ORIGINAL CONTRIBUTION

ONLINE FIRST

Resequencing of 29 Candidate Genes in Patients With Familial and Sporadic Amyotrophic Lateral Sclerosis

Hussein Daoud, PhD; Paul N. Valdmanis, PhD; Francois Gros-Louis, PhD; Véronique Belzil, MSc; Dan Spiegelman, MSc; Edouard Henrion, MSc; Ousmane Diallo, BSc; Anne Desjarlais, BSc; Julie Gauthier, PhD; William Camu, MD; Patrick A. Dion, PhD; Guy A. Rouleau, MD, PhD, FRCPC

Objective: To identify novel disease-causing genes for amyotrophic lateral sclerosis (ALS).

Design, Setting, and Patients: We carried out a systematic mutation screening of the entire coding regions of 29 candidate genes encoding critically important proteins for proper differentiation and development of corticospinal motor neurons in 190 patients with familial and sporadic ALS.

Main Outcome Measures: We focused our analysis on coding variants and evaluated the distribution of nonsynonymous and synonymous variants in our cohort of patients with ALS.

Results: We identified 40 novel nonsynonymous variants and showed a significant excess of unique nonsynonymous variants in our cohort of patients with ALS, which suggests the presence of ALS-predisposing mutations.

Conclusions: Using a multifaceted approach based on the functional prediction of missense variants, the conservation of the altered amino acid, and the cosegregation of the variants identified in familial cases, we identified several promising novel genes for ALS such as *LUM* and *CRYM*. We have also highlighted the analytical challenges of large-scale sequencing screens to detect disease-causing variants.

Arch Neurol. 2011;68(5):587-593. Published online January 10, 2011. doi:10.1001/archneurol.2010.351

Author Affiliations are listed at the end of this article.

AMYOTROPHIC LATERAL SCLEROSIS (ALS) is a fatal neurodegenerative disease characterized by the selective loss of motor neurons in the spinal cord and motor cortex, resulting in progressive muscle weakness and atrophy. It typically leads to death within 3 to 5 years following onset.[1]

Among cases of ALS, most are sporadic ALS (SALS), while approximately 5% of patients have a positive family history with clear mendelian inheritance (familial ALS [FALS]). Mutations in the superoxide dismutase 1 gene (*SOD1*) are the most common known cause of ALS. They account for 15% to 20% of cases of FALS, which is most frequently inherited in an autosomal dominant manner, and 1% to 2% of all cases.[2] While several mechanisms have been proposed to explain the toxic effects of mutant SOD1, the mechanism by which SOD1 leads to selective motor neuron death remains unclear.[3] More recently, mutations in *TARDBP* and *FUS*, encoding 2 multifunctional DNA or RNA binding proteins, were identified in patients with FALS[4-7] and were subsequently reported in patients with SALS.[8,9] Several other rare causative genes linked to FALS have been identified, including *ALS2*, *SETX*, *VAPB*, *ANG*,[10] and, more recently, *OPTN*.[11]

Although the discovery of these genes has led to significant new insights into the causes of ALS, the basic pathogenic mechanism and the genetic cause of most ALS cases remain unknown. This emphasizes our need to identify additional ALS-causing genes so that we can better understand the pathogenesis of this disease.

A set of genes specifically expressed at different stages of development in mouse corticospinal motor neurons has been identified, including 29 biologically relevant genes.[12] These genes are of particular interest as they might be involved in different aspects of corticospinal motor neuron development and are likely to govern the ability of developing brain cells to connect properly to the spinal cord, which may have significant relevance to ALS. Four of these 29 genes are located within or near previously identified loci for ALS and 1 of these (*NEFH*) has previously been implicated in ALS,[13,14] suggesting that these newly identified genes may be good candidates for ALS.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

Table 1. Characteristics of the Amyotrophic Lateral Sclerosis Cases Studied

Characteristic	Affected Participants (n=190)
Origin, No.	
France	90
Quebec, Canada	100
Sex, No.	
Male	104
Female	86
Onset site, No.[a]	
Bulbar	25
Spinal	107
Age at onset, mean (SD), y[a]	55.4 (13.1)
Duration, mean (SD), y[a]	3.4 (2.6)

[a] Information is not available for each patient.

Therefore, in an effort to identify novel ALS disease-causing genes, we have selected human orthologs for these 29 corticospinal motor neuron mouse genes and carried out a mutation screening of the entire coding regions by direct sequencing in 190 unrelated patients with ALS (80 with FALS and 110 with SALS).

METHODS

PATIENTS AND CONTROL SUBJECTS

A total of 190 patients with ALS (80 with FALS and 110 with SALS) and 190 unrelated, neurologically healthy individuals matched for age and ethnicity were included in this study. All individuals were recruited through clinics in France and Quebec, Canada. All patients were diagnosed with probable or definite ALS according to El Escorial criteria.[15] The clinical characteristics of this cohort are summarized in **Table 1**. All patients with FALS tested negative for mutations in *SOD1*, *TARDBP*, *FUS*, *VAPB*, and *ANG*. Informed written consent was obtained from each participant, and the study was approved by the ethics committees and institutional review boards of the relevant institutions. Blood samples were obtained from patients and control subjects, and genomic DNA was extracted from peripheral blood cells using standard methods.

AMPLICON DESIGN AND GENE SCREENING

We identified the full messenger RNA sequences of all potential transcriptional isoforms and determined the exon and intron structures based on alignment with the UCSC Genome Browser (Hg18 build; http://genome.ucsc.edu/) (**Table 2**). These genes have been classified into 5 different groups based on expression profiles suggestive of a specific role in corticospinal motor neuron development. Primers were designed using the ExonPrimer software (http://ihg2.helmholtz-muenchen.de/ihg/ExonPrimer.html) from the UCSC Genome Browser to amplify 200- to 700–base pair (bp) fragments that cover coding exons plus 50 bp from each of the flanking introns. Polymerase chain reaction (PCR) was performed in 384-well plates using 5 ng of genomic DNA and AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, California) as per the manufacturer's instructions. The PCR products were sequenced on a 3730XL DNA analyzer (Applied Biosystems). A fragment was considered successfully sequenced if the sequence quality of more than 90% of the traces was sufficient to permit analysis.

ANALYSIS OF SEQUENCE VARIANTS

Sequence variants were called automatically using PolyPhred version 5.04 (http://droog.gs.washington.edu/polyphred) and PolySCAN version 3.0 (http://genome.wustl.edu/tools/genome_center_software/polyscan). We verified all identified mutations by manual inspection of the trace files using Mutation Surveyor software (SoftGenetics, State College, Pennsylvania). All potential truncating variants were further investigated by resequencing the appropriate sample from a newly amplified PCR product using forward and reverse primers. For variants identified in FALS cases, we tested the cosegregation by sequencing the corresponding fragment in available additional family members. For novel nonsynonymous variants, we resequenced the corresponding fragment in 190 control subjects. Finally, the potential disruptiveness of missense mutations as well as the conservation of the altered amino acids were analyzed by bioinformatics programs PANTHER (http://www.pantherdb.org/tools/csnpScoreForm.jsp), PolyPhen (http://genetics.bwh.harvard.edu/pph/), and SIFT (http://sift.jcvi.org/www/SIFT_BLink_submit.html) (eAppendix, http://www.archneurol.com).

RESULTS

SCREENING OF CANDIDATE GENES

A total of 278 PCR primer pairs were designed to amplify the coding exons of the 29 genes included in this study. The PCR conditions could be optimized for 257 amplicons, corresponding to an average success rate of 92.45% (eTable 1). Given that the overall success rate was 92.45% and that we screened 15 168 codons (Table 2), we estimate the amount of coding sequence screened per sample to be 42 kilobases and the total amount of coding sequence screened in our ALS cohort to be approximately 8 megabases.

CODING VARIANTS IDENTIFIED IN THE SCREEN

We focused our analysis on coding variants because coding exons are believed to harbor a substantial amount of functional variations that cause mendelian diseases.[16] We identified a total of 142 different coding variants, of which 78 were synonymous changes and 64 were nonsynonymous changes (**Table 3** and eTable 2). No variants that affect highly conserved consensus splice sites and no insertions and deletions were found. Synonymous and nonsynonymous variants were found in approximately equal proportion, which is consistent with the genetic variations recently found following complete sequencing of 12 human exomes.[17]

NOVEL NONSYNONYMOUS VARIANTS IDENTIFIED

Among the 64 nonsynonymous variants identified, 24 were previously described in the dbSNP database (National Center for Biotechnology Information, Bethesda, Maryland) and were not considered further (eTable 2). The other 40 novel nonsynonymous variants are reported in **Table 4**. Of these, 10 were found in control subjects, suggesting that these variants are not associated with ALS. Among the remaining 30 variants, 21 have

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

Table 2. Genes Sequenced in the Study

Gene Type	Mouse Genes: Gene Name	Mouse Genes: GenBank Accession No.	Human Orthologs: Gene Name	Human Orthologs: GenBank Accession No.	Human Orthologs: Chromosome	Human Orthologs: Exon Count, No.	Human Orthologs: mRNA Length, bp	Human Orthologs: Protein Length, No. of Amino Acids
CSMN general identity genes	*riken B430320C24*	AK046712	*C5orf23*	NM_024563	5p13.3	1	2895	121
	fez	NM_080433	*FEZF2*	NM_018008	3p14.2	5	2095	459
	daf1	NM_010016	*CD55*	NM_000574	1q32	10	2812	381
	doc2b	NM_007873	*DOC2B*	NM_003585	17p13.3	9	2023	412
	contactin 6	NM_017383	*CNTN6*	NM_014461	3p26-p25	23	3530	1028
	cadherin 13	NM_019707	*CDH13*	NM_001257	16q24.2-q24.3	14	3842	713
CSMN early development genes	*sox5*	NM_011444	*SOX5*	NM_006940	12p12.1	15	4333	763
	ctip2	AF186019	*BCL11B*	NM_138576	14q32.2	4	7816	894
	clim1/ldb2	NM_010698	*LBD2*	NM_001290	4p16	8	2560	373
	riken 2010001O09	NM_025909	*OMA1*	NM_145243	1p32.2-p32.1	9	1928	524
	grb14	NM_016719	*GRB14*	NM_004490	2q22-q24	14	2402	540
	synaptotagmin 9	NM_021889	*SYT9*	NM_175733	11p15.4	7	3958	491
CSMN intermediate development genes	*neto1*	NM_144946	*NETO1*	NM_153181	18q22-q23	11	2328	532
	crim1	XM_128751	*CRIM1*	NM_016441	2p21	17	5601	1036
	lumican	NM_008524	*LUM*	NM_002345	12q21.3-q22	3	2116	338
	stk39	NM_016866	*STK39*	NM_013233	2q24.3	18	3288	545
	mu-crystallin	NM_016669	*CRYM*	NM_001888	16p13.11-p12.3	8	1303	314
	riken 2610024A01	XM_129450	*TMEM163*	NM_030923	2q21.3	8	1901	289
CSMN late development genes	*bcl6*	BC052315	*BCL6*	NM_001706	3q27	10	3537	706
	encephalopsin	AF140241	*OPN3*	NM_014322	1q43	4	2632	402
	cadherin 22	AB019618	*CDH22*	NM_021248	20q13.1	11	3675	828
	pcp4	NM_008791	*PCP4*	NM_006198	21q22.2	3	566	62
	netrin-G1	NM_030699	*NTNG1*	NM_014917	1p13.3	6	3015	438
CSMN-specific vs corticotectal neurons	*ramp3*	NM_019511	*RAMP3*	NM_005856	7p13-p12	3	1376	148
	igfbp4	NM_010517	*IGFBP4*	NM_001552	17q12-q21.1	4	2246	258
	neurofilament, heavy polypeptide	M35131	*NEFH*	NM_021076	22q12.2	4	3721	1020
	diap3	NM_019670	*DIAPH3*	NM_001042517	13q21.2	28	4812	1193
	itm2a	NM_008409	*ITM2a*	NM_004867	Xq13.3-Xq21.2	6	1851	263
	S100a10	NM_009112	*S100a10*	NM_002966	1q21	3	1069	97

Abbreviations: bp, base pairs; CSMN, corticospinal motor neuron; mRNA, messenger RNA.

Table 3. Summary of All Coding Sequence Variants Identified in Our Cohort of Patients With Amyotrophic Lateral Sclerosis[a]

Variant Type	Variants, No.: Total	Novel	Known	*P* Value	Unique	Frequent	*P* Value
Single-nucleotide changes							
Silent substitutions	**78**	42	36		31	47	
Missense substitutions	**63**	39	24	.19	37	26	.01
Nonsense substitutions	**1**	1	0		1	0	
Splice	**0**	0	0		0	0	
Insertions and deletions	**0**	0	0		0	0	
Total	**142**	**82**	**60**		**69**	**73**	

[a] The novel group includes all variants that were not reported in any database; the known group contains all variants previously reported in the dbSNP database. The unique variants included all variants observed in only 1 patient with amyotrophic lateral sclerosis; the frequent group comprised all other variants.

been found exclusively in SALS cases, 8 have been found exclusively in FALS cases, and only 1 has been found in both SALS and FALS cases. On the other hand, 10 genes exhibited no variant, 10 genes had only 1 variant each, and 9 genes had more than 1 variant.

We identified 21 different nonsynonymous variants in SALS cases, including 20 missense variants and 1 nonsense variant (Table 4). Four missense variants (Arg65Cys, Gly113Arg, Arg246Trp, and Glu367Gln) were found in the *CDH13* gene, which encodes a calcium-dependent cell-cell adhesion glycoprotein. None of these affect a functional domain of the protein, and only Arg246Trp is predicted to be damaging by the bioinformatics programs.

Two missense variants (Ile234Thr and Pro578Leu) as well as 1 nonsense variant (Arg1191X) were found in the *DIAPH3* gene. The Ile234Thr variant affects the guanine triphosphatase–binding domain, which is required for the protein activation but is not predicted to affect the protein function. The Pro578Leu variant affects a proline residue located in a proline stretch and is predicted to affect the protein function only by PANTHER and PolyPhen. The Arg1191X variant is the only nonsense variation identified in this study and removes the last 3 amino acids of this protein.

Two missense variants (Lys867Asn and Glu918Gly) in the *NEFH* gene, encoding the heavy neurofilament protein, were identified in 2 patients with SALS. These do

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

Table 4. Novel Nonsynonymous Variants Identified in the Screening

Gene	Nucleotide Variant	Amino Acid Variant	Htz or Hmz	Frequency in Cases (n=190)	Frequency in Control Subjects (n=190)	PANTHER Prediction	SIFT Prediction	PolyPhen Prediction	Conservation Score	Disease	Cosegregation With FALS
C5ORF23	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
FEZF2	c.860G>A	Gly188Asp	Htz	1	0	Deleterious	Affects protein function	Possibly damaging	33	FALS	No
CD55	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
DOC2B	c.628G>T	Arg209Leu	Htz	1	0	Deleterious	Tolerated	Possibly damaging	12	SALS	NA
CNTN6	c.1207T>G	Phe314Val	Htz	1	2	Deleterious	Tolerated	Benign	33	SALS	NA
	c.3128A>T	Glu954Val	Htz	1	0	Nondeleterious	Tolerated	Possibly damaging	43	FALS	No
CDH13	c.313C>T	Arg65Cys	Htz	1	0	NA	Affects protein function	Possibly damaging	8	SALS	NA
	c.428C>T	Ala103Val	Htz	1	0	NA	Tolerated	Benign	1	FALS	Unknown
	c.457G>C	Gly113Arg	Htz	1	0	NA	Tolerated	Benign	33	SALS	NA
	c.856C>T	Arg246Trp	Htz	1	0	Deleterious	Affects protein function	Possibly damaging	11	SALS	NA
	c.1099G>C	Glu367Gln	Htz	1	0	Nondeleterious	Tolerated	Benign	38	SALS	NA
SOX5	c.1189A>C	Gln362Pro	Htz	1	0	NA	Tolerated	Probably damaging	40	SALS	NA
BCL11B	c.364A>T	Glu32Val	Htz	1	0	NA	Affects protein function	Possibly damaging	36	FALS	Unknown
	c.956C>T	Pro229Ser	Htz	1	0	Nondeleterious	Tolerated	Possibly damaging	40	FALS	Unknown
LBD2	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
OMA1	c.297C>T	His69Tyr	Htz	2	0	NA	Tolerated	Benign	6	SALS	NA
	c.905A>G	Glu272Gly	Htz	1	0	Nondeleterious	Tolerated	Probably Damaging	25	SALS	NA
	c.1185G>T	Asp365Tyr	Htz	1	5	Nondeleterious	Affects protein function	Probably damaging	52	FALS	No
GRB14	c.2064C>T	His507Tyr	Htz	1	0	Nondeleterious	Affects protein function	Possibly damaging	35	FALS	No
SYT9	c.697G>A	Val154Met	Htz	12	1	Nondeleterious	Tolerated	Benign	2	FALS and SALS	No
	c.949A>G	Ile238Val	Htz	1	1	NA	Tolerated	Benign	5	SALS	NA
	c.1294C>G	Leu353Val	Htz	3	3	Nondeleterious	Tolerated	Possibly damaging	12	SALS	NA
NETO1	c.1204G>A	Gly401Arg	Htz	1	1	NA	Tolerated	Benign	2	SALS	NA
CRIM1	c.1256A>G	Asn406Ser	Htz	1	0	Nondeleterious	Tolerated	Benign	19	SALS	NA
LUM	c.987T>C	Leu199Pro	Htz	4	0	Deleterious	Affects protein function	Possibly damaging	43	FALS and SALS	Unknown
STK39	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
CRYM	c.594C>T	Arg169Cys	Htz	1	0	Deleterious	Affects protein function	Probably damaging	39	SALS	NA
TMEM163	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
BCL6	c.640A>G	Asn73Ser	Htz	1	0	NA	Tolerated	Benign	38	SALS	NA
OPN3	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
CDH22	c.679G>A	Glu92Lys	Htz	1	0	Nondeleterious	Tolerated	Benign	43	SALS	NA
	c.804C>T	Thr134Met	Htz	1	0	Deleterious	Tolerated	Possibly damaging	40	SALS	NA
	c2001G>A	Arg533His	Htz	3	1	Nondeleterious	Tolerated	Benign	1	FALS and SALS	No
PCP4	c.253G>T	Ala58Ser	Htz	1	0	NA	Tolerated	Benign	16	SALS	NA
NTNG1	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
RAMP3	c.393T>C	Leu111Ser	Htz	1	0	Nondeleterious	Tolerated	Possibly damaging	5	SALS	NA
	c.410G>A	Glu117Lys	Htz	1	0	Nondeleterious	Tolerated	Benign	36	SALS	NA
IGFBP4	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
NEFH	c.2634G>T	Lys867Asn	Htz	1	0	Nondeleterious	Tolerated	Benign	7	SALS	NA
	c.2786A>G	Glu918Gly	Htz	1	0	Nondeleterious	Tolerated	Benign	22	SALS	NA
	c.1070G>A	Arg346His	Htz	1	0	Nondeleterious	Tolerated	Possibly damaging	36	FALS	Unknown
	c.152C>T	Ala40Val	Htz	1	1	Nondeleterious	Tolerated	Benign	12	SALS	NA
DIAPH3	c.922T>C	Ile234Thr	Htz	1	0	Nondeleterious	Tolerated	Benign	4	SALS	NA
	c.1954C>T	Pro578Leu	Htz	1	0	Deleterious	Tolerated	Probably damaging	11	SALS	NA
	c.1984C>T	Pro588Leu	Htz and Hmz	10	13	Deleterious	Affects protein function	Probably damaging	11	FALS and SALS	No
	c.2008C>T	Pro596Leu	Htz	1	3	Nondeleterious	Affects protein function	Probably damaging	33	FALS and SALS	No
	c.3125G>A	Arg1042His	Htz	1	0	Deleterious	Tolerated	Possibly damaging	41	FALS	No
	c.3794C>T	Arg1191X	Htz	1	0	NA	NA	NA	27	SALS	NA
ITM2A	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
S100A10	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .

Abbreviations: FALS, familial amyotrophic lateral sclerosis; Hmz, change found in a homozygous state; Htz, change found in a heterozygous state; NA, not applicable; SALS, sporadic amyotrophic lateral sclerosis; ellipses, no mutations were found.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

not affect a conserved residue, and neither is predicted to affect the protein function.

Two missense variants were identified in each of the *OMA1*, *RAMP3*, and *CDH22* genes. The His65Tyr and Glu272Gly variants found in *OMA1* are predicted to be tolerated and do not affect highly conserved residues. Likewise, in *RAMP3*, the Leu111Ser and Glu117Lys missense variants do not affect a functional domain of the encoded protein and are predicted to be tolerated. By contrast, the Glu92Lys and Thr134Met missense variants found in *CDH22* affect highly conserved residues located in the first cadherin domain, and the Thr134Met missense variant is predicted to be damaging by PANTHER and PolyPhen. Of the 6 remaining missense variants identified in patients with SALS, only the Arg169Cys variant found in the *CRYM* gene, encoding a taxon-specific crystallin protein that binds thyroid hormone, is predicted to be damaging by all 3 programs used and affects a highly conserved residue, which is conserved in 34 of 35 vertebrates from humans to fish.

Eight missense variants were identified in patients with FALS (Table 4). We were unfortunately unable to test the cosegregation of 4 of these missense variants (Ala103Val in *CDH13*, Glu32Val and Pro229Ser in *BCL11B*, and Arg346His in *NEFH*) as no additional family members were available. Three missense variants (Gly188Asp in *FEZF2*, Glu954Val in *CNTN6*, and His507Tyr in *GRB14*) were also found in unaffected siblings, suggesting that these are not associated with ALS in these families. The Arg1042His missense variant identified in *DIAPH3* was also found in the only available unaffected brother, suggesting that this missense variant could be causative with incomplete penetrance or could be unrelated to the patient's ALS phenotype.

Only 1 missense variant (Leu199Pro) found in *LUM* was identified in 1 patient with FALS and 3 patients with SALS. Unfortunately, we were unable to test the cosegregation of this mutation in the familial case as no additional family members were available. This missense variant affects a residue located within a leucine-rich repeat domain, is highly conserved, and is predicted to be deleterious by all 3 programs.

Finally, no variation was found in *C5ORF23*, *CD55*, *LBD2*, *STK39*, *TMEM163*, *OPN3*, *NTNG1*, *IGFBP4*, *ITM2A*, and *S100A10*, suggesting that these genes may not have a genetic role in ALS.

EXCESS OF UNIQUE NONSYNONYMOUS VARIANTS

Given that nonsynonymous variants are more likely to have an effect on protein function and that synonymous variants are likely to represent neutral benign polymorphisms, we evaluated the distribution of nonsynonymous and synonymous variants in our ALS cohort by classifying them into 2 categories. The first category includes novel as well as known variants that were previously reported in the dbSNP database. The second category includes unique variants that were observed in only 1 patient with ALS as well as frequent variants. Of the 142 coding sequence variants, 82 (57.7%) were novel. However, the distribution of novel vs known variants was not significantly different between nonsynonymous and synonymous variants ($P=.19$). On the other hand, the grouping of variants into unique and frequent variants revealed a significant difference ($P=.01$) in the distribution of nonsynonymous and synonymous variants toward the identification of more unique nonsynonymous variants than expected (Table 3).

COMMENT

We report here the resequencing of 29 candidate genes in a cohort of 190 patients with ALS. We identified 142 coding variants and prioritized and genotyped those most likely to have a detrimental effect. To identify mutations in genes that cause or predispose to ALS, we focused our analysis on nonsynonymous variants as these are more likely to have a substantial effect on protein function. We identified 40 novel nonsynonymous variants, including 39 missense variants and 1 nonsense variant. Most importantly, we showed a significant excess of unique nonsynonymous variants compared with unique synonymous variants in our cohort of patients with ALS. This excess suggests that these unique nonsynonymous variants may include ALS-predisposing mutations because converging evidence shows that a large number of individually rare, highly pathogenic mutations cause ALS.[4,8,9] However, we believe that only a small fraction of these unique nonsynonymous variants identified here are likely to predispose to ALS. Thus, the identification of ALS genes in which unique missense variants are responsible for the disease is challenging. We therefore adopted a multifaceted approach based on the functional prediction of missense variants, the conservation of the altered amino acid, and the cosegregation of the variants identified in familial cases to highlight missense variations potentially involved in ALS. Genes with missense variants that also scored highly based on these secondary criteria were considered excellent candidates. Using these combined strategies, we were able to identify genes with possible involvement in ALS.

The *LUM* gene could be considered an excellent candidate for ALS as the Leu199Met missense variant was identified in 4 patients with ALS but not in 190 control subjects, affects a highly conserved residue located within a leucine-rich repeat domain, and is predicted to be deleterious based on bioinformatics programs. This gene encodes lumican, a keratan sulfate proteoglycan involved in collagen fibril organization, epithelial cell migration, and tissue repair.[18] Based on our combined approach, the *LUM* gene constitutes a promising candidate for ALS; however, screening of this gene in additional cases will help determine its role in ALS pathogenesis.

We identified 5 different missense variants in the *CDH13* gene, which is a member of the cadherin superfamily and encodes a calcium-dependent cell-cell adhesion glycoprotein.[19] We were unable to assess the cosegregation of the Ala103Val variant found in a patient with FALS as no additional family members were available. The other missense variants were found in SALS cases; among these, only the Arg246Trp variant is predicted to be deleterious. Therefore, mutations in this gene may play a role in ALS, but further investigation is warranted.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

Three missense variants and 1 nonsense variant were identified in the *DIAPH3* gene, which encodes a protein involved in actin cytoskeleton reorganization. The Arg1042His variant was found in a patient with FALS and an unaffected brother who had passed the average age at onset for ALS, suggesting that this variant could represent a partially penetrant allele or could simply be a rare benign polymorphism. The 2 other missense variants, Ile234Thr and Pro578Leu, were found in patients with SALS and affect the guanine triphosphatase–binding domain and a proline stretch, respectively. The Ile234Thr variant is predicted to be tolerated, whereas Pro578Leu is predicted to be deleterious by PANTHER and PolyPhen. The Arg1191X nonsense variant was found in a patient with SALS and occurs at the C-terminal part of the protein, removing the last 3 amino acids. However, we do not believe this variant is pathogenic as another truncating variant (Leu1190X) was identified in 4 control subjects. Nonetheless, screening of this gene in additional patients is needed to determine its role in ALS.

Four missense variants were also found in the *NEFH* gene. Three were found in patients with SALS and 1 was found in a patient with FALS. No missense variants found in patients with SALS (Ala40Val, Lys867Asn, and Glu918Gly) are predicted to affect the protein function. We were unable to assess the cosegregation of the Arg346His variant found in the patient with FALS as no additional family members were available. Nevertheless, deletions of *NEFH* were previously identified in both patients with FALS and those with SALS,[13,14] and the transgenic mice that overexpress *NEFH* develop motor neuron pathology.[20] Altogether, this suggests that point mutations in *NEFH* may also play a role in the pathogenesis of ALS in some cases.

Two missense variants were identified in each of the *BCL11B*, *OMA1*, *RAMP3*, and *CDH22* genes. In *BCL11B*, 2 missense variants were found in 2 patients with FALS, but we were unable to assess the cosegregation as no family members were available. In the 3 other genes, missense variants were identified in patients with SALS, and none are predicted to be damaging by the 3 bioinformatics programs. Still, sequencing of these genes in additional patients with ALS is needed to help determine their role in ALS.

Finally, among the remaining missense variants identified in SALS cases, only Arg169Cys found in the *CRYM* gene scored high based on our combined strategy as it affects a highly conserved residue and is predicted to be damaging by the 3 programs. *CRYM* encodes the nicotinamide adenine dinucleotide phosphate–regulated thyroid hormone–binding protein, a taxon-specific crystallin protein that binds thyroid hormone for possible regulatory or developmental roles. Mutations in *CRYM* have been identified as a cause of nonsyndromic deafness.[21] However, increased expression of *CRYM* has been recently observed in microglial cells of a transgenic mouse carrying the Leu126delTT mutation in *SOD1*, which exhibited distinct ALS-like motor symptoms and pathological findings.[22] This overexpression was not found in control mice, suggesting a link with the pathogenesis of FALS. Thus, the Arg169Cys missense variant identified here might be considered as contributing to ALS. However, as for the other missense variants identified in SALS cases, we are aware that it is still difficult to claim this missense variant causes disease and that its pathogenicity regarding ALS remains to be established.

In summary, we have sequenced 29 candidate genes in 190 patients with ALS with an average success rate of 92.45%. This rate constitutes a limitation of our study as we could have missed some potentially important variants. Nevertheless, the analysis of these genes led to the identification of promising novel ALS genes such as *LUM* and *CRYM*. Moreover, this study also highlights some analytical challenges of large-scale sequencing screens to detect disease-causing variants. Consequently, all promising variants identified through this study require careful evaluation and cannot be regarded as disease causing on their own. Further screening and functional studies of these variants are needed to confirm their implication in ALS. As the sequencing of the complete human genome or exome in large numbers of individuals has been shown to be feasible,[23,24] the use of such approaches in the field of ALS could accelerate the identification of novel ALS genes. Overall, this could open new avenues for research into the pathogenesis of this disease and offer leads to the development of new treatment strategies.

Accepted for Publication: October 21, 2010.
Published Online: January 10, 2011. doi:10.1001/archneurol.2010.351
Author Affiliations: Centre of Excellence in Neuromics of Université de Montréal, Centre hospitalier de l'Université de Montréal Research Center (Drs Daoud, Valdmanis, Gauthier, Dion, and Rouleau, Mss Belzil and Desjarlais, and Messrs Spiegelman, Henrion, and Diallo), Department of Medicine, Faculty of Medicine (Dr Dion) and Department of Pediatrics and Biochemistry (Dr Rouleau), Université de Montréal, and Research Center, Centre hospitalier universitaire Sainte-Justine (Dr Rouleau), Montreal, and Research Centre of Centre hospitalier universitaire de Québec, Department of Psychiatry and Neuroscience, Laval University, Quebec City (Dr Gros-Louis), Quebec, Canada; and Unité de Neurologie Comportementale et Dégénérative, Institute of Biology, Montpellier, France (Dr Camu).
Correspondence: Guy A. Rouleau, MD, PhD, FRCPC, Centre of Excellence in Neuromics, Centre hospitalier de l'Université de Montréal Research Center, 2099 Alexandre De-Seve St, Montreal, QC H2L 2W5, Canada (guy.rouleau@umontreal.ca).
Authors Contributions: *Study concept and design:* Daoud, Gros-Louis, Dion, and Rouleau. *Acquisition of data:* Daoud, Valdmanis, Spiegelman, Henrion, Diallo, Desjarlais, and Gauthier. *Analysis and interpretation of data:* Daoud, Valdmanis, Belzil, Spiegelman, and Camu. *Drafting of the manuscript:* Daoud. *Critical revision of the manuscript for important intellectual content:* Valdmanis, Gros-Louis, Belzil, Spiegelman, Henrion, Diallo, Desjarlais, Gauthier, Camu, Dion, and Rouleau. *Obtained funding:* Gros-Louis and Rouleau. *Administrative, technical, and material support:* Daoud, Valdmanis, Belzil, Spiegelman, Henrion, Diallo, Desjarlais, Gauthier, and Camu. *Study supervision:* Dion and Rouleau.
Financial Disclosure: None reported.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013

Funding/Support: This study was supported by the Muscular Dystrophy Association and the ALS Association. Dr Daoud is supported by a postdoctoral fellowship and Ms Belzil is supported by a doctoral fellowship from the Canadian Institutes of Health Research. Dr Rouleau holds a Canada Research Chair in Genetics of the Nervous System and the Jeanne-et-J-Louis-Lévesque Chair in Genetics of Brain Diseases.
Role of the Sponsors: The sponsors had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Online-Only Material: The eAppendix and eTables are available at http://www.archneurol.com.
Additional Contributions: We thank the patients and their families involved in this study. Annie Levert, Annie Raymond, BSc, Pascale Thibodeau, BSc, Karine Lachapelle, Philippe Jolivet, and Sylvia Dobrzeniecka, MSc, provided technical assistance.

REFERENCES

1. Tandan R, Bradley WG. Amyotrophic lateral sclerosis: part 1, clinical features, pathology, and ethical issues in management. *Ann Neurol*. 1985;18(3):271-280.
2. Andersen PM. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr Neurol Neurosci Rep*. 2006;6(1):37-46.
3. Rothstein JD. Current hypotheses for the underlying biology of amyotrophic lateral sclerosis. *Ann Neurol*. 2009;65(suppl 1):S3-S9.
4. Kabashi E, Valdmanis PN, Dion P, et al. *TARDBP* mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet*. 2008;40(5):572-574.
5. Sreedharan J, Blair IP, Tripathi VB, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. 2008;319(5870):1668-1672.
6. Vance C, Rogelj B, Hortobágyi T, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*. 2009;323(5918):1208-1211.
7. Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, et al. Mutations in the *FUS/TLS* gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. 2009;323(5918):1205-1208.
8. Daoud H, Valdmanis PN, Kabashi E, et al. Contribution of *TARDBP* mutations to sporadic amyotrophic lateral sclerosis. *J Med Genet*. 2009;46(2):112-114.
9. Belzil VV, Valdmanis PN, Dion PA, et al; S2D Team. Mutations in *FUS* cause FALS and SALS in French and French Canadian populations. *Neurology*. 2009;73(15):1176-1179.
10. Dion PA, Daoud H, Rouleau GA. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat Rev Genet*. 2009;10(11):769-782.
11. Maruyama H, Morino H, Ito H, et al. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature*. 2010;465(7295):223-226.
12. Arlotta P, Molyneaux BJ, Chen J, Inoue J, Kominami R, Macklis JD. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron*. 2005;45(2):207-221.
13. Figlewicz DA, Krizus A, Martinoli MG, et al. Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis. *Hum Mol Genet*. 1994;3(10):1757-1761.
14. Al-Chalabi A, Andersen PM, Nilsson P, et al. Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis. *Hum Mol Genet*. 1999;8(2):157-164.
15. Brooks BR; Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical Limits of Amyotrophic Lateral Sclerosis" Workshop Contributors. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. *J Neurol Sci*. 1994;124(suppl):96-107.
16. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet*. 2003;33(suppl):228-237.
17. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. 2009;461(7261):272-276.
18. Chakravarti S, Magnuson T. Localization of mouse lumican (keratan sulfate proteoglycan) to distal chromosome 10. *Mamm Genome*. 1995;6(5):367-368.
19. Takeuchi T, Ohtsuki Y. Recent progress in T-cadherin (CDH13, H-cadherin) research. *Histol Histopathol*. 2001;16(4):1287-1293.
20. Julien JP, Côté F, Collard JF. Mice overexpressing the human neurofilament heavy gene as a model of ALS. *Neurobiol Aging*. 1995;16(3):487-492.
21. Abe S, Katagiri T, Saito-Hisaminato A, et al. Identification of *CRYM* as a candidate responsible for nonsyndromic deafness, through cDNA microarray analysis of human cochlear and vestibular tissues. *Am J Hum Genet*. 2003;72(1):73-82.
22. Fukada Y, Yasui K, Kitayama M, et al. Gene expression analysis of the murine model of amyotrophic lateral sclerosis: studies of the Leu126delTT mutation in *SOD1*. *Brain Res*. 2007;1160:1-10.
23. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2010;42(1):30-35.
24. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*. 2010;362(13):1181-1191.

Announcement

Visit www.archneurol.com. You can send an e-mail to a friend that includes a link to an article and a note if you wish. Links will go to short versions of articles whenever possible.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://jamanetwork.com/ on 05/21/2013