

## ORIGINAL ARTICLE

# Whole-Genome Analysis of Sporadic Amyotrophic Lateral Sclerosis

Travis Dunckley, Ph.D., Matthew J. Huentelman, Ph.D., David W. Craig, Ph.D., John V. Pearson, B.Sc., Szabolcs Szelinger, B.S., Keta Joshipura, B.S., Rebecca F. Halperin, B.Sc., Chelsea Stamper, B.S., Kendall R. Jensen, Ph.D., David Letizia, M.S., Sharon E. Hesterlee, Ph.D., Alan Pestronk, M.D., Todd Levine, M.D., Tulio Bertorini, M.D., Michael C. Graves, M.D., Tahseen Mozaffar, M.D., Carlyne E. Jackson, M.D., Peter Bosch, M.D., April McVey, M.D., Arthur Dick, M.D., Richard Barohn, M.D., Catherine Lomen-Hoerth, M.D., Jeffrey Rosenfeld, M.D., Daniel T. O'Connor, M.D., Kuixing Zhang, M.D., Ph.D., Richard Crook, Ph.D., Henrik Ryberg, Ph.D., Michael Hutton, Ph.D., Jonathan Katz, M.D., Ericka P. Simpson, M.D., Hiroshi Mitsumoto, M.D., Robert Bowser, Ph.D., Robert G. Miller, M.D., Stanley H. Appel, M.D., and Dietrich A. Stephan, Ph.D.

## ABSTRACT

**BACKGROUND**

Approximately 90% of persons with amyotrophic lateral sclerosis (ALS) have the sporadic form, which may be caused by the interaction of multiple environmental factors and previously unknown genes.

**METHODS**

We performed a genomewide association analysis using 766,955 single-nucleotide polymorphisms (SNPs) found in 386 white patients with sporadic ALS and 542 neurologically normal white controls (the discovery series). Associations of SNPs with sporadic ALS were confirmed in two independent replication populations: replication series 1, with 766 case patients with the disease and 750 neurologically normal controls, and replication series 2, with 135 case patients and 275 controls.

**RESULTS**

We identified 10 genetic loci that are significantly associated ( $P < 0.05$ ) with sporadic ALS in three independent series of case patients and controls and an additional 41 loci that had significant associations in two of the three series. The most significant association with disease in white case patients as compared with controls was found for a SNP near an uncharacterized gene known as *FLJ10986* ( $P = 3.0 \times 10^{-4}$ ; odds ratio for having the genotype in patients vs. controls, 1.35; 95% confidence interval, 1.13 to 1.62). The *FLJ10986* protein was found to be expressed in the spinal cord and cerebrospinal fluid of patients and of controls. Specific SNPs seem to be associated with sex, age at onset, and site of onset of sporadic ALS.

**CONCLUSIONS**

Variants of *FLJ10986* may confer susceptibility to sporadic ALS. *FLJ10986* and 50 other candidate loci warrant further investigation for their potential role in conferring susceptibility to the disease.

From the Translational Genomics Research Inst., Phoenix, AZ (T.D., M.J.H., D.W.C., J.V.P., S.S., K.J., R.F.H., C.S., K.R.J., D.L., D.A.S.); Muscular Dystrophy Association, Tucson, AZ (S.E.H.); Washington Univ. School of Medicine, St. Louis (A.P.); Phoenix Neurological Associates, Phoenix, AZ (T.L.); Univ. of Tennessee, Memphis (T.B.); Univ. of California, Los Angeles (M.C.G.); Univ. of California, Irvine (T.M.); Univ. of Texas Health Science Center, San Antonio (C.E.J.); Mayo Clinic, Scottsdale, AZ (P.B.); Univ. of Kansas Medical Center, Kansas City (A.M., A.D., R.B.); Univ. of California, San Francisco (C.L.-H.); Carolinas Medical Center, Charlotte, NC (J.R.); Univ. of California at San Diego School of Medicine, La Jolla (D.T.O., K.Z.); Mayo Clinic College of Medicine, Jacksonville, FL (R.C., M.H.); Univ. of Pittsburgh Medical Center, Pittsburgh (H.R., R.B.); California Pacific Medical Center, San Francisco (J.K., R.G.M.); Methodist Neurological Inst., Houston (E.P.S., S.H.A.); and Columbia Univ. Medical Center, New York (H.M.). Address reprint requests to Dr. Stephan at the Translational Genomics Research Inst., 445 N. Fifth St., Phoenix, AZ 85004, or at dstephan@tgen.org.

This article (10.1056/NEJMoa070174) was published at [www.nejm.org](http://www.nejm.org) on August 1, 2007.

N Engl J Med 2007;357.

Copyright © 2007 Massachusetts Medical Society.

**A**MYOTROPHIC LATERAL SCLEROSIS (ALS) is the most common motor neuron disease, with an incidence of 1 or 2 cases per 100,000 persons and a lifetime risk of 1 case per 800 persons. It is characterized by a progressive loss of motor neurons from the spinal cord, brain stem, and cerebral cortex,<sup>1</sup> leading to paralysis and death within 2 to 5 years after diagnosis, without intensive physiological support. Ten percent of ALS cases are familial forms resulting from highly penetrant, monogenic mutations that cause disease. Familial ALS is caused by mutation of many genes that have been identified,<sup>2</sup> including *SOD1* (encoding superoxide dismutase 1); five other dominant loci — ALS3, ALS4 (with *SETX* the causative gene at this locus), ALS6, ALS7, and ALS8 (with *VAPB* the causative gene); and two recessive loci — ALS2 (encoding alsin) and ALS5. Four additional dominantly inherited loci have been linked with variant forms of motor neuron disease: two loci for ALS with frontotemporal dementia (one on chromosome 9q21–22 and the other on 9p21.3–13.3), one locus for ALS with dementia and Parkinson's disease (*MAPT* on chromosome 17q21), and one locus for progressive lower motor neuron disease (*DCTN1* on chromosome 2p13).<sup>2</sup>

Little is known about the specific genes that contribute to the development of sporadic ALS. Moreover, despite extensive study of familial ALS-causing mutations in vitro and in animal models of the disease, the key events in the initiation and progression of sporadic ALS remain unclear. Pathologically, sporadic ALS is characterized by loss of motor neurons from the motor cortex, brain stem, and ventral horns of the spinal cord. Ubiquitinated inclusions (covalent bonds between ubiquitin and other proteins that mark them for degradation) have been observed in the lower motor neurons, although their role in the initiation and progression of disease is unclear.<sup>3</sup> Numerous mechanisms have been implicated in the selective degeneration of motor neurons in patients with sporadic ALS, including oxidative damage, excitotoxicity, apoptosis, cytoskeletal dysfunction, axonal-transport defects, inflammation, protein-processing and degradation defects, and mitochondrial dysfunction.<sup>1,4</sup>

Identification of the specific genetic variants associated with sporadic ALS will improve our understanding of fundamental disease mechanisms. To this end, genomewide association stud-

ies provide a comprehensive, unbiased approach to screen groups of patients with sporadic ALS and groups of controls for genetic markers that are more common in patients with ALS and thus may reside in or near predisposition genes. To identify such markers, we carried out a genomewide case-control association study.

---

## METHODS

---

### ACQUISITION OF SAMPLES

The overall study was approved at the Translational Genomics Research Institute by the Western Institutional Review Board and by the appropriate institutional review board at each participating site. Written informed consent was obtained from all participants. Between April 27 and October 6, 2006, we prospectively collected 1251 DNA samples from patients with a diagnosis of laboratory-supported probable, probable, or definite sporadic ALS, according to the El Escorial diagnostic criteria, and used the Motor Neuron Disease Clinical Data Elements form of the National Institute of Neurological Disorders and Stroke to facilitate data sharing with the community (see [www.alsrg.org](http://www.alsrg.org) and Table 1 in the Supplementary Appendix, available with the full text of this article at [www.nejm.org](http://www.nejm.org)).<sup>5</sup> Patients and controls were recruited and enrolled from all participating clinical sites.

We prospectively collected 1251 DNA samples from patients with sporadic ALS. Among these were 231 DNA samples obtained from the Coriell Cell Repositories. The 231 samples were cross-referenced with those collected from our prospectively enrolled patients, and we removed three samples from patients for whom we already had a sample. All clinical information for every enrolled patient was entered — in an anonymous, coded format — and tracked in a custom online database (designed by SAM Solutions) that is fully compliant with the Health Insurance Portability and Accountability Act of 1996. A total of 1152 DNA samples were of sufficient quality to be genotyped. These were from 824 whites, 87 Hispanics, 35 blacks, 8 Asians, 3 American Indians, 3 Pacific Islanders, and 192 persons of unknown ethnic group (as self-reported in the presence of the physician; see the Supplementary Appendix for details). Of these patients, 692 were men and 460 were women, with a mean age of 59 years and a mean Amyotrophic Lateral Sclerosis Functional

Rating Scale–Revised score of 30.37 (range, 4.00 to 48.00, with higher scores indicating less severe disease (see Table 1 in the Supplementary Appendix for more clinical details).

We divided these 1152 patients into a discovery series and an independent replication series (replication series 1). The discovery series was made up of 386 white patients and 542 controls who all were white, older than 65 years of age, and neurologically normal on clinical assessment. There was no evidence of population stratification in the control group.<sup>6</sup> Replication series 1 consisted of 766 patients, 308 of whom were women and 458 of whom were men; there were 438 whites, 136 non-whites, and 192 patients of unknown ancestral origin. The mean age was 59 years and the mean Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised score was 29.94. The 750 control DNA samples for series 1 were obtained from neurologically normal elderly persons (353 white women and 397 white men, with a mean age of 66.1 years) and were purchased from the Rutgers University Cell and DNA Repository. For the second independent validation series (replication series 2), we obtained data from Schymick et al.<sup>7</sup> (see below).

#### STUDY DESIGN AND GENOTYPING

We first carried out a pooled analysis of DNA samples extracted from the blood or cell lines from each of the 386 white patients with sporadic ALS and 542 white controls in the discovery series. (Details about DNA extraction are given in the Supplementary Appendix.) We divided the case subgroup of 386 patients in half and pooled the DNA samples in each, ensuring that DNA was present in equimolar amounts in each pair of pooled samples. We prepared three pairs of pooled samples, for a total of six independently created pooled samples. A pooled DNA sample for the 542 controls was also created in triplicate, to control for pipetting errors.

Each of the pooled samples was hybridized to three GeneChip Human Mapping 500K Array Sets (Affymetrix) and two Infinium II HumanHap300 Genotyping BeadChip arrays (Illumina) according to the manufacturers' protocols for genotyping individual DNA samples, yielding a total of 27 Affymetrix arrays and 18 Illumina arrays. The Illumina chip is made up of probes that query HapMap-defined tag single-nucleotide polymorphisms (SNPs), whereas the Affymetrix platform

has probes that query relatively evenly spaced SNPs. Using these two platforms, we genotyped 766,955 unique SNPs, with an average intermarker distance of 3.9 kb.

We ranked the SNPs from the Affymetrix arrays and from the Illumina arrays according to the P values for the minor allele frequency for each SNP in the case group as compared with the control group, with the most significant P value corresponding to the highest rank and a cutoff P value of 0.05.<sup>8</sup> The top-ranked 192 SNPs, according to P value, from each genotyping platform (a total of 384 SNPs) were selected for validation in replication series 1 (Fig. 1 in the Supplementary Appendix). For each of the 384 validation SNPs, we selected and genotyped in replication series 1 an additional linked SNP — that is, one that resides in the same haplotype block defined by the HapMap CEU data (from persons of Northern and Western European ancestry), as indicated by strong linkage disequilibrium ( $r^2 > 0.8$ ) — thereby protecting against errors in genotyping or assay failure at any one locus, as well as increasing the odds of having at least one informative SNP in the ethnically diverse subgroup of replication series 1. Thus, 2 SNPs per associated locus (768 SNPs in total) from the initial genome screen were tested in replication series 1. Genotyping of these SNPs was contracted to KBiosciences, which uses a proprietary variation of primer extension referred to as KASPar.

Schymick et al.<sup>7</sup> have made their data set (created with the use of the Infinium II 550K platform [Illumina]) publicly available at [https://queue.coriell.org/Q/snp\\_index.asp](https://queue.coriell.org/Q/snp_index.asp). They genotyped DNA samples from white patients with sporadic ALS and white controls, all residing in the United States, from the Coriell Cell Repositories. There was no evidence for population stratification in this series, on the basis of analysis with the use of STRUCTURE software. We created our replication series 2 using SNP genotype data from the 135 patients with sporadic ALS who were unique to this data set (after removing Coriell samples that were already represented in our replication series 1), as well as 275 unique white controls.

To confirm the presence of the FLJ10986 protein in patients with sporadic ALS, we immunoprecipitated protein from cerebrospinal fluid, using anti-FLJ10986 antibody, and analyzed the immunoprecipitate on a sodium dodecyl sulfate polyacrylamide gel. The 45-kDa and 48-kDa bands

**Table 1. Single-Nucleotide Polymorphisms (SNPs) Significantly Associated with Sporadic Amyotrophic Lateral Sclerosis (ALS)\***

dbSNP No.†	Chromosome	Base Pair	Minor Allele	Replication Series 1				Replication Series 2‡				
				MAF in Controls (N=750)	MAF in Case Patients (N=766)	P Value for Whites	P Value for Nonwhites	Overall P Value	Gene or EST§	Candidate Mechanism of ALS	SNP	P Value
<b>rs6690993</b>	<b>1</b>	<b>59,416,003</b>	<b>G</b>	<b>0.316</b>	<b>0.389</b>	<b>3.0×10<sup>-4</sup></b>	<b>0.11</b>	<b>2.0×10<sup>-4</sup></b>	<b>0.048</b>	<b>rs6587852</b>	<b>0.048</b>	<b>59,417,468</b>
<b>rs6700125</b>	<b>1</b>	<b>59,414,818</b>	<b>T</b>	<b>0.319</b>	<b>0.405</b>	<b>6.0×10<sup>-4</sup></b>	<b>6.0×10<sup>-4</sup></b>	<b>1.8×10<sup>-5</sup></b>	<b>0.043</b>	<b>rs7531917</b>	<b>0.043</b>	<b>59,406,632</b>
rs7074175	10	20,556,984	T	0.601	0.667	0.002	0.16	0.001	0.08	rs878584	0.08	20,550,170
rs4827700	X	145,052,081	G	0.567	0.634	0.003	0.06	0.001				
rs6036180	20	22,627,977	A	0.542	0.611	0.003	0.03	0.001	0.37	rs6082791	0.37	22,616,509
<b>rs2836061</b>	<b>21</b>	<b>38,247,104</b>	<b>C</b>	<b>0.805</b>	<b>0.839</b>	<b>0.003</b>	<b>0.02</b>	<b>0.03</b>	<b>0.048</b>	<b>rs2836072</b>	<b>0.048</b>	<b>38,263,430</b>
rs2279605	15	55,611,622	A	0.595	0.648	0.004	0.86	0.008	0.07	rs16977585	0.07	55,618,699
rs4756063	11	33,822,142	G	0.693	0.734	0.005	0.04	0.03	0.30	rs4756063	0.30	25,835,257
rs11018623	11	88,837,360	G	0.510	0.567	0.005	0.52	0.006	0.11	rs10501705	0.11	88,827,923
rs4629724	6	121,250,591	T	0.337	0.409	0.005	1.1×10 <sup>-5</sup>	3.0×10 <sup>-4</sup>	0.23	rs10499101	0.23	121,262,670
rs4704336	5	75,899,375	G	0.466	0.537	0.006	4.0×10 <sup>-4</sup>	7.0×10 <sup>-4</sup>	0.14	rs10942784	0.14	75,889,806
rs5970919	X	22,639,221	A	0.581	0.636	0.006	0.51	0.008				
rs5929816	X	136,099,981	A	0.208	0.263	0.007	0.007	0.002				
rs2279607	15	55,611,764	T	0.596	0.642	0.009	0.49	0.02	0.07	rs16977585	0.07	55,618,699
rs7003876	8	1,135,748	T	0.474	0.533	0.009	0.08	0.004	0.12	rs11986875	0.12	1,136,821
<b>rs988213</b>	<b>18</b>	<b>42,378,965</b>	<b>A</b>	<b>0.558</b>	<b>0.604</b>	<b>0.009</b>	<b>0.41</b>	<b>0.02</b>	<b>0.03</b>	<b>rs4614822</b>	<b>0.03</b>	<b>42,360,329</b>
rs2036535	17	28,775,126	T	0.472	0.516	0.011	0.34	0.03		None		
rs5925683	X	22,629,374	C	0.588	0.638	0.011	0.69	0.01				
rs10499100	6	121,250,044	T	0.338	0.404	0.013	3.4×10 <sup>-5</sup>	0.001	0.23	rs10499101	0.23	121,262,670
rs1172149	1	201,956,415	T	0.372	0.426	0.014	0.10	0.007	0.37	rs1172161	0.37	201,938,152
rs3810715	X	150,555,188	G	0.881	0.912	0.015	0.47	0.01				
<b>rs13036957</b>	<b>20</b>	<b>41,255,110</b>	<b>G</b>	<b>0.749</b>	<b>0.795</b>	<b>0.016</b>	<b>0.18</b>	<b>0.01</b>	<b>0.002</b>	<b>rs13036957</b>	<b>0.002</b>	<b>41,255,110</b>
rs752257	20	22,630,289	G	0.558	0.614	0.019	0.04	0.007	0.37	rs6082791	0.37	22,616,509
<b>rs17027230</b>	<b>2</b>	<b>102,537,848</b>	<b>C</b>	<b>0.697</b>	<b>0.745</b>	<b>0.03</b>	<b>0.03</b>	<b>0.009</b>	<b>0.02</b>	<b>rs11690532</b>	<b>0.02</b>	<b>102,534,944</b>
<b>rs757863</b>	<b>7</b>	<b>77,316,032</b>	<b>A</b>	<b>0.537</b>	<b>0.587</b>	<b>0.03</b>	<b>0.16</b>	<b>0.02</b>	<b>0.03</b>	<b>rs3807797</b>	<b>0.03</b>	<b>77,303,892</b>

rs10740320	10	70,840,449	G	0.538	0.579	0.03	0.64	0.048	TACR2	Unknown	rs5030922	0.29	70,852,424
rs4263905	X	145,052,983	T	0.570	0.618	0.03	0.32	0.02					
rs10942784	5	75,889,806	A	0.448	0.510	0.03	4.8×10 <sup>-5</sup>	0.003	IQGAP2	Cytoskeletal regulation	rs10942784	0.14	75,889,806
rs10809959	9	13,497,924	C	0.375	0.430	0.03	0.007	0.007			rs4741304	0.23	13,478,608
rs10762294	10	70,840,387	C	0.537	0.577	0.03	0.81	0.048	TACR2	Unknown	rs5030922	0.29	70,852,424
rs1466471	8	61,478,245	G	0.897	0.929	0.03	0.01	0.008			rs10957133	0.11	61,463,064
rs3744477	17	40,183,199	T	0.725	0.763	0.03	0.80	0.04	DBF4B	Unknown	None		
rs10748358	12	42,149,850	T	0.663	0.712	0.03	0.04	0.01	ADAMTS20	Cytoskeletal regulation	rs11182091	0.06	42,150,040
rs12119273	1	61,655,314	G	0.972	0.985	0.03	0.85	0.04	NFIA	Unknown	rs7540743	0.18	61,657,202
rs10834819	11	25,821,137	G	0.655	0.698	0.03	0.37	0.03			rs2033979	0.13	25,835,257
rs10506228	12	42,150,219	T	0.663	0.711	0.03	0.049	0.01	ADAMTS20	Cytoskeletal regulation	rs11182091	0.06	42,150,040
rs12995017	2	205,046,522	A	0.853	0.887	0.04	0.09	0.02			rs6750362	0.17	205,044,709
rs945699	1	224,400,054	G	0.234	0.293	0.04	2.0×10 <sup>-8</sup>	0.001	WNT9A	Unknown	rs4653525	0.25	224,395,006
rs1554914	X	150,549,225	T	0.760	0.797	0.04	0.50	0.03					
rs4287603	17	2,722,492	G	0.486	0.541	0.04	0.003	0.008	GARNL4	Unknown	None		
rs1027615	12	41,998,556	A	0.660	0.713	0.04	4.0×10 <sup>-4</sup>	0.006			rs905079	0.003	41,995,086
rs666481	18	10,010,682	C	0.802	0.840	0.04	0.05	0.02			rs571890	0.15	10,016,087
rs1447830	3	74,695,861	C	0.817	0.853	0.04	0.07	0.02			rs4677414	0.36	74,689,710
rs12473579	2	203,030,073	G	0.630	0.683	0.04	8.0×10 <sup>-4</sup>	0.007			rs6715945	0.004	203,048,801
rs905080	12	41,995,195	G	0.659	0.712	0.04	7.0×10 <sup>-4</sup>	0.006			rs905079	0.003	41,995,086
rs2205545	X	150,677,351	A	0.336	0.380	0.04	0.22	0.03					
rs3771150	2	102,519,369	C	0.701	0.747	0.04	0.01	0.01	IL18RAP	Neuroinflammation	rs11690532	0.02	102,534,944
rs1891592	1	148,367,576	A	0.297	0.342	0.046	0.046	0.02	TUFT1	Unknown	rs7554707	0.08	148,354,788
rs3749870	6	155,646,464	G	0.498	0.551	0.046	0.003	0.01	TIAM2	Neurodevelopment	rs2882936	0.11	155,648,582
rs12279181	11	25,819,399	A	0.656	0.696	0.046	0.52	0.04			rs2033979	0.13	25,835,257
rs11172457	12	56,752,884	G	0.359	0.409	0.0498	0.004	0.02			rs1506884	0.59	56,725,220
rs733281	20	41,264,461	T	0.742	0.777	0.05	0.52	0.048	PTPRT	Cytoskeletal regulation	rs13036957	0.002	41,255,110
rs4819840	22	18,096,320	A	0.272	0.317	0.063	0.01	0.02			rs9618678	0.047	18,099,167
rs4491817	22	18,097,369	G	0.274	0.321	0.06	0.003	0.01			rs9618678	0.047	18,099,167

Table 1. (Continued.)		Replication Series 1										Replication Series 2†			
		Chromosome No.‡	Chromosome	Base Pair	Minor Allele	MAF in Controls (N = 750)	MAF in Case Patients (N = 766)	P Value for Whites	P Value for Nonwhites	Overall P Value	Gene or EST§	Candidate Mechanism of ALS	SNP	P Value	Base Pair
rs1314625	18	26,844,530	C	0.217	0.252	0.06	0.22	0.04	DSC3	Cytoskeletal regulation	rs17799201	0.046	26,864,504		
rs4516412	2	203,029,371	G	0.631	0.681	0.06	$8.0 \times 10^{-4}$	0.01	DSC3	Cytoskeletal regulation	rs6715945	0.004	203,048,801		
rs879012	20	957,788	G	0.281	0.323	0.07	0.02	0.03	PARP8	Unknown	rs11699432	0.048	965,232		
rs27628	5	50,266,128	T	0.279	0.320	0.08	0.009	0.03	PARP8	Unknown	rs1354171	0.04	50,267,305		
rs276915	18	26,853,979	A	0.405	0.447	0.09	0.08	0.04	DSC3	Cytoskeletal regulation	rs17799201	0.046	26,864,504		
rs38271	7	14,080,271	C	0.410	0.454	0.09	0.01	0.03	DGKB	Unknown	rs12537544	0.02	14,061,813		
rs276916	18	26,854,159	C	0.409	0.450	0.10	0.06	0.045	DSC3	Cytoskeletal regulation	rs17799201	0.046	26,864,504		
rs772593	6	106,451,750	T	0.068	0.092	0.10	0.005	0.03	NELL1	Unknown	rs772593	0.02	106,451,750		
rs7937375	11	21,698,795	A	0.423	0.467	0.12	0.004	0.03	NELL1	Unknown	rs1945359	0.03	21,695,027		
rs27248	5	50,268,304	A	0.274	0.311	0.13	0.02	0.049	PARP8	Unknown	rs1354171	0.04	50,267,305		
rs4622670	2	29,357,853	G	0.634	0.677	0.16	$2.0 \times 10^{-4}$	0.03	ALK	Neurodevelopment	rs4666184	0.006	29,367,475		
rs7818421	8	8,328,291	C	0.024	0.039	0.33	$1.02 \times 10^{-6}$	0.045			rs12680238	0.001	8,308,793		

\* Shown are all SNPs from the genomewide screen that were significantly associated ( $P < 0.05$ ) with sporadic ALS. Bold rows indicate SNPs that were significantly associated with ALS in the discovery series as well as replication series 1 and 2. The P values, calculated with the use of the chi-square test, are given for the minor allele frequency (MAF) of each SNP for the white case patients as compared with the controls, all of whom were white (P value for whites), for the nonwhite case patients as compared with the controls (P value for nonwhites), and for all case patients as compared with the controls (overall P value). SNPs are listed in order of increasing P values for whites.

† Replication series 2 was assembled from an independent study of patients with sporadic ALS versus controls.<sup>7</sup> The locus in replication series 2 corresponding to each SNP in replication series 1 was defined as the sequence within 25 kb on either side of the series 1 SNP. The SNP given for replication series 2 is the SNP with the most significant P value for the given locus.

‡ The dbSNP number is the reference number from the SNP database of the National Center for Biotechnology Information.

§ The gene or expressed-sequence tag (EST) is listed if one was annotated within 50 kb on either side of the given SNP.

were excised from the gel, the proteins were eluted and digested with trypsin, and the resultant fragments were sequenced. (Details about Western blotting and other assays are given in the Supplementary Appendix.)

#### STATISTICAL ANALYSIS

Bonferroni correction for multiple testing was performed, with rs6700125 being the sole SNP to retain a significant association in replication series 1. Our method of assessing significance is dependent on defining association signals that rank highest across the genome (given the high-density coverage of the genome) and replicate across two or more independent cohorts. Because both the Affymetrix and Illumina platforms were used for our discovery series and replication series 1, many of our top-ranked SNPs were not analyzed in replication series 2. We therefore used a locus-specific validation method: we identified the SNPs found by Schymick et al. on the Infinium II 550K array that were also present within 25 kb on either side of each of our top-ranked loci and then calculated P values for the minor allele frequency for each SNP in the case group as compared with the control group in replication series 2, reporting the most significant P value (Table 1). Odds ratios were calculated by means of the DeFinetti program (<http://ihg.gsf.de/cgi-bin/hw/hwa1.pl>), and methods were adapted from Sasieni.<sup>9</sup> Further details are given in the Supplementary Appendix. P values less than 0.05 were considered to indicate statistical significance.

---

## RESULTS

---

#### GENOMEWIDE ASSOCIATIONS

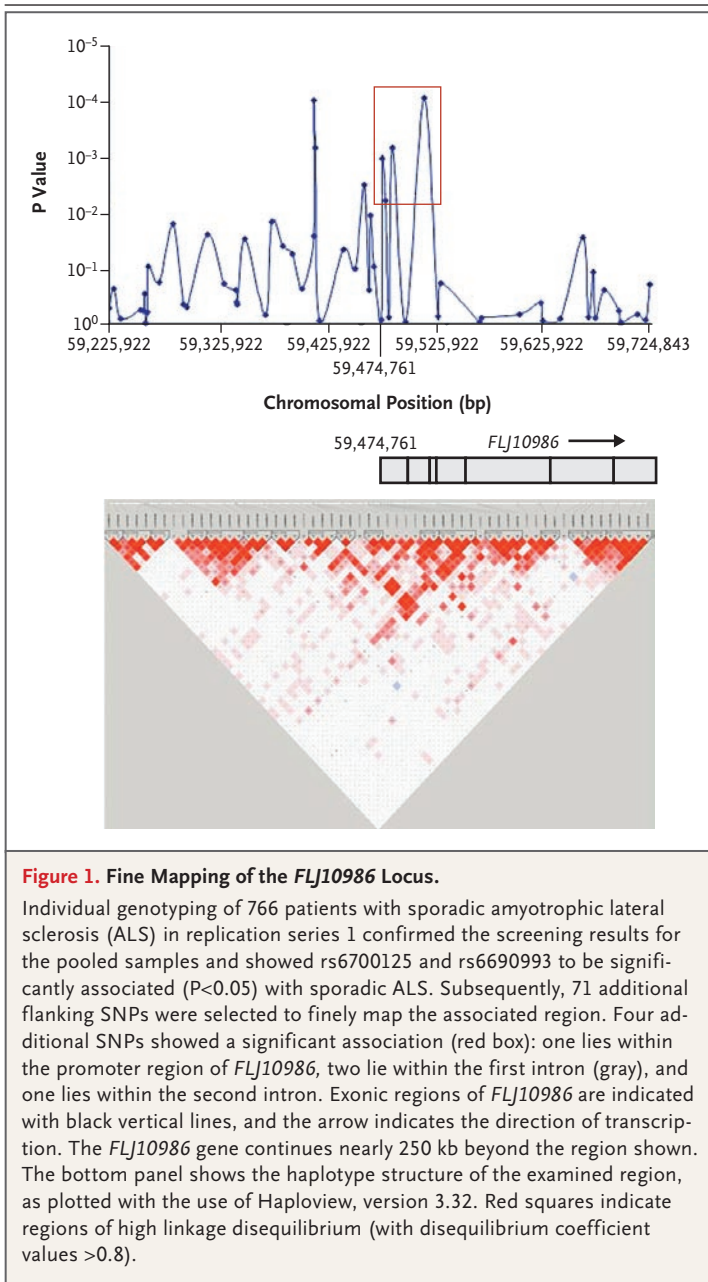
There was no population stratification in the discovery series (Fig. 2 in the Supplementary Appendix). Genotypes from a screen of the 386 patients with sporadic ALS (155 white women and 231 white men, with a mean age of 59 years and a mean Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised score of 30.80) were ranked in comparison with the genotypes of the 542 controls (all white, with a mean age of 68 years and equal numbers of men and women), with the highest rank assigned to the SNP with most significant P value. The rank-ordered SNPs from both platforms are listed in Tables 2 and 3 in the Supplementary Appendix.

#### VALIDATION OF SIGNIFICANT ASSOCIATIONS

Individual genotype data for 768 SNPs in replication series 1 showed a significant overall association of 66 SNPs with sporadic ALS, representing 51 unique loci (Table 1). Numerous loci had significant associations for both tag SNPs, suggesting that error in genotyping did not contribute to false positive associations. Of the 51 loci, 28 are intragenic or within approximately 50 kb upstream or downstream of annotated genes. The remaining 23 loci are not associated with a known gene within 50 kb upstream or downstream of the SNP. Of the 28 annotated genes, 9 have functions related to cytoskeletal regulation or neurodevelopment, suggesting that differences in these processes underlie predisposition to sporadic ALS.

We also found no population stratification in replication series 2 (Fig. 2 in the Supplementary Appendix). The most significant P value for each locus, the SNP in replication series 2 associated with that P value, and the chromosomal position of that SNP within the locus are listed in Table 1. The results show that there are 10 loci significantly associated with sporadic ALS among white members of all three independent series. An additional 41 loci are significantly associated in whites in two of the three series.

*FLJ10986* was found to be the gene most significantly associated with sporadic ALS, without a specific association with any clinical subclass (see below), suggesting that it is an early and common predisposition gene for the disease, independent of clinical course. There is different haplotype-block structure between whites and persons of other ancestries in replication series 1, as evidenced by the P values for the *FLJ10986* SNP rs6700125, which are the same for both the white case patients and the nonwhite case patients (87 Hispanics, 35 blacks, 8 Asians, 3 American Indians, 3 Pacific Islanders, and 192 for whom ethnic group is unknown). In contrast, the second SNP in this gene, rs6690993, which is in strong linkage disequilibrium with rs6700125 ( $r^2 > 0.8$ ) and is significantly associated with disease ( $P = 3.0 \times 10^{-4}$ ) among whites in replication series 1, is not associated with disease ( $P = 0.11$ ) in the group of nonwhite case patients in replication series 1 (Table 1). This indicates a difference in allele frequency and haplotype-block structure between the two groups and underscores the wisdom of using more than one SNP per locus until the Hap-



**Figure 1. Fine Mapping of the *FLJ10986* Locus.**

Individual genotyping of 766 patients with sporadic amyotrophic lateral sclerosis (ALS) in replication series 1 confirmed the screening results for the pooled samples and showed rs6700125 and rs6690993 to be significantly associated ( $P < 0.05$ ) with sporadic ALS. Subsequently, 71 additional flanking SNPs were selected to finely map the associated region. Four additional SNPs showed a significant association (red box): one lies within the promoter region of *FLJ10986*, two lie within the first intron (gray), and one lies within the second intron. Exonic regions of *FLJ10986* are indicated with black vertical lines, and the arrow indicates the direction of transcription. The *FLJ10986* gene continues nearly 250 kb beyond the region shown. The bottom panel shows the haplotype structure of the examined region, as plotted with the use of Haploview, version 3.32. Red squares indicate regions of high linkage disequilibrium (with disequilibrium coefficient values  $> 0.8$ ).

Map database is populated with data from all ancestral and admixed populations.

Complete odds-ratio calculations for the significant SNPs from replication series 1 (Table 1), as well as all other SNPs, are presented in Table 4 in the Supplementary Appendix. The genes (and specific SNPs) in white case patients that were significantly associated with sporadic ALS in both replication series were *FLJ10986* (rs6700125: odds ratio for having the genotype in patients vs. con-

trols, 1.38; 95% confidence interval [CI], 1.16 to 1.65; rs6690993: odds ratio, 1.35; 95% CI, 1.13 to 1.62), *PTPRT* (rs13036957: odds ratio, 1.28; 95% CI, 1.04 to 1.56), *IL18RAP* (rs3771150; odds ratio, 1.21; 95% CI, 1.00 to 1.46), *MAGI2* (rs757863: odds ratio, 1.23; 95% CI, 1.04 to 1.46), and *LOXHD1* (rs988213: odds ratio, 1.31; 95% CI, 1.10 to 1.55). An additional five chromosomal loci for which no gene has been annotated were significantly associated with the disease in both replication series: 12q12 (rs1027615: odds ratio, 1.18; 95% CI, 0.98 to 1.41), 2q33.1 (rs12473579: odds ratio, 1.29; 95% CI, 1.09 to 1.53), 2q12.1 (rs17027230: odds ratio, 1.23; 95% CI, 1.02 to 1.48), 21q22.13 (rs2836061: odds ratio, 1.41; 95% CI, 1.13 to 1.77), and 12q12 (rs905080: odds ratio, 1.18; 95% CI, 0.98 to 1.41).

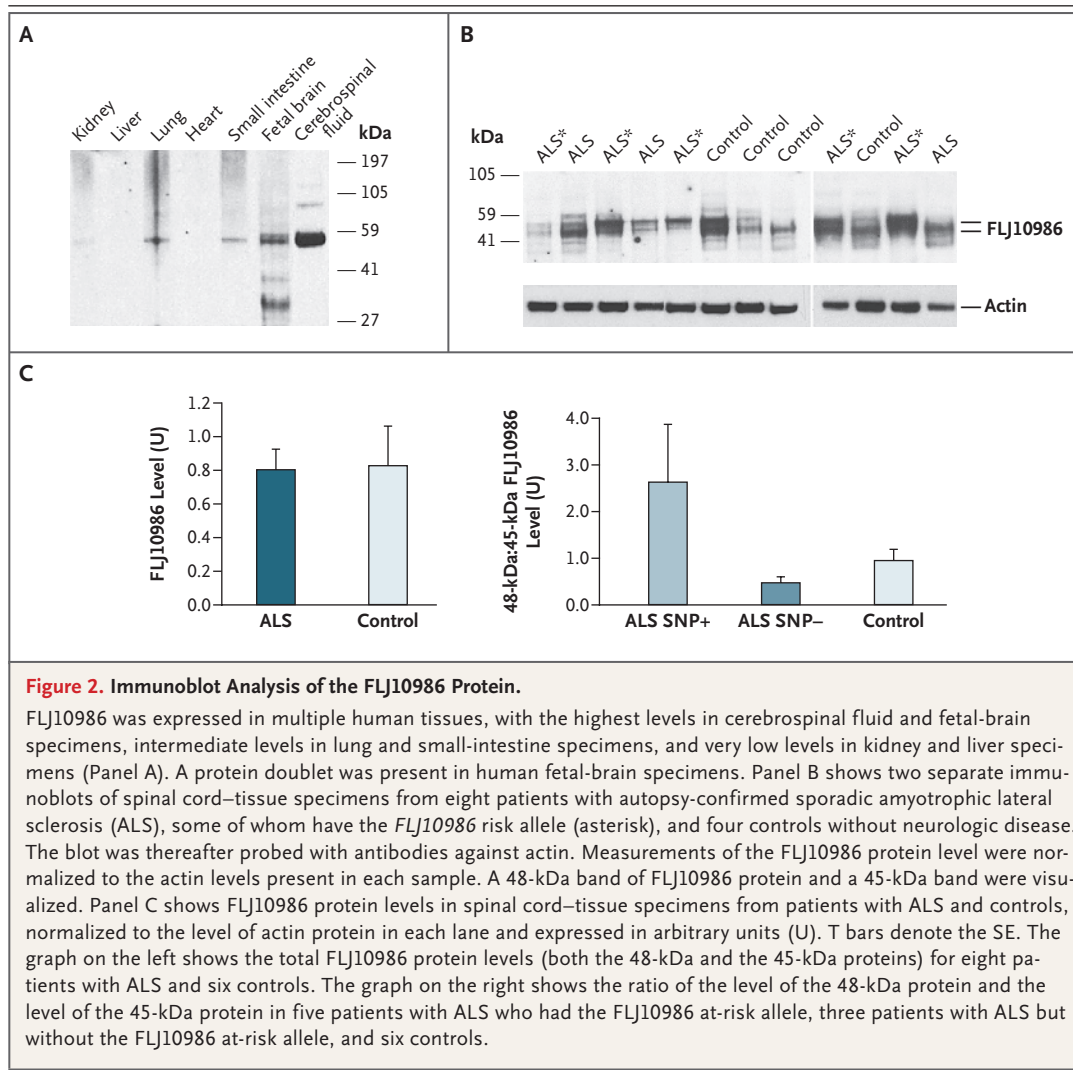
The most significant associations in our analyses of whites were with rs6700125 ( $P = 6.0 \times 10^{-4}$ ) and rs6690993 ( $P = 3.0 \times 10^{-4}$ ), which lie approximately 60 kb upstream of the uncharacterized gene *FLJ10986* (Table 1). To confirm the association of this gene with sporadic ALS, we genotyped 71 additional flanking SNPs representing haplotype blocks defined by the HapMap data for whites from replication series 1 that spanned a total of 500 kb across the locus (Fig. 1). We found four additional SNPs that were significantly associated with disease (rs10493256,  $P = 0.003$ ; rs6587852,  $P = 0.001$ ; rs1470407,  $P = 7.0 \times 10^{-4}$ ; rs333662,  $P = 9.0 \times 10^{-5}$ ) and that lie in the promoter region or the first two exons or introns of the *FLJ10986* gene.

#### EXPRESSION OF THE *FLJ10986* PROTEIN

The predicted molecular mass of the *FLJ10986* protein is 48 kDa, and we found a protein of this approximate size in kidney, lung, and small-intestine specimens of unaffected persons, with a lower level of expression in the liver (Fig. 2). A protein doublet of approximately 48 and 50 kDa was found in human fetal brain specimens, along with species of lower molecular weight. Intense *FLJ10986* immunoreactivity was also apparent in cerebrospinal fluid (Fig. 2A). Blotting with secondary antibody alone failed to detect these bands.

An *FLJ10986*-protein doublet of approximately 45 kDa and 48 kDa was evident in the spinal cord specimens of controls and patients with sporadic ALS (Fig. 2B). Each of these bands contained *FLJ10986* amino-acid sequences, indicating that they contained *FLJ10986* protein.

In exploratory analyses, although the amount of *FLJ10986* protein in spinal cord samples from



**Figure 2. Immunoblot Analysis of the FLJ10986 Protein.**

FLJ10986 was expressed in multiple human tissues, with the highest levels in cerebrospinal fluid and fetal-brain specimens, intermediate levels in lung and small-intestine specimens, and very low levels in kidney and liver specimens (Panel A). A protein doublet was present in human fetal-brain specimens. Panel B shows two separate immunoblots of spinal cord–tissue specimens from eight patients with autopsy-confirmed sporadic amyotrophic lateral sclerosis (ALS), some of whom have the *FLJ10986* risk allele (asterisk), and four controls without neurologic disease. The blot was thereafter probed with antibodies against actin. Measurements of the FLJ10986 protein level were normalized to the actin levels present in each sample. A 48-kDa band of FLJ10986 protein and a 45-kDa band were visualized. Panel C shows FLJ10986 protein levels in spinal cord–tissue specimens from patients with ALS and controls, normalized to the level of actin protein in each lane and expressed in arbitrary units (U). T bars denote the SE. The graph on the left shows the total FLJ10986 protein levels (both the 48-kDa and the 45-kDa proteins) for eight patients with ALS and six controls. The graph on the right shows the ratio of the level of the 48-kDa protein and the level of the 45-kDa protein in five patients with ALS who had the FLJ10986 at-risk allele, three patients with ALS but without the FLJ10986 at-risk allele, and six controls.

controls and patients with sporadic ALS was found to be equal when normalized to the level of actin present in each, the relative ratio of the levels of 48-kDa FLJ10986 and 45-kDa FLJ10986 was greatest in patients with sporadic ALS who carried the risk allele in the rs6700125 or rs6690993 FLJ10986 SNP (Fig. 2C) ( $P=0.049$ ). The ratios in controls and in patients who did not have a FLJ10986 risk genotype were similar.

#### CLINICAL SUBCLASSES OF SPORADIC ALS

Sporadic ALS is clinically heterogeneous, and thus genetic heterogeneity may underpin the disease subclasses. It is therefore likely that the relevance of overall P values for the association of SNPs with sporadic ALS is reduced in analyses of a genetically and clinically diverse series of patients. In post hoc

exploratory analyses, we performed association analyses of data for subgroups with early onset of sporadic ALS (at  $\leq 45$  years) as compared with late onset (at  $\geq 60$  years) (Table 2), male patients with sporadic ALS as compared with female patients (Table 3), and sporadic ALS with a bulbar onset as compared with a limb onset (Table 4).

Heterogeneity was evident; for example, there were significant differences in the frequencies of alleles in rs735888 between female patients with sporadic ALS and male patients ( $P=0.007$ ). Separate comparisons of female patients as compared with controls and male patients as compared with controls were performed to determine whether the SNP is specifically associated with sporadic ALS in either sex (Table 3). The SNP rs735888 was associated with sporadic ALS in women ( $P=0.02$ ) but

**Table 2. SNPs Significantly Associated with Early Onset (at ≤45 Years) of Sporadic Amyotrophic Lateral Sclerosis (ALS), as Compared with Late Onset (at ≥60 Years), in Replication Series 1.\***

dbSNP No.†	Chromosome	Base Pair	Gene‡	Overall P Value§		Late Onset (N = 289) vs. Controls (N = 750)			Early Onset (N = 61) vs. Controls (N = 750)			Early Onset (N = 61) vs. Late Onset (N = 289)		
				P Value	Minor Allele	Allelic Ratio	P Value	Minor Allele	Allelic Ratio	P Value	Minor Allele	Allelic Ratio	P Value	
rs12471471	2	213,848,557	ZNFN1A2	0.06	G	321:257, 765:675	0.33	A	76:46, 675:765	0.001	A	76:46, 257:321	3.0×10 <sup>-4</sup>	
rs7569588	2	45,331,732	NGNL6975	0.04	G	409:159, 1020:450	0.25	T	51:65, 450:1020	0.003	T	51:65, 159:409	7.0×10 <sup>-4</sup>	
rs12929266	16	49,453,731	BM924006	0.47	C	465:121, 1103:353	0.08	T	43:83, 353:1103	0.01	T	43:83, 121:465	0.001	
rs1390762	16	49,452,674	BM924006	0.28	G	464:120, 1126:350	0.12	A	43:85, 350:1126	0.01	A	43:85, 120:464	0.002	
rs11096490	2	17,949,476		0.16	A	544:42, 1369:111	0.79	G	20:106, 111:1369	0.001	G	20:106, 42:544	0.002	
rs4245528	6	106,480,927	PRDM1	0.14	A	20:576, 47:1435	0.83	A	12:114, 47:1435	3.0×10 <sup>-4</sup>	A	12:114, 20:576	0.002	
rs17118549	1	59,196,347	LOXHD1	0.02	T	1:589, 2:1480	0.85	T	3:123, 2:1480	1.4×10 <sup>-5</sup>	T	3:123, 1:589	0.003	
rs16983965	2	17,951,571		0.44	C	555:43, 1368:112	0.77	T	19:105, 112:1368	0.003	T	19:105, 43:555	0.003	
rs10438441	15	90,663,620	SLCO3A1	0.62	C	419:149, 1059:401	0.57	T	45:71, 401:1059	0.009	T	45:71, 149:419	0.006	
rs2919708	11	70,660,625	DA629334	0.17	A	507:81, 1232:212	0.60	G	28:92, 212:1232	0.01	G	28:92, 81:507	0.008	
rs11089823	22	35,833,678	IL2RB	0.16	T	463:129, 1141:337	0.62	C	41:87, 337:1141	0.02	C	41:87, 129:463	0.01	
rs38271	7	14,080,270	DGKB	0.03	C	248:332, 590:854	0.43	C	68:56, 590:854	0.003	C	68:56, 248:332	0.01	
rs838732	2	234,103,751	DGKD	0.33	C	239:351, 568:888	0.53	T	91:37, 888:568	0.02	T	91:37, 351:239	0.01	
rs2010435	11	82,528,143	PCF11	0.36	G	430:158, 1042:396	0.76	A	48:80, 396:1042	0.02	A	48:80, 158:430	0.02	
rs11233487	11	82,529,791	PCF11	0.10	A	165:421, 406:1064	0.81	A	49:77, 406:1064	0.007	A	49:77, 165:421	0.02	
rs7171883	15	90,664,487	SLCO3A1	0.67	C	270:308, 641:771	0.59	A	83:45, 771:641	0.03	A	83:45, 308:270	0.02	
rs2093689	13	94,150,134	SOX21	0.10	T	174:418, 423:1023	0.95	A	102:24, 1023:423	0.01	A	102:24, 418:174	0.02	
rs11914132	22	35,833,586	IL2RB	0.14	C	467:129, 1145:333	0.66	T	40:88, 333:1145	0.03	T	40:88, 129:467	0.02	
rs9558712	13	105,646,374		0.13	G	174:406, 442:1036	0.97	G	51:75, 442:1036	0.01	G	51:75, 174:406	0.02	
rs3020040	11	70,661,830	DA629334	0.21	G	508:78, 1247:207	0.59	A	26:96, 207:1247	0.03	A	26:96, 78:508	0.02	
rs838731	2	234,097,362	DGKD	0.22	T	233:343, 565:899	0.44	C	86:36, 899:565	0.047	C	86:36, 343:233	0.02	
rs11751085	6	155,653,676	TIAM2	0.02	C	303:271, 745:733	0.33	C	79:45, 745:733	0.004	C	79:45, 303:271	0.03	
rs10224956	7	32,969,593	PTHBI	0.06	A	495:93, 1144:318	0.002	G	29:91, 318:1144	0.54	G	29:91, 93:495	0.03	
rs3936139	19	2,538,575	GNG7	0.32	T	203:391, 430:1054	0.02	C	97:31, 1054:430	0.25	C	97:31, 391:203	0.03	
rs7467398	9	7,392,207		0.86	A	169:425, 405:1075	0.62	G	102:24, 1075:405	0.04	G	102:24, 425:169	0.03	
rs6772591	3	171,997,451	EIF5A	0.11	A	121:449, 238:1184	0.02	G	108:16, 1184:238	0.27	G	108:16, 449:121	0.03	
rs13236414	7	32,969,673	PTHBI	0.05	T	498:96, 1138:322	0.003	G	30:96, 322:1138	0.65	G	30:96, 96:498	0.04	
rs1943934	18	69,938,052	FBXO15	0.67	A	78:516, 189:1303	0.77	A	25:99, 189:1303	0.02	A	25:99, 78:516	0.04	

\* Bold data correspond to SNPs that were significantly associated (P<0.05) with ALS for the given allele. Only associations that were significant for at least one allelic comparison for the given SNP are shown. Additional comparisons with data for the 750 controls in replication series 1 were performed to determine whether the difference was primarily driven by a particular subgroup. Subclass status was not known for some patients. The allelic ratio is the ratio of the minor allele frequency (MAF) and the major allele frequency, given for case patients as well as for controls. SNPs are listed in order of increasing P values for the association with early onset as compared with late onset.  
† The dbSNP number is the reference number from the SNP database of the National Center for Biotechnology Information.  
‡ The gene is listed if one was annotated within 50 kb on either side of the given SNP.  
§ The overall P values, calculated with the use of the chi-square test, are given for the MAF of each SNP for the 766 case patients as compared with the controls.

**Table 3. SNPs Significantly Associated with Female Patients with Sporadic Amyotrophic Lateral Sclerosis (ALS), as Compared with Male Patients, in Replication Series 1.\***

dbSNP No.†	Chromosome	Base Pair	Gene‡	Overall P Value§	Female Sex (N=136) vs. Controls (N=750)		Male Sex (N=215) vs. Controls (N=750)		Female Sex (N=136) vs. Male Sex (N=215)			
					Minor Allele	Allelic Ratio	P Value	Minor Allele	Allelic Ratio	P Value	Allelic Ratio	P Value
rs735888	2	47,111,675	TTC7A	0.53	G	116:154, 526:944	0.02	A	283:139, 944:526	G	117:155, 138:282	0.007
rs11775313	8	24,765,856	DA418610	0.17	G	85:183, 367:1101	0.02	T	330:98, 1101:367	G	86:184, 97:329	0.008
rs926105	8	24,776,395	DA418610	0.17	C	87:189, 364:1118	0.02	T	340:102, 1118:364	C	88:190, 101:339	0.01
rs735887	2	47,111,596	TTC7A	0.91	T	120:162, 525:953	0.02	G	295:147, 953:525	T	121:163, 146:294	0.01
rs11172457	12	56,752,883	KUB3	0.01	G	132:146, 527:951	2.0x10 <sup>-4</sup>	G	165:267, 527:951	G	133:147, 164:266	0.01
rs7943838	11	96,991,874		0.43	C	216:62, 1105:387	0.20	T	136:302, 387:1105	C	217:63, 301:135	0.01
rs648735	13	109,654,180	COL4A1	0.53	G	205:77, 959:525	0.009	A	162:278, 525:959	G	205:79, 278:160	0.02
rs4781528	16	13,673,792		0.29	A	112:156, 553:883	0.31	G	278:132, 883:553	A	112:158, 132:276	0.02
rs4722196	7	22,696,330	DA833177	0.08	A	140:136, 598:884	0.001	A	178:256, 598:884	A	140:138, 178:254	0.02
rs10499100	6	121,250,043	C6orf170	0.001	T	117:157, 505:973	0.007	T	147:277, 505:973	T	119:157, 145:277	0.02
rs11711863	3	185,808,656	EPHB3	0.25	T	70:208, 273:1193	0.01	C	352:76, 1193:273	T	70:210, 76:350	0.02
rs1326005	1	5,313,957		0.93	C	186:94, 886:588	0.047	T	186:254, 588:886	C	187:95, 253:185	0.02
rs11172427	12	56,707,319	KUB3	0.07	G	97:175, 388:1066	0.003	G	120:310, 388:1066	G	98:176, 119:309	0.03
rs9558712	13	105,646,374		0.13	G	100:176, 442:1036	0.04	A	305:125, 1036:442	G	102:176, 123:305	0.03
rs4334421	19	2,531,056	GNG7	0.81	G	88:188, 346:1040	0.02	A	315:103, 1040:346	G	89:189, 102:314	0.03
rs648705	13	109,654,153	COL4A1	0.57	C	198:78, 939:515	0.02	A	156:268, 515:939	C	198:80, 268:154	0.03
rs4629724	6	121,250,590	C6orf170	3.0x10 <sup>-4</sup>	T	118:160, 504:972	0.008	T	157:285, 504:972	T	120:160, 155:285	0.04
rs7082776	10	125,061,868	GPR26	0.03	A	182:84, 873:605	0.004	A	253:161, 873:605	A	184:84, 251:161	0.04
rs1565774	16	17,745,336	BX109972	0.12	T	87:197, 349:1131	0.01	T	105:331, 349:1131	T	88:198, 104:330	0.04
rs13315088	3	151,167,835	PFN2	0.79	A	196:82, 1026:464	0.59	G	159:279, 464:1026	A	198:82, 277:159	0.047

\* Bold data correspond to SNPs that were significantly associated (P<0.05) with ALS for the given allele. Only associations that were significant for at least one allelic comparison for the given SNP are shown. Additional comparisons with data for the 750 controls in replication series 1 were performed to determine whether the difference was primarily driven by a particular subgroup. Subclass status was not known for some patients. The allelic ratio is the ratio of the minor allele frequency (MAF) and the major allele frequency, given for case patients as well as for controls. SNPs are listed in order of increasing P values for the association with female sex as compared with male sex.

† The dbSNP number is the reference number from the SNP database of the National Center for Biotechnology Information.

‡ The gene is listed if one was annotated within 50 kb on either side of the given SNP.

§ The overall P values, calculated with the use of the chi-square test, are given for the MAF of each SNP for the 766 case patients as compared with the controls.

**Table 4. SNPs Significantly Associated with Bulbar Onset of Sporadic Amyotrophic Lateral Sclerosis (ALS), as Compared with Limb Onset, in Replication Series 1.\***

dbSNP No.†	Chromosome	Base Pair	Gene‡	Overall P Value§	Bulbar Onset (N = 119) vs. Controls (N = 750)			Limb Onset (N = 179) vs. Controls (N = 750)			Bulbar Onset (N = 119) vs. Limb Onset (N = 179)		
					Minor Allele	Allelic Ratio	P Value	Minor Allele	Allelic Ratio	P Value	Minor Allele	Allelic Ratio	P Value
rs12695988	3	154,604,997	BM927666	0.32	A	118:110, 649:827	0.03	C	280:178, 827:649	0.05	A	118:108, 178:280	9.0×10 <sup>-4</sup>
rs4680060	3	154,601,610	BM927666	0.38	T	119:117, 631:817	0.0497	A	298:188, 817:631	0.06	T	119:115, 188:298	0.002
rs988213	18	42,378,964	LOXHD1	0.02	G	114:118, 641:827	0.12	A	294:178, 827:641	0.02	G	112:118, 178:294	0.006
rs10884751	10	111,100,812		0.21	A	213:19, 1280:170	0.11	G	75:407, 170:1280	0.03	A	211:19, 407:75	0.007
rs7806370	7	38,461,063	AMPH	0.35	C	37:195, 141:1315	0.004	T	423:43, 1315:141	0.77	C	37:193, 43:423	0.008
rs6677714	1	236,530,180	BU173572	0.95	A	193:43, 1065:413	0.002	A	348:132, 1065:413	0.85	A	191:43, 348:132	0.008
rs2247691	18	41,199,732	DA290737	0.36	T	166:74, 907:543	0.049	C	195:285, 543:907	0.21	T	165:73, 285:195	0.009
rs11233487	11	82,529,791	PCF11	0.10	T	179:55, 1064:406	0.19	A	159:321, 406:1064	0.02	T	177:55, 321:159	0.01
rs17667053	16	70,704,931	DHX38	0.96	C	7:229, 29:1449	0.32	G	477:3, 1449:29	0.04	C	7:227, 3:477	0.01
rs7193888	16	82,653,630	MBTPS1	0.53	T	86:152, 392:1006	0.01	C	352:132, 1006:392	0.75	T	86:150, 132:352	0.01
rs27628	5	50,266,127	PARP8	0.03	T	89:151, 412:1066	0.004	T	135:347, 412:1066	0.96	T	87:151, 135:347	0.02
rs27248	5	50,268,303	PARP8	0.049	A	86:152, 405:1069	0.006	G	349:131, 1069:405	0.94	A	84:152, 131:349	0.02
rs17741655	2	127,147,541	GYPC	0.11	G	181:57, 1007:471	0.01	A	156:328, 471:1007	0.88	G	179:57, 328:156	0.03
rs4745434	9	75,515,725	PCSK5	0.37	T	106:126, 530:936	0.005	T	178:296, 530:936	0.58	T	106:124, 178:296	0.03
rs13398914	2	127,152,871	GYPC	0.09	A	181:57, 1004:466	0.02	A	320:148, 1004:466	0.98	A	179:57, 320:148	0.04
rs7740727	6	5,654,334	FARS2	0.02	G	100:136, 622:854	0.95	G	242:238, 622:854	0.002	A	135:99, 238:242	0.04
rs11711863	3	185,808,656	EPHB3	0.25	C	191:39, 1193:273	0.54	T	109:369, 273:1193	0.046	C	191:37, 369:109	0.04
rs3944131	4	92,386,146	AB051467	0.18	C	140:92, 859:595	0.72	C	321:151, 859:595	5.0×10 <sup>-4</sup>	T	91:139, 151:321	0.048

\* Bold data correspond to SNPs that were significantly associated (P<0.05) with ALS for the given allele. Only associations that were significant for at least one allelic comparison for the given SNP are shown. Additional comparisons with data for the 750 controls in replication series 1 were performed to determine whether the difference was primarily driven by a particular subgroup. Subclass status was not known for some patients. The allelic ratio is the ratio of the minor allele frequency (MAF) and the major allele frequency, given for case patients as well as for controls. SNPs are listed in order of increasing P values for the association with bulbar onset as compared with limb onset.

† The dbSNP number is the reference number from the SNP database of the National Center for Biotechnology Information.

‡ The gene is listed if one was annotated within 50 kb on either side of the given SNP.

§ The overall P values, calculated with the use of the chi-square test, are given for the MAF of each SNP for the 766 case patients as compared with the controls.

not in men ( $P=0.28$ ). Some of the overall associations that we found in replication series 1 (Table 1) were significant in certain subgroup comparisons and not in others, suggesting that clinical subclasses of sporadic ALS were driving the overall associations for these SNPs. Validation of genes underlying these loci may provide clues about the mechanisms of disease in each subgroup and may ultimately contribute to the development of genotype-specific therapeutic interventions.

---

## DISCUSSION

---

We calculated the odds ratios for SNPs using three independent series selected to contain only whites of European descent. We present  $P$  values for all case patients in our ancestrally diverse replication series 1, since there will not be stratification at all loci, and the data may therefore assist future efforts at replication. Our replication series 2 was considerably smaller than the other two series and was underpowered to detect subtle allelic associations. We believe that the associations in replication series 2 that were not significant should still be considered, with increased confidence placed in those that were significant.

Biologic factors implicated by our data include cytoskeletal regulation (Table 1), a finding that suggests that aspects of cytoskeletal dysfunction may be central to the initiation or progression of sporadic ALS. This idea is congruous with emerging models of neurodegeneration in sporadic ALS, Alzheimer's disease, and spinal muscular atrophy, in which the loss of synaptic efficacy, aberrant axon pruning, and concomitant "dying back" of the neurons from the synaptic sites inward toward the cell bodies are thought to be among the earliest pathologic events.<sup>10-12</sup> In particular, in considering the ongoing process of reorganizing the neuromuscular junction, the association of variants of anaplastic lymphoma kinase and nuclear factor 1 $\alpha$  are interesting given their roles in neurite outgrowth<sup>13</sup> and neuronal differentiation,<sup>14</sup> respectively. Anaplastic lymphoma kinase has recently been shown to be critical for pleiotrophin-mediated axonal regeneration in motor neurons in the spinal cord, is necessary for neuroprotection in response to glutamate excitotoxicity, and is not expressed in the spinal cord of patients with sporadic ALS.<sup>15</sup> Restoration of the function of anaplastic lymphoma kinase either directly, downstream in its intracellular signaling pathway, or

by increasing the amount of its ligand pleiotrophin could result in protective effects. Retinoic acid induces the expression of pleiotrophin and may therefore provide some therapeutic benefit to patients with sporadic ALS, particularly if administered early in the disease process — although this is speculative, at best, and the association between variants of anaplastic lymphoma kinase and sporadic ALS awaits replication and refinement by others. *NOX4* has been previously implicated in sporadic ALS,<sup>16</sup> which lends support not only to our finding of association between *NOX4* and disease but also, by extension, to the other genetic associations that we report (Table 1).

Little is known about the function of FLJ10986. However, about 80% of its 439 amino acids make up the FGGY family of carbohydrate kinase domains. These domains are found in a family of proteins including L-fucolokinase, gluconokinase, glycerol kinase, and xylulokinase, which phosphorylate fucose, gluconate, glycerol, and xylulose, respectively, and have roles in energy metabolism and glycolysis. The potential substrate or substrates for FLJ10986 are unknown, and how the activity of the protein may be relevant to the pathogenesis of sporadic ALS is also unclear. We found an FLJ10986-protein doublet in tissues of the nervous system. Further work is required to determine whether FLJ10986 is alternatively spliced or is subject to post-translational modifications within the nervous system and to confirm that the higher-molecular-weight species of FLJ10986 is more abundant than the lower-molecular-weight species in patients with sporadic ALS who have a SNP associated with the disease.

Our findings suggest that there is no single, overwhelming genetic association underlying sporadic ALS, and this is consistent with a model in which sporadic ALS results from a complex interplay of environmental factors and numerous low-risk susceptibility loci. Unraveling the network of causes will probably require substantial effort once the genes involved have been identified. However, the identification of the candidate susceptibility loci in this and other studies is an essential first step.

Supported by a grant from the Muscular Dystrophy Association Augie's Quest initiative and the Dorrance Family Foundation.

Dr. Levine reports receiving speaker fees from Eli Lilly, Boehringer Ingelheim, and GlaxoSmithKline; Dr. Bertorini, consulting fees from Serono, Pfizer, Teva, and Allergan; Dr. Graves, consulting fees from Avanir, Novartis, Sanofi-Aventis, the Muscular Dystrophy Association, the Guillain-Barré Syndrome Sup-

port Group, and Pharmacia; lecture fees from Avanir; and grant support from Pharmacia; Dr. Mozzafar, consulting fees from Celgene, Genzyme, and Allergen and lecture fees from Genzyme and Crescent Healthcare; Dr. Lomen-Hoerth, consulting fees from Kinemed, Neurological Biological Institute, Rinat, Pfizer, Alta Bates, Celgene, Allergen, and Columbia and grant support from Avanir Pharmaceuticals and holding equity inter-

ests in Hewlett-Packard; Dr. Mitsumoto, consulting fees from Eisai, lecture fees from Avanir, and grant support from Aventis, Aeolus, and Avanir; Dr. Miller, consulting fees from Celgene and Novartis; and Dr. Appel, consulting fees for his position on the medical advisory board of Vasogen and holding equity interests in Cyberomics. No other potential conflict of interest relevant to this article was reported.

#### REFERENCES

- Cleveland DW, Rothstein JD. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. *Nat Rev Neurosci* 2001;2:806-19.
- Pasinelli P, Brown RH. Molecular biology of amyotrophic lateral sclerosis: insights from genetics. *Nat Rev Neurosci* 2006;7:710-23.
- Ince PG. Neuropathology. In: Brown RH Jr, Meininger K, Swash M, eds. *Amyotrophic lateral sclerosis*. London: Martin Dunitz, 2000:83-112.
- Bruijn LI, Miller TM, Cleveland DW. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu Rev Neurosci* 2004;27:723-49.
- Brooks BR, Miller RG, Swash M, Mun-sat TL. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000;1:293-9.
- Reiman EM, Webster JA, Myers AJ, et al. GAB2 alleles modify Alzheimer's risk in APOE varepsilon4 carriers. *Neuron* 2007;54:713-20.
- Schymick JC, Scholz SW, Fung HC, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol* 2007;6:322-8.
- Pearson JV, Huentelman MJ, Halperin RF, et al. Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies. *Am J Hum Genet* 2007;80:126-39.
- Sasieni PD. From genotypes to genes: doubling the sample size. *Biometrics* 1997;53:1253-61.
- Coleman PD, Yao PJ. Synaptic slaughter in Alzheimer's disease. *Neurobiol Aging* 2003;24:1023-7.
- Frey D, Schneider C, Xu L, Borg J, Spoo-ren W, Caroni P. Early and selective loss of neuromuscular synapse subtypes with low sprouting competence in motoneuron disease. *J Neurosci* 2000;20:2534-42.
- Cifuentes-Diaz C, Nicole S, Velasco ME, et al. Neurofilament accumulation at the motor endplate and lack of axonal sprouting in a spinal muscular atrophy mouse model. *Hum Mol Genet* 2002;11:1439-47.
- Motegi A, Fujimoto J, Kotani M, Sakuraba H, Yamamoto T. ALK receptor tyrosine kinase promotes cell growth and neurite outgrowth. *J Cell Sci* 2004;117:3319-29.
- Wang W, Stock RE, Gronostajski RM, Wong YW, Schachner M, Kilpatrick DL. A role for nuclear factor 1 in the intrinsic control of cerebellar granule neuron gene expression. *J Biol Chem* 2004;279:53491-7.
- Mi R, Chen W, Höke A. Pleiotrophin is a neurotrophic factor for spinal motor neurons. *Proc Natl Acad Sci U S A* 2007;104:4664-9.
- Wu DC, Ré DB, Nagai M, Ischiropoulos H, Przedborski S. The inflammatory NADPH oxidase enzyme modulates motor neuron degeneration in amyotrophic lateral sclerosis mice. *Proc Natl Acad Sci U S A* 2006;103:12132-7.

Copyright © 2007 Massachusetts Medical Society.