University of Szeged
Albert Szent-Györgyi Medical School
Doctoral School of Interdisciplinary Medicine

Comprehensive genetic analysis of a Hungarian amyotrophic lateral sclerosis cohort

Summary of the Ph.D. thesis

Zsófia Flóra Nagy MD

Supervisors:
Márta Széll DSc, habil, professor
Margit Pál Ph.D

Szeged,
2024

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a rare fatal neurodegenerative disease characterized by the loss of motor neurons in the motor cortex as well as in the brainstem. Its incidence in European populations varies between 0.5-3.6/100 000/year and it is constantly rising.

ALS affects both the upper- and lower motor neurons, its hallmark symptoms include muscle weakness or spasticity, muscle wasting, fasciculations which spread from one region of the body to another. Over the course of the disease bulbar symptoms such as dysphagia, dysarthria may also develop. Respiratory insufficiency due to the weakness of the respiratory muscles is the leading cause of death among ALS patients.

Currently only two drugs are available in ALS: riluzole and edaravone which only moderately prolong survival. A genetically based antisense oligonucleotide based therapy for patients harbouring a *SOD1* gene mutation was granted marketing authorization in 2024.

Despite all therapeutic efforts, ALS still has a poor survival rate with the median survival time from disease onset to death ranging from 1.5-4 years.

1.1 Genetic background of ALS

Around 10% of ALS cases show a positive family history (fALS), while the other 90% remains of sporadic etiology (sALS). The first gene to be associated with ALS was superoxide dismutase 1 (*SOD1*) in 1993 by a linkage analysis study. Currently there are more than 130 genes linked to ALS in the most comprehensive database. The application of next-generation sequencing methods in the last 20 years has led to a boom in the number of ALS genes, which may be split into two categories: causative/major ALS genes and genetic risk factors or clinical modifiers of the disease.

2. Aims

In my thesis I aimed to summarize the genetic examinations carried out by our research group to explore the complex genetic architecture of the Hungarian amyotrophic lateral sclerosis population and to establish genotype-phenotype correlations.

3. Patients and methods

3.1 Patients

183 unrelated patients of Hungarian origin diagnosed with ALS between 2008 and 2021 were recruited for our study. A written informed consent document was signed by all participating patients. All patients fulfilled the revised Awaji-Shima and El Escorial criteria. Our study was performed in accordance with the Helsinki Declaration and its later amendments and was approved by the Ethical Board of the University of Szeged. One patient had a positive family history of ALS.

3.2 Methods

Genomic DNA extraction was performed from peripheral blood samples of ALS patients with a commercially available DNA extraction kit following the manufacturer's protocol.

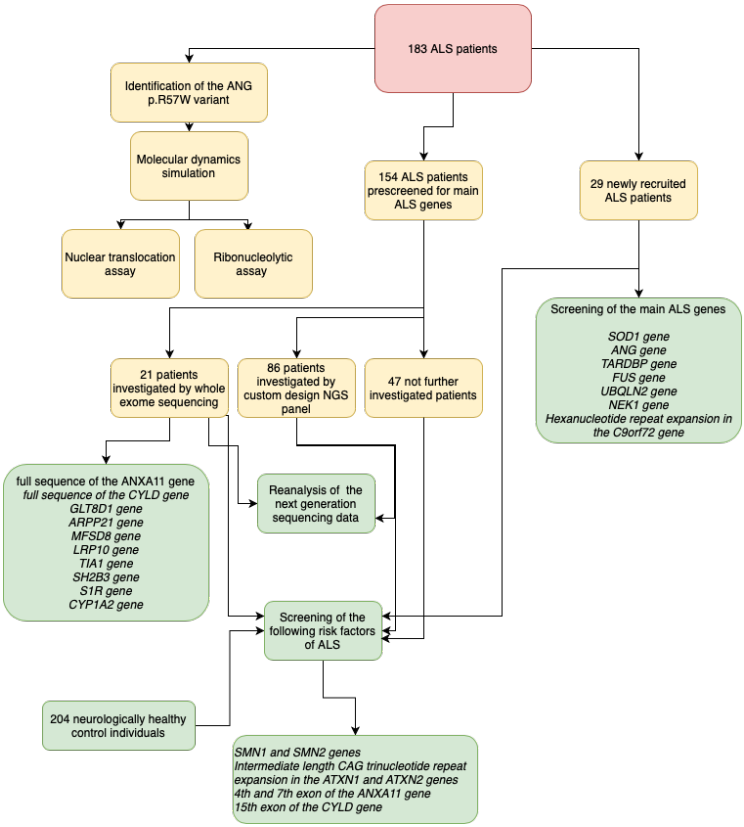

Figure 1. Workflow of our study.

Next generation sequencing techniques were used according to standard practices to assess 107 patients prescreened for variants in the most important ALS genes. Whole exome sequencing was carried out in 21 ALS patients, whilst a custom design gene panel was applied in 86 further

patients. The panel contained 294 genes which were analyzed in 3 gene sets: “Major ALS genes set” (35 genes), “Candidate or minor ALS genes set” (99 genes) and “Further neurodegenerative and/or neuromuscular diseases gene set” (160 genes). Variant interpretation was carried out in accordance with the 2015 guidelines of the American College of Medical Genetics and Genomics.

Sanger sequencing was used to assess the 6 following major ALS genes: *SOD1, TARDBP, ANG, FUS, UBQLN2* and *NEK1*. The G4C2 sequence found in the intronic region of the *C9orf72* gene was investigated by long read repeat primed PCR method. The CAG repeat containing regions of the *ATXN1* and *ATXN2* genes were amplified by polymerase chain reactions using specific 6-FAM fluorophore marked primers. Copy number variation of the *SMN1* and *SMN2* genes were analyzed by multiplex ligation-dependent probe-amplification assay.

To confirm the pathogenicity of the novel *ANG* p.R57W variant, *in silico* modelling and computational simulations were conducted. A ribonucleolytic activity analysis was also carried out using wild type and p.R57W mutant angiogenin proteins on yeast tRNA substrate. A nuclear translocation assay was performed by incubating human umbilical vein endothelial cells

with wild type and with p.R57W angiogenin protein. Subsequent immunostaining was also carried out for quantifying the results.

4. Results

4.1 Main ALS genes

A known pathogenic *SOD1* mutation was detected in heterozygous state in a 70-year-old female patient. The c.435G>C; p.L145F variant has already been found in two previously investigated Hungarian ALS patients, thus, it is the most common causative variant in our cohort.

Altogether 9.56% (13/136) of the patients carried the hexanucleotide repeat expansion of the *C9orf72* gene.

4.2 Next generation sequencing (NGS) results

In the 35 genes included in the "major ALS genes" gene set 25 variants were identified: one pathogenic, three likely pathogenic variants and 21 variants of uncertain significance (VUS).

The pathogenic variant was identified in the *ALS2* gene, while two likely pathogenic variants were detected in the

SQSTM1 gene, and one likely pathogenic variant was uncovered in the *UBQLN2* gene.

A re-categorization was also performed 5 years after the initial data analysis to assess whether new evidence has surfaced to aid the categorization of the clinical significance of the variants. In three cases the re-analysis suggested changes in the classification.

Thirty-two variants in 25 candidate or minor ALS genes were detected in 25 patients. One variant is classified as pathogenic, another one as likely pathogenic, and 30 as VUSs.

In the "Further neurodegenerative and/or neuromuscular diseases gene set" 45 variants in 29 genes were revealed (two pathogenic and 43 VUS-s).

By re-analyzing the NGS results and focusing on the genes which have been linked to ALS since the initial assessment we observed 3 VUSs and a likely pathogenic sequence alteration. The likely pathogenic *MFSD8* c.910C>T, p.Q304X could lead to premature termination of the translation.

4.3 Functional studies of the *ANG* c.169C>T p.R57W variant

In the *in silico* simulation study of the variant, the histidine in the 114th position, a catalytic site amino acid showed

a conformational switch. This conformational change implicates a reduced ribonucleolytic activity.

Furthermore, a significant difference could be observed between the nuclear translocation ability of the wild type and mutant protein which suggests that the p.R57W mutation impairs the ability of the protein to undergo nuclear translocation.

4.4 Screening of ALS genetic risk factors

Almost 9% (8.79%, 16/182) of the investigated ALS patients were found to carry the intermediate length CAG repeat expansion in the *ATXN1* gene, while only 1.12% (2/178) control individuals carried an intermediate length allele.

We detected *ATXN2* intermediate length alleles with a frequency of 18.3% (26/153) in the investigated ALS patients and 9.23% (18/195) in the control population.

Duplications of the *SMN1* gene were present in 4.05% of the ALS patients and in 4.73% of control patients (6/148, and 7/148 respectively). Homozygous deletions of the *SMN2* pseudogene were observed in 4.73% of ALS patients and in 6.76% of control individuals (7/148 and 10/148 respectively).

5. Discussion

In our study we performed a thorough genetic characterization of 183 Hungarian amyotrophic lateral sclerosis patients, diagnosed between 2008 and 2021. By utilizing Sanger sequencing, multiplex ligation-dependent probe-amplification assay, next-generation sequencing methods and fragment length analysis we screened the major and the minor ALS genes, genes associated with other neurodegenerative and neuromuscular diseases and additionally provided information on the genetic risk factors of ALS.

5.1 *C9orf72* repeat expansion and *SOD1* mutations are the most common genetic alterations in Hungarian ALS patients

The *SOD1* c.435G>C, p.L145F variant is the most commonly detected SOD1 variant in the Hungarian ALS population. Studies associate the variant with a later onset of disease and a slower progression. The phenotype of our lastly diagnosed patient with the variant fits the literature data.

C9orf72 hexanucleotide repeat expansion was detected in 9.56% (13/136) of the examined Hungarian patients which is higher than reported from European populations. *C9orf72* positive patients showed an earlier age at onset than *C9orf72*

negative patients and an abundance of bulbar symptoms was also observed.

5.2 Re-analysis of the panel sequencing data re-categorizes variants and detects rare damaging ALS associated gene variants

In the “major ALS genes” gene set a known pathogenic *ALS2* variant (c.3529G>T, p.G1177Ter) was uncovered in heterozygous form. When the variant is found in biallelic state infantile-onset ascending spastic paralysis is the caused phenotype, while heterozygous variants are associated with an adult-onset motor neuron disease. Two likely pathogenic variants were identified in the *SQSTM1* gene. Both variants localize to the highly conserved ubiquitin binding domain. A likely pathogenic variant was identified in the *UBQLN2* gene. The c.1174A>G, p.M392V variant was reported before by a Chinese research group. In our study 5.61% of patients were observed to carry more than one relevant variant in major ALS genes. Co-occurrence of different variants have been noted by others as well in ALS patients.

In the “candidate/minor ALS genes” gene set and in the “Further neurodegenerative and/or neuromuscular diseases

gene" gene set altogether 77 relevant variants were identified. However, in some cases the association between the disease and the detected variant is tenuous, since the mode of inheritance does not fit the zygosity of the variant.

During the re-analysis of the formerly acquired data we also assessed genes recently linked to ALS. A likely pathogenic nonsense *MFSD8* variant (c.910C>T, p.Q304X) was uncovered which might result in a truncated protein product. Furthermore, a variant (c.13A>G, p.K5E) altering an amino acid in the Golgi localization signal was detected in the *GLT8D1* gene. We hypothesized the protein to mislocalize and thus damage the normal ganglioside metabolism.

5.3 A novel angiogenin variant impairs both the ribonucleolytic activity and the nuclear translocation of the encoded protein

Our ribonucleolytic assay showed that the c.169C>T, p.R57W variant possesses a bit more than half of the wild type protein's ribonucleolytic activity.

Due to the amino acid change, the nuclear localization sequence of the mutant protein forms a more closed off structure compared to the wild type protein's open structure. Thus, the nuclear translocation activity of the protein shrinks

due to the decreased solvent accessible surface area of the variant protein.

Our results support the pathogenicity of the *ANG* c.169C>T, p.R57W variant.

5.4 Intermediate length repeat expansion of the *ATXN1* and *ATXN2* genes are common ALS genetic risk factors in the Hungarian ALS population

ATXN1 intermediate length alleles were uncovered in 8.79% of ALS patients and in 1.12% of control individuals which is higher than reported before. Repeat expansion in the *ATXN1* and in the *C9orf72* gene have been found to coexist frequently; in our study one patient with an aggressive, bulbar onset disease was found to harbour both alterations.

Similarly, *ATXN2* intermediate length repeat expansions were also more common among Hungarian ALS patients than in patients of other European ancestry. The reason behind this phenomenon remains unclear.

Our results suggest that neither *SMN1* duplications nor *SMN2* homozygous deletions are a potent genetic risk factors of ALS, since the frequency of these genetic variations are similar in the ALS and the control group.

6. Conclusion

In this PhD dissertation we report the first detailed genetic investigations in the Hungarian amyotrophic lateral sclerosis population that resulted in the identification of novel putative ALS-associated genetic factors.

Within the frameworks of our project, a novel angiogenin mutation was detected and functionally characterized to prove its disease-causing role. We also performed the whole exome sequencing of 21 ALS patients and carried out a thorough custom design panel sequencing of additional 86 ALS patients. These panels were reanalyzed 5 years later. Furthermore, we investigated the role of established genetic risk factors of ALS in the Hungarian ALS patient's cohort.

Our project identified the genetic background of ALS in almost 41% of our patients and our work also revealed novel genotype-phenotype correlations. This is the first through genetic study in Hungarian ALS population and with this we may contribute to the understanding of the population specific genetic determinant of the disease.