Identification of *TARDBP* Gly298Ser as a Founder Mutation for Amyotrophic Lateral Sclerosis in Southern China

Fanxi Xu

Xuanwu Hospital of Capital Medical University

Sen Huang

First Affiliated Hospital of Sun Yat-sen University

Xu-Ying Li

Xuanwu Hospital of Capital Medical University

Jianing Lin

First Affiliated Hospital of Sun Yat-sen University

Xiuli Feng

National Human Genome Center in Beijing

Shu Xie

National Human Genome Center in Beijing

Zhanjun Wang

Xuanwu Hospital of Capital Medical University

Xian Li

Xuanwu Hospital of Capital Medical University

Junge Zhu

Xuanwu Hospital of Capital Medical University

Hong Lai

Xuanwu Hospital of Capital Medical University

Yanming Xu

West China Hospital of Sichuan University

Xusheng Huang

Chinese PLA General Hospital

Xiaoli Yao

First Affiliated Hospital of Sun Yat-sen University

Chaodong Wang (✉ cdongwang01@126.com)

Xuanwu Hospital of Capital Medical University

Research Article

Keywords: Amyotrophic lateral sclerosis, TARDBP, G298S mutation, founder effect, China

Posted Date: February 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1172380/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease characterized by predominant impairment of upper and lower motor neurons. Over 50 TARDBP mutations have been reported in both familial (FALS) and sporadic ALS (SALS). Some mutations in TARDBP, e.g. A382T and G294V, have genetic founder effects in certain geographic regions. However, such prevalence and founder effect have not been reported in Chinese.

Methods: Whole-exome sequencing (WES) was performed in 16 Chinese FALS patients, followed by Sanger sequencing for the TARDBP p.Gly298Ser mutation (G298S) in 798 SALS patients and 1,325 controls. Haplotype analysis using microsatellites flanking TARDBP was conducted in the G298S-carrying patients and noncarriers. The geographic distribution and phenotypic correlation of the TARDBP mutations reported worldwide were reviewed.

Results: WES detected the TARDBP G298S mutation in 8 FALS patients, and Sanger sequencing found additional 8 SALS cases, but no controls, carrying this mutation. All the 16 cases came from Southern China, and 7 of these patients shared the 117-286-257-145-246-270 allele for the D1S2736-D1S1151-D1S2667-D1S489-D1S434-D1S2697 markers, which was not found in the 92 non-carrier patients (0/92) ($p < 0.0001$). The A382T and G298S mutations were prevalent in Caucasians and Eastern Asians, respectively. Additionally, carriers for the M337V mutation are dominated by bulbar onset with a long survival, whereas those for G298S are dominated by limb onset with a short survival.

Conclusions: Some prevalent TARDBP mutations are distributed in a geographic pattern and related to clinical profiles. TARDBP G298S mutation is a founder mutation in the Southern Chinese ALS population.

Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive, fatal neurodegenerative disease characterized by predominant impairment of upper and lower motor neurons, with an incidence of approximately 1-2 per 100,000 people. Patients usually suffer from progressive muscle weakness and atrophy and die of respiratory failure 3-5 years after the onset [1,2]. The etiology of ALS is not fully understood. Appropriately 90% of patients appear sporadically (sporadic ALS, SALS), while ~10% of patients have a family history positive for ALS (familial ALS, FALS) or frontotemporal dementia (FTD). With the advancement of sequencing technologies, rapid increasing number of genes and mutations have been identified in ALS, suggesting that genetic factors play important roles in its pathogenesis.

To date, variants in 25 genes that fulfill adequate criteria for causation for ALS have been reported [3,4]. Mutations in these genes have been identified in approximately two-thirds of FALS and 10% of SALS cases [5]. Recent studies have identified *SOD1*, *FUS*, *TARDBP* and *C9orf72* as the major ALS-related genes in both European and Asian populations [6]. Mutations in *SOD1* are the most common cause for ALS in the world, and are detected in ~20% of FALS and 3% of sporadic ALS (SALS) [7]. *C9orf72* mutation is a prevalent cause for ALS in Caucasians, but was rarely reported in Eastern Asians [8,9]. *FUS* mutations seem to be the most frequent genetic cause in early-onset sporadic ALS patients [10,11]. A characteristic feature of degenerating neurons in ALS patients is the presence of cytoplasmic insoluble and ubiquitinated inclusions containing abnormal aggregates of TAR DNA-binding protein (TDP-43), encoded by the *TARDBP* gene [12]. However, it is worthwhile to explore that since the first report of TDP-43 positive aggregates in ALS and FTD, they were also reported in other neurodegenerative diseases as a secondary feature [13]. TDP-43 is involved in RNA processing, splicing and transport. It consists of an N-terminal nuclear localization signal followed by two RNA recognition motifs and a C-terminal glycine rich domain. To date, more than 50 *TARDBP* mutations have been reported which explain approximately 4% of FALS cases and a smaller proportion of FTD cases [14]. Of course, this does not mean that all mutations are actually pathogenic mutations. But, all *TARDBP* pathogenic mutations exhibit an autosomal dominant pattern of inheritance [15]. The disease-associated mutations in *TARDBP* are distributed in the C-terminal glycine-rich domain, indicating critical involvement of this region in the pathogenesis of TDP-43 proteinopathy [16,17]. Indeed, emerging evidence supports that TDP-43 may act like a prion to initiate cascades of protein misfolding [18].

Some mutations are in high frequency in certain geographic regions and genetic founder effects have been examined for their associations with ALS. The distribution of ALS causing genes varies widely across populations and may differ widely between two seemingly similar countries. Founder mutations for ALS have been established in Italians (*SOD1*: D124G and G41S) [19-21], North American (*SOD1*: A4V) [22], Polish (*SOD1*: L144S) [23], Brazil (*VAPB*: P56S) [24] and Germany (*SOD1*: R115G) [25]. Hexanucleotide (G₄C₂) repeat expansion (HRE) in *C9orf72* has a single founder and is the most common mutation in familial and sporadic ALS in Europe [26]. While several founder mutations have been reported, by far the most common is the *SOD1* D90A (highly prevalent in Sweden and Finland but are rare in neighboring countries) [27]. In *TARDBP*, the most commonly reported missense mutations are A382T and M337V, and some of the most well-studied mutations are A315T, Q331K, M337V, D169G, G294A/V, Q343R, etc. Moreover, ALS cases carrying the A382T and G295S mutation of *TARDBP* and the *C9orf72* repeat expansion shared distinct haplotypes across these loci in Sardinia [28]. In Chinese ALS cases, H47R, R521 and M337V are the most frequent mutations in *SOD1*, *FUS* and *TARDBP*, respectively [29]. However, there is still no haplotype analysis to support the founder effect of these mutations.

Here, we identified a prevalent heterozygous Gly298Ser mutation (G298S, as follows) in *TARDBP* in both FALS and SALS in Guangdong and Guangxi, two neighboring Southern Chinese provinces. Haplotype analysis with the microsatellites surrounding the gene further confirmed a founder effect of this mutation in Southern, but not Northern Chinese. We also reviewed the global distribution of the mutations in *TARDBP* and revealed geographic distribution patterns and clinical relevance of some prevalent mutations.

Materials And Methods

Subjects

A total of 798 ALS cases were recruited between 2016 to 2021, including 462 from The First Affiliated Hospital of Sun Yet-Sen University, 232 from Xuanwu Hospital of Capital Medical University and the Chinese PLA General Hospital, and 104 from West China Hospital of Sichuan University. Among these, 16 had a family history. All the patients were diagnosed as definite, probable or possible ALS by at least 2 neuromuscular specialists, based on the revised El Escorial criteria (2000) [30]. During the same period, 1,325 ethnicity-matched controls without any neurological diseases were enrolled. To avoid recruiting presymptomatic patients, we tended to recruit older people as healthy controls. Therefore, all healthy controls were older than 55 years.

Ethnics approvals and patient consents

Written informed consent was obtained from all participants, as approved by the Ethics Committee and the Expert Committee of Xuanwu Hospital of Capital Medical University (Clinical Research Audit: [2021] 034) and The China Human Genetic Resource Administration Office (China Ministry of Science and Technology Genetics Audit: [2021] CJ1167).

Screening of C9orf72 (GGGGCC)n repeat

Two-step polymerase chain reaction (PCR) was performed to detect *C9orf72* GGGGCC hexanucleotide repeat expansion (HRE) as previously described [31]. Briefly, fluorescent fragment-length analysis was performed with genotyping primers. The samples with a homozygous peak pattern were analyzed by fluorescent repeat-primed PCR to identify HREs. The HRE was defined as repeat number >30 indicated by the typical “saw-tooth” pattern seen by repeat-primed PCR [32].

Whole-exome sequencing for FALS

As it is easier to find causative mutations in FALS, whole-exome sequencing (WES) was performed in all the 16 FALS cases whose DNA samples were available. The WES was performed as the methods described previously (see details in Supplementary Methods) [33]. We only selected the pathogenic, likely pathogenic or uncertain significance variants according to the 2015 American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) criteria [34]. For all the variants, mutations in 25 known ALS causative genes, including *SOD1*, *ALS2*, *SETX*, *FUS*, *VAPB*, *ANG*, *TARDBP*, *FIG4*, *OPTN*, *VCP*, *UBQLN2*, *KIF5A*, *TIA1*, *ANXA11*, *SIGMAR1*, *CHMP2B*, *PFN1*, *AR*, *DCTN1*, *NEFH*, *PRPH*, *DAO*, *TFG*, *TAF15* and *GRN*, were analyzed. Further Sanger sequencing was used to confirm the mutation.

Sanger sequencing for the TARDBP G298S mutation in the SALS patients and controls

In order to confirm whether the *TARDBP* c.892G>A (p.Gly298Ser) mutation was also present in SALS patients and controls, validation of the mutation was performed in 782 sporadic cases (including 511 Southern and 271 Northern) and 1,325 controls. The specific methods can be found in the Supplementary Material.

Haplotype analysis using the microsatellite markers

To test whether the *TARDBP* G298S mutation carriers share the common ancestry, we performed haplotype analysis for the ALS patients with the mutation using the eight microsatellite markers in 1p36.22 spanning a region of about 7Mb (D1S450, D1S244, D1S2736, D1S1151, D1S2667, D1S434, D1S489, D1S2697) surrounding the *TARDBP* gene, according to Orrù et al [35]. They were genotyped in the 16 ALS individuals carrying the G298S mutation and 92 cases not carrying the mutation, using the fluorescent-labeled primers designed and listed in Supplemental Table 1. The alleles were typed by electrophoresis on an ABI 3730 (Applied Biosystems, Foster City, CA, USA) and analyzed using GeneMarker Demo software V2.6.3 (SoftGenetics, State College, PA, USA). Haplotype frequencies and association statistics for the markers were constructed using PHASE version 2 software [36].

Statistical analysis

The results of all continuous data are presented in this report as mean $\pm$ standard deviation. The difference in haplotype distribution between mutation carriers and noncarriers was evaluated by χ^2 and Fisher's exact tests.

Results

Demographics and clinical phenotype of the patients

A total of 798 ALS patients were included in this study, including 16 FALS (from 15 pedigrees) and 782 SALS. All patients were of Chinese Han origin. Among the 798 cases diagnosed, 482 were males and 316 were females. The mean age of ALS onset was 52.77 ± 10.40 years. For initial symptoms, 615 patients (77.07%) first involved the spinal cord, while 183 (22.93%) presented with a bulbar-onset. The survival time ranged from 6-122 months. In the controls, 749 were males and 576 were females, and the mean age were 69.23 ± 14.62 years.

Mutations identified in the FALS patients by whole-exome sequencing

In total, we detected 134 variants of all types in the 25 selected genes. After sequential screening and filtering by population frequency (max frequency <0.01) and selection of variant types (missense, stop gained, frameshift, in-frame and splicing) and functional /conservation prediction (SIFT, CADD, Polyphen2, GERP++), 7 rare damaging variants were identified in the 16 FALS patients. Among all the rare damaging variants, 5 were pathogenic and 2 were likely pathogenic, according to the ACMG/AMP guidelines. Among the 16 patients, 10 carried mutations in *TARDBP* (G298S, N378D and N345K), 3 carried mutations in *SOD1* (C112Y and G142A), 1 carried a *FUS* mutation (R521C) and 1 carried a *VCP* mutation (R155C), and no mutation was identified in one patient (No.12) (Table 1). The GGGGCC hexanucleotide repeats in *C9orf72* were 2 to 11 copies and not pathogenic in the cases. Strikingly, among all the mutations, G298S

was the most frequently (8/16, 50%) identified, and all of these 8 patients were from 7 pedigrees in Guangdong Province, a region in Southern China (Fig. 1). This is a high frequency mutation that has not been widely reported.

Table 1 Clinical data and mutations identified in the 16 FALS probands

Patient ^a	Gender	Birthplace (provinces)	AAO (y)	Onset site	Duration (months)	Genetic variant	ACMG
1(F1)	M	Guangdong	50	Upper limb	48 (alive)	TARDBP, c.892G>A, p. Gly298Ser	P
2(F2)	F	Guangdong	50	Lower limb	15	TARDBP, c.892G>A, p. Gly298Ser	P
3	F	Guangdong	38	Lower limb	NA	SOD1, c.425G>C, p. Gly142Ala	P
4(F3-1)	F	Guangdong	62	Upper limb	13	TARDBP, c.892G>A, p. Gly298Ser	P
5	F	Sichuan	43	Bulbar	NA	FUS, c.1561C>T, p. Arg521Cys	P
6	F	Guangxi	60	Bulbar	NA	TARDBP, c.1035C>A, p. Asn345Lys	LP
7	M	Guangdong	33	Lower limb	NA	VCP, c.463C>T, p. Arg155Cys	P
8	F	Hunan	50	Upper limb	NA	SOD1, c.335G>A, p. Cys112Tyr	P
9(F6)	F	Guangdong	53	Upper limb	20	TARDBP, c.892G>A, p. Gly298Ser	P
10	M	Guangdong	40	Upper limb	12 (alive)	TARDBP, c.1132A>G, p. Asn378Asp	LP
11(F3-2)	M	Guangdong	54	Bulbar	20	TARDBP, c.892G>A, p. Gly298Ser	P
12	F	Guangdong	66	Upper limb	NA	Not identified	-
13(F4)	F	Guangdong	73	Upper limb	10	TARDBP, c.892G>A, p. Gly298Ser	P
14(F5)	F	Guangdong	49	Bulbar	11	TARDBP, c.892G>A, p. Gly298Ser	P
15(F7)	M	Guangdong	46	Bulbar	14	TARDBP, c.892G>A, p. Gly298Ser	P
16	F	Hunan	32	Upper limb	NA	SOD1, c.335G>A, p. Cys112Tyr	P

NA stands for “not available.” LP, Likely Pathogenic. VUS, Uncertain significance. AAO, Age at onset. Patient^a, Numbers with F in brackets represent the FALS in our G298S carriers.

Carriers for TARDBP G298S mutation in the SALS patients

To further identify the carriers for G298S mutation in *TARDBP*, we screened it by Sanger sequencing in 782 SALS patients (including 511 Southern and 271 Northern Chinese) and 1,325 ethnically matched healthy controls. In all the cases, 8 were found to carry the *TARDBP* G298S mutation. Ultimately, we found 16 patients carried the same G298S mutation, in which 13 were from Guangdong Province and 3 were from its neighboring province—Guangxi. However, the mutation was not detected in the 232 Northern or 104 Western Chinese ALS patients, neither in the 1,325 controls. And no mutation was found in these 16 patients in other ALS-causing gene, such as *SOD1*, *FUS*, *NEK1*, *VAPB*, *ALS2*, *OPTN*, *VCP*, etc. Given the strong geographic preference of the mutation, we hypothesize that G298S might be a founder mutation in the Guangdong-Guangxi region.

The TARDBP G298S mutation was associated with an ancestral haplotype

To test whether all the 16 cases carrying the *TARDBP* G298S mutations were inherited from the same ancestor, we genotyped the cases and 92 sporadic ALS cases (Southern 40, Northern 52) not carrying the mutation, using eight microsatellite markers (D1S450, D1S244, D1S2736, D1S1151, D1S2667, D1S489, D1S434, D1S2697) flanking the *TARDBP* gene (Table 2). The 7 of 16 patients shared the 117-286-257-145-246-270 (bps) allele at markers D1S2736-D1S1151-D1S2667-D1S489-D1S434-D1S2697, which was not detected in the 92 non-carrier cases (0/92) ($p < 0.0001$). The large haplotype spans a 5.8 Mb fragment that includes the *TARDBP* gene. For each marker, the number of alleles observed in the control group and the frequency of the shared alleles were shown in Table 2. In particular, the 286 and 257 bp alleles for the D1S1151 and D1S2667, which were very close (0.3Mb) to the *TARDBP* gene, were shared by 9 mutation-carriers but were not frequent in the non-carriers (10% and 18%). These data suggest that the G298S mutation may come from a common ancestral chromosome.

Table 2 Genotypes of microsatellites flanking TARDBP in the ALS patients carrying the G298S mutation

Marker	D1S450	D1S244	D1S2736	TARDBP	D1S1151	D1S2667	D1S489	D1S434	D1S2697
(Mb) ^a	(9.585)	(10.574)	(10.615)	(11.083)	(11.464)	(11.487)	(12.048)	(12.332)	(16.419)
F1	250 /252	286/ 290	118/124		278/290	262/268	136/144	238/ 246	270 /270
F2	251/257	288/ 290	115/ 117		286 /310	257 /265	139/ 145	246 /248	270 /270
F3-1	248/ 250	298/300	117 /119		260/ 286	257 /259	139/ 145	244/ 246	270 /274
F3-2	248/ 250	288/ 290	117 /119		260/ 286	257 /259	139/ 145	244/ 246	270 /274
F4	250 /254	286/ 290	117 /125		282/290	258/266	136/144	244/ 246	268 /270
F5	248/254	288/ 290	119/121		268/290	261/277	141/147	246 /254	270 /270
F6	248/256	290 /294	119/125		275/293	257 /269	139/ 145	246 /256	270 /270
F7	250 /256	288/ 290	117 /119		278/ 286	257 /259	136/138	246 /248	270 /270
S1	254/256	284/ 290	117 /123		268/ 286	257 /259	143/ 145	246 /248	268/ 270
S2	244/ 250	286/288	117 /119		274/ 286	257 /265	143/ 145	246 /248	270 /270
S3	250 /256	288/ 290	117 /119		286 /292	257 /265	143/ 145	242/ 246	270 /274
S4	248/ 250	286/ 290	117 /119		290/310	257 /271	137/ 145	242/ 246	270 /270
S5	248/ 250	282/ 290	115/ 117		252/ 286	257 /259	139/ 145	246 /248	270 /270
S6	248/ 250	286/ 290	117 /119		290/306	257 /265	138/146	242/ 246	270 /270
S7	244/ 250	286/ 290	118/122		286 /302	257 /269	139/143	246 /256	270 /270
S8	250 /258	298/300	117 /119		282/290	257 /271	137/ 145	238/ 246	268/ 270
Shared allele ^b	250	290	117		286	257	145	246	270
Frequency ^c (n=92)	0.28	0.58	0.61		0.10	0.18	0.33	0.77	0.95

Alleles in the haplotype are presented as length of PCR product in base pairs; a Mb, megabases from chromosome 1p telomere (UCSC Genome Browser on Human March 2006 Assembly); b Allele shared by the largest number of patients with the G298S mutation; c Shared allele frequencies in the control group for each marker. Shared alleles are indicated in bold.

Geographical distribution and clinical features of TARDBP mutations reported worldwide

To date, over 50 mutations of *TARDBP* have been described in both familial and sporadic ALS cases (Supplemental Table 2). A geographical map of the mutations highlights the qualitative distribution patterns (Fig. 2A). Most reported mutations were detected in countries of North America (US, Canada), Europe (France, Italy, Norway, Germany, Denmark) and Asia (India, Japan, Korea and China). Moreover, mutations were mostly detected in the exon 6 of *TARDBP* and dominantly reported in different geographic regions, such as I383V and S375G in the US, G348C and A382T in France, A382T and G295S in Italy, N352S, G357S and N378D in Japan, and M337V and G298S in China. According to the reported cases, most of the mutations, such as M337V, G348C, N352S and I383V, showed a distribution in multiple ethnicities, while some mutations were detected in specific geographic regions. For example, the A382T mutation was the most frequent mutation in Europe and America but not in Eastern Asians, while G298S was only detected in Eastern Asians. Among the 11 mutations reported in Chinese cases, M337V was most frequently reported in Fujian and Taiwan. In contrast, the G298S mutation was common in Guangdong and Guangxi (Fig. 2B, Table 3).

In demographic and clinical characteristics, except for G287S, most of the mutations were detected in both familial and sporadic cases, and most of the cases displayed onset of limb weakness. However, carriers for M337V mainly developed the disease with a bulbar-onset (bulbar/limb: 37/22) and had the longest survival (75.9 months), while the G298S carriers displayed a limb onset and the shortest survival (19.3 months). At onset age, carriers for three mutations had a later onset (>60 years): G287S, G294V and G295S (Table 3).

Table 3 Clinical characteristics of the hotspot mutations in TARDBP

Characteristics	G298S	M337V	A382T	G287S	G294V	G295S	G348C	I383V	N352S
Birth place ^a	East Asia	Admixed	Europe and America	Europe and America	Italy	Italy	Admixed	Admixed	Admixed
Family History ^b (F/S)	10/10	1/15	78/95	0/9	8/8	4/10	5/4	6/2	11/4
Gender (M/F)	13/9	28/30	122/60	2/0	9/6	2/1	9/6	1/2	1/8
Age at onset (mean years, SD)	51.3 (8.3)	51.7 (7.1)	52.7 (12.9)	67.5 (3.5)	61.1 (11.2)	60.0 (5.3)	49.9 (11.3)	58.3 (11.4)	57.1 (12.7)
Onset site	L:15; B:6	L:22; B:37	L:141; B:36	L:1; B:1	L:7; B:8	L:3; B:0	L:13; B:2	L:1; L+B:2	L:7; B:0
Disease duration (mean months)	19.3	75.9	57.3	57.5	19.1	32.0	49.1	36.0	63.1
Cognitive impairment	No	Yes	Yes	No	Yes	Yes	No	Yes	No

Birthplace ^a: admixed: the mutation can be detected in both Asian and European patients

Family History ^b: F, familial ALS; S, sporadic ALS

Onset site: L, limb; B, bulbar

Clinical characteristics of ALS patients with the TARDBP G298S mutation

The 16 FALS and SALS patients carrying the G298S mutation had similar clinical features (Table 4). The onset age varied considerably, cases S3 and S4 showed early onset at 38 and 39 years old, while F4 had onset at 73 years old. Of all the patients, 11 showed limb-onset, while 5 showed bulbar-onset. Most of the patients displayed fasciculation and signs of dysfunction in both upper motor neuron (UMN) and lower motor neuron (LMN). No signs of dementia were reported in these patients. The severity of neurological functions varies: the scores for ALS functional rating scale-revised (ALSF_{RS}-R) ranged 20-48. Electromyography (EMG) examination was performed in 12 patients, all showed fibrillations, fasciculations and positive sharp waves. Except F-1, all the patients had survived no longer than 24 months. In particular, the survival time of S4 and S7 was 7-8 months since onset. Most of the cases did not have histories of drinking, smoking or pesticide exposure. Five patients (F6, F7, S5, S6 and S7) carried another rare variant in the ALS-causing genes (*ALS2*, *SPG11*, *FIG4*, *NEK1*), but all these variants were classified as variant of unknown significance (VUS), according to the ACMG guidelines.

Table 4 The demographic and clinical features of the ALS patients carrying the TARDBP G298S mutation

Patient ^a	F1	F2	F3-1	F3-2	F4	F5	F6	F7	S1	S2	S3	S4	S5	S6	S7	S8
Gender	F	F	F	M	F	F	F	M	F	M	F	F	M	M	M	M
Birthplace ^b	GD	GD	GD	GD	GD	GD	GD	GD	GX	GD	GD	GD	GX	GD	GX	GD
Diagnosis level ^e	Def	Def	Def	Def	Def	Prob	Poss	Def	Def	Def	Prob	Def	Poss	Def	Poss	De
AAO (years)	50	50	62	54	73	49	53	46	43	53	38	39	59	49	52	59
Site of onset ^c	U	L	U	B	U	B	U	B	U	U	U	B	U	U	B	U+
UMN signs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Ye
LMN signs	No	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Ye
Fasciculation	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Ye
ALSFRS-R	48	46	44	NA	40	37	37	45	42	42	NA	NA	NA	46	40	20
Survival (months)	48	15	13	20	10	11	20	14	NA	14	NA	7	NA	24	8	NA
EMG ^d	NA	Fib	Fib	Fib	Fib	Fib	Fib	Fib	Fib	Fib	NA	Fib	NA	Fib	Fib	NA
		Fas	Fas	Fas	Fas	Fas	Fas	Fas	Fas	Fas		Fas		Fas	Fas	
		PSW	PSW	PSW	PSW	PSW	PSW	PSW	PSW	PSW		PSW		PSW	PSW	
Other variants	No	No	No	No	No	No	ALS2 A1550T	SPG11 L1982S	No	No	No	No	FIG4 I220V	NEK1 D1208N	No	No
Smoking	No	No	No	No	No	No	No	No	No	Yes	NA	No	NA	Yes	No	No
Drinking	No	No	No	No	No	No	No	No	No	No	NA	No	NA	No	No	No
Pesticide	No	No	No	No	No	No	No	Yes	No	No	NA	No	NA	Yes	No	No

Patient a: F, familial case, S=sporadic case, F3-1 and F3-2 are two cases in the third family; Birthplace b: GD, Guangdong; GX, Guangxi; Site of onset c: U, upper limb; L, Lower limb; B, Bulbar; EMG d: Fib, Fibrillations, Fas, fasciculations, PSW, positive sharp waves; Diagnosis level: Def, definite; Prob, probable; Poss, possible. NA, not available.

Discussion

In this study, we identified G298S as a frequent mutation that were merely detected in patients from the Southern Chinese provinces, and haplotype analysis revealed that inheritance of the G298S mutation was attributable to a founder effect. We also reviewed the *TARDBP* mutations reported worldwide and found that some prevalent mutations were apparently distributed in geographical prevalence and associated with specific clinical features.

Increasing studies have reported the correlation of mutations in ALS genes with specific haplotypes. For *TARDBP* variants, haplotype analyses for the A382T [19,37], N352S [38] and G294V [39] mutations have been documented. However, these mutations had been mostly reported in Westerners but scarcely reported in Easterners. Before this report, the G298S mutation has only been described in Chinese and Japanese FALS families [40-42]. Our study adds the most cases carrying this mutation and demonstrates that the high prevalence of the G298S mutation in Southern Chinese patients is due to a founder effect. The analysis of microsatellite markers surrounding the *TARDBP* gene in cases carrying the mutation showed that they were inherited from a common ancestor with the D1S2736-D1S1151-D1S2667-D1S489-D1S434-D1S2697 haplotype. Clinically, no patient with the G298S mutation exhibited overt cognitive impairment in our study, although the TDP-43 protein inclusions being originally identified in FTD cases [43]. However, a limitation of our study is that we were unable to perform postmortem studies to investigate this. Interestingly, as reported here and in the Chinese and Japanese FALS cases, most of the patients carrying this mutation displayed a relatively short survival (<25 months). These data suggest that G298S is a mutation with unique genetic basis and clinical relevance. However, due to the rarity of the mutation carriers, it remains unclear where was this founder effect originated and transmitted. Population-based studies that would be taken into account biologic and sociocultural factors would help in the understanding of these questions.

The consequences of *TARDBP* mutations on TDP-43 function and neuro-degeneration remain unclear. Studies have suggested that TDP-43 alters RNA metabolism in the cytoplasm through a toxic gain of function whereas the nuclear depletion of TDP-43 could lead to aberrant RNA metabolism through a loss of function [17]. More than 15 mouse models have been reported in which WT and ALS mutant of human TDP-43 are expressed, but its roles in disease mechanisms is uncertain [44, 45]. In addition, although the mutations have been involved in alteration of the aggregation propensity and cytoplasmic mislocalization in neurons, their pathological effects remain to be established [15]. A recent study using the knock-in mice expressing the M337V and G298S mutations showed that the hemizygous mutations did not influence TDP-43 levels or changed its nuclear localization by 2 years of age, while most homozygous knock-in animals display asymmetric denervation of tibialis anterior muscle at 2.5 years, demonstrating the dose dependent MN toxicity of the TDP-43 mutant protein. Moreover, the TDP-43 knock-in mice displayed varying degrees of denervation, consistent with the incomplete penetrance of *TARDBP*

mutations in ALS patients. These observations suggest that other factors, such as environmental or genetic, may contribute to TDP-43 pathology [46]. However, such factors have not been documented and could be identified via studies using larger cohorts and more advanced sequencing technologies.

Although the genotype-phenotype correlations of *TARDBP* mutations have been reported, the inter-mutational differences in distribution and clinical features remain elusive. We reviewed the literature-reported mutations and revealed a geographic pattern of the recurrent mutations. Interestingly, the M337V, G348C, N352S and I383V mutations were detected in cases all over the world, whereas the A382T was predominantly detected in European (especially Sardinian) [28] and American, but not in Asians. The clinical phenotype encompasses patients with bulbar and limb onset and with short and long disease durations. Although most of the cases with mutations displayed limb onset, a bulbar symptoms-dominant onset and the longest disease duration was reported in the carriers for M337V mutation [47]. In contrast, a limb onset and the shortest survival was observed in our carriers for G298S mutation. These data suggest the distinct geographic distribution and clinical features of different mutations, and the differences in the site of onset in patients with different mutations does not seem to explain the differences in survival.

Conclusion

In conclusion, our study identified G298S as a unique mutation showing founder effect in cases from Southern China. However, the number of cases carrying the mutation were relatively small, and more G298S carriers are needed in the future in order to gain a clearer genotype-phenotype picture of this rare mutation. In addition, we reviewed the correlation of geographic and clinical features with *TARDBP* mutations reported worldwide, highlighting the mutation hotspots in Western and Eastern countries.

Abbreviations

ALS: Amyotrophic lateral sclerosis; FTD: frontotemporal dementia; PCR: polymerase chain reaction; HRE: hexanucleotide repeat expansion; WES: whole-exome sequencing; ACMG/AMP: American College of Medical Genetics and Genomics and the Association for Molecular Pathology; SIFT: Sorting intolerant from tolerant; CADD: Combined annotation-dependent depletion; GERP: Genomic evolutionary rate profiling; UMN: upper motor neuron; LMN: lower motor neuron; ALSFRS-R: ALS functional rating scale-revised; EMG: Electromyography; VUS: variant of unknown significance.

Declarations

Ethics approval and consent to participate

All research participants or their legal representatives signed informed consent forms for participation in clinical and genetic research. This study was approved by the institutional ethics board of Xuanwu Hospital of the Capital Medical University (Clinical Research Audit: [2021]034) and The China Human Genetic Resource Administration Office (China Ministry of Science and Technology Genetics Audit: [2021] CJ1167). In this study, all methods were performed in accordance with the relevant guidelines and regulations in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Availability of data and materials

Detailed mutational and clinical data of all patients are submitted as Table 1 and Table 4. Genotypes of microsatellites flanking *TARDBP* in the ALS patients carrying the G298S mutation are submitted as Table 2. The primer sequences are shown in Additional file: Supplementary Materials. Geography-mutation analyses were based on published reports or databases. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive [48] in National Genomics Data Center [49], China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA001905) that are publicly accessible at <https://ngdc.cncb.ac.cn/gsa-human>.

Competing interests

The authors declare no conflicts of interest.

Funding

CW was supported by grants from the National Natural Science Foundation [82171412], Ministry of Science and Technology [2016YFC1306000], Special Fund from Key laboratory of Neurodegenerative Diseases, Ministry of Education of China [PXM2019_026283_000002]. XY was supported by was supported by grants from National Key Research and Development Program of China [2017YFC0907703], Guangdong Provincial Key Laboratory of Diagnosis and Treatment of Major Neurological Diseases [2020B1212060017], Guangdong Provincial Clinical Research Center for Neurological Diseases [2020B111170002], the Southern China International Cooperation Base for Early Intervention and Functional Rehabilitation of Neurological Diseases [2015B050501003, 2020A0505020004].

Authors' contributions

FX, XY and CW conceived the project and designed the study. FX wrote the original draft. FX, SH, XYL, XF, ZW, XY and CW analysed and annotated the data. FX, SH, XYL, XF, SX, ZW and XL performed methodology. XY and CW acquired funding and supervised the whole project. All authors read and approved the final manuscript.

Acknowledgments

All authors are grateful for the participation and contribution of all patients and their families.

References

1. Brown RH, Al-Chalabi A. Amyotrophic Lateral Sclerosis. *N Engl J Med*. 2017;377:162-172.
2. van Es MA, Hardiman O, Chio A, et al. Amyotrophic lateral sclerosis. *Lancet*. 2017;390:2084-2098.
3. Morgan S, Shoai M, Fratta P, et al. Investigation of next-generation sequencing technologies as a diagnostic tool for amyotrophic lateral sclerosis. *Neurobiol Aging*. 2015;36:1600.e5-1600.e8.
4. Nguyen HP, Van Broeckhoven C, van der Zee J. ALS Genes in the Genomic Era and their Implications for FTD. *Trends Genet*. 2018;34:404-423.
5. Zou ZY, Zhou ZR, Che CH, et al. Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry*. 2017;88:540-549.
6. Renton AE, Chiò A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci*. 2014;17:17–23.
7. Deng HX, Hentati A, Tainer JA, et al. Amyotrophic lateral sclerosis and structural defects in Cu, Zn superoxide dismutase. *Science*. 1993;261:1047-1051.
8. Jiao B, Guo JF, Wang YQ, et al. C9orf72 mutation is rare in Alzheimer's disease, Parkinson's disease, and essential tremor in China. *Front Cell Neurosci*. 2013;7, 164.
9. Kim EJ, Kwon JC, Park KH, et al. Clinical and genetic analysis of MAPT, GRN, and C9orf72 genes in Korean patients with frontotemporal dementia. *Neurobiol Aging*. 2014;35:1213.e13-1213.e17.
10. Zou ZY, Liu MS, Li XG, et al. Mutations in FUS are the most frequent genetic cause in juvenile sporadic ALS patients of Chinese origin. *Amyotroph Lateral Scler Frontotemporal Degener*. 2016;17:249-252.
11. Liu ZJ, Lin HX, Liu GL, et al. The investigation of genetic and clinical features in Chinese patients with juvenile amyotrophic lateral sclerosis. *Clin Genet*. 2017;92:267-273.
12. Arai T, Hasegawa M, Akiyama H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun*. 2006;351:602–611.
13. Liachko N F, Saxton A D, McMillan P J, et al. The phosphatase calcineurin regulates pathological TDP-43 phosphorylation. *Acta Neuropathol*. 2016;132:545-561.
14. Sreedharan J, Blair IP, Tripathi VB, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. 2008;319:1668-1672.
15. Rutherford NJ, Zhang YJ, Baker M, et al. Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet*. 2008;4:e1000193.
16. Pesiridis GS, Lee VM, Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009;18:R156–162;
17. Lee EB, Lee VM, Trojanowski JQ. Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nat Rev Neurosci*. 2011;13:38–50.
18. McAlary L, Plotkin S S, Yerbury J J, et al. Prion-like propagation of protein misfolding and aggregation in amyotrophic lateral sclerosis. *Front Mol Neurosci*. 2019;12:262.
19. Chiò A, Borghero G, Pugliatti M, et al. Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. *Arch Neurol*. 2011;68:594-598.
20. Ricci C, Giannini F, Intini E, et al. Genotype-phenotype correlation and evidence for a common ancestor in two Italian ALS patients with the D124G SOD1 mutation. *Amyotroph Lateral Scler Frontotemporal Degener*. 2019;20:611-614.
21. Battistini S, Ricci C, Giannini F, et al. G41S SOD1 mutation: A common ancestor for six ALS Italian families with an aggressive phenotype. *Amyotroph Lateral Scler*. 2010;11:210-215.
22. Saeed M, Yang Y, Deng HX, et al. Age and founder effect of SOD1 A4V mutation causing ALS. *Neurology*. 2009;72:1634-1639.
23. Kuźma-Kozakiewicz M, Andersen PM, Elahi E, et al. Putative founder effect in the Polish, Iranian and United States populations for the L144S SOD1 mutation associated with slowly uniform phenotype of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener*. 2021;22:80-85.
24. Nishimura AL, Al-Chalabi A, Zatz M. A common founder for amyotrophic lateral sclerosis type 8 (ALS8) in the Brazilian population. *Hum Genet*. 2005;118:499-500.
25. Niemann S, Joos H, Meyer T, et al. Familial ALS in Germany: origin of the R115G SOD1 mutation by a founder effect. *Journal of neurology, neurosurgery, and psychiatry*. 2004;75:1186-1188.
26. 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;21:102-108.
27. Andersen P M. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr Neurol Neurosci Rep*. 2006;6:37-46.
28. Borghero G, Pugliatti M, Marrosu F, et al. Genetic architecture of ALS in Sardinia. *Neurobiol Aging*. 2014;35:2882.e7-2882.e12.
29. Liu ZJ, Lin HX, Wei Q, et al. Genetic Spectrum and Variability in Chinese Patients with Amyotrophic Lateral Sclerosis. *Aging Dis*. 2019;10:1199-1206.
30. Deng J, Wu W, Xie Z, et al. Novel and Recurrent Mutations in a Cohort of Chinese Patients with Young-Onset Amyotrophic Lateral Sclerosis. *Front Neurosci*. 2019;13:1289.

31. 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.
32. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011;72: 257-268.
33. Li XY, Cui Y, Jing D, et al. Novel PSEN1 and PSEN2 Mutations Identified in Sporadic Early-onset Alzheimer Disease and Posterior Cortical Atrophy. *Alzheimer Dis Assoc Disord*. 2021;35:208-213.
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405-424.
35. Orrù S, Manolakos E, Orrù N, et al. High frequency of the TARDBP p.Ala382Thr mutation in Sardinian patients with amyotrophic lateral sclerosis. *Clin Genet*. 2012;81:172-178.
36. Stephens M and Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*. 2003;73:1162-1169.
37. Orrù S, Manolakos E, Orrù N, et al. High frequency of the TARDBP p.Ala382Thr mutation in Sardinian patients with amyotrophic lateral sclerosis. *Clin Genet*. 2012;81:172-178.
38. van Blitterswijk M, van Es MA, Hennekam EA, et al. Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum Mol Genet*. 2012;21:3776-3784.
39. Corrado L, Pensato V, Croce R, et al. The first case of the TARDBP p.G294V mutation in a homozygous state: is a single pathogenic allele sufficient to cause ALS? *Amyotroph Lateral Scler Frontotemporal Degener*. 2020;21:273-279.
40. Lin J, Chen W, Huang P, et al. The distinct manifestation of young-onset amyotrophic lateral sclerosis in China. *Amyotroph Lateral Scler Frontotemporal Degener*. 2021;22:30-37.
41. Van Deerlin VM, Leverenz JB, Bekris LM, et al. TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol*. 2008;7:409-416.
42. Nozaki I, Arai M, Takahashi K, et al. Familial ALS with G298S mutation in TARDBP: a comparison of CSF tau protein levels with those in sporadic ALS. *Intern Med*. 2010;49:1209-1212.
43. Synofzik M, Born C, Rominger A, et al. Targeted high-throughput sequencing identifies a TARDBP mutation as a cause of early-onset FTD without motor neuron disease. *Neurobiol Aging*. 2014;35:1212.e1-1212.e5.
44. McGoldrick P, Joyce PI, Fisher EM, et al. Rodent models of amyotrophic lateral sclerosis. *Biochim Biophys Acta*. 2013;1832:1421–1436.
45. Xu YF, Gendron TF, Zhang YJ, et al. Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J Neurosci*. 2010;30:10851–10859.
46. Ebstein SY, Yagudayeva I, Shneider NA. Mutant TDP-43 Causes Early-Stage Dose- Dependent Motor Neuron Degeneration in a TARDBP Knockin Mouse Model of ALS. *Cell Rep*. 2019;26:364-373.e4.
47. Liu W, Li X, Sun Y, et al. Genotype-phenotype correlations in a chinese population with familial amyotrophic lateral sclerosis. *Neurol Res*. 2021;1-11.
48. Chen T, Chen X, Zhang S, et al. The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. *Genomics Proteomics Bioinformatics*. 2021; S1672-0229: 00163-00167.
49. CNGB-NGDC Members and Partners: Xue Y, Bao Y, Zhang Z, et al. Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. *Nucleic Acids Res* 2022;50: D27-D38.

Figures

Figure 1

Pedigree map of the TARDBP G298S families. Pedigrees showing 7 families detected to carry the TARDBP G298S mutations. Black symbols represent patients affected with ALS, white symbols represent unaffected individuals. Arrowheads indicate the probands.

Figure 2

Distribution of TARDDBP mutations reported in the world and China. (A) Mutations reported in the world. Each mutation was represented for one number. The three prevalent mutations were denoted in dots with different colors: M337V (orange), A382T (yellow) and G298S (purple), and other mutations were denoted in blue dots. The amplified gray square regions represent West Europe which was enlarged on upper right. Five pie charts show the countries in which mutations were more frequently detected. Maps were obtained through My Maps (<https://www.google.com/mymaps>, Accessed in 05.30.2021). (B) Map of China with reported TARDDBP mutations. Different mutations were denoted with different colors. The number on the dot represents the number of reported cases.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterials.docx](#)