Impact of a frequent nearsplice SOD1 variant in Amyotrophic Lateral Sclerosis: optimizing SOD1 genetic screening for gene therapy opportunities

François Muratet, Elisa Teyssou, Aude Chiot, Séverine Boillée, Christian S
Lobsiger, Delphine Bohl, Beata Gyorgy, Justine Guegan, Yannick Marie,
Maria-Del-Mar Amador, et al.

HAL Id: hal-03353407
https://hal.sorbonne-universite.fr/hal-03353407

Submitted on 24 Sep 2021

Impact of a frequent nearsplice *SOD1* variant in Amyotrophic Lateral Sclerosis: optimizing *SOD1* genetic screening for gene therapy opportunities

François Muratet[1], Elisa Teyssou[1], Aude Chiot[1], Séverine Boillée[1], Christian S. Lobsiger[1], Delphine Bohl[1], Beata Gyorgy[1], Justine Guegan[1,] Yannick Marie[1], Maria-Del-Mar Amador[1,2], François Salachas[1,2], Vincent Meininger[3], Emilien Bernard[4,5], Jean-Christophe Antoine[6], Jean-Philippe Camdessanché[6], William Camu[7], Cécile Cazeneuve[8], Anne-Laure Fauret-Amsellem[8], Eric Leguern[1,8], Kevin Mouzat[9], Claire Guissart[9], Serge Lumbroso[9], Philippe Corcia[10,11], Patrick Vourc'h[11,12], Aude-Marie Grapperon[13], Shahram Attarian[13], Annie Verschueren[13], Danielle Seilhean[1,14*] and Stéphanie Millecamps[1*#]

[1]ICM, Institut du Cerveau, Inserm U1127, CNRS UMR7225, Sorbonne Université, Paris.

[2]AP-HP, Département de Neurologie, Centre de référence SLA Ile de France, Hôpital de la Pitié-Salpêtrière, Paris.

[3]Hôpital des Peupliers, Ramsay Générale de Santé, Paris.

[4]Centre de référence SLA, Hôpital Neurologique Pierre Wertheimer, Hospices Civils de Lyon, Université de Lyon, Bron.

[5]Institut NeuroMyoGène, CNRS UMR5310, INSERM U1217, Faculté de Médecine Rockefeller, Université Claude Bernard Lyon I, Lyon.

[6]Service de Neurologie, Centre de Ressource et de Compétence SLA, Hôpital Nord, CHU de Saint-Etienne, Saint-Etienne.

[7]Centre de référence SLA, Hôpital Gui de Chauliac, CHU et Université de Montpellier, Montpellier

[8]Département de Génétique et Cytogénétique, Unité Fonctionnelle de neurogénétique moléculaire et cellulaire, APHP, Hôpital Pitié-Salpêtrière, Paris.

[9]Laboratoire de Biochimie et Biologie Moleculaire, CHU Nimes, Nîmes, Motoneuron Disease: Pathophysiology and Therapy, INM, INSERM, Univ. Montpellier, Montpellier.

[10]Centre SLA, CHU Tours, Tours.

[11]UMR 1253, Université de Tours, Inserm, Tours.

[12]Service de Biochimie et Biologie Moléculaire, CHU Tours, Tours.

[13]Centre de Référence pour les Maladies Neuromusculaire et la SLA, CHU de Marseille, Hôpital de la Timone, Marseille.

[14]AP-HP, Département de Neuropathologie, Hôpital Pitié-Salpêtrière, Paris.

[*]These authors jointly directed this work

[#]correspondence should be addressed to

Dr Stéphanie Millecamps, Institut du Cerveau, Hôpital Pitié-Salpêtrière, 47-83 Bd de l'Hôpital, CS21414, 75646 Paris Cedex, France, tel : +33 1 57 27 43 41, e-mail: stephanie.millecamps@upmc.fr

Abstract

Objective. Mutations in superoxide dismutase 1 gene (*SOD1)*, encoding copper/zinc superoxide dismutase protein, are the second most frequent high penetrant genetic cause for Amyotrophic Lateral Sclerosis (ALS) motor neuron disease in populations of European descent. More than 200 missense variants are reported along the SOD1 protein. To limit the production of these aberrant and deleterious SOD1 species, antisense oligonucleotide approaches have recently emerged and showed promising effects in clinical trials. To offer the possibility to any patient with SOD1-ALS to benefit of such a gene therapy, it is necessary to ascertain whether any variant of unknown significance (VUS), detected for example in *SOD1* non-coding sequences, is pathogenic.

Methods. We analyzed SOD1 mutation distribution after SOD1 sequencing in a large cohort of 470 French familial ALS (fALS) index cases.

Results. We identified a total of 27 SOD1 variants in 38 families including two SOD1 variants located in nearsplice or intronic regions of the gene. The pathogenicity of the c.358-10T>G nearsplice *SOD1* variant was corroborated based on its high frequency (as the second most frequent SOD1 variant) in French fALS, the segregation analysis confirmed in eight affected members of a large pedigree, the typical SOD1-related phenotype observed (with lower limb onset and prominent lower motor neuron involvement), and findings on post-mortem tissues showing SOD1 mis-accumulation.

Conclusions. Our results highlighted nearsplice/intronic mutations in *SOD1* are responsible for a significant portion of French fALS and suggested the systematic analysis of the *SOD1* mRNA sequence could become the method of choice for *SOD1* screening, not to miss these specific cases.

INTRODUCTION

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative motor neuron disease that usually develops in the sixth decade and is uniformly fatal, with a median of survival of 36 months.[1] It leads to limb paralysis, swallowing/speech impairments and ultimate respiratory failure. In 10% of cases it is associated with family history (familial ALS, fALS), with multiple genes involved.[1] Superoxide dismutase 1 gene *(SOD1)* was the first genetic cause discovered more than 25 years ago.[2 3] In the discovery study, eleven missense variants were identified at the heterozygous state in 13 families with ALS, supporting the autosomal dominant implication of *SOD1* in this motor neuron disease.[3] Exception for this dominant transmission is known for the recessive D90A (corresponding to the p.Asp91Ala) variant with either frequent homozygosity in ALS populations of Swedish and Finnish origin[4] or compound heterozygosity with the D96N (corresponding to the p.Asp97Asn) variant.[5] Incomplete penetrance has been observed for dominant SOD1 variants, including the I113T mutation, and the unaffected transmitters were more frequently women than men in those pedigrees.[6] Rare *de novo* occurrence has also been reported for sporadic case.[7] To date, about 200 variants,[8] most of them being missense, have been identified, distributed all along the *SOD1* gene which encodes the superoxide dismutase 1 protein (SOD1), a 32 kDa homo-dimeric metalloenzyme composed of 154 amino acids highly conserved among species (however, as the methionine start codon was not previously considered as part of SOD1, the number of amino acids is also considered to be 153 and the numbering of mutations is, in this case, shifted of one amino acid). It is well admitted that mutant SOD1 forms are deleterious through a gain of toxic function although the exact mechanisms leading to neurodegeneration are not fully elucidated.[9] *SOD1* variant frequency in ALS patients is different according to the studied population: they account for 30% of fALS patients in Asia, while they explain 23.5% of US fALS[10 11] and 12% of French fALS.[12] This

therefore makes variants in *SOD1* either the first or second most frequent familial cause of ALS in these regions.

ALS patients with *SOD1* variants usually have disease onset in the lower limbs, prominent lower motor neuron signs resembling progressive muscular atrophy and rarely present frontotemporal dementia (FTD).[12][13] Disease duration is bimodal with patients having rather rapid (<3 years) ALS while others having rather slow (>7 years) disease course,[12,14] with a either consistent prognosis for all patients carrying a specific mutation across all families (for example the short survival conferred by the p.Ala5Val, A4V mutation) or with a significant intra-family variability in survival, documented for example, for the p.Asp84Gly, D83G or the p.Leu145Phe, L144F mutations, the latest with the longest survival reported to date in SOD1 linked family (44 years).[12][15]

Neuropathological analysis of post-mortem tissues with *SOD1* variants showed that TDP-43 inclusions, the pathological hallmark usually detected within motor neurons of most ALS patients,[16] are often missing.[17] Ubiquitinated SOD1 aggregates and neurofilament conglomerates are typically detected in these tissues.[18]

However, whether all rare variants in the nearsplice/intronic regions of *SOD1* can cause the disease, which is of prime importance for new gene therapy opportunities, remained unsolved. In this study, we documented a frequent *SOD1* nearsplice variant in French fALS localized 10 bp upstream of *SOD1* exon 5 for which we aimed to undoubtedly prove the pathogenicity before considering SOD1 targeted therapeutic strategy[8] in these families.

METHODS

Methods for patient and control cohorts, genetic analyses, mRNA extraction and splicing analysis, prediction for the protein secondary structure, neuropathological examination and protein level analysis in patient spinal cord are detailed in supplementary material.

RESULTS

The c.358-10T>G is the second most frequent variant in French fALS and segregates with the disease in a large pedigree.

We identified a total of 27 SOD1 variants in 38 families in our cohort of French fALS, accounting for a frequency of 8.1%. These variants are detailed in Fig.S1. Among these variants, the c.280G>T, p.Gly94Cys (G93C), occurred in 5 different families and was the most frequent (Fig.S1). The second most frequent SOD1 variant was the nearsplice c.358-10T>G, which was recurrently observed in 4 different families with no obvious family relationship, with a frequency of 10.5%, in these rankings. This c.358-10T>G variant has a frequency of 0.0018% in gnomAD non Finnish (NFE) population, is recorded in dbSNP (rs1197141604) and ClinVar (RCV000015893.26) databases. In addition, the c.443G>A, p.Gly148Asp (G147D) occurred in 3 families and both c.124G>A, p.Gly42Ser (G41S) and c.418A>G, p.Asn140Asp (N139D) variants were identified in two different families. None of the exonic, nearsplice or intronic SOD1 variants we identified in ALS patients (Fig.S1) was detected in our cohort of 525 healthy French controls.

One of these four c.358-10T>G families extends on 6 generations and includes 22 affected members: we could gather the DNA collected in different French ALS reference centers for 8 affected patients, over 4 generations, and the c.358-10T>G variant was identified at the heterozygous state in all of them (Fig.1). Obligate carriers were evidenced in this pedigree, however missing information about the cause and the age of death for most of them did not allow to firmly conclude about incomplete penetrance for the c.358-10T>G variant. For example one of these obligate carriers (FALS1II2) deceased with no motor neuron symptom at 50 years and could have developed the disease if he would have lived longer. Nevertheless the penetrance of the c.358-10T>G variant seems to be high in the three last generations (with

6

full knowledge about the cause of death and no censored data) which is an important information for genetic counselling to be delivered to the concerned families. Pedigrees of the three other unrelated fALS are presented in Fig.1. Overall, the c.358-10T>G variant was identified in 11 French fALS patients of four unrelated families. The most extended family (FALS1) was included in a linkage analysis using the genotype at the mutation site as marker. This analysis resulted in a LOD score of 3.67 which exceeds the genome-wide significance threshold confirming the association of the disease with the nearsplice SOD1 variant in this pedigree. Whole Exome Sequencing (WES) analysis was performed for all patients carrying the c.358-10T>G variant (with available DNA) to check for the absence of other pathogenic variants in any other ALS-linked genes (Table S1).

Clinical presentation of the patients with the c.358-10T>G variant is typical of SOD1-linked ALS.

Age at onset and site at onset could be retrieved for 17 members of these c.358-10T>G families (10 males and 7 females). Mean age at onset was 60 years old (±3 years), with a range from 41 to 83 years of age. No difference was noticed between males (60±4y, range 41-83y) and females (59±5y; range 42-71y). The first symptoms appeared in legs for most (14 out of 17; 82%) of these patients (Table 1). Mean disease duration was 43 months (±7 months) ranging from 8 to 96 months with no gender effect. Age at death was available for 20 patients (11 males and 9 females) and mean was 65±4y for both gender groups. Transcranial magnetic stimulation (TMS) recordings performed for some c.358-10T>G patients did not show any abnormality (Table 1).

Pedigree	Patient	Sex	Age at onset, (range of y)	Site of onset	Disease duration, mo	TMS recordings	DNA availability	Lymphoblast availability	Post-mortem tissue availability
FALS1	III1	M	61-80	LL	96	-	-	-	-
	III8	M	≥81	UL	30	normal	+	+	-
	IV1	F	61-80	LL	36	-	+	-	-
	IV12	M	61-80	UL	35	normal	+	+	-
	IV13	F	41-60	LL	84	-	-	-	-
	IV14	M	41-60	LL	57	normal	+	+	-
	IV20	F	61-80	LL	80	normal	+	+	-
	V1	F	41-60	LL	60	-	-	-	-
	V2	M	41-60	UL	17	-	+	+	-
	V3	F	41-60	LL	16	-	+	-	-
	VI1	M	41-60	LL	-	-	+	-	-
FALS2	II2	M	41-60	LL	24	-	-	-	-
	II3	F	61-80	LL	12	-	-	-	-
	III1	M	41-60	LL	8	-	+	-	+
FALS3	I1	F	61-80	LL	48	-	-	-	-
	II1	M	61-80	LL	30	-	+	-	-
FALS4	II1	M	41-60	LL	53	-	+	+	-

Table 1. Clinical features of patients with the SOD1 c.358-10T>G variant.

Age at onset is indicated as a range of years to preserve individual anonymization. fALS, familial amyotrophic lateral sclerosis; LL, lower limb; TMS, Transcranial Magnetic Stimulation UL, upper limb.

Analysis of the splicing effect of the c.358-10T>G variant on patient tissue and cells

The variant is a nearsplice point substitution 10 bp upstream exon 5, c.358-10T>G (Fig.2A) which is presumed to affect the intron 4 acceptor splicing site according to prediction software tools accessed through Alamut Visual 2.15 (Interactive Biosoftware): it is predicted to abolish the natural acceptor splice site according to SSF-like and GeneSplicer and to create a new acceptor splice site according to SSF-like, MaxEntScan and Human Splicing Finder system.

We used the post-mortem tissue available for one of the patient with c.358-10T>G variant (FALS2III1, Fig.1) to confirm these predictions and analyzed the mRNA products after amplification of a fragment spanning exon 4 and 5 from the *SOD1* cDNA: two longer additional

bands were visualized for tissue carrying the c.358-10T>G variant compared to the single expected band amplified for other patients (Fig.2B). The longest one is presumably a heteroduplex formed between the wild-type (WT) and the mutant strands because only two strands were detected by the fluorescent fragment-length analysis used to determine the exact length of these amplified products (Fig.2C). A similar heteroduplex DNA structure was previously detected after amplifying mRNA produced in ALS fibroblasts carrying the c.240-7T>G, another nearsplice/intronic *SOD1* variant.[19] Thus using patient post-mortem tissue, we showed this nearsplice variant alters the splicing of *SOD1*. Sequencing of the amplified products revealed that this c.358-10T>G substitution led to a 9 nucleotide insertion between exon 4 and 5 in the mRNA, resulting in the insertion of 3 amino acids (Phe-Leu-Gln, FLQ) in the protein sequence (Fig.2A). Similar results were obtained on mRNA products extracted from lymphoblasts available for 6 other ALS patients carrying this *SOD1* nearsplice variant (including 5 members of FALS1 and the index case of FALS4, Table 1). For these patients, the PCR products spanning exon 4 and 5, amplified from the *SOD1* cDNA, migrated with a "three band" profile (Fig.S2A) and the presence of the 9 nucleotide insertion could be confirmed after sequencing these fragments (Fig.S2B). Thus the insertion of the 3 amino acids FLQ in SOD1 sequence could be ascertained in cells/tissue for 7 patients from 3 different families. No cell or tissue were available for the last family (FALS3).

This 3 amino acid insertion is predicted to modify the secondary structure of the resulting mutant SOD1 protein according to JPred 4 analysis, turning a beta sheet into an alpha helix, which could directly impact a dimerization site and a copper binding site (Fig.2D).

Neuropathological analysis of spinal cord carrying the *SOD1* c.358-10T>G variant.

The most characteristic lesions found at autopsy in one *SOD1* c.358-10T>G patient (FALS2III1) was severe neuronal loss associated with numerous swollen neurites in the

anterior and lateral horns of the spinal cord and in motor nuclei of the medulla. Bulky SOD1 intracytoplasmic inclusions with grape cluster shape were found in the neuronal cell bodies and in the proximal part of swollen axons (Fig.3D, 3H). There were some diffuse positivities in some interneurons and glial cells. Diffuse staining was also present in posterior and intermediate columns of the spinal cord. No staining was observed in the white matter of the posterior cords. Using the same antibody, SOD1 aggregates were observed in the motor neurons with the c.251A>G, p.Asp84Gly, D83G (Fig.3B, 3F) or c.281G>A, p.Gly94Asp, G93D (Fig.3C, 3G) SOD1 exonic variant. No SOD1 aggregates were found in the cell bodies of patient tissue with no SOD1 variant (Fig.3A, 3E). Some perivascular SOD1 positive swellings were observed in SOD1 mutated patients and controls. Typical inclusions were observed with phosphorylated or non-phosphorylated neurofilament immunohistochemistry in SOD1 variant carriers (Fig.3I-J, M-N). Most of these inclusions were slightly labeled with p62. Cystatin C immunohistochemistry showed some Bunina bodies (Fig.3O) and numerous large granular aggregates (Fig.3P). TDP-43 immunohistochemistry showed no abnormal positivity in the regions tested for patients with nearsplice or exonic SOD1 mutation (Fig.3K-L).

There were no visible abnormalities in the motor cortex where large Betz cells were present. Nevertheless inconspicuous axonal rarefaction was found in the corticospinal tracts. Rare intracytoplasmic SOD1 diffuse positive inclusions were observed in the frontal associative cortex (Fig.S3A-B) and in the cerebellum (Fig.S3C). A punctiform, neuropil positivity was observed in the frontal, motor and associative cortices examined.

Analysis of SOD1 forms expressed in spinal cord with the *SOD1* c.358-10T>G variant.

In mouse models of ALS expressing human mutant SOD1, the formation of SOD1 aggregates was proposed to be triggered by the disulfide-reduced forms of SOD1, more sensitive to

oxydative stress and having the property to misfold and oligomerize.[20] We hypothesized that these detergent-resistant, disulfide linked complexes could represent a significant proportion of SOD1 in the aggregates detected in motor neurons of ALS patients with SOD1 mutation. These disulfide cross-linked SOD1 forms are not detectable using classic Western blot procedures. We adapted a previously reported protocol combining the omission of reducing agents (expected to reduce disulfide cross-linked species) and the use of iodoacetamide, a thiol-specific modifying reagent capable of preventing degradation of these disulfide-reduced forms of SOD1 during insoluble protein extraction, to maximize their detection during the western blot procedure. We compared SOD1 levels in the soluble and insoluble fractions of protein extracts prepared from spinal cords of the FALS2III1 carrying the nearsplice c.358-10T>G SOD1 variant, a patient carrying the D83G SOD1 mutation and a patient without any SOD1 variant. In the soluble fraction, a single band of 16 kDa was detected using the anti-SOD1 antibody. The level of expression of this soluble SOD1 monomeric form was similar for all patients, carrying or not a SOD1 mutation. In contrast, in the iodoacetamide treated insoluble fraction, the level of expression of the 16 kDa SOD1 was decreased for the patients with the nearsplice SOD1 variant and the D83G mutation. After overexposure of the signal, additional species (between 30 and 60kDa) were observed. These bands could be oligomerized disulfide-reduced forms of SOD1 and some of them are shared for the D83G SOD1 exonic mutation, suggesting their contribution to the SOD1 toxicity.

DISCUSSION

The development of therapeutic strategies against SOD1[8 21] have modified the appreciation level for SOD1 variants. The validation of the pathogenicity of all SOD1 variants, including nearsplice/intronic ones, has become a paradigm before envisaging any gene therapy intervention targeting *SOD1* by antisense oligonucleotide or microRNA delivered by adeno-associated virus (AAV) to prevent SOD1 accumulation.[8 21] Our study was dedicated to a variant

in the last intron of the gene, the c.358-10T>G substitution, particularly frequent in French patients. Even if the distribution of *SOD1* variants is known to be heterogeneous depending on geographic area, illustrated for instance for the most frequent A4V variant in the U.S. rarely found in other world areas,[2 14 22] the frequency of the c.358-10T>G variant among French fALS deserves to be highlighted. Additional arguments in favor of the pathogenicity of this nearsplice variant is the segregation and the linkage analyses provided, the typical SOD1-linked disease presentation,[12] and our observations in post-mortem tissue. Our mRNA product analyses are in line with a previous study having detected the c.358-10T>G variant in a single fALS case and performed analyses on cDNA extracted from lymphoblastoid cell lines.[23] Moreover, the JPred4 analysis we performed predicts that FLQ introduction impacted the SOD1 secondary structure by creating a new alpha helix. We showed strong SOD1 accumulation in FLQ-SOD1 patient motor neurons. Moreover using a method that preserve the disulfide status of proteins, to analyze the SOD1 forms in spinal cord protein extracts, we observed a decreased in insoluble SOD1 monomer levels in patients carrying mutations in SOD1 (either the nearsplice variant or a validated exonic SOD1 mutation) with a concomitant appearance of high molecular weight species of SOD1, that could be disulfide cross-linked multimers of SOD1, the misfolding and oligomerization of which are proposed to lead to the formation of insoluble aggregates.[20] Such SOD1 aggregates are proposed to be pathogenic and to contribute to the initiation and the spreading of the disease through prion-like properties.[24] The design of an antibody specific for the FLQ-SOD1 could be valuable to ascertain the presence of this specific form of SOD1 within the aggregates we observed in motor neurons. However our findings are strongly supportive of the pathogenicity of this variant, comfort the SOD1 diagnosis and suggested that gene therapy intervention targeting *SOD1* by antisense oligonucleotide to prevent SOD1 accumulation, would be valuable for these patients.

At least ten nearsplice/intronic *SOD1* variants, affecting all the *SOD1* introns, have been reported in ALS patients (Fig.S4A-B). The impact on mRNA and/or protein sequence was determined for only 4 of them (Fig.S4C), ascertaining 2 in-frame introductions of few amino acids in the SOD1 sequence or 2 frameshifts leading to the synthesis of a truncated SOD1 protein missing the last 33 amino acids with a modified C-terminal part compared to the WT sequence.[19] [23] [25] [26] Other insertions/deletions within the *SOD1* coding sequence are also frameshift variants identified in ALS patients (Fig.S4C). Most of the SOD1 forms with premature stop codon, notably those identified in the last 5th exon of *SOD1*, are not supposed to be targeted by the Nonsense-mediated mRNA Decay (NMD) ensuring the rapid degradation of mRNAs containing premature translation termination codon.[27] Thus the resulting truncated SOD1 proteins are supposed to be toxic through abnormal aggregation within neurons and the therapeutic strategy would be to reduce the production of these abnormal SOD1 forms. In line with this assumption, deposits of SOD1 were detected in the spinal cord and the brainstem of two patients carrying the truncated p.Gly142* (G141X) variant showing these truncated forms accumulated in neurons and glial cells.[28]

Nevertheless two recent studies have nuanced this paradigm and reported a new homozygous truncating variant: c.335dupG, p.Cys112Trpfs*11 in 2 different families with consanguinity.[29] [30] Analysis on patient lymphoblasts showed this mutant mRNA escapes the NMD process: the resulting protein misses 43 amino acids at the end of the 4th exon and contains 10 additional ones.[29] However, parents, heterozygous for the variant, did not present any motor phenotype. In contrast all infants, homozygous for the variant, presented severe spastic motor neuron phenotype with mild cerebellar atrophy, which was suggested to be due to the complete loss of SOD1 enzymatic activity.[29,30] These observations on recessive SOD1 variants transmittance make more complex the genetic counseling and the design of therapeutic strategies based on the SOD1 silencing.

Nowadays SOD1 genetic screening is performed through targeted gene panels, exome or even genome sequencing. However attention should be drawn on filters applied to select variants of interest. Many nearsplice/intronic variants not directly qualified as "splicing site" and with a supposed "low impact" can be missed with a too drastic filter query on those targeted gene panel, exome and genome sequencing databases. Indeed "splice variant" is often selected as a change occurring within the region of the splice site. Splice site region definition is therefore crucial. For example, if this splicing region query covers only, the 1-3 bases of the exon and 3-8 bases of the intron, most nearsplice/intronic SOD1 variants reported to date would not be selected and would be missed (Fig.S4A). Previous analyses performed on exome sequencing data did not report such nearsplice/intronic variants in ALS genes: for example a c.2993-14G>T variant in the intron 26 of the *KIF5A* gene and affecting *KIF5A* exon 27 splicing, recently identified after genome sequencing, was not selected, after the exome sequencing initially performed in the same patient DNA.[31][32] Indeed a p.Phe115Cys variant in *MATR3* was rather initially proposed as the causing variant in this individual and conclusions have changed, after the segregation analysis could recently be performed in this family, to incriminate the *KIF5A* variant.[32]

Novel variants could be identified in under explored intronic regions of the *SOD1* gene. As it becomes crucial to avoid excluding any SOD1-linked patients from upcoming clinical trial studies targeting the *SOD1* gene,[8][21] a genetic screening based on the amplification of the entire SOD1 transcript prepared from patient blood cells could become an alternative approach for *SOD1* diagnosis. Indeed *SOD1* transcript is short and encompasses only 465 pb. We designed specific primers and PCR amplification conditions to optimize the direct detection of abnormal *SOD1* transcripts using a single step amplification (Fig.4). These conditions allowed the perfect detection of both c.358-10T>G and c.358-304C>G frameshifts on cDNA sequence and would be useful for the detection of novel deep intronic SOD1 mutants. Such an approach in routine

diagnosis could (i) rapidly give, to the concerned ALS patients, the possibility to beneficiate from antisense therapeutic proposals, (ii) identify undiscovered nearsplice/intronic variants in *SOD1* responsible for the disease in broader populations and, (iii) as the mRNA sequence would be precisely defined, it would specify whether the available antisense strategy is designed for these patients or if they would need other specific antisense.

Altogether, our results undoubtedly evidence the pathogenicity of the c.358-10T>G *SOD1* variant and point out the attention to be payed on variants detected in nearsplice/intronic regions of the *SOD1* gene, even if they are not qualified as evident splice site by interfaces developed to analyze targeted gene panel, exome or genome sequencing data. Our study further suggests that systematic analysis of the complete *SOD1* cDNA sequence could become the method of choice for *SOD1* gene screening in diagnosis to avoid missing these specific cases. This could be performed for example on mRNA prepared from any tissue available including blood, collected on PAXgene[TM] Blood RNA Tube. Indeed a complete and detailed screening of the *SOD1* transcript would be useful to maximize the number of patients and pre-symptomatic carriers to be part of ongoing or future clinical trials targeting SOD1 by antisense oligonucleotide (ASO)[8] or microRNA delivered by AAV[21] and would be beneficial to design future strategy if required. A similar approach could be applied for the screening of the *FUS* gene in which many frameshifts in the C-terminal part of the protein have been reported to cause early onset ALS.[33]

Acknowledgements

References

1   Al-Chalabi A, Hardiman O. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat Rev Neurol* 2013;**9**:617–28.

2   Deng HX, Hentati A, Tainer JA, *et al.* Amyotrophic lateral sclerosis and structural defects in Cu,Zn superoxide dismutase. *Science* 1993;**261**:1047–51.

3   Rosen DR, Siddique T, Patterson D, *et al.* Mutations in Cu/Zn Superoxide Dismutase Gene Are Associated with Familial Amyotrophic Lateral Sclerosis. *Nature* 1993;**362**:59–62.

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   Hand CK, Mayeux-Portas V, Khoris J, *et al.* Compound heterozygous D90A and D96N SOD1 mutations in a recessive amyotrophic lateral sclerosis family. *Ann Neurol* 2001;**49**:267–71.

6   Andersen PM, Sims KB, Xin WW, *et al.* Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2003;**4**:62–73.

7   Alexander MD, Traynor BJ, Miller N, *et al.* "True" sporadic ALS associated with a novel SOD-1 mutation. *Ann Neurol* 2002;**52**:680–3.

8   Miller T, Cudkowicz M, Shaw PJ, *et al.* Phase 1–2 Trial of Antisense Oligonucleotide Tofersen for *SOD1* ALS. *N Engl J Med* 2020;**383**:109–19.

9   Bunton-Stasyshyn RKA, Saccon RA, Fratta P, *et al.* SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology: New and Renascent Themes. *Neurosci Rev J Bringing Neurobiol Neurol Psychiatry* 2015;**21**:519–29.

10 Battistini S, Giannini F, Greco G, *et al.* SOD1 mutations in amyotrophic lateral sclerosis: Results from a multicenter Italian study. *J Neurol* 2005;**252**:782–8.

11 Zou Z-Y, Zhou Z-R, Che C-H, *et al.* Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry* 2017;**88**:540–9.

12 Millecamps S, Salachas F, Cazeneuve C, *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. *J Med Genet* 2010;**47**:554–60.

13 Millecamps S, Boillée S, Le Ber I, *et al.* Phenotype difference between ALS patients with expanded repeats in *C9ORF72* and patients with mutations in other ALS-related genes. *J Med Genet* 2012;**49**:258–63.

14 Cudkowicz ME, McKenna-Yasek D, Sapp PE, *et al.* Epidemiology of mutations in superoxide dismutase in amyotrophic lateal sclerosis. *Ann Neurol* 1997;**41**:210–21.

15 Weber M, Neuwirth C, Thierbach J, *et al.* ALS patients with SOD1 mutations in Switzerland show very diverse phenotypes and extremely long survival. *J Neurol Neurosurg Psychiatry* 2012;**83**:351–3.

16 Neumann M, Sampathu DM, Kwong LK, *et al.* Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science* 2006;**314**:130–3.

17 Tan C-F, Eguchi H, Tagawa A, *et al.* TDP-43 immunoreactivity in neuronal inclusions in familial amyotrophic lateral sclerosis with or without SOD1 gene mutation. *Acta Neuropathol (Berl)* 2007;**113**:535–42.

18 Kato S, Takikawa M, Nakashima K, *et al.* New consensus research on neuropathological aspects of familial amyotrophic lateral sclerosis with superoxide dismutase 1 (SOD1) gene mutations: Inclusions containing SOD1 in neurons and astrocytes. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000;**1**:163–84.

19 Birve A, Neuwirth C, Weber M, *et al.* A novel SOD1 splice site mutation associated with familial ALS revealed by SOD activity analysis. *Hum Mol Genet* 2010;**19**:4201–6.

20 Furukawa Y, Fu R, Deng H-X, *et al.* Disulfide cross-linked protein represents a significant fraction of ALS-associated Cu, Zn-superoxide dismutase aggregates in spinal cords of model mice. *Proc Natl Acad Sci* 2006;**103**:7148–53.

21 Mueller C, Berry JD, McKenna-Yasek DM, *et al. SOD1* Suppression with Adeno-Associated Virus and MicroRNA in Familial ALS. *N Engl J Med* 2020;**383**:151–8.

22 Juneja T, Pericak-Vance MA, Laing NG, *et al.* Prognosis in familial amyotrophic lateral sclerosis: progression and survival in patients with glu100gly and ala4val mutations in Cu,Zn superoxide dismutase. *Neurology* 1997;**48**:55–7.

23 Sapp PC, Rosen DR, Hosler BA, *et al.* Identification of three novel mutations in the gene for CuZn superoxide dismutase in patients with familial amyotrophic lateral sclerosis. *Neuromuscul Disord* 1995;**5**:353–7.

24 Ekhtiari Bidhendi E, Bergh J, Zetterström P, *et al.* Mutant superoxide dismutase aggregates from human spinal cord transmit amyotrophic lateral sclerosis. *Acta Neuropathol (Berl)* 2018;**136**:939–53.

25 Valdmanis PN, Belzil VV, Lee J, *et al.* A Mutation That Creates a Pseudoexon in SOD1 Causes Familial ALS. *Ann Hum Genet* 2009;**73**:652–7.

26 Zu JS, Deng H-X, Lo TP, *et al.* Exon 5 encoded domain is not required for the toxic function of mutant SOD1 but essential for the dismutase activity: identification and characterization of two new SOD1 mutations associated with familial amyotrophic lateral sclerosis. *neurogenetics* 1997;**1**:65–71.

27 Zhang Z, Xin D, Wang P, *et al.* Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay. *BMC Biol* 2009;**7**:23.

28 Nakamura M, Bieniek KF, Lin W-L, *et al.* A truncating SOD1 mutation, p.Gly141X, is associated with clinical and pathologic heterogeneity, including frontotemporal lobar degeneration. *Acta Neuropathol (Berl)* 2015;**130**:145–57.

29 Andersen PM, Nordström U, Tsiakas K, *et al.* Phenotype in an Infant with *SOD1* Homozygous Truncating Mutation. *N Engl J Med* 2019;**381**:486–8.

30 Park JH, Elpers C, Reunert J, *et al.* SOD1 deficiency: a novel syndrome distinct from amyotrophic lateral sclerosis. *Brain* 2019;**142**:2230–7.

31 Johnson JO, Pioro EP, Boehringer A, *et al.* Mutations in the Matrin 3 Gene Cause Familial Amyotrophic Lateral Sclerosis. *Nat Neurosci* 2014;**17**:664–6.

32 Saez-Atienzar S, Dalgard CL, Ding J, *et al.* Identification of a pathogenic intronic KIF5A mutation in an ALS-FTD kindred. *Neurology* 2020;**95**:1015–8.

33 Kapeli K, Martinez FJ, Yeo GW. Genetic mutations in RNA-binding proteins and their roles in ALS. *Hum Genet* 2017;**136**:1193–214.
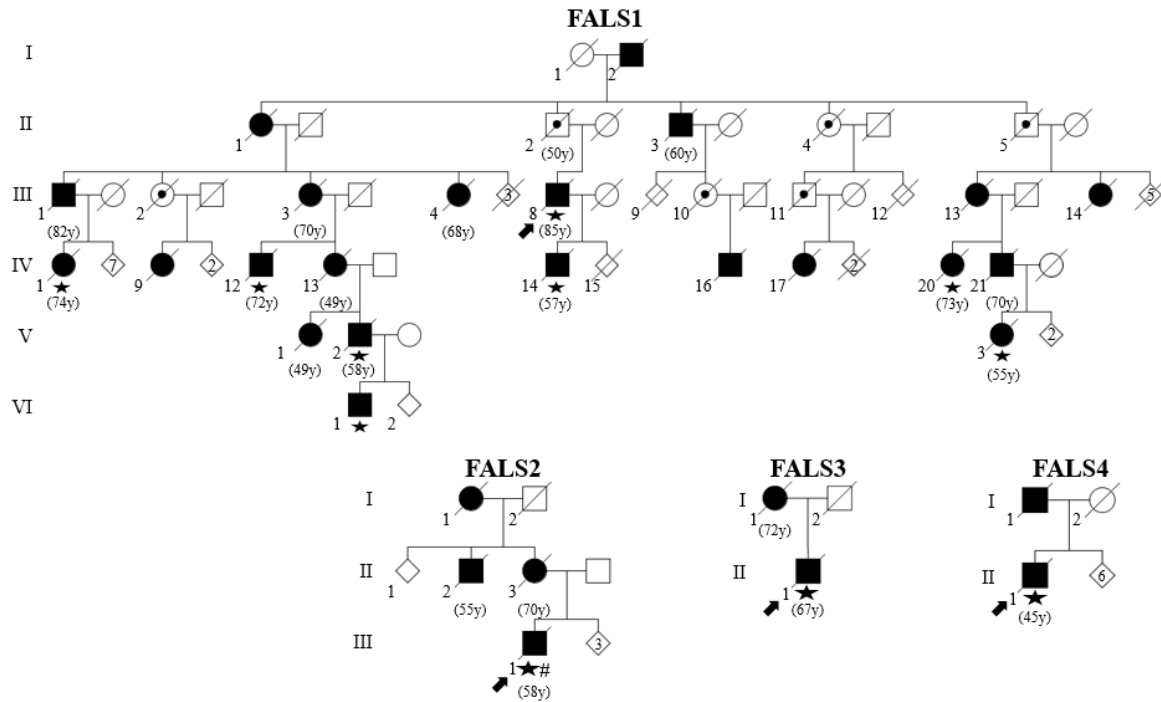
Fig.1. Pedigrees of French fALS with the SOD1 c.358-10T>G variant.

The variant segregates in 8 affected members of the FALS1 family with available DNA (pointed by a STAR). When available, the age at death is indicated below the symbol representing the patient (in brackets). Arrows indicate index cases. Black fill: Amyotrophic Lateral Sclerosis (ALS) case; black dot: unaffected obligate carrier. To ensure de-identification of the families diamonds for some individuals/branches hide the corresponding gender. #: post-mortem tissue analyzed.
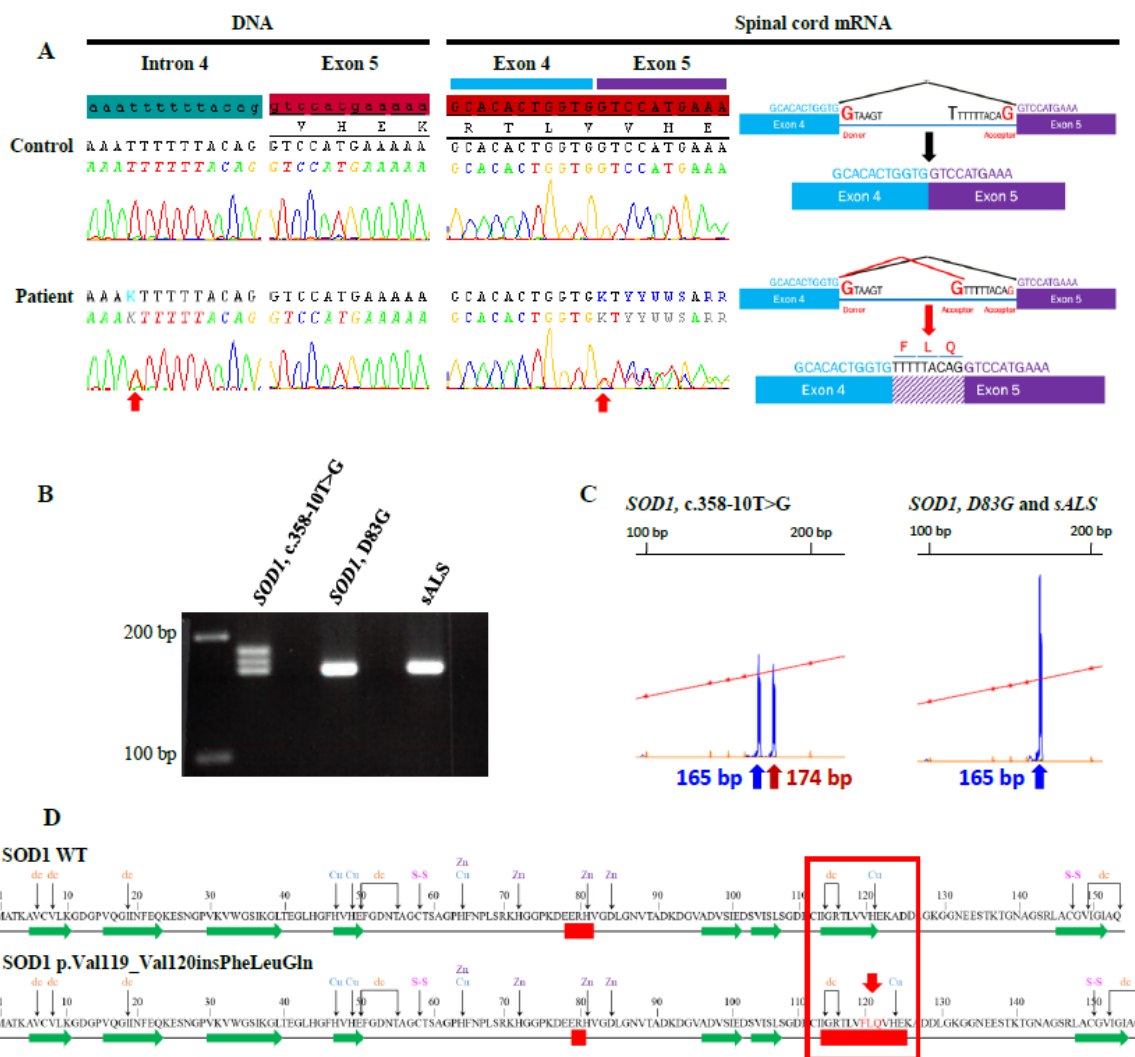
Fig.2. The SOD1 c.358-10T>G variant disturbs splicing and SOD1 protein secondary structure.

(A) Parts of chromatograms showing the position of the c.358-10T>G variant (red arrows) in the DNA and the spinal cord cDNA sequences (amplified from mRNA) compared to that of a control. The c.358-10T>G variant creates a new acceptor splice site leading to the addition of three amino acids Phenylalanine-Leucine-Glutamine, (Phe-Leu-Gln, FLQ) between exons 4 and 5. (B) PCR products amplified from patient spinal cord cDNA with primers located in exons 4 and 5 of the SOD1 gene. Three fragments were detected in the SOD1 c.358-10T>G patient and a single fragment of 165 bp in other ALS patients, including one sporadic ALS, sALS, and one fALS carrying the c.251A>G, p.Asp84Gly, D83G, SOD1 variant. (C) Fluorescent fragment length analyses of the PCR fragments, requiring formamide addition to destabilize double-stranded DNA structures and visualized using Peak Scanner software v1.0, revealed the presence of two fragments of 165 bp and 174 bp, in the SOD1 c.358-10T>G patient (left), and one fragment of 165 bp in the other patients (right). Thus the longest PCR band detected in (B) is presumably a heteroduplex formed between the wild-type (WT) and the mutant strands. The red line shows the GeneScan 500 LIZ dye Size standard (Applied Biosystems). (D) Schematization of WT and mutated SOD1 protein secondary structure compiled from the Jpred4 consensus prediction scores (provided by the JPred4 website) and 2. The insertion of the three amino acids changes a beta sheet (green arrow) into an alpha helix (red rectangle), affecting a dimerization site (orange) and a copper binding site (blue). Zinc binding (purple) and disulfide bonds (pink) sites are also indicated.SOS1, superoxide dismutase 1.
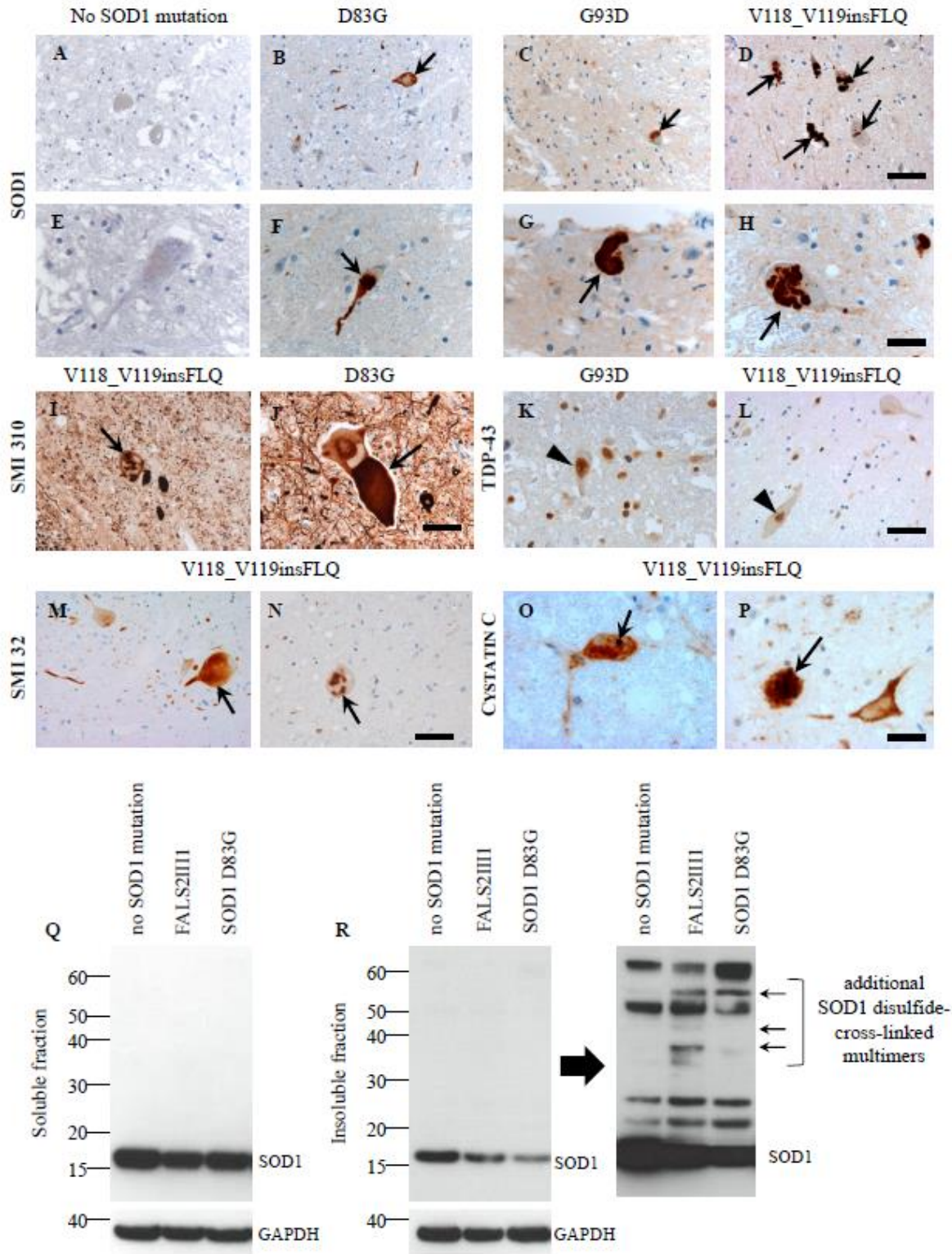
Fig.3. SOD1 accumulates and multimerizes in spinal cord of patient with the c.358-10T>G nearsplice SOD1 variant.

Spinal cord sections of ALS patients with no SOD1 mutation (A, E) or carrying the D83G (B, F, J), the G93D (C, G, K) or the V118_V119insFLQ SOD1 (D, H, I, L-P) variant (FALS2III1 patient) were stained with anti-SOD1 (A-H), anti-SMI 310 (I-J), anti-TDP-43 (K-L), anti-SMI 32 (M-N) and anti-Cystatin C (O-P) antibodies. Cytoplasmic SOD1 accumulations (A, arrows) were observed in the motor neurons with the D83G (B, F, arrows)

and the G93D (C, G, arrows) SOD1 exonic mutations. Clusters of immunopositive SOD1 were noted in the motor neurons of the patient (FALS2III1) with the nearsplice SOD1 variant (D, H, arrows). No such SOD1 accumulations were visible in patients with no mutation in SOD1 (A, E). Phosphorylated (I, arrow) or unphosphorylated (M, N, arrow) neurofilament inclusions were stained in the axonal swellings for the patient with the nearsplice SOD1 variant as well as for SOD1 exonic mutations (J). Anti-Cystatin C staining showed Bunina bodies (O, arrow) and large round deposits (P, arrow). A normal nuclear localization of TDP-43 was noted for SOD1 patients (K-L, arrowhead). Scale bar: 20 µm (A-D, I,-N), 10 µm (E-H, O-P). Immunoblot analyses of soluble (Q) and insoluble (R) protein extracts from transversal sections of spinal cord prepared from one ALS patient with no variant in SOD1, patient FALS2III1 with the nearsplice SOD1 variant (V118_V119insFLQ) or patient carrying the D83G SOD1 mutation, using anti-SOD1 and anti-GAPDH antibodies. In insoluble fraction, the band of SOD1 monomers (16 kDa) was decreased in patients with SOD1 variants (R, left panel). Prolonged exposition (R, right panel) revealed the presence of additional SOD1 multimers in extracts prepared from patient spinal cord with SOD1 variants compared to those observed without mutation in SOD1.
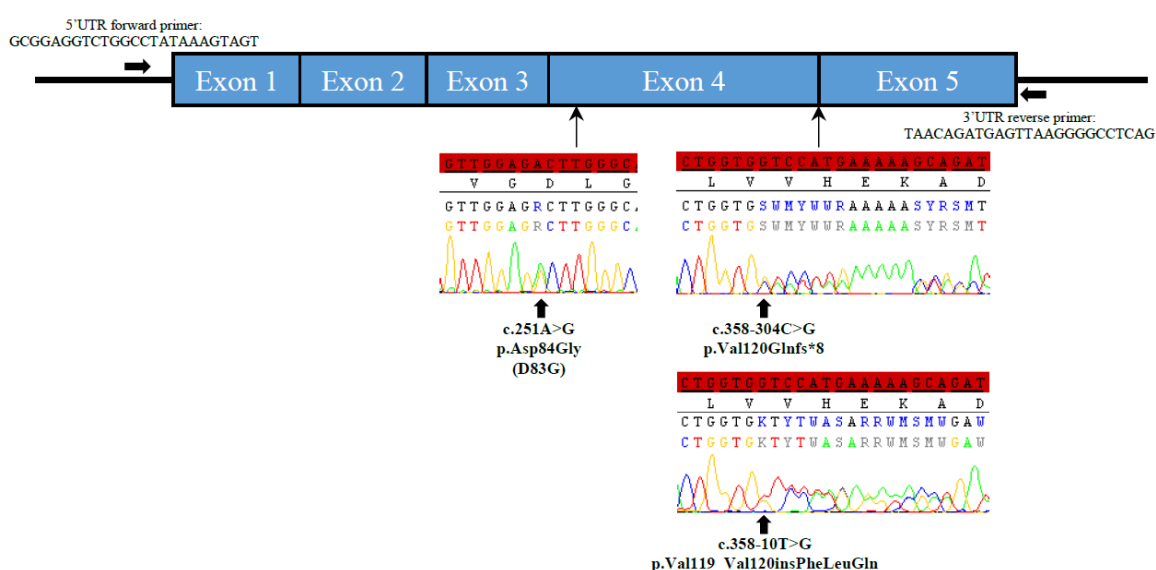
Fig.4. One-step *SOD1* mRNA amplification allows the identification of *SOD1* frameshift variants.

Schematic representation of the position of primers, localized in the 5'UTR and 3'UTR regions, used for the amplification of the *SOD1* cDNA prepared from the mRNA extracted from patient lymphoblasts. Part of chromatograms showed the c.251A>G, p.Asp84Gly (D83G); c.358-304C>G, p.Val120Glnfs*8 and c.358-10T>G, p.Val119_Val120insPheLeuGln variants obtained after sequencing the resulting fragments with the forward primer.

SUPPLEMENTARY MATERIALS

1. Supplemental methods

1.1. Cohorts

DNA samples from fALS index cases (n=470) were collected during 20 years (between 1994-2014) at 10 French ALS National Reference Centers (including Bordeaux, Clermont-Ferrand, Lille, Limoges, Lyon, Marseille, Nancy, Paris, St-Etienne and Strasbourg). DNA of some relatives of these families have then been gathered for segregation analysis. Control samples were obtained from 525 healthy Caucasian individuals matched to the sex and age of the patients. All patients signed a consent form for the genetic research and protocols were approved by the Medical Research Ethics Committee of "Assistance Publique Hôpitaux de Paris" (#A75). Autopsied patients were enrolled in the NeuroCEB brain donation program declared to the Ministry of Research and Universities, as requested by French Authorities (#AC-2013-1887). An explicit consent was signed by the patient himself, or by the next of kin, in the name of the patient, in accordance with the French Bioethical Laws. This document includes consent for genetic analyses. No recognizable individual data are presented in this study and, to ensure de-identification of the families, only abbreviated pedigrees are presented with diamonds for some individuals/branch hiding the corresponding gender.

1.2. Genetic analyses

1.2.1. Sanger sequencing

All *SOD1* variants are referred according to NM_000454.4 transcript. In addition to the standard HGVS nomenclature, all *SOD1* variants are also indicated according to the alternative nomenclature commonly found in the literature reported in the ALS Online Database lacking the methionine start codon. *SOD1* variant screening or confirmation was performed by Sanger sequencing as described.[12] Screening of the *SOD1* rs1555836889 in intron 4[25] was performed using the following primers: 5'-TCTTGAGCCATAGTTGTAAAGCAGAAAAATC-3' and 5'-AAACATTAGGTAAGTGTTGCTGTCTAAATGAAATA-3' in all patients. All these patients were also systematically screened for *C9orf72* expansions, *TARDBP*, *FUS*, and *UBQLN2* variants.[13 34]

1.2.2. Whole exome sequencing analysis

Whole exome sequencing analysis was performed as previously described[35] for all patient with the SOD1 nearsplice variant and available DNA. These exome databases were interrogated for

37 ALS related genes (including *ALS2, ANG, ANXA11, CCNF, CHCHD10, CHMP2B, DAO, DCTN1, DNAJC7, ERLIN2, FIG4, FUS, GLE1, GLT8D1, HNRNPA1, HNRNPA2B1, KIF5A, MAPT, MATR3, NEK1, OPTN, PFN1, PRPH, SETX, SIGMAR1, SOD1, SPG11, SQSTM1, SS18L1, TAF15, TARDBP, TBK1, TIA1, TUBA4A, UBQLN2, VAPB* and *VCP*) to select variants with a minor allele frequency (MAF) <0.005% in dbSNP, Hapmap, 1000genome, Exome Variant Server and gnomAD databases.

1.2.3. Linkage analysis

Linkage analysis was performed using MERLIN (Multipoint Engine for Rapid Likelihood Inference) interface (version 1.1.2).[36] We assumed an autosomal dominant model for the analysis. The phenocopy rate was set at 0.003, the penetrance at 0.7, the frequency of the deleterious allele at 0.0001 and the marker allele frequency at 0.001.

1.3. mRNA extraction and splicing analysis

The splicing prediction score was provided by Alamut Visual 2.15 (Interactive Biosoftware). These predictions are compiled from SpliceSiteFinder-like (SSF-like) based on Shapiro-Senapathy algorithm,[37] GeneSplicer Web interface,[38] MaxEntScan algorithm,[39] Human Splicing Finder system[40] and NNSPLICE Web interface.[41]

Total RNAs were extracted from patient lymphoblasts and post-mortem tissue (whole transverse sections of spinal cord) using the Trizol reagent (ThermoFisher). First strand cDNA synthesis was performed using SuperScript IV First-Strand system (ThermoFisher) according to the manufacturer's instructions.

The exon 4 and 5 junction of *SOD1* cDNA sequence was amplified using the following primers: 5'-TGTGACTGCTGACAAAGATGGTG-3' (exon 4) and 5'-CCAGCGTTTCCTGTCTTTGTACT-3' (exon 5) with annealing temperature of 56°C and sequenced by Sanger method. Amplification of the complete cDNA *SOD1* transcript was performed using the following primers: 5'-GCGGAGGTCTGGCCTATAAAGTAGT-3' (5'UTR) and 5'-TAACAGATGAGTTAAGGGGCCTCAG-3' (3'UTR) with annealing temperature of 56°C and sequenced using forward primer. To determine the length of amplified fragments, a Genescan fluorescent fragment length analysis was performed using forward primer in *SOD1* exon 4 labelled with FAM (Fluorescein AMidites) and reverse primer in *SOD1* exon 5. Amplified fluorescent fragments were separated on an ABI 3730 XL sequencer after

adding GeneScan500LIZ dye Size standard and HIDI formamide (Applied Biosystems). Resulting data were analyzed with Peak Scanner Software (v1.0).

1.4. Prediction for the protein secondary structure

Secondary structures were provided by the JPred4 Protein Secondary Structure Prediction Server (http://www.compbio.dundee.ac.uk/jpred/), performed on a multiple alignment of 15 SOD1 orthologues (including NP_000445.1 *Homo sapiens*, XP_004062735.1 *Gorilla gorilla gorilla*, NP_001027976.1 *Macaca mulatta*, NP_001272335.1 *Macaca fascicularis*, NP_001009025.1 *Pan troglodytes*, NP_777040.1 *Bos taurus*, NP_035564.1 *Mus musculus*, NP_058746.1 *Rattus norvegicus*, NP_001075295.1 *Equus caballus*, NP_001003035.1 *Canis lupus familiaris*, NP_001177351.1 *Sus scrofa*, XP_003467296.1 *Cavia porcellus*, NP_990395.1 *Gallus gallus*, NP_001016252.1 *Xenopus tropicalis*, NP_571369.1 *Danio rerio*) using MultAlin interface.[42]

1.5. Neuropathological examination

The brain and spinal cord were sampled following the usual protocol for analysis of cases of ALS. Samples were taken from orbitofrontal and midfrontal gyrus, motor cortex, temporal, parietal and occipital cortices, hippocampus, amygdala, striatum, thalamus, mesencephalon, pons, medulla, cerebellum and three levels of spinal cord (cervical, thoracic and lumbar), fixed in formaldehyde and sampled for neuropathological examination. After inclusion in paraffin, the slides were cut to 3 µm thick and stained with hematoxylin–eosin and Luxol fast blue. Immunohistochemical detections were performed on motor cortex, frontal, temporal, parietal and occipital cortices, hippocampus, mesencephalon, cerebellum, medulla and spinal cord. Three sections were stained for each region. Immunostainings against SOD1 (#07-403-I, Merck Millipore, 1:1000) were performed after deparaffinization of 2 µm thick sections by an automatic slide stainer (Benchmark XT Ventana staining system). Slides were pre-treated at 95°C in CC1 (pH 8) proprietary retrieval buffer (Ventana Medical Systems). Antibodies were targeted with a biotin-free detection system (Ventana Medical Systems ultraView universal DAB Detection Kit). Additional antibodies used included rabbit anti-TDP-43 (10782-2-AP, ProteinTech Group), Phosphorylated TDP-43 pS409/410-2 (TIP-PTD-P02, Cosmo Bio Co, LTD), Cystatin C (HPA013143, Sigma Aldrich), mouse anti-p62 (610833, BD Transduction Laboratories), Neurofilament H (NF-H), non-phosphorylated (clone SMI32, 801703, Biolegend) and Neurofilament H & M (NF-H/NF-M), phosphorylated (SMI 310, 837701, Biolegend). All slides were examined blinded to the genotype.

1.6. Protein level analysis in patient spinal cord

Procedure for detection of multimerized SOD1 was adapted from 20. Transversal sections of spinal cords were successively homogenized with Polytron PT1200E (Kinematica) to extract the soluble fraction then, after centrifugation, to extract the insoluble fraction, in a buffer containing 50 mM HEPES, 0.1mM EDTA, cOmplete Mini EDTA-free protease inhibitor cocktail and PhosStop phosphatase inhibitors with 1% NP40 at room temperature (for soluble fraction) or 2% Sodium dodecyl sulfate (SDS), 0.1M iodoacetamide and one hour incubation at 37°C (for insoluble fraction). All reagents mentioned above are from Sigma-Aldrich. Both soluble and insoluble tissue extracts were centrifuged at 13,000 rpm for 10 min and protein concentration of the corresponding supernatants was estimated by the bicinchoninic acid assay (Sigma-Aldrich). Soluble and insoluble proteins (20 µg) were separated on NuPAGE 4%-12% Bis-Tris Gel in reducing buffer (for soluble) or non reducing buffer (for insoluble) conditions (Life Technologies). After electrophoresis, gel with insoluble protein fraction was shaken for 15 min in 25 ml of the SDS/PAGE running buffer with 5 mM Tris(2-carboxyethyl) phosphine (Sigma-Aldrich) to reduce disulfide bonds of SOD1 proteins in the gel and enhance the detection of SOD1 band intensities. Proteins were electrophoretically transferred to nitrocellulose (for soluble) or poly(vinylidene difluoride), PVDF, (for insoluble extract) membranes. Membranes were incubated overnight with rabbit anti-SOD1 (#07-403-I MerckMillipore) and anti-GAPDH (#D16H11, Cell Signaling Technology) antibodies in phosphate-buffered saline (dilution: 1/5000), 5% milk, and 0.1% Tween 20, followed by one hour of incubation with peroxidase-conjugated goat anti-rabbit with minimal cross-reaction to human serum proteins (the Jackson Immunoresearch Laboratories). Signals were detected using ECL Prime Western Blotting Detection Reagent (GE Healthcare SA) coupled with Clarity Max Western ECL Substrate (Biorad) to enhance the signal in insoluble conditions. Signal intensity was analyzed with ImageJ 1.53 software.

2. Supplemental Table

Pedigrees	Patients	SOD1 variants	Variants in other ALS genes	rs number in dbSNP	MAF in gnomAD (%)	MAF in Project MinE (%)
FALS1	III8	c.358-10T>G	No			
	IV1	c.358-10T>G	SPG11, c.4274A>G, p.Asn1425Ser	rs142567993	ALL: 0.0053 NFE: 0.011	Cases: 0.0115 Controls: 0.0273
			TAF15, c.1692_1715del, p.Gly565_Gly572del	rs537726014	ALL: 0.026 NFE: 0.026	-
	IV14	c.358-10T>G	ALS2, c.4261C>T, p.Arg1421Ter	rs863225293	ALL: 0.0004 NFE: 0.0009	-
	IV20	c.358-10T>G	No			
	V2	c.358-10T>G	No			
	V3	c.358-10T>G	No			
	VII1	c.358-10T>G	No			
FALS2	III1	c.358-10T>G	SETX, c.3347A>G, p.Asn1116Ser	rs148550755	ALL: 0.0004	-
FALS3	II1	c.358-10T>G	No			
FALS4	II1	c.358-10T>G	DCTN1, c.3029+19G>A	rs201665169	ALL: 0.011 NFE: 0.021	-

Table S1. Variants identified in ALS genes after Whole Exome Sequencing (WES) analysis

All variants identified are heterozygous. Four software packages (SIFT, Polyphen2, CADD Phred score, Mutation Taster) investigating in silico pathogenicity have been used to classify variants identified in other ALS genes after whole exome sequencing: none of them predicted a pathogenic effect of these variants. Minor allele frequency (MAF) in gnomAD control populations and dbSNP identification number are indicated. MAF for ALS patients in Project MinE is specified. All identified variants are absent of ALSdb (ALS Data Browser) database. dbSNP, Single Nucleotide Polymorphism Database; MAF, minor allele frequency; NFE, Non-Finnish European; -: absent.

3. Supplemental figures
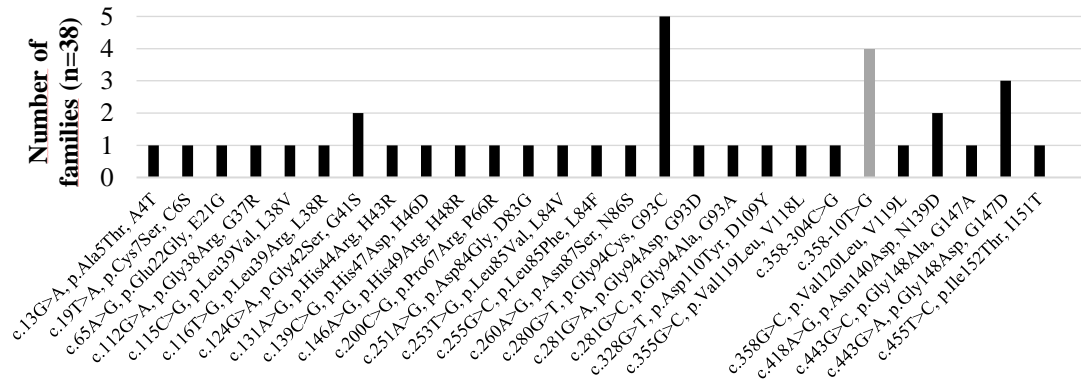
**Fig.S1. The SOD1 c.358-10T>G variant was frequent in French FALS.**

Distribution of the 27 SOD1 variants identified among the French FALS (n=38): the c.358-10T>G variant (grey) was identified in 4 different French pedigrees presented in Fig.1.
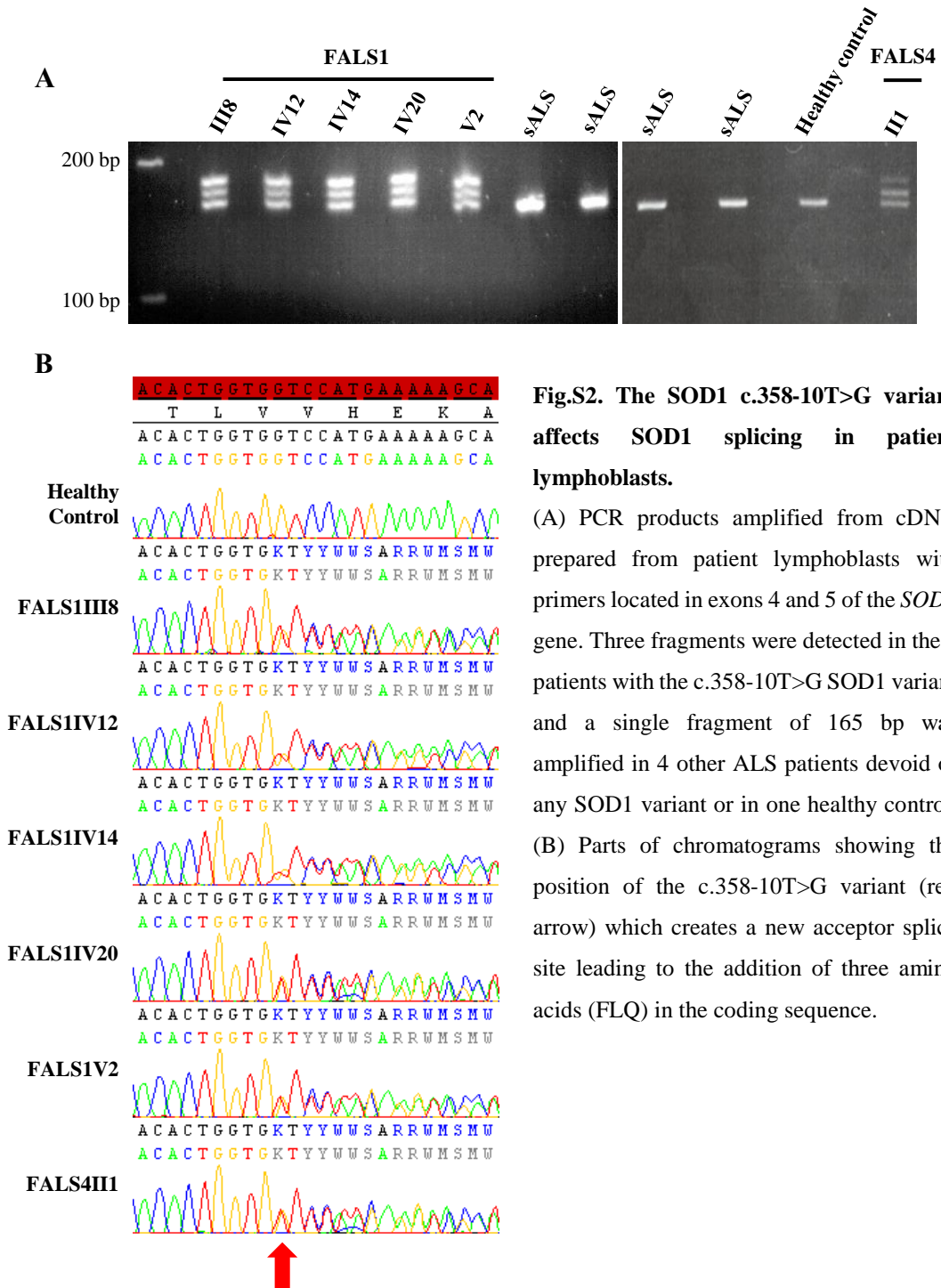
Fig.S2. The SOD1 c.358-10T>G variant affects SOD1 splicing in patient lymphoblasts.

(A) PCR products amplified from cDNA prepared from patient lymphoblasts with primers located in exons 4 and 5 of the *SOD1* gene. Three fragments were detected in the 6 patients with the c.358-10T>G SOD1 variant and a single fragment of 165 bp was amplified in 4 other ALS patients devoid of any SOD1 variant or in one healthy control. (B) Parts of chromatograms showing the position of the c.358-10T>G variant (red arrow) which creates a new acceptor splice site leading to the addition of three amino acids (FLQ) in the coding sequence.
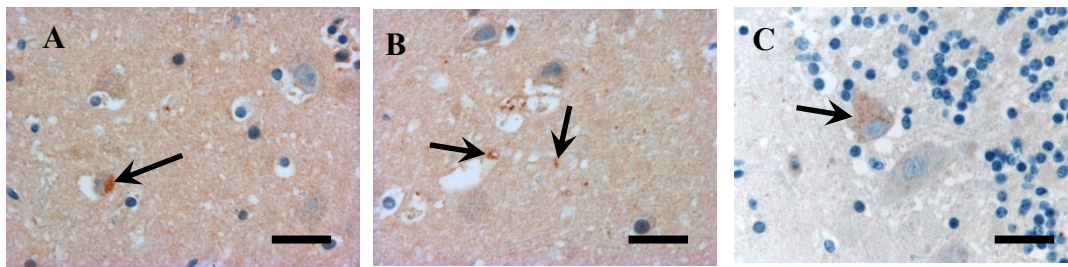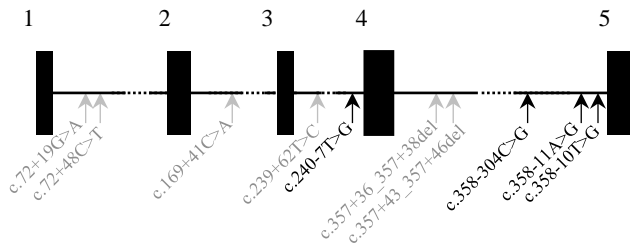
Fig.S3. Immunodetection of SOD1 in brain regions of the FALS2III1 patient with the SOD1 c.358-10T>G variant.

Rare diffuse SOD1 positivities were observed in neurons (A, arrow) and neuropils (B, arrows) in the frontal associative cortex and in some Purkinje cells of the cerebellum (C, arrow). Scale bar: 10 µm.

A

Exons

B

| | | | Muratet et al INNP 2021 | |
cDNA position	rs number in dbSNP	MAF in gnomAD (%)	Splicing prediction	References
c.72+19G>A	rs201574089	0.048	-	43
c.72+48C>T	rs540569412	0.0085	-	44
c.169+41C>A	No	Absent	-	45
c.239+62T>C	rs1568810368	Absent	-	10
c.240-7T>G	rs1568810602	Absent	+	19
c.357+36_357+38del	rs772727110	0.012	-	46
c.357+43_357+46del	No	Absent	-	47
c.358-304C>G	rs1555836889	Absent	-	25
c.358-11A>G	rs369600566	0.0008	+	26
c.358-10T>G	rs1197141604	0.0008	+	23

C

33
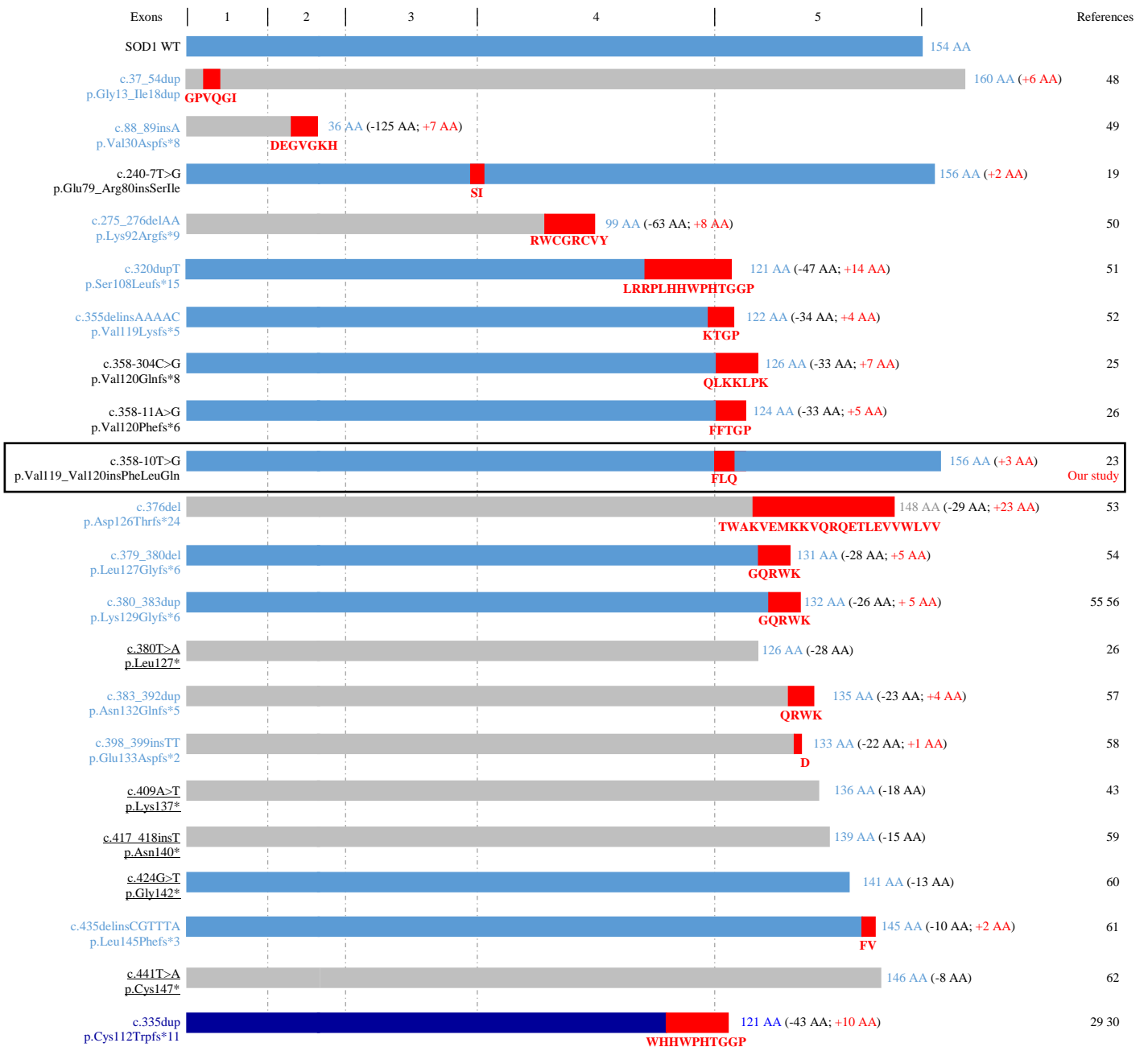
Fig.S4. SOD1 amino acid sequence length is altered by nearsplice, intronic, frameshift and nonsense variants.

(A) Schematic representation of the SOD1 gene with the position of rare nearsplice/intronic variants reported in the literature. Impact on mRNA and/or protein sequences have been experimentally studied for those in black. (B) SOD1 nearsplice/intronic variants reported in ALS patients. Single Nucleotide Polymorphism Database (dbSNP) identification number and Minor Allele Frequency (MAF) in gnomAD total control populations are indicated. +, predicted to impact splicing using Alamut software; -, no predicted splicing impact. None of these rare variants were reported in ALS data browser (ALSdb) or Project MinE database. SOD1 nearsplice/intronic variants with a MAF>1% were not included in this Table. (C) Representation of WT and mutated SOD1 proteins resulting from nearsplice/intronic (black), frameshift (blue) and nonsense (black underlined) variants indicated on the left. Variant detailed in the present study is boxed. Blue blocks indicate that the SOD1 sequence was validated at the mRNA levels on patient cells. Red parts represent the mutated amino acid sequence modelized by Mutalyzer software (https://mutalyzer.nl/) for each variant. Total amino acid (AA) length is indicated for each SOD1 form in blue, as well as the number of WT amino acids deleted (in black) and mutated amino acids inserted (in red), in brackets. Dark blue mutant protein was reported as responsible of a particular recessive motor neuron syndrome in infants. Data are compiled from publications indicated at the right. For better reading, in-frame insertions and deletions in the SOD1 coding sequence are not represented. Some of the variants included in this figure were renumbered according to HGVS nomenclature. Thus the initial reported numbering has changed from c.18_35 dup18, p.19_24[48] to c.37_54dup, p.Gly13_Ile18dup; c.275_276delAA, p.Lys91ArgfsTer8[50] to c.275_276delAA, p.Lys92Argfs*9; c.320_321insT, p.Ser108LeufsTer15[51] to c.320dupT, p.Ser108LeufsTer15; Gly127ins4TGGG[55] [56] to c.380_383dup, p.Lys129Glyfs*6; c.392_393insGCAAAGGTGG, p.N132Qfs*5[57] to c.383_392dup, p.Asn132Glnfs*5; 132insTT[58] to c.383_392dup, p.Asn132Glnfs*5; G141X[60] to c.424G>T, p.Gly142*; and L144FVX[61] to c.435delinsCGTTTA, p.Leu145Phefs*3.

4. **Supplemental references**

34 Teyssou E, Chartier L, Amador M-D-M, *et al.* Novel UBQLN2 mutations linked to amyotrophic lateral sclerosis and atypical hereditary spastic paraplegia phenotype through defective HSP70-mediated proteolysis. *Neurobiol Aging* 2017;**58**:239.e11-239.e20.

35 Amador M-D-M, Muratet F, Teyssou E, *et al.* Spastic paraplegia due to recessive or dominant mutations in *ERLIN2* can convert to ALS. *Neurol Genet* 2019;**5**:e374.

36 Abecasis GR, Cherny SS, Cookson WO, *et al.* Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;**30**:97–101.

37 Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 1987;**15**:7155–74.

38 Pertea M. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res* 2001;**29**:1185–90.

39 Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol J Comput Mol Cell Biol* 2004;**11**:377–94.

40 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**:e67–e67.

41 Reese MG, Eeckman FH, Kulp D, *et al.* Improved splice site detection in Genie. *J Comput Biol J Comput Mol Cell Biol* 1997;**4**:311–23.

42 Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988;**16**:10881–90.

43 Aguirre T, Matthijs G, Robberecht W, et al. Mutational analysis of the Cu/Zn superoxide dismutase gene in 23 familial and 69 sporadic cases of amyotrophic lateral sclerosis in Belgium. Eur J Hum Genet 1999;7:599–602.

44 Corrado L, D'Alfonso S, Bergamaschi L, et al. SOD1 gene mutations in Italian patients with Sporadic Amyotrophic Lateral Sclerosis (ALS). Neuromuscul Disord 2006;16:800–4.

45 Vildan C, Sule D, Turker B, et al. Genetic alterations of C9orf72, SOD1, TARDBP, FUS, and UBQLN2 genes in patients with Amyotrophic Lateral Sclerosis. Cogent Med 2019;6.

46 Boukaftane Y, Khoris J, Moulard B, et al. Identification of six novel SOD1 gene mutations in familial amyotrophic lateral sclerosis. Can J Neurol Sci J Can Sci Neurol 1998;25:192–6.

47 Conforti FL, Magariello A, Mazzei R, et al. Abnormally high levels of SOD1 mRNA in a patient with amyotrophic lateral sclerosis. Muscle Nerve 2004;29:610–1.

48 Nakamura A, Kuru S, Hineno A, et al. Slowly progressing lower motor neuron disease caused by a novel duplication mutation in exon 1 of the SOD1 gene. Neurobiol Aging 2014;35:2420.e7-2420.e12.

49 Hu J, Chen K, Ni B, et al. A novel SOD1 mutation in amyotrophic lateral sclerosis with a distinct clinical phenotype. Amyotroph Lateral Scler 2012;13:149–54.

50 Tripolszki K, Csányi B, Nagy D, et al. Genetic analysis of the SOD1 and C9ORF72 genes in Hungarian patients with amyotrophic lateral sclerosis. Neurobiol Aging 2017;53:195.e1-195.e5.

51 Canosa A, De Marco G, Lomartire A, et al. A novel p.Ser108LeufsTer15 SOD1 mutation leading to the formation of a premature stop codon in an apparently sporadic ALS patient: insights into the underlying pathomechanisms. Neurobiol Aging 2018;72:189.e11-189.e17.

52 Jackson M, Al-Chalabi A, Enayat ZE, et al. Copper/zinc superoxide dismutase 1 and sporadic amyotrophic lateral sclerosis: Analysis of 155 cases and identification of novel insertion mutation. Ann Neurol 1997;42:803–7.

53 Brown JA, Min J, Staropoli JF, et al. SOD1, ANG, TARDBP and FUS mutations in amyotrophic lateral sclerosis: a United States clinical testing lab experience. Amyotroph Lateral Scler Off Publ World Fed Neurol Res Group Mot Neuron Dis 2012;13:217–22.

54 Pramatarova A, Goto J, Nanba E, et al. A two basepair deletion in the SOD 1 gene causes familial amyotrophic lateral sclerosis. Hum Mol Genet 1994;3:2061–2.

55 Andersen PM, Nilsson P, Keränen ML, et al. Phenotypic heterogeneity in motor neuron disease patients with CuZn-superoxide dismutase mutations in Scandinavia. Brain J Neurol 1997;120 (Pt 10):1723–37.

56 Ekhtiari Bidhendi E, Bergh J, Zetterström P, et al. Mutant superoxide dismutase aggregates from human spinal cord transmit amyotrophic lateral sclerosis. Acta Neuropathol (Berl) 2018;136:939–53.

57 Chen S, Li M, Zhu W, et al. A novel 10-base pair insertion mutation in exon 5 of the SOD1 gene in a Chinese family with amyotrophic lateral sclerosis. Neurobiol Aging 2016;45:212.e1-212.e4.

58 Orrell RW, Habgood JJ, Gardiner I, et al. Clinical and functional investigation of 10 missense mutations and a novel frameshift insertion mutation of the gene for copper-zinc superoxide dismutase in UK families with amyotrophic lateral sclerosis. Neurology 1997;48:746–51.

59 Zhan Z, Shu Y, Zhao Y, et al. A novel nonsense SOD1 mutation (p.Asn140Ter) in a sporadic amyotrophic lateral sclerosis case with rapid progression. Amyotroph Lateral Scler Front Degener 2021;15:1–3.

60 Andersen PM, Sims KB, Xin WW, et al. Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes. Amyotroph Lateral Scler Other Motor Neuron Disord 2003;4:62–73.

61 Kawamata C, Morita M, Shibata N, et al. Familial amyotrophic lateral sclerosis (FALS) with a novel SOD1 gene mutation: a clinicopathological study. Rinsho Shinkeigaku 2007;47:211–6.

62 Wu J, Shen E, Shi D, et al. Identification of a novel Cys146X mutation of SOD1 in familial amyotrophic lateral sclerosis by whole-exome sequencing. Genet Med 2012;14:823–6.