study is not without limitations. Cognitive and behavioral function PET studies were not available for all cohort patients. However, the 2 subcohorts were similar to the whole cohort for all relevant clinical factors, and patients underwent cognitive testing and PET within 2 months of diagnosis, limiting the selection bias. Second, this study is based on a recessive genotype model. However, the recessive model both in our cohort and in literature has been consistently reported to be the best fit. The genotypic model through which UNC13A affects ALS will be determined only when we know its exact pathophysiologic mechanism. A notable feature of the study is its population-based nature, including some 80% of incident patients.**

**In conclusion, we have found that the C/C rs12608932 genotype of UNC13A is associated with a specific motor and cognitive/behavioral phenotype, which reflects on PET findings. We confirmed the polymorphism's effect on survival and added novel characteristics to its phenotype, such as the greater weight loss independent from bulbar involvement. We also extended its cognitive phenotype to encompass social cognition. Our observations highlight the importance of adding the UNC13A rs12608932 variant to the ALS genetic panel to refine the individual prognostic prediction and reduce heterogeneity in clinical trials.**

## 6.6 References


**1. van Es MA, Hardiman O, Chio A, et al.. Lancet. 2017;390(10107):2084-2098. doi: 10.1016/S0140-6736(17)31287-4.**

**2. Grassano M, Calvo A, Moglia C, et al., American Genomic Center. Mutational analysis of known ALS genes in an Italian population-based cohort. Neurology. 2021;96(4):e600-e609. doi: 10.1212/WNL.0000000000011209.**

**3. Chiò A, Moglia C, Canosa A, et al.. ALS phenotype is influenced by age, sex, and genetics: a population-based study. Neurology. 2020;94(8):e802-e810. doi: 10.1212/WNL.0000000000008869.**

**4. Chiò A, Benzi G, Dossena M, Mutani R, Mora G. Severely increased risk of amyotrophic lateral sclerosis among Italian professional football players. Brain. 2005;128(pt 3):472-476. doi: 10.1093/brain/awh373.**

**5. Bandres-Ciga S, Noyce AJ, Hemani G, et al.. Shared polygenic risk and causal inferences in amyotrophic lateral sclerosis. Ann Neurol. 2019;85(4):470-481. doi: 10.1002/ana.25431.**

**6. Vasta R, Chia R, Traynor BJ, Chiò A. Unraveling the complex interplay between genes, environment, and climate in ALS. EBioMedicine. 2022;75:103795. doi: 10.1016/j.ebiom.2021.103795.**

**7. Diekstra FP, van Vught PWJ, van Rheenen W, et al.. UNC13A is a modifier of survival in amyotrophic lateral sclerosis. Neurobiol Aging. 2012;33(3):630.e3–8. doi: 10.1016/j.neurobiolaging.2011.10.029.**

**8. Chiò A, Mora G, Restagno G, et al.. UNC13A influences survival in Italian amyotrophic lateral sclerosis patients: a population-based study. Neurobiol Aging. 2013;34(1):357.e1–5. doi: 10.1016/j.neurobiolaging.2012.07.016.**

**9. Yang B, Jiang H, Wang F, et al.. UNC13A variant rs12608932 is associated with increased risk of amyotrophic lateral sclerosis and reduced patient survival: a meta-analysis. Neurol Sci. 2019;40(11):2293-2302. doi: 10.1007/s10072-019-03951-y.**

**10. van Eijk RPA, Jones AR, Sproviero W, et al., For UKMND-LiCALS and LITALS Study Group. Meta-analysis of pharmacogenetic interactions in amyotrophic lateral sclerosis clinical trials. Neurology. 2017;89(18):1915-1922. doi: 10.1212/WNL.0000000000004606.**

**11. Tan HHG, Westeneng HJ, van der Burgh HK, et al.. The distinct traits of the UNC13A polymorphism in amyotrophic lateral sclerosis. Ann Neurol. 2020;88(4):796-806. doi: 10.1002/ana.25841.**

**12. Chiò A, Mora G, Moglia C, et al., Piemonte and Valle d'Aosta Register for ALS PARALS. Secular trends of amyotrophic lateral sclerosis: the Piemonte and Valle d'Aosta register. JAMA Neurol. 2017;74(9):1097-1104. doi: 10.1001/jamaneurol.2017.1387.**

**13. Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Mot Neuron Disord. 2000;1(5):293-299. doi: 10.1080/146608200300079536.**

**14. Torrieri MC, Manera U, Mora G, et al.. Tailoring patients' enrollment in ALS clinical trials: the effect of disease duration and vital capacity cutoffs. Amyotroph Lateral Scler Frontotemporal Degener. 2022;23(1-2):108-115. doi: 10.1080/21678421.2021.1936063.**

**15. Iazzolino B, Pain D, Peotta L, et al.. Validation of the revised classification of cognitive and behavioural impairment in ALS. J Neurol Neurosurg Psychiatry. 2019;90(7):734-739. doi: 10.1136/jnnp-2018-319696.**

**16. Iazzolino B, Peotta L, Zucchetti JP, et al.. Differential neuropsychological profile of patients with amyotrophic lateral sclerosis with and without C9orf72 mutation. Neurology. 2021;96(1):e141-e152. doi: 10.1212/WNL.0000000000011093.**

**17. Strong MJ, Abrahams S, Goldstein LH, et al.. Amyotrophic lateral sclerosis–frontotemporal spectrum disorder (ALS-FTSD): revised diagnostic criteria. Amyotroph Lateral Scler Frontotemporal Degener. 2017;18(3-4):153-174. doi: 10.1080/21678421.2016.1267768.**

**18. Montuschi A, Iazzolino B, Calvo A, et al.. Cognitive correlates in amyotrophic lateral sclerosis: a population-based study in Italy. J Neurol Neurosurg Psychiatry. 2015;86(2):168-173. doi: 10.1136/jnnp-2013-307223.**

**19. Chiò A, Calvo A, Mazzini L, et al., PARALS. Extensive genetics of ALS: a population-based study in Italy. Neurology. 2012;79(19):1983-1989. doi: 10.1212/WNL.0b013e3182735d36.**

**20. Della Rosa PA, Cerami C, Gallivanone F, et al., EADC-PET Consortium. A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. Neuroinformatics. 2014;12(4):575-593. doi: 10.1007/s12021-014-9235-4.**

**21. Bouteloup C, Desport JC, Clavelou P, et al.. Hypermetabolism in ALS patients: an early and persistent phenomenon. J Neurol. 2009;256(8):1236-1242. doi: 10.1007/s00415-009-5100-z.**

**22. Gorges M, Vercruysse P, Müller HP, et al.. Hypothalamic atrophy is related to body mass index and age at onset in amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry. 2017;88(12):1033-1041. doi: 10.1136/jnnp-2017-315795.**

**23. Placek K, Baer GM, Elman L, et al.. UNC13A polymorphism contributes to frontotemporal disease in sporadic amyotrophic lateral sclerosis. Neurobiol Aging. 2019;73:190-199. doi: 10.1016/j.neurobiolaging.2018.09.031.**

**24. Dodich A, Cerami C, Canessa N, et al.. A novel task assessing intention and emotion attribution: Italian standardization and normative data of the Story-based Empathy Task. Neurol Sci. 2015;36(10):1907-1912. doi: 10.1007/s10072-015-2281-3.**

**25. Vidal-Taboada JM, Lopez-Lopez A, Salvado M, et al.. UNC13A confers risk for sporadic ALS and influences survival in a Spanish cohort. J Neurol. 2015;262(10):2285-2292. doi: 10.1007/s00415-015-7843-z.**

**26. Chiò A, Logroscino G, Hardiman O, et al., Eurals Consortium. Prognostic factors in ALS: a critical review. Amyotroph Lateral Scler. 2009;10(5-6):310-323. doi: 10.3109/17482960802566824.**

**27. Moglia C, Calvo A, Grassano M, et al., Piemonte and Valle d'Aosta Register for ALS PARALS. Early weight loss in amyotrophic lateral sclerosis: outcome relevance and clinical correlates in a population-based cohort. J Neurol Neurosurg Psychiatry. 2019;90(6):666-673. doi: 10.1136/jnnp-2018-319611.**

**28. Carluer L, Mondou A, Buhour MS, et al.. Neural substrate of cognitive theory of mind impairment in amyotrophic lateral sclerosis. Cortex. 2015;65:19-30. doi: 10.1016/j.cortex.2014.12.010.**

**29. Canosa A, Pagani M, Cistaro A, et al.. 18F-FDG-PET correlates of cognitive impairment in ALS. Neurology. 2016;86(1):44-49. doi: 10.1212/WNL.0000000000002242.**

**30. Cerami C, Dodich A, Canessa N, et al.. Emotional empathy in amyotrophic lateral sclerosis: a behavioural and voxel-based morphometry study. Amyotroph Lateral Scler Frontotemporal Degener. 2014;15(1-2):21-29. doi: 10.3109/21678421.2013.785568.**

**31. Trojsi F, Di Nardo F, Santangelo G, et al.. Resting state fMRI correlates of Theory of Mind impairment in amyotrophic lateral sclerosis. Cortex. 2017;97:1-16. doi: 10.1016/j.cortex.2017.09.016.**

**32. Brown AL, Wilkins OG, Keuss MJ, et al.. Common ALS/FTD risk variants in UNC13A exacerbate its cryptic splicing and loss upon TDP-43 mislocalization. Nature. 2022;603(7899):131-137. doi: 10.1038/s41586-022-04436-3.**

## 6.7 Tables And Figures

### Table 1

**Clinical and Demographic Characteristics of ALS Patients With A/A + A/C vs C/C rs12608932 UNC13A Genotypes**

| | A/A + A/C n = 1,225 | C/C n = 184 | *p* Value |
|---|---|---|---|
| **Age at onset (y, median, IQR)** | 67.7 (59.5–73.8) | 69.4 (62.5–75.4) | 0.02 |
| **Sex (female)** | 539 (44.0%) | 90 (48.9%) | 0.21 |
| **Site of onset (bulbar)** | 361 (29.5%) | 77 (41.8%) | 0.001 |
| **Bulbar symptoms at diagnosis** | 624 (50.9%) | 124 (67.4%) | 0.001 |
| **Diagnostic delay (mo, median, IQR)** | 9.0 (5.1–13.2) | 7.9 (5.1–12.0) | 0.14 |
| **Education (median, IQR)** | 8.0 (5.0–13.0) | 8.0 (5.0–11.0) | 0.18 |
| **ALSFRS-R at diagnosis (median, IQR)** | 42 (38–45) | 42 (36–45) | 0.35 |
| **FVC% at diagnosis[a] (median, IQR)** | 91 (71–104) | 85 (68–101) | 0.09 |
| **BMI at diagnosis[b] (median, IQR)** | 24.2 (22.0–26.8) | 24.4 (22.0–26.8) | 0.91 |
| **ΔALSFRS-R (points/mo, median, IQR)** | 0.63 (0.30–1.24) | 0.72 (0.35–1.52) | 0.10 |
| **ΔWeight[b] (kg/mo, median, IQR)** | 0.25 (0.0–0.85) | 0.37 (0.0–1.12) | 0.004 |
| **ALS-FTD[c]** | 43 (8.4%) | 9 (10.8%) | 0.30 |
| **King stage (1/2/3/4A + 4B) at diagnosis** | 541/385/248/45 | 77/50/51/6 | 0.15 |
| **MiToS stage (0/1/2/3/4) at diagnosis** | 833/336/38/9/3 | 130/45/6/3/0 | 0.62 |

Abbreviations: ALS = amyotrophic lateral sclerosis; ALSFRS-R = amyotrophic lateral sclerosis functional rating scale - revised; BMI = body mass index; FVC = forced vital capacity; IQR = interquartile range; MiToS = Milano-Torino Staging.
*p* values are calculated with $\chi^2$ (discrete variables) or Mann-Whitney *U* test (continuous variables).
[a] Available for 1,314 cases (A/A + A/C, 1,141; C/C, 173).
[b] Available for 1,373 cases (A/A + A/C, 1,193; C/C, 180).
[c] Available for 592 cases (509 A/A + A/C and 83 C/C).

## Table 2

**Median Values (Interquartile Range) of Age-Corrected and Education-Corrected Scores of Cognitive Tests in ALS Patients With A/A + A/C vs C/C UNC13A rs12608932 Genotypes**

| Test | A/a + A/C (n = 467) | C/C (n = 74) | *p* Valu |
|---|---|---|---|
| MMSE | 27.9 (26.7–30.0) n = 466 | 27.7 (26.2–29.4) n = 73 | 0.17 |
| FAS | 29.5 (23.6–37.2) n = 456 | 26.7 (18.5–34.1) n = 72 | 0.01 |
| CAT | 19.8 (16.3–22.3) n = 456 | 19.6 (15.5–22.0) n = 72 | 0.58 |
| FAB | 15.3 (13.7–16.9) n = 401 | 14.9 (13.3–16.3) n = 64 | 0.13 |
| Digit span FW | 5.7 (5.1–6.3) n = 420 | 5.5 (4.9–6.5) n = 63 | 0.28 |
| Digit span BW | 4.0 (3.5–4.5) n = 420 | 4.0 (3.3–4.6) n = 63 | 0.74 |
| TMT A | 36 (24–53.25) n = 422 | 41 (29–60) n = 63 | 0.13 |
| TMT B | 68.5 (34.25–146) n = 422 | 70 (48–188) n = 63 | 0.09 |
| TMT B-A | 34 (8–81) n = 422 | 38 (13–135) n = 63 | 0.21 |
| RAVL-IR | 40.2 (34–46.3) n = 238 | 38.9 (30–45.3) n = 43 | 0.08 |
| RAVL-DR | 7.0 (4.0–9.0) n = 238 | 6.0 (3.0–9.0) n = 43 | 0.45 |
| BSRT-IR | 5.7 (4.5–6.9) n = 246 | 5.7 (4.7–7.4) n = 36 | 0.64 |
| BRRT-DR | 6.7 (5.2–7.8) n = 246 | 6.5 (5.5–8.0) n = 36 | 0.68 |
| ROCF-IR | 32.0 (28.8–34.5) n = 335 | 31.9 (27.4–34.3) n = 54 | 0.52 |
| ROCF-DR | 12.3 (8.3–16.7) n = 335 | 12.2 (6.9–16.8) n = 54 | 0.66 |
| CPM47 | 29.1 (25.3–32.0) n = 449 | 28.7 (24.3–32.7) n = 65 | 0.69 |
| Clock[a] | 4 (3–5) n = 392 | 4 (3–5) n = 58 | 0.78 |
| SET-IA | 4.9 (3.3–6.0) n = 78 | 4.2 (3.1–5.7) n = 17 | 0.49 |
| SET−CI | 4.6 (3.4–5.3) n = 78 | 4.0 (2.9–5.1) n = 17 | 0.16 |
| SET-EA | 5.0 (3.4–6.0) n = 78 | 3.2 (2.2–5.1) n = 17 | 0.01 |
| SET-GS | 14.0 (10.3–16.2) n = 78 | 10.1 (8.9–14.7) n = 17 | 0.04 |
| HADS-A | 7 (5–10) n = 370 | 8 (5–10.5) n = 57 | 0.27 |
| HADS-D | 5 (3–8) n = 370 | 5 (3–9) n = 57 | 0.73 |

Fifty-one patients with FTD have been excluded.
Abbreviations: ALS = amyotrophic lateral sclerosis; BSRT-DR = babcock story recall test-delayed recall; BSRT-IR = babcock story recall test-immediate Recall CAT = category fluency test; Clock = Clock Drawing Test; CPM47 = Raven Colored Progressive Matrices; Digit Span BW = digit span backward; Digit Span FW = digi span forward; FAB = frontal assessment battery; FAS = letter fluency test; HADS-A = Hospital Anxiety and Depression Scale-Anxiety; HADS-D = Hospital Anxiety an Depression Scale-Depression; MMSE = Mini-Mental State Examination; RAVL-DR = rey auditory verbal learning test-delayed recall; RAVL-IR = rey auditory verba learning test-immediate recall; ROCF-DR = babcock story recall test-delayed recall; ROCF-IR = babcock story recall test-immediate recall; SET−CI = Story-Base Empathy Task–Causal Inference; SET-EA = Story-Based Empathy Task-Emotion Attribution; SET-GS = Story-Based Empathy Task-Global Score; SET-IA = Story-Base Empathy Task-Intention Attribution; TMT A = Trail Making Test A; TMT B = Trail Making Test B; TMT B-A = Trail Making Test B-A.
*p* values are calculated with the Mann-Whitney *U* test.
[a]The scores of the Clock test are not corrected by age and education because no Italian normative data are available.

## Table 3

**Median Values (Interquartile Range) of Age-Corrected and Education-Corrected Scores of ECAS in ALS Patients With A/A + A/C vs C/C UNC13A rs12608932 Genotypes**

**Table 3** Median Values (Interquartile Range) of Age-Corrected and Education-Corrected Scores of ECAS in ALS Patients With A/A + A/C vs C/C *UNC13A* rs12608932 Genotypes

| Domains | A/A + A/C (n = 194) | C/C (n = 29) | *p* Value |
|---|---|---|---|
| Language | 26 (23–27) | 25 (21.5–26.5) | 0.17 |
| Fluency | 16 (13.5–18) | 14 (9–16) | 0.014 |
| Executive | 30 (23–37) | 30 (22.5–35) | 0.56 |
| Memory | 17 (14–19.5) | 16 (13–19.5) | 0.20 |
| Visuospatial | 12 (11–12) | 11 (11–12) | 0.24 |
| ALS-specific functions | 72 (60–80) | 69 (52.5–77.5) | 0.17 |
| ALS nonspecific functions | 29 (25–31) | 27 (23.5–32) | 0.40 |
| ECAS total score | 100 (86–112) | 96 (76.5–110) | 0.19 |

Abbreviations: ALS = amyotrophic lateral sclerosis; ECAS = Edinburgh Cognitive and Behavioural ALS Screen; FTD = frontotemporal dementia. Patients with FTD have been excluded. *p* values are calculated with the Mann-Whitney *U* test.

## Figure 1

**Survival From Onset According to a Recessive Model in Patients With ALS**

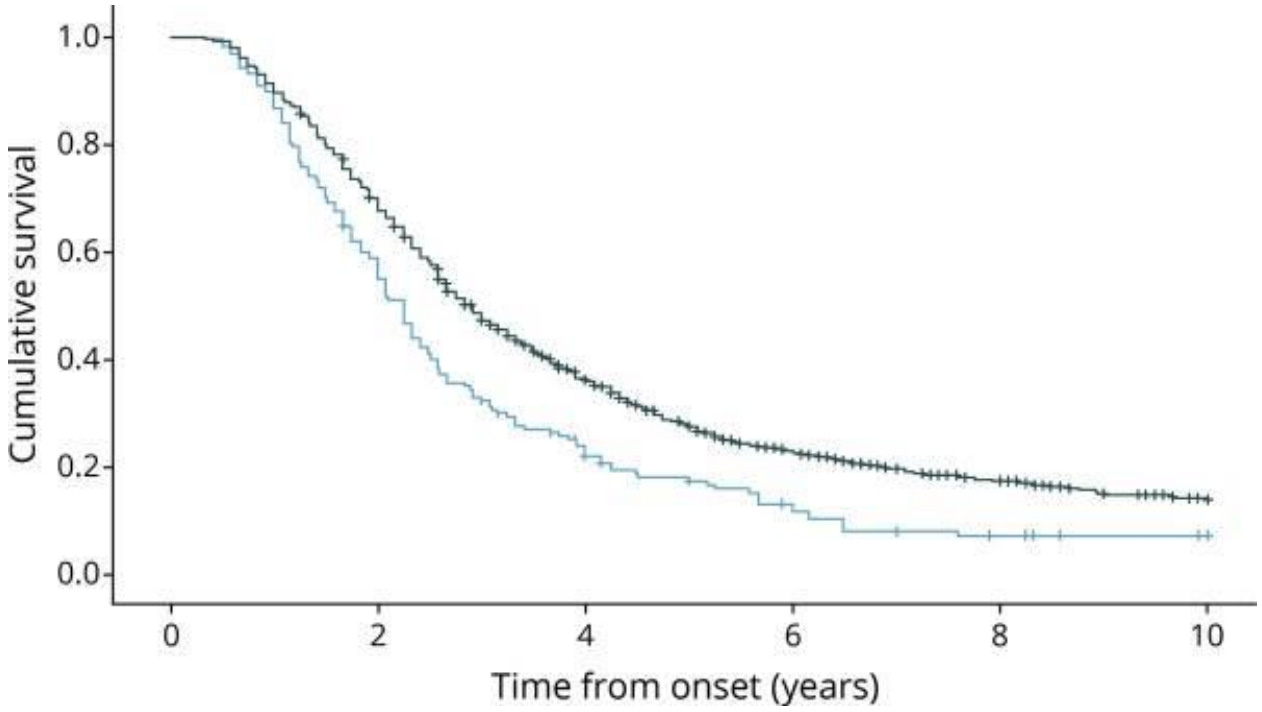
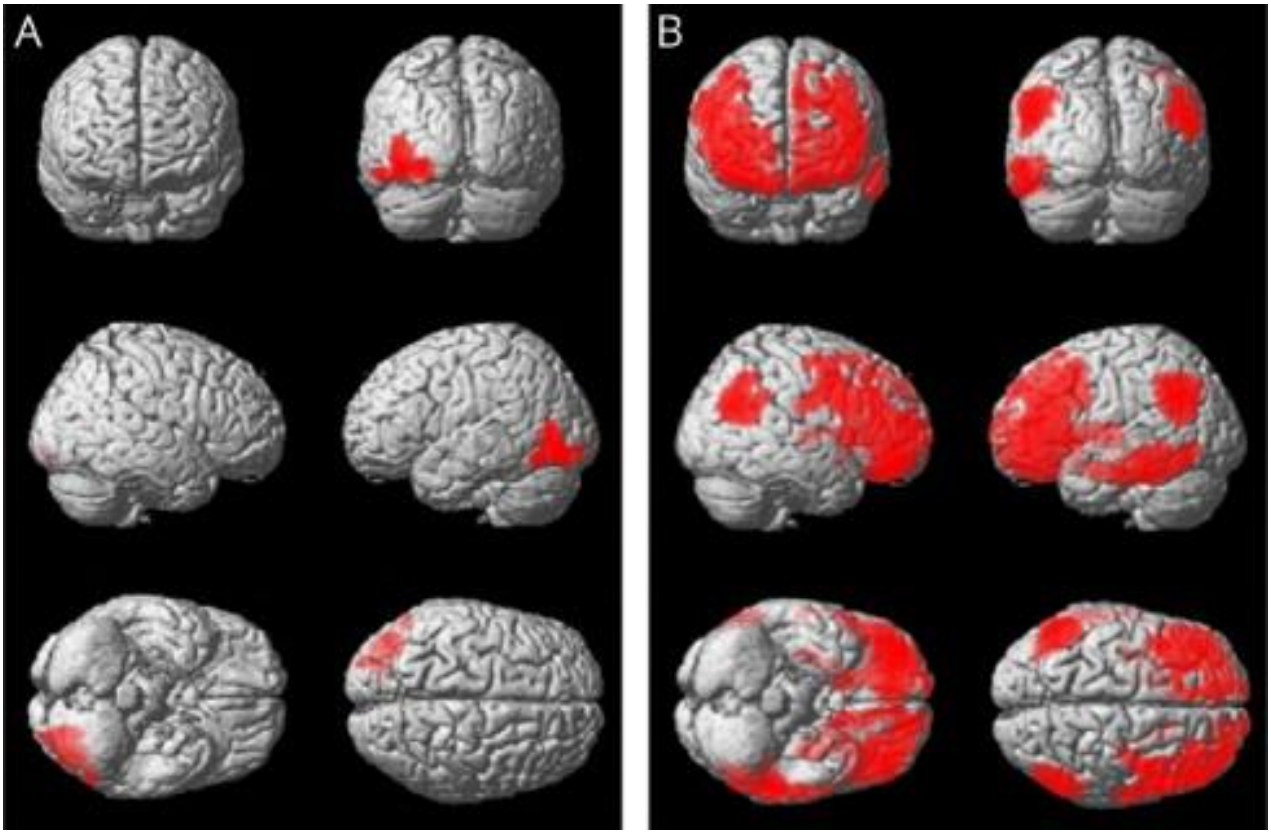



**Figure 2**

**18FDG-PET Imaging of Cases With ALS vs Controls**

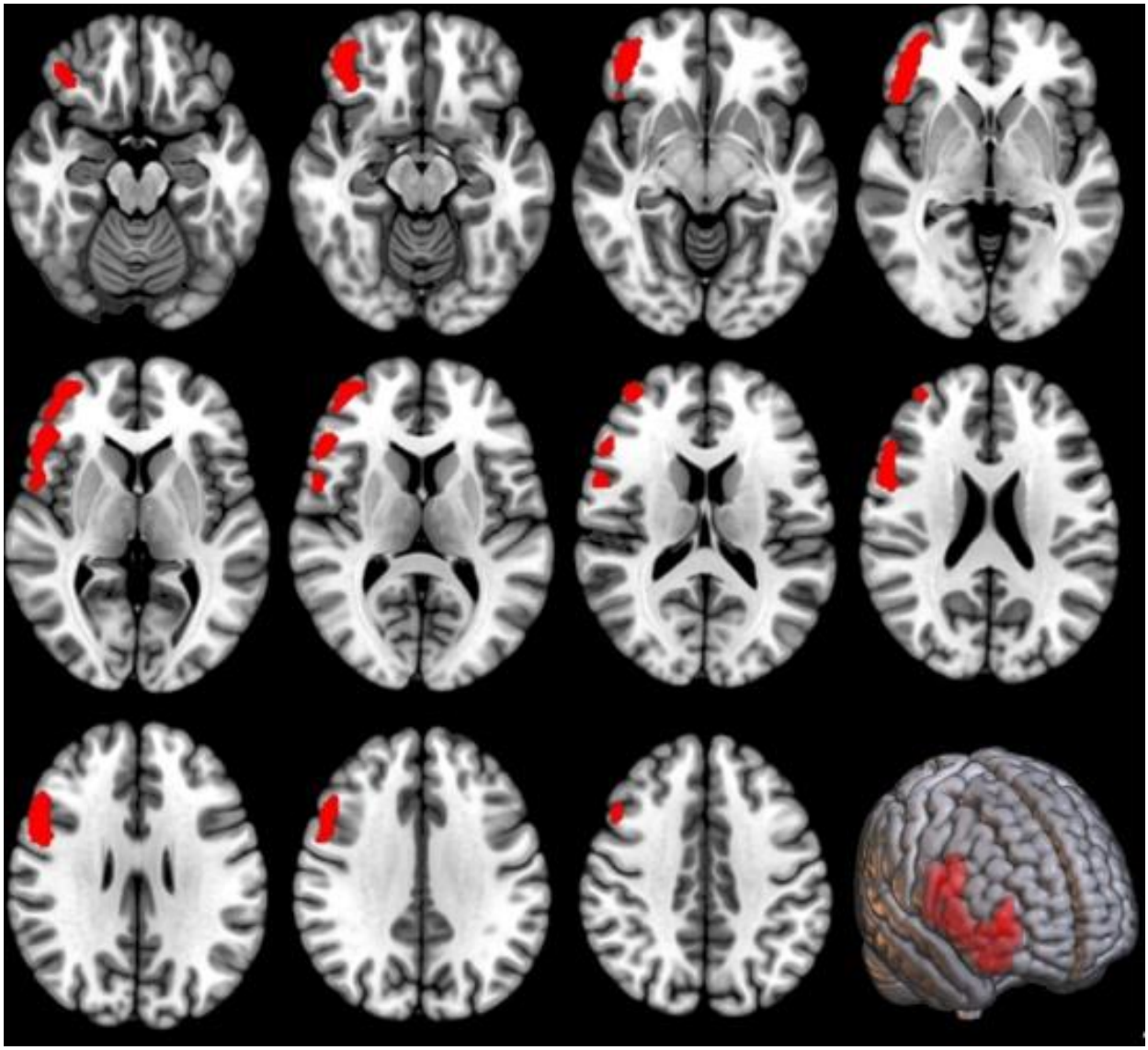
**(A) A/A + A/C Group: the regions showing a statistically significant relative hypometabolism in the A/A + A/C group compared with those in controls are marked in red. (B) C/C Group: the regions showing a statistically significant relative hypometabolism in the C/C group compared with those in controls are marked in red. ALS = amyotrophic lateral sclerosis; 18F-FDG-PET = 18F-2-fluoro-2-deoxy-d-glucose-PET.**



## Figure 3

**18FDG-PET Imaging of UNC13A A/A+A/C vs C/C Group**

**The regions showing a statistically significant relative hypometabolism in the C/C group compared with those in the A/A + A/C group are marked in red. They are reported on axial sections of a brain MRI template and the brain surface of a glass brain rendering (bottom right). ALS = amyotrophic lateral sclerosis; 18F-FDG-PET = 18F-2-fluoro-2-deoxy-d-glucose-PET.**



# 6. Exploring the phenotype of Italian patients with ALS with intermediate ATXN2 polyQ repeats

**As published in *Chio, A., Moglia, C., Canosa, A., Manera, U., Grassano, M., Vasta, R., ... & Calvo, A. (2022). Exploring the phenotype of Italian patients with ALS with intermediate ATXN2 polyQ repeats. Journal of Neurology, Neurosurgery & Psychiatry, 93(11), 1216-1220*.**

**AUTHORS**

**Adriano Chio1,2, Cristina Moglia1,2, Antonio Canosa1,2, Umberto Manera1,2, Maurizio Grassano1, Rosario Vasta1, Francesca Palumbo1, Salvatore Gallone2, Maura Brunetti1, Marco Barberis3, Fabiola De Marchi4, Clifton Dalgard5,6, Ruth Chia7, Gabriele Mora1, Barbara Iazzolino1, Laura Peotta1, Bryan Traynor8,9, Lucia Corrado10, Sandra D'Alfonso10, Letizia Mazzini4, Andrea Calvo1,2**

**Author Affiliations**

**1 Rita Levi Montalcini' Department of Neuroscience, University of Turin, Torino, Italy**

**2 Neurology 1, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Torino, Italy**

**3 Genetics, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Torino, Italy**

**4 Neurology, Azienda Ospedaliero-Universitaria Maggiore della Carità, Novara, Italy**

**5 Department of Anatomy, Physiology & Genetics, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA**

**6 The American Genome Center, Collaborative Health Initiative Research Program, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA**

**7 Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Porter Neuroscience Research Center, Bethesda, Maryland, USA**

**8 Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, Maryland, USA**

**9 Department of Neurology, Johns Hopkins, Baltimore, Maryland, USA**

**10 Department of Health Sciences Interdisciplinary Research Center of Autoimmune Diseases, University of Eastern Piedmont Amedeo Avogadro School of Medicine, Novara, Italy**

## 7.1 Abstract

### Objective

**To detect the clinical characteristics of patients with amyotrophic lateral sclerosis (ALS) carrying an intermediate ATXN2 polyQ number of repeats in a large population-based series of Italian patients with ALS.**

### Methods

**The study population includes 1330 patients with ALS identified through the Piemonte and Valle d'Aosta Register for ALS, diagnosed between 2007 and 2019 and not carrying C9orf72, SOD1, TARDBP and FUS mutations. Controls were 1274 age, sex and geographically matched Italian subjects, identified through patients' general practitioners.**

### Results

**We found 42 cases and 4 controls with≥31 polyQ repeats, corresponding to an estimated OR of 10.4 (95% CI 3.3 to 29.0). Patients with≥31 polyQ repeats (ATXN2+) compared with those without repeat expansion (ATXN2−) had more frequently a spinal onset (p=0.05), a shorter diagnostic delay (p=0.004), a faster rate of ALSFRS-R progression (p=0.004) and King's progression (p=0.004), and comorbid frontotemporal dementia (7 (28.0%) vs 121 (13.4%), p=0.037). ATXN2+ patients had a 1-year shorter survival (ATXN2+ patients 1.82 years, 95% CI 1.08 to 2.51; ATXN2− 2.84 years, 95% CI 1.67 to 5.58, p=0.0001). ATXN2 polyQ intermediate repeats was independently related to a worse outcome in Cox multivariable analysis (p=0.006).**

### Conclusions

**In our population-based cohort, ATXN2+ patients with ALS have a distinctive phenotype, characterised by a more rapid disease course and a shorter survival. In addition, ATXN2+ patients have a more severe impairment of cognitive functions. These findings have relevant implications on clinical practice, including the possibility of refining the individual prognostic prediction and improving the design of ALS clinical trials, in particular as regards as those targeted explicitly to ATXN2.**

## 7.2 Introduction

**Amyotrophic lateral sclerosis (ALS) is a multisystem disorder of adult life characterised by progressive degeneration of upper and lower motor neurons and frontotemporal cortex neurons. Several genes have been related to this fatal neurodegenerative disorder, accounting for 10%–20% of ALS cases.1 2 Among these, an intermediate-length CAG number of repeats (encoding ≥31 glutamines, polyQ) in the ataxin 2 (ATXN2) gene, already known as the cause of spinocerebellar ataxia type 2 (characterised by a number of polyQ≥38), is recognised to be associated with an increased risk of developing ALS and has been reported to be a modifier of survival.3 4 More recently, it has been reported that an intermediate number of ATXN2 polyQ repeats can also be a modifier of frontotemporal dementia (FTD) phenotype.5 6 Nevertheless, the phenotypic characteristics of patients with ALS with intermediate-length CAG repeats in the ATXN2 gene are still incompletely understood. This study aimed at detecting the clinical characteristics of patients with ALS carrying an intermediate number of ATXN2 polyQ repeats in a large population-based series of Italian patients with ALS.**

## 7.3 Methods

**The study population includes 1487 patients with ALS identified through the Piemonte and Valle d'Aosta Register for ALS, a prospective population-based register active since 1995. The characteristics of the register have been reported elsewhere.7 For the present paper, we considered ALS cases diagnosed between 2007 and 2019. Patients met the El Escorial revised diagnostic criteria for definite, probable and probable laboratory-supported ALS.8**

**ALSFRS-R mean monthly decline (ΔALSFRS-R) was calculated using the following formula: (48 – ALSFRS-R score at diagnosis)/(months from onset to diagnosis). Similarly, weight mean monthly decline (ΔWeight) as (Weight at diagnosis – healthy body weight)/(months from onset to diagnosis). Finally, to have a proxy of disease spread, we calculated the mean/monthly decline of King's staging as (King's staging at diagnosis)/(months from onset to diagnosis).**

**A total of 928 patients underwent an extensive cognitive battery at the time of diagnosis. These cases were classified into five categories according to the Consensus Criteria for diagnosing**

**frontotemporal cognitive and behavioural syndromes in ALS.9 The battery assessed executive function, memory, visuospatial function, language and social cognition using the following tests Letter Fluency test; Category Fluency Test; Frontal Assessment Battery; Trail Making Test A, B and B-A; Rey-Osterrieth Complex Figure Test, immediate (IR) and delayed recall (DR); Rey Auditory Verbal Learning Test (RAVL), immediate (IR) and DR; BSRT, immediate (IR) and DR; Digit Span Forward and Backward; Raven's Colored Progressive Matrices (CPM47); Story-Based Empathy Task; and Mini-Mental State Examination. The raw scores of all tests were age, sex and education-corrected using the more recent Italian normative.10**

**Neurobehavioral dysfunction was determined with the Frontal Systems Behavior Scale (FrSBe), using the Family-form evaluated by a close relative/caregiver (scores: normal≤59, borderline 60–64; pathological≥65). For this study, we considered the change in points for each of the three domains of FrSBe (apathy, disinhibition, executive) from before-disease to disease scores. If a subject had scores reflecting a frontal systems abnormality both in the premorbid and post-illness forms, they were considered pathological only if there was an increase of≥10 points at the T-score between the two states.10 Anxiety and depression were assessed using the Hospital Anxiety and Depression Scale; the item 'I feel slowed down' was discussed with patients to have them not refer to physical disability.11**

### Controls

**Controls were 1274 age, sex and geographically matched Italian subjects, identified through patients' general practitioners.**

### Genetic analysis

**All patients included in the study were tested for SOD1 (all exons), TARDBP (exon 6), FUS (exons 14 and 15) mutations, and C9ORF72 intronic expansion using the methods described elsewhere.12 However, since 1180 cases underwent whole-genome sequencing, no mutation in the other exons of TARDBP and FUS were found.**

### ATXN2 CAG repeat analysis

**In 434 cases and 509 controls, ATXN2 CAG repeat in exon 1 (NM_002973.3) was amplified using a fluorescent primer and sized by capillary electrophoresis on an ABI3130 genetic analyzer (Applied Biosystems, Foster City, California, USA). In 1043 cases and 765 controls, ATXN2 CAG repeats were detected through next-generation sequencing. All subjects were screened for the C9orf72 intronic expansion using a standard repeat-primed PCR13 (Renton et al. 2011 PMID: 21944779). Repeat lengths of≥30 units with the characteristic sawtooth pattern were considered pathogenic. Whole-**

**genome sequencing was performed at The American Genome Center located at the Uniformed Services University, Bethesda, Maryland, USA. Libraries were prepared using TruSeq DNA PCR-Free High Throughput Library Prep Kit (Illumina) per the manufacturer's instructions. Sequencing was performed on an Illumina HiSeqX10 sequencer using paired-end 150 base pair reads. A significant advantage of next-generation sequencing is its ability to reliably assay repeat expansions, such as those in C9orf72 and ATXN2. ATXN2 CAG repeats were deemed intermediate if they were in the 31–38 range. ExpansionHunter—Targeted software (V.3.0.1) was used to estimate repeat lengths of known, disease-causing expansions in samples that had undergone whole-genome sequencing.14 This algorithm has been validated using experimentally confirmed samples carrying the C9orf72 and ATXN2 repeat expansions. In particular, cases with≥31 polyQ repeats were identified correctly with both PCR and ExpansionHunter.**

### Statistical methods

**Multivariable analysis for survival was performed with the Cox proportional hazards model (stepwise backward) with a retention criterion of p value<0.1. A p value<0.05 was considered significant. Statistical analyses were carried out using the SPSS V.26.0 statistical package (SPSS).**

## 7.4 Results

**We assessed for intermediate ATXN2 polyQ repeats 1487 patients with ALS diagnosed in Piemonte e Valle d'Aosta between 2007 and 2019. Of these, 157 were subsequently excluded from the present analysis because they carried a genetic mutation of one of the most common ALS-related genes (C9orf72, 97; SOD1, 32; TARDBP, 21; FUS, 6). We decided to exclude C9orf72 patients from this analysis because we found in a previous paper that in patients with this mutation intermediate ATXN2 polyQ repeats do not modify survival.15 As for SOD1, TARDBP and FUS, we chose to also exclude these cases because of the strong influence of these genes on survival. The study was therefore performed on 1330 patients. Controls were 1274 subjects matched to cases by age, gender and geographical origin. Patients and controls did not differ for the main demographic variables (online supplemental table 1).ATXN2.**

**In our cohort, we found 46 cases and 44 controls with 27 to 30 polyQ repeats, and 42 cases and 4 controls with≥31 polyQ repeats. The estimated OR for 27–30 polyQ was 0.99 (95% CI 0.66 to 1.52). The estimated OR for≥31 polyQ was 10.4 (95% CI 3.3 to 29.0). The 4 controls had 31 (3) and 32 (1) polyQ repeats. Their age ranged between 55 and 72 years. They were neurologically normal and**

**they had no family history for neurodegenerative diseases or ataxia. The number of polyQ repeats in patients was 20 (31 repeats), 12 (32 repeats), 4 (33 repeats), 3 (34 repeats), 2 (35 repeats) and 1 (38 repeats).**

## Phenotype of ATXN2+ patients (table 1).

**Patients with intermediate ATXN2 polyQ repeats≥31 polyQ repeats (ATXN2+) and those without expansion (<31 polyQ repeats) (ATXN2−) did not differ for the age at onset but ATXN2+ had more frequent spinal onset (p=0.05). In addition, ATXN+ patients had a shorter diagnostic delay (p=0.004) and a faster rate of progression as measured by ΔALSFRS-R (p=0.004) and ΔKing's (p=0.004). A total of 928 patients (25 ATXN2+ and 903 ATXN2−) and 129 controls underwent cognitive examination. The demographic and clinical characteristics of these patients did not differ from that of the overall cohort (online supplemental table 2). ATXN2+ patients were more commonly diagnosed as ALS-FTD (7 (28.0%) vs 121 (13.4%), p=0.037) (table 2). When assessing the differences in specific tests, excluding the 128 patients with ALS-FTD, we found a significantly worse performance of ATXN2+ patients in the RAVL − Delayed Recall (p=0.023), the BSRT − Immediate Recall (p=0.044), and in the Executive Functions domain of ECAS (p=0.034) (online supplemental tables 3 and 4). No differences were found in the behavioural function assessed with FrSBe. Finally, no differences were found in anxiety, while depression was significantly more severe in ATXN2+ patients (p=0.04).**

### Survival analysis

**ATXN2+ patients had a 1 year shorter survival (median survival time for ATXN2+ patients = 1.82 years, 95% CI 1.08 to 2.51; ATXN2− =2.84 years, 95% CI 1.67 to 5.58, p=0.0001, figure 1). The Cox multivariable analysis confirmed that the ATXN2 polyQ intermediate number of repeats was independently related to a worse outcome compared with non-expanded patients in our cohort (p=0.006) (table 3). A separate analysis including only the patients not assessed in our previous paper3 gave similar results (data not shown).**

### Analysis of oligogenicity

**All but two patients with≥31 polyQ repeats underwent whole-genome sequencing. We extracted variant information for 47 genes previously implicated in ALS pathogenesis (see online supplemental table 5 for the list of extracted genes). The following six mutations of ALS-related genes, all in heterozygosis, of ALS-related genes were detected (one for each patient): OPTN: p.L111R; SETX: p.H476R; CCNF: p.R123H; EWSR1: p.A132S; SETX: p.V919I; DTCN1: p.A354V. Allele frequencies and predicted functional effects of identified genetic variants are reported in online supplemental table 6. Patients carrying one of these mutations (n=6) did not differ from those not**

**carrying other genetic mutations (n=34) for any clinical characteristic but the median weight decline (Δweight) (see online supplemental table 7).**

## 7.5 Discussion

**We have assessed a large population-based cohort of Italian patients with ALS without mutations for C9orf72, SOD1, TARDBP and FUS genes to identify the clinical signature of ATXN2 polyQ intermediate number of repeats. We have found that ATXN2+ patients, despite having more commonly spinal onset, were characterised by a more rapid clinical course, as shown by (1) a 1.5-fold ΔALSFRS-R at the time of diagnosis compared with ATXN2- patients; (2) a shorter diagnostic delay, a factor related to a faster disease progression; and (3) a significantly higher ΔKing's, indicating a more rapid spreading of symptoms from one to three body regions. The greater aggressiveness of ALS in subjects with ATXN2 polyQ intermediate number of repeats is reflected in the 1 year shorter survival (median survival time, 1.82 vs 2.84), confirming our previous findings.3 This result was independent of relevant prognostic factors in Cox multivariable analysis. Finally, patients with ATXN2 PolyQ intermediate number of repeats were more frequently affected by frontotemporal impairment.**

**ATXN2 polyQ intermediate number of repeats were first recognised as a risk factor for ALS in 2010 in a cohort of US patients (using a cut-off≥27)16 and subsequently confirmed in populations of different ethnic origin,3 with the only exception of South Africans.17 Although a length of 27–33 polyQ was initially considered significantly associated with ALS, later studies have shown that the cut-off is≥31 polyQ repeats. In our population, we found that the best risk cut-off for ALS is≥31 polyQ since the distribution of alleles in the 27–30 polyQ repeats range was substantially similar among patients and controls.**

**ATXN2 is an RNA binding protein with an essential function in the nucleocytoplasmic shuttling of RNA and regulation of transcription.18 However, the pathogenic mechanism of ATXN2 in ALS is unknown. It has been reported that ATXN2 induces an increase of phosphorylated TDP-43 in the spinal anterior horn but not in motor cortex neurons of patients with ALS.19 Interestingly, our data show that the spinal phenotype is more common in patients with ATXN2 polyQ intermediate number of repeats than those without expansion.**

**A recent paper has reported that ATXN+patients (cut-off limits≥31) did not show any survival difference compared with ATXN-.20 The discrepancy with the present study is likely related to the**

**different nature of the two cohorts, that is, the prevalent, referral centers, population in the US study (as indicated by the young age at onset (~60 years), the median survival (~3.5 years) and finally the very low percentage of patients with ALS-FTD (3.8%)) and the incident population in our study. It is therefore possible that at least a part of fast progressor patients, including those who are ATXN+, have not been caught in the US study. Several papers have demonstrated that prevalent and incident populations strongly differ from the clinical point of view, including survival, supporting the notion that studies derived from clinical cohorts should be cautiously interpreted.21 22**

**A novel finding in this paper is the identification of a correlation between ATXN2 polyQ intermediate number of repeats in ALS and cognitive impairment. ALS-FTD was two times more frequent among the ATXN2+ patients (28% vs 13.4%). Furthermore, some cognitive tests related to frontal function (RAVL – delayed recall, BSRT – immediate recall, and Executive Functions domain of ECAS) were significantly more compromised in ATXN2+ patients. Interestingly, ATXN2 polyQ intermediate number of repeats (using a cut-off of≥27) have been recently proposed to be a modifier of the behavioural variant FTD phenotype, with earlier age at onset and more frequent parkinsonian and psychotic symptoms.5 6 23 However, no increased risk of developing behavioural variant FTD was reported by other studies.24 25 An antagonistic pleiotropic role in cognition of ATXN2 has been identified, with a positive influence on verbal–numerical reasoning, reaction time, educational attainment and cognitive resilience,26 27 while in spinocerebellar ataxia 2 polyQ expansions are related to cognitive impairment in executive functions, memory and visuoconstructive skills.28 Finally, a postmortem study in patients with non-fluent primary progressive aphasia with 39 polyQ repeat expansion showed neuronal loss and gliosis, associated with a superficial laminar spongiosis, were severe in the superficial layers of the middle frontal gyrus, motor cortex, supramarginal gyrus, CA1 and the subiculum, but not in the cerebellum.29**

**One limitation of this study is that not all patients were tested for cognitive function. However, the clinical characteristics of tested and non-tested patients were similar, excluding a selection bias. A remarkable feature of our study is its population-based nature, including some 80% of incident patients.**

**In conclusion, in our population-based cohort of patients of Italian ancestry, we found that patients with ALS carrying an intermediate ATXN2 polyQ number of repeats≥31 have a distinctive phenotype, characterised by a more rapid disease course, with a 1.5-fold increase of ΔALSFRS-R rate and a higher ΔKing's. Compared with ATXN2− patients, this greater aggressiveness resulted in a 1 year reduction in survival. In addition, ATXN2+ patients have a more severe impairment of**

**cognitive functions, with relative preservation of the behavioural domain. Identifying the specific phenotypic characteristics of patients with ALS with ATXN2 polyQ intermediate number of repeats has many implications. These include the possibility of refining the individual prognostic prediction and improving the design of ALS clinical trials, including those targeted explicitly to ATXN2.**

## 7.6 References


**1 Goutman SA , Hardiman O , Al-Chalabi A , et al . Emerging insights into the complex genetics and pathophysiology of amyotrophic lateral sclerosis. Lancet Neurol 2022;21:465–79. doi:10.1016/S1474-4422(21)00414-2**

**2 Grassano M , Calvo A , Moglia C , et al . Mutational analysis of known ALS genes in an Italian population-based cohort. Neurology 2021;96:e600–9.doi:10.1212/WNL.0000000000011209**

**3 Chiò A , Calvo A , Moglia C , et al . ATXN2 polyQ intermediate repeats are a modifier of ALS survival. Neurology 2015;84:251–8. doi:10.1212/WNL.0000000000001159**

**4 Sproviero W , Shatunov A , Stahl D , et al . ATXN2 trinucleotide repeat length correlates with risk of ALS. Neurobiol Aging 2017;51:178.e1–178.e9.doi:10.1016/j.neurobiolaging.2016.11.010**

**5 Fournier C , Anquetil V , Camuzat A , et al . Interrupted CAG expansions in ATXN2 gene expand the genetic spectrum of frontotemporal dementias. Acta Neuropathol Commun 2018;6:41.doi:10.1186/s40478-018-0547-8**

**6 Rubino E , Mancini C , Boschi S , et al . ATXN2 intermediate repeat expansions influence the clinical phenotype in frontotemporal dementia. Neurobiol Aging 2019;73:231.e7–231.e9. doi:10.1016/j.neurobiolaging.2018.09.009**

**7 Chiò A , Mora G , Moglia C , et al . Secular trends of amyotrophic lateral sclerosis: the Piemonte and Valle d'Aosta register. JAMA Neurol 2017;74:1097–104. doi:10.1001/jamaneurol.2017.1387**

**8 Brooks BR , Miller RG , Swash M . World Federation of Neurology Research group on motor neuron diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 2000;1:293–9. doi:10.1080/146608200300079536**

**9 Strong MJ , Abrahams S , Goldstein LH . Revised diagnostic criteria. Amyotroph Lateral Scler Frontotemporal Degener 2017;18:153–74. doi:10.1080/21678421.2016.1267768**

**10 Iazzolino B , Peotta L , Zucchetti JP , et al . Differential neuropsychological profile of patients with amyotrophic lateral sclerosis with and without C9orf72 mutation. Neurology 2021;96:e141–52. doi:10.1212/WNL.0000000000011093**

**11 Montuschi A , Iazzolino B , Calvo A , et al . Cognitive correlates in amyotrophic lateral sclerosis: a population-based study in Italy. J Neurol Neurosurg Psychiatry 2015;86:168–73.doi:10.1136/jnnp-2013-307223**

**12 Chiò A , Calvo A , Mazzini L , et al . Extensive genetics of ALS: a population-based study in Italy. Neurology 2012;79:1983–9.doi:10.1212/WNL.0b013e3182735d36**

**13 Renton AE , Majounie E , Waite A , et al . A hexanucleotide repeat expansion in C9orf72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011;72:257–68.doi:10.1016/j.neuron.2011.09.010**

**14 Dolzhenko E , Deshpande V , Schlesinger F , et al . ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics 2019;35:4754–6.doi:10.1093/bioinformatics/btz431**

**15 Chiò A , Mora G , Sabatelli M , et al . ATNX2 is not a regulatory gene in Italian amyotrophic lateral sclerosis patients with C9orf72 GGGGCC expansion. Neurobiol Aging 2016;39:218.e5–218.e8.doi:10.1016/j.neurobiolaging.2015.11.027**

**16 Elden AC , Kim H-J , Hart MP , et al . Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 2010;466:1069–75.doi:10.1038/nature09320**

**17 Nel M , Mavundla T , Gultig K , et al . Repeats expansions in ATXN2, NOP56, NIPA1 and ATXN1 are not associated with ALS in Africans. IBRO Neurosci Rep 2021;10:130–5.doi:10.1016/j.ibneur.2021.02.002**

**18 Zhao M , Kim JR , van Bruggen R , et al . RNA-binding proteins in amyotrophic lateral sclerosis. Mol Cells 2018;41:818–29.doi:10.14348/molcells.2018.0243**

**19 Yang Y , Halliday GM , Kiernan MC , et al . TDP-43 levels in the brain tissue of ALS cases with and without C9ORF72 or ATXN2 gene expansions. Neurology 2019;93:e1748–55.doi:10.1212/WNL.0000000000008439**

**20 Glass JD , Dewan R , Ding J , et al . ATXN2 intermediate expansions in amyotrophic lateral sclerosis. Brain 2022:awac167.doi:10.1093/brain/awac167**

**21 Sorenson EJ , Mandrekar J , Crum B , et al . Effect of referral bias on assessing survival in ALS. Neurology 2007;68:600–2.doi:10.1212/01.wnl.0000254501.58158.e7**

**22 Logroscino G , Marin B , Piccininni M , et al . Referral bias in ALS epidemiological studies. PLoS One 2018;13:e0195821.doi:10.1371/journal.pone.0195821**

**23 Rosas I , Martínez C , Clarimón J , et al . Role for ATXN1, ATXN2, and HTT intermediate repeats in frontotemporal dementia and Alzheimer's disease. Neurobiol Aging 2020;87:139.e1–139.e7.doi:10.1016/j.neurobiolaging.2019.10.017**

**24 Ross OA , Rutherford NJ , Baker M , et al . Ataxin-2 repeat-length variation and neurodegeneration. Hum Mol Genet 2011;20:3207–12.doi:10.1093/hmg/ddr227**

**25 Lattante S , Millecamps S , Stevanin G , et al . Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. Neurology 2014;83:990–5.doi:10.1212/WNL.0000000000000778**

**26 Fitzgerald J , Fahey L , Holleran L , et al . Thirteen independent genetic loci associated with preserved processing speed in a study of cognitive resilience in 330,097 individuals in the UK Biobank. Genes 2022;13:122.doi:10.3390/genes13010122**

**27 Davies G , Marioni RE , Liewald DC , et al . Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). Mol Psychiatry 2016;21:758–67.doi:10.1038/mp.2016.45**

**28 Gigante AF , Lelli G , Romano R , et al . The relationships between ataxia and cognition in spinocerebellar ataxia type 2. Cerebellum 2020;19:40–7.doi:10.1007/s12311-019-01079-5**

**29 Fournier C , Anquetil V , Camuzat A , et al . Interrupted CAG expansions in ATXN2 gene expand the genetic spectrum of frontotemporal dementias. Acta Neuropathol Commun 2018;6:41.doi:10.1186/s40478-018-0547-8**

## 7.7 Tables And Figures

### Table 1

**Demographic and clinical characteristics of patients according to ATXN2 PolyQ intermediate number of repeats (ATXN2+, PolyQ≥31; ATXN2−, PolyQ≤30)**

| | ATXN2+ (n=42) | ATXN2− (n=1288) | P value |
|---|---|---|---|
| Age at onset (median, IQR) | 69.6 (63.5–75.7) | 68.3 (60.3–74.4) | 0.15 |
| Gender (female) | 15 (35.7%) | 575 (44.6%) | 0.25 |
| Site of onset (bulbar) | 7 (16.7%) | 393 (30.5%) | 0.05 |
| Diagnostic delay (months, IQR) | 6.0 (3.94–10.03) | 9.04 (5.88–13.97) | 0.004 |
| ALSFRS-R score at diagnosis (median, IQR) | 42 (34.75–44) | 42 (37–45) | 0.13 |
| ΔALSFRS-R (median points/month, IQR) | 1.00 (0.50–1.99) | 0.66 (0.31–1.33) | 0.004 |
| FVC% at diagnosis (median, IQR)* | 89 (74–105) | 91 (72–104) | 0.81 |
| ΔWeight (kg/month, median, IQR)† | 0.50 (0–1.26) | 0.28 (0–0.96) | 0.87 |
| MiToS stage at diagnosis (0/1/2/3/4) | 27/12/2/1/0 | 858/374/42/10/2 | 0.80 |
| King's state at diagnosis (1/2/3/4) | 14/14/11/3 | 530/409/293/50 | 0.59 |
| ΔKing's (median points/month, IQR) | 0.25 (0.17–0.53) | 0.19 (0.10–0.34) | 0.004 |
| ALS-FTD‡ | 7 (28.0%) | 121 (13.4%) | 0.037 |

- ***FVC, 1222 (ATXN2+, 38; ATXN2−, 1184).**
- **†Weight 1288 (ATXN2+, 40; ATXN2−, 1246).**
- **‡928 cases (ATXN2+, 25, ATXN2−, 903)**

## Table 2

**Frequency of cognitive impairment classified according to the consensus criteria for the diagnosis of frontotemporal cognitive and behavioural syndromes in ALS9**

| | ATXN2+ (n=25) | ATXN2- (n=903) |
|---|---|---|
| Cognitively normal ALS | 10 (40%) | 483 (53.5%) |
| ALSbi | – | 72 (8%) |
| ALSci | 7 (28%) | 164 (18.2%) |
| ALScbi | 1 (4%) | 63 (7%) |
| ALS-FTD | 7 (28%) | 121 (13.4%) |

- In patients with ALS with *ATXN2* PolyQ intermediate number of repeats ≥31 (ATXN2+) compared with patients with PolyQ ≤30 (ATXN2-). ALS-FTD was significantly more frequent in ATXN2+ patients (p=0.037).
- ALS, amyotrophic lateral sclerosis; ALSbi, ALS with behavioural impairment; ALScbi, ALS with cognitive and behavioural impairment; ALSci, ALS with cognitive impairment; ALS-FTD, ALS with comorbid FTD.

**Table 3**

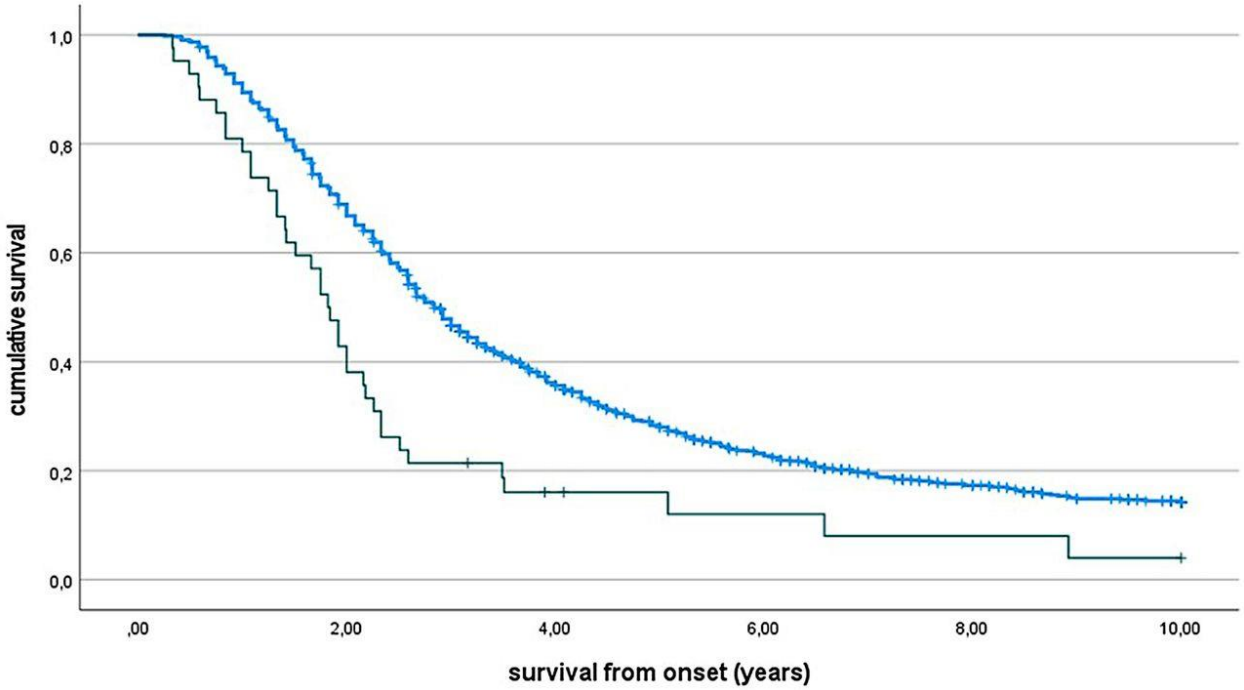
**Patients' survival**

| Variable | Value | HR (95% CI) | P value |
|---|---|---|---|
| Age at onset (years) | Per each year of age at onset | 1.030 (1.024 to 1.037) | 0.0001 |
| Diagnostic delay (months) | Per each month of delay | 0.953 (0.945 to 0.962) | 0.0001 |
| Site of onset | Spinal<br>Bulbar | 1 (reference)<br>1.45 (1.27–1.67) | 0.0001 |
| ΔALSFRS-R | Per each point loss/month | 1.25 (1.20 to 1.29) | 0.0001 |
| ΔKing's | Per each point loss/month | 1.56 (1.16 to 2.09) | 0.003 |
| *ATXN2* polyQ | <31<br>≥31 | 1 (reference)<br>1.58 (1.14 to 2.19) | 0.006 |

- Cox multivariable analysis.

**Figure 1**

**Survival from onset according to ATXN2 polyQ intermediate number of repeats. PolyQ≥31 (green line) versus PolyQ≤30 (blue line). Ticks indicate censored patients. P<0.0001.**




## 7.8 Supplementary

**Supplementary Table 1. Comparison of demographic and genotype data of ALS cases and controls**

| | ALS cases n=1330 | Controls n=1274 | P value |
|---|---|---|---|
| Median age (IQR), years | 68.4 (60.6-74.4) | 66.9 (59.0-71.8) | 0.15 |
| Sex (female, %) | 590 (44.3%) | 590 (46.3%) | 0.32 |
| Median education (IQR), years | 8.0 (5.0-12.0) | 8.2 (4.9-12.2) | 0.32 |
| ATXN2 polyQ ≥31 | 42 (3.3%) | 4 (0.3%) | 0.0001 |

**Supplementary Table 2. Comparison of demographic and genotype data of ALS cases according to the study cohorts**

| | ALS cases – whole cohort (n=1330) | ALS cases-cognitive cohort (n=928) | p |
|---|---|---|---|
| **Age at onset (median, IQR)** | 68.3 (60.4-74.4) | 68.1 (60.3-73.9) | 0.85 |
| **Gender (female)** | 590 (44.4%) | 403 (43.4%) | 0.66 |
| **Site of onset (bulbar)** | 400 (30.1%) | 292 (31.5%) | 0.48 |
| **Diagnostic delay (months, IQR)** | 9.04 (5.10-13.97) | 9.04 (5.11-13.97) | 0.94 |
| **ALSFRS-R score at diagnosis (median,IQR)** | 42 (37-45) | 42 (38-45) | 0.43 |
| **FVC% at diagnosis (median, IQW)** | 91 (72-104) | 93 (74-106) | 0.21 |
| **ΔALSFRS-R (median points/month,IQR)** | 0.67 (0.33-1.34) | 0.60 (0.30-1.18) | 0.12 |
| **ΔWeight (Kg/month, median, IQR)** | 0.28 (0-0.98) | 0.25 (0-0.87) | 0.37 |
| **MiToS stage at diagnosis (0/1/2/3/4)** | 887/386/44/11/2 | 650/252/20/5/1 | 0.30 |
| **King's state at diagnosis (1/2/3/4)** | 550/423/304/53 | 408/304/192/24 | 0.15 |
| **ΔKing's (median points/month, IQR)** | 0.20 (0.10-0.34) | 0.19 (0.10-0.33) | 0.28 |
| **ATXN2 PolyQ (n, %)** | 42 (3.2%) | 25 (2.7%) | 0.61 |

**Supplementary Table 3. Median values (interquartile range) of age-, and education-corrected scores of cognitive tests in ALS patient according to ATXN2 status. 128 patients with FTD (7 ATXN2+ and 121 ATXN2-) have been excluded). P values are calculated with the MannWhitney U test.**

| Test | ATXN- (n=903) | ATXN+ (n=25) | p |
|---|---|---|---|
| **MMSE** | 27.5 (26.2-29.5) n=903 | 28.2 (26.9-30) n=25 | 0.30 |
| **FAS** | 28.8 (22.4-35.9) n=836 | 28.7 (22.4-40.8) n=22 | 0.68 |
| **CAT** | 19.3 (15.5-22.0) n=836 | 18.8 (16.0-25.8) n=22 | 0.51 |
| **FAB** | 14.9 (13.2-16.5) n=778 | 15.2 (13.5-17.8) n=22 | 0.50 |
| **Digit Span FW** | 5.7 (5.0-6.3) n=789 | 5.6 (4.9-6.4) n=19 | 0.79 |
| **Digit Span BW** | 3.9 (3.4-4.5) n=789 | 4.3 (2.6-4.7) n=19 | 0.78 |
| **TMT A** | 39 (25-61) n=781 | 40 (26-76) n=18 | 0.73 |
| **TMT B** | 81 (39-174) n=781 | 74 (43-307) n=18 | 0.74 |
| **TMT B-A** | 44 (12-115) n=781 | 52 (12-1735) n=18 | 0.73 |
| **RAVL-IR** | 39.2 (32.9-45.5) n=569 | 37.0 (34.1 -39.1) n=14 | 0.18 |
| **RAVL-DR** | 6.0 (4.0-9.0) n=569 | 3.0 (2.0-6.0) n=14 | **0.023** |
| **BSRT-IR** | 5.6 (4.5-6.8) n=543 | 4.6 (3.5-5.4) n=14 | **0.044** |
| **BRRT-DR** | 6.5 (5.0-7.5) n=543 | 5.4 (4.3-7.2) n=14 | 0.07 |
| **ROCF-IR**<br>**ROCF-DR** | 31.5 (26.3-34.3) n=676<br>11.7 (7.7-16.4) n=676 | 31.9 (27.4-34.3), n=15<br>12.2 (6.9-16.8) n=15 | 0.95<br>0.31 |
| **CPM47** | 28.5 (24.2-31.5) n=823 | 30.0 (26.5-33.0) n=15 | 0.20 |
| **Clock** | 4 (3-5) n=755 | 4 (3-5) n=22 | 0.78 |
| **SET-IA** | 4.2 (3.2-5.3) n=150 | 3.2 (1.5-5.0) n=6 | 0.30 |
| **SET-CI** | 4.3 (3.3-5.2) n=150 | 4.7 (2.6-6.0) n=6 | 0.75 |
| **SET-EA** | 4.2 (3.2-5.9) n=150 | 3.3 (2.5-4.0) n=6 | 0.18 |
| **SET-GS** | 13.1 (9.3-15.3) n=150 | 10.4 (6.6-14.8) n=6 | 0.38 |
| **HADS-A** | 7 (5-10) n=795 | 10 (2-14) n=16 | 0.43 |
| **HADS-D** | 5 (3-8) n=795 | 9 (3.8-11) n=16 | **0.04** |

*** The scores of the Clock test are not corrected by age and education because no Italian normative data are available**

**MMSE, Mini-Mental State Examination; FAS, Letter Fluency test; CAT, Category Fluency Test; FAB, Frontal Assessment Battery; Digit Span FW, Digit Span Forward; Digit Span BW, Digit Span Backward; TMT A, Trail Making Test A; TMT B, Trail Making Test B; TMT B-A, Trail Making Test B-A; RAVL-IR, Rey Auditory Verbal Learning Test, Immediate Recall; RAVLDR, Rey Auditory Verbal Learning Test, Delayed Recall; BSRT-IR, Babcock Story Recall Test, Immediate Recall; BSRT-DR, Babcock Story Recall Test, Delayed Recall ; ROCF-IR,**

**Babcock Story Recall Test, Immediate Recall ; ROCF-DR, Babcock Story Recall Test, Delayed Recall ; CPM47, Raven's Colored Progressive Matrices; Clock, Clock Drawing Test; SET-IA, Story-Based Empathy Task – Intention Attribution; SET-CI, Story-Based Empathy Task – Causal Inference; SET-EA, Story-Based Empathy Task – Emotion Attribution; SET-GS, Story-Based Empathy Task – Global Score; HADS-A, Hospital Anxiety and Depression Scale - Anxiety; HADS-D, Hospital Anxiety and Depression Scale - Depression**

**Supplementary Table 4. Median values (interquartile range) of age-, and education-corrected scores of ECAS in ALS patient according to ATXN2 status. Patients with FTD have been excluded. P values are calculated with the Mann-Whitney U test.**

| **Domains** | **ATXN2-** **(n=291)** | **ATXN2+** **(n=6)** | **p** |
|---|---|---|---|
| **Language** | 25 (22-27) | 25.5 (24.2-27.5) | 0.44 |
| **Fluency** | 16 (12-18) | 15 (14-17.5) | 0.92 |
| **Executive** | 30 (22-38) | 20.5 (20-24) | **0.034** |
| **Memory** | 17 (14-19) | 14.5 (13.2-16.5) | 0.061 |
| **Visuospatial** | 12 (11-12) | 11.5 (9.5-12) | 0.67 |
| **ALS specific functions** | 70.5 (57-79) | 62 (60-66) | 0.11 |
| **ALS non-specific functions** | 28 (25-31) | 26 (23.5-27.5) | 0.25 |
| **ECAS total score** | 97 (82-110) | 87.5 (84.8-93.3) | 0.19 |

**Supplementary Table 5. List of extracted genes**

| **GENES** | | | |
|---|---|---|---|
| ***ALS2*** | ***CHMP2B*** | ***MATR3*** | ***SPAST*** |
| ***ANG*** | ***DAO*** | ***NEFH*** | ***SPG11*** |
| ***ANXA11*** | ***DCTN1*** | ***NEK1*** | ***SS18LL1*** |
| ***ATXN2*** | ***ERBB4*** | ***OPTN*** | ***TAF15*** |
| ***C21orf2*** | ***FIG4*** | ***PFN1*** | ***TBK1*** |
| ***C9orf72*** | ***FUS*** | ***PRPH*** | ***TUBA4A*** |
| ***CAMTA1*** | ***HNRNPA1*** | ***PGRN*** | ***UBQLN2*** |

| | | | |
|---|---|---|---|
| ***CCNF*** | ***HNRNPA2B1*** | ***SETX*** | ***UNC13A*** |
| ***CEP112*** | ***KIF5A*** | ***SIGMAR1*** | ***VAPB*** |
| ***CHCHD10*** | ***MAPT*** | ***SOD1*** | ***VCP*** |

**Supplementary Table 6. Allele frequencies and predicted functional effects of identified genetic variants**

| Gene | AAChange | dbSNP | gnomAD NFE | SIFT | Polyphen2 | LRT | Mutation Taster | FATHMM | MetaSVM | MetaLR | CADD |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *CCNF* | exon5 c.G368A p.R123H | rs371699142 | 4.52E-05 | T | P | D | D | T | T | T | 25.1 |
| *DCTN1* | exon11 c.C1061T p.A354V | . | . | D | P | D | D | T | D | D | 28.4 |
| *EWSR1* | exon5 c.G394T p.A132S | . | . | T | B | U | D | T | T | T | 24.2 |
| *OPTN* | exon3 c.T332G p.L111R | . | . | D | B | N | D | D | D | D | 26.4 |
| *SETX* | exon10 c.A1427G p.H476R | rs779742691 | 9.34E-06 | D | D | D | D | D | D | D | 24.8 |
| *SETX* | exon10 c.G2755A p.V919I | . | 0 | D | B | N | N | D | T | T | 10.35 |

**SIFT: D, Deleterious; T, tolerated: PolyPhen-2: D, Damaging; P, Possibly Damaging; B, Benign; LRT: D, Deleterious; N, neutral; U, Unknown; Mutation Taster, D, Deleterious; N, Neutral; FATHMM: D, Deleterious; T, Tolerated: MetaSVM: D, Deleterious; T, Tolerated; MetaLR: D, Deleterious; T, Tolerated**

**Supplementary Table 7. Demographic and clinical characteristics of patients with ≥31 ATXN2 PolyQ intermediate number of repeats according to the presence of mutations in other genes**

| | **With other gene mutations (n=6)** | **Without other gene mutations (n=34)** | **p** |
|---|---|---|---|
| Age at onset (median, IQR) | 71.5 (65.5-77.9) | 68.3 (62.9-75.8) | 0.52 |
| Gender (female) | 2 (33.3%) | 12 (35.3%) | 0.65 |
| Site of onset (bulbar) | 1 (14.7%) | 5 (16.7%) | 0.65 |
| Diagnostic delay (months, IQR) | 5.7 (2.5-10.5) | 6.0 (3.9-9.9) | 0.62 |
| ALSFRS-R score at diagnosis (median, IQR) | 40 (30-45) | 42 (35-44) | 0.96 |
| ΔALSFRS-R (median points/month, IQR) | 1.37 (0.55-4.41) | 1.00 (0.49-1.85) | 0.45 |
| FVC% at diagnosis (median, IQR) * | 77 (51-98) | 89 (79-106) | 0.19 |
| ΔWeight (Kg/month, median, IQR) § | 1.94 (0.4-3.10) | 0.29 (0-1.03) | 0.026 |
| MiToS stage at diagnosis (0/1/2/3/4) | 5/0/1/0 | 21/11/1/1 | 0.22 |
| King's state at diagnosis (1/2/3/4) | 2/1/2/1 | 12/12/8/2 | 0.67 |
| ΔKing's (median points/month, IQR) | 0.49 (0.19-1.10) | 0.25 (0.17-0.51) | 0.31 |
| ALS-FTD ^ | 1 (33.3%) | 5 (23.8%) | 0.037 |
| Median survival (years, IQR) | 1.08 (0.84-2.26) | 1.82 (1.25-2.59) | 0.30 |

* **available for 36 patients (5 with other mutations and 31 without other mutations)**

§ **available for 38 patients (6 with other mutations and 32 without other mutations)**

^ **available for 24 patients (3 with other mutations and 21 without other mutations)**

# 7. Association of **Intermediate HTT** CAG Repeat Sizes With Increased **Disease** Risk **and Shorter Survival in** Amyotrophic Lateral Sclerosis

## 8.1 Abstract

**Recent research has revealed an intriguing connection between trinucleotide CAG repeat expansions in the Huntingtin gene (HTT) and Amyotrophic Lateral Sclerosis (ALS). However, whether HTT expansions in the intermediate range contribute to ALS risk has not been fully investigated. In this study, we investigate the occurrence of intermediate HTT repeats in a cohort of 904 ALS patients and 757 healthy controls and evaluate their potential role in ALS pathogenesis. We revealed that 6.80% of ALS cases and 4.62% of controls possessed intermediate-length HTT CAG repeats, indicating a slightly increased risk of ALS among intermediate-repeat carriers. Meta-analysis of previously published studies further confirmed the association between intermediate HTT expansions and ALS risk. Phenotypic analysis showed that ALS patients carrying intermediate HTT expansions exhibited an older age at onset, accelerated disease progression, shorter survival, and increased diagnostic delay compared to non-carrier patients. These findings provide compelling evidence supporting the pathogenic role of intermediate HTT expansions in ALS. The results highlight the increased risk and more severe disease course associated with intermediate HTT repeats, underscoring the importance of genetic factors in ALS pathogenesis. The study contributes to our understanding of ALS and opens new avenues for targeted therapies and interventions. Further replication in larger cohorts and diverse populations is warranted to validate these findings and elucidate the underlying mechanisms.**

## 8.2 Introduction

**Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disorder characterized by the progressive degeneration of motor neurons[1]. The genetic landscape of ALS is highly complex, with numerous contributing factors identified, including more than 40 genes associated with the disease[2]. Recent research has suggested a compelling connection between ALS and Huntington's disease, an inherited neurodegenerative condition caused by a trinucleotide CAG repeat expansion within exon 1 of the Huntingtin gene (HTT; MIM:613004) [3,4].**

**In fact, the same CAG repeat expansions observed in the HTT gene have been implicated in patients with frontotemporal dementia/amyotrophic lateral sclerosis (FTD/ALS). These patients exhibit fully**

**penetrant, pathogenic expansions consisting of 40 or more CAG repeats, while healthy subjects typically have 26 or fewer CAG repeats. Notably, carriers of repeats in the range from 27 to 35 may be at risk of transmitting a pathogenic allele (≥40 repeats) due to instability in the CAG tract [5]. Moreover, intermediate expansions in other polyglutamine repeat expansions, particularly in the ataxin 2 gene (ATXN2), have been identified as a known risk factor for developing ALS [6].**

**Given these intriguing observations, we conducted a comprehensive analysis to investigate the occurrence of intermediate HTT repeats in a cohort of Italian ALS patients and assess their potential role in the pathogenesis of ALS.**

## 8.3 Methods

### Study population

**The study population consists of 904 patients diagnosed with ALS from 2009 to 2016 identified through the Piemonte and Valle d'Aosta Register for ALS. The characteristics of the register have been reported elsewhere[7], and these samples were part of the original report of *HTT* repeat expansions in patients with FTD/ALS.[3] Briefly, patients met the El Escorial revised diagnostic criteria for definite, probable and probable laboratory-supported ALS.[8] ALSFRS-R mean monthly decline (ΔALSFRS-R) was calculated using the following formula: (48 – ALSFRS-R score at diagnosis)/(months from onset to diagnosis). As a proxy of disease spread at diagnosis, we calculated the mean/monthly decline of King's staging as (King's staging at diagnosis)/(months from onset to diagnosis) was calculated.[9] A total of 965 patients (72.9%) underwent an extensive cognitive battery at the time of diagnosis and were classified according to the Consensus Criteria for diagnosing frontotemporal cognitive and behavioral syndromes in ALS.[10] Characteristics of cognitive testing are extensively described elsewhere.[11]**

### Controls

**A total of 757 controls were age, sex and geographically matched Italian subjects were included in the study. These participants were identified through patients' general practitioners. They were neurologically normal and had no reported family history of neurodegenerative diseases. The control samples were also part of the original report of *HTT* repeat expansions in patients with FTD/ALS.[3]**

### Genetic analysis

**All patients included in the study were tested for mutations in *SOD1*, *TARDBP, FUS* and other major ALS genes, as previously reported[12,13]. The presence of the *C9orf72* intronic expansion was evaluated using the methods described elsewhere.[12,14] We also tested for *ATXN2* CAG expansions and *UNC13A* rs12608932 A/C, as recently reported.[13,15,16]**

### HTT CAG repeat analysis

***HTT* CAG repeats were detected in both cases and controls through the whole-genome sequencing methods described elsewhere.[3,13] We used ExpansionHunter—Targeted software (V.3.0.1) to estimate repeat lengths of *HTT* expansions as previously described.[17] This algorithm has been validated using experimentally confirmed samples carrying the *C9orf72* repeat expansions. We previously demonstrated that ExpansionHunter correctly identified the expansions in C9orf72 or ATXN2 genes present in our cohort.[3,13,15] *HTT* CAG repeats were deemed full penetrant if ≥ 40, incompletely penetrant if 36-39 and intermediate if they were in the 27–35 range.**

### Statistical methods

**The association of *HTT* CAG expansions with disease risk was assessed using logistic regression analysis. Estimates from the model for the contribution of *HTT* intermediate repeats were converted to odds ratio by exponentiating the estimates. The logistic models were also corrected for sex and age at DNA collection. We assessed the effect of *HTT* CAG expansion on age at disease onset, cognitive impairment and disease progression at diagnosis using appropriate linear or logistic regression models. Survival was calculated from onset to death, tracheostomy, or censoring date (December 31, 2021), using the Kaplan-Meier method, and compared with the log-rank test. No patients were lost to follow-up. Multivariable survival analysis was performed with the Cox proportional hazards model (stepwise backward) with a retention criterion of p value<0.1. We included sex, age and site of symptoms onset whenever necessary as covariates. We also included the diagnostic delay, the ALSFRS-R mean monthly decline and the *UNC13A* rs12608932 A/C status as a covariate in the survival analysis. The significance level was set for all models at $p<0.05$. Statistical analyses were carried out using R v4.0 and the glm() function, the logisdf package (v.1.23) and the survfit() or coxph() function of the survival package.**

### Inclusion criteria for published studies

**Literature review and meta-analysis were conducted in accordance with the PRISMA[18] (Preferred Reporting Items for Systematic reviews and Meta-Analyses) group guidelines. We included case-control studies that evaluated whether intermediate expansion of the *HTT* gene conferred risk for**

**ALS. Series of cases and descriptive reports evaluating *HTT* expansions in the incomplete penetrance range (repeats of size 36 to 39) were excluded from this study.**

### Meta-analysis

**Mantel–Haenszel method meta-analysis of odds ratios (ORs) was performed using the 'metafor' 2.0 package. Additionally, we combined our data with the data extracted from different studies in a single logistic regression analysis (joint analysis) with cohort and method of genotyping as fixed-effects covariates.**

## 8.4 Results

### Study population

**Our study cohort consisted of 1324 ALS cases: 730 males (55.1%) and 594 females (44.9%). The mean age at disease onset was 66.2 years, without differences between men (65.8 years) and women (66.7 years). One-third of cases reported their initial symptoms to be in the bulbar musculature (n=407, 31.5%); as expected, bulbar onset was more common in women (41.2%) than men (23.6%). The *C9ORF72* repeated expansion was detected in 92 patients (6.9%). Additionally, 51 ALS cases (3.1%) had a pathogenic variant in *SOD1, TARDBP* or *FUS*. The demographic characteristics of the cohort are summarized in the previous chapters.**

### HTT intermediate expansion and ALS risk

**As previously reported, we did not observe an *HTT* polyQ repeat length greater than 40 in controls or ALS patients. Two healthy individuals carried incomplete penetrance CAG alleles (35-39 repeats, 0.26% of the control cohort), and three (0.22%) ALS cases fell within this range.**

**We found that 90 of 1324 ALS cases (6.80%) harbored an intermediate-length *HTT* CAG, whereas 35 of 757 healthy controls (4.62%) possessed an intermediate-length *HTT* CAG repeat, corresponding to a slightly increased risk of ALS among intermediate-length HTT repeats carriers (OR = 1.5, 95% C.I. = 1.0–2.3, *P* = 0.0459). Eight ALS cases carried both a *C9ORF72* pathogenic expansion and an intermediate-length *HTT* repeat (0.6%). Additionally, three ALS cases with known pathogenic mutations (*SOD1*:p.L144F, *TARDBP*:p.G295S and *FUS*:p.R495X) also carried an intermediate-length *HTT* repeat. The analysis excluding all ALS cases mutated in *C9ORF72, SOD1, TARDBP* or *FUS* revealed a similar increased risk (OR = 1.5, 95% C.I. = 0.99–2.3, *P* = 0.0590). These data suggest that intermediate-length *HTT* CAG repeat expansions are nominally associated with an increase in ALS risk.**

### Meta-analysis of previously published studies

**Based on literature searches, we identified four case-control studies examining intermediate range *HTT* CAG expansion in ALS (Supplementary Materials)[19,20,21]. The frequencies of intermediate expansions carriers were extracted for both cases and controls from the selected papers. The random-effect meta-analysis of these four cohorts indicated an association between the presence of intermediate expansions in *HTT* and ALS status with an OR = 1.8 (95% CI = 1.4–2.2, P <0.0001; Fig. 1). In addition, joint analysis on the combined data of all four datasets confirmed the increased risk related to HTT intermediate expansions (OR = 1.8 (95% CI = 1.5–2.2, P <0.0001).**

### HTT intermediate expansion effect on ALS phenotype

***HTT* intermediate expansion carriers had a significantly older age at onset than non-carriers (66.0 years in those ≤26 versus 69.6 years in those ≥27, difference = 3.6 years, *P* = 0.0023). No differences in the type of disease onset (spinal versus bulbar) were observed comparing carrier and non-carrier ALS patients (*P* = 0.942; Table 1). Fourteen (22.6%) patients in this *HTT* intermediate group had ALS-FTD compared with 139 (15.3%) ALS-FTD among non-carrier patients. However, logistic regression shows that despite the higher rate of FTD, *HTT* intermediate expansions carriers did not exhibit a higher burden of cognitive impairment (OR = 1.49, 97.5% C.I. = 0.76-2.75, *P* = 0.2180), including when accounting for sex, age and site of disease onset.**

**Regarding disease severity, patients carrying intermediate *HTT* expansions had a shorter diagnostic delay (-2.7 months, 97.5% C.I. = -5.1 – -0.4, *P* = 0.0218) and a faster progression rate measured by ΔALSFRS-R (+0.2 points per months, 97.5% C.I. = 0.0 – 0.4, *P* = 0.0398) and ΔKing's (+0.06, 97.5% C.I. = 0.00 – 0.12, *P* = 0.0835) (Table 1 and Figure 1). We also observed that patients carrying intermediate expansions had shorter survival compared with non-expanded patients (median survival time for *HTT*-carrier ALS patients = 33 months, 95% CI = 31–35; median survival time for non-carriers ALS patients = 27 months, 95% C.I. = 22-29; p=0.0017, Figure 2). Cox multivariable analysis confirmed that the *HTT* CAG intermediate number of repeats was independently related to a worse outcome (HR = 1.4, 97.5% C.I. = 1.1 – 1.8, *P* = 0.0018, Table 1).**

## 8.5 Discussion

**In this study, we investigated the role of *HTT* expansions within the intermediate range in ALS. Our findings support the association between intermediate-range *HTT* expansions and increased risk and poorer prognosis in ALS patients. The prevalence of *HTT* intermediate expansions in our ALS cohort aligns with previous studies [19, 20], while the frequency of intermediate HTT repeat expansions**

**in our control cohort is consistent with what has been reported in the general population [22]. Importantly, our meta-analysis of different cohorts further confirms that intermediate-length polyglutamine expansions in *HTT* significantly elevate the risk of ALS.**

**In addition to the increased risk, we observed a more severe disease course among ALS patients carrying intermediate repeat expansions, characterized by accelerated progression and reduced survival rates. These findings provide compelling evidence supporting the pathogenic role of HTT expansions in the pathogenesis of FTD/ALS. Indeed, the clinical severity observed in patients carrying HTT intermediate repeats further suggests their contribution to disease pathogenesis.**

**While interpreting our data, we acknowledge the limitation of a relatively small number of cases available for phenotype analysis. Therefore, replication of our results in larger cohorts, encompassing diverse ancestries, will be essential to establish the pathogenic role of intermediate HTT repeat expansions in ALS and their impact on the phenotype. However, it is worth noting that our findings are supported by orthogonal biological evidence.**

**Pathological CAG expansions in *HTT* can induce TDP-43 and huntingtin co-aggregation in vitro, likely through a direct interaction between poly-Q residues and the TDP-43 C-terminal domain.[23] This interaction leads to the sequestration of TDP-43 into polyglutamine aggregates, disrupting TDP-43-mediated splicing in the nucleus. Thus, HTT expansion serves as an additional example of the pathogenic link connecting poly-Q expansions and TDP-43 proteinopathies, including ALS/FTD.[23]**

**Our study sheds light on the impact of intermediate range HTT expansions on ALS, revealing their association with increased disease risk. Furthermore, we demonstrate that ALS patients carrying intermediate HTT CAG expansions (≥27) exhibit a distinct phenotype characterized by a more severe disease course. By establishing the association between intermediate HTT expansions and a more severe ALS phenotype, our study underscores the importance of genetic factors in determining the natural history of ALS. These findings deepen our understanding of ALS and provide a foundation for future research on targeted therapies. The identification of intermediate HTT expansions as risk factors opens new avenues for intervention and treatment strategies, offering hope for improved outcomes for ALS patients.**

**In summary, our study contributes to the growing body of evidence linking intermediate HTT expansions to ALS pathology and emphasizes the crucial role of genetics in shaping disease risk and progression. This knowledge paves the way for the development of innovative therapeutic approaches aimed at mitigating the impact of ALS.**

## 8.6 References


**1. Feldman EL, Goutman SA, Petri S, Mazzini L, Savelieff MG, Shaw PJ, Sobue G. Amyotrophic lateral sclerosis. Lancet. 2022 Oct 15;400(10360):1363-1380. doi: 10.1016/S0140-6736(22)01272-7. Epub 2022 Sep 15.**

**2. Suzuki, N., Nishiyama, A., Warita, H. et al. Genetics of amyotrophic lateral sclerosis: seeking therapeutic targets in the era of gene therapy. J Hum Genet 68, 131–152 (2023). https://doi.org/10.1038/s10038-022-01055-8**

**3. Dewan R, Chia R, Ding J, et al. Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. Neuron. 2021;109(3):448-460.e4. doi:10.1016/j.neuron.2020.11.005**

**4. Hickman RA, Dewan R, Cortes E, Traynor BJ, Marder K, Vonsattel JP. Amyotrophic lateral sclerosis is over-represented in two Huntington's disease brain bank cohorts: further evidence to support genetic pleiotropy of pathogenic HTT gene expansion. Acta Neuropathol. 2022;143(1):105-108. doi:10.1007/s00401-021-02385-1**

**5. Semaka A, Kay C, Doty C, Collins JA, Bijlsma EK, Richards F, Goldberg YP, Hayden MR. CAG size-specific risk estimates for intermediate allele repeat instability in Huntington disease. J Med Genet. 2013 Oct;50(10):696-703. doi: 10.1136/jmedgenet-2013-101796. Epub 2013 Jul 29. PMID: 23896435.**

**6. Neuenschwander, A. G., Thai, K. K., Figueroa, K. P., & Pulst, S. M. (2014). Amyotrophic lateral sclerosis risk for spinocerebellar ataxia type 2 ATXN2 CAG repeat alleles: a meta-analysis. JAMA neurology, 71(12), 1529-1534.**

**7. Chiò A, Mora G, Moglia C, et al. Secular Trends of Amyotrophic Lateral Sclerosis. JAMA Neurology. 2017;74(9). doi:10.1001/jamaneurol.2017.1387**

**8. Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders. 2000;1(5). doi:10.1080/146608200300079536**

**9. Roche JC, Rojas-Garcia R, Scott KM, et al. A proposed staging system for amyotrophic lateral sclerosis. Brain. 2012;135(3). doi:10.1093/brain/awr351**

**10. Strong MJ, Abrahams S, Goldstein LH, et al. Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. 2017;18(3-4). doi:10.1080/21678421.2016.1267768**

**11. Chiò A, Moglia C, Canosa A, et al. Cognitive impairment across ALS clinical stages in a population-based cohort. Neurology. 2019;93(10):e984-e994. doi:10.1212/WNL.0000000000008063**

**12. Grassano M, Calvo A, Moglia C, et al. Mutational analysis of known ALS genes in an Italian population-based cohort. Neurology. Published online November 2020. doi:10.1212/WNL.0000000000011209**

**13. Grassano M, Calvo A, Moglia C, et al. Systematic evaluation of genetic mutations in ALS: a population-based study. J Neurol Neurosurg Psychiatry. Published online July 27, 2022:jnnp-2022-328931. doi:10.1136/jnnp-2022-328931**

**14. Chiò A, Calvo A, Mazzini L, et al. Extensive genetics of ALS: a population-based study in Italy. Neurology. 2012;79(19):1983-1989. doi:10.1212/WNL.0b013e3182735d36**

**15. Chio A, Moglia C, Canosa A, et al. Exploring the phenotype of Italian patients with ALS with intermediate ATXN2 polyQ repeats. J Neurol Neurosurg Psychiatry. Published online August 25, 2022:jnnp-2022-329376. doi:10.1136/jnnp-2022-329376**

**16. Calvo A, Canosa A, Moglia C, et al. Clinical and Metabolic Signature of UNC13A rs12608932 Variant in Amyotrophic Lateral Sclerosis. Neurol Genet. 2022;8(6):e200033. doi:10.1212/NXG.0000000000200033**

**17. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics. 2019;35(22):4754-4756. doi:10.1093/bioinformatics/btz431**

**18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71. PMID: 33782057; PMCID: PMC8005924.**

**19. Manini A, Gagliardi D, Meneri M, et al. Analysis of HTT CAG repeat expansion in Italian patients with amyotrophic lateral sclerosis. Ann Clin Transl Neurol. 2022;9(11):1820-1825. doi:10.1002/acn3.51673**

**20. Ramos EM, Keagle P, Gillis T, et al. Prevalence of Huntington's disease gene CAG repeat alleles in sporadic amyotrophic lateral sclerosis patients. Amyotroph Lateral Scler. 2012 May;13(3):265-9. doi: 10.3109/17482968.2011.653573. Epub 2012 Mar 13. PMID: 22409360.**

**21. Lee T, Li YR, Chesi A, Hart MP, et al.. Evaluating the prevalence of polyglutamine repeat expansions in amyotrophic lateral sclerosis. Neurology. 2011 Jun 14;76(24):2062-5. doi: 10.1212/WNL.0b013e31821f4447. Epub 2011 May 11. PMID: 21562248; PMCID: PMC3111240.**

**22. 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. 2019;76(6):650–656. doi:10.1001/jamaneurol.2019.0423**

**23. Fuentealba RA, Udan M, Bell S, et al. Interaction with Polyglutamine Aggregates Reveals a Q/N-rich Domain in TDP-43. Journal of Biological Chemistry. 2010;285(34):26304-26314. doi:10.1074/jbc.M110.125039**

## 8.7 Tables And Figures

### Table 1.

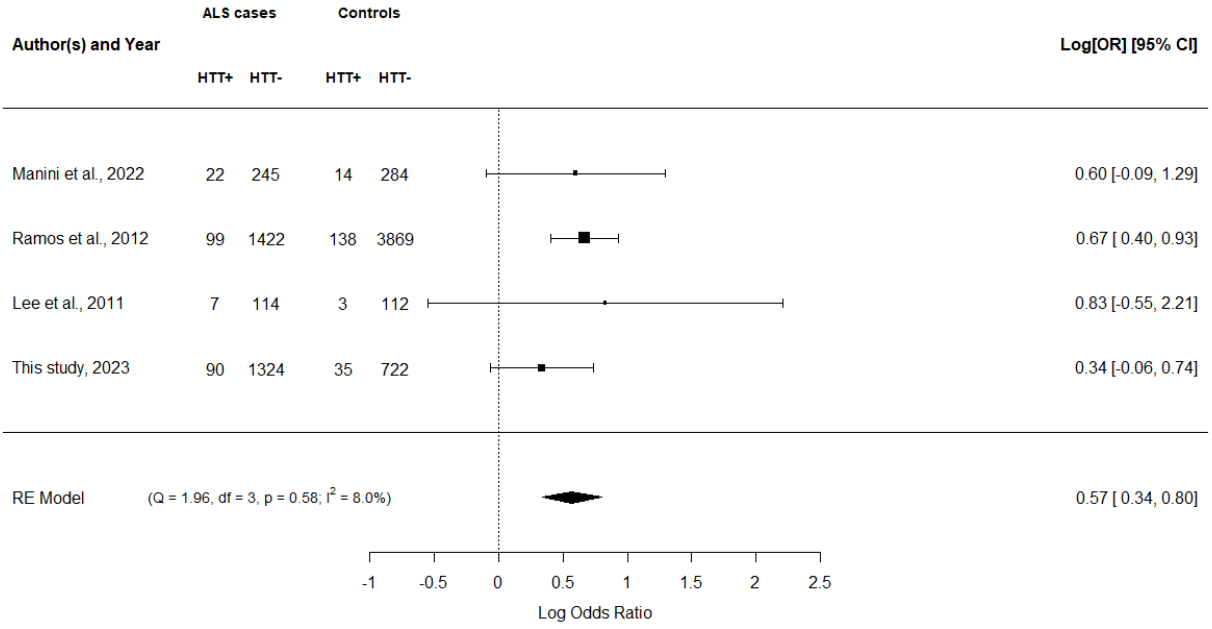
**Demographic and clinical characteristics of patients according to *HTT* CAG intermediate number of repeats. HTT+, CAG 27-35; HTT−, CAG ≤26.**

| | | HTT+ | HTT- | p-value |
|---|---|---|---|---|
| **Age at disease onset (years, SD)** | | 69.6 (10.4) | 66.0 (11.4) | **0.0023** |
| **Sex (F,%)** | | 36 (40.0%) | 557 (45.1%) | 0.3910 |
| **Site of onset (Bulbar,%)** | | 31 (34.4%) | 386 (31.2%) | 0.4870 |
| **Cognitive** | **Normal (%)** | 29 (46.8%) | 464 (51.4%) | **0.0219** |
| | **ALS-bi (%)** | 5 (8.1%) | 75 (8.3%) | |
| | **ALS-ci (%)** | 11 (17.7%) | 167 (18.5%) | |
| | **ALS-cbi (%)** | 3 (4.8%) | 60 (6.7%) | |
| | **ALS-FTD (%)** | 14 (22.6%) | 138 (15.3%) | |
| **Diagnostic delay (months, mean, SD)** | | 9.3 (6.7) | 12.0 (11.13) | **0.0210** |
| **ΔALSFRS-R (points/mo, mean, SD)** | | 1.13 (1.02) | 0.93 (0.87) | **0.0398** |
| **ΔKing's (points/mo, mean, SD)** | | 0.34 (0.38) | 0.28 (0.30) | **0.0835** |
| **Survival (years, median, IQR)** | | 2.3 (1.2-3.6) | 2.8 (1.7-5.3) | **0.0017** |

### Figure 1.

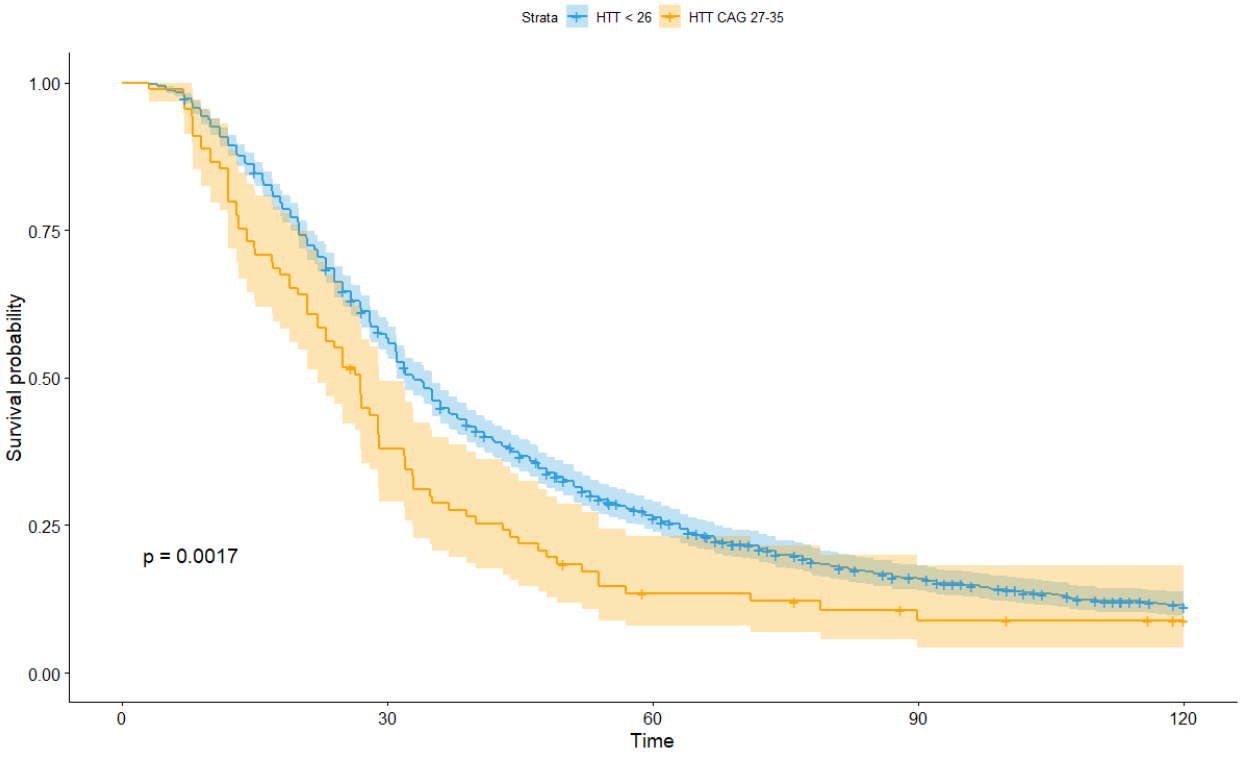
**HTT intermediate repeat expansion meta-analysis. Forest plot for the random-effect Mantel–Haenszel meta-analysis of the effect of presence of an intermediate expanded (27-35) HTT CAG repeats on ALS risk in four different cohorts. In addition, individual-level data of all four datasets were combined in a joint logistic regression analysis, which was corrected for the country of origin and method of genotyping. Weights depending on the number of participants. CI, confidence interval. *Ramos et al. used the normal chromosome (the one with the shorter nonexpanded**

**repeat, i.e. with less than 35 HTT CAG repeats) of large cohort of patients with HD as the control group, based on the assumption that it would reflect the population distribution of the normally polymorphic HTT CAG allele. Since this control cohort differed from the other studies, we performed a sensitivity analysis excluding the cases from Ramos et al.: results from the meta-analysis and the joint analysis confirmed the statistical association of intermediate HTT expansions with disease risk (Appendix).**



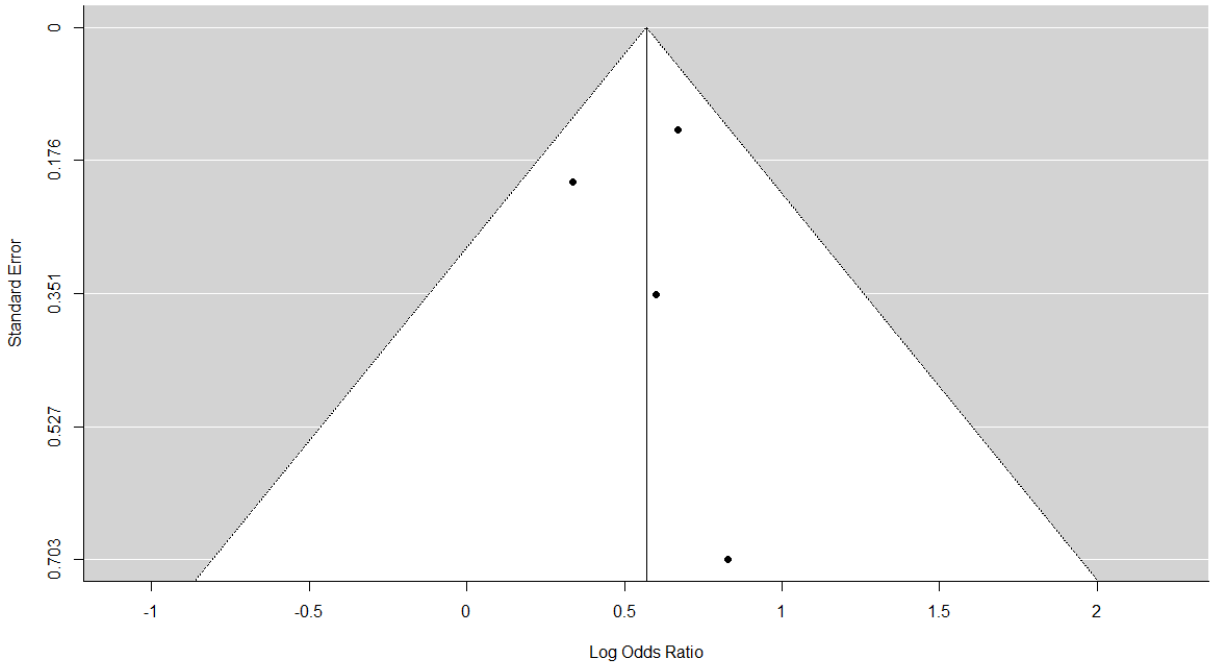
## Figure 2.

**Survival from onset according to HTT CAG intermediate number of repeats. CAG 27-35 (orange line) versus CAG ≤26 (blue line).**



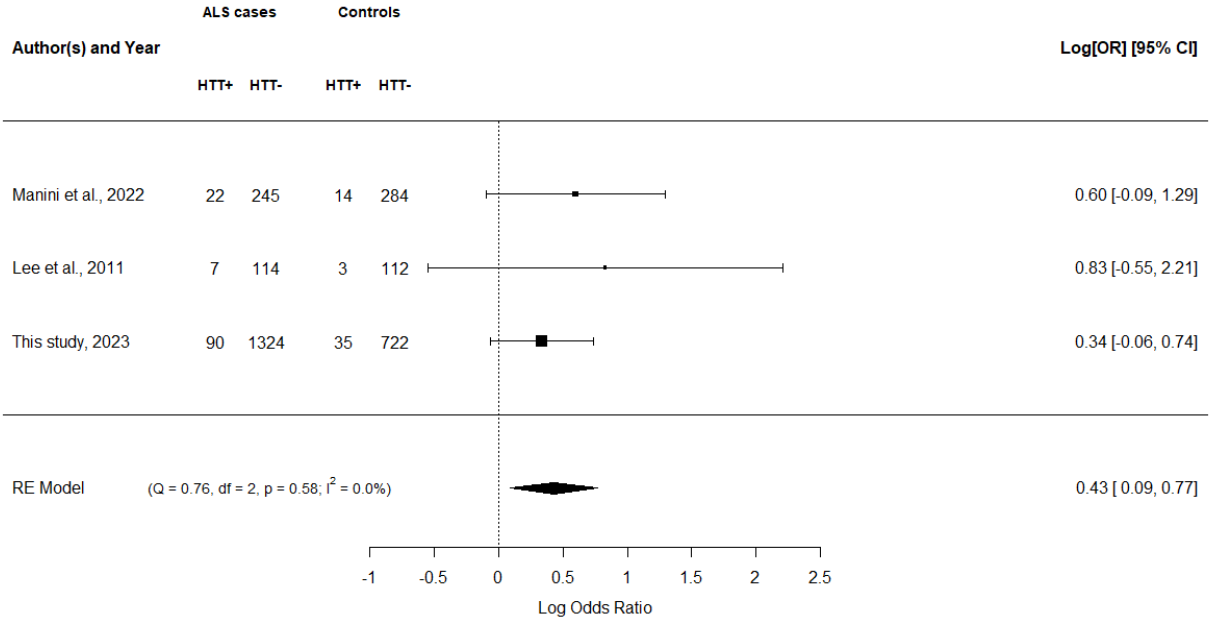
## 8.8 Appendix

**Appendix Figure 1. Funnel plot of the random-effect Mantel–Haenszel meta-analysis**

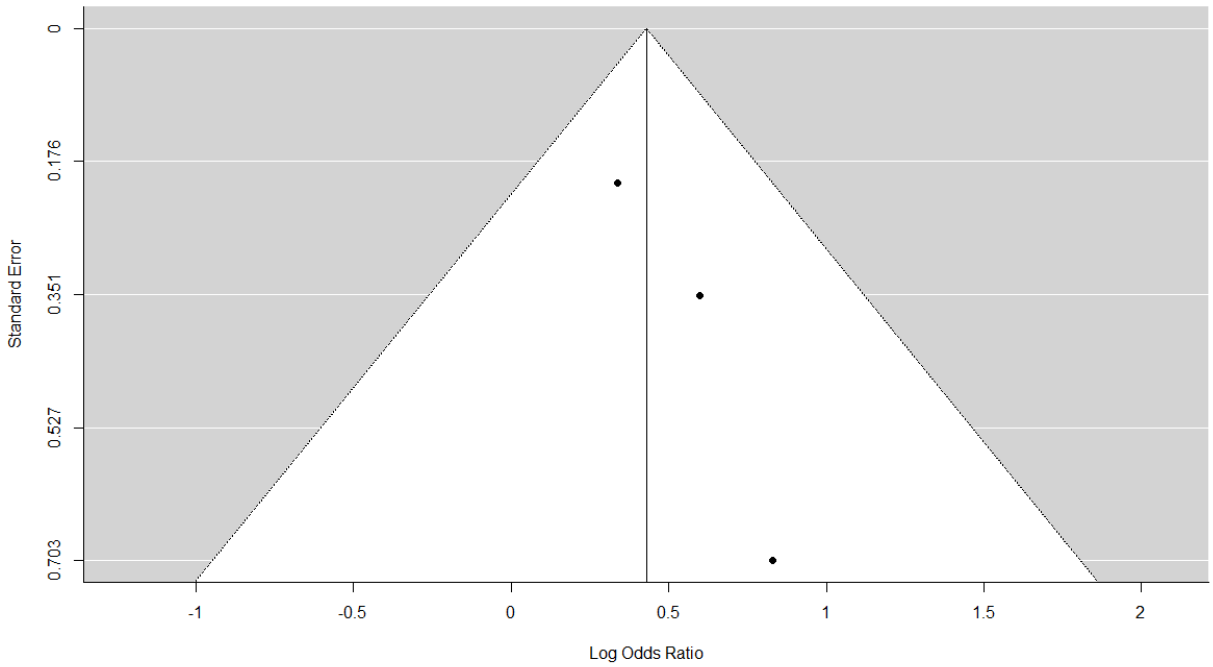


Appendix Figure 2. Forest plot of the sensitivity random-effect Mantel–Haenszel meta-analysis, performed excluding the study from Ramos **et al., 2012.**

**The random-effect meta-analysis of the three remaining cohorts showed that the association between intermediate expansions in *HTT* and increased ALS risk persisted (OR , 95% CI = 1.4–2.2, P=0124). The joint analysis on the combined data of all three datasets confirmed the increased risk related to HTT intermediate expansions (OR = 1.5 (95% CI = 1.1–2.2, P=0.0131)**



**Appendix Figure 3. Funnel plot of the sensitivity random-effect Mantel–Haenszel meta-analysis, performed excluding the study from Ramos et al., 2012.**



# 8. DISCUSSION

## 9.1 Overview

**Within this Dissertation, I have comprehensively characterized the genetic determinants of the population-based cohort from the Piemonte and Valle d'Aosta Register for ALS (PARALS). Throughout this work I analyzed the data I produced using whole-genome sequencing to elucidate the genetic contributors to ALS diagnoses and prognosis. More specifically, I assessed the contribution of both common and rare variants, and evaluated the utility of short tandem repeat analysis from WGS data in a clinical setting. I investigated the variants associated with an increased ALS risk, identified disease-causing variants of high penetrance in low-frequency genes and determined the genetic contribution to phenotypic presentation. I also presented two studies demonstrating how clinico-epidemiological and genomic data could be leveraged to detect novel genetic factor involved in disease pathogenesis and progression.**

## 9.2 Summary of results

### Mutational analysis of known ALS genes in a large Italian population-based cohort

**In this chapter, I performed single-variant analysis in our population-based ALS cohort and made an intriguing discovery of low-frequency variants in NEFH that exhibited a protective effect, a rare phenomenon documented in the literature. This study also showcased the remarkable utility of genome sequencing in a rapid and cost-effective manner, potentially surpassing the insights provided by targeted sequencing and exome sequencing. Gene-burden analysis revealed the genes with the most significant associations with ALS in our cohort, providing a deeper understanding of the genetic architecture of ALS.**

### A systematic evaluation of genetic mutations in ALS: a population-based study

**After establishing the genes contributing to disease architecture in our cohort, I used a comprehensive approac to define disease-causing mutations and provide an updated picture of the genetic determinants of ALS in the general population. Indeed, I identified disease-causive variants in a higher proportion of ALS cases than what was previously hypothezed, particularly in familial and early-onset cases. The C9orf72 mutation was the most common, followed by mutations in SOD1, NEK1, TARDBP, and KIF5A. Additionally, I examined the prevalence of oligogenic cases and determined that oligogenic combination of rare variants is not a common factor in ALS genetics. The**

**therapeutic implications for ALS patients strongly support the advocacy for whole-genome sequencing as a routine test. Indeed, when considering ATXN2, C9orf72, SOD1, and FUS—genes currently under investigation for ASO-based treatments—almost one out of seven ALS patients in our cohort could potentially benefit from gene-based therapy.**

### GBA variants influence cognitive status in amyotrophic lateral sclerosis

**In this chapter, I investigated whether variants in the GBA gene, which are well-established risk factor for neurodegenerative disease with additional effect on cognitive phenotypes, could represnt a pleiotropic genetic factor involved in ALS. I demonstrate that deletrious GBA variants are enriched in ALS patients with cognitive impairment. The findings highlights the potential role of GBA variants in ALS and expand our knowledge of the genetic factors influencing the susceptibility of ALS patients to cognitive dysfunction. Furthermore, they reinforce the significance of lysosomal impairment in the neurodegenerative mechanisms involved in ALS. Our study emphasizes that genes can not only impact the risk of developing ALS but also influence various aspects of its phenotype. Bridging the gap in our understanding of the role of genetic modifiers in ALS is crucial for improving diagnosis, prognosis, and the development of effective therapies.**

### Neurodegeneration-associated rare variants contribute to ALS risk and clinical phenotypeIn this chapter I further explored the role of rare variants in neurodegeneration-associated genes in ALS pathogenesis.

**I observed that individuals carrying rare or novel high-impact variants in neurodegenerative disease-associated genes have a higher risk of ALS. I also investigated the overlap of rare genetic variants between ALS and other neurodegenerative diseases, and observed a genetic overlaps of Parkinson Disease and motor neuropathy with ALS. Other than predisposing to motor neuron degeneration, rare variants in neurodegenerative pathways are also associated with a higher burden of cognitive impairment in ALS patients. These findings suggest that the presence of rare variants increases the overall neurodegenerative burden, potentially leading to a more complex neurodegenerative phenotype. This also highlights the genetic heterogeneity underlying ALS phenotypes and underscores the importance of understanding the genotypic correlates of clinical phenotypes to improve patient classification and our knowledge of the disease's biological basis. In summary, the results of this study further illuminates the intricate genetic landscape of motor neuron degeneration. By uncovering the impact of rare variants and emphasizing their role as modifiers of**

**the ALS phenotype, I provided a foundation for considering a more comprehensive evaluation of neurodegenerative disease genes in genetic counseling.**

### Clinical and Metabolic Signature of UNC13A rs12608932 Variant in Amyotrophic Lateral Sclerosis

**In this chapter I analyzed the UNC13A rs12608932 variant and its association with ALS risk. The C/C genotype and the C allele were both associated with an increased risk of ALS, but the recessive inheritance model provided the best fit for the data and was associated with a higher effect on ALS risk and phenotype. When comparing the clinical features of patients with, I revealed confirmed that UNC13A rs12608932 determined a shorter survival and revealed novel associations of the variant with age at onset, site of onset, and weight loss.**

### Exploring the phenotype of Italian patients with ALS with intermediate ATXN2 polyQ repeats

**In this study, I assessed the role of intermediate ATXN2 polyglutamine expansions in ALS. Among our cohort, the presence of ≥31 polyQ repeats was strongly associated with increased ALS risk. When comparing ALS patients with intermediate ATXN2 polyQ repeats to those without expansion, I found that intermediate ATXN2 polyQ repeats patients had a higher frequency of spinal onset and faster disease progression. Survival analysis confirmed that intermediate ATXN2 polyQ repeats were independently associated with worse outcomes. Furthermore, intermediate ATXN2 polyQ repeats were more frequent among ALS individual with frontotemporal dementia (ALS-FTD).**

### Association of Intermediate HTT CAG Repeat Sizes With Increased Disease Risk and Shorter Survival in Amyotrophic Lateral Sclerosis

**In this chapter I investigated the association between intermediate HTT CAG repeat expansions and ALS risk and phenotype. I found a slightly increased risk of ALS among intermediate-length HTT repeat carriers, that I confirmed through a meta-analysis of four previously published. Regarding ALS phenotype, ALS patients carrying intermediate HTT expansions had a significantly shorter diagnostic delay and faster disease progression, as measured by ALSFRS-R and King's scores. They also had shorter median survival compared to non-carrier ALS patients, and Cox multivariable analysis confirmed that HTT CAG intermediate repeat number was independently associated with worse outcomes. With this study I provided additional compelling evidence for the pathogenic role of HTT expansions in ALS and FTD. These findings emphasize the importance of genetic factors in**

**determining the natural history of ALS and provide insights for future research on targeted therapies. The identification of intermediate HTT expansions as risk factors opens new possibilities for intervention and treatment strategies, offering hope for improved outcomes in ALS patients.**

## 9.3 Strengths and limitations

**Throughout this dissertation, various strengths and limitations have been discussed in each relevant data chapter. Here, I will provide an overview of the strengths and limitations that are applicable to the overall results presented in Chapters 2-8.**

### Strengths

**One notable strength of this dissertation is the study design of PARALS, which is population-based. This approach reduces the likelihood of referral bias and provides a more representative sample of the general ALS population. By including patients with atypical presentations, later disease onset, or faster disease progression, the analysis encompasses a broader range of cases often excluded from multicenter case-control studies. Additionally, the prospective nature of the cohort allows for longitudinal evaluation of ALS natural history, providing a unique opportunity to investigate genetic factors influencing disease progression.**

**Another strength is the utilization of whole-genome sequencing data across a wide range of neurodegenerative-related genes. This approach offers valuable insights into overlapping genetic risk factors for various neurodegenerative diseases. The work expands our understanding of the complex genetic architecture of ALS and enhances our knowledge of rare genetic variants contributing to ALS development and phenotypes. The identification of novel gene-disease associations in Chapters 2, 4, 5, and 6 demonstrates the potential of this approach.**

**Moreover, the combined data in this dissertation emphasize the importance of exploring genetic factors associated with other diseases in ALS patients. By studying overlapping disease features, novel genetic mechanisms in ALS can be discovered, overcoming the limitations encountered by genome-wide association studies in terms of statistical power and replication.**

### Limitations

**It is essential to consider the limitations when interpreting the results and conclusions presented in this dissertation. One fundamental limitation is the relatively modest sample size of the PARALS cohort. Genome-wide genetic analyses require larger sample sizes to achieve sufficient statistical power, especially when investigating rare variants. Although the deep phenotyping available in the PARALS cohort strengthens the study, it also complicates comparisons with other cohorts, limiting the generalizability of the findings. Most of the analyses included only a single cohort, and replication in independent cohorts was not possible. Replication cohorts are crucial in genomics studies to validate associations and mitigate the risk of spurious findings. Although efforts were made to address this limitation by demonstrating replication with publicly available data (Chapter 5) and performing joint analyses of different cohorts (Chapter 8), caution should be exercised until replication is achieved in cohorts of diverse ancestry to ensure the generalizability of the results.**

## 9.4 Applications and future directions

### Whole-genome sequencing as a diagnostic tools

**The results presented in this dissertation highlights the needs to implent WGS as a first-tier genetic test in ALS and the importance of expanding genetic analysis to a greater number of neurodegenerative genes. In the fight to cure ALS it is imperative to gain a better understanding of the genetic factors underlying the disease: WGS could allow us to study variants not considered in this dissertation, like non coding variants or structural variation, that could be important factors to improve ALS diagnosis and genetic counselling.**

### Influence of genetic variation on disease phenotypes

**A better understanding of the genetic determinants of ALS natural history may allow for a more accurate and patient tailored progression prediction. Additionally, the study of genetic modifiers may provide necessary insights to novel therapeutic approaches and to prioritize the potential targets based on their overall impact on several aspects of ALS manifestation. Ultimately, in the study of such complex diseases, this will be an important component for future personalized medicine**

### Identification of novel genetic association and gene-disease relationships

**I provided additional demonstration that applying a gene-set approach based on clinical or functional annotations of genetic variants could mitigate the limitations of rare-variants association tests in term of statistical power. Finally, I demonstrated that most ALS risk factor also influence the clinical phenotype of the disease, and by combining a case-control study with phenotype analysis we can discover novel relatively common genetic factors associated with ALS.**