The evolving genetic risk for sporadic ALS

ARTICLES

Summer B. Gibson, MD*
Jonathan M. Downie, BS*
Spyridoula Tsetsou, MD
Julie E. Feusier, BS
Karla P. Figueroa, MS
Mark B. Bromberg, MD, PhD
Lynn B. Jorde, PhD†
Stefan M. Pulst, MD, DrMed†

ABSTRACT

Objective: To estimate the genetic risk conferred by known amyotrophic lateral sclerosis (ALS)-associated genes to the pathogenesis of sporadic ALS (SALS) using variant allele frequencies combined with predicted variant pathogenicity.

Methods: Whole exome sequencing and repeat expansion PCR of C9orf72 and ATXN2 were performed on 87 patients of European ancestry with SALS seen at the University of Utah. DNA variants that change the protein coding sequence of 31 ALS-associated genes were annotated to determine which were rare and deleterious as predicted by MetaSVM. The percentage of patients with SALS with a rare and deleterious variant or repeat expansion in an ALS-associated gene was calculated. An odds ratio analysis was performed comparing the burden of ALS-associated genes in patients with SALS vs 324 normal controls.

Results: Nineteen rare nonsynonymous variants in an ALS-associated gene, 2 of which were found in 2 different individuals, were identified in 21 patients with SALS. Further, 5 deleterious C9orf72 and 2 ATXN2 repeat expansions were identified. A total of 17.2% of patients with SALS had a rare and deleterious variant or repeat expansion in an ALS-associated gene. The genetic burden of ALS-associated genes in patients with SALS as predicted by MetaSVM was significantly higher than in normal controls.

Conclusions: Previous analyses have identified SALS-predisposing variants only in terms of their rarity in normal control populations. By incorporating variant pathogenicity as well as variant frequency, we demonstrated that the genetic risk contributed by these genes for SALS is substantially lower than previous estimates. Neurology® 2017;89:226–233

GLOSSARY

ALS = amyotrophic lateral sclerosis; ALSoD = Amyotrophic Lateral Sclerosis Online Database; dbNSFP = database for nonsynonymous single nucleotide polymorphism functional predictions; dbSNP = Single Nucleotide Polymorphism database; ExAC = Exome Aggregation Consortium; FALS = familial amyotrophic lateral sclerosis; FTD = frontotemporal dementia; HGMD = Human Gene Mutation Database; MAF = minor allele frequency; OR = odds ratio; FCA = principal components analysis; SALS = sporadic amyotrophic lateral sclerosis; SNV = single nucleotide variant; SSC = Simons Simplex Collection; VQSR = variant quality score recalibration.

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease of the upper and lower motor neurons, which eventually leads to death within an average of 3–5 years after symptom onset. ALS is classified as familial (FALS) when a clear family history of ALS exists and sporadic (SALS) when it does not. No clinical features reliably distinguish FALS from SALS. Genetic research on ALS has largely been focused on FALS, which represents 10% of ALS cases. Most FALS is inherited in autosomal dominant fashion. However, this transmission pattern can be complicated by the early death of unrecognized affected family members due to non-ALS causes, misdiagnoses in older affected individuals, small family sizes, incomplete penetrance of genetic risk factors, and the development of disorders associated with ALS, such as frontotemporal dementia (FTD). Thus, sporadic and familial forms of ALS can be difficult to distinguish, and much remains unknown about the roles of genetic factors in FALS and especially in SALS.

*These authors contributed equally to this work.
†Co–senior authors.

From the Departments of Neurology (S.B.G., S.T., K.P.F., M.B.B., S.M.P.) and Human Genetics (J.M.D., J.E.F., L.B.J.), University of Utah School of Medicine, Salt Lake City.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.
The discovery of the pathogenic \((G\_C_2)_n\) hexanucleotide repeat expansion of \(C9orf72\) in a large percentage of FALS and SALS patients,\(^2\)\(^–\)\(^4\) as well as the identification of other ALS genes in patients with SALS,\(^5\)\(^–\)\(^6\) has highlighted the importance of genetic risk factors in SALS pathogenesis. The significance of genetics in SALS is further supported by ALS genome-wide association studies, which estimate the heritability of ALS to be at least 21.0%.\(^7\)

With the growing affordability and availability of next-generation sequencing technologies, along with the advent of specific treatments for certain genetic forms of ALS,\(^8\) it is increasingly important to understand the genetic factors in causing SALS. Currently, there is considerable variation in estimates of the percentage of SALS cases caused by genetic variants, ranging from 11%\(^5\) to 28%\(^9\) in populations of European ancestry. This variation is due largely to differences in estimation methods. In one large study,\(^9\) the percentage was derived by calculating the portion of SALS cases with a rare (minor allele frequency [MAF]<0.01), protein-altering variant in a set of known ALS genes. Using variant rarity as the main criterion for pathogenicity may have inflated the risk estimate as the majority of rare nonsynonymous variants are not thought to be pathogenic.\(^10\)

In this study, we sought to better estimate the percentage of SALS cases that have an identifiable genetic factor likely responsible for disease pathogenesis. To address this, a joint approach utilizing both allele frequency and variant pathogenicity prediction was used to determine the percentage of SALS cases that possess a potentially deleterious genetic variant in an ALS-associated gene.

**METHODS** Standard protocol approvals, registrations, and patient consents. The sample collection and study design we performed was approved by the University of Utah institutional review board. Written informed consent for disease-specific genetic studies was obtained from each patient who participated in this study.

Participants. Patients with ALS diagnosed at the University of Utah from 2011 to 2013 were invited to participate in genetic studies. All participants were seen by neuromuscular specialists and diagnosed with probable or definite ALS according to revised El Escorial criteria.\(^7\) These patients were followed longitudinally in our motor neuron disease clinic. Patients with SALS were identified as having no self-reported family history of ALS, probable ALS, or FTLD. In total, 96 patients with SALS were enrolled in this study. DNA was obtained from whole blood of each participant using the Gentra Puregene Blood Kit (Qiagen, Venlo, Netherlands).

Identification of deleterious \(ATXN2\) and \(C9orf72\) repeat expansions. \(ATXN2\) CAG repeat size was determined by fluorescent PCR amplification. Repeat lengths between 29 and 33 were considered to be of intermediate length and deleterious.\(^12\) The detection of \(C9orf72\) GGGGCC repeat expansions was performed by using previously established repeat primed-PCR and amplicon length analysis criteria.\(^13\)

Whole exome sequencing. Patient DNA was exome enriched by the SeqCap EZ Exome Enrichment Kit v3.0 (Roche [Basel, Switzerland] NimbleGen) and sequenced by the Illumina (San Diego, CA) HiSeq platform to generate 101-bp, paired-end reads that covered target regions to an average depth ranging from 41× to 224× per sample. Reads were then aligned to the GRCh37 reference genome using BWA-MEM v0.7.12. Picard Tools v1.150 was used to perform indexing, coordinate sorting, and duplicate read marking of all aligned genomic reads. Variant calling and quality filtering were performed using Genome Analysis Toolkit’s (v3.3-0) HaplotypeCaller and variant quality score recalibration (VQSR) methods.\(^14\) In order to properly power VQSR filtering, 99 CEU (Utah residents [CEPH] with northern and western European ancestry) and 92 GBR (British in England and Scotland) individuals from the 1000 Genomes Project\(^7\) with exome sequencing data were included in the genotyping steps.

Quality control. Utah’s population is outbred and genetically resembles other populations of northern European ancestry.\(^15\)\(^–\)\(^17\) As a result, we focused our analysis on patients with SALS who were of European ancestry in order to limit population stratification effects. Patient ancestry derived from genetic data is more reliable than self-reported ancestry, which has been used in previous ALS studies.\(^5\)\(^–\)\(^8\)\(^18\) An Admixture\(^19\) analysis (K = 3) was performed to determine the genetic ancestry of each patient with SALS. Any participants with less than 90% European ancestry were removed from further analysis. Next, principal components analysis (PCA) was performed using smartPCA\(^20\) to remove poor-quality samples. Any sample with an eigenvector value more than 6 SDs from the mean for the first 10 principal components was discarded. Finally, the sex of each patient with SALS was inferred by PLINK 1.9\(^21\) and compared to the reported sex to identify sample identification errors.

Variant annotation. SnpEff\(^22\) (v4.1) was used to identify protein-coding and splice-site altering genetic variants. These variants were then annotated with information from the database for nonsynonymous single nucleotide polymorphism functional predictions (dbNSFP; sites.google.com/site/jpopgen/dbNSFP) v2.9.\(^23\) dbNSFP contains 11 different in silico functional prediction methods that determine which single nucleotide variants (SNVs) are likely to alter protein function. MetaSVM was chosen as the primary method to determine variant pathogenicity as it has been shown to have a better predictive ability than other methods.\(^14\) Insertion, deletion, and splice-site acceptor/donor variants were classified as deleterious. Variants were also annotated with European-specific MAF estimates from the Exome Aggregation Consortium (ExAC)\(^24\) by dbNSFP. A manual search of the Amyotrophic Lateral Sclerosis Online Database (ALSod),\(^25\) the Single Nucleotide Polymorphism database (dbSNP),\(^7\) and the Human Gene Mutation Database...
Figure 1

Admixture and principal components analysis (PCA) plots show the ancestry and sample quality of the sporadic amyotrophic lateral sclerosis (SALS) cohort.

(A) An Admixture plot where each bar represents a patient with SALS (in total 96 patients). The height of each colored bar represents the amount of ancestry each individual derives from. Blue = European (CEU), green = East Asian (CHB + JPT), and red = African (YRI). Individuals with less than 90% European ancestry (yellow bar) were removed from further analysis. The 9 patients with SALS with less than 90% European ancestry are indicated with a red asterisk. (B) PCA plot of all 96 individuals with 1,000 genomes data (CEU = Utah residents [CEPH] with northern and western European ancestry; CHB = Han Chinese in Beijing, China; JPT = Japanese in Tokyo, Japan; YRI = Yoruba in Ibadan, Nigeria). Shaded areas represent the area over which the kernel density of each respective 1000 genomes population spans. SALS samples are listed as purple circles. Arrows indicate non-European individuals who were removed from further analysis.

(Raw Text)

Known ALS-associated genetic variants. We identified pathogenic C9orf72 hexanucleotide repeat expansions in 5 of 87 patients with SALS (5.7%). Two patients with SALS (2.3%) possessed ALS-associated trinucleotide repeat expansions in ATXN2 (31 and 32 repeats in length, respectively). We compared the rare (European MAF <0.001) coding variants in 31 ALS-associated genes uncovered in the SALS cohort to known ALS risk variants contained in ALSoD, dbSNP, and HGMD. This comparison revealed only one known ALS-associated rare SVN (SOD1:p.Asp91Ala), which was found in 2 heterozygous patients (table 2).
Sporadic amyotrophic lateral sclerosis. (Abbreviations: dbSNP, Single Nucleotide Polymorphism database; ExAC, Exome Aggregation Consortium; MAF, minor allele frequency; SALS, sporadic amyotrophic lateral sclerosis.)

Rate of progression data were available for 76 participants overall and for 69 analyzed.

Survival data were available for 78 participants overall and for 72 analyzed.

Table 1  Detailed summary of the sporadic amyotrophic lateral sclerosis cohort before and after selecting for European patients

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall (n = 96)</th>
<th>Analyzed (n = 87)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex, % (n)</td>
<td>61.5 (59)</td>
<td>62 (54)</td>
</tr>
<tr>
<td>Bulbar onset, % (n)</td>
<td>28.1 (27)</td>
<td>27.6 (24)</td>
</tr>
<tr>
<td>Age at onset, y, mean ± SD</td>
<td>58.7 ± 12.1</td>
<td>59.3 ± 11.6</td>
</tr>
<tr>
<td>Age at onset, y, range</td>
<td>19–83</td>
<td>19–83</td>
</tr>
<tr>
<td>Survival, y, mean ± SD*</td>
<td>2.9 ± 1.7</td>
<td>3 ± 1.8</td>
</tr>
<tr>
<td>Survival, y, range</td>
<td>0.5–12</td>
<td>0.5–12</td>
</tr>
<tr>
<td>Rate of progression, ALSFRS-R/y, mean</td>
<td>-13.1 ± 9.7</td>
<td>-12.4 ± 9.5</td>
</tr>
<tr>
<td>Rate of progression, ALSFRS-R/y, range</td>
<td>-1 to –60</td>
<td>-1 to –60</td>
</tr>
</tbody>
</table>

Abbreviation: ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-revised.
*Survival data were available for 78 participants overall and for 72 analyzed.
Rate of progression data were available for 76 participants overall and for 69 analyzed.

Potentially novel ALS variants. After examining the 31 ALS-associated genes in our patient cohort, we identified 18 rare coding variants (European MAF <0.001) not previously described in ALS (table 2). Of these variants, 10 were not found in dbSNP (v141). Furthermore, 6 were novel, as they were not found in ExAC, the 1000 Genomes Project dataset, or the National Heart, Lung, and Blood Institute Exome Sequencing Project dataset.

When combined with the 5 C9orf72 and 2 ATXN2 deleterious repeat expansions, 28 rare variants or repeat expansions were found in 25 patients across all ALS-associated genes. Three patients had 2 rare variants in an ALS gene. One patient had a GLE1:p.Met134Val (chr9:131277886A>G) variant in addition to a pathologic C9orf72 repeat expansion. Another patient had a C9orf72 repeat expansion in addition to a rare SETXp.Thr2507Ala (chr9:135140228T>C) missense variant. Finally, one patient possessed SPAST:p.Pro42His (chr2:32289025C>A) and ERBB4:p.Thr643Ile (chr2:212522497 G>A) missense variants.

These 28 rare variants were used to calculate the proportion of patients with SALS who have a rare mutation in at least one ALS-associated gene. A total of 25 patients with SALS (28.7%) had at least one

Table 2  The 19 rare nonsynonymous variants found in the 31 amyotrophic lateral sclerosis–associated genes

<table>
<thead>
<tr>
<th>Chromosome:position (GRCh37)</th>
<th>dbSNP141 ID</th>
<th>Amino acid change</th>
<th>Gene</th>
<th>MetaSVM prediction</th>
<th>ExAC European MAF</th>
<th>No. of patients with SALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:32289025C&gt;A</td>
<td>p.Pro42His</td>
<td>SPAST</td>
<td>Tolerated</td>
<td>0.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2:745934847T&gt;A</td>
<td>p.Ser883Cys</td>
<td>DCTN1</td>
<td>Tolerated</td>
<td>4.50E-05</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2:212251806T&gt;C</td>
<td>rs143251275</td>
<td>p.Thr1085Ala</td>
<td>ERBB4</td>
<td>Tolerated</td>
<td>4.50E-05</td>
<td>1</td>
</tr>
<tr>
<td>2:212483999C&gt;A*</td>
<td>p.Gly735Val*</td>
<td>ERBB4*</td>
<td>Damaging</td>
<td>0.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2:212522497G&gt;A</td>
<td>p.Thr643Ile</td>
<td>ERBB4</td>
<td>Tolerated</td>
<td>0.0001049</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4:170400673C&gt;G*</td>
<td>p.Gly646Arg*</td>
<td>NEK1*</td>
<td>Damaging</td>
<td>4.54E-05</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4:170428944C&gt;T*</td>
<td>c.1750-1G&gt;A (splice variant)*</td>
<td>NEK1*</td>
<td>NA*</td>
<td>0.0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5:138629745G&gt;T</td>
<td>p.Ala265Ser</td>
<td>MATR3</td>
<td>Tolerated</td>
<td>0.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5:138652744G&gt;A</td>
<td>rs201075828</td>
<td>p.Ala378Thr</td>
<td>MATR3</td>
<td>Tolerated</td>
<td>0.0001978</td>
<td>1</td>
</tr>
<tr>
<td>5:179263453A&gt;AT*</td>
<td>p.Glu396fs*</td>
<td>SQSTM1*</td>
<td>NA*</td>
<td>0.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6:110106211T&gt;A</td>
<td>p.Leu643*</td>
<td>FIG4</td>
<td>Tolerated</td>
<td>0.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9:131277886A&gt;G</td>
<td>p.Met134Val</td>
<td>GLE1</td>
<td>Tolerated</td>
<td>4.65E-05</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9:1351402287C</td>
<td>rs142303658</td>
<td>p.Thr2507Ala</td>
<td>SETX</td>
<td>Tolerated</td>
<td>7.49E-05</td>
<td>1</td>
</tr>
<tr>
<td>9:135144866G&gt;A</td>
<td>rs375949756</td>
<td>p.Pro2433Leu</td>
<td>SETX</td>
<td>Tolerated</td>
<td>1.57E-05</td>
<td>1</td>
</tr>
<tr>
<td>9:135203159G&gt;C</td>
<td>rs148604312</td>
<td>p.Gln1276Glu</td>
<td>SETX</td>
<td>Tolerated</td>
<td>0.0004495</td>
<td>1</td>
</tr>
<tr>
<td>9:1352032752G&gt;A</td>
<td>rs139559547</td>
<td>p.Ser1087Phe</td>
<td>SETX</td>
<td>Tolerated</td>
<td>1.50E-05</td>
<td>1</td>
</tr>
<tr>
<td>16:31195253T&gt;A</td>
<td>rs372638663</td>
<td>p.Ser89Thr</td>
<td>FUS</td>
<td>Tolerated</td>
<td>3.00E-05</td>
<td>1</td>
</tr>
<tr>
<td>16:312022284G&gt;A*</td>
<td>rs141684472*</td>
<td>p.Gly465Glu*</td>
<td>FUS*</td>
<td>Damaging</td>
<td>0.0001352*</td>
<td>1*</td>
</tr>
<tr>
<td>21:33009603A&gt;C*</td>
<td>rs80265967*</td>
<td>p.Asp91Ala*</td>
<td>SOD1*</td>
<td>Damaging</td>
<td>0.00087*</td>
<td>2*</td>
</tr>
</tbody>
</table>

Abbreviations: dbSNP = Single Nucleotide Polymorphism database; ExAC = Exome Aggregation Consortium; MAF = minor allele frequency; SALS = sporadic amyotrophic lateral sclerosis.
*Variants that were considered to be deleterious by MetaSVM. Indel and splice-site variants were automatically considered deleterious.

Neurology 89 July 18, 2017 229
© 2017 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.
rare variant or pathogenic repeat expansion in an ALS gene when variant deleteriousness was not considered. However, only 4 of the 17 SNVs annotated with MetaSVM were considered deleterious (table 2). As a result, the proportion of patients with SALS with a rare and deleterious SNP or repeat expansion in an ALS-associated gene was 17.2% (15/87 patients; figure 2). Variant predictions from 10 other methods were also used (table e-1 at Neurology.org), which yielded proportions ranging between 14.9% and 21.8%.

**OR analysis.** The genetic burden of ALS-associated genes in patients with SALS was compared with the burden among 324 SSC controls. Using only rare variant frequency as a criterion for assessing burden, patients with SALS had a modest increase in burden compared to controls (OR 1.90, p < 0.025; table 3). However, when variant pathogenicity was added by incorporating MetaSVM results and variant frequency, SALS cases showed a much higher burden in ALS-associated genes compared to controls (OR 4.98, p < 9 × 10^{-2}; table 3). Other variant prediction methods in dbNSFP yielded similar findings, but the OR analysis using MetaSVM predictions resulted in the highest p value (table e-2).

**DISCUSSION** We report findings from a genomic analysis of 87 patients with SALS of European origin. In total, 28 rare variants were found in 33 ALS genes in our patient cohort. Only one non-repeat variant that has been previously described in ALS pathogenesis was observed (SOD1:p.Asp91Ala). This variant is known to cause autosomal recessive ALS and was predicted to be deleterious by MetaSVM. SOD1:p.Asp91Ala has also been suggested to act in a dominant fashion; however, few instances of this have been reported. In addition, we identified 18 rare variants in ALS-associated genes that have not been described previously in patients with ALS. Of these 18 variants, 5 either caused a protein loss of function or were predicted to be deleterious by MetaSVM. One is a frameshift variant in the ubiquitin-associated domain of SQSTM1. This change likely ablates SQSTM1’s ability to bind ubiquitinated substrates, which is often seen in SQSTM1 variants that cause ALS. NEK1:c.1750-1G>A is a novel loss of function SNV that ablates the splice acceptor site of intron 19, which is located approximately in the middle of the gene. NEK1:p.Gly646Arg is another damaging variant that was discovered in NEK1; however, it does not occur in a defined protein domain. FUS:p.Gly465Glu is predicted to be damaging by MetaSVM and affects an amino acid one position upstream from previously reported SALS variant (FUS:p.Met464Ile). ERBB4:p.Gly735Val is an SNV predicted to be deleterious and occurs in the

---

**Table 3** Odds ratio (OR) analyses comparing the genetic burden of amyotrophic lateral sclerosis (ALS)-associated genes of patients with sporadic ALS (SALS) vs controls

<table>
<thead>
<tr>
<th>Variant prediction model</th>
<th>No. of patients with SALS with a rare and deleterious mutation in an ALS-associated gene/number without</th>
<th>No. of SSC individuals with a rare and deleterious mutation in an ALS-associated gene/number without</th>
<th>Odds ratio (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rare (MAF &lt; 0.001)</td>
<td>22/65</td>
<td>49/275</td>
<td>1.90 (1.07–3.38)</td>
<td>0.022</td>
</tr>
<tr>
<td>Rare + MetaSVM</td>
<td>15/72</td>
<td>13/311</td>
<td>4.98 (2.27–10.94)</td>
<td>8.90 × 10^{-4}</td>
</tr>
</tbody>
</table>

Abbreviations: CI = confidence interval; MAF = minor allele frequency; SSC = Simons Simplex Collection.

The incorporation of MetaSVM predictions of deleteriousness shows a much higher genetic burden of ALS-associated genes in patients with SALS than by considering variant rarity alone.
tyrosine kinase domain of erbB-4. The tyrosine kinase function of erbB-4 is required for protein autophosphorylation and triggering downstream signaling cascades upon activation. A variant in the tyrosine kinase domain of erbB-4, which was identified from a FALS family, has been shown to reduce protein autophosphorylation and likely causes ALS.24 Additional studies will determine the functional importance of these variants on cellular and molecular mechanisms.

We have demonstrated that using variant pathogenicity predictions is more reliable than variant frequency alone to determine the proportion of patients with SALS whose disease is likely caused by a variant in an ALS-associated gene. The relative effect of ALS-associated genes is stronger when variant pathogenicity is considered instead of only variant rarity. This follows from the fact that an appreciable proportion of rare nonsynonymous variants are not predicted to be functionally damaging.10 Thus, only a subset of rare variants in ALS-associated genes are pathogenic.

Our approach to estimating the genetic contribution of a large panel of known ALS-associated genes by directly predicting variant pathogenicity differs from earlier approaches. The first attempts at determining the proportion of genetically caused SALS cases did so by calculating the proportion of patients who had a protein-coding variant in a panel of 5–7 ALS-associated genes.23,33–35 These analyses yielded estimates ranging from 2.8% to 11%, which are lower than our estimate of 17.2%. A more recent study, in which variant rarity (MAF <1%) was used as the sole criterion for pathogenicity in a panel of 17 ALS-associated genes, concluded that genetic factors may cause 27.8% of SALS cases,27 a figure similar to our estimate when only variant rarity is considered (table e-1). However, these variants (MAF <1%) are not significantly more common in our patients with SALS than in unaffected controls (OR 1.25, p > 0.25), suggesting that many of them are not pathogenic. The same 17 ALS-associated genes are significantly more burdened in patients with SALS than controls when variant rarity (MAF <1%) and pathogenicity (estimated by MetaSVM) are combined (OR 2.61, p < 0.02). These OR differences support our conclusion that variant frequency alone is not a sufficient predictor of SALS risk.

Another analysis of 33 ALS-associated genes defined only novel and extremely rare variants (MAF ≈ 0.0002) as pathogenic and found that 14.5% of SALS cases could be attributed to genetic causes.39 In our sample, the genetic burden of ALS-associated genes in patients with SALS is less when pathogenicity is defined in the same way than when MetaSVM is integrated (OR 2.24, p < 0.01 vs OR 5.52, p < 2 × 10−4). These results demonstrate that direct predictions of variant pathogenicity are important for defining genetic risk in SALS and other genetic diseases.

Our results also highlight that genetic factors play an important role in the disease, the clinical relevance of which will become even more important as genetic specific treatments become available. Further, exome or targeted sequencing of patients with SALS and their family members is likely warranted to provide adequate genetic counseling. In addition, our results suggest the distinction between SALS and FALS may be problematic as heritable risk factors are found in a significant proportion of patients with SALS. Future genetic investigations of patients with SALS are needed to broaden the scope of SALS-associated loci. Studies with larger patient cohorts that incorporate measures of variant pathogenicity will also be needed to further pinpoint the proportion of SALS cases with an identifiable probable genetic cause of disease, especially as more ALS-associated genetic loci are discovered.

Our study has several limitations. First, the size of the SALS cohort was limited, especially given the genetically heterogeneous nature of ALS. Second, because we focused on individuals of European ancestry, our findings may not be completely applicable to ALS found in other populations. Third, 13 of the 324 (4.0%) healthy control samples used in this study had at least one rare and deleterious variant in ALS-associated genes as predicted by MetaSVM (table 3; table e-3). The mean age of these individuals was 41.76 (SD 5.92) years, which is much lower than the average age at onset of SALS at 56 years of age.40 It is possible that some of the control individuals with these variants could develop ALS later in life.

**AUTHOR CONTRIBUTIONS**

Dr. Gibson made a substantive contribution to the design and conceptualization of the study, analysis and interpretation of the data, and drafting and revising of the manuscript. J.M. Downie made a substantive contribution to the design and conceptualization of the study, analysis and interpretation of the data, and drafting and revising of the manuscript. Dr. Tsetsou made a substantive contribution to the drafting and revising of the manuscript. K.P. Figueroa made a substantive contribution to the analysis and interpretation of the data, and drafting and revising of the manuscript. Dr. Bromberg made a substantive contribution to the design and conceptualization of the study, interpretation of the data, and drafting of the manuscript. J.E. Feusier made a substantive contribution to the analysis and interpretation of the data and revising of the manuscript. K.P. Figueroa made a substantive contribution to the analysis and interpretation of the data as well as revising of the manuscript. Dr. Broome made a substantive contribution to the revising of the manuscript. Dr. Pulst made a substantive contribution to the design and conceptualization of the study, interpretation of the data, and revising of the manuscript.

**STUDY FUNDING**

Supported by NIH (TR001066, GM59290, GM104990, GM118335, and NS033123) and Target ALS. Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the NIH under award TR001066. The content is solely the responsibility of the authors and does not necessarily represent the official views of the authors and does not necessarily represent the official views of the NIH.
views of the NIH. Partial funding support for S.M.P. was provided by Targer ALS and by the NIH under award NS033125. L.B.J. was supported by NIH grants GM59290, GM104830, and GM118335. The control data used in this analysis can be found in the National Database for Autism Research (ndar.nih.gov/) under study DOI:10.15154/1149697.

DISCLOSURE
S. Gibson is a shareholder of Recursion Pharmaceuticals, J. Downie, S. Tsetsou, J. Feusier, K. Figueroa, M. Bromberg, and L. Jorde report no disclosures relevant to the manuscript. S. Pulst reports the following disclosures: Progenitor Life Sciences, shareholder; Cedars-Sinai, royalties; University of Utah, royalties; and Ataxon Therapeutics, consultant. Go to Neurology.org for full disclosures. Received December 8, 2016. Accepted in final form March 17, 2017.

REFERENCES
32. Al-Chalabi A, Andersen PM, Chiaota B, et al. Recessive amyotrophic lateral sclerosis families with the D90A SOD1