Genetic investigations in Hungarian patients affected by amyotrophic lateral sclerosis

Ph.D. Thesis

Kornélia Tripolszki M.Sc.

Graduate School of Clinical Medicine

University of Szeged

Supervisor:

Nikoletta Nagy MD, Ph.D.

Department of Medical Genetics

University of Szeged

Szeged

2017

TABLE OF CONTENTS

1 INTRODUCTION 6

1.1 Amyotrophic lateral sclerosis 6
1.1.1 Clinical symptoms 6
1.1.2 Genetic background 8
1.1.2.1 Superoxide dismutase 1 (*SOD1*) 10
1.1.2.2 TAR DNA binding protein (*TARDBP*) 10
1.1.2.3 Angiogenin (*ANG*) 11
1.1.2.4 Chromosome 9 open reading frame 72 (*C9ORF72*) 11
1.1.2.5 Senataxin (*SETX*) 12
1.1.2.6 Fused in sarcoma (*FUS*) 12

1.2 Aims 13

2 PATIENTS AND METHODS 14

2.1 Patients 14

2.2 Methods 14
2.2.1 DNA isolation 14
2.2.2 Mutation Screening – Sanger Sequencing 14
2.2.3 Repeat expansion analysis 15
2.2.3.1 Repeat-primed PCR and Fragment length analysis 15
2.2.3.2 Genotyping 15
2.2.4 Targeted High-troughput sequencing 16
2.2.4.1 Library preparation and sequencing 16
2.2.4.2 Bioinformatic analysis 16

3 RESULTS 18

3.1 *SOD1* gene 18
3.1.1 Novel mutation detected in the *SOD1* gene 18
3.1.1.1 Clinical features of the patient with the p.Lys91ArgfsTer8 *SOD1* mutation 19
3.1.2 Detected known mutations of the *SOD1* gene 20

3.2 *TARDBP* gene 23

3.3 *ANG* gene 23

3.4 *C9orf72* gene 25
3.4.1 Clinical symptoms of the patient with the *C9orf72* RE 26

3.5 *SETX* gene 27
3.5.1 Clinical characterictics of the patient with the novel p.Asn264Ser *SETX* mutation 28

4 DISCUSSION 30

4.1 Mutations in the *SOD1* gene 31

4.2 Mutations in the *TARDBP* gene 33

4.3 Mutations in the *ANG* gene 33

4.4 Repeat expansion of the *C9orf72* gene 34

4.5 Mutations of the *SETX* gene 35

4.6 Mutations in the *FUS* gene 38

5 SUMMARY 39

6 ACKNOWLEDGEMENTS 41

7 ELECTRONIC DATABASE INFORMATION 42

8 REFERENCES 44

9 APPENDIX 50

LIST OF PUBLICATIONS

Publications providing the basis of the dissertation

I. **Tripolszki K**, Csányi B, Nagy D, Ratti A, Tiloca C, Silani V, Kereszty É, Török N, Vécsei L, Engelhardt IJ, Klivényi P, Széll M, Nagy N. Genetic analysis of the SOD1 and C9orf72 genes in Hungarian patients with amyotrophic lateral sclerosis. Neurobiol Aging 2017. Published online: 2017.01.29. **IF: 5.153**

II. **Tripolszki K,** Török D, Goudenege D, Farkas K, Sulák A, Török N, Engelhardt J, Klivényi P, Procaccio V, Nagy N, Széll M. High-troughput sequencing revealed a novel SETX mutation in a Hungarian ALS patient. Brain Behav 2017. Accepted manuscript **IF: 2.128**

III. **Tripolszki K,** Nagy ZsF, Nagy D, Török N, Klivényi P, Engelhardt IJ, Vécsei L, Nagy N, Széll M. *ANG* and *TARDBP* mutations in Hungarian patients with amyotrophic lateral sclerosis. Manuscript in preparation.

Publications indirectly related to the subject of the dissertation

I. Kinyo A, Valyi P, Farkas K, Nagy N, Gergely B, **Tripolszki K,** Torok D, Bata-Csorgo Z, Kemeny L, Szell M. A newly identified missense mutation of the EDA1 gene in a Hungarian patient with Christ–Siemens–Touraine syndrome. Arch Derm Res 2014; 306(1):97-100. **IF:2.270**

II. Nagy N, Valyi P, Csoma Z, Sulak A, **Tripolszki K**, Farkas K, Paschali E, Papp F, Toth L, Fabos B, Kemeny L, Nagy K, Szell M. CTSC and Papillon–Lefèvre syndrome: detection of recurrent mutations in Hungarian patients, a review of published variants and database update. Molecular Genetics & Genomic Medicine 2014; 2(3):217-228.

III. Vályi P, Farkas K, **Tripolszki K**, Sulák A, Széll M, Nagy N, Nagy K. Rekurrens európai misszensz mutáció egy magyar Papillon-Lefévre szindrómában szenvedő családban. Fogorvosi Szemle 2014; 107(3):87-92.

IV. Nagy N, Farkas K, **Tripolszki K**, Sulák A, Kemény L, Széll M. A cylindromatosis gén mutációi által okozott genodermatosisok. Bőr Vener Szemle 2014; 90:(5) 185-193.

V. Sulak A, Toth L, Farkas K, **Tripolszki K**, Fabos B, Kemeny L, Valyi P, Nagy K, Nagy N, Szell M. One mutation, two phenotypes: a single nonsense mutation of the CTSC gene causes two clinically distinct phenotypes. Clin Exp Dermatol 2016; 41(2):190-195. **IF: 1.092**

VI. **Tripolszki K**, Knox R, Parker V, Semple R, Farkas K, Sulák A, Horváth E, Széll M, Nagy N. Somatic mosaicism of the PIK3CA gene identified in a Hungarian girl with macrodactyly and syndactyly. Eur J Med Genet Apr;59(4):223-6. **IF: 1.81**

VII. **Tripolszki K**, Farkas K, Sulák A, Duga B, Melegh B, Knox R, Parker V, Semple R, Kemény L, Széll M, Nagy N. Atypical neurofibromatosis type 1 with unilateral limb hypertrophy mimicking overgrowth syndrome. Accepted manuscript. Clin Exp Dermatol. Accepted on 28 June 2016. **IF: 1.315**

1 INTRODUCTION

1.1 Amyotrophic lateral sclerosis

1.1.1 Clinical symptoms

Amyotrophic lateral sclerosis (ALS; ORPHA803) is a fatal, neurodegenerative disorder characterized by the death of motor neurons in the brain, brainstem and spinal cord. The characteristic clinical course of ALS is a progressive loss of voluntary movement, with symptoms spreading to more distant locations, resulting in paralysis and death from respiratory failure (Morrison and Harding, 1994). *Amyotrophic* refers to the lack (*a*-) of muscle (-*myo*-) nourishment (-*trophic*), resulting in wasting of the fibers; *Lateral* refers to the lateral corticospinal tract of affected neurons between the brain and spinal cord; *Sclerosis* is the resultant hardening of the tissue. ALS is also termed Charcot's disease after Jean Martin Charcot (French clinician) who first described its features in 1869 (Charcot, 1869) or Lou Gehrig's disease in the United States, after the famous New York Yankees baseball player afflicted with the condition.

Originally, the spectrum of adult motor neuron diseases (MND) included ALS with the combined degeneration of upper (cortical) and lower (pontobulbar and spinal) motor neurons, primary lateral sclerosis (PLS) with only upper motor neuron lesion and lower MND with the damage of pontobulbar and/or spinal motor neurons (Victor and Ropper, 2000). In a recent work, Finsterer and Burgunder (2014) suggested that ALS can manifest clinically as a continuum ranging from exclusive impairment of the upper motor neurons to exclusive impairment of the lower motor neurons. These observations have resulted in the now commonly held belief that ALS is a syndrome rather than a specific disease. As a result of the heterogeneity of presentation the exclusion of ALS mimics is an extremely important part of diagnosis (Chio *et al.* 2011; Chang RCC, 2011).

Exclusive impairment of the upper motor neurons can manifest in PLS characterized by spastic paresis of the striated muscles without atrophy and denervation, increased

deep tendon reflexes, pathological reflexes designating upper motor neuron lesion (Finsterer and Burgunder, 2014). Exclusive impairment of the lower motor neurons can contribute to the development of lower MNDs observed as atrophy of the striated muscles of the body due to degeneration of lower motor neurons with increasing weakness, decreased muscle tone, electrophysiological signs of denervation (fasciculations, fibrillations), lack of the deep tendon reflexes and, finally, inevitable death from respiratory failure (Finsterer and Burgunder, 2014; Victor and Ropper, 2000). The onset is usually asymmetrical, and a smaller group of patients have "bulbar onset" ALS, which targets muscles of the face and neck. This form has a worse prognosis and it is more common in women (Forbes, et al., 2004; McCombe and Henderson, 2010).

Familial forms account for about 10% of ALS cases, although higher levels have been reported in certain geographical regions (Murros and Fogelholm, 1983). The mean age of onset (at approximately 48-52 years) for familial ALS (FALS) is earlier than the mean age (at about 56 years) for sporadic ALS (SALS) and the survival period is also shorter for FALS than for SALS (Mulder *et al.* 1986; Li *et al.,* 1988; Strong *et al.,* 1991; Hewitt *et al.,* 2010). A small minority of ALS patients show juvenile onset (typically before age 25), which is usually characterized by longer survival (Aggarwal and Shashiraj, 2006). The male to female ratio is 1.6 to 1. The lifetime risk for developing the disease is approximately 1/400 (Johnston *et al.,* 2006). ALS has a very poor prognosis, with a median survival period of three years after the onset of the initial symptoms (Hardiman *et al.,* 2011). There is no cure for this relentless disease, although the drug riluzole, which blocks the release of glutamate, has been shown to slow disease progression (Bensimon *et al.,* 1994). Despite numerous attempts at identifying other drugs, no further agents have risen to the level of significance.

1.1.2 Genetic background

The genetics of ALS are of research interest because it helps to uncover the mechanism of cell death in ALS. Neuronal cytoplasmic protein aggregation and defective RNA metabolism show to be frequent pathogenic mechanisms involved in ALS. Regarding its genetic background, more than 20 major ALS genes (Table 1) have been implicated in the Mendelian ALS forms and further about 100 genes have been associated as predisposing factors with the non-Mendelian variants (ALSoD Database, http://alsod.iop.kcl.ac.uk). ALS can be inherited in an autosomal dominant, autosomal recessive or X-linked manner. Familial forms are mainly transmitted in a Mendelian pattern of autosomal dominant inheritance (Hardiman *et al.,* 2011). Several twin studies have estimated the genetic contribution to the risk for ALS to be quite high (61%), but the genetic background remains poorly understood (Al-Chalabi *et al.,* 2010). Non-genetic risk factors, including environmental factors and physical injury, are being actively investigated. Although the contribution of these risks is difficult to expain (e.g. the prevalence of ALS in athletes), the involvement of these factors may enhance underlying, genetic predispositions to the disease.

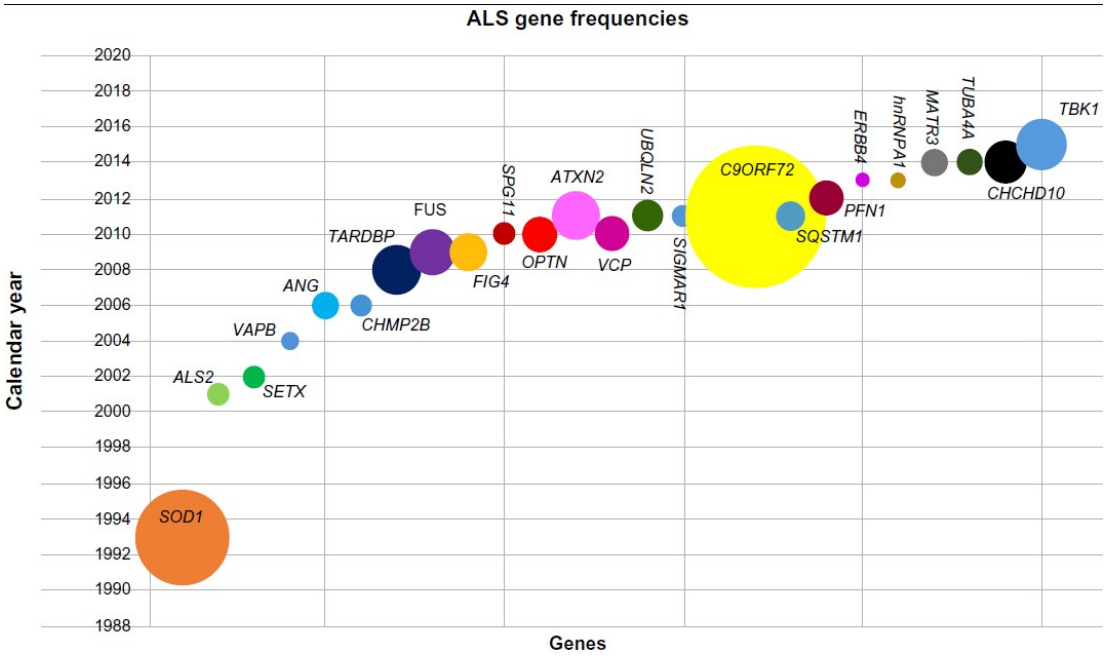

Figure 1. Gene frequencies in ALS - each gene is plotted against the year it was found; the size of the circles signifies the frequency of mutations in FALS or ALS as cited in the literature (Alsultan, *et al.*, 2016)

Genetic subtype	Gene	Inheritance	Chromosome
ALS1	*SOD1*	AD, AR	21q22.11
ALS2	*Alsin*	AR	2q33.2
ALS3	*Unknown*	AD	18q21
ALS4	*SETX*	AD	9q34.13
ALS5	*SPG11*	AR	15q14
ALS6	*FUS*	AD, AR	16p11.2
ALS7	*Unknown*	AD	20p13
ALS8	*VAPB*	AD	20q13.33
ALS9	*ANG*	AD	14q11.1
ALS10	*TARDBP*	AD,AR	1p36.22
ALS11	*FIG4*	AD	6q21
ALS12	*OPTN*	AD, AR	10p13
ALS13	*ATXN2*	AD	12q23-q24.1
ALS14	*VCP*	AD	9p13
ALS15	*UBQLN2*	SD	Xp11.21
ALS16	*SIGMAR1*	AD	9p13
ALS17	*CHMP2B*	AD	3p12.1
ALS18	*PFN1*	AD	17p13.3
ALS19	*ERBB4*	AD	2q33.3-q34
ALS20	*HNRNPA1*	AD	12q13.1
ALS21	*MATR3*	AD	5q31.2
ALS-FTD1	*C9orf72*	AD	22q11.23
ALS-FTD2	*CHCHD10*	AD	9p21.2
ALS	*TBK1*	AD	12q14.2

Table 1. The major ALS genes that are implicated in the Mendelian forms of ALS. (ALSoD Database). The genes investigated in this study are shaded.

1.1.2.1 Superoxide dismutase 1 (*SOD1*)

Among the ALS causative genes, superoxide dismutase 1 (*SOD1*) is one of the most commonly mutated genes and accounts for approximately 12–23% of the familial and up to 7% of the sporadic ALS forms (Andersen, 2006). *SOD1* gene encodes the Cu/Zn superoxide dismutase enzyme, which catalyzes the inactivation of superoxide into oxygen and hydrogen peroxide, providing antioxidant defense (Smirnoff, 1993). The SOD1 enzyme is widely expressed and constitutes approximately 0.5–0.8% of the soluble proteins in the brain (Pardo *et al.,* 1995). Proper coordination of the Cu^{++} and Zn^{++} ions provides conformational steadiness to the dimeric SOD1 protein as well as extreme resistance against denaturing conditions (Arnesano *et al.,* 2004). To date, more than 170 mutations have been reported for *SOD1* in the Amyotrophic Lateral Sclerosis Online Genetics Database (ALSoD Database; Abel *et al.,* 2012) since the gene was firstly associated to ALS in 1993 (Rosen *et al.,* 1993). *SOD1* mutations occur in all the five exons of the gene.

1.1.2.2 TAR DNA binding protein (*TARDBP*)

Human TDP-43 was discovered in a screen for transcriptional repressors of the transactive response (TAR) DNA binding element of the HIV-1 virus in 1995 (Ou *et al.,* 1995). TDP-43 is a 414-amino acid nuclear protein encoded by the *TARDBP* gene on chromosome 1p36.22 and, because of its molecular weight of 43kDa it was named TDP-43 (Buratti and Baralle, 2008). Since the TDP-43 association in ALS was initially described in 2006 (Neumann *et al.,* 2006) more than 40 mutations have been found in familial and sporadic ALS cases (ALSoD Database; Lattante *et al.,* 2013). The *TARDBP* gene contains 6 exons, and with one exception (p.D169G in exon 4), all the identified mutations are located in the last exon (Kabashi *et al.,* 2008; ALSoD Database). TDP-43 is mainly expressed in the nucleus and contains two RNA recognition motifs (RRM1 and RRM2) and a glycine-rich C-terminal region (GRR). TDP-43 is a DNA/RNA binding protein, and its functions include involvement in gene transcription, RNA splicing, microRNA processing and transport of mRNA (Chen *et al.,* 2013). *TARDBP* gene is highly conserved from human to *C. elegans*. According to in situ hibridization studies described by Shakaran *et al.* (2008), TDP-43 is expressed early in the spinal cord and brain of

zebrafish. Ubiquinated inclusions were found in the motor neurons of ALS patients and TDP-43 was the major component of these inclusions (Neumann *et al.,* 2006).

1.1.2.3 Angiogenin (*ANG*)

The *ANG* gene encodes angiogenin, a 123-residue, 14.1-kDa protein and it is located on chromosome 14q11.2. The ANG protein belongs to pancreatic ribonuclease superfamily, and it plays a role in rRNA biogenesis, cellular proliferation and has a crucial role in inhibiting protein translation by cleaving tRNA. It is known that ANG mediates neovascularization and promotes neurite outgrowth during early embryonic development and it protects against hypoxia-induced motor neuron death (Sebastia *et al.,* 2009; Chen *et al.,* 2013).

Greenway *et al.* (2004) identified *ANG* as a candidate gene for amyotrophic lateral sclerosis in Irish and Scottish populations. ALS caused by mutations in the *ANG* gene is named as ALS9, an autosomal dominant adult onset disease. To date, 29 variants have been described for *ANG* in the ALSoD database. Mutations in ANG gene cause loss of ribonucleolytic activity and nuclear translocation activity (Chen *et al.,* 2013). The ANG (123-residue protein) is synthesized with a signal peptide of 24 amino acids that is cleaved to form the mature protein. The mature protein contains a receptor binding region, an RNase A canonical domain, a nuclear localization signal sequence, and a catalytic site responsible for a low ribonuclease activity (Sheng and Xu, 2016).

1.1.2.4 Chromosome 9 open reading frame 72 (*C9ORF72*)

A frequently mutated ALS gene is chromosome 9 open reading frame 72 (*C9orf72*), which – in addition to the *SOD1* mutations – is now recognized as the main cause of familial and sporadic ALS (Majounie *et al.,* 2012; Gijselinck *et al.,* 2012; Ratti *et al.,* 2012; Smith *et al.,* 2013). A hexanucleotide (GGGGCC) repeat expansion (RE) located in the non coding region of the gene that can reach up to 4400 units (normal range: 2-23 units) has been identified in patients with ALS and/or frontotemporal dementia. The GGGGCC RE contributes to 23–47% of familial ALS and to 4–5% of sporadic cases (Renton *et al.,* 2011; DeJesus-Hernandez *et al.,* 2011; Byrne *et al.,*

2012; Ratti *et al.,* 2012), with a frequency depending on geographical origin. Although the pathomechanism with which the hexanucleotide RE leads to the development of ALS has not been elucidated completely, both C9orf72 haploinsufficiency gain of function mechanisms (driven by toxicity of sense and antisense RNA transcripts and derived dipeptide repeat proteins) have been reported (Taylor *et al.,* 2016). Apart from the repeat expansion a splice site mutation has been identified in the *C9orf72* gene (Liu *et al.,* 2016).

1.1.2.5 Senataxin (*SETX*)

Mutations in the senataxin (*SETX*) gene have been identified at a lower frequency than in the *SOD1* gene in ALS. *SETX* encodes a helicase protein involved in DNA repair and RNA production. Homozygous or compound heterozygous *SETX* mutations are associated with the development of autosomal recessive ataxia with oculomotor apraxia type 2 (AOA2) (Anheim *et al*., 2009). In addition, heterozygous *SETX* mutations have been associated with the autosomal dominant form of juvenile-onset ALS (ALS4) (Chen *et al.,* 2004).

1.1.2.6 Fused in sarcoma (*FUS*)

The fused in sarcoma (*FUS*) gene is also associated with the Mendelian forms of ALS. *FUS* is located on chromosome 16p11.2 and encodes for a 526 amino acid protein. FUS is a nucleoprotein that plays a role in transcription, splicing, and shuttling of RNA from the nucleus to the cytoplasm. Fused in sarcoma also acts as a transcriptional regulatory sensor of DNA damage signals and therefore is required in maintaining the integrity of the genome (Zinszner *et al.,* 1997). To date, 80 mutations of the *FUS* gene have been described (ALSoD Database) since the gene was firstly associated to ALS in 2009 (Vance *et al.,* 2009; Kwiatkowski *et al.,* 2009).

1.2 Aims

In my thesis, the primary aim was to summarize the results of my genetic investigations in a group of Hungarian patients (n=66) suffering from ALS. I investigated the contributions of six commonly mutated ALS genes, *SOD1, TARDBP, ANG, FUS, SETX* and *C9orf72*. Besides, the genetic investigations aiming to identify the disease-causing mutations, it was also among my goal to compare these variants with the reported ones in the literature in order to define genotype-phenotype correlations and Hungarian population specific mutations.

This study represents the first genetic screening of ALS in Hungary, which adds novel data to the genetic and phenotypic diversity of this disease.

2 PATIENTS AND METHODS

2.1 Patients

The unrelated patients (n=66) included in this study were recruited from the Department of Neurology, University of Szeged, Szeged, Hungary, between 2010 and 2016. The group consisted of 45 females and 21 males and the median of the age at onset was 60. All patients fulfilled the revised El Escorial and the Awaji-shima criteria for ALS (de Carvalho *et al.,* 2009; Ludolph *et al.,* 2015). According to the revised El Escorial criteria, the lower motor neuron disease (progressive muscular atrophy: PMA) is determined as one of the ”restricted phenotypes” of ALS, therefore one patient with only lower motor neuron involvement was also diagnosed as ALS. All patients and age- and sex-matched healthy controls (n=110) were of Hungarian ancestry. The investigation was approved by the Internal Ethical Review Board of the University of Szeged. Written informed consent was obtained from patients and healthy controls, and the study was conducted according to the Principles of the Declaration of Helsinki.

2.2 Methods

2.2.1 DNA isolation

Blood samples from all the enrolled individuals (n=66) were collected after the individuals had signed an informed consent document. Genomic DNA was isolated using a QIAamp DNA Blood Mini Kit (QIAGEN; Hilden, Germany).

2.2.2 Mutation Screening – Sanger Sequencing

The entire coding region and the flanking introns of the *SOD1, TARDBP* and *ANG* genes were amplified (primer sequences used were taken from the UCSC Genome Browser www.genome.ucsc.edu). PCR was performed using Dream Taq Green PCR Master Mix (Fermentas) according to the manufacturer’s instructions. Direct sequencing of the PCR products was performed on an ABI 3100 sequencer and compared with the wild-type gene sequences at the Ensemble Genome Browser (http://ensemble.org). To identify known variations, I used ALS Online Genetics

Database (http://alsod.iop.kcl.ac.uk/) (Abel et al., 2012), 1000 Genomes Database (www.1000genomes.org/), dbSNP (http://www.ncbi.nlm.nih.gov/project/SNP) and Exome Aggregation Consortium (ExAC) database (http://exac.broadinstitute.org). To predict the functional effects of novel mutations, the sequence variations were assessed by in silico prediction programs, such as SIFT (http://sift.bii.a-star.edu.sg/), Polyphen-2 (http://genetics.bwh.harvard.edu/pph2) and Mutation Taster (http://mutationtaster.org). To examine possible effects of the mutations on the three-dimensional (3-D) structure of the SOD1 protein, I used the Swiss-Model protein structure homology-modeling server (http://swissmodel.expasy.org/; SOD1 Protein Data Bank accession number 4b3e.1.A).

Fragments with identified mutations were independently re-amplified and resequenced from both ends.

2.2.3 Repeat expansion analysis

2.2.3.1 Repeat-primed PCR and Fragment length analysis

A two-step protocol proposed by Akimoto *et al*. (2014) was followed for the detection of the GGGGCC hexanucleotide RE in the *C9orf72* gene. Fragment length analysis was performed using GeneMapper ID v3.2.1, and the samples producing a single peak product were further analyzed in the second step by repeat-primed PCR using 310 ABI Prism Genetic Analyzer (Applied Biosystems). The peaks were visualized using GeneMapper ID v3.2.1 software (Akimoto *et al.,* 2014). The presence of the GGGGCC RE was observed as a saw-tooth pattern with 6-base pair periodicity.

2.2.3.2 Genotyping

To determine whether the single individual carrying the GGGGCC RE identified in this study also carried the “risk” haplotype, I selected the rs3849942 variant to be used as a marker for the “risk” haplotype for the patient and control genotypes. Rs3849942 genotyping was based on allelic discrimination assays using TaqMan chemistry following the manufacturer’s instructions (Life Technologies; Budapest, Hungary).

2.2.4 Targeted High-troughput sequencing

2.2.4.1 Library preparation and sequencing

Amplicons (n=56) were designed (range 519–704 bp; mean: 612 bp) to cover the coding regions and the flanking introns of the investigated *FUS, SETX* and *C9orf72* genes, and amplicon libraries were prepared according to the Amplicon Library Preparation Manual for the Roche Junior 454 next generation sequencing system (Roche; Budaörs, Hungary). Template-specific primers were used in the first round PCR, which represented the target-specific step of the library preparation. The target specific primers provided ’universal tails’. Then a second round of PCR was carried out, targeting the ’universal tails’ and adding the MID-labeled Primer A and Primer B sequences to barcode the samples. For each DNA sample, separated amplicon PCR amplifications were performed using FastStart High Fidelity PCR System (Roche; Budaörs, Hungary) with 50ng genomic DNA per reaction. PCR products were visualized on a 2.0% agarose gel by electrophoresis.Amplicons were purified with the Agencourt AMPure XP kit (Beckman Coulter; Budapest, Hungary), quantified with the Quant-iT PicoGreen Assay (Life Technologies; Budapest, Hungary), diluted separately to $1x10^7$ molecules/µl and pooled. Emulsion PCR and next generation sequencing were performed according to the manufacturers’ protocols (Roche).

2.2.4.2 Bioinformatic analysis

To improve the efficiency of the Roche pipeline, composed of the Roche 454 GS Reference Mapper for mapping (on UCSC human reference genome hg19) and Amplicon Variant Analyzer (v2.5p1) for variant calling, an additional in house pipeline was used. Sequencing reads were aligned to the reference genome (UCSC hg19) using Roche 454 GS Reference Mapper and Amplicon Variant Analyzer (version 2.5p1). BAM files were converted to FASTQ and realigned to hg19 using Bowtie2 V.2.2.6 (Langmead *et al.,* 2012). Samtools V.0.1.19 and Python V.2.7.11 were used for additional file handling. Variant calling was performed combining GATK Unified Genotyper (Genome Analysis Tool Kit v.3.4) (McKenna *et al.,* 2010) and Platypus (v.0.8.1) (Rimmer *et al.,* 2014). The annotation and the prioritization

steps were executed with ANNOVAR (15.06.07) (Wang *et al.,* 2010). Finally, the target region reads depth was controlled using BEDTools (v.2.25.0) (Quinlan *et al.,* 2010) and the alternative allele frequency was computed using pysamstats (v.0.24.2) (https://github.com/). Candidate variants were filtered according to read depth, allele frequency and prevalence in genomic variant databases such as ExAc (v.0.3), ClinVar, Kaviar (including 1000g-phase3, dbSNP146) (Glusman *et al.,* 2011; Landrum et al., 2016). Variant prioritization tools (PolyPhen2, SIFT, LRT, MutationTaster, Mutation Assessor) were used to predict the functional impact and then focused on variants with predicted deleterious consequences (nonsense SNVs, frameshift indels, essential splice variants and complex indels). The putative effect on splicing efficiency was predicted using the Human Splicing Finder (Desmet *et al.,* 2009). The alignments were visualized with IGV V.2.0.34 (Robinson *et al.,* 2011). All the identified disease-causing candidate variants were confirmed by direct sequencing.

3 RESULTS

3.1 SOD1 gene

The direct DNA sequencing approach identified four different mutations of the *SOD1* gene in 5 ALS patients: three known heterozygous missense mutations (c.43G>A p.Val14Met; c.272A>C p.Asp90Ala; c.435G>C p.Leu144Phe) and one novel mutation (c.275_276delAA, p.Lys91ArgfsTer8) (Table 2.).

3.1.1 Novel mutation detected in the *SOD1* gene

The detected novel heterozygous mutation (c.275_276delAA, p.Lys91ArgfsTer8) is located in the fourth exon of the *SOD1* gene (Figure 2) and led to a frameshift with the insertion of 8 novel amino acids and the formation of premature stop codon at the new amino acid position 99. Analysis using Mutation Taster software predicted that the p.Lys91ArgfsTer8 mutation causes severe truncation of the encoded enzyme, and, thus, I hypothesize that this mutation is likely to be pathogenic (Figure 3). The pathogenic role of the p.Lys91ArgfsTer8 mutation is further supported by the fact that it interferes with the integrity of the Cys57-Cys146 disulfide bond, and results in the weakening of the dimer interface.

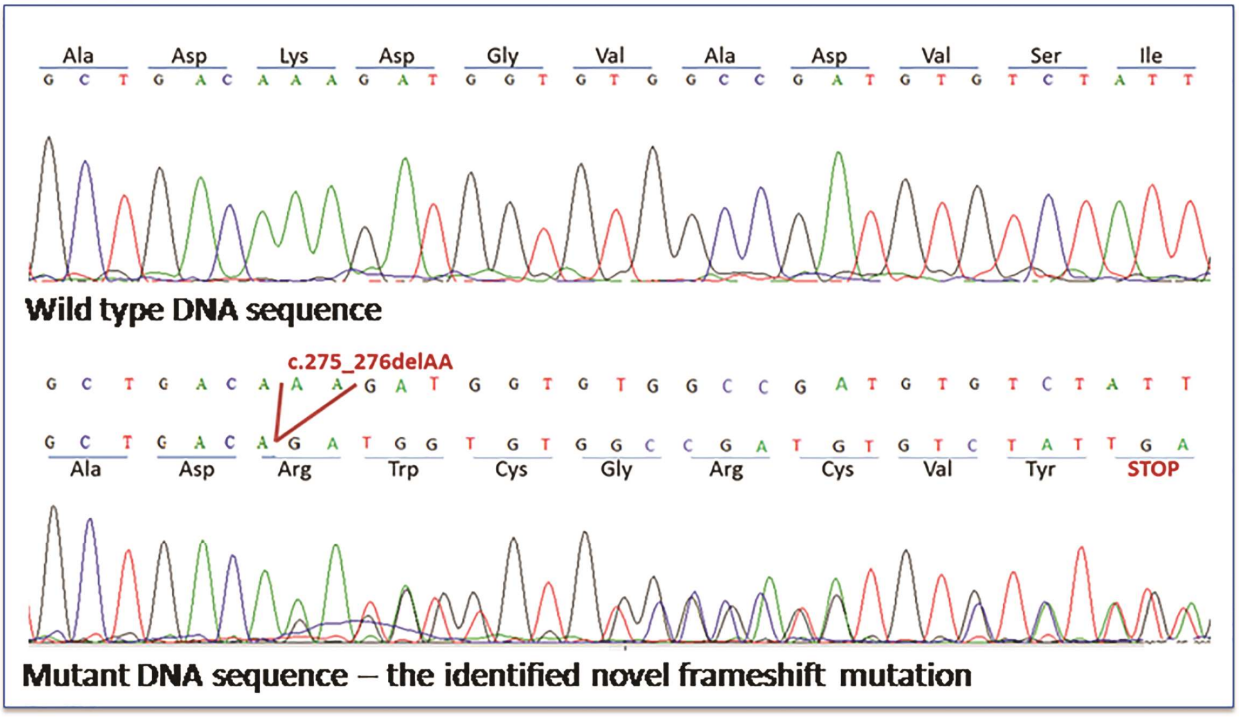

Figure 2. A novel p.Lys91ArgfsTer8 mutation in the Cu-Zn superoxide dismutase 1 (*SOD1*) gene identified by direct sequencing (Tripolszki *et al.,* 2017a)

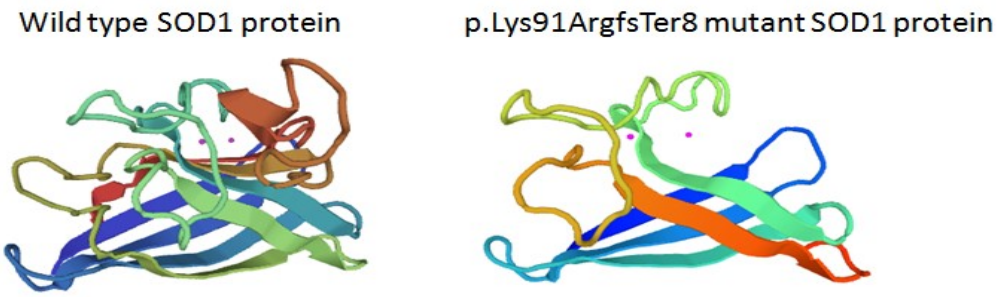

Figure 3.The crystallographic model shows that the novel frameshift mutation causes a severe truncation of the protein (Tripolszki *et al.,* 2017a)

The novel p.Lys91ArgfsTer8 *SOD1* mutation was not present in 110 healthy controls of Hungarian ancestry that I investigated, neither it is represented in mutation databases, including Single Nucleotide Polymorphism Database (http://www.ncbi.nlm.nih.gov/projects/SNP/), 1000 Genomes (www.1000genomes.org/), Exome Aggregation Consortium (ExAC) database (http://exac.broadinstitute.org) and the ALS Online Genetics Database (http://alsod.iop.kcl.ac.uk/) (Abel *et al.,* 2012). Unfortunately, the segregation of this mutation could not be investigated because DNA samples from other family members were not available.

3.1.1.1 Clinical features of the patient with the p.Lys91ArgfsTer8 SOD1 mutation

This novel mutation was carried by a female Hungarian patient suffering from sporadic ALS, who reported a story of breast cancer eight years before the onset of ALS disease. The breast cancer, located on the left side, was surgically removed and irradiation therapy and anti-estrogen medication was introduced after the operation. On subsequent follow up investigation, the patient was tumor-free. However, eight years after the breast cancer, her lower extremities became clumsy and weak because of spasticity and muscle atrophy as a result of right peroneal palsy. On examination, the patellar reflexes were hyperactive while the Achilles reflexes were diminished. The weakness showed distal predominance. The muscle strength of the upper extremities were preserved, but the deep tendon reflexes were exaggerated. She also had dysphagia with increased gag reflex, dysarthria and bilateral recurrent nerve

palsy with coarse voice. The calculated ALS Functional Rating Scale R (ALSFRS-R score; Cedarbaum *et al.,* 1999) was 39/48 at the first examination, which was four months after the appearance of the first symptoms. Fasciculations of the striated muscles in all the extremities were seen. Electromyography (EMG) showed signs of denervation (fasciculations, fibrillations, positive sharp waves), a complete interference pattern did not develop in full effort and several enlarged motor units were recorded. Electroneurography (ENG) exhibited motor axonal loss. Serum level of creatine kinase was 238 U/l (normal < 170). The serum levels of thyroid hormones and parathyroid hormone were normal; however, the anti-thyroglobulin IgG level was almost five times higher than the normal value: 543.5 iu/ml (normal < 115). There were no alterations in the immunoglobulin levels of the serum, and monoclonal gammopathy or elevated anti-ganglioside IgG or IgM were not detected. Magnetic resonance imaging (MRI) demonstrated spondylosis and multiple protrusions of the intervertebral discs in the cervical and lumbar regions of the spinal cord without stenosis of the spinal canal, myelopathy and root lesions. MRI of the scull revealed only a few lacunar infarcts in the white matter of the cerebrum. The patient had been treated for hypertension and hypercholesterolemia for several years. She was given 2x50 mg riluzole/day; subsequently her status deteriorated quickly, and she became tetraplegic with dysarthria and dysphagia. Respiratory failure developed because of the weak respiratory muscles and she died at home.

3.1.2 Detected known mutations of the *SOD1* gene

The *SOD1* p.Leu144Phe mutation is located in the fifth exon of the gene and was identified in two female ALS patients (Figure 4). One of the patients reported that her maternal grandfather had suffered from non-progressive paralysis and weakness for 20 years. The other patient reported that her paternal grandmother had suffered from a disease similar to her own. In this study, only the latter patient was therefore considered as having a familial form of the disease.

The p.Val14Met mutation is located in the first exon of the *SOD1* gene and was present in an affected female patient who reported no family history of ALS (Figure 5). ALS symptoms first appeared at the age of 62. This patient showed lower and upper motor neuron signs. The disease course was progressive and led to the patient's death within one year after disease onset.

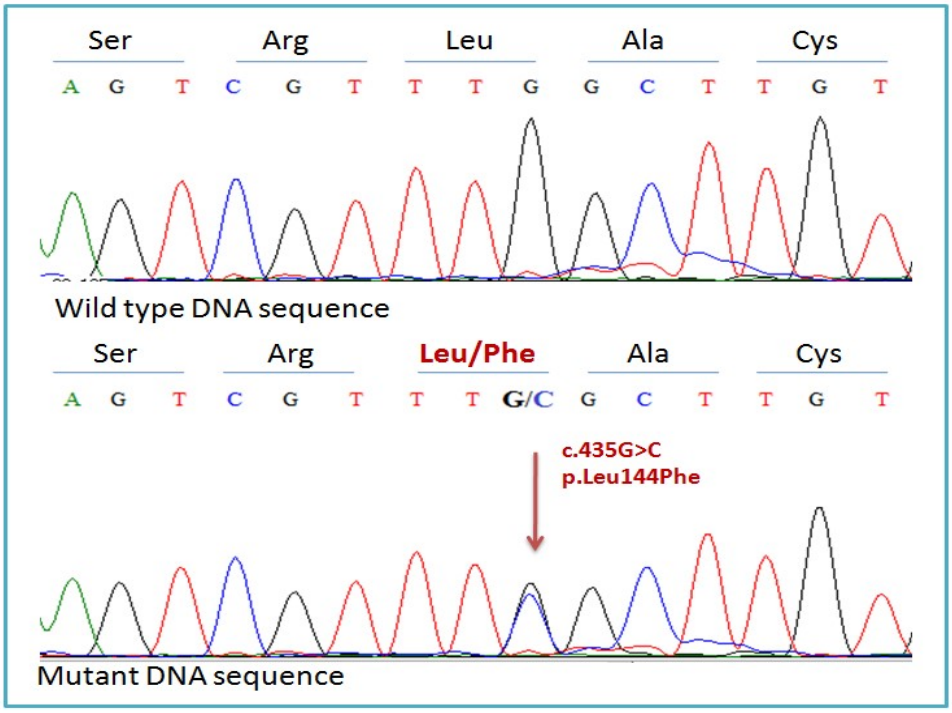

Figure 4. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Leu144Phe mutation.

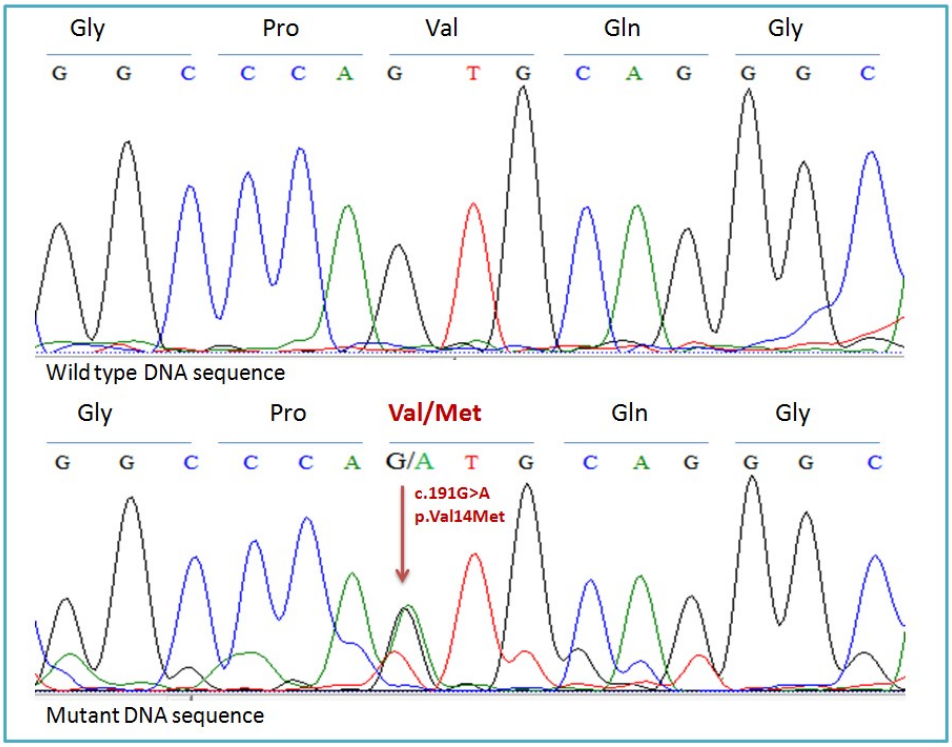

Figure 5. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Val14Met mutation.

The p.Asp90Ala mutation, which is the most prevalent *SOD1* mutation in Europe (Andersen *et al.,* 1995; Al-Chalabi *et al.,* 1998; Andersen, 2001), is located in the fourth exon of the *SOD1* gene and was present in a female patient (Figure 6). Interestingly, the patient also carried the rs111273304 splice-donor variant (c.239+2T>A), which is of unknown significance, in heterozygous form. This patient had clinical features typically associated with this genotype, including a relatively long survival after onset (Andersen, 2006).

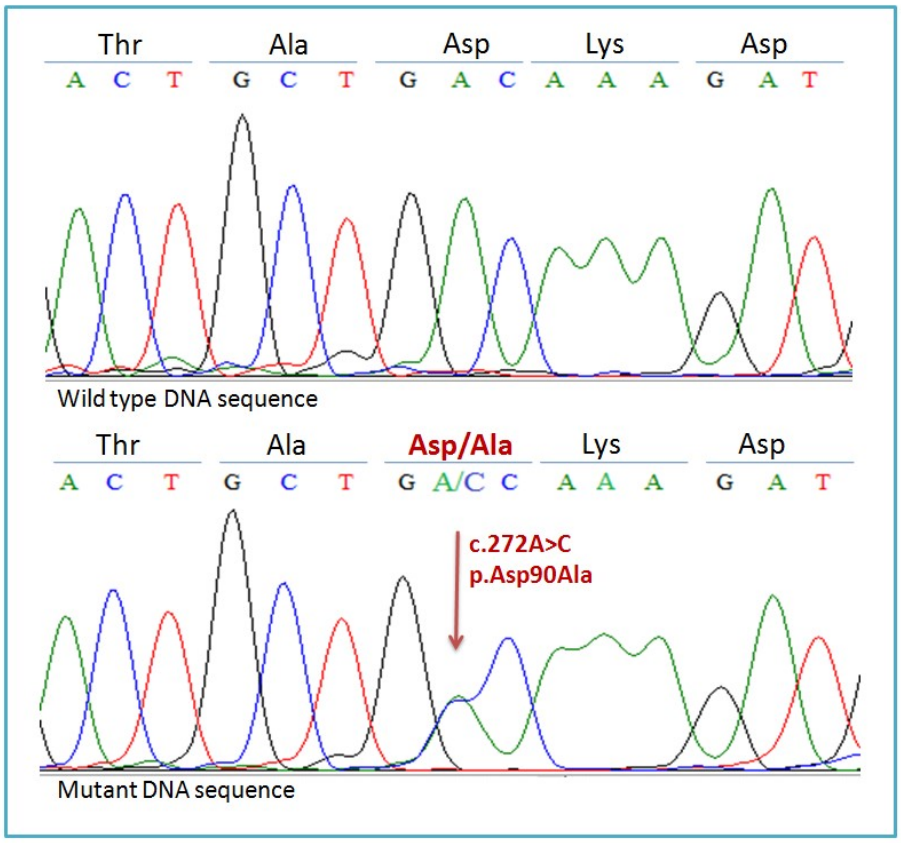

Figure 6. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Asp90Ala mutation.

3.2 TARDBP gene

Direct sequencing of the *TARDBP* gene revealed a previously described missense mutation in exon 6 (c.931A>G, p.Met311Val) (Figure 7). The variant was first described by Lemmens *et al.* (2009) in a family of Belgian origin. The Hungarian patient with the p.Met311Val variant reported no family history and the ALS symptoms first appeared at the age of 62.

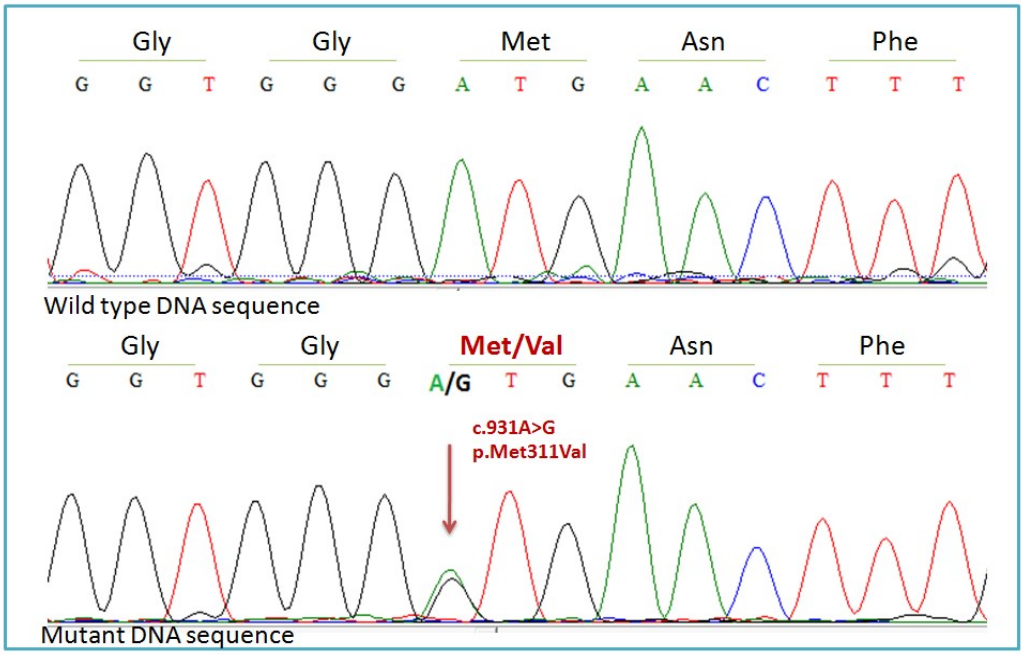

Figure 7. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Met311Val mutation.

3.3 ANG gene

Mutation analysis of the angiogenin gene revealed two heterozygous mutations (c.3G>A, p.Met-24Ile[#], p.Met1Ile[*]; c.169C>T, p.Arg33Trp[#], p.Arg57Trp[*]).

The p.Met-24Ile mutation is located in the signal peptide region of the protein, and it was initially described by Conforti *et al.* (2008) in a sporadic patient from Italy. This mutation (Figure 8) was detected in the same Hungarian patient who carried the SOD1 p.Val14Met mutation.

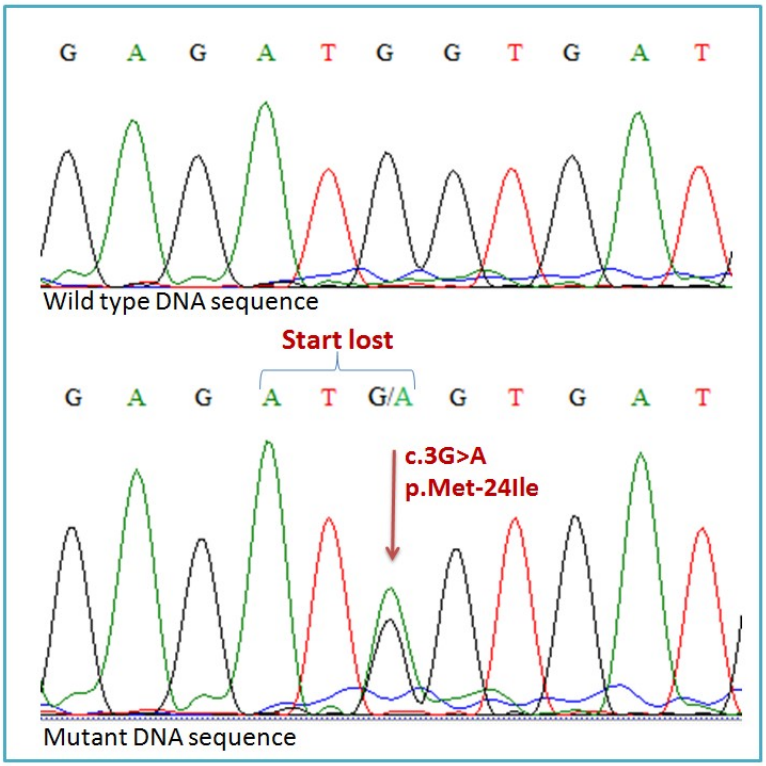

Figure 8. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Met-24Ile mutation.

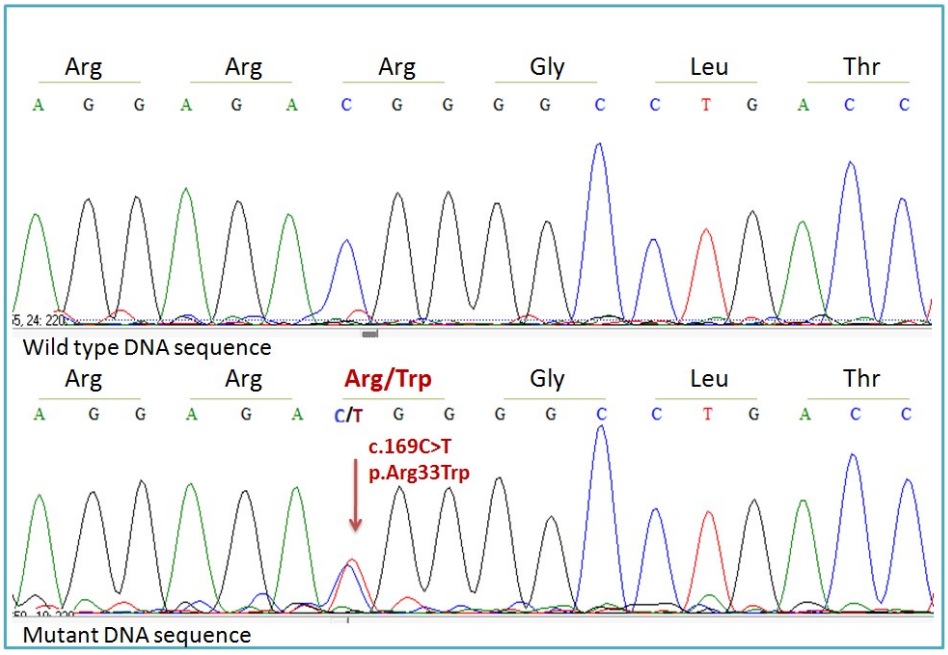

Figure 9. The wild type DNA sequence and the DNA sequence of the affected patient that shows the p.Arg33Trp mutation.

The second *ANG* mutation detected in this cohort (c.169C>T, p.Arg33Trp, p.Arg57Trp[*]) affects the nuclear translocation signal of the angiogenin (Figure 9). The mutation (p.Arg33Trp) is not reported in the ALSoD database and the only information about this variant was that according to the ExAc database it has been found in one individual (out of 121.410 individuals).

**Amino acid numbering according to Human Genome Variation Society guidelines (http://www.hgvs.org/).*
[#]Amino acid numbering according to the ALS Online Database (http://alsod.iop.kcl.ac.uk/), which has been used in the previous published reports on ANG mutations.

3.4 C9orf72 gene

Analysis of the *C9orf72* gene identified GGGGCC RE in one (Figure 10) out of 66 ALS patients. The average repeat number based on fragment-length analysis was 5 (range 2–17 repeats) in the remaining 65 patients, none of whom carried repeat expansion. The patient with the repeat expansion also carried the rs3849942 risk allele, which was previously described as a part of the Finnish "risk" haplotype (Laaksovirta *et al.,* 2010).

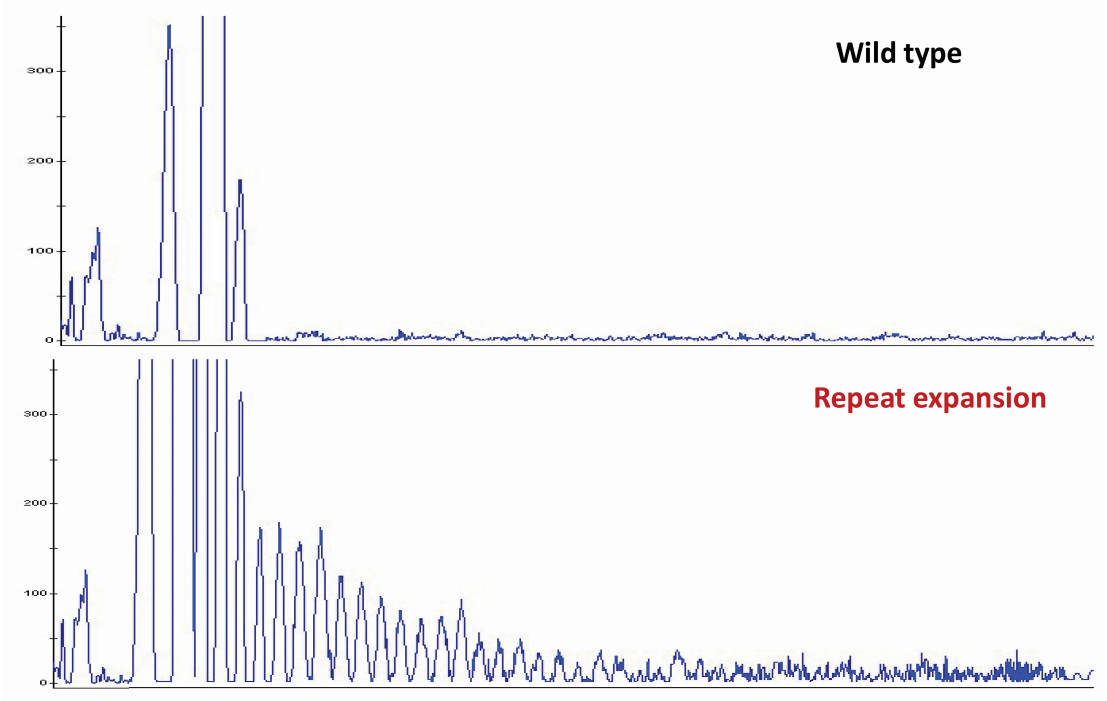

Figure 10. PCR products of repeat-primed PCR reactions separated on 310 ABI Prism Genetic Analyzer and visualized by GeneMapper software. The difference between the wild type and the shutter amplification of the sample carrying repeat expansion is shown.

3.4.1 Clinical symptoms of the patient with the *C9orf72* RE

This latter patient presented to the neurological unit at age 68 with gradually increasing foot drop on the left side due to peroneal weakness. Neurological examination revealed mild dysarthria, dysphagia, atrophy, fasciculations and fibrillations in the muscles of the tongue. She reported 15 kg weight loss in the previous six months and generalized fatigue. She had moderate paraparesis (4/5) with diminished deep tendon reflexes and muscle atrophy with fasciculations all over the muscles of the lower extremities together with complete left peroneal palsy. The signs of lower motor neuron lesion (muscle atrophy, fasciculations) were accompanied with brisk deep tendon reflexes in the upper extremities as the only sign of upper motor neuron lesion. The ALSFRS-R was 44/48. She was treated with duloxetine (60 mg/day) and alprazolam (0.5 mg/day) for depression as well as perindopril (3.34 mg/day) for hypertension. Three years before admission, an adenoma of one of the parathyroid glands causing hyperparathyroidism was surgically removed. On admission, the levels of serum calcium, phosphate and parathyroid hormone were within normal ranges. The level of creatine phosphokinase was minimally raised (193, normal<170 U/l). There were no other abnormal laboratory values observed. Several palpable clumps in the subcutaneous tissue were observed all over the body. Histological examination of one such excised clump proved to be lipoma. Because of the presence of multiple lipomas, lipomatosis was diagnosed; however, mutational analysis of the *HMGIC* gene (Schoenmakers *et al.,* 1995) has not been performed. Other entities which can cause multiple lipomas have also been considered but not proved. A lumbar IV-V disc protrusion was present, but it did not explain the neurological signs and symptoms. The motor and sensory nerve conduction velocities were all normal. EMG identified fasciculations, fibrillations, positive sharp waves, polyphasic high amplitude (maximum 5 mV) motor units without full interference pattern in full effort in several examined muscles of all four extremities. There were no sensory abnormalities and no involvement of the external eye muscles and sphincter functions. Malignant tumor was excluded by several examinations. MRI of the central nervous system did not confirm the presence of other diseases. There were no paraproteins or elevated anti-ganglioside IgG or IgM found in the serum. Parathyroid dysfunction could not be proved; however, it is noted

that debate whether hyperparathyroidism can mimic ALS has been ongoing for the last 50 years (Jackson *et al.,* 1998).

The next generation sequencing approach did not detect mutations in the investigated coding regions of the *C9orf72* gene.

3.5 SETX gene

A novel heterozygous missense mutation (c.791A>G, p.Asn264Ser) of the *SETX* gene (Figure 11a) was identified in a Hungarian female patient. The novel c.791A>G, p.Asn264Ser *SETX* mutation identified in this study was confirmed with direct sequencing (Figure 11b). SIFT, Polyphen and Mutation Taster analyses determined this variant as a damaging, disease-causing mutation. The mutation was absent in the patient's clinically healthy 45-year-old son and in the age- and gender-matched healthy control individuals. The heterozygous p.Asn264Ser mutation affects the N-terminal region of the SETX protein (Figure 12a) and is located in a region which is conserved among mammals (Figure 12b). According to the results of the Human Splicing Finder online prediction tool, the c.791A>G mutation could result in activation of an exonic cryptic donor site, creation of an exonic silencer site (ESS) or alteration of an exonic enhancer site (ESE).

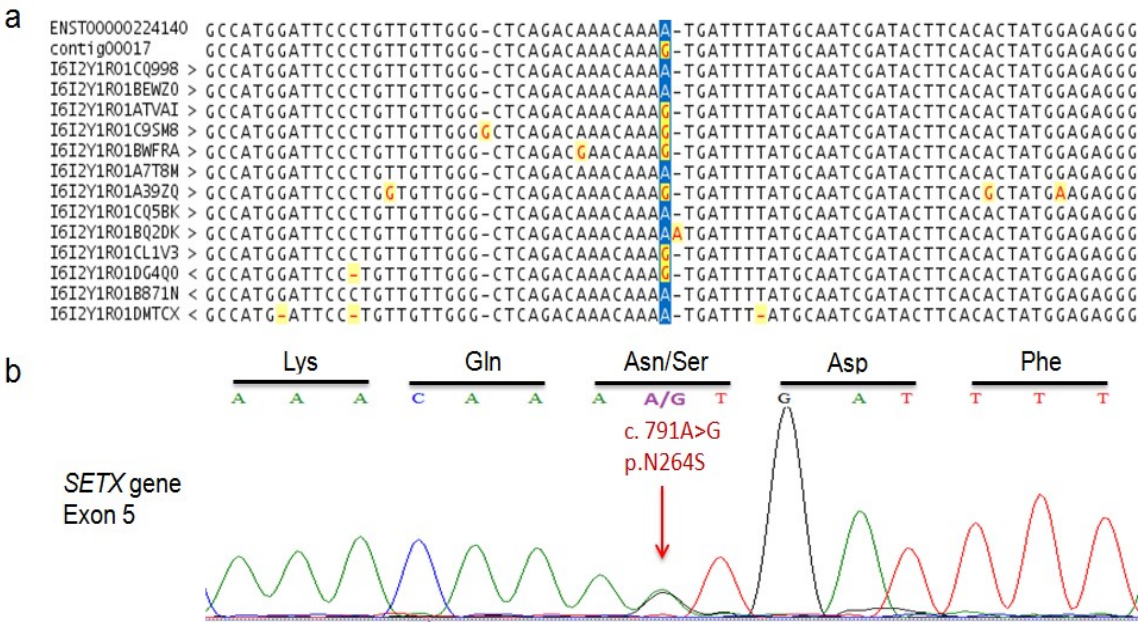

Figure 11. A novel mutation in the *SETX* gene. (a) Next generation sequencing identified the novel heterozygous missense *SETX* mutation. (b) The presence of the mutation was confirmed using direct sequencing. (Tripolszki *et al.,* 2017b).

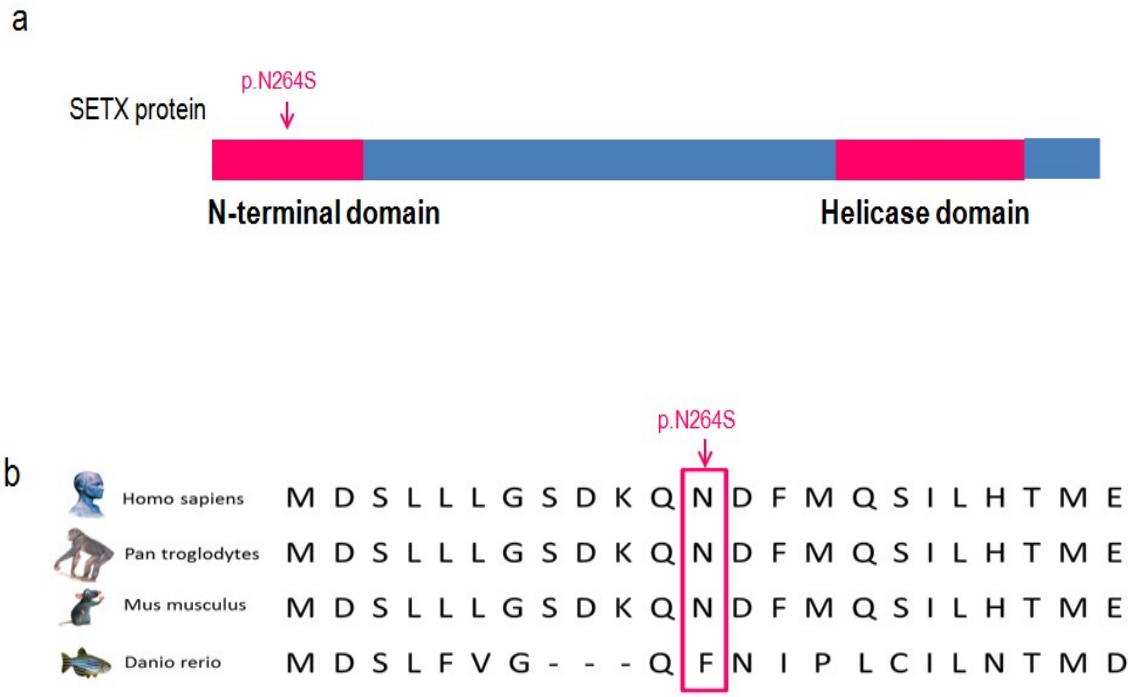

Figure 12. (a) The mutation is located within the N-terminal domain of the encoded protein. (b) The region of the mutation is highly conserved in mammals (Tripolszki *et al.,* 2017b)

3.5.1 Clinical characterictics of the patient with the novel p.Asn264Ser *SETX* mutation

This patient presented with a lower motor neuron disease phenotype, which started with an unsteady gait at the age of 65. Electroneurography revealed motor axonal loss in the right tibial and peroneal nerves with spared sensory innervation. The electromyography indicated subacute and chronic signs of denervation and reinnervation. One year later, follow up examinations indicated that denervation had spread to several muscles of both lower extremities, and the clinical and electrophysiological signs of denervation appeared in the muscles of the upper extremities. At the time blood was obtained, the interosseous muscles were wasted, dysarthria and dysphagia had developed with diminished gag reflex and lack of deep tendon reflexes were observed all over the body. Muscle strength was generally 3/5

in the upper extremities and 2/5 in the lower extremities. Sensory deficit was not observed and the extraocular muscles were not involved in the disease process. The external sphincter muscles were also spared. During the 3.5 year course of the disease, corticospinal or corticobulbar signs were not detected with physical examination and with magnetic resonance images of the central nervous system. Progressive weakness was observed. Other diseases causing similar symptoms were all excluded (thyroid and parathyroid diseases, paraneoplastic syndromes, gammopathies, multifocal motor neuropathy with ganglioside antibodies). Ataxia and oculomotor apraxia were not noted in the follow up examinations. The cerebrospinal fluid proved to be normal. The patient passed away at the age of 69 and the cause of death was determined to be respiratory failure due to weakness of the respiratory muscles. No other members of the family were affected. The autopsy revealed a decreased number of motor neurons in the spinal cord and diminished number of axons in the ventral roots. Hyalineinclusions were noted in the remaining spinal motor neurons. Histological examination of the striated muscles (proximal and distal in the extremities, in the diaphragm and in the tongue) revealed signs of neurogenic atrophy. There were no signs of degeneration in the corticospinal tracts. The motor cortex remained intact, and other parts of the brain (temporal and parietal lobes, hippocampus, cerebellum, thalamus, pons and the medulla) seemed to be normal except some cell loss in the nuclei of the hypoglossal and facial nerve. Together with the clinical symptoms, the disease course and the autopsy findings confirmed the diagnosis of an atypical ALS form with lower motor neuron involvement.

4 DISCUSSION

In this study, I analyzed *SOD1, TARDBP, ANG, C9orf72, SETX* and *FUS* genes in a cohort of 66 Hungarian ALS patients, including one single case with a reported familial history for the disease (Table 2).

Table 2. Clinical data of Hungarian ALS patients carrying pathogenic variants in *SOD1, ANG, TARDBP, C9orf72* and *SETX* genes

Gene	Age of onset/gender	Mutation	Disease signs	Disease duration, years	Other diseases	ALS family history	Reference
SOD1	63/Female	p.Asp90Ala	LMN, UMN, B	12	Lumbar disc protrusions	No	Andersen *et al*. 1995
SOD1	67/Female	p.Lys91ArgfsTer8	LMN, UMN, B, PB	1	Breast cancer, hypertension, hypercholesterolemia, spondylosis and lumbar disc protrusions	No	*This study*
SOD1	29/Female	p.Leu144Phe	LMN, UMN	4	None reported	Yes	Deng *et al*. 1993
SOD1	46/Female	p.Leu144Phe	LMN, UMN, B	3	Lumbar disc protrusions	No	Deng *et al*. 1993
SOD1, ANG	62/Female	p.Val14Met, p.Met-24Ile	LMN, UMN	1	Lumbar disc protrusions, atherosclerosis	No	Deng *et al*. 1995
ANG	54/Male	p.Arg33Trp	UMN, LMN	3	Spondylarthrosis and disc protrusions	No	*This study*
TARDBP	62/Male	p.Met311Val	UMN, LMN, B, PB	3	Lumbar spondylosis	No	Lemmens *et al*. 2009
C9orf72	65/Female	Repeat Expansion	LMN, UMN	0,5	Hyperparathyroidism, multiple lipomas, hypertension, osteoporosis	No	Renton *et al*., 2011; DeJesus *et al*.,2011
SETX	66/Female	p.Asn264Ser	LMN, B	3	Hypercholesterolaemia, spondylarthrosis, discopathia lumbalis, small white matter infarct in the cerebrum	No	*This study*

4.1 Mutations in the SOD1 gene

I identified a novel and likely disease-causing heterozygous frameshift mutation (p.Lys91ArgfsTer8) in the *SOD1* gene (Figure 2). Three other heterozygous recurrent *SOD1* missense mutations (p.Val14Met, p.Asp90Ala and p.Leu144Phe) were also detected (Figure 13). *SOD1* mutations were detected in 7.5% (5/66) of this cohort, in line with literature data reporting that *SOD1* mutations account for approximately 0–7% patients with sporadic disease (Andersen, 2006).

The identified novel p.Lys91ArgfsTer8 *SOD1* mutation is associated with a typical ALS phenotype characterized by lower and upper motor neuron signs; in addition, the patient had also suffered from breast cancer, hypertension, hypercholesterolemia, spondylosis and lumbar disc protrusions before ALS developed. The disease had a late onset at the age of 67 years and progressive course leading to the death of the patient within one year after the onset of the signs and symptoms of the disease. The two-base-pair deletion of the p.Lys91ArgfsTer8 *SOD1* mutation leads to a frameshift with the formation of a premature stop codon after the insertion of eight novel amino acids, causing a severe truncation of the protein (Figure 2). This truncation abolishes the integrity of the intrachain C57–C146 disulfide bridge. Although most mutations in the *SOD1* gene are missense ones, a few deletions and insertions have been described previously causing truncations of different sizes and in most cases affecting also the C57-C146 disulfide bridge (http://alsod.iop.kcl.ac.uk/; Abel *et al.,* 2012). The most comprehensively studied truncation mutant is the Gly127insTGGG (G127X), carriers of which developed signs of motor neuron degeneration with a rapid disease course (Jonsson *et al.,* 2004). These observations correlate well with the clinical manifestations and course of ALS in the investigated Hungarian patient with the truncating p.Lys91ArgfsTer8 *SOD1* mutation. SOD1 protein with truncating mutations exhibits structural instability causing misfolding in the mutated enzyme (Jonsson *et al.,* 2004) which can consequently aggregate in motor neurons and lead to the development of ALS (Forsberg *et al.,* 2011).

With the exception of the p.Asp90Ala *SOD1* mutation, the detected recurrent mutations are all associated with typical ALS phenotypes, characterized by lower and upper motor neuron signs, late onset and progressive disease course. In the case of

the patient with the p.Asp90Ala heterozygous mutation, lower limb involvement and relatively long duration of the disease course was detected similarly to cases reported previously (Robberecht *et al.,* 1996; Andersen *et al.,* 2006). The p.Asp90Ala mutation is the most common *SOD1* mutation in Europe (Andersen 2001), and it can be inherited in either a dominant or recessive manner (Robberecht *et al.,* 1996).

The p.Leu144Phe missense mutation, which is the most prevalent mutation in the Balkan region (ALSoD), was detected in a familial and in a sporadic case. The clinical symptoms and the course of the disease were similar in these two patients. To note, the patient with a positive family history developed ALS symptoms at a relatively early age, whereas onset for the other patient was late. Although a significant number of patients with this mutation have lower limb onset (Corcia *et al.,* 2011), both of the patients reported in this study developed upper and lower motor neuron signs. The rare p.Val14Met mutation was detected in an apparently sporadic case in a female patient with upper and lower motor neuron signs. None of the *SOD1* mutations detected in these ALS patients was identified in 110 healthy Hungarian controls.

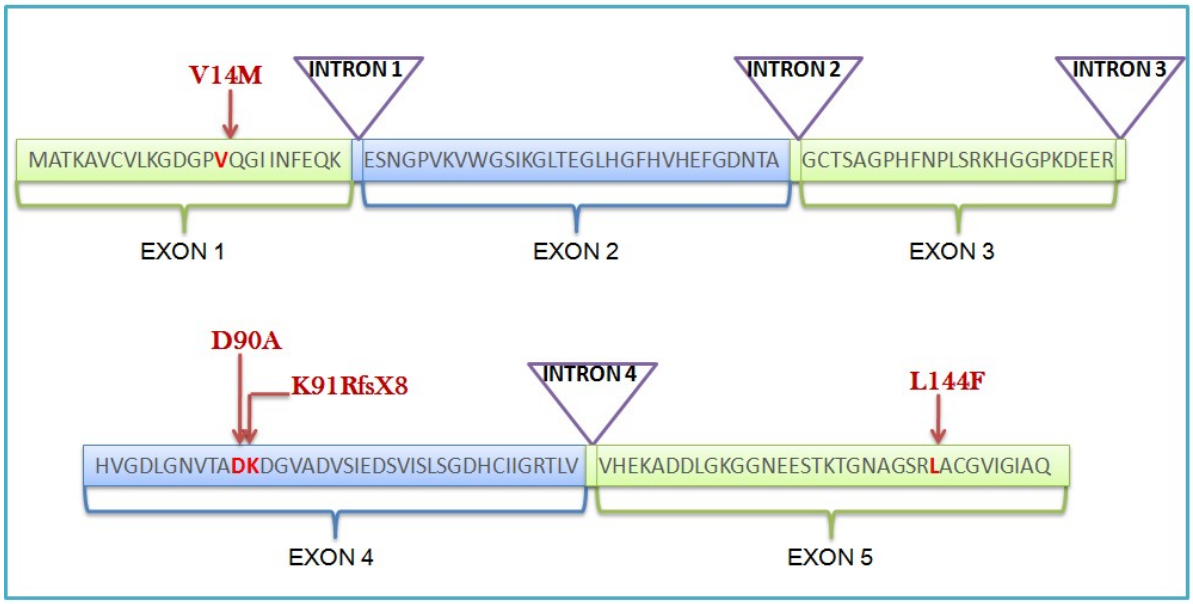

Figure 13. Diagram of SOD1 mutations detected in Hungarian patients. The amino acid sequence is shown, with the location of introns.

4.2 Mutations in the TARDBP gene

A known disease-causing mutation (p.Met311Val) has been detected in the last exon of the *TARDBP* gene. With the exeption of one mutation (p.Asp169Gly in exon 4), all the identified mutations are localized across the entire C-terminal sequence of TDP-43 with a tendency of clustering in specific positions. The p.Met311Val mutation is located in the C-terminal part of the protein, the glycine-rich region that may mediate interactions with other proteins, including heterogenous ribonucleoproteins (Buratti *et al.,* 2005). This affected methionine is evolutionary well conserved from *Homo sapiens* to *Xenopus tropicalis*.

4.3 Mutations in the ANG gene

Two heterozygous mutations were detected in the *ANG* gene (c.3G>A, p.Met-24Ile, p.Met1Ile*; c.169C>T, p.Arg33Trp, p.Arg57Trp*).

The p.Met-24Ile mutation, located in the signal peptide region of the protein, was first described by Conforti *et al.* (2008). This mutation was detected in the same Hungarian patient who carried the *SOD1* p.Val14Met mutation. The biological function of the signal peptide is not entirely understood, hence it is difficult to predict the effect of the p.Met-24Ile mutation. Nonetheless, the Met-24I mutation affects the start codon (ATG) of the gene, which may influence the correct translation of the protein. The co-occurrence of different variants in ALS associated genes is also detected in other ALS cohort screenings (Kenna *et al.,* 2013; Cady *et al.,* 2015; Cirulli *et al.,* 2015; Krüger *et al.,* 2016). As it was previously mentioned in the text, this patient with the *SOD1* p.Val14Met mutation showed a very progressive disease course that could be explained by the fact that she carried disease-causing mutations in two ALS associated genes. Cady *et al.* (2015) described that 3,8% of the investigated patients carried pathogenic mutations in more than one gene and those patients had disease onset 10 years earlier, which supports a model of ALS where the additive or synergistic effects of multiple defective genes influence disease phenotype. On the contrary, Cirulli *et al.* (2015) did not report an influence of the number of variants on the disease phenotype. Due to the small sample size, I could

not make statistically significant observations on a possible correlation between the number of pathogenic variants and disease phenotype.

The p.Arg33Trp (p.Arg57Trp*) affects the nuclear translocation signal of the angiogenin. Amino acids ^{29}IMRRRGL^{35} represent the nuclear localization signal and helps in nuclear translocation of angiogenin. The arginine residues ^{31}RRR^{33} are critical in governing the nuclear translocation activity of angiogenin. The R^{33} is essential to nuclear translocation and residues ^{31}RR^{32} modulate this process. Moroinau et al. (1994) described that R33A mutant is not translocated to the nucleus and lacks of angiogenic activity (Moroianu and Riordan, 1994). P.Arg33Trp is not reported in the ALSoD database and to my knowledge, there is no report about tis variant in the literature. The only previous information about this variant is that it has been found in one out of 121.410 individuals in the Exome Aggregation Consortium.

4.4 Repeat expansion of the C9orf72 gene

C9orf72 repeat expansion was detected only in one patient of 66. The patient carrying the RE variant also carried the previously described rs3849942 risk allele (Laaksovirta *et al.,* 2010; DeJesus-Hernandez *et al.,* 2011) in heterozygous form. According to earlier findings, the "A" allele of SNP rs3849942 was significantly associated to the expanded *C9orf72* allele (DeJesus-Hernandez *et al.,* 2011). The presence of the risk allele suggests that the common founder is also shared by the investigated Hungarian patient with the *C9orf72* repeat expansion. I screened 110 controls of Hungarian origin to establish the allele frequencies in this genomic position. The allele frequency of the minor allele in the Hungarian population (MAF (A) = 0.18) correlates well with the data from 1000 Genomes Database (www.1000genomes.org/). I observed that the average repeat number was 5 (range: 2–17 repeats) in the remaining 65 patients who did not carry repeat expansions.

In this study, I identified the hexanucleotide RE of the *C9orf72* gene only in one sporadic patient. These results suggest that the frequency of RE observed in the Hungarian ALS patients (1,5%) is significantly lower than in Western European populations (Majounie *et al.,* 2012; Ratti *et al.,* 2012; Fogh *et al.,* 2014), further demonstrating that the frequency of genetic factors for ALS varies among different geographic regions.

I did not identify any mutations in the coding regions of the *C9orf72* gene. These results correlate well with the data reported in the literature. Recently, Liu *et al.* (2016) reported a splice site mutation in one patient with ALS, which is the second disease-causing variant reported in the *C9orf72* gene. In addition, there are different studies that failed to detect any disease-causing mutations in the coding regions of the *C9orf72* gene (Koppers *et al.,* 2013; Kenna *et al*., 2013).

4.5 Mutations of the SETX gene

A novel, disease-causing heterozygous missense mutation has been identified (p.N264S) in the *SETX* gene. To date, only 17 pathogenic *SETX* mutations (Table 4) have been implicated in the development of the autosomal dominant juvenile form, ALS4. The typical phenotype, present in the majority of the ALS4 patients, presents as a combination of lower and upper motor neuron impairments (Avemaria *et al.,* 2011; Hirano *et al.,* 2011; Saracchi *et al.,* 2014).

The identified novel p.Asn264Ser *SETX* mutation is associated with an unusual ALS phenotype characterized by lower motor neuron impairment only. These results further support the accepted view that ALS and lower MNDs are not different disease entities, but that, instead, they are clinical variants of the same disease spectrum. The phenotypic variation can be explained by yet unidentified genetic modifiers, environmental and/or life style factors. The unusual phenotype observed in the reported Hungarian patient is not unique in the literature of pathogenic heterozygous *SETX* mutations. A similar atypical phenotype has been described in a Chinese patient carrying the p.T1118I heterozygous missense mutation of the *SETX* gene (Zhao *et al.,* 2009). Although the p.N264S mutation is located within the N-terminal region of the protein and the p.T1118I mutation is located in a region of yet unknown function, both of these mutations result in the development of the same unusual ALS4 phenotype. Further studies are needed to elucidate the underlying mechanism responsible for the observed unusual phenotypes associated with the heterozygous p.N264S and p.T1118I missense mutations of the *SETX* gene (Table 3).

The 17 known heterozygous missense mutations of the SETX gene are located in exons 1, 5, 8, 11, 12, 13, 17 and 24 (Table 4).

Table 3. Clinical features of the atypical ALS4 phenotype.

SETX Mutation	Amino Acid Substitution	Ethnic Origin	Age of Onset	Lower Motor Neuron Sign	Upper Motor Neuron Sign	Bulbar Sign	Sensory Impairment	Reference
c.791A>G	p.N264S	Hungarian	65	+	-	+	-	*This study*
c.3353C>T	p.T1118I	Chinese (Han)	42	+	-	+	-	*Zhao et al., 2009*

Four of the 17 mutations are located within the helicase domain of the SETX protein (Chen *et al.,* 2004; Hirano *et al.,* 2011; Kenna *et al.,* 2013; Saracchi *et al.,* 2014). In AOA2, homozygous or compound heterozygous missense mutations located within the helicase domain are generally associated with less severe phenotypes than the mutations affecting other regions of the protein (Moreira *et al.,* 2004; Anheim *et al.,* 2009). A similar genotype–phenotype association has not been observed for the heterozygous missense mutations associated with ALS4. However, in ALS4, mutations impairing the region of the helicase domain are associated with cortical and spinal motor neuron impairment, whereas others located outside of this region can lead to the development of either spinal or bulbar motor neuron involvement (Table 4).

It has also been demonstrated that, in AOA2, missense mutations affecting the N-terminal or the helicase domain of SETX cause phenotypes similar to those caused by deletions, nonsense or frameshift mutations (Chen *et al.,* 2006). However, only missense *SETX* mutations have been implicated in the development of ALS4 to date (Avemaria *et al.,* 2011; Hirano *et al.,* 2011; Kenna *et al.,* 2013; Saracchi *et al.,* 2014); deletions, nonsense, frameshift or other types of mutations have not yet been detected for this disease (Table 4).

Table 4. Clinical manifestations of *SETX* mutations in ALS4 (Tripolszki *et al.,* 2017b)

SETX Mutation	Amino Acid Substitution	Affected Exon	Impaired Region of the Protein	Age of Onset	Affected sites	Reference
c.8C>T	p.T3I	1	N-terminal Domain	8	Spinal	*Chen et al., 2004*
c.791A>G	**p.N264S**	**5**	**N-terminal Domain**	**65**	**Both**	***This study***
c.814C>G	p.H272D	5	N-terminal Domain	66	Bulbar	*Kenna et al., 2013*
c.1166T>C	p.L389S	8	N-terminal Domain	17	Spinal	*Chen et al., 2004,*
				24	Spinal	*Avemaria et al., 2011*
c.2755G>C	p.V919L	8	Unknown Function	68	Both	*Kenna et al., 2013*
rc.2842C>A	p.P948T	8	Unknown Function	58	Bulbar	*Kenna et al., 2013*
c.2975A>G	p.K992R	8	Unknown Function	78	Bulbar	*Kenna et al., 2013*
c.3353C>T	p.T1118I	8	Unknown Function	42	Both	*Zhao et al., 2009*
c.4660T>G	p.C1554G	8	Unknown Function	6	Spinal	*Hirano et al., 2011*
c.5587A>G	p.T1863A	11	Unknown Function	61	Other	*Kenna et al., 2013*
c.5842A>G	p.M1948V	12	Helicase Domain	47	Spinal	*Kenna et al., 2013*
c.6052A>G	p.K2018E	13	Helicase Domain	54	Spinal	*Saracchi et al., 2014*
c.6085C>G	p.K2029E	13	Helicase Domain	35	Spinal	*Hirano et al., 2011*
c.6407G>A	p.R2136H	17	Helicase Domain	6	Spinal	*Chen et al., 2004*
c.7640T>C	p.I2547T	24	Unknown Function	30	Spinal	*Hirano et al., 2011*
c.7645G>A	p.V2549I	24	Unknown Function	64	Spinal	*Kenna et al., 2013*
c.7682C>T	p.S2561L	24	Unknown Function	84	Spinal	*Kenna et al., 2013*

This study further widens the geographic range for the origin of disease-causing heterozygous missense mutations of the *SETX* gene, which have already been implicated in ALS in patients from different countries (Avemaria *et al.,* 2011; Hirano *et al.,* 2011; Kenna *et al.,* 2013; Saracchi *et al.,* 2014). To my knowledge, my study is the first demonstrating a novel *SETX* mutation in the Hungarian ALS population (Table 4).

4.6 Mutations in the FUS gene

No mutations in *FUS*, were detected in the Hungarian patients. *FUS* mutations have been detected in two regions of the gene that consequently affect the encoded protein: approximately one-third are located within exon 3 and 6 and impair the glutamine-glycine-serine-tyrosine-rich and the arginine-glycine-glycine-rich domains of the protein. Two-thirds of the mutations are located within the region including exons 12 to 15, which encodes a zinc-finger and arginine-glycine-glycine-rich domains (Shang and Huang, 2016). Based on my results, I hypothesize that FUS mutations in the Hungarian ALS population might be very rare. This hypothesis is supported by the literature, as FUS mutations have been reported to contribute to the development of the disease in approximately 0.5% of the patients (Kenna *et al.,* 2013).

Based on my results and the results of previous studies, I also emphasize that the genetic screening of the Mendelian ALS-associated genes might not elucidate the causative genetic variant in the majority of ALS cases (Kenna *et al.,* 2013; Cirulli *et al.,* 2015). The genetic heterogeneity of ALS is extremely complex: rare mutations of Mendelian genes and common variants of non-Mendelian genes can also contribute to the development of the disease (Abel *et al.,* 2012). This comprehensive study adds novel data to the genetic and phenotypic diversity of ALS and indicates that complex approaches, high-throughput methods and large-scale studies are needed to understand the genetic heterogeneity of this disease.

5 SUMMARY

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease characterized by the degeneration of the motor neurons. Approximately 90% of ALS cases are sporadic the remaining 10% are familial. Due to the complex genetic background of the disease, the underlying disease-causing variant is rarely established for individual cases. More than 20 major ALS genes are implicated in the Mendelian ALS forms and about 100 genes are associated with increased susceptibility for the development of the disease.

The aim of this study was to investigate Mendelian-disease causing genes in a group of Hungarian ALS patients (n=66). Considering the onset, the symptoms and the course of the disease, six genes (superoxide dismutase, *SOD1*; TAR DNA binding protein, *TARDBP*; angiogenin, *ANG*; senataxin, *SETX*; fused in sarcoma, *FUS* and chromosome 9 open reading frame 72, *C9orf72*) were selected for mutation screening.

The patients (n=66) participating in this study were recruited from the Department of Neurology, University of Szeged, Hungary. All patients and age- and sex-matched healthy controls (n=110) were of Hungarian ancestry. Genomic DNA was isolated from blood. *SOD1*, *TARDBP* and *ANG* genes were analysed by Sanger sequencing. The GGGGCC hexanucleotide repeat in *C9orf72* was analyzed by a two-step protocol, including a first step of polymerase chain reaction amplification using genotyping primers. The fragment length analysis was performed using GeneScan. Only samples presenting with a single peak/amplification product were further analyzed in the second step by repeat-primed polymerase chain reaction. Targeted high-throughput sequencing was used for the analysis of the coding regions of the *SETX, FUS* and *C9orf72* genes.

Direct sequencing revealed a novel heterozygous frameshift mutation (c.275_276delAA, p.Lys92ArgfsX8) and three recurrent missense mutations (p.Val14Met; p.Asp90Ala and p.Leu144Phe in two patients) in the *SOD1* gene in five patients. The novel heterozygous *SOD1* mutation (p.Lys91ArgfsTer8) is located in the fourth exon of the gene and led to a frameshift with the insertion of 8 novel amino acids and the formation of premature stop codon at the new amino acid

position 99. A known missense mutation (p.Met311Val) was detected in the *TARDBP* gene, which is located in the C-terminal part of the protein, the glycine-rich region that may mediate interactions with other proteins. Two heterozygous mutations were identified in the *ANG* gene (c.3G>A, p.Met-24Ile; c.169C>T, p.Arg33Trp). The p.Met-24Ile mutation was detected in the same Hungarian patient who carried the SOD1 p.Val14Met mutation, and is located in the signal peptide region of the protein. The second *ANG* mutation (c.169C>T, p.Arg33Trp) affects the nuclear translocation signal of the angiogenin. C9orf72 repeat expansion was detected in one sporadic female patient. She also carried the previously described rs3849942 risk allele in heterozygous form. Targeted high-troughput sequencing revealed a novel heterozygous missense mutation (c.791A>G, p.Asn264Ser) in the *SETX* gene. The mutation affects the N-terminal region of the SETX protein and is located in a region which is evolutionary conserved among mammals.

In conclusion, I performed the first genetic analysis of *SOD1, TARDBP, ANG, C9orf72, SETX* and *FUS* genes in a cohort of Hungarian ALS patients. My study further widens the geographic range for the origin of disease-causing heterozygous missense and frameshift mutations of the *SOD1, TARDBP, ANG, C9orf72* and *SETX* genes, which have already been implicated in ALS patients from different countries of origin.

6 ACKNOWLEDGEMENTS

I would like to thank to Dr. Nikoletta Nagy for her great supervising activity.

I would like to thank to Prof. Dr. Márta Széll for the opportunity to perform the genetic investigations of this study in the molecular laboratory of the Department of Medical Genetics, University of Szeged.

I am grateful to Dr. József I Engelhardt, Dr. Péter Klivényi, Nóra Török and Dr. László Vécsei for the neurological examinations and for enrolling the pateints into the study.

Special thanks to Dr. Antonia Ratti, Dr. Cinzia Tiloca and Dr. Vincenzo Silani (Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan, Italy) for providing the *C9orf72* repeat expansion positive samples.

I am extremely grateful to Dr. Bernadett Csányi for her help in the *C9orf72* repeat expansion analysis. I would like to thank to Dr. Éva Kereszty for the opportunity to perform the *C9orf72* repeat expansion analysis in the laboratory of the Department of Forensic Medicine, University of Szeged.

I would like to express my gratitude to Dr. Vincent Procaccio and Dr. David Goudenege (University of Angers, France) for their guidance and help in the bioinformatic analysis of the next generation sequencing data.

Special thanks to all my colleagues for their kind help at the Department of Medical Genetics, University of Szeged.

7 ELECTRONIC DATABASE INFORMATION

- **Ensemble Genome Browser**

  http://www.ensembl.org/Homo_sapiens/Info/Index

- **Online Mendelian Inheritance in Man** https://www.omim.org
- **UCSC Genome Browser, Primer3** https://genome.ucsc.edu
- **ALS Online Genetics Database** http://alsod.iop.kcl.ac.uk/
- **1000 Genomes Database** (www.1000genomes.org/)
- **dbSNP** http://www.ncbi.nlm.nih.gov/project/SNP
- **Exome Aggregation Consortium (ExAC) database**

  http://exac.broadinstitute.org

- **SIFT** http://sift.bii.a-star.edu.sg/
- **Polyphen-2** http://genetics.bwh.harvard.edu/pph2
- **Mutation Taster** http://mutationtaster.org
- **Swiss-Model protein structure homology-modeling server**

  http://swissmodel.expasy.org/

- **Human Splicing Finder** http://www.umd.be/HSF3/

LIST OF ABBREVIATIONS

ALS	Amyotrophic lateral sclerosis
ALSFRS-R	ALS Functional Rating Scale R
ANG	Angiogenin
AOA2	Ataxia with oculomotor apraxia type 2
C9orf72	Chromosome 9 open reading frame 72
EMG	Electrimyography
ENG	Electroneurography
FUS	Fused in sarcoma
LMN	Lower motor neuron
MND	Motor neuron disease
MRI	Magnetic Resonance Imaging
NGS	Next-generation sequencing
PCR	Polimerase chain reaction
PLS	Primary lateral sclerosis
PMA	Progressive muscular atrophy
RE	Repeat expansion
SETX	Senataxin
SNP	Single nucleotide polymorphism
SOD1	Superoxide dismutase 1
TARDBP	Transactive response DNA binding protein
UMN	Upper motor neuron

8 REFERENCES

1. **Abel O, Powell JF, Andersen PM, Al-Chalabi A.** ALSoD: a user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. *Hum Mutat*. 2012;33:1345-1351.
2. **Aggarwal A and Shashiraj**. Juvenile amyotrophic lateral sclerosis. *Indian J Pediatr*. 2006;73:225-226.
3. **Alsultan AA, Waller R, Heath PR, Kirby J.** The genetics of amyotrophic lateral sclerosis: current insights. *Degen Neurol Neuromusc Dis.*2016;6:49-64.
4. **Andersen PM, Nilsson P, Ala-Hurula V, *et al.*** Amyotrophic lateral sclerosis associated with homozygosity for an Asp90Ala mutation in CuZn-superoxide dismutase. *Nat Genet.* 1995;10:61-6.
5. **Andersen PM, Spitsyn VA, Makarov SV, *et al.*** The geographical and ethnic distribution of the D90A CuZn-SOD mutation in the Russian Federation. *Amyotroph Lateral Scler Other Motor Neuron Disord.* 2001;2:63-9.
6. **Andersen PM.** Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr Neurol Neurosci Rep.* 2006;6:37-46.
7. **Andersen PM, Borasio GD, Dengler R, *et al.*** Good practice in the management of amyotrophic lateral sclerosis: clinical guidelines. An evidence-based review with good practice points. EALSC Working Group. *Amyotroph Lateral Scler.* 2007;8:195e213.
8. **Akimoto C, Volk AE, van Blitterswijk M, *et al.*** A blinded international study on the reliability of genetic testing for GGGGCC-repeat expansions in *C9orf72* reveals marked differences in results among 14 laboratories*.* *J Med Genet*. 2014;51:419-24.
9. **Al-Chalabi A, Andersen PM, Chioza B, *et al.*** Recessive amyotrophic lateral sclerosis families with the D90A SOD1 mutation share a common founder: evidence for a linked protective factor. *Hum Mol Genet.* 1998;7:2045–2050.
10. **AvemariaF, Lunetta C, Tarlarini C, *et al.*** Mutation in the senataxin gene found in a patient affected by familial ALS with juvenile onset and slow progression. *Amyotroph Lateral Scler.* 2011;12:228-230.
11. **Anheim M, Monga B, Fleury M, *et al.*** Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. *Brain*. 2009;132:2688-2698.
12. **Al-Chalabi A, Fang F, Hanby MF, *et al.*** An estimate of amyotrophic lateral sclerosis heritability using twin data. *J Neurol Neurosurg Psychiatry*. 2010;81:1324-1326.
13. **Arnesano F, Banci L, Bertini I, *et al.*** The Unusually Stable Quaternary Structure of Human Cu, Zn-Superoxide Dismutase 1 Is Controlled by Both Metal Occupancy and Disulfide Status. *J Biol Chem*. 2004;279:47998-8003.
14. **Bensimon G, Lacomblez L, Meininger V.** A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. *N Engl J Med*. 1994;330:585-591.

15. **Buratti E, Brindisi A, Giombi M, *et al.*** TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J Biol Chem*. 2005;280:37572-37584.
16. **Buratti E and Baralle FE.** Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front Biosci*. 2008;13:867-878.
17. **Byrne S, Elamin M, Bede P, *et al.*** Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a *C9orf72* repeat expansion: a population-based cohort study. *Lancet Neurol*. 2012;11:232–240.
18. **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. *Ann Neurol*. 2015;77:100–113.
19. **Cedarbaum JM, Stambler N, Malta E, *et al.*** The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci*. 1999;169:13-21.
20. **Charcot JM.** Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arch Physiol Neurol Pathol.* 1869*;*2:744-754.
21. **Chen YZ, Bennett CL, Huynh HM, *et al.*** DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet*. 2004;74: 1128-1135.
22. **Chen YZ, Hashemi SH, Anderson SK, *et al.*** Senataxin, the yeast Sen1p orthologue: characterization of aunique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. *Neurobiol Dis.* 2006;23: 97-108.
23. **Chen S, Sayana P, Zhang X, Le W.** Genetics of amyotrophic lateral sclerosis: an update. *Mol Neurodegener.* 2013;8:28.
24. **Chio A, Calvo C, Moglia C, *et al.*** Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study. *J Neurol Neurosurg Psychiatry*. 2011;82:740-746.
25. **Cirulli ET, Lasseigne BN, Petrovski S, *et al.*** Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*. 2015;347:1436–1441.
26. **Conforti FL, Sprovieri T, Mazzei R, *et al.*** A novel Angiogenin gene mutation in a sporadic patient with amyotrophic lateral sclerosis from southern Italy. *Neuromuscul Disord*. 2008;18:68-70.
27. **Corcia P, Petiot P, Stevic Z, *et al.*** Respiratory onset in an ALS family with L144F *SOD1* mutation. *J Neurol Neurosurg Psychiatry*. 2011;82:747e749.
28. **de Carvalho M, Swash M.** Awaji diagnostic algorithm increases sensitivity of El Escorial criteria for ALS diagnosis. *Amyotroph Lateral Scler*. 2009;10:53-7.
29. **DeJesus-Hernandez M, Mackenzie IR, Boeve BF, *et al.*** Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron*. 2011;72:245-256.

30. **Deng HX, Tainer JA, Mitsumoto H, et al.** Two novel SOD1 mutations in patients with familial amyotrophic lateral sclerosis. *Hum Mol Genet.* 1995;4:1113-1116.
31. **Deng HX, Hentati A, Tainer JA, et al.** Amyotrophic Lateral Sclerosis and Structural Defects in Cu,Zn Superoxide Dismutase. *Science.* 1993;20:1047-1051.
32. **Desmet F O, Hamroun D, Lalande M, *et al.*** Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37:67.
33. **Finsterer J, Burgunder JM.** Recent progress in the genetics of motor neuron disease. *Eur J Med Genet.* 2014;57:103-12.
34. **Fogh I, Ratti A, Gellera C*, et al.*** A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. *Hum Mol Genet.* 2014; 23:2220-2231.
35. **Forbes RB, Colville S, Swingler RJ.** The epidemiology of amyotrophic lateral sclerosis (ALS/MND) in people aged 80 or over. *Age Ageing*. 2004;33:131-134.
36. **Forsberg K, Andersen PM, Marklund SL, Brännström T.** Glial nuclear aggregates of superoxide dismutase-1 are regularly present in patients with amyotrophic lateral sclerosis. *Acta Neuropathol*. 2011;121:623-34.
37. **Gijselinck I, Van Langenhove T, van der Zee J, *et al.*** A *C9orf72* promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol*. 2012;11:54–65.
38. **Glusman G, Caballero J, Mauldin DE, Hood L, Roach J.** KAVIAR: an accessible system for testing SNV novelty. *Bioinformatics*. 2011;27:3216-3217.
39. **Greenway MJ, Alexander MD, Ennis S, *et al.*** A novel candidate region for ALS on chromosome 14q11.2. *Neurology*. 2004;63:1936-1938.
40. **Hardiman O, van den Berg LH, Kiernan MC.** Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nat Rev Neurol*. 2011;7:639-49.
41. **Hewitt C, Kirby J, Highley JR, *et al.*** Novel FUS/TLS mutations and pathology in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol*. 2010;67: 455-461.
42. **Hirano M, Quinzii CM, Mitsumoto H, *et al.*** Senataxin mutations and amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2011;12:223-227.
43. **Jackson CE, Amato AA, Bryan WW, *et al.*** Primary hyperparathyroidism and ALS, Is there a relation? *Neurology*. 1998;50:1795-99.
44. **Johnston CA, Stanton BR, Turner MR, *et al.*** Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *J Neurol*. 2006;253:1642-1643.
45. **Jonsson PA, Ernhill K, Andersen PM, *et al.*** Minute quantities of misfolded mutant superoxide dismutase-1 cause amyotrophic lateral sclerosis. *Brain*. 2004;127:73-88.
46. **Kabashi E, Valdmanis PN, Dion P, *et al.*** *TARDBP* mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet*. 2008;40:572-574.

47. **Krüger S, Battke F, Sprecher A, *et al.*** Rare Variants in Neurodegeneration Associated Genes Revealed by Targeted Panel Sequencing in a GermanALS Cohort. *Front Mol Neurosci.* 2016;9:92.
48. **Koppers M, Groen EJ, van Vught PW, *et al.*** Screening for rare variants in the coding region of ALS-associated genes at 9p21.2 and 19p13.3. *Neurobiol Aging.* 2013;34:1518.e5-7.
49. **Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, *et al.*** Mutations in the *FUS*/*TLS* gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. 2009;323:1205-1208.
50. **Laaksovirta H, Peuralinna T, Schymick JC, *et al.*** Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol.* 2010;9:978-985.
51. **Landrum MJ, Lee JM, Benson M, *et al.*** ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44:862-868.
52. **Langmead B, Salzberg SL.** Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357-9.
53. **Lattante S, Rouleau GA, Kabashi E.** *TARDBP* and *FUS* mutations associated with amyotrophic lateral sclerosis: summary and update. *Hum Mutat.* 2013;34:812-826.
54. **Lemmens R, Race V, Hersmus N, *et al.*** TDP-43 M311V mutation in familial amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry.* 2009;80:354-355.
55. **Li TM, Alberman E, Swash M.** Comparison of sporadic and familial disease amongst 580 cases of motor neuron disease. *J Neurol Neurosurg Psychiatry*. 1988;51:778-784.
56. **Liu F, Liu Q, Lu CX, *et al.*** Identification of a novel loss-of-function *C9orf72* splice site mutation in a patient with amyotrophic lateral sclerosis. *Neurobiol Aging*. 2016;47:219.e1-219.e5.
57. **Ludolph A, Drory V, Hardiman O, Nakano I, *et al.*** A revision of the El Escorial criteria-2015. *Am Lat Scler Frontotemp Degen*. 2015;16:291-292.
58. **Majounie E, Renton AE, Mok K, *et al.*** Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol*. 2012;11:323–30.
59. **McCombe PA and Henderson RD.** Effects of gender in amyotrophic lateral sclerosis. *Gend Med*. 2010;7:557-570.
60. **McKenna A, Hanna M, Banks E, *et al.*** The Genome Analysis Toolkit: a Map Reduce framework for analyzing next-generation DNA sequencing data. *Genome Res*.2010;20:1297-303.
61. **Moreira MC, Klur S, Watanabe M, *et al.*** Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet.* 2004;36:225-227.
62. **Moroianu J and Riordan JF.** Identification of the nucleolar targeting signal of human angiogenin. *Biochem Biophys Res Commun.* 1994;203:1765-1772.
63. **Morrison KE, Harding AE.** Disorders of the motor neurone. Baillieres Clin Neurol. 1994;3: 431–445.Neumann M, Sampathu DM, Kwong LK, et al. Ubiquitinated TDP-43 in

frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 2006;314:130-133.

64. **Mulder DW, Kurland LT, Offord KP, Beard CM.** Familial Adult Motor-Neuron Disease – Amyotrophic lateral sclerosis. *Neurology*. 1986;36:511-517.
65. **Murros K and Fogelholm.** Amyotrophic lateral sclerosis in Middle-Finland: an epidemiological study. *Acta Neurol Scand.* 1983;67:41-47.
66. **Ou SH, Wu F, Harrich D, *et al.*** Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol.* 1995;69:3584-3596.
67. **Pardo CA, Xu Z, Borchelt DR, *et al.*** Superoxide dismutase is an abundant component in cell bodies, dendrites, and axons of motor neurons and in a subset of other neurons. *Proc Natl Acad Sci USA*. 1995;92:954-8.
68. **Pearson JP, Williams NM, Majounie E, *et al.*** Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *J Neurol*. 2011;258: 647-655.
69. **Quinlan AR, Hall IM.** BED Tools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841-842.
70. **Rimmer A, Phan H, Mathieson I, *et al.*** Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* 2014;46:912–918.
71. **Robinson JT, Thorvaldsdóttir H, Winckler W, *et al.*** IntegrativeGenomicsViewer. *Nature Biotechnology.* 2011;29:24-26.
72. **Saracchi E, Castelli M, Bassi MT, *et al.*** A novel heterozygous *SETX* mutation in a patient presenting with chorea and motor neuron disease. *Amyotroph Lateral Scler Frontotemporal Degener*. 2014;15:138-140.
73. **Sebastià J, Kieran D, Breen B, *et al.*** Angiogenin protects motoneurons against hypoxic injury. *Cell Death Differ*. 2009;16:1238-1247.
74. **Shang Y, Huang EJ.** Mechanisms of *FUS* mutations in familial amyotrophic lateral sclerosis. *Brain Res.* 2016;1647:65-78.
75. **Shankaran SS, Capell A, Hruscha AT, *et al.*** Missense mutations inthe progranulin gene linked to frontotemporal lobar degeneration with ubiquitin-immunoreactive inclusions reduce progranulin production and secretion. *J Biol Chem*. 2008;283:1744-1753.
76. **Sheng J and Xu Z.** Three decades of research on angiogenin: a review and perspective. *Acta Biochim Biophys Sin*. 2016;48:399-410.
77. **Shin SD, Pratt AJ, Getzoff ED and Perry JJP.** Amyotrophic Lateral Sclerosis, Advanced Understanding of Neurodegenerative Diseases, Dr Raymond Chuen-Chung Chang (Ed.). Rijeka, Croatia: Intech. 2011;Chapter 20. 417-442.
78. **Ratti A, Corrado L, Castellotti B, *et al.*** *C9ORF72* repeat expansion in a large Italian ALS cohort: evidence of a founder effect. *Neurobiol Aging*. 2012;33:2528.e7–2528.e14.

79. **Renton AE, Majounie E, Waite A, *et al.*** A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome9p21-linked ALS-FTD. *Neuron*. 2011;72:257-268.
80. **Robberecht W, Aguirre T, Van den Bosch L, *et al.*** D90A heterozygosity in the *SOD1* gene is associated with familial and apparently sporadic amyotrophic lateral sclerosis. *Neurology*. 1996;47:1336-9.
81. **Rosen DR, Siddique T, Patterson D, *et al.*** Mutationsin Cu/Zn superoxide dismutase gene are associated with familial amyotrophiclateral sclerosis. *Nature*. 1993;362:59-62.
82. **Schoenmakers EF, Wanschura S, Mols R, *et al.*** Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. *Nature Genetics*. 1995;10:436-44.
83. **Smirnoff, N.** The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytologist*. 1993;125:27–58.
84. **Smith BN, Newhouse S, Shatunov A, *et al.*** The *C9ORF72* expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet*. 2013;21:102–8.
85. **Strong MJ, Hudson AJ, Alvord WG.** Familial amyotrophic lateral sclerosis, 1850–1989: a statistical analysis of the world literature. *Can J Neurol Sci*. 1991;18:45-58.
86. **Taylor JP, Brown RH Jr, Cleveland DW.** Decoding ALS: from genes to mechanism. *Nature*. 2016;539:197-206.
87. **Vance C, Rogelj B, Hortobágyi T, *et al.*** Mutations in *FUS*, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*. 2009;323:1208-1211.
88. **Victor M, Ropper AH, editors.** Adams and Victor's: Principles of Neurology. Degenerative diseases of the nervous system. 7th edn. *McGraw-Hill Professional*. 2000;1106-1174.
89. **Wang K, Li M, Hakonarson H.** ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164.
90. **Zhao ZH, Chen WZ, Wu ZY, *et al.*** A novel mutation in the senataxin gene identified in a Chinese patient with sporadic amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2009;10:118-122.
91. **Zinszner H, Sok J, Immanuel D, *et al.*** TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J Cell Sci*. 1997;110:1741-1750.

9 APPENDIX

I.

II.