ATXN2 polyQ intermediate repeats are a modifier of ALS survival.

Original Citation:

Published version:
DOI:10.1212/WNL.0000000000001159

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)
ATXN2 polyQ intermediate repeats are a modifier of ALS survival

Adriano Chiò, MD, FAAN, Andrea Calvo, MD, PhD, Cristina Moglia, MD, Antonio Canosa, MD, Maura Brunetti, BSc, Marco Barberis, BSc, Gabriella Restagno, MD, Amelia Conte, MD, Giulia Bisogni, MD, Giuseppe Marangi, MD, Alice Moncada, MD, Serena Lattante, MD, Marcella Zollino, MD, Mario Sabatelli, MD, Alessandra Bagarotti, PhD Lucia Corrado, PhD, Gabriele Mora, MD, Enrica Bersano, MD, Letizia Mazzini, MD; Sandra D’Alfonso, PhD for the PARALS*

From: the ALS Center, ‘Rita Levi Montalcini’ Department of Neuroscience, University of Torino, Torino, Italy (Chiò, Calvo, Moglia, Canosa, Brunetti, Barberis); the Azienda Ospedaliero Universitaria Città della Scienza di Torino, Torino, Italy (Chiò, Calvo); the Neuroscience Institute of Torino (NIT), Torino, Italy (Chiò, Calvo); the Department of Neurosciences, Ophthalmology, Genetics, Rehabilitation and Child Health, University of Genoa (Canosa); the Laboratory of Molecular Genetics, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Torino (Brunetti, Barberis, Restagno); the Neurological Institute, Catholic University of the Sacred Heart, Rome, Italy (Conte, Bisogni, Sabatelli); the Institute of Medical Genetics, Catholic University of the Sacred Heart, Rome, Italy (Marangi, Moncada, Lattante, Zollino); the Department of Health Sciences, Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), ‘Amedeo Avogadro’ University of Eastern Piedmont, Novara, Italy (Bagarotti, Corrado, D’Alfonso); the Salvatore Maugeri Foundation, IRCSS; Scientific Institute of Milan, Milan, Italy (Mora); the Department of Neurology, ‘Amedeo Avogadro’ University of Eastern Piedmont and Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy (Bersano, Mazzini).

* A list of the members of PARALS is provided in the Co-investigator Appendix
Abstract word count: 247

Text work count: 2643

Character count for title: 63

Number of references: 40

**Corresponding author:** Prof. Adriano Chiò, ‘Rita Levi Montalcini’ Department of Neuroscience, Via Cherasco 15, I-10126 Torino, Italy, E-mail: achio@usa.net

**Running title:** ATXN2 polyQ repeats in ALS

**Disclosures**

Dr. Chiò serves on the editorial advisory board of Amyotrophic Lateral Sclerosis and has received research support from the Italian Ministry of Health (Ricerca Finalizzata), Regione Piemonte (Ricerca Finalizzata), University of Torino, Federazione Italiana Giuoco Calcio, Fondazione Vialli e Mauro onlus, and the European Commission (Health Seventh Framework Programme); he serves on scientific advisory boards for Biogen Idec and Cytokinetics. Dr. Calvo has received research support from the Italian Ministry of Health (Ricerca Finalizzata). Dr. Moglia reports no disclosures. Dr. Canosa reports no disclosures. Dr. Brunetti reports no disclosures. Dr. Barberis reports no disclosures. Dr. Restagno has received research support from the Italian Ministry of Health (Ricerca Finalizzata) and Regione Piemonte (Ricerca Finalizzata). Dr. Conte reports no disclosures. Dr. Bisogni reports no disclosures. Dr. Marangi reports no disclosures. Dr. Moncada reports no
disclosures. Dr. Lattante reports no disclosures. Dr. Zollino reports no disclosures. Dr. Sabatelli
reports no disclosures. Dr. Bagarotti reports no disclosures. Dr. Corrado reports no disclosures. Dr.
Mora reports no disclosures. Dr. Bersano reports no disclosures. Dr. Mazzini reports no disclosures.
Dr. D’Alfonso reports no disclosures.

Author Contributions: Study concept and design: Chiò, Calvo, Restagno, Zollino, Sabatelli, Mora,
Mazzini, D’Alfonso. Acquisition of data: Moglia, Canosa, Brunetti, Barberis, Conte, Bisogni,
Marangi, Moncada, Lattante, Bagarotti, Corrado, Bersano. Analysis and interpretation of data:
Chiò, Calvo, Restagno, Zollinjo, Sabatelli, Mora, Mazzini, D’Alfonso. Drafting of the manuscript:
Chiò, Restagno, Sabatelli, Mora, Mazzini, D’Alfonso. Critical revision of the manuscript for
important intellectual content: Chiò, Calvo, Restagno, Zollinjo, Sabatelli, Mora, Mazzini,
D’Alfonso, Moglia, Canosa, Brunetti, Barberis, Conte, Bisogni, Marangi, Moncada, Lattante,
Bagarotti, Corrado, Bersano. Obtained funding: Chiò, Sabatelli. Administrative, technical, and
material support: Moglia, Canosa, Brunetti, Barberis, Conte, Bisogni, Marangi, Moncada, Lattante,
Bagarotti, Corrado, Bersano. Study supervision: Chiò, Restagno, Sabatelli, Mora, Mazzini,
D’Alfonso.

Adriano Chiò had full access to all of the data in the study and takes responsibility for the integrity
of the data and the accuracy of the data analysis. All authors have approved the submitted version of
the paper.
Abstract

Objective. To analyze the frequency and clinical characteristics of ALS patients with intermediate-length (CAG) expansion (encoding 27-33 glutamines, polyQ) in the \textit{ATXN2} gene, in a population-based cohort of Italian ALS patients (discovery cohort), and to replicate the findings in an independent cohort of consecutive patients from an ALS tertiary center (validation cohort).

Methods. PolyQ repeats were assessed in 672 ALS patients incident in Piemonte and Valle d’Aosta regions, Italy, in the 2007-2012 period (discovery cohort); controls were 509 neurologically healthy age- and gender-matched subjects resident in the study area. The validation cohort included 661 ALS patients consecutively seen between 2001 and 2013 in the ALS Clinic Center of the Catholic University in Rome, Italy.

Results. In the discovery cohort the frequency of \( \geq 31 \) polyQ \textit{ATXN2} repeats was significantly more common in ALS cases (19 patients \textit{vs}. 1 control, \( p=0.0001 \); odds ratio 14.8, 95% confidence interval, 1.9-110.8). Patients with an increased number of polyQ repeats had a shorter survival than those with <31 repeats (median survival, polyQ \( \geq 31 \), 1.8 years, interquartile range [IQR] 1.3-2.2; polyQ <31, 2.7 years, IQR 1.6-5.1) (\( p=0.001 \)). An increased number of polyQ repeats remained independently significant also at multivariable analysis. In the validation cohort, patients with \( \geq 31 \) polyQ repeats had a shorter survival than those with <31 repeats (median survival, polyQ \( \geq 31 \), 2.0 years, IQR 1.5-3.4; polyQ <31, 3.2 years, IQR 2.0-6.4; \( p=0.007 \)).

Conclusions. \textit{ATXN2} polyQ intermediate-length repeat is a modifier of ALS survival. Disease-modifying therapies targeted to \textit{ATXN2} represent a promising therapeutic approach for ALS.

Keywords: Amyotrophic Lateral Sclerosis; Ataxin2 gene; prognosis; phenotype
Amyotrophic lateral sclerosis (ALS) is characterized by a progressive degeneration of upper and lower motor neurons, leading to loss of motor function and, eventually, to complete palsy of limb muscles, loss of speech and swallowing, and respiratory failure, usually within 2 to 3 years after onset. Several prognostic factors are known in ALS, in particular age at onset, bulbar onset, diagnostic delay, and cognitive impairment.

With the exception of riluzole, which slightly increases survival, no disease-modifying drug is available. The cause of ALS is still unknown, but at least 20 genes have been related to the disease; the most common in Caucasian populations are C9ORF72, SOD1, TARDBP and FUS, which are present in about 10% of patients. In 2010, an intermediate-length (CAG) expansion (encoding 27-33 glutamines, polyQ) in the ataxin 2 (ATXN2) gene, already known as the cause of spinocerebellar ataxia type 2 (SCA2),\textsuperscript{1,2,3} was reported to be associated to an increased risk of developing ALS.\textsuperscript{4} The association was subsequently confirmed in clinically-based series,\textsuperscript{5-16} but the clinical characteristics of patients with this expansion still remain to be fully investigated.

The aim of this paper was to analyze the frequency of intermediate polyQ expansion in the ATXN2 gene in a population-based cohort of ALS patients, with an in-depth assessment of their clinical and prognostic characteristics. The study findings would then be replicated in an independent replication cohort of consecutive patients from an ALS tertiary center.

**Methods**

**Discovery cohort.** The study population included all ALS cases diagnosed in Piemonte and Valle d’Aosta, Italy, during the 6-year period January 1 2007, to December 30 2012. ALS cases were recruited through the PARALS, a prospective epidemiologic register involving all the neurologic departments of the two regions of northern Italy. Epidemiologic data regarding the 1995–2004 period have been published elsewhere.\textsuperscript{17} Both familial and apparently sporadic ALS patients were
include in the present study. The diagnosis of ALS was based on El Escorial revised criteria.\textsuperscript{18} Patients with definite, probable, and probable laboratory-supported ALS were included in the register. Controls were (1) 395 regionally matched unrelated Italian subjects, mainly blood donors, already reported in a previous paper;\textsuperscript{6} (2) 114 matched subjects identified through the patients’ general practitioners (population-based controls). The two series of controls were similar for demographic characteristics and had a substantially similar frequency of ATXN2 intermediate polyQ expansions (0.4% and 0.3%, respectively).

**Validation cohort.** This cohort included 661 patients consecutively admitted between January 2001 and December 2013 to the ALS Clinic Center of the Catholic University in Rome, Italy. All patients were resident in the central or southern regions of Italy. Diagnostic criteria were identical to those of the discovery cohort. Data on 528 of the patients were already reported in a study on genetics of ALS patients.\textsuperscript{19}

**Genetic analysis.** Genomic DNA was isolated from peripheral blood lymphocytes using a standard protocol. The ATXN-2 CAG repeat in exon 1 (Ref Seq NM_002973.3) was amplified using a fluorescent primer and sized by capillary electrophoresis on an ABI 3130 genetic analyzer (Applied Biosystem, Foster City, CA, USA).\textsuperscript{20} As reported in recent guidelines for molecular genetic testing of SCA,\textsuperscript{21} capillary electrophoresis is the preferred method. This method allows to size alleles accurately and to resolve alleles one triplet apart in size. As a quality control, 20 samples have been genotyped in the different laboratories that performed the molecular genetic testing for the present study. The comparison of the results showed a consistent allele assignment for all the samples. Receiver operating characteristics (ROC) analysis showed that a cutoff $\geq$31 polyQ repeats in ATXN2 had the greatest sensitivity and specificity for discriminating ALS patients versus controls. However, also data related to a repeat size 27-30 were assessed.
All ALS cases of both cohorts were also tested for SOD1 (all exons), TARDBP (exon 6), FUS (exons 14 and 15), ANG and C9ORF72 using the methodology described elsewhere.\textsuperscript{19,22} Familial ALS patients were also tested for OPTN (exons 5, 9, 12 and 14).

**Statistical methods.** Two-tailed Fisher’s exact test was used to evaluate the genetic association between \textit{ATXN2} polyQ repeat size and ALS. Survival was calculated from onset to death, tracheostomy or censoring date (December 31, 2013), using the Kaplan-Meier method, and compared with the log-rank test. No patients were lost to follow-up. Multivariable analysis was performed with the Cox proportional hazards model (stepwise backward) with a retention criterion of \( p<0.1 \). Significance level was set at \( p<0.05 \). Data were processed using SPSS statistical package version 21 (IBM Corporation, Chicago, IL, USA).

**Ethical approval.** The study was approved by the ethical committees of the participating centers. All patients and controls signed a written informed consent. Databases were treated with due respect for Italian privacy regulations.

**Results**

**Discovery cohort**

Out of a total of 869 ALS cases incident in the period January 1\textsuperscript{st} 2007 to December 31\textsuperscript{st} 2012, 672 (77.3\%) patients were included in the study and tested for DNA. Data on 28 of them had already been reported in a previous paper.\textsuperscript{6} Of the 197 patients not included in the study, 59 did not give their consent for genetic analysis, 60 died before blood sampling and 78 were found only through secondary sources and were therefore not tested for DNA. The patients not included in the genetic analysis had an older mean age at onset (67.9 [SD 10.7] vs. 65.6 [SD 10.8] years, \( p<0.02 \)), but their site of onset and clinical phenotype were similar (data not shown). A total of 71 (10.6\%) patients
carried a mutation of one of the ALS-related genes (C9ORF72=43; SOD1=15; TARDBP=10; FUS =2; OPTN=1).

**ATXN2 polyQ repeat size at risk.** The size of the ATXN2 repeats in ALS patients compared to the control group is reported in Figure 1. The more common alleles (22 and 23) were identified in 98.9% of controls’ chromosomes and 95.9% of cases’ chromosomes. ATXN2 repeats ≤30 were similarly distributed between ALS cases and controls, while those ≥31 were significantly more common in cases (19 cases and 1 control, \( p=0.0001 \)); the odds ratio for ALS of having ≥31 repeats was 14.8 (95% confidence interval, 1.9-110.8) (Figure 1, insert). The second allele in patients with ≥27 repeats was 22 in all but one case, who had 29/24.

**Clinical characteristics of patients according to ATXN2 polyQ repeat size.** Demographic and clinical characteristics of patients with ATXN2 polyQ repeats ≥31 vs. patients with polyQ repeats <31 are presented in Table 1. No significant differences were found in gender, age at onset, and presence of co-morbid FTD. Conversely, bulbar onset was present in only one case with ATXN2 polyQ repeats ≥31 (~5%) versus about one third of patients with polyQ repeats <31 (\( p=0.0001 \)). No patients with ATXN2 polyQ repeats ≥31 had a positive family history for ALS or FTD. Of the 71 patients carrying a known genetic mutation, 69 were homozygous (22-22 repeats), one p.D90N SOD1 patient had 22-24 repeats and one C9ORF72 patient had 22-27 repeats. Clinical characteristics of ALS patients with ≥31 polyQ repeats of the ATXN2 gene are detailed in Table e-1. The single patient carrying 39 polyQ repeats had features of both ALS and SCA. She developed gait ataxia at the age of 55 and spinal ALS 12 years later; her mother was also affected by SCA. She underwent percutaneous endoscopic gastrostomy 12 months after the onset of ALS and died of respiratory failure at age 69.

Patients with ≥31 polyQ repeats had a shorter median survival than those with <31 repeats (1.8 years, interquartile range [IQR] 1.3-2.2 vs. 2.7 years, IQR 1.6-5.1; \( p=0.001 \)) (Figure 2A). An increased number of polyQ repeats remained independently significant also at multivariable
analysis (Table 2). A second analysis was performed assessing separately the survival of patients with 27-30 polyQ repeats: this group had an intermediate survival time between those with ≥31 and those with <27 polyQ repeats, indicating a possible dose-response relationship (Figure 2C).

Repetition cohort

Clinical and demographic characteristics of the 661 ALS patients included in this cohort are reported in Table 1. All were tested for ALS genes and 80 (12.1%) carried a mutation of one of the ALS-related genes (C9ORF72=34; SOD1=20; TARDBP=13; ANG=5, FUS=4; OPTN=3; VCP=1). Most of these cases were reported in a previous paper. The more common alleles (22 and 23) were identified in 95.8% of chromosomes (Figure e-1). Sixteen cases had ≥31ATXN2 polyQ repeats. The second allele in patients with ≥27 repeats was 22 in all cases. Of the 80 patients carrying a known genetic mutation, 63 were homozygous (22-22 repeats); one C9ORF72 patient and one p.I380V TARDBP patient had 19-22 repeats; two C9ORF72 patients had 21-22 repeats; 3 C9ORF72 patients, one p.R521C FUS patient, one p.D11Y SOD1 patient and one p.S134N SOD1 patient had 22-23 repeats; one heterozygous D90A SOD1 patient had 22-26 repeats; one p.N87K ANG patient had 22-27 repeats; one p.G93D SOD1 patient had 22-28 repeats; two C9ORF72 patients had 22-29 repeats; and one p.N390S TARDBP patient had 22-30 repeats.

Out of the 16 patients with ≥31ATXN2 polyQ repeats, 2 had a bulbar onset and 2 had a positive family history for ALS or FTD (Table e-1). Cognitive impairment was not systematically assessed in patients of this series. Patients with ≥31polyQ repeats had a shorter median survival than those with <31 repeats (2.0 years, IQR 1.5-3.4 vs. 3.2 years, IQR 2.0-6.4; p=0.007) (Figure 2B). Multivariable analysis confirmed that ≥31polyQ repeats was an independent prognostic factor (data not shown). Also in this cohort, patients with 27-30 polyQ repeats had an intermediate survival time between those with ≥31 and those with <27 polyQ repeats, indicating a possible dose-response relationship (Figure 2D).
Discussion

In this population-based study, which included ~80% of ALS patients incident in Piemonte/Valle d’Aosta over a 6-year period, we found that ≥31 polyQ repeats in the \textit{ATXN2} gene represents a risk factor for ALS, with an odds ratio of 14.8. All but one patient with intermediate-length polyQ repeats had a spinal onset and all the patients who were positive for mutations of ALS-related genes had normal size polyQ repeats. To the best of our knowledge, the effect of an increased number of polyQ repeats in ALS outcome has never been studied. In our series, the presence of ≥31 polyQ repeats reduced the median survival by ~1 year, patients with 27-31 repeats showing an intermediate survival time between that of those with ≥31 and those with <27 polyQ repeats, and this effect persisted after adjusting for other known prognostic factors at multivariable analysis. These survival data were confirmed in a validation cohort of 661 consecutive patients from a referral ALS center in central Italy.

The correlation between \textit{ATXN2} intermediate polyQ repeats and a shorter survival, observed in both our series, is further supported by finding that almost all patients in both series have a spinal onset. In fact, spinal onset is considered a positive prognostic factor in ALS,\textsuperscript{23} and therefore these patients were expected to have a better prognosis.

Besides the correlation with spinal onset, \textit{ATXN2} intermediate polyQ repeats in both our series did not influence the age at onset of ALS, differently from the original observation of a lower age at onset observed in the first report of ATXN2 in ALS,\textsuperscript{4} but in keeping with all subsequent papers\textsuperscript{5-16}.

In the last few years several papers have highlighted the phenotypic variability of ALS, characterized by a wide range of age at onset, different clinical phenotypes, variable involvement of upper and lower motor neuron signs, possible impairment of cognitive function, and an extremely variable length of survival.\textsuperscript{22-24} An increased understanding of the genetic background of ALS has
partly clarified its pathogenetic mechanisms, but has provided few clues about its clinical heterogeneity, with the possible exception of the frequent cognitive involvement in patients with C9ORF72 mutations and the young age at onset and severe course of most patients carrying FUS mutations. The search for genes modulating the clinical expression of ALS has produced few, albeit interesting, findings. For example, polymorphisms of Unc-13 homolog A (UNC13A)\textsuperscript{25,26} and Non-Imprinted in Prader–Willi/Angelman syndrome 1 (NI-PA1)\textsuperscript{27} have been associated to a shorter survival, while a locus on 1p34.1\textsuperscript{28} has been associated to a younger age at onset. In our population-based series we found that intermediate-length polyQ repeats in the ATXN2 gene are a strong modifier of ALS outcome in an Italian subgroup, reducing survival by one year, without modifying age at onset. This finding was replicated in an independent validation cohort from an ALS tertiary center. In addition, ALS patients carrying ≥31 polyQ repeats had more frequently a spinal onset. A meta-analysis of published case series (Table 3) revealed that 70 (88.6%) cases had a spinal onset, significantly more than the expected 65%.\textsuperscript{29}

The mechanism underlying the increased risk due to an intermediate-length polyQ repeat in the ATXN2 gene in ALS is still unclear. It was initially hypothesized that ATXN2 intermediate-length polyQ repeats interact with TDP-43 augmenting its toxicity\textsuperscript{4} or increase the translocation of TDP-43 and mutant FUS to the cytoplasm enhancing their pathogenicity.\textsuperscript{30,31} More recent studies found that ATXN2 intermediate-length polyQ repeats may interact with the NADPH oxidase enzyme, increasing ROS production, and may also interfere with RNA metabolism, sequestering essential RNA-binding proteins.\textsuperscript{32}

≥31 polyQ repeats in the ATNX2 gene have also been identified in patients with FTD (4/641, 0.6%),\textsuperscript{33} but in no patients in a small series of subjects with FTD-ALS.\textsuperscript{15} In our discovery cohort, behavioral FTD was found in two patients (10.5%) with ≥31 polyQ repeats, compared to 76 patients (12.4%) with <31 repeats (p=n.s.), confirming that cognitive impairment can be present also in association to intermediate-length polyQ repeat expansions in ALS, but the frequency is similar to
that found in general in ALS.\textsuperscript{24,34} Due to incomplete cognitive testing in part of our patients, we could not include in the multivariable analysis intermediate forms of cognitive impairment, in particular isolate dysexecutive function.\textsuperscript{24,34}

A meta-analysis of the published literature (Table 3) provides some clues regarding a possible different frequency of polyQ expansions, in particular a north-to-south increasing trend, ranging from <1\% in patients of northern European ancestry\textsuperscript{7,8,12,15} to ~2\% in French/French Canadian\textsuperscript{5} and ~2.5\% in Italian patients.\textsuperscript{6,10,11,13} In Turkish\textsuperscript{14} and in Chinese,\textsuperscript{9,16} the frequency of patients with $\geq$31 polyQ repeats is ~1\%. These ethnic differences in polyQ repeats parallel the different frequencies of other ALS genes, such as the north-to-south decreasing gradient of $C9ORF72$ in Europe\textsuperscript{35} or the rarity of $SOD1$ mutations in Irish\textsuperscript{36} and Dutch\textsuperscript{37} compared to Scandinavian patients.\textsuperscript{38} Interestingly, a north-to-south increasing gradient of the frequency of full-length polyQ repeats of the $ATXN2$ gene has been also described in SCA2.\textsuperscript{39}

A recent multicenter paper performed on an US population has assessed polyQ intermediate repeats of $ATXN2$ in a series of 331 patients carrying a $C9ORF72$ expansion compared to 376 control subjects and found the $ATXN2$ expansion was more significantly frequent in $C9ORF72$ mutation carriers with ALS or ALS-FTD phenotype but not in those with pure FTD phenotype.\textsuperscript{40} Combining our series, out of 79 ALS patients with $C9ORF72$ mutation, 3 (3.8\%) had $\geq$ 27 polyQ repeats, not significantly different from patients not carrying the $C9ORF72$ mutation (7.4\%) or controls (1.6\%). This finding reinforce the differential characteristics of $ATXN2$ in population with different ethnic background.

A strength of our discovery cohort is that it includes ~80\% of patients diagnosed in Piemonte/Valle d’Aosta in the 2007-2011 period. Moreover, all patients were prospectively followed and none was lost. Finally, all patients were assessed for major ALS genes ($C9ORF72$, $SOD1$, $TARDBP$, $FUS$, $ANG$ and $OPTN$). The validation cohort included a large series of ALS patients consecutively seen
in an ALS referral center, and, although not population-based, it is fairly representative of the ALS population.

We found that an ATXN2 intermediate-length polyQ repeat is a significant risk factor for ALS, is correlated to a spinal phenotype and associated to shorter survival. Disease-modifying therapies targeted to ATXN2 represent a promising therapeutic approach for a devastating disease such as ALS; possible strategies may be the use of antisense oligonucleotides, transcription activator-like effectors nuclease, and clustered regularly interspaced short palindromic repeats.\(^{32}\)

**Acknowledgment**

The authors thank the patients and their families for their collaboration in this study.

**Study Funding.** This work was in part supported by the Italian Ministry of Health (Ministero della Salute, Ricerca Sanitaria Finalizzata, 2010, grant RF-2010-2309849), the European Community’s Health Seventh Framework Programme (FP7/2007-2013 under grant agreement 259867), the Joint Programme - Neurodegenerative Disease Research (Italian Ministry of Education and University) (Sophia and ALS-Care Projects), the Agenzia Italiana per la Ricerca sulla SLA (ARISLA) (RepeatALS project), the Associazione Piemontese per l’Assistenza alla SLA (APASLA), Torino, Italy, the Uniti per la Ricerca sulla Sclerosi Laterale Amiotrofica (URSLA) Association, Novara, Italy, and the Fondazione Mario e Anna Magnetto, Alpignano, Torino
References


**Table 1.** Demographic and clinical characteristics of patients according to \textit{ATXN2} repeat size

<table>
<thead>
<tr>
<th>Factor</th>
<th>Discovery cohort</th>
<th>Validation cohort</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;31 (n=653)</td>
<td>≥31 (n=19)</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>65.5 (10.8)</td>
<td>68.8 (8.1)</td>
<td>0.19</td>
</tr>
<tr>
<td>Age at onset (years, SD)</td>
<td>60.6 (12.1)</td>
<td>62.2 (10.4)</td>
<td>0.61</td>
</tr>
<tr>
<td>Gender, female (%)</td>
<td>300 (45.9%)</td>
<td>8 (42.1%)</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>274 (42.5%)</td>
<td>6 (37.5%)</td>
<td>0.45</td>
</tr>
<tr>
<td>Site of onset, bulbar (%)</td>
<td>217 (33.2%)</td>
<td>1 (5.3%)</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>167 (25.9%)</td>
<td>2 (12.5%)</td>
<td>0.18</td>
</tr>
<tr>
<td>Positive family history for ALS or FTD</td>
<td>59 (9.0%)</td>
<td>0</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>64 (9.9%)</td>
<td>2 (12.5%)</td>
<td>0.74</td>
</tr>
</tbody>
</table>
Table 2. Cox Multivariable analysis

<table>
<thead>
<tr>
<th>Factor</th>
<th>Level</th>
<th>Hazard ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at onset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20-49</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>1.49 (0.90-2.33)</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>1.90 (1.23-2.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70-79</td>
<td>2.38 (1.54-3.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>80-99</td>
<td>4.00 (2.39-6.72)</td>
<td></td>
</tr>
<tr>
<td>Site of onset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spinal</td>
<td>1</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Bulbar</td>
<td>1.55 (1.25-1.90)</td>
<td></td>
</tr>
<tr>
<td>ATXN2 polyQ repeats</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;31</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥31</td>
<td>2.79 (1.67-4.64)</td>
<td>0.0001</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Frequency, odds ratio and clinical presentation of patients with ≥31 CAG repeats in the \textit{ATXN2} gene.

<table>
<thead>
<tr>
<th>Reference</th>
<th>N. of ALS cases</th>
<th>N. of expanded cases (≥31)</th>
<th>% of expanded cases</th>
<th>N. of controls</th>
<th>N. of expanded controls (≥31)</th>
<th>Odds ratio (95% c.i.)</th>
<th>N. of spinal onset</th>
<th>N. of bulbar onset</th>
<th>Ethnic origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>[9]</td>
<td>1294</td>
<td>13</td>
<td>1.00</td>
<td>679</td>
<td>0</td>
<td>6.9 (0.9-52.8)</td>
<td>NS</td>
<td>NS</td>
<td>Mostly north European</td>
</tr>
<tr>
<td>[7]</td>
<td>1948</td>
<td>10$^\text{a}$</td>
<td>0.51</td>
<td>2002</td>
<td>2</td>
<td>5.2 (1.1-23.6)</td>
<td>9</td>
<td>1</td>
<td>Dutch/Flemish Belgian</td>
</tr>
<tr>
<td>[12]</td>
<td>559</td>
<td>4</td>
<td>0.72</td>
<td>1369</td>
<td>2</td>
<td>5.0 (0.9-26.9)</td>
<td>NS</td>
<td>NS</td>
<td>German</td>
</tr>
<tr>
<td>[15]</td>
<td>72</td>
<td>1</td>
<td>1.39</td>
<td>810</td>
<td>1</td>
<td>11.4 (0.7-184.1)</td>
<td>NS</td>
<td>NS</td>
<td>Flemish Belgian</td>
</tr>
<tr>
<td>North European</td>
<td>3873</td>
<td>28</td>
<td>0.72</td>
<td>4860</td>
<td>5</td>
<td>7.1 (2.7-18.3)</td>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>French/French Canadian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[6]</td>
<td>556</td>
<td>11</td>
<td>1.98</td>
<td>471</td>
<td>1</td>
<td>9.8 (1.2-73.7)</td>
<td>8$^\text{a}$</td>
<td>0</td>
<td>French/French Canadian</td>
</tr>
<tr>
<td>French</td>
<td>556</td>
<td>11</td>
<td>1.98</td>
<td>471</td>
<td>1</td>
<td>9.8 (1.2-73.7)</td>
<td>8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>South European</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[4]</td>
<td>#</td>
<td>232#</td>
<td>3.02</td>
<td>395#</td>
<td>0#</td>
<td>12.3 (1.5-100.5)</td>
<td>6#</td>
<td>1#</td>
<td>Northern Italian</td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Cases</td>
<td>Controls</td>
<td>OR (95% CI)</td>
<td>P-value</td>
<td>N</td>
<td>MAF</td>
<td>Population</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-------</td>
<td>----------</td>
<td>-------------</td>
<td>---------</td>
<td>---</td>
<td>-----</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>Present paper – discovery cohort</td>
<td>672</td>
<td>19</td>
<td>509</td>
<td>14.8 (1.9-110.8)</td>
<td>18</td>
<td>1</td>
<td>2.83</td>
<td>Northern Italian</td>
<td></td>
</tr>
<tr>
<td>Present paper – validation cohort</td>
<td>661</td>
<td>16</td>
<td>-</td>
<td>-</td>
<td>14</td>
<td>2</td>
<td>2.42</td>
<td>Central Italian</td>
<td></td>
</tr>
<tr>
<td>South European</td>
<td>3014</td>
<td>73</td>
<td>1712</td>
<td>8.5 (3.9-21.1)</td>
<td>48</td>
<td>7</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>236</td>
<td>4</td>
<td>420</td>
<td>7.2 (0.8-65-2)</td>
<td>2</td>
<td>2**</td>
<td>1.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>345</td>
<td>4</td>
<td>350</td>
<td>4.1 (0.5-36.9)</td>
<td>NS</td>
<td>NS</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1067</td>
<td>17</td>
<td>506</td>
<td>8.2 (1.1-61.7)</td>
<td>13*</td>
<td>1</td>
<td>1.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1412</td>
<td>21</td>
<td>856</td>
<td>12.9 (1.7-96.5)</td>
<td>13</td>
<td>1</td>
<td>1.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>9640</td>
<td>147</td>
<td>1.52</td>
<td>14908</td>
<td>20</td>
<td>11.5 (7.2-18.4)</td>
<td>80</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>----</td>
<td>----------------</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>

*unknown in 1 case; # cases and controls also included in the present paper are excluded from the total count; *unknown in 3 cases; **1 mixed (spinal and bulbar); ≥32 repeats; NS: not stated
**Figure legends**

**Figure 1.** Discovery cohort. The distribution of ATXN-2 polyQ repeat lengths in ALS and control cases. In the insert, data concerning cases and controls with ≥27 repeats are magnified. PolyQ lengths ≥31 are significantly more frequent in ALS cases (p=0.0001) (blue, ALS patients; red, controls).

**Figure 2.** Kaplan-Meier survival estimates from onset to death/tracheostomy. A. Discovery cohort. Blue line, <31 polyQ repeats; green line, ≥31 polyQ repeats. p=0.0001. B. Validation cohort. Blue line, <31 polyQ repeats; green line, ≥31 polyQ repeats. p=0.009. C. Discovery cohort. Kaplan-Meier survival estimation from onset to death/tracheostomy. Blue line, <27 polyQ repeats; red line, 27-30 polyQ repeats green line, ≥31 polyQ repeats. p=0.0001. D. Validation cohort. Kaplan-Meier survival estimation from onset to death/tracheostomy. Blue line, <27 polyQ repeats; red line, 27-30 polyQ repeats green line, ≥31 polyQ repeats. p=0.003.
Figure 1
Figure 2

A

B

C

D

Cumulative survival vs. time from onset (years)