King's Research Portal

DOI: 10.1001/jamaneurol.2016.1114

Document Version
Publisher's PDF, also known as Version of record

Link to publication record in King's Research Portal

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher’s definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher’s website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 08. Mar. 2024
Association of a Locus in the CAMTA1 Gene With Survival in Patients With Sporadic Amyotrophic Lateral Sclerosis

Isabella Fogh, PhD; Kuang Lin, PhD; Cinzia Tiloca, PhD; James Rooney, MD; Cinzia Gellera, PhD; Frank P. Dielstra, MD; Antonia Ratti, PhD; Aleksey Shatunov, PhD; Michael A. van Es, MD, PhD; Petroula Proitsi, PhD; Ashley Jones, PhD; William Sproviero, PhD; Adriano Chiò, MD; Russell Lewis McLaughlin, PhD; Gianni Soraru, MD, PhD; Lucia Corrado, PhD; Daniel Stahl, PhD; Roberto Del Bo, PhD; Cristina Cereda, PhD; Barbara Castellotti, PhD; Jonathan D. Glass, MD; Steven Newhouse, PhD; Richard Dobson, PhD; Bradley N. Smith, PhD; Simon Topp, MSc; Wouter van Rheenen, MD; Vincent Meininger, MD, PhD; Judith Melki, MD, PhD; Karen E. Morrison, MD; Pamela J. Shaw, MD; P. Nigel Leigh, MD, PhD; Peter M. Andersen, MD, MSc; Giacomo P. Comi, MD; Nicola Ticozzi, MD; Letizia Mazzini, MD; Sandra D'Alfonso, PhD; Bryan J. Traynor, MD, PhD; Philip Van Damme, MD, PhD; Wim Robberecht, MD, PhD; Robert H. Brown, MD, DPhil; John E. Landers, PhD; Orla Hardiman, MD, FRCPI; Cathryn M. Lewis, PhD; Leonard H. van den Berg, MD, PhD; Christopher E. Shaw, MD; Jan H. Veldink, MD, PhD; Vincenzo Silani, MD; Ammar Al-Chalabi, PhD, FRCP; John Powell, PhD

IMPORTANCE Amyotrophic lateral sclerosis (ALS) is a devastating adult-onset neurodegenerative disorder with a poor prognosis and a median survival of 3 years. However, a significant proportion of patients survive more than 10 years from symptom onset.

OBJECTIVE To identify gene variants influencing survival in ALS.

DESIGN, SETTING, AND PARTICIPANTS This genome-wide association study (GWAS) analyzed survival in data sets from several European countries and the United States that were collected by the Italian Consortium for the Genetics of ALS and the International Consortium on Amyotrophic Lateral Sclerosis Genetics. The study population included 4256 patients with ALS (3125 [73.4%] deceased) with genotype data extended to 7,174,392 variants by imputation analysis. Samples of DNA were collected from January 1, 1993, to December 31, 2009, and analyzed from March 1, 2014, to February 28, 2015.

MAIN OUTCOMES AND MEASURES Cox proportional hazards regression under an additive model with adjustment for age at onset, sex, and the first 4 principal components of ancestry, followed by meta-analysis, were used to analyze data. Survival distributions for the most associated genetic variants were assessed by Kaplan-Meier analysis.

RESULTS Among the 4256 patients included in the analysis (2589 male [60.8%] and 1667 female [39.2%]; mean [SD] age at onset, 59 [12] years), the following 2 novel loci were significantly associated with ALS survival: at 10q23 (rs139550538; \( P = 1.87 \times 10^{-9} \)) and in the CAMTA1 gene at 1p36 (rs2412208, \( P = 3.53 \times 10^{-8} \)). At locus 10q23, the adjusted hazard ratio for patients with the rs139550538 AA or AT genotype was 1.61 (95% CI, 1.38-1.89; \( P = 1.87 \times 10^{-9} \)), corresponding to an 8-month reduction in survival compared with TT carriers. For rs2412208 CAMTA1, the adjusted hazard ratio for patients with the GG or GT genotype was 1.17 (95% CI, 1.11-1.24; \( P = 3.53 \times 10^{-8} \)), corresponding to a 4-month reduction in survival compared with TT carriers.

CONCLUSIONS AND RELEVANCE This GWAS robustly identified 2 loci at genome-wide levels of significance that influence survival in patients with ALS. Because ALS is a rare disease and prevention is not feasible, treatment that modifies survival is the most realistic strategy. Therefore, identification of modifier genes that might influence ALS survival could improve the understanding of the biology of the disease and suggest biological targets for pharmaceutical intervention. In addition, genetic risk scores for survival could be used as an adjunct to clinical trials to account for the genetic contribution to survival.
Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons in which relentlessly progressive weakness of voluntary muscles usually leads to death within 3 to 5 years of symptom onset. Amyotrophic lateral sclerosis is a heterogeneous disease with a poorly understood cause. Phenotypic variability in ALS is remarkable, consisting of heterogeneity in disease duration, age at onset, onset site, and type of motor neuron affected. Several ALS genes have been identified. Of these, a massive hexanucleotide repeat expansion in the chromosome 9 open reading frame 72 (C9orf72 [NCBI Entrez Gene 203228]) gene is the most common mutation in patients with the familial and sporadic ALS variants. Large genome-wide association studies (GWAS) have identified a number of susceptibility genes, including Unc-13 homologue A (UNC13A [NCBI Entrez Gene 23025]), CAMTA1 (NCBI Entrez Gene 23098), and sterile alpha and TIR motif containing 1 (SARM1 [NCBI Entrez Gene 203228]).

Despite the poor prognosis of ALS, about 5% of patients may survive more than 10 years. Long-term survivors are more likely to have primary lateral sclerosis, but all phenotypic patterns are represented. Younger age at onset correlates with longer survival, and other prognostic factors include disease progression rate at diagnosis, site of involvement at onset, certain phenotypic patterns (flail limb variants), cognitive impairment, and respiratory involvement. However, yet unknown factors are also likely to influence survival.

Previous studies have reported an association of survival with single-nucleotide polymorphisms (SNPs) in the kinesin-associated protein 3 (KIFAP3 [NCBI Entrez Gene 22920]) and UNC13A genes, although the KIFAP3 finding has not been replicated. Identification of gene variants influencing survival is thus crucial, particularly because these factors may provide important targets for disease-modifying therapies. To identify modifier genes that might influence ALS survival, we performed a GWAS using Cox proportional hazards regression modeling that included age at onset and onset site as covariates, followed by meta-analysis.

**Methods**

**Samples and Data**

Genotypes were obtained from previously published GWAS of patients with sporadic ALS from Italy, the United States, the United Kingdom, Ireland, Sweden, Belgium, and France, collected by the Italian Consortium for the Genetics of ALS (SLAGEN) and International Consortium on Amyotrophic Lateral Sclerosis Genetics (ALSGEN) (eMethods and eTable 1 in the Supplement). Participating patients fulfilled the El Escorial revised criteria for ALS without a reported family history of motor neuron diseases. Individuals included were of European ancestry by self-declaration. Clinical information was collected from medical notes, including the date of last consultation, and survival data from death certificates or hospital or public records. The site of onset was defined as bulbar for those in whom first weakness affected speech and swallowing and as spinal for those with limb or respiratory symptoms at onset. Symptom onset was defined as the date of first weakness or speech or swallowing disturbances. Survival duration was defined as the difference between the date of death or tracheostomy and the date of symptom onset and, for those still alive, as the difference between the censor date and the date of symptom onset. The censor date was taken as date of the last follow-up. This study was approved by the ethics boards of the participating institutions, and all patients or their representatives provided written informed consent.

**Genotyping, Quality Control, and Imputation Analysis**

Samples of DNA were collected from January 1, 1993, to December 31, 2009. We collected genotype raw data of previously published ALS GWAS with 12,426 individuals (6389 cases and 6037 controls) from 7 different cohorts (eMethods and eTable 1 in the Supplement). Quality control, imputation, and genotyped or inferred SNP-filtering procedures were performed separately per cohort, including data for cases and controls. In total, 11,136 individuals (5846 cases and 5290 controls) passed stringent quality control (eMethods and eTable 2 in the Supplement). To improve accuracy of the downstream imputation analysis, cleaned, genotyped SNPs were first aligned to hg19 coordinates and phased by estimation of the samples’ haplotype structures according to the 1000 Genomes Project reference (phase 1, version 3, NCBI build 37, hg19 coordinates, August 2012) using ShapeIt2 software (eMethods in the Supplement). Finally, aligned and prephased genotyped SNP data were imputed genome wide using the IMPUTE2 toolset with 1000 Genomes phase 1 as the reference panel. After imputation analysis, genotype data of cases only were extracted from each cohort; in total, 4256 (72.8%) of 5846 patients with sporadic ALS had complete clinical information (eTable 3 in the Supplement) and therefore were included in the present study. Given the low SNP coverage present in some original commercial arrays (eTable 1 in the Supplement), the proportion of uncertain inferred genotypes was high, with a mean of 52.3% of SNPs per cohort (eMethods and eTable 2 in the Supplement) that did not pass the stringent quality control threshold by posterior probability greater than 0.9, information metric greater than 0.4, and minor allele frequency (MAF) from 0.4% to 2% (eMethods and eTables 2 and 4 in the Supplement). In total, 7,174,392 overlapping variants within each cohort were tested for association with ALS survival using Cox proportional hazards regression analysis.
Statistical Analysis
Data were analyzed from March 1, 2014, to February 28, 2015. Multivariate Cox proportional hazards regression was modeled to estimate crude hazard ratios (HRs) and build by backward elimination (Wald test) estimation of HRs with 95% CI. The Cox proportional hazards regression baseline model included age at onset (as a continuous variable) and sex and onset site (bulbar vs spinal) as factor variables (eTable 4 in the Supplement). We tested the proportional hazards assumption by comparing the hazard curves stratified by sex, age at onset, and onset site. All tests were 2-tailed, and significance was assessed at P < .05 and performed in SPSS software (version 22; IBM Corporation).

The Cox proportional hazards model was applied genome-wide to filtered imputed data in each population with the following independent variables: SNP genotype under a log-additive model, the 4 principal components of ancestry, sex, and age at onset. To maximize power in the exploratory analysis, onset site was omitted in the final model owing to the smaller numbers of patients (3438 [80.8%]) with this information. The model was built by backward elimination using the pacoroph program in the ProbABEL21 toolset to estimate the HR with 95% CI, model, and covariate P values for each SNP. Statistical significance was assessed at the genome-wide level (P = 5 × 10^-8).

Summary statistics for 7 174 392 overlapping SNPs were combined in a meta-analysis using METAL software22 weighted by β coefficients and the inverse of the corresponding standard errors; the fixed-effects model was applied to adjust data by β coefficients and the inverse of the corresponding standard errors. Statistical significance was assessed at the genome-wide level (P = 5 × 10^-8).

Results
The international ALS cohort analyzed in the present study included 4256 patients (2589 male [60.8%] and 1667 female [39.2%]), of whom 3125 (73.4%) had died after a median survival of 32.8 (interquartile range [IQR], 22.2-49.2) months. The mean (SD) age at onset, including censored individuals, was 59.1 (12.1) years (eTable 6 in the Supplement).

Data for onset site were available in a subset of 3438 patients (80.8% [2066 male and 1372 female]; 1025 [29.8%] had bulbar onset, with a mean (SD) age at onset of 62.4 (11.4) years compared with spinal onset at a mean (SD) age of 57.7 (12.5) years. The median survival was 27.5 (IQR, 19.8-39.5) and 35.9 (IQR, 22.9-56.4) months in patients with bulbar and spinal onset, respectively. Full details are reported in eTable 7 in the Supplement.

A total of 7 174 392 SNPs had genotypes that passed quality control measures. Two loci exceeded the genome-wide significance threshold: one on chromosome 10q23 and one on chromosome 1p36 (Figure 1A and Table 1). At locus 10q23, the top-ranked SNP was rs139550538, with a hazard ratio of 1.61 (95% CI, 1.38-1.89; P = 1.87 × 10^-9). This variant is moderately rare (MAF, 0.03) and intronic within the insulin-degrading enzyme (IDE [NCBI Entrez Gene 3416]) gene (Figure 1C).

At the 1p36 locus, 4 SNPs exceeded genome-wide significance, with the top-ranked SNP being rs2412208 (HR, 1.17; 95% CI, 1.11-1.24; P = 3.53 × 10^-8), followed by 87 SNPs in strong linkage disequilibrium with rs2412208. All these SNPs fell within a 90-kilobase region encompassing introns 3 to 4 of the calmodulin-binding transcription activator 1 (CAMTA1 [NCBI Entrez Gene 23261]) gene (Figure 1B and Table 1). Cox proportional hazards regression analyses conditioning on the most associated SNPs in both loci showed no evidence of residual association.

Because rs139550538 is rare, Kaplan-Meier analysis was performed under a dominant model (226 patients [5.3%] carried ≥1 A allele). The AA or AT genotype was associated with ALS survival (log-rank P = 1.3 × 10^-7) and a median survival of 30.7 months compared with 36.7 months for the TT homozygotes (Figure 2A and Table 2). Kaplan-Meier analysis of SNP rs2412208 under an additive model showed that carrying a G allele (1909 patients [44.9%]) was significantly associated with a decreased survival (log-rank P = 3.5 × 10^-8), with median survivals of 36.0 months (GG) and 36.8 months (GT) in contrast to 40.8 months in TT carriers (Figure 2B and Table 2). The HR estimates were consistent across the 7 data sets analyzed (Figure 3). Under a dominant model, the results were similar (Figure 2C) and a χ² likelihood ratio test comparing the 2 models was not significant (P = .12), showing that either could be valid.

We tested whether observed effect sizes (β) of the most associated SNPs from the combined Cox proportional hazards regression analysis were homogeneous across cases. We found some evidence of heterogeneity across the different data sets (rs2412208; F² = 57.2%; P = .03) (eTable 5 in the Supplement).

In the subset of 3439 patients with ALS with clinical data including onset site, Cox proportional hazards regression was modeled with this variable as an additional covariate. The top-ranked SNP was rs2412208 at 1p36 with the combined HR of 1.19 (95% CI, 1.27-1.12; P = 5.11 × 10^-8) (eFigure 2 and eTable 8 in the Supplement), confirming association of the CAMTA1 locus with ALS survival identified by the larger sample size when this covariate was not included in the model. The SNP rs139550538 in IDE gene was less significant (HR, 1.51; 95% CI, 1.27-1.78; P = 2.24 × 10^-5), possibly because of the lower frequency of this SNP in a reduced sample. In addition, in linear regression analysis testing bulbar vs spinal phenotypes, this variant was not significantly associated (P = .50), indicating that the onset site was unlikely to confound the 10q23 association with survival.

Kaplan-Meier distribution of rs2412208 genotypes indicated that risk allele G was associated (log-rank P = 5 × 10^-5) with a shorter survival of 3.5 months, which corresponded to a 19% increased rate of mortality compared with the TT ho-
Figure 1. Genome-wide Association Study of Survival in Patients With Sporadic Amyotrophic Lateral Sclerosis (ALS)

A, Manhattan plot of the combined (METAL software) Cox proportional hazards regression analysis. The threshold for genome-wide significance after correction for multiple testing was set at \( P = 5 \times 10^{-8} \) (horizontal blue line). Loci significantly associated with ALS survival are highlighted in red and labeled according to the corresponding genes. At locus 10q23, the most associated single-nucleotide polymorphism (SNP), rs139550538 \( (P = 1.87 \times 10^{-9}) \), was moderately rare with a minor allele frequency (MAF) of 0.03, whereas at the 1p36 locus, the 4 SNPs significantly associated (rs2412208 \( P = 3.53 \times 10^{-8} \), rs4584415 \( P = 3.68 \times 10^{-8} \), rs35447019 \( P = 3.86 \times 10^{-8} \), and rs4409676 \( P = 4.48 \times 10^{-8} \)) were common (MAF > 0.26). B and C, Regional linkage disequilibrium (LD) plots of the 2 regions significantly associated with ALS survival. At the 1p36 locus, 4 SNPs passed the genome-wide significant threshold, followed by 87 tagged proxies suggestively associated \( (P < 10^{-4}) \). All the associated SNPs mapped within introns 3 to 4 of the CAMTA1 gene. At the 10q23 locus, the top-ranked SNP, rs139550538, intronic to the IDE gene, was in weak \( (r^2 < 0.4) \) LD with the tagged proxies that were located in the neighbor gene, KIF11.

mozygotes (eFigure 3 in the Supplement). We examined previously reported candidate genes for ALS survival. The SNP rs1541160 in the KIFAP3 gene was not significantly associated with survival in this study \( (HR, 1.04; 95\% CI, 0.98-1.1; P = .42) \) (eFigure 4 in the Supplement), which confirmed previous findings.\(^{15,16}\) The SNP rs2608932 in the UNCI3A gene showed suggestive association \( (HR, 1.17; 95\% CI, 1.1-1.24; P = .003) \), but coverage for this SNP was limited to a reduced subset of patients \( (n = 3574) \) (eFigure 4 in the Supplement), and further studies on a larger scale are needed to validate the genetic effect of UNCI3A as survival modifier. Of 105 SNPs tested in the D-amino acid oxidase \( (DAO \text{ NCBI Entrez Gene 1610}) \) gene,\(^{23}\) none passed Bonferroni correction for multiple testing, with the top-ranked SNP being rs4623951 \( (HR, 1.07; 95\% CI, 1.02-1.13; \text{uncorrected } P = .005) \). To replicate the association of the EPH receptor A4 \( (EPHA4 \text{ NCBI Entrez Gene 2043}) \) gene with ALS survival,\(^{24}\) we analyzed 1743 SNPs within the gene region. None of these variants reached the genome-wide significance \( (P = 5 \times 10^{-8}) \) or passed Bonferroni correction for multiple testing; the top-ranked SNP rs6436254 (MAF, 0.47) was intronic in EPHA4 and associated with ALS survival \( (HR, 1.07; 95\% CI, 1.02-1.26; P = .007) \).

Discussion

We have identified 2 loci associated with survival in patients with ALS at genome-wide significance in a large meta-analysis using Cox proportional hazards regression analysis. The discovery of gene variants within the IDE and CAMTA1
The effect of size of the variants found is comparable to that of riluzole, a drug shown to improve survival in ALS, for which the HR for those not taking vs taking riluzole is 1.14. A weakness of our study is that the extent of riluzole use was not available to include in the analysis. Generally, rates of prescription are higher in countries in which access to health care is free or reimbursed than in those where private insurance is required, and if such differences correlate with allele frequency differences, a spurious association might arise. We mitigated against this association by accounting for differences in allele frequency by ancestry using principal components and if such differences correlate with allele frequency differences, a spurious association might arise.
performing a meta-analysis stratified by country. Multidisciplinary clinic attendance has also been reported to increase survival, which also may vary across countries, but we have accounted for this possibility through the country-stratified meta-analysis.

The most associated polymorphism at the 10q23 locus was a low-frequency variant within the IDE gene, a zinc metallopeptidase that degrades intracellular insulin and other peptides, such as β-amyloid. Tagged proxies for this polymorphism were in weak linkage disequilibrium and median survival of 36.0 and 36.8 months, respectively, compared with 40.8 months in 2347 TT carriers. C, Variant CAMTA1 rs2412208 genotypes under a dominant model show survival in 1909 GG/GT and 2347 TT carriers; TT homozygotes have a life span extended by more than 4 months. Kaplan-Meier curves report patients’ survival up to 10 years, plotted in SPSS software.

Kaplan-Meier curves report patients’ survival up to 10 years, plotted in SPSS software.

Kaplan-Meier curves plot survival distribution. A, IDE rs139550538 distribution of genotypes under a dominant model. Survival in the 226 AA/AT carriers was compared with that in 4030 TT carriers, showing that the presence of at least 1 A allele is associated with a median survival of 30.7 months compared with 36.7 months in TT homozygotes. B, Variant CAMTA1 rs2412208 genotypes under an additive genetic model show 265 GG and 1644 GT carriers with a median survival of 36.0 and 36.8 months, respectively, compared with 40.8 months in 2347 TT carriers. C, Variant CAMTA1 rs2412208 genotypes under a dominant model show survival in 1909 GG/GT and 2347 TT carriers; TT homozygotes have a life span extended by more than 4 months. Kaplan-Meier curves report patients’ survival up to 10 years, plotted in SPSS software.
Table 2. Genetic Effect of the Most Associated SNPs in the Summary Cox Proportional Hazards Regression Analysis

<table>
<thead>
<tr>
<th>SNP</th>
<th>No. (%) of Patients</th>
<th>Median Survival, mo</th>
<th>HR (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs139550538 (Chr 10q23; IDE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additive genetic model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>2 (0.05)</td>
<td>19.0</td>
<td>1.61 (1.38-1.89)</td>
<td>1.87 × 10^{-9}</td>
</tr>
<tr>
<td>TA</td>
<td>224 (5.3)</td>
<td>31.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>4030 (94.7)</td>
<td>39.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant genetic model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA/AT</td>
<td>226 (5.3)</td>
<td>30.7</td>
<td>1.52 (1.31-1.77)</td>
<td>1.3 × 10^{-7}</td>
</tr>
<tr>
<td>TT</td>
<td>4030 (94.7)</td>
<td>36.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2412208 (Chr 1p36; CAMTA1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additive genetic model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>265 (6.2)</td>
<td>36.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>1644 (38.6)</td>
<td>36.8</td>
<td>1.17 (1.11-1.24)</td>
<td>3.5 × 10^{-8}</td>
</tr>
<tr>
<td>TT</td>
<td>2347 (55.1)</td>
<td>40.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant genetic model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG/GT</td>
<td>1909 (44.9)</td>
<td>36.6</td>
<td>1.18 (1.09-1.26)</td>
<td>4.6 × 10^{-6}</td>
</tr>
<tr>
<td>TT</td>
<td>2347 (55.1)</td>
<td>40.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Forest Plot of CAMTA1 rs2412208 Hazard Ratio (HR) Estimates

Hazard ratio estimates are measured under an additive genetic model across the 7 genome-wide association study (GWAS) data sets (described in detail in eTable 1 in the Supplement). The HR in each cohort was estimated in a multivariate log-additive genetic model using the rmeta program adjusted by age at onset, sex, and population stratification, whereas the summary HR was calculated by fixed-effects meta-analysis using R library rmeta. Genotype raw data included in each study and combined in the meta-analysis were collected from exons 3 to 7 (6825092 to 7640553 base pair; GRCh37/hg19 Assembly) within CAMTA1. Intragenic CAMTA1 microrearrangements disrupting a CG-1 DNA-binding domain have been reported to cosegregate with nonprogressive congenital cerebellar ataxia (NPCA) and gait instability in several unrelated families. CSF CAMTA1 variants within CAMTA1 have also been reported to be associated with variation in human episodic memory. Mutant CAMTA1 knockout mice, disrupted in the CG-1 domain, show severe ataxia and neuronal atrophy approximating the phenotype of haploinsufficiency observed in patients with NPCA. Furthermore, the identification of the consensus sequences of the DNA-binding site of the CG-1 domain combined with expression analyses in CAMTA1 knockout mice have shown more than 80 neural-related genes regulated by CAMTA1. The finding of a gene involved in cerebellar disease in ALS is not surprising given that trinucleotide repeat expansion in the ataxin 2 (ATXN2) gene causes spinocerebellar ataxia or ALS, the finding of C9orf72 pathologic mechanisms in the cerebellum of patients with ALS, and the discovery of abnormal eye gaze in patients with ALS. Increasing evidence suggests an association between ALS and cerebellar degeneration that is currently underrecognized, in the same way as the association between ALS and frontotemporal dementia remained undetected until recently.

A strength of our study is, to our knowledge, the use of the largest data set for sample size (n = 4256) and genotyped SNP coverage (>7 million) analyzed to date. In addition, the use of a Cox proportional hazards regression model allowed us to include 1131 patients still alive (26.6%), meaning the study was not biased by the restriction of a linear regression method limited to patients who have died. In contrast, a potential weakness of our study is the difficulty of imputing low-frequency variants; a reference panel including disease-specific geno- type data will improve imputation of rare variants.

Iterative and accelerated forward stepwise multivariate analyses in a Cox proportional hazards regression model allowed us to include 1131 patients still alive (26.6%), meaning the study was not biased by the restriction of a linear regression method limited to patients who have died. In contrast, a potential weakness of our study is the difficulty of imputing low-frequency variants; a reference panel including disease-specific genotype data will improve imputation of rare variants.

located in a neighboring gene, the kinesin family member 11 (KIF11 [NCBI Entrez Gene 3832]), a motor kinesinlike protein involved in the spindle function during cell mitosis (Figure IC). The biological basis of this association therefore unclear.

The most associated 87 variants in the CAMTA1 gene (P ≤ 10^{-4}) map to a small 90-kilobase region within introns 3 to 4 (Figure 1B) encompassing the CG-1 DNA-binding domain. The CG-1 motif is a functional domain with a nuclear localization signal and transcriptional regulation properties that extends from exons 3 to 7 (6825092 to 7640553 base pair; GRCh37/hg19 Assembly) within CAMTA1. Intragenic CAMTA1 microrearrangements disrupting a CG-1 DNA-binding domain have been reported to cosegregate with nonprogressive congenital cerebellar ataxia (NPCA) and gait instability in several unrelated families. Common variants within CAMTA1 have also been reported to be associated with variation in human episodic memory. Mutant CAMTA1 knockout mice, disrupted in the CG-1 domain, show severe ataxia and neuronal atrophy approximating the phenotype of haploinsufficiency observed in patients with NPCA. Furthermore, the identification of the consensus sequences of the DNA-binding site of the CG-1 domain combined with expression analyses in CAMTA1 knockout mice have shown more than 80 neural-related genes regulated by CAMTA1. The finding of a gene involved in cerebellar disease in ALS is not surprising given that trinucleotide repeat expansion in the ataxin 2 (ATXN2) gene causes spinocerebellar ataxia or ALS, the finding of C9orf72 pathologic mechanisms in the cerebellum of patients with ALS, and the discovery of abnormal eye gaze in patients with ALS. Increasing evidence suggests an association between ALS and cerebellar degeneration that is currently underrecognized, in the same way as the association between ALS and frontotemporal dementia remained undetected until recently.

A strength of our study is, to our knowledge, the use of the largest data set for sample size (n = 4256) and genotyped SNP coverage (>7 million) analyzed to date. In addition, the use of a Cox proportional hazards regression model allowed us to include 1131 patients still alive (26.6%), meaning the study was not biased by the restriction of a linear regression method limited to patients who have died. In contrast, a potential weakness of our study is the difficulty of imputing low-frequency variants; a reference panel including disease-specific genotype data will improve imputation of rare variants.

Iterative and accelerated forward stepwise multivariate analyses in a Cox proportional hazards regression model allowed us to include 1131 patients still alive (26.6%), meaning the study was not biased by the restriction of a linear regression method limited to patients who have died. In contrast, a potential weakness of our study is the difficulty of imputing low-frequency variants; a reference panel including disease-specific genotype data will improve imputation of rare variants.
Conclusions

We have identified genetic variants that have a statistically significant association with survival. The promise of this research is not only to improve our understanding of the biology of the disease and suggest biological targets for pharmaceutical intervention to extend the survival time of the patients but also to use genetic risk scores as an adjunct to clinical trials to account for the genetic contribution to survival.
REFERENCES


