

**Background:** The SQSTM1 gene encodes p62, a major pathologic protein involved in neurodegeneration.

**Objective:** To examine whether SQSTM1 mutations contribute to familial and sporadic amyotrophic lateral sclerosis (ALS).

**Design:** Case-control study.

**Setting:** Academic research.

**Patients:** A cohort of 546 patients with familial (n=340) or sporadic (n=206) ALS seen at a major academic referral center were screened for SQSTM1 mutations.

**Main Outcome Measures:** We evaluated the distribution of missense, deletion, silent, and intronic variants in SQSTM1 among our cohort of patients with ALS.

**Results:** We identified 10 novel SQSTM1 mutations (9 heterozygous missense and 1 deletion) in 15 patients (6 with familial ALS and 9 with sporadic ALS). Predictive in silico analysis classified 8 of 9 missense variants as pathogenic.

**Conclusions:** Using candidate gene identification based on prior biological knowledge and the functional prediction of rare variants, we identified several novel SQSTM1 mutations in patients with ALS. Our findings provide evidence of a direct genetic role for p62 in ALS pathogenesis and suggest that regulation of protein degradation pathways may represent an important therapeutic target in motor neuron degeneration.

In silico analysis of variants was performed to predict alterations in p62 structure and function.

**Arch Neurol. 2011;68(11):1440-1446**
hances aggregate formation, and this effect is significantly diminished when the ubiquitin-associated (UBA) domain (UBA) of p62 is deleted. Mutant SOD1 can be recognized by p62 in a ubiquitin-independent fashion and targeted for autophagy.19 Furthermore, p62 colocalizes with FUS and TDP-43 in brains of patients having frontotemporal lobe degeneration with motor neuron disease.20 It was recently shown that p62 colocalizes with FUS and TDP-43 in ubiquitinated inclusions among motor neurons in spinal cords from patients with SALS, non-SOD1 FALS, and ALS with dementia.21 In addition, overexpression of p62 reduces TDP-43 aggregation in an autophagy- and pro tease-dependent manner.22 Also, p62 knockout mice develop memory loss after neurodegeneration caused by accumulation of hyperphosphorylated tau and neurofibrillary tangles.23 Although some proteins may participate in pathogenic aggregates in a wide variety of neurodegenerative disorders, recent investigations indicate that they cause very specific disease phenotypes when mutant. Choosing candidate genes based on prior biological knowledge can identify causative genes for neurodegenerative disorders, including ALS. If present in ALS, SQSTM1 mutations may have an increased aggregation potential. To investigate the role of p62 in ALS, we screened a large cohort of patients with ALS for SQSTM1 mutations.

Table 1. SQSTM1 Variants in Patients With Familial Amyotrophic Lateral Sclerosis (FALS) and in Patients With Sporadic Amyotrophic Lateral Sclerosis (SALS)

<table>
<thead>
<tr>
<th>Exon</th>
<th>Change, bp</th>
<th>Variant</th>
<th>Patients With FALS</th>
<th>Patients With SALS</th>
<th>Patients With SALS or FALS</th>
<th>Control Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c.98G&gt;T</td>
<td>A33V</td>
<td>1/340</td>
<td>2/206</td>
<td>3/546</td>
<td>0/724</td>
</tr>
<tr>
<td>2</td>
<td>g.3+7G&gt;C</td>
<td>Intrinsic</td>
<td>1/340</td>
<td>0/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>3</td>
<td>c.457G&gt;A</td>
<td>V153I</td>
<td>0/340</td>
<td>2/206</td>
<td>2/546</td>
<td>0/724</td>
</tr>
<tr>
<td>4</td>
<td>g.5-37G&gt;T</td>
<td>Intrinsic</td>
<td>1/340</td>
<td>2/206</td>
<td>3/546</td>
<td>0/724</td>
</tr>
<tr>
<td>5</td>
<td>c.83R&gt;T</td>
<td>P228L</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>6</td>
<td>c.702G&gt;A</td>
<td>V234V</td>
<td>1/340</td>
<td>0/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>7</td>
<td>c.714-716delGAA</td>
<td>K238del</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>8</td>
<td>c.783C&gt;T</td>
<td>H261H</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>9</td>
<td>c.952T&gt;C</td>
<td>S318P</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>10</td>
<td>c.961C&gt;T</td>
<td>R321C</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>11</td>
<td>c.1108T&gt;C</td>
<td>S370P</td>
<td>1/340</td>
<td>0/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>12</td>
<td>c.1175C&gt;T</td>
<td>P392L</td>
<td>2/340</td>
<td>1/206</td>
<td>3/546</td>
<td>0/724</td>
</tr>
<tr>
<td>13</td>
<td>c.1231G&gt;A</td>
<td>G411S</td>
<td>1/340</td>
<td>0/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
<tr>
<td>14</td>
<td>c.1273G&gt;A</td>
<td>G429R</td>
<td>0/340</td>
<td>1/206</td>
<td>1/546</td>
<td>0/724</td>
</tr>
</tbody>
</table>

Abbreviation: bp, base pair.

Genomic DNA was extracted from transformed lymphoblastoid cell lines or whole blood using standard protocols (Qiagen, Valencia, California). Intronic primers covering the coding sequence were designed at least 50 base pairs away from the intron and exon boundaries. Primers were designed using several products (Oligo Analyzer [IDT, Coralville, Iowa], ExonPrimer [Institute of Human Genetics, Cologne, Germany], and UCSC Genome Browser [Center for Biomolecular Science & Engineering, Santa Cruz, California]). Genomic DNA was amplified according to standard protocols. Unconsumed deoxyribonucleotide triphosphates and primers were digested (ExoSAP-IT; USB, Cleveland, Ohio). Fluorescent dye-labeled single-strand DNA was amplified using sequencing reagents (GenomeLab DTCS Quick Start Kit; Beckman Coulter, Fullerton, California), followed by single-pass bidirectional sequencing (CEQ 8000 Genetic Analysis System, Beckman Coulter). Forward primer was used for mutation screening, and all variations were confirmed by reverse sequencing. When a variant was identified, it was first excluded in the dbSNP Short Genetic Variations (http://www.ncbi.nlm.nih.gov/snp) and 1000 Genomes Project databases, and then more than 724 normal control DNA samples (Table 1) were analyzed to exclude the possibility of a common polymorphism.

Bioinformatics

An Internet server (NetPhos 2.0 [http://www.cbs.dtu.dk/services/NetPhos/]) was used to predict changes in phosphorylation sites among variants identified during sequencing. Variants were also analyzed using programs (SIFT27 and PMUT28) to predict the effect of the mutations on p62. Three-dimensional modeling was performed (Swiss-PdbViewer29) using 1Q02A backbone as a template.

Methods

Participants

A cohort of 340 patients with FALS, 206 patients with SALS, and 738 neurologically healthy control subjects was ascertained from our neurological diseases registry, in which participants are enrolled after informed consent is obtained. Pedigrees and clinical data were collected according to protocols approved by our institutional review board and met Health Insurance Portability and Accountability Act standards of confidentiality and disclosure. All the patients were diagnosed by board-certified neurologists and met the revised El Escorial World Federation of Neurology criteria for diagnosis of clinically definite, probable, or laboratory-supported probable ALS.24 All the ALS cases were negative for mutations in the SOD1, TARDBP, and FUS genes. By self-reported race/ethnicity, 93.9% of the patients with ALS were white (European American), 2.5% Asian, 1.9% African American, and 1.3% Latino. The race/ethnicity of 2 patients was unknown. By self-report, 97.6% of controls were white, 1.0% Latino, 0.8% Asian, and 0.6% African American.
Statistical Analysis

Data were analyzed using available software (PSPP for Windows [http://pspp.awardspace.com/]). Case-control genotype associations were assessed using χ² analyses, and odds ratios were calculated. Estimates of departure from Hardy-Weinberg equilibrium were calculated using the χ² test. A statistical program (QuickCalcs; GraphPad Software Inc, La Jolla, California) was used to perform the 2-tailed Fisher exact test for comparing rare variant frequencies. Clinical data were analyzed using the same program, and Kaplan-Meier analysis was performed using another software package (Epi Info; Centers for Disease Control and Prevention, Atlanta, Georgia).

Results

SQSTM1 is located on chromosome 5q35, with 8 coding exons (Figure, A). To identify DNA mutations that predispose to ALS, the entire coding region of SQSTM1 was sequenced in a cohort of 546 patients with ALS (340 had FALS and 206 had SALS, representing 1929 chromosomes). Ten novel SQSTM1 mutations (9 heterozygous missense and 1 deletion) were identified in 15 patients (6 with FALS and 9 with SALS) (Table 1, Table 2, and eFigure [http://www.archneurol.com]). None of these changes were present in more than 724 controls (representing 1448 chromosomes), the dbSNP Short Genetic Variations database, or the 1000 Genomes Project database (Table 1). There was a personal history of parkinsonism in 2 patients. The A33V mutation was present in 1 patient with FALS and in 2 patients with SALS. The P392L substitution was found in 2 patients with FALS, but segregation analysis was impossible because of lack of samples from additional family members. The P392L variant was also present in 1 patient with SALS. The V153I change was present in 2 patients with SALS. The frequency of these variants in our cohort of patients with ALS was 2.8%.

We also identified 2 silent and 2 intronic variants exclusively in our ALS cohort that were not present in controls (Table 1). Several other rare and common variants were identified in cases and controls or were reported in the dbSNP Short Genetic Variations database, and their allele frequencies are given in eTable 1 and eTable 2 (http://www.archneurol.com). We defined rare variants as variations having frequencies of less than 1.0%. A few rare variants were observed exclusively in controls (eTable 1). We observed a statistically significant difference in the frequency of all rare variants exclusively present in patients with ALS vs in controls (22 of 1092 vs 11 of 1448; \( P = .007, 2\)-tailed Fisher exact test) (Table 3). Moreover, we also observed a statistically significant difference in the frequency of only rare missense or deletion variants exclusively present in patients with ALS vs in controls (16 of 1092 vs 9 of 1448; \( P = .04, 2\)-tailed Fisher exact test).

As shown in panel C of the Figure, p62 is highly conserved in mammals. All the mutations identified in our ALS cohort were located in conserved regions of p62. Four of 10 mutations observed in our ALS cohort were fully conserved across 7 species examined. Four mutated residues were conserved in 5 species. The V153 residue was conserved in 4 species. Eight of 9 missense variants were predicted to have a harmful effect on the structure and function of p62 by at least 1 of 2 protein conformation prediction methods.
prediction programs used (eTable 3). The frequency of functionally relevant rare variants exclusively present in patients with ALS was significantly higher than that in controls (13 of 1092 vs 6 of 1448; \( P = .03 \), 2-tailed Fisher exact test) (Table 3 and eTable 3).

The A33V substitution occurs in the SH2-binding domain (Figure, B). These domains are generally about 100 residues in length and are known to associate with phosphorylated tyrosine residues. The A33V change may affect phosphotyrosine ligand binding and specificity, which may lead to altered function of p62 in protein tyrosine kinase pathways. Indeed, several mutations in SH2 domain proteins are associated with human diseases. The V153I mutation occurs in the ZZ-type zinc finger domain, which is thought to be involved in protein–protein interactions and is present in proteins like dystrophin. The P228L and K238del mutations occur in the binding site for the tumor necrosis factor receptor–associated factor 6. The S318P and R321C mutations are not present in any known domains and may lead to abnormal protein folding and aggregation of p62. The S370P variant occurs in a PEST domain, which is a region enriched in proline, glutamic acid, serine, and threonine residues, and phosphorylation in this domain marks proteins for proteolysis. Moreover, the S318 residue occurs between 2 PEST domains and may remove a crucial phosphorylation site and make p62 more prone to aggregation. In fact, 4 of 10 mutations seen in our ALS cohort were predicted to have an effect on p62 phosphorylation (data not shown). Particularly, the S318P and S370P substitutions remove serine residues that are predicted to be highly probable phosphorylation sites. The UBA domain of p62 forms a compact 3-helix bundle (eFigure). The P392L and G411S substitutions are present just outside the hydrophobic patch of helix 1 and helix 2, whereas the G425R change occurs within the hydrophobic patch of helix 3. These UBA domain mutations may affect binding of p62 to ubiquitin or ubiquitinated proteins and may lead to accumulation of the ubiquitin-positive protein aggregates that are characteristic of ALS.

We obtained clinical data on 14 patients having SQSTM1 mutations for comparison with a cohort of patients having SOD1, TARDBP, and FUS mutations. The mean (SD) age at symptom onset among 14 patients with SQSTM1 mutations (54.6 [10.9] years) was similar to that among 34 patients with TARDBP mutations (54.7 [15.3] years).
years) but was later than that for 54 patients with FUS (43.6 [15.8] years, P = .02) or for 164 patients with SOD1 mutations (47.7 [13.0] years, P = .05) (2-tailed t test for both). We further tested the association between age at onset and different ALS-linked genes by comparing Kaplan-Meier survival curves and then evaluated the homogeneity of the survival curves by using the log-rank test and Wilcoxon test (data not shown). Using the log-rank test, we observed no significant differences between patients with SQSTM1 mutations vs patients with SOD1, FUS, or TARDBP mutations. However, differences between patients with SQSTM1 vs FUS mutations were significant using the Wilcoxon test (P = .01), which is more sensitive than the log-rank test to differences between groups that occur at earlier time points. The mean duration of symptoms was longer for 14 patients having SQSTM1 mutations (6.3 [5.3] years) compared with 44 patients having FUS mutations (3.4 [5.7] years), 144 patients having SOD1 mutations (4.1 [4.9] years), or 30 patients having TARDBP mutations (3.3 [2.3] years). The mean duration of symptoms in patients with SQSTM1 mutations was almost twice as long as that in patients with TARDBP mutations (P = .01, 2-tailed t test). The duration of symptoms varied widely. However, 64.3% of patients having SQSTM1 mutations survived beyond 4 years, which was remarkably higher compared with 11.4% of patients having FUS mutations, 29.9% of patients having SOD1 mutations, and 30.0% of patients having TARDBP mutations. In comparing the site of symptom onset, the proportion of patients with bulbar-onset symptoms was similar among patients having SQSTM1 mutations (28.6%) vs patients having FUS mutations (33.3%) or TARDBP mutations (32.1%) but was markedly higher than that among patients having SOD1 mutations (7.6%) (P = .05, 2-tailed Fisher exact test).

Herein, we describe SQSTM1 mutations in approximately 2% to 3% of our large cohort of patients with FALS and SALS from unrelated families. Although this frequency needs to be confirmed in future independent cohorts, it is similar to what has been reported for other genes involved in ALS, namely, FUS, TARDBP, VCP, and ANG. SQSTM1 mutations may confer a toxic gain of function through novel protein interactions and subsequent deregulation of cell signaling pathways. They may also lead to protein misfolding and aggregation. The SQSTM1 mutations described in our study may have low penetrance, as most were present in patients with small pedigree structure of familial aggregates or sporadic cases and not in large families. Other genes implicated in ALS, such as the PON genes and ANG, may cause disease by low-penetrance mutations, as they have been described in SALS and familial aggregates rather than in large multigenerational pedigrees. Some low-penetrance mutations in SOD1, FUS, and TARDBP have been reported in apparent SALS. Because our approach was based on candidate gene sequencing as opposed to linkage analysis, the possibility exists that the identified changes represent rare, possibly functional, variants conferring increased risk rather than pathogenic mutations. One criterion suggesting that a group of rare variants in a certain gene influences inherited susceptibility is that they are overrepresented in disease vs control groups. We observed statistically significant differences between ALS and controls whether we considered all exclusive variants, only missense and deletion variants, or only functionally relevant variants in SQSTM1. We also report several lines of evidence that suggest that the variants found in our ALS cohort may indeed be pathogenic. First, none of the variants present in our ALS cohort were detected in more than 724 controls (representing 1448 chromosomes), the dbSNP Short Genetic Variations database, or the 1000 Genomes Project database. Second, all these variants affect amino acids that show a high level of evolutionary conservation. Third, in silico analysis predicts that almost all these variants will have a deleterious effect on the structure and function of p62.

Histopathologic studies have shown that p62 is present in ubiquitinated inclusions of SOD1-positive FALS and other forms of ALS, suggesting a common pathogenic mechanism. Our study presents a parallel between p62 and other proteins linked to neurodegeneration, such as TDP-43, FUS, optineurin, β-amyloid, α-synuclein, and tau. These proteins may aggregate in a wide variety of neurological disorders, but mutations in their genes cause very specific phenotypes in rare families. Such rare but pathogenic mutations provide a novel approach in which the gene and its product can be investigated in molecular pathways at epigenetic, genetic, and posttranslational levels for relevance to sporadic disease.

Genes linked to 2 distinct clinical syndromes are well known. For instance, mutations in TRPV4 that were previously linked to bony dysplasias were recently linked to axonal neuropathies. Moreover, mutations in the gene encoding valosin-containing protein (VCP) have been implicated to cause human neurodegeneration in the syndrome of inclusion body myopathy with Paget disease of bone (PDB) and/or frontotemporal lobe degeneration (IBMPFD). Recently, VCP mutations were described by Johnson et al in patients with ALS. Notably, one of the mutations (R191Q) described by Johnson and colleagues in their ALS cohort had already been described in families with IBMPFD, and 2 other mutations from the same study (R159G and R155H) involved codons that had been found to be mutated in IBMPFD, highlighting the ability of the same mutation to confer variable clinical phenotypes. Furthermore, optineurin (mutations of which were recently linked to ALS) was identified as a genetic risk factor for PDB in a recent genome-wide association study. The 3 UBA domain mutations described in our ALS cohort have been previously identified in familial and sporadic PDB. This is intriguing because the coexistence of PDB and ALS, although not widely recognized, has been previously reported, suggesting a possible common link between these diseases. It is possible that this coexistence is underreported because PDB, like ALS, is rarely diagnosed before age 40 years, when symptoms of ALS (being more severe and lethal) would preclude PDB diagnosis. We found no evidence of a family or personal history of PDB in our cohort, and these mutations were absent in our control popu-
loration. It has been noted that affected individuals from the same family with PDB have variable and sometimes no expressivity of the disease even with a mutated copy of SQSTM1. Moreover, reported transgenic mouse models of PDB do not develop bone disease. This suggests that specific environmental factors or other modifier loci in addition to SQSTM1 mutations may be important in determining the specificity of the disease phenotype.

This study represents the first comprehensive genetic screen of SQSTM1 in ALS. Further screening and functional studies of these variants are needed to confirm their implication in ALS. Our data widen the clinical spectrum associated with SQSTM1 mutations to include ALS and suggest that patients with ALS should be monitored for features of PDB and, more broadly, altered bone metabolism. We can hypothesize that SQSTM1 mutations may act directly to cause ALS or through a modifier effect involving additional genes or environmental factors. The specific effects of these mutations in SQSTM1 on protein degradation pathways are important to resolve, as this may identify molecular targets for developing novel therapeutics in ALS.

Accepted for Publication: April 13, 2011.

Correspondence: Teepu Siddique, MD, Division of Neuromuscular Medicine, Ken and Ruth Davee Department of Neurology and Clinical Neurological Sciences, Feinberg School of Medicine, Northwestern University, Room 13-715, Tarry Bldg, 303 E Chicago Ave, Chicago, IL 60611 (t-siddique@northwestern.edu).

Author Contributions: Dr T. Siddique had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Fecto, Deng, and T. Siddique. Acquisition of data: Fecto, Yan, Vemula, Liu, Yang, Chen, Zheng, Shi, N. Siddique, Arrat, Donkervoort, Ajroud-Driss, Sufit, Heller, and Deng, T. Siddique. Analysis and interpretation of data: Fecto, Deng, and T. Siddique. Drafting of the manuscript: Fecto and T. Siddique. Critical revision of the manuscript for important intellectual content: Yan, Vemula, Liu, Yang, Chen, Zheng, Shi, N. Siddique, Arrat, Donkervoort, Ajroud-Driss, Sufit, Heller, and Deng. Obtained funding: T. Siddique. Study supervision: T. Siddique.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grants NS050641 and T32 AG20506 from the National Institutes of Health and by the Les Turner ALS Foundation, the Vena E. Schiff ALS Research Fund, the Harold Post Research Professorship, the Herbert and Florence C. Wenske Foundation, the Ralph and Marian Falk Medical Research Trust, the David C. Asselin MD Memorial Fund, and the Les Turner ALS Foundation/Herbert C. Wenske Foundation Professorship.

Previous Presentation: This study was presented at the 2011 Annual Meeting of the American Academy of Neurology; April 4, 2011; Honolulu, Hawaii.

Additional Contributions: We thank the patients and families who participated in the study.

Online-Only Material: The online-only eTables 1 through 3 and the eFigure are available at http://www.archneurol.com.

REFERENCES

24. Durbin RM, Abecasis GR, Altshuler DL, et al; 1000 Genomes Project Consor-


Incorrect Positioning. In the Observation titled “Novel POLG Splice Site Mutation and Optic Atrophy” by Milone et al, published in the June issue of the Archives (2011;68[6]:806-811), on page 807 Figure 1A and B should have been turned 90° counterclockwise for proper positioning.