Genome-wide Analyses Identify KIF5A as a Novel ALS Gene

Aude Nicolas[1,2], Kevin P. Kenna[2,3], Alan E. Renton[2,4,5], Nicola Ticozzi[2,6,7], Faraz Faghri[2,8,9], Ruth Chia[1,2], Janice A. Dominov[10], Brendan J. Kenna[3], Mike A. Nalls[8,11], Pamela Keagle[3], Alberto M. Rivera[1], Wouter van Rheenen[12], Natalie A. Murphy[1], Joke J.F.A. van Vugt[13], Joshua T. Geiger[14], Rick A. Van der Spek[13], Hannah A. Pliner[1], Shankaracharya[3], Bradley N. Smith[15], Giuseppe Marangi[1,16], Simon D. Topp[15], Yevgeniya Abramzon[1,17], Athina Soragia Gkazi[15], John D. Eicher[18], Aoife Kenna[3], ITALSGEN Consortium, Gabriele Mora[19], Andrea Calvo[20], Letizia Mazzini[21], Nilo Riva[22], Jessica Mandrioli[23], Claudia Caponnetto[24], Stefania Battistini[25], Paolo Volanti[26], Vincenzo La Bella[27], Francesca L. Conforti[28], Giuseppe Borghero[29], Sonia Messina[30], Isabella L. Simone[31], Francesca Trojsi[32], Fabrizio Salvi[33], Francesco O. Logullo[34], Sandra D'Alfonso[35], Lucia Corrado[35], Margherita Capasso[36], Luigi Ferrucci[37], Genomic Translation for ALS Care (GTAC) Consortium, Cristiane de Araujo Martins Moreno[38], Sitharthan Kamalakaran[38,39], David B. Goldstein[39], The ALS Sequencing Consortium, Aaron D. Gitler[40], Tim Harris[41], Richard M. Myers[42], NYGC ALS Consortium, Hemali Phatnani[43], Rajeeva Lochan Musunuri[44], Uday Shankar Evani[44], Avinash Abhyankar[44], Michael Charles Zody[44], Answer ALS Foundation, Julia Kaye[45], Steven Finkbeiner[45,46], Stacia Wyman[47], Alexander LeNail[48], Leandro Lima[45], Ernest Fraenkel[48,49], Clive N. Svendsen[50,51], Leslie M. Thompson[52,53], Jennifer E. Van Eyk[54], James D. Berry[55,56], Timothy M. Miller[57], Stephen J. Kolb[58,59], Merit Cudkowicz[60,61], Emily Baxi[62], Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium, Michael Benatar[63], J. Paul Taylor[64,65], Evadnie Rampersaud[66], Gang Wu[66], Joanne Wuu[63], SLAGEN Consortium, Giuseppe Lauria[67], Federico Verde[68], Isabella Fogh[68,69], Cinzia Tiloca[68], Giacomo P. Comi[70], Gianni Sorarù[71], Cristina Cereda[72], French ALS Consortium, Philippe Corcia[73], Hannu Laaksovirta[74], Liisa Myllykangas[75], Lilja Jansson[74], Miko Valori[76], John Ealing[77], Hesham Hamdallah[77], Sara Rollinson[78], Stuart Pickering-Brown[78], Richard W. Orrell[79], Katie C. Sidle[80], Andrea Malaspina[81], John Hardy[80], Andrew B. Singleton[8], Janel O. Johnson[1], Sampath Arepalli[82], Peter C. Sapp[10], Diane McKenna-Yasek[10], Meraida Polak[83], Seneshaw Asress[83], Safa Al-Sarraj[15], Andrew King[15], Claire Troakes[15], Caroline Vance[15], Jacqueline de Belleroche[84], Frank Baas[85], Anneloor LMA ten Asbroek[86], José Luis Muñoz-Blanco[87], Dena G. Hernandez[82], Jinhui Ding[88], J. Raphael Gibbs[88], Sonja W. Scholz[89,62], Mary Kay Floeter[90], Roy H. Campbell[9], Francesco Landi[91], Robert Bowser[92], Stefan M. Pulst[93], John M. Ravits[94], Daniel J.L. MacGowan[95], Janine Kirby[96], Erik Pioro[97], Roger Pamphlett[98], James Broach[99], Glenn Gerhard[100], Travis L. Dunckley[100], Christopher B. Brady[101,102], Neil W. Kowall[103], Juan C. Troncoso[104], Isabelle LE BER[105], Kevin Mouzat[106], Serge Lumbroso[106], Terry D. Heiman-Patterson[107,108], Freya Kamel[109], Ludo Van Den Bosch[110,111], Robert H. Baloh[112], Tim M. Strom[113], Thomas Meitinger[114], Aleksey Shatunov[115], Kristel R. Van Eijk[13], Mamede de Carvalho[116,117], Maarten Kooyman[118], Bas Middelkoop[119], Matthieu Moisse[120,121], Russell L. McLaughlin[122], Michael A. Van Es[13], Markus Weber[123], Kevin B. Boylan[124], Marka Van Blitterswijk[125], Rosa Rademakers[125], Karen E. Morrison[126], A. Nazli Basak[127], Jesús S. Mora[128], Vivian E. Drory[129], Pamela J. Shaw[130], Martin R. Turner[131], Kevin Talbot[131,132], Orla Hardiman[133], Kelly L Williams[134], Jennifer A. Fifita[134], Garth A. Nicholson[134,135], Ian P. Blair[134], Guy A. Rouleau[136], Jesús Esteban-Pérez[137], Alberto García-Redondo[137], Ammar Al-Chalabi[69], Project MinE ALS Sequencing Consortium, Ekaterina Rogaeva[138], Lorne Zinman[139], Lyle Ostrow[62], Nicholas J. Maragakis[62], Jeffrey D. Rothstein[62], Zachary Simmons[140], Johnathan Cooper-Knock[141], Alexis Brice[105], Stephen A. Goutman[142], Eva L. Feldman[142], Summer B. Gibson[93], Franco Taroni[143], Antonia Ratti[7,144], Cinzia Gellera[143], Philip Van Damme[121,145], Wim Robberecht[121,145], Pietro Fratta[17], Mario Sabatelli[146], Christian Lunetta[147], Albert C. Ludolph[148], Peter M. Andersen[149], Jochen H. Weishaupt[148], William Camu[150,151], John Q Trojanowski[152], Vivianna M. Van Deerlin[152], Robert H. Brown, Jr.[10], Leonard H. van den Berg[119], Jan H. Veldink[119], Matthew B. Harms[153], Jonathan D. Glass[154], David J. Stone[2,155], Pentti

Tienari[2,74], Vincenzo Silani[2,7,144], Adriano Chiò[2,20,156], Christopher E. Shaw[2,15], Bryan J. Traynor[1,2,62*], John E. Landers[2,10,157*]

Author affiliations and consortia members are located in the Supplemental Material.

SUMMARY

To identify novel genes associated with ALS, we undertook two lines of investigation. We carried out a genome-wide association study comparing 20,806 ALS cases and 59,804 controls. Independently, we performed a rare variant burden analysis comparing 1,138 index familial ALS cases and 19,494 controls. Through both approaches, we identified *kinesin family member 5A (KIF5A)* as a novel gene associated with ALS. Interestingly, mutations predominantly in the N-terminal motor domain of KIF5A are causative for two neurodegenerative diseases, hereditary spastic paraplegia (SPG10) and Charcot-Marie-Tooth Type 2 (CMT2). In contrast, ALS associated mutations are primarily located at the C-terminal cargo-binding tail domain and patients harboring loss of function mutations displayed an extended survival relative to typical ALS cases. Taken together, these results broaden the phenotype spectrum resulting from mutations in *KIF5A* and strengthen the role of cytoskeletal defects in the pathogenesis of ALS.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS, OMIM #105400) is a neurodegenerative disorder clinically characterized by rapidly progressive muscle weakness and death due to respiratory failure, typically within two to four years of symptom onset (van Es et al., 2017). Although ALS is perceived as being rare, approximately 6,000 Americans die annually from the condition (Hirtz et al., 2007). Furthermore, the number of ALS cases across the globe will increase to nearly 400,000 in 2040, predominantly due to aging of the population (Arthur et al., 2016). This increase is anticipated to place an enormous socioeconomic burden on global healthcare systems, in particular because the annual healthcare cost per patient with ALS is among the highest for any neurological disease (Gladman and Zinman, 2015).

Approximately 10% of ALS display a family history (FALS) whereas the remaining 90% of ALS cases are sporadic (SALS) in nature. Driven in large part by advances in genotyping and sequencing technology, the genetic etiology of two-thirds of familial cases and about 10% of sporadic ALS cases is now known (Chia et al., 2018; Renton et al., 2014). Mutations in *SOD1* were the first identified cause of ALS (Rosen et al., 1993) contributing to ~20% of FALS and ~2% of SALS. More recently, pathogenic hexanucleotide repeat expansions located within the first intron of the *C9orf72* gene on chromosome 9p21 were identified as the most common cause of both FALS (~40%) and SALS (~7%) (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Interestingly, this repeat expansion contributes to ~10% of all FTD cases thus genetically explaining much of the overlap between these clinical syndromes (Majounie et al., 2012). As a result of these major discoveries, there are several ongoing efforts towards directed silencing of these mutant genes which could result in a therapeutic treatment for up to 10% of all ALS cases and for a similar portion of FTD cases.

In addition to the insights provided by each novel ALS gene, the collective knowledge gained from genetic factors provides a more comprehensive understanding of the interacting pathways underlying motor neuron degeneration. For example, the identification of ALS genes has revealed at least three pathways believed to contribute to the development of ALS: (1) RNA metabolism (based on the observation of mutations in *C9orf72*, *TDP-43*, *FUS*, *HNRNPA1*, and *MATR3*); (2) protein homeostasis (*UBQLN2*, *VCP*, *OPTN*, *VAPB*); (3) cytoskeletal dynamics (*PFN1*, *TUBA4A*, *DCTN1*) (Chia et al., 2018; Robberecht and Eykens, 2015; Taylor et al., 2016). Understanding the mechanisms leading to disease pathogenesis again provides targets for therapeutic intervention that may be applicable to all forms of ALS.

Due to the decreased accessibility of multiple affected family members with unknown genetic etiology, there has been an increased focus on the identification of ALS associated genes with moderate to low impact. Despite their lower effect, such genes continue to provide valuable insight into ALS pathogenesis. For example, the product of the risk factor *TBK1* is known to interact with the product of ALS associated gene *OPTN*, further solidifying the role of autophagy and protein homeostasis in disease development (Cirulli et al., 2015; Freischmidt et al., 2015; Maruyama et al., 2010; Morton et al., 2008). Similarly, the risk factor *NEK1*, identified through a rare variant burden analysis of index FALS (i.e., one affected sample per family), is a known binding partner of *C21orf2*, an ALS risk factor found through genome-wide association studies (GWAS) (Cirulli et al., 2015; Kenna et al., 2016; Malovannaya et al., 2011; van Rheenen et al., 2016). The interaction of these two proteins is required for efficient DNA damage repair (Fang et al., 2015), a pathway which is becoming increasingly implicated as a contributing factor in ALS and other neurodegenerative diseases (Coppedè and Migliore, 2015; Lopez-Gonzalez et al., 2016; Madabhushi et al., 2014; Wang et al., 2013).

RESULTS

Genome-wide Association Studies Identify *KIF5A* as a Novel ALS Associated Gene

To identify new susceptibility loci operating in ALS, we undertook a large-scale GWAS involving 12,663 patients diagnosed with ALS and 53,439 control subjects (**Table S1, S2**). Our data were then incorporated into a meta-analysis with a recently published GWAS involving 12,577 ALS cases and 23,475 control subjects (van Rheenen et al., 2016). After imputation and quality-control measures (see **Methods**, **Figure S1** for the workflow and **Figure S2** for the multidimensional scaling plot), 10,031,630 genotyped and imputed variants from 20,806 ALS cases and 59,804 control samples were available for association analysis (**Figure 1A)**. Quantile-quantile plots did not show evidence of significant population stratification ($\lambda_{1000} = 1.001$, **Figure S3**). Single nucleotide polymorphisms (SNPs) achieving genome-wide significance ($P < 5.0\text{x}10^{-8}$) are listed in **Table 1, Table S3** and suggestive loci with SNPs associated at $P < 5.0\times10^{-7}$ are listed in **Table S4.**

Our analysis revealed five previously identified loci that achieved genome-wide significance (loci including *TNIP1*, *C9orf72*, *TBK1*, *UNC13A*, *C21orf2*) (Benyamin et al., 2017; Laaksovirta et al., 2010; Shatunov et al., 2010; van Es et al., 2009; van Rheenen et al., 2016). In addition, we observed a strong association signal for five SNPs in linkage disequilibrium on chromosome 12q14.1 that reached genome-wide statistical significance (**Table 1, Figure 1B)** spanning a region of several hundred kilobases. Of the five SNPs, two of them resided in close proximity to each other within a large intergenic region and two in proximity to *short-chain dehydrogenase/reductase family 9C member 7* (*SDR9C7*), a gene expressed primarily in skin. However, one SNP (rs113247976) results in a p.Pro986Leu coding change within the *kinesin family member 5A* (*KIF5A*) gene ($P = 6.4\text{x}10^{-10}$, OR = 1.38, 95% CI = 1.24-1.53). The case:control allele frequencies for the combined discovery cohort were

2.07%:1.55% and genotype counts were 5 : 529 : 12,043 to 7 : 786 : 22,682 (homozygotes alternative allele : heterozygotes : homozygous reference allele, **Figure 2**). Calculations based on our cohort size as well as the OR and allele frequency of rs113247976 result in a ~99.5% power to detect this as an ALS associated SNP.

Rare Variant Burden Analysis Identifies *KIF5A* as an ALS gene

In an independent line of investigation, we attempted to identify novel ALS genes through exome-wide rare variant burden analysis (RVB). In brief, RVB compares the frequency of variants within each gene below a user defined frequency threshold in a case-control cohort. As the last two ALS associated genes identified by this methodology (*TBK1*, *NEK1*) displayed an increased frequency of loss of function (LOF) variants, we focused our initial analysis on such variants (consisting of nonsense and predicted splice-altering) (Cirulli et al., 2015; Freischmidt et al., 2015; Kenna et al., 2016).

Towards this end, we performed RVB testing for association of LOF variants in a cohort of 1,138 index FALS cases and 19,494 controls, after applying quality control filters (**Experimental Methods, Figure S4, Table S5, S6)**. Genes displaying $P < 5.0 \times 10^{-4}$ are shown in **Table 2**. The previously identified ALS genes, *TBK1* ($P = 5.58 \times 10^{-7}$, OR = 15.11, 95% CI = 5.81-38.69) and *NEK1* ($P = 1.68 \times 10^{-6}$, OR = 6.64, 95% CI = 3.32-12.51), yielded strong associations with ALS reaching exome-wide significance (**Figure 3**). In addition, we observed a single novel gene reaching exome-wide significance, *KIF5A* ($P = 5.55 \times 10^{-7}$; OR = 32.07, 95% CI = 9.05-135.27). Within this gene, we observed 6 LOF variants in our 1,138 cases (0.53%) compared to 3 such variants in our comparison cohort of 19,494 controls (0.015%, **Table 2**). There was no evidence of genomic inflation (λ =0.93, **Figure S5**), sequencing center or other sub-cohort bias (**Figure S6**), or call rate bias (**Figure S7)** in our analysis. Of the index FALS cases

carrying *KIF5A* LOF mutations, we obtained DNA from two siblings of the proband carrying a c.2993-3C>T, exon 27 - 5' splice junction variant, and from a sibling of a different proband carrying a c.3020+2T>A, exon 27 - 3' splice junction variant. These variants segregated with disease within each of these families. Sanger sequencing was performed for 8 of the 9 LOF variant containing samples and was validated in every case.

Interestingly, when we investigated the location of the six ALS associated variants present in *KIF5A*, all occurred within a 34 bp stretch of DNA and were predicted to effect splicing of exon 27 that encodes amino acids 998-1007 (**Table 3, Figure 4A**). Five of the six variants were located on sequential base pairs on the 3' end of the exon, whereas one was located 5' end of the exon. We used the application ASSEDA (Automated Splice Site and Exon Definition Analyses) to predict any mutant mRNA splice isoforms resulting from these variants (Tompson et al., 2007). This algorithm was chosen as it is known to have high performance in splice prediction (Caminsky et al., 2014). ASSEDA predicted a complete skipping of exon 27 for all variants, yielding a transcript with a frameshift at coding amino acid 998, the deletion of the normal C-terminal 34 amino acids of the cargo-binding domain, and the extension of an aberrant 39 amino acids to the C-terminus (**Table 3, Figure 4B, 4C**). The presence of transcripts with skipped exon 27 was demonstrated by performing RT-PCR in two patients carrying exon 27 - 3' splice junction variants (c.3020+2T>A and c.3020+1G>A) using RNA from lymphoblasts and peripheral blood mononuclear cells, respectively. This splice form was not detected in four control lines (**Figure 4D**). Sequence analysis of the smaller RT-PCR products obtained from the patient cells confirmed the exon 26-28 splicing event. Material for RT-PCR was not available for any other patient carrying a *KIF5A* LOF variant.

Our initial RVB was restricted to single nucleotide variants due to the limited sensitivity and comparatively high false positive rates associated with identifying small insertions and deletions (indels) within exome sequencing data (Fang et al., 2014). Based on our discovery of increased LOF variants within *KIF5A*, we re-evaluated this region for the presence of indels. Our analysis revealed two (0.026%) indels within our cohort of 1,138 FALS cases, compared to zero (0%) indels among 19,494 control samples. Both of these indels (p.Asp996fs, p.Asn999fs) resulted in a frameshift of the KIF5A protein coding sequence, and were located close to the splice junction variants that we previously observed to cause skipping of exon 27 resulting in a frameshift at amino acid 998 (**Table 3**). Sanger sequencing confirmed the presence of both indels. Combining the results of the single nucleotide and indel variant analysis yielded a highly significant *P* of 3.8×10^{-9} (OR = 41.16, 95% CI = 12.61-167.57). We failed to detect any signals of RVB association for rare missense variants across *KIF5A* or within any sub-domain of the gene (**Table S8**).

Replication Analysis of rs113247976 and LOF Variants in *KIF5A*

Given the strong signal of the missense variant identified by our GWAS (p.Pro986Leu, rs113247976) and its close proximity to the LOF variants identified by our RVB analysis (amino acids 996-999), we attempted to replicate its association with ALS by analyzing additional cohorts. To accomplish this, we evaluated this variant in a cohort of 4,159 ALS cases and 18,650 controls that were non-overlapping with our GWAS discovery analysis (**Methods, Figure S8**). This included non-overlapping samples from our RVB analysis (673 FALS, 17,696 controls). Analysis of the cohort revealed an allele frequency of 1.78% in cases and 1.32% in controls (rs113247976, $P = 3.82 \times 10^{-4}$, OR = 1.42, 95% CI = 1.17-1.70), thereby replicating the association of the original GWAS. A meta-analysis of the GWAS and replication cohort (n = 24,965 cases, 78,454 controls) yielded a highly significant *P* of 7.09×10^{-13} (OR = 1.39, 95%

CI = 1.33-1.45, **Figure 2)**. These results support the association of *KIF5A* p.Pro986Leu with ALS. However, at this point we cannot definitely state that the missense variant is the primary risk factor, as we cannot rule out other variants in linkage disequilibrium.

We next performed mutational screening of *KIF5A* in an additional cohort of 9,046 ALS cases that had not been included in our original RVB analysis. This revealed three additional carriers of C-terminal variants. One sporadic patient harbored an exon 26 frameshift mutation (p.Asn997fs) and a second sporadic patient harbored an exon 27 splice altering mutation (c.2993-1G>A, **Table 3**). The third patient carried a p.Arg1007Lys (c.3020G>A) mutation and had a familial history of ALS. This mutation was also observed in a FALS patients from our RVB analysis, however, a comparison of 240,715 common variant sites between the two patients failed to reveal a familial relationship (**Experimental Methods**). Additionally, one patient was observed to carry a predicted splice altering variant proximal to exon 3 (c.291+5G>A). However, this variant was not supported as creating an aberrant transcript by ASSEDA. The cohort used for this analysis was comprised mainly frequency of sporadic ALS cases. LOF variants were not observed in a follow up panel of 1,955 controls. Comparison of the LOF variants in sporadic patients (2/9,046 cases, 0.022%) with either the 1,955 replication controls or all controls analyzed in this study (21,449 controls) both yielded insignificant *P* values (0.868 and 0.423, respectively). Interestingly, the frequency of LOF variants in sporadic cases is lower than that observed in our original FALS cohort (0.703%), suggesting that *KIF5A* LOF variants display a high penetrance. Furthermore, the rate of LOF variants reported in the Exome Aggregation Consortium (ExAC) database is lower than we observed in the control samples used in our discovery cohort (0.007% versus 0.015%).

ALS-Associated Mutations in *KIF5A* are Distinct from SPG10/CMT2 Mutations

Missense mutations within *KIF5A* are a known cause of hereditary spastic paraparesis (spastic paraplegia type 10, autosomal dominant; OMIM #604187) and of Charcot-Marie-Tooth disease Type 2 (CMT2) (Crimella et al., 2011; Jennings et al., 2017; Liu et al., 2014; Reid et al., 2002). Although SPG10 and CMT2 share clinical features with ALS, a careful examination of the clinical records of the ALS cases with LOF mutations in *KIF5A* ruled out misdiagnosis. Furthermore, we detected no variants previously associated with SPG10 or CMT2 in our FALS cohort (Liu et al., 2014).

To further elucidate genotype-phenotype relationships, we evaluated the location of mutations within *KIF5A*. Interestingly, mutations contributing to SPG10 and to CMT2 are almost exclusively to be missense changes and are located in the N-terminal motor domain (amino acids 9-327) of KIF5A (**Figure 5**). In contrast, the mutations identified as contributing to ALS are found predominantly in the C-terminal cargo binding region of KIF5A (amino acids 907-1032) with the highly penetrant FALS mutations showing LOF. These results indicate that the functional domain mutated in KIF5A dictates the clinical phenotype, resulting in distinct yet overlapping neurodegenerative diseases.

Patients with *KIF5A* LOF Mutations Display Younger Age at Onset and Longer Survival

To establish the existence of any commonalities between patients with LOF mutations in the C-terminal region of KIF5A, we evaluated their clinical phenotype. Cases with LOF mutations exhibited a median age of onset at 46.5 years (n = 19, **Table S7**). This is lower than the age of onset reported for ALS in epidemiological studies (65.2 years, interquartile range 56.0 – 72.2) (ALSGEN Consortium et al., 2013). Interestingly, we also observed an increased disease duration (survival) in patients harboring these LOF mutations. The median survival time of ALS patients is 20 – 36 months (ALSGEN Consortium et al., 2013). In contrast, cases with LOF mutations exhibited a median survival of nearly 10 years (117

months, n = 17, **Table S7**). ALS patients with symptom onset before 40 years of age have been shown to have longer survival, often exceeding 10 years (Chio et al., 2009). In contrast, patients with uncomplicated types of hereditary spastic paraparesis and CMT2 display a normal life expectancy (Patzkó and Shy, 2011).

DISCUSSION

We previously identified *KIF5A* as a candidate gene for ALS in our prior study that lacked the power to draw a definitive conclusion (Kenna et al., 2016). *KIF5A* was also a candidate ALS gene in a previous GWAS, though it similarly failed to reach genome-wide significance (McLaughlin et al., 2017; van Rheenen et al., 2016) as well as a single gene study selected based on the *a priori* knowledge of its role in HSP/CMT2 and cytoskeletal function (Brenner et al., 2018). Here, we have confirmed *KIF5A* as an ALS-associated gene with genome-wide significance through two independent approaches. By performing a GWAS involving ~80,000 samples, in addition to replicating five previously published loci, as well as the previously reported locus *SCFD1* using a linear mixed model analysis (data not shown), we identified a missense variant within the *KIF5A* gene that reached genome-wide significance for association with ALS risk. It should be stated though that, as with all GWAS, we cannot rule out that other variants in linkage disequilibrium represent the primary risk factor. In an independent line of investigation, we applied RVB analysis to exome sequencing of ~21,000 samples and identified an exome-wide significant association between FALS risk and rare *KIF5A* LOF variants. Analyses of *KIF5A* in independent replication cohorts confirmed our initial finding for both the p.Pro986Leu variant and revealed three additional carriers of LOF variants in 9,046 ALS cases. Taken together our results indicate that the p.Pro986Leu *KIF5A* variant may represent a relatively common, but low penetrance

risk allele for ALS, while LOF variants constitute rare, but high penetrance risk factors.

Kinesins are microtubule-based motor proteins involved in intracellular transport of organelles within eukaryotic cells. In mammals, there are three heavy chain isoforms of KIF5: KIF5A, KIF5B and KIF5C (Miki et al., 2001). The three proteins homo- and heterodimerize through their coiled-coiled stalk domain, and create a complex with two kinesin light chains via binding to the tail domain (Hirokawa et al., 1989). All three KIF5 genes are expressed in neurons (Kanai et al., 2000) and function to transport many cargos by binding to distinct adaptor proteins.

The central role of kinesins in axonal transport leads us to speculate that mutations in *KIF5A* cause disease by disrupting axonal transport. Indeed, defects in axonal transport are a common observation in ALS patients and are already known to directly contribute to motor neuron degeneration pathogenesis (Chevalier-Larsen and Holzbaur, 2006; Hirokawa et al., 2010; Millecamps and Julien, 2013). KIF5 mediates the transport of granules containing both RNA and RNA binding proteins within neuronal dendrites and axons (Kanai et al., 2004). Among these cargos are the ALS-associated proteins FUS and hnRNPA1
(Guo et al., 2017; Kim et al., 2013; Kwiatkowski et al., 2009; Vance et al., 2009). Similarly, KIF5 mediates the transport of VAPB through the adaptor protein protrudin (Matsuzaki et al., 2011), and mutations in the *VAPB* gene have been identified in ALS and late-onset spinal muscular atrophy (Nishimura et al., 2005; 2004). KIF5 is responsible for the axonal transport of neurofilaments (Wang and Brown, 2010) and KIF5A knockout mice display abnormal transport of neurofilaments (Xia et al., 2003). Abnormal accumulation of neurofilaments are a pathological hallmark of ALS and rare mutations in neurofilament heavy polypeptide (NEFH) are associated with ALS (Al-Chalabi et al., 1999).

KIF5 also contributes to the transport of mitochondria (Kanai et al., 2000; Tanaka et al., 1998) and motor neurons derived from KIF5A$^{-/-}$ mice display transport deficits and reduced survival (Karle et al., 2012). Impaired transport and dysfunction of mitochondria represent another common hallmark observed in ALS patients (Chevalier-Larsen and Holzbaur, 2006; Guo et al., 2017; Palomo and Manfredi, 2015; Smith et al., 2017). KIF5 also contributes to the transport of AMPA-type (Heisler et al., 2014; Setou et al., 2002) and $GABA_A$ receptors (Nakajima et al., 2012). In keeping with reported ALS genes such as *NEK1* (Thiel et al., 2011) and *PFN1* (Wu et al., 2012), modulation of *KIF5A* expression has been shown to influence the formation of neurite like membrane protrusions (Matsuzaki et al., 2011). Given its critical interactions with the cytoskeleton, the identification of *KIF5A* mutations further extends the list of cytoskeletal related proteins implicated in ALS pathogenesis, such as PFN1, TUBA4A, NEFH and peripherin (Al-Chalabi et al., 1999; Gros-Louis, 2004; Smith et al., 2014; Wu et al., 2012).

An important question raised by the current study is why variation within the C-terminal cargo binding domain is associated with ALS, while missense variations of the N-terminal motor domain are associated with hereditary spastic paraparesis and Charcot-Marie-Tooth, type 2. Missense mutations within this latter domain have been shown to affect microtubule binding and/or ATP hydrolysis, resulting in a defective KIF5A-mediated anterograde transport of cargo along dendrites and axons. This, in turn, leads to the axonal retrograde degeneration observed both in hereditary spastic paraparesis and Charcot-Marie-Tooth, type 2, two length-dependent axonopathies (Ebbing et al., 2008). In contrast, the primary cellular lesion in ALS is believed to occur within motor neuron cell bodies, where cytoplasmic protein aggregates are consistently observed, and to propagate anterograde along neurites. We anticipate

that LOF variants within the C-terminal domain of KIF5A will disrupt binding with specific cargo proteins. This is supported by a study in zebrafish in which truncation of the C-terminal resulted in a dramatic disruption of axonal localization of mitochondria (Campbell et al., 2014). One possible mechanism is that disruption of binding to cargo may possibly lead to their accumulation and seed aggregation within the cell body resulting in a deficiency at neurite terminals. Deficiency in KIF5A expression and cargo binding has been associated with accumulation of phosphorylated neurofilaments and amyloid precursor protein within neuronal cell bodies, and subsequent neurodegeneration, in patients with multiple sclerosis (Hares et al., 2016). While differences in KIF5A kinetics and KIF5A interactions constitute one possibility to explain the phenotypic heterogeneity, it is also possible C-terminal and N-terminal variants act through a common mechanism, but that a difference in the relative extent of loss or gain of function toxicities leads to milder (i.e. hereditary spastic paraplegia, Charcot-Marie-Tooth, type 2) or more severe (i.e. ALS) phenotypes.

ACKNOWLEDGEMENTS

Author Contributions:

Sample Collection, Preparation and Clinical Evaluation: N.T., B.J.K., P.K., W.v.R., J.J.v.V., R.A.V.d.S., B.N.S., G.Ma., S.D.T., A.S.G., A.Ke., ITALSGEN Consortium, G.M., A.Ca., L.Ma., N.R., J.M., C.Ca., S.B., P.V., V.L.B., F.L.C., G.B., S.M., I.L.S., F.Tr., F.S., F.O.L., S.D., L.C., M.Ca., L.F., Genomic Translation for ALS Care (GTAC) Consortium, C.d.A.M.M., S.K., D.B.G., A.D.G., T.H., R.M.M., NYGC ALS Consortium, H.P., R.L.Mu., U.S.E., A.Ab., M.C.Z., J.Ka., S.F., S.W., A.L., L.L., E.F., C.N.S., L.M.T., J.E.V.E., J.D.Be., T.M.M., S.J.K., M.C., E.B., Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium , M.B., J.W., SLAGEN Consortium, G.L., F.V., I.F., C.Ti., G.P.C., G.S., French ALS Consortium, P.C., H.L., L.M., L.J., M.V., J.Ea., H.H.,

S.R., S.P., R.W.O., K.C.S., A.M., J.H., P.C.S., D.M., M.P., A.K., C.T., C.V., J.d.B., F.B., A.L.t.A., J.L.M., S.W.S., M.K.F., F.L., R.B., S.M.P., J.M.R., D.J.M., J.K., E.P., R.P., J.B., G.G., T.L.D., C.B.B., N.W.K., J.C.T., I.L.B., K.M., S.L., T.D.H., F.K., L.V.D.B., R.H.B., T.M.S., T.M., A.S., K.R.V.E., M.d.C., M.K., B.M., M.M., R.L.M., M.A.V.E., M.W., K.B.B., M.V.B., R.R., K.E.M., A.N.B., J.S.M., V.E.D., P.J.S., M.R.T., K.T., O.H., K.L.W., J.A.F., G.A.N., I.P.B., G.A.R., J.E., A.G., A.A., Project MinE ALS Sequencing Consortium, E.R., L.Z., L.O., N.J.M., J.D.R., Z.S., J.C., A.B., S.A.G., E.L.F., S.B.G., F.T., A.R., C.G., P.V.D., W.R., P.F., M.S., C.L., A.C.L., P.M.A., J.H.W., W.C., J.Q.T., V.M.V.D., R.H.B.J., L.H.v.d.B., J.H.V., M.B.H., J.D.G., D.J.S., P.T., V.S., A.C., C.E.S., B.J.T., J.E.L.

Performed Experiments and Data Analysis: A.N., K.P.K., A.E.R., N.T., F.F., R.C., J.A.D., B.J.K., M.A.N., P.K., A.M.R., W.v.R., N.A.M., J.J.v.V., J.T.G., R.A.V.d.S., H.A.P., F.N.S., B.N.S., G.Ma., S.D.T., Y.A., A.S.G., J.D.E., A.Ke., Genomic Translation for ALS Care (GTAC) Consortium, C.d.A.M.M., S.K., D.B.G., A.D.G., T.H., R.M.M., NYGC ALS Consortium, H.P., R.L.Mu., U.S.E., A.Ab., M.C.Z., J.Ka., S.F., S.W., A.L., L.L., E.F., C.N.S., L.M.T., J.E.V.E., J.D.Be., T.M.M., S.J.K., M.C., E.B., Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium , E.Ra., G.W., SLAGEN Consortium, G.L., F.V., I.F., C.Ti., G.P.C., G.S., A.B.S., J.O.J., S.Ar., P.C.S., D.M., M.P., C.T., C.V., J.d.B., F.B., A.L.t.A., J.L.M., D.G.H., J.D., J.R.G., S.W.S., R.H.C., T.D.H., L.V.D.B., R.H.B., T.M.S., T.M., A.S., K.R.V.E., M.d.C., M.K., B.M., M.M., R.L.M., M.A.V.E., M.W., K.B.B., M.V.B., R.R., K.E.M., A.N.B., J.S.M., V.E.D., P.J.S., M.R.T., K.T., O.H., K.L.W., J.A.F., G.A.N., I.P.B., G.A.R., J.E., A.G., A.A., Project MinE ALS Sequencing Consortium, E.R., L.O., N.J.M., J.D.R., Z.S., J.C., S.A.G., E.L.F., F.T., A.R., C.G., P.V.D., W.R., A.C.L., P.M.A., J.H.W., W.C., J.Q.T., V.M.V.D., R.H.B.J., L.H.v.d.B., J.H.V., M.B.H., J.D.G., D.J.S., P.T., V.S., A.C., C.E.S., B.J.T., J.E.L.

Scientific Planning and Direction: A.N., K.P.K., A.E.R., N.T., F.F., R.C., J.A.D., B.J.K., M.A.N., P.K., A.M.R., W.v.R., N.A.M., J.J.v.V., J.T.G., R.A.V.d.S., H.A.P., F.N.S., B.N.S., G.Ma., S.D.T., Y.A., J.D.E., ITALSGEN Consortium, G.M., A.Ca., L.Ma., N.R., J.M., C.Ca., S.B., P.V., V.L.B., F.L.C., G.B., S.M., I.L.S., F.Tr., F.S., F.O.L., S.D., L.C., M.Ca., L.F., Genomic Translation for ALS Care (GTAC) Consortium, D.B.G., A.D.G., T.H., R.M.M., NYGC ALS Consortium, H.P., J.Ka., S.F., S.W., A.L., L.L., E.F., C.N.S., L.M.T., J.E.V.E., J.D.Be., T.M.M., S.J.K., M.C., E.B., Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium , M.B., J.P.T., J.W., French ALS Consortium, P.C., H.L., L.M., L.J., M.V., J.Ea., H.H., S.R., S.P., R.W.O., K.C.S., A.M., J.H., A.B.S., C.T., C.V., J.d.B., F.B., A.L.t.A., J.L.M., D.G.H., J.R.G., S.W.S., M.K.F., F.L., R.B., S.M.P., J.M.R., D.J.M., J.K., E.P., R.P., J.B., G.G., T.L.D., C.B.B., N.W.K., J.C.T., I.L.B., K.M., S.L., T.D.H., F.K., L.V.D.B., R.H.B., T.M.S., T.M., A.S., K.R.V.E., M.d.C., M.K., B.M., M.M., R.L.M., M.A.V.E., M.W., K.B.B., M.V.B., R.R., K.E.M., A.N.B., J.S.M., V.E.D., P.J.S., M.R.T., K.T., O.H., K.L.W., J.A.F., G.A.N., I.P.B., G.A.R., J.E., A.G., A.A., Project MinE ALS Sequencing Consortium, E.R., L.Z., L.O., N.J.M., J.D.R., Z.S., J.C., A.B., S.A.G., E.L.F., S.B.G., F.T., A.R., C.G., P.V.D., W.R., P.F., M.S., C.L., A.C.L., P.M.A., J.H.W., W.C., J.Q.T., V.M.V.D., R.H.B.J., L.H.v.d.B., J.H.V., M.B.H., J.D.G., D.J.S., P.T., V.S., A.C., C.E.S., B.J.T., J.E.L.

Initial Manuscript Preparation: A.N., K.P.K., A.E.R., N.T., F.F., R.C., J.A.D., M.A.N., A.B.S., S.W.S., J.H.V., D.J.S., P.T., V.S., A.C., C.E.S., B.J.T., J.E.L.

Conflict of Interest Disclosures: None of the authors report any conflicts of interest. J.D.Be.is a consultant to Neuraltus Pharmaceuticals and Denali Therapeutics, and held a research fellow position

funded by Voyager Therapeutics. M.C. has been a consultant for Eli Lilly and Company, Mitsubishi Tanabe Pharma America (MT Pharma America), Denali Therapeutics, Karyopharm Therapeutics and Cytokinetics. S.A.G. has served as a consultant and received research support from Cytokinetics. O.H. has received speaking honoraria from Novarits, Biogen Idec, Sanofi Aventis and Merck-Serono and has been a member of advisory panels for Biogen Idec, Allergen, Ono Pharmaceuticals, Novartis, Cytokinetics and Sanofi Aventis. O.H. serves as Editor-in-Chief of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. L.H.v.d.B. serves on scientific advisory boards for the Prinses Beatrix Spierfonds, Thierry Latran Foundation, Biogen and Cytokinetics. Serves on the editorial board of Amyotrophic Lateral Sclerosis And Frontotemporal Degeneration and The Journal of Neurology, Neurosurgery, and Psychiatry. J.H.V. reports that his institute received consultancy fees from Vertex Pharmaceuticals. A.C. serves on scientific advisory boards for Biogen Idec, Cytokinetics, Italfarmaco, and Neuraltus.

Funding/Support: P.V.D. is a senior investigator of FWO-Vlaanderen. Project MinE Belgium has been supported by ALS liga België, Flanders Innovation & Enterpreneurship (IWT grant Project MinE), the Belgian National Lottery and a grant from Opening the Future Fund (KU Leuven). W.v.R. is supported through the E. von Behring Chair for Neuromuscular and Neurodegenerative Disorders and ERC (grant agreement n° 340429). Additional funding support includes NINDS R35 NS097261 (R.R.) and P01NS084974 (R.R. and K.B.B.). A.N.B. thanks Suna and Inan Kirac Foundation for its generous support throughout this study. Funding for this work was provided by the Heaton-Ellis Trust, The Middlemass Family, Motor Neurone Disease Association, Medical Research Council, Medical Research Foundation, The Psychiatry Research Trust of the Institute of Psychiatry, Guy's and St. Thomas' Charity, the Wellcome Trust, and the Noreen Murray Foundation (C.E.S). This work was also supported

by the UK Dementia Research Institute which is funded by the Medical Research Council, Alzheimer's Society and Alzheimer's Research UK (C.E.S.). The salary for B.N.S was funded by the Medical Research Foundation (MRF) (MRF-060-0003-RG-SMITH). P.C.S. was supported through the auspices of Dr. H. Robert Horvitz (Massachusetts Institute of Technology), an Investigator of the Howard Hughes Institute. Support from the Department of Veterans Affairs and NIH (P30AG13846) to N.W.K. I.P.B. is supported by the Motor Neurone Disease Research Institute of Australia and the National Health and Medical Research Council of Australia (1107644, 1095215). Research support from NIH/NIEHS (K23ES027221), the ALS Association, Target ALS, and Cytokinetics was provided to S.A.G. M.C. was awrded funding from ALS Finding a Cure. N.T., C.Ti., C.G., V.S. and J.E.L. received research support from AriSLA - Fondazione Italiana di Ricerca per la SLA (grants EXOMEFALS and NOVALS) and the Italian Ministry of Health (grant GR-2011-02347820 - IRisALS). R.L.M. was supported by Science Foundation Ireland and the MND Association of England, Wales and Northern Ireland. O.H. is funded by the Health Research Board Clinician Scientist Programme and Science Foundation Ireland. P.J.S. is supported as an NIHR Senior Investigator (NF-SI-0512-10082). P.J.S. and J.K. are supported by the Sheffield NIHR Biomedical Research Centre for Translational Neuroscience (IS-BRC-1215-20017). A.C. receives research support from the Italian Ministry of Health (Ricerca Finalizzata), Regione Piemonte (Ricerca Finalizzata), University of Turin, Fondazione Vialli e Mauro onlus, and the European Commission (Health Seventh Framework Programme). R.B. received funding support from NINDS/NS061867 and Target ALS. R.H.B.J. received funding from the Angel Fund, Project ALS/P2ALS, and the ALS Therapy Alliance. Additional funding was provided by the NIH/NINDS (R01NS073873, J.E.L.), the American ALS Association (N.T., V.S. C.E.S., R.H.B.J., J.E.L.), and the MND Association (N.T., V.S., C.E.S., J.E.L.).

The sequencing activities at NYGC were supported by the ALS Association (ALSA) and TOW Foundation. The CReATe consortium (U54NS092091) is part of Rare Diseases Clinical Research Network (RDCRN), an initiative of the Office of Rare Diseases Research (ORDR), NCATS. This consortium is funded through collaboration between NCATS, and the NINDS. Additional support is provided by the ALS Association (17-LGCA-331). The InCHIANTI study baseline (1998–2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001–2003) was funded by the U.S. National Institute on Aging (Contracts: N.1-AG-1-1 and N.1-AG-1-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004–2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002). This work was supported in part by the Intramural Research Programs of the US National Institutes of Health (NIH), National Institute on Aging (Z01-AG000949-02); by the National Institute of Neurological Disorders and Stroke; and by Merck & Co., Inc. The work was also supported by the Center for Disease Control and Prevention, the Muscular Dystrophy Association, Microsoft Research, the Packard Center for ALS Research at Johns Hopkins, the ALS Association, UK MND Association, Medical Research Council (MRC) UK, Wellcome Trust/MRC Joint Call in Neurodegeneration Award, MRC Neuromuscular Centre, UK National Institute for Health Research Biomedical Research Unit, Italian Health Ministry (Ricerca Sanitaria Finalizzata 2007), Fondazione Vialli e Mauro Onlus, Compagnia di San Paolo, European Community's Health Seventh Framework Programme (FP7/2007-2013) under grant agreements 259867.

Role of the Sponsors: The sponsors did not participate in the design and conduct of the study; collection, management, analysis or interpretation of the data; or preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Additional contributions: We thank the patients and research subjects who contributed samples for this study. This study used DNA samples and clinical data from the Target ALS Human Postmortem Tissue Core, the NINDS Repository at Coriell, the North East ALS (NEALS) Consortium Biorepository, the New York Brain Bank-The Taub Institute, Columbia University, Department of Veterans Affairs Biorepository Brain Bank (grant #BX002466; C.B.B.), the Baltimore Longitudinal Study on Aging at the Johns Hopkins School of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland and the Australian MND DNA Bank. We also thank Crystal Pacut, Blake Swihart, and Jayna Duell, RN for assistance in study coordination (S.A.G.). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Maryland (http://biowulf.nih.gov) and the Massachusetts Green High Performance Computing Center at the University of Massachusetts Medical School (https://www.mghpcc.org). This study also used genotype and clinical data from the Wellcome Trust Case Control Consortium.

REFERENCES

Al-Chalabi, A., Andersen, P.M., Nilsson, P., Chioza, B., Andersson, J.L., Russ, C., Shaw, C.E., Powell, J.F., and Leigh, P.N. (1999). Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis. Human Molecular Genetics *8*, 157–164.

ALSGEN Consortium, Ahmeti, K.B., Ajroud-Driss, S., Al-Chalabi, A., Andersen, P.M., Armstrong, J., Birve, A., Blauw, H.M., Brown, R.H., Bruijn, L., et al. (2013). Age of onset of amyotrophic lateral sclerosis is modulated by a locus on 1p34.1. Neurobiol. Aging *34*, 357.e7–.e19.

Arthur, K.C., Calvo, A., Price, T.R., Geiger, J.T., Chio, A., and Traynor, B.J. (2016). Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. Nat Commun *7*, 12408.

Benyamin, B., He, J., Zhao, Q., Gratten, J., Garton, F., Leo, P.J., Liu, Z., Mangelsdorf, M., Al-Chalabi, A., Anderson, L., et al. (2017). Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis. Nat Commun *8*, 611.

Brenner, D., Yilmaz, R., Müller, K., Grehl, T., Petri, S., Meyer, T., Grosskreutz, J., Weydt, P., Ruf, W., Neuwirth, C., et al. (2018). Hot-spot KIF5A mutations cause familial ALS. Brain. doi: 10.1093/brain/awx370.

Brooks, B.R. (1994). El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. 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. J. Neurol. Sci. *124 Suppl*, 96–107.

Caminsky, N.G., Mucaki, E.J., and Rogan, P.K. (2014). Interpretation of mRNA splicing mutations in genetic disease: review of the literature and guidelines for information-theoretical analysis. F1000Res *3*, 282.

Campbell, P.D., Shen, K., Sapio, M.R., Glenn, T.D., Talbot, W.S., and Marlow, F.L. (2014). Unique function of Kinesin Kif5A in localization of mitochondria in axons. Journal of Neuroscience *34*, 14717–14732.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 1–16.

Chevalier-Larsen, E., and Holzbaur, E.L.F. (2006). Axonal transport and neurodegenerative disease. Biochim. Biophys. Acta *1762*, 1094–1108.

Chia, R., Chio, A., and Traynor, B.J. (2018). Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. Lancet Neurol *17*, 94–102.

Chio, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., Traynor, B.G., and On Behalf of the Eurals Consortium (2009). Prognostic factors in ALS: A critical review. Amyotroph Lateral Scler *10*, 310–323.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.

Cirulli, E.T., Lasseigne, B.N., Petrovski, S., Sapp, P.C., Dion, P.A., Leblond, C.S., Couthouis, J., Lu, Y.-F., Wang, Q., Krueger, B.J., et al. (2015). Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science *347*, 1436–1441.

Coppedè, F., and Migliore, L. (2015). Mutation Research/Fundamental and MolecularMechanisms of Mutagenesis. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis *776*, 84–97.

Crimella, C., Baschirotto, C., Arnoldi, A., Tonelli, A., Tenderini, E., Airoldi, G., Martinuzzi, A., Trabacca, A., Losito, L., Scarlato, M., et al. (2011). Mutations in the motor and stalk domains of KIF5A in spastic paraplegia type 10 and in axonal Charcot-Marie-Tooth type 2. Clin. Genet. *82*, 157–164.

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat. Genet. *48*, 1284–1287.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256.

Delaneau, O., Zagury, J.-F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. Nat Meth *10*, 5–6.

Ebbing, B., Mann, K., Starosta, A., Jaud, J., Schols, L., Schule, R., and Woehlke, G. (2008). Effect of spastic paraplegia mutations in KIF5A kinesin on transport activity. Human Molecular Genetics *17*, 1245–1252.

Fang, H., Wu, Y., Narzisi, G., O'Rawe, J.A., Barrón, L.T.J., Rosenbaum, J., Ronemus, M., Iossifov, I., Schatz, M.C., and Lyon, G.J. (2014). Reducing INDEL calling errors in whole genome and exome sequencing data. Genome Med *6*, 89.

Fang, X., Lin, H., Wang, X., Zuo, Q., Qin, J., and Zhang, P. (2015). The NEK1 interactor, C21ORF2, is required for efficient DNA damage repair. Acta Biochim Biophys Sin *47*, 834–841.

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., Marroquin, N., Nordin, F., Hübers, A., Weydt, P., et al. (2015). Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. Nat Neurosci *18*, 631–636.

Gladman, M., and Zinman, L. (2015). The economic impact of amyotrophic lateral sclerosis: a systematic review. Expert Rev Pharmacoecon Outcomes Res *15*, 439–450.

Gros-Louis, F. (2004). A Frameshift Deletion in Peripherin Gene Associated with Amyotrophic Lateral Sclerosis. Journal of Biological Chemistry *279*, 45951–45956.

Guo, W., Naujock, M., Fumagalli, L., Vandoorne, T., Baatsen, P., Boon, R., Ordovás, L., Patel, A., Welters, M., Vanwelden, T., et al. (2017). HDAC6 inhibition reverses axonal transport defects in motor neurons derived from FUS-ALS patients. Nat Commun *8*, 861.

Hares, K., Redondo, J., Kemp, K., Rice, C., Scolding, N., and Wilkins, A. (2016). Axonal motor protein KIF5A and associated cargo deficits in multiple sclerosis lesional and normal-appearing white matter. Neuropathol Appl Neurobiol *43*, 227–241.

Heisler, F.F., Lee, H.K., Gromova, K.V., Pechmann, Y., Schurek, B., Ruschkies, L., Schroeder, M., Schweizer, M., and Kneussel, M. (2014). GRIP1 interlinks N-cadherin and AMPA receptors at vesicles to promote combined cargo transport into dendrites. Proceedings of the National Academy of Sciences *111*, 5030–5035.

Hirokawa, N., Pfister, K.K., Yorifuji, H., Wagner, M.C., Brady, S.T., and Bloom, G.S. (1989). Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. Cell *56*, 867–878.

Hirokawa, N., Niwa, S., and Tanaka, Y. (2010). Molecular Motors in Neurons:Transport Mechanisms and Rolesin Brain Function, Development, and Disease. Neuron *68*, 610–638.

Hirtz, D., Thurman, D.J., Gwinn-Hardy, K., Mohamed, M., Chaudhuri, A.R., and Zalutsky, R. (2007). How common are the "common" neurologic disorders? Neurology *68*, 326–337.

Jennings, S., Chenevert, M., Liu, L., Mottamal, M., Wojcik, E.J., and Huckaba, T.M. (2017). Characterization of kinesin switch I mutations that cause hereditary spastic paraplegia. PLoS ONE *12*, e0180353.

Jian, X., Boerwinkle, E., and Liu, X. (2014). In silico prediction of splice-altering single nucleotide variants in the human genome. Nucleic Acids Research *42*, 13534–13544.

Kanai, Y., Okada, Y., Tanaka, Y., Harada, A., Terada, S., and Hirokawa, N. (2000). KIF5C, a novel neuronal kinesin enriched in motor neurons. J. Neurosci. *20*, 6374–6384.

Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. Neuron *43*, 513–525.

Karle, K.N., Möckel, D., Reid, E., and Schols, L. (2012). Axonal transport deficit in a KIF5A –/– mouse model. Neurogenetics *13*, 169–179.

Kenna, K.P., van Doormaal, P.T.C., Dekker, A.M., Ticozzi, N., Kenna, B.J., Diekstra, F.P., van Rheenen, W., van Eijk, K.R., Jones, A.R., Keagle, P., et al. (2016). NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. Nat. Genet. *48*, 1037–1042.

Kim, H.J., Kim, N.C., Wang, Y.-D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature *495*, 467–473.

Kwiatkowski, T.J., Bosco, D.A., LeClerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A.,

Gilchrist, J., Kasarskis, E.J., Munsat, T., et al. (2009). Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. Science *323*, 1205–1208.

Laaksovirta, H., Peuralinna, T., Schymick, J.C., Scholz, S.W., Lai, S.-L., Myllykangas, L., Sulkava, R., Jansson, L., Hernandez, D.G., Gibbs, J.R., et al. (2010). Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol *9*, 978–985.

Liu, X., Jian, X., and Boerwinkle, E. (2013). dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum. Mutat. *34*, E2393–E2402.

Liu, Y.-T., Laurá, M., Hersheson, J., Horga, A., Jaunmuktane, Z., Brandner, S., Pittman, A., Hughes, D., Polke, J.M., Sweeney, M.G., et al. (2014). Extended phenotypic spectrum of KIF5A mutations: From spastic paraplegia to axonal neuropathy. Neurology *83*, 612–619.

Lopez-Gonzalez, R., Lu, Y., Gendron, T.F., Karydas, A., Tran, H., Yang, D., Petrucelli, L., Miller, B.L., Almeida, S., and Gao, F.-B. (2016). Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. Neuron *92*, 383–391.

Madabhushi, R., Pan, L., and Tsai, L.-H. (2014). DNA Damage and Its Links to Neurodegeneration. Neuron *83*, 266–282.

Majounie, E., Renton, A.E., Mok, K., Dopper, E.G.P., Waite, A., Rollinson, S., Chio, A., Restagno, G., Nicolaou, N., Simón-Sánchez, J., et al. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol *11*, 323–330.

Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W., Ding, C., Shi, Y., Yucer, N., Krenciute, G., et al. (2011). Analysis of the human endogenous coregulator complexome. Cell *145*, 787–799.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873.

Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. Nat Rev Genet *11*, 499–511.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., et al. (2010). Mutations of optineurin in amyotrophic lateral sclerosis. Nature *465*, 223–226.

Matsuzaki, F., Shirane, M., Matsumoto, M., and Nakayama, K.I. (2011). Protrudin serves as an adaptor molecule that connects KIF5 and its cargoes in vesicular transport during process formation. Molecular Biology of the Cell *22*, 4602–4620.

McLaughlin, R.L., Schijven, D., van Rheenen, W., van Eijk, K.R., O'Brien, M., Kahn, R.S., Ophoff, R.A., Goris, A., Bradley, D.G., Al-Chalabi, A., et al. (2017). Genetic correlation between amyotrophic lateral sclerosis and schizophrenia. Nat Commun *8*, 14774.

Miki, H., Setou, M., Kaneshiro, K., and Hirokawa, N. (2001). All kinesin superfamily protein, KIF, genes in mouse and human. Proc. Natl. Acad. Sci. U.S.a. *98*, 7004–7011.

Millecamps, S., and Julien, J.-P. (2013). Axonal transport deficits and neurodegenerative diseases. Nat. Rev. Neurosci. *14*, 161–176.

Morton, S., Hesson, L., Peggie, M., and Cohen, P. (2008). Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. FEBS Lett. *582*, 997–1002.

Nakajima, K., Yin, X., Takei, Y., Seog, D.-H., Homma, N., and Hirokawa, N. (2012). Molecular Motor KIF5A Is Essential for GABA. Neuron *76*, 945–961.

Nishimura, A.L., Al-Chalabi, A., and Zatz, M. (2005). A common founder for amyotrophic lateral sclerosis type 8 (ALS8) in the Brazilian population. Hum Genet *118*, 499–500.

Nishimura, A.L., Mitne-Neto, M., Silva, H.C.A., Richieri-Costa, A., Middleton, S., Cascio, D., Kok, F., Oliveira, J.R.M., Gillingwater, T., Webb, J., et al. (2004). A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. The American Journal of Human Genetics *75*, 822–831.

Palomo, G.M., and Manfredi, G. (2015). Exploring new pathways of neurodegeneration in ALS: the role of mitochondria quality control. Brain Research *1607*, 36–46.

Patzkó, A., and Shy, M.E. (2011). Update on Charcot-Marie-Tooth disease. Curr Neurol Neurosci Rep *11*, 78–88.

Pe'er, I., Yelensky, R., Altshuler, D., and Daly, M.J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet. Epidemiol. *32*, 381–385.

Reid, E., Kloos, M., Ashley-Koch, A., Hughes, L., Bevan, S., Svenson, I.K., Graham, F.L., Gaskell, P.C., Dearlove, A., Pericak-Vance, M.A., et al. (2002). A kinesin heavy chain (KIF5A) mutation in hereditary spastic paraplegia (SPG10). The American Journal of Human Genetics *71*, 1189–1194.

Renton, A.E., Chio, A., and Traynor, B.J. (2014). State of play in amyotrophic lateral sclerosis genetics. Nat Neurosci *17*, 17–23.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268.

Robberecht, W., and Eykens, C. (2015). The genetic basis of amyotrophic lateral sclerosis. Agg 327–345.

Rosen, D.R., Siddique, T., Patterson, D., Figlewicz, D.A., Sapp, P., Hentati, A., Donaldson, D., Goto, J., O'Regan, J.P., and Deng, H.X. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature *362*, 59–62.

Salvi, E., Kutalik, Z., Glorioso, N., Benaglio, P., Frau, F., Kuznetsova, T., Arima, H., Hoggart, C.,

Tichet, J., Nikitin, Y.P., et al. (2012). Genomewide association study using a high-density single nucleotide polymorphism array and case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of endothelial NO synthase. Hypertension *59*, 248–255.

Setou, M., Seog, D.-H., Tanaka, Y., Kanai, Y., Takei, Y., Kawagishi, M., and Hirokawa, N. (2002). Glutamate-receptor-interacting protein GRIP1 directly steers kinesin to dendrites. Nature *417*, 83–87.

Shatunov, A., Mok, K., Newhouse, S., Weale, M.E., Smith, B., Vance, C., Johnson, L., Veldink, J.H., van Es, M.A., van den Berg, L.H., et al. (2010). Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. Lancet Neurol *9*, 986–994.

Smith, B.N., Ticozzi, N., Fallini, C., Gkazi, A.-S., Topp, S., Kenna, K.P., Scotter, E.L., Kost, J., Keagle, P., Miller, J.W., et al. (2014). Exome-wide Rare Variant Analysis Identifies TUBA4A Mutations Associated with Familial ALS. Neuron *84*, 324–331.

Smith, E.F., Shaw, P.J., and De Vos, K.J. (2017). The role of mitochondria in amyotrophic lateral sclerosis. Neuroscience Letters. doi: 10.1016/j.neulet.2017.06.052.

Tanaka, Y., Kanai, Y., Okada, Y., Nonaka, S., Takeda, S., Harada, A., and Hirokawa, N. (1998). Targeted disruption of mouse conventional kinesin heavy chain, kif5B, results in abnormal perinuclear clustering of mitochondria. Cell *93*, 1147–1158.

Taylor, J.P., Brown, R.H., and Cleveland, D.W. (2016). Decoding ALS: from genes to mechanism. Nature *539*, 197–206.

Thiel, C., Kessler, K., Giessl, A., Dimmler, A., Shalev, S.A., Haar, von der, S., Zenker, M., Zahnleiter, D., Stöss, H., Beinder, E., et al. (2011). REPOR TNEK1 Mutations CauseShort-Rib Polydactyly Syndrome Type Majewski. The American Journal of Human Genetics *88*, 106–114.

Tompson, S.W.J., Ruiz-Perez, V.L., Blair, H.J., Barton, S., Navarro, V., Robson, J.L., Wright, M.J., and Goodship, J.A. (2007). Sequencing EVC and EVC2 identifies mutations in two-thirds of Ellis-van Creveld syndrome patients. Hum Genet *120*, 663–670.

van Es, M.A., Hardiman, O., Chio, A., Al-Chalabi, A., Pasterkamp, R.J., Veldink, J.H., and van den Berg, L.H. (2017). Amyotrophic lateral sclerosis. Lancet *390*, 2084–2098.

van Es, M.A., Veldink, J.H., Saris, C.G.J., Blauw, H.M., van Vught, P.W.J., Birve, A., Lemmens, R., Schelhaas, H.J., Groen, E.J.N., Huisman, M.H.B., et al. (2009). Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat. Genet. *41*, 1083–1087.

van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A.A., Võsa, U., de Jong, S., Robinson, M.R., et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat. Genet. *48*, 1043–1048.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B.,

Ruddy, D., Wright, P., et al. (2009). Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. Science *323*, 1208–1211.

Wang, L., and Brown, A. (2010). A hereditary spastic paraplegia mutation in kinesin-1A/KIF5A disrupts neurofilament transport. Mol Neurodegeneration *5*, 52–64.

Wang, W.-Y., Pan, L., Su, S.C., Quinn, E.J., Sasaki, M., Jimenez, J.C., Mackenzie, I.R.A., Huang, E.J., and Tsai, L.-H. (2013). Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. Nat Neurosci *16*, 1383–1391.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.

Wu, C.-H., Fallini, C., Ticozzi, N., Keagle, P.J., Sapp, P.C., Piotrowska, K., Lowe, P., Koppers, M., McKenna-Yasek, D., Baron, D.M., et al. (2012). Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. Nature *488*, 499–503.

Xia, C.-H., Roberts, E.A., Her, L.-S., Liu, X., Williams, D.S., Cleveland, D.W., and Goldstein, L.S.B. (2003). Abnormal neurofilament transport caused by targeted disruption of neuronal kinesin heavy chain KIF5A. The Journal of Cell Biology *161*, 55–66.

MAIN FIGURE TITLES AND LEGENDS

Figure 1. Identification of association between *KIF5A* locus and ALS risk through GWAS. (A) Manhattan plot showing *P* values from the discovery set GWAS. Analysis of a combined set of 20,806 cases and 59,804 controls is shown. The dashed red line denotes the threshold for genome-wide significance after multiple test correction ($P < 5.0x10^{-8}$). Five previously reported ALS associated loci are labeled in grey and one novel loci, containing the *KIF5A* gene, is labeled in black. (B) Regional association plot of the *KIF5A* locus. Recombination rates are from HapMap phase 2 European ancestry samples. The R^2 pattern is based on the rs113247976 SNP using 85 European ancestry samples (CEU) from the November 2010 release of the 1000 Genomes Project dataset. R^2 of the p.Pro986Leu (rs113247976) with additional SNPs achieving genome-wide significance was 0.544 (rs117027576), 0.544 (rs118082508), 0.741 (rs116900480), and 0.347 (rs142321490).

Figure 2. Discovery and replication for the association of the *KIF5A* p.Pro986Leu (rs113247976) variant with ALS. Analysis of the p.Pro986Leu (rs113247976) variant within each of the described cohorts is shown. Allelic association for all subcohorts were analyzed by logistic regression followed by a fixed-effects meta-analysis. The Forest plot (right) displays the distribution of OR estimates across study cohorts with the vertical dotted line denoting the OR estimated under the meta-analysis.

Figure 3. Identification of association between *KIF5A* and ALS risk through rare variant burden analysis of exome sequencing. Manhattan plot showing gene-level *P* values from an exome-wide rare variant burden analysis. Analyses of 1,138 index FALS cases versus 19,494 controls were restricted to rare LOF variants (splice altering/nonsense, MAF < 0.001). A minimum of 3 LOF gene variants were

required for analysis. The dashed red line denotes the threshold for exome-wide significance after correction for 11,472 genes ($4.36x10^{-6}$). Previously reported (grey) and novel (black) genes exhibiting a significant excess of rare LOF variants in patients are shown.

Figure 4. ALS associated loss of function variants of *KIF5A* disrupt C-terminal sequence by inducing skipping of exon 27. (A) Single nucleotide variants (SNVs) within *KIF5A* identified in ALS patients are clustered at the 5' and 3' splice junctions of exon 27. The consensus splice sequence is shown. (B) ALS associated SNVs are predicted to induce skipping of exon 27 and result in an aberrant mRNA transcript. (C) The skipping of exon 27 of *KIF5A* yields an out-of-frame and extended disrupted C-terminal peptide sequence. The amino acids in red signify the divergence from the normal protein. (D) RT-PCR was performed using RNA derived from ALS patients with the indicated LOF variant or without (controls) using primers to either amplify both wild-type (155 bp) and mutant (127 bp) splice forms or specifically the mutant splice form (80 bp, right panel). The arrow represents the position of the mutant specific product. The tick marks represent 200 bp (upper) and 100 bp (lower) markers.

Figure 5. *KIF5A* ALS mutations show distinct localization from missense mutations previously associated with SPG10 and CMT2. Causative mutations for SPG10 and CMT2 described within the literature (Crimella et al., 2011; Jennings et al., 2017; Liu et al., 2014; Reid et al., 2002) and ALS associated mutations identified within this study are shown. As illustrated, mutations causative for SPG10/CMT2 are predominantly missense changes located in the N-terminal motor domain. In contrast, ALS mutations are primarily located at the C-terminal motor domain and are LOF.

MAIN TABLES AND CORRESPONDING TITLES AND LEGENDS

SNP Information				Present Study (8,229 Cases / 36,329 Controls)				Van Rheenen *et al.* (12,577 Cases / 23,475 Controls)				Combined Discovery Set (20,806 Cases / 59,804 Controls)			
SNP	**Chr**	**Position**	**Gene**	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***
Novel Loci															
rs117027576	12	57,316,603	*KIF5A*	1.55%	1.27%	1.45 [1.20-1.76]	$1.1x10^{-4}$	1.98%	1.59%	1.33 [1.16-1.53]	$4.3x10^{-5}$	1.81%	1.40%	1.37 [1.23-1.54]	$2.3x10^{-8}$
rs118082508	12	57,318,819	*KIF5A*	1.56%	1.28%	1.45 [1.20-1.76]	$1.0x10^{-4}$	1.98%	1.60%	1.33 [1.16-1.53]	$3.8x10^{-5}$	1.81%	1.41%	1.37 [1.23-1.54]	$2.0x10^{-8}$
rs113247976*	12	57,975,700	*KIF5A*	1.83%	1.42%	1.46 [1.23-1.74]	$9.2x10^{-6}$	2.14%	1.70%	1.33 [1.17-1.52]	$1.1x10^{-5}$	2.02%	1.53%	1.38 [1.24-1.53]	$6.4x10^{-10}$
rs116900480	12	58,656,105	*KIF5A*	1.75%	1.46%	1.42 [1.21-1.68]	$1.9x10^{-5}$	2.08%	1.66%	1.34 [1.18-1.53]	$7.1x10^{-6}$	1.95%	1.54%	1.37 [1.24-1.52]	$6.6x10^{-10}$
rs142321490	12	58,676,132	*KIF5A*	1.79%	1.48%	1.43 [1.21-1.68]	$1.5x10^{-5}$	2.08%	1.66%	1.34 [1.18-1.53]	$8.0x10^{-6}$	1.97%	1.55%	1.37 [1.24-1.52]	$6.1x10^{-10}$
Previously Published Loci															
rs10463311	5	150,410,835	*TNIP1*	73.19%	74.84%	0.94 [0.89-0.98]	$7.8x10^{-3}$	73.34%	75.79%	0.91 [0.87-0.94]	$8.5x10^{-7}$	73.28%	75.21%	0.92 [0.89-0.95]	$4.0x10^{-8}$
rs3849943	9	27,543,382	*C9orf72*	71.79%	76.31%	0.84 [0.80-0.88]	$1.4x10^{-12}$	72.78%	76.5%	0.83 [0.80-0.87]	$4.0x10^{-19}$	72.39%	76.38%	0.84 [0.81-0.86]	$3.8x10^{-30}$
rs74654358	12	64,881,967	*TBK1*	3.77%	4.01%	1.20 [1.07-1.34]	$1.6x10^{-3}$	5.12%	4.61%	1.23 [1.13-1.34]	$7.7x10^{-7}$	4.59%	4.25%	1.22 [1.14-1.30]	$4.7x10^{-9}$
rs12973192	19	17,753,239	*UNC13A*	67.62%	69.37%	0.86 [0.82-0.91]	$1.3x10^{-8}$	64.52%	66.00%	0.9 [0.87-0.93]	$2.4x10^{-8}$	65.75%	68.05%	0.89 [0.86-0.91]	$3.9x10^{-15}$
rs75087725	21	45,753,117	*C21orf2*	0.70%	0.46%	1.99 [1.44-2.75]	$2.2x10^{-5}$	1.83%	1.27%	1.61 [1.39-1.87]	$8.7x10^{-11}$	1.38%	0.78%	1.67 [1.46-1.91]	$1.8x10^{-14}$

Table 1. SNPs achieving genome-wide significance in the discovery GWAS. Position is based on Human Genome Assembly build 37. Nearest gene or previously published gene names are included. Chr, chromosome; MAF, minor allele frequency; OR, odds ratio; 95% CI, confidence interval; *, rs113247976 represents the p.Pro986Leu variant in *KIF5A* (NM_004984.2).

Gene	FALS	Control	OR	P
KIF5A	6 (0.53%)	3 (0.02%)	32.07 (9.05-135.27)	5.55×10^{-7}
TBK1	8 (0.70%)	9 (0.05%)	15.11 (5.81-38.69)	5.58×10^{-7}
NEK1	12 (1.05%)	32 (0.16%)	6.64 (3.32-12.51)	1.68×10^{-6}
CALHM2	7 (0.62%)	9 (0.05%)	12.13 (4.47-31.79)	9.19×10^{-6}
COL14A1	8 (0.70%)	16 (0.08%)	8.04 (3.32-18.08)	2.72×10^{-5}
AK1	10 (0.88%)	34 (0.17%)	5.37 (2.55-10.41)	5.62×10^{-5}
ATRN	5 (0.44%)	9 (0.05%)	11.06 (3.57-31.02)	1.66×10^{-4}
VLDLR	5 (0.44%)	9 (0.05%)	10.87 (3.51-30.43)	1.79×10^{-4}
FUS	4 (0.35%)	4 (0.02%)	16.53 (4.25-64.33)	2.08×10^{-4}
ZMYND12	6 (0.53%)	12 (0.06%)	7.92 (2.86-19.96)	2.61×10^{-4}

Table 2. Top ALS associations identified through RVB of FALS and control exome sequencing results

Position	Variant	Exon	cDNA	Description	Predicted Exon Skipping	Gender	Age of Onset (years)	Site of Onset	Survival (months)	Alive (yes/no)
Control Variants										
57,963,470	A>G	11	c.1117+4A>G	3' Splice Junction	P	M	n/a	n/a	n/a	n/a
57,966,423	C>T	15	c.1630C>T	p.Arg544*	-	F	n/a	n/a	n/a	n/a
57,976,884	G>C	28	c.3021G>C	5' Splice Junction	N	F	n/a	n/a	n/a	n/a
FALS Variants										
57,975,729	GA>A	26	c.2987delA	p.Asp996fs	-	M	45	n/a	n/a	n/a
57,976,382	C>T	27	c.2993-3C>T	5' Splice Junction	Y	M	29	L	>264	Y
57,976,385	GA>G	27	c.2996delA	p.Asn999fs	-	M	42	L	>12	Y
57,976,411	A>G	27	c.3019A>G	p.Arg1007Gly	Y	F	53	L	45	N
57,976,412	G>A	27	c.3020G>A	p.Arg1007Lys	Y	M	50	L	>108	Y
57,976,412	G>A	27	c.3020G>A	p.Arg1007Lys	Y	F	50	n/a	>240	Y
57,976,413	G>A	27	c.3020+1G>A	3' Splice Junction	Y	M	45	B	>220	Y
57,976,414	T>A	27	c.3020+2T>A	3' Splice Junction	Y	M	46	B	124	N
57,976,415	A>G	27	c.3020+3A>G	3' Splice Junction	Y	M	50	B	54	N
SALS Variants										
57,957,481	G>A	3	c.291+5G>A	3' Splice Junction	N	n/a	n/a	n/a	n/a	n/a
57,975,731	CA>C	26	c.2989delA	p.Asn997fs	-	F	50	L	>96	Y
57,976,384	G>A	27	c.2993-1G>A	5' Splice Junction	Y	n/a	52	B	n/a	n/a

Table 3. Loss of function variants within *KIF5A* identified in probands. P, possible; Y, yes; N, no; M, male; F, female; L, limb onset; B, bulbar onset, n/a, not available or applicable. Note, ASSEDA does not predict exon skipping based on frameshifts or nonsense mutations (Tompson et al., 2007).

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, John Landers (john.landers@umassmed.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study cohorts

GWAS cohort I. We undertook a GWAS of patients diagnosed with ALS (case cohort) and neurologically normal control individuals (control cohort). DNA was extracted from either whole blood or frozen brain tissue samples using standard procedures. All 12,663 patients included in the case cohort had been diagnosed with ALS according to the El Escorial criteria (Brooks, 1994) by a neurologist specializing in ALS, had onset of symptoms after age 18 years, and were of non-Hispanic white race/ethnicity. Both patients with familial ALS and patients with sporadic ALS were included in the analysis.

For the control cohort, we used genotype data obtained from (a) the database of Genotypes and Phenotypes (dbGaP) web repository (n = 44,017 US samples); (b) the HYPERGENES Project (n = 887 Italian samples) (Salvi et al., 2012); and (c) the Wellcome Trust Case Control Consortium (n = 5,663 British samples). An additional 2,112 US and Italian control samples were genotyped in the Laboratory of Neurogenetics, National Institute on Aging. The control cohort was matched to the case cohort for race and ethnicity, but not for age or sex. A detailed description of the cohorts is available in **Table S1, S2**.

Written consent was obtained from all individuals enrolled in this study, and the study was approved by the institutional review board approval of the National Institute on Aging (protocol number 03-AG-N329).

GWAS cohort II. Summary statistics from a recently published GWAS based on logistic regression analysis involving 12,577 cases and 23,475 controls were downloaded from the Project MinE Variant Browser. Additional details of the cohorts used in this study are available in van Rheenen *et al* (van Rheenen et al., 2016).

FALS discovery cohort. A total of 1,463 FALS patients were included in the initial cohort (pre-QC). Patients were recruited at specialist clinics in Australia (n = 92), Belgium (n = 13), Canada (n = 34), Germany (n = 228), Ireland (n = 18), Israel (n = 26), Italy (n = 230), Netherlands (n = 50), Spain (n = 60), Turkey (n = 72), UK (n = 223), and USA (n = 417). All samples were exome sequenced except those from the Netherlands which were whole genome sequenced. Familial history was considered positive for ALS if the proband had at least one affected relative within three degrees of relatedness.

Control discovery cohort. Read level sequencing data were obtained from dbGAP and the European Genome-Phenome Archive (EGA) and are listed in **Table S5**.

ALS WXS/WGS replication cohort. Replication analyses included sequencing data for a further 9,046 ALS cases and 1,955 non-ALS controls that were not also represented in the FALS discovery set. These samples included 2,742 cases subjected to WXS by the ALS Sequencing Consortium, as described previously (Cirulli et al., 2015); 719 cases subjected to WXS by the Laboratory of Neurogenetics,

National Institute on Aging; 307 cases and 296 controls subjected to WGS by the Laboratory of Neurogenetics, National Institute on Aging; 155 cases subjected to WGS by the CReATe Consortium; 1,017 cases subjected to WGS by the NYGC ALS Consortium, Genomic Translation for ALS Care (GTAC) Consortium and Answer ALS Foundation; and 4,100 cases and 1,659 controls subjected to WGS by the Project MinE Sequencing Consortium.

All samples included in the case cohort had been diagnosed with ALS according to the El Escorial criteria (Brooks, 1994) by a neurologist specializing in ALS. We received approval for this study from the institutional review boards of the participating centers, and written informed consent was obtained from all patients (consent for research).

METHOD DETAILS

Data generation and pre-processing

*Generation of SNP array callset***.** The case cohort (n = 12,663 samples) and part of the control cohort (n = 2,112) were genotyped in the Laboratory of Neurogenetics, National Institute on Aging, using HumanOmniExpress BeadChips (version 1.0, Illumina Inc., San Diego, CA) according to the manufacturer's protocol. These SNP genotyping arrays assay 716,503 SNPs across the genome. Individual-level genotypes for these samples are available on the dbGaP web portal (accession number phs000101.v4.p1). The remainder of the control cohort had been previously genotyped on HumanOmni BeadChips (Illumina) as part of other GWAS efforts (see **Table S2**). Analyses were confined to the 595,692 autosomal SNPs that were common across the SNP genotyping arrays.

Generation of FALS case-control callset for exome-wide RVB discovery analysis. Exome sequencing of cases was performed as previously described (Kenna et al., 2016). Control exome sequences were generated as described under the relevant dbGAP and EGA project accessions. Sequence reads were aligned to human reference GRCh37 using BWA (Burrows-Wheeler Aligner) and processed according to recommended Genome Analysis Toolkit's (GATK) best practices (https://software.broadinstitute.org/gatk/best-practices/). Joint variant detection and genotyping of all samples were performed using the GATK HaplotypeCaller. Variant quality control was performed using the GATK variant quality score recalibration method with default filters. A minimum variant quality by depth (QD) score of 2 was also imposed and all genotypes associated with genotype quality (GQ)< 20 were reset to missing. Variants were also excluded in the event of case or control call rates <70% (post genotype QC). Identified variants can be viewed through our web based ALS Variant Server (see link below).

Generation of ALS case-control callset for KIF5A replication analysis. Data for the *KIF5A* locus was extracted from all independently generated sequencing datasets and remapped to GRCh37. Variant calling was performed using the GATK haplotype caller as described above. In addition to the *KIF5A* locus, data was also extracted for a panel of 240,715 common variant sites and used to perform a single unified sample QC as described below.

Functional annotation of variants identified by WXS/WGS

Variant calls were assigned predicted functional consequences using snpEFF (Single Nucleotide Polymorphism Effect)(Cingolani et al., 2012), dbNSFP (A Database of Human Non-synonymous SNVs and Their Functional Predictions and Annotations)(Liu et al., 2013) and dbscSNV (database of splice

site consequences of Single Nucleotide Variants)(Jian et al., 2014), which is incorporated into dbNSFP. Variants were classified as "loss of function" (LOF) where the sequence change was predicted to encode a premature stop codon, a frameshift causing insertion-deletion or a splice site disrupting SNV. Variants were classified as potentially splice altering if assigned an "ada" or "rf" score >0.7 by dbscSNV. Splice variants of potential interest were further assessed for putative effects on exon skipping using a secondary algorithm - automated splice site and exon definition server (ASSEDA)(Tompson et al., 2007).

RT-PCR Analysis

Total RNA was prepared from lymphoblast lines using Trizol reagent. Reverse transcription using Applied Biosystems RNA to cDNA kit (# 4368814) was performed with 0.5 ug with RNAse inhibitor in a 20 ul reaction according to the manufacturer's protocol. PCR was carried out using New England Biolabs One Taq Hot Start DNA Polymerase (# M0481S), 2 ul RT reaction (representing 50 ng input RNA) and forward and reverse primer (0.15 uM each) in a 20 ul reaction volume. Amplification conditions were as follows: 94°C for 30 seconds, {94°C for 20 seconds, 58°C for 20 seconds, 68°C for 1 minute} x 35 cycles, followed by an extension stage of 68°C for 5 minutes and a 4°C hold. Amplification of both normal and mutant splice forms used primers F1 (CAGTGGAGCCACATCTTCTG) and R1 (TCTCTTGGTGGAGAGGGAAA). Primers used for the specific amplification of the mutant splice form were F2 (CCAACATGGACAATGGAGTGA), which spans exons 26 and 28, and R1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses

Analysis of SNP array genotypes. Standard quality-control procedures were applied to our genotype data using PLINK software package (version 1.9)(Chang et al., 2015), and a summary of the workflow is shown in **Figure S1**. We excluded samples that demonstrated: call rates of less than 97.5%; non-European ancestry; abnormal *F* inbreeding coefficient; mismatch between phenotypic and genotypic gender; or, cryptic relatedness defined as identity-by-descent proportion of inheritance (pi_hat from PLINK) greater than 0.125. Samples in common between our study and van Rheenen's study were identified using the checksum program id_geno_checksum and were removed from our analyses. We excluded palindromic SNPs, as well as SNPs with: call rates less than 95% in the US and Italian cohorts or less than 99% in the UK, French and Belgium cohorts; minor allele frequency less than 0.05 in the control cohorts; Hardy-Weinberg equilibrium *P* less than 10^{-7} in the US and Italian control cohorts and less than 10^{-5} in the UK, French and Belgium cohorts; missingness by case-control status *P* less than 10^{-5}; or SNPs associated between the UK and French control cohorts with *P* less than 5.0×10^{-8}. After quality control, 8,229 case and 36,329 control samples were included in the analysis, and 436,746 SNPs were available for imputation in the USA and Italy cohorts, and 420,131 SNPs were available in the UK, French and Belgium cohorts.

Estimation of the haplotypes was performed with SHAPEIT (version 2.r790)(Delaneau et al., 2013). Imputation was performed for individual batches based on ethnicity using the 1000 Genomes Project dataset (phase 3, version 5a, release 2013-05-02, 1000genomes.org) as reference and using Minimac3 software (version 1.0.11)(Das et al., 2016) with default settings. After imputation, principal components were calculated using PLINK software after removing known hypervariable regions and the 1 MB surrounding the *C9orf72* region. After analysis of the Scree plots, 2 to 4 principal components were

retained per cohort as covariates in the association analyses to compensate for any residual population stratification.

Logistic regression was performed per batch using mach2dat software (version 1.0.24)(Marchini and Howie, 2010) incorporating 2 to 4 principal components, age and gender as covariates, with dosage of imputed SNPs selected based on a Minimac3 R^2 value of imputation accuracy greater than 0.3. SNPs with an absolute beta coefficient value above 5 or with a minor allele frequency less than 0.01 were excluded from meta-analysis. Meta-analysis was then performed combining the association results of the 13 batches of our individual-level studies with van Rheenen's study summary statistics using METAL software (version 2011-03-25)(Willer et al., 2010) under an inverse-weighted, fixed effect model. A threshold P of 5.0×10^{-8} was set for genome-wide significance after Bonferroni correction for multiple testing in the GWAS (Pe'er et al., 2008).

The programming code used to analyze these data is freely available on GitHub (see link below), and GWAS summary statistics results for all tested SNPs are available from (link to be supplied at publication).

Analysis of WXS/WGS genotypes. For both the discovery and replication phases, samples were excluded from the study in the event of failing to meet standard genotype call rate, heterozygosity, duplication, relatedness or population stratification filters as summarized in **Table S6**. Each of these filters was performed using a set of autosomal markers meeting all of the following criteria: call rate > 0.95, minor allele frequency (MAF) > 0.01, $P > 0.001$ for deviation from Hardy-Weinberg equilibrium, linkage disequilibrium pruning ($R^2 < 0.5$, window size = 50, step = 5). Filtering of autosomal markers, sample

call rate assessments and sample heterozygosity assessments were performed using PLINK software. Study duplicates and sample relatedness within the WXS/WGS cohorts was identified using KING software (Manichaikul et al., 2010). Study duplicates between WXS/WGS cohorts and GWAS datasets were identified using the checksum program id_geno_checksum. LASER was used to generate PCA coordinates for samples from the Human Genome Diversity Panel (HGDP). Samples from the FALS discovery cohort were then mapped to this reference co-ordinate space. The discovery cohort was restricted to cases and controls occurring within 3 standard deviations of the mean for European HGDP samples along principal components 1-4.

RVB analyses were performed by penalized logistic regression of case-control status with respect to number of minor alleles observed per sample per gene with and MAF <0.001. Analyses were only performed where the dataset contained more than 3 variant allele occurrences. Replication analyses of rs113247976 were performed using the same logistic regression protocol as used for RVB analyses. All analyses were conditioned on the first 4 eigenvectors generated by principal components analysis of common variant profiles. Genomic inflation factors were calculated using genome-wide association analysis for quantitative, binary and time-till-event traits using GenABEL software. Candidate associations were tested for signs of call-rate or subcohort biases as outlined in **Figures S6, S7**. Meta-analysis of rs113247976 association results between sequencing and GWAS was performed using METAL. Unless otherwise indicated, all statistical analyses were performed using R (version 3.2.0).

Control-control analyses. To identify genes potentially subject to confounding biases in FALS RVB analyses and to assess the potential impact of batch effects with non-ALS-related data, population or phenotypic stratifiers, the control sample cohort was divided into 28 pseudo case-control groups based

on the sequencing center or associated dbGaP / EGA project (**Table S5**). Genes shown in gray achieve for possible confounder association. Loci achieving a minimu $P < 1 \times 10^{-3}$ were deemed as displaying possible association with non-ALS related batch effects.

DATA AND SOFTWARE AVAILABILITY

Datasets

The programming code used to analyze the GWAS data including the imputation with SHAPEIT and Minimac3, individual-based association analysis using Mach2dat and a meta-analysis using METAL is freely available on GitHub: https://github.com/AudeDN/ALS_GWAS_1000G_mach2dat_2017). GWAS summary statistics results for all tested SNPs and identified SNVs from our 1,138 FALS cohort used for the RVB analysis can be viewed through our web based ALS Variant Server (http://als.umassmed.edu). For each variant, information on over 50 annotation fields and the results can be downloaded directly into Excel.

Data Resources and Databases

1000 Genomes Project dataset: http://www.internationalgenome.org

Database of Genotypes and Phenotypes (dbGaP): www.ncbi.nlm.nih.gov/gap

dbNSFP: https://sites.google.com/site/jpopgen/dbNSFP

dbscSNV: incorporated into dbNSFP (see previous link).

European Genome-phenome Archive (EGA): https://ega-archive.org

HapMap project: http://www.sanger.ac.uk/resources/downloads/human/hapmap3.html.

Human Genome Diversity Panel (HGDP): http://www.hagsc.org/hgdp/

HYPERGENES Project: http://www.hypergenes.eu

Project MinE Variant Browser: http://databrowser.projectmine.com

snpEFF: http://snpeff.sourceforge.net/SnpEff.html

Wellcome Trust Case Control Consortium: www.wtccc.org.uk

Software

ASSEDA: http://www.cytognomix.com/?post_type=duka&p=2670

BWA: http://bio-bwa.sourceforge.net

GenABEL: http://www.genabel.org

GATK: https://software.broadinstitute.org/gatk/

id_geno_checksum: https://personal.broadinstitute.org/sripke/share_links/checksums_download/

KING: http://people.virginia.edu/~wc9c/KING/

LASER: http://csg.sph.umich.edu/chaolong/LASER/

Mach2dat: https://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output

METAL: http://csg.sph.umich.edu/abecasis/metal/index.html

Minimac3: https://genome.sph.umich.edu/wiki/Minimac3

PLINK: http://zzz.bwh.harvard.edu/plink/

R: https://www.r-project.org

SHAPEIT: https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

SUPPLEMENTAL ITEM TITLES

Figure S1. Related to Figure 1; Workflow showing the quality control procedures applied to the present study.

Figure S2. Related to Figure 1; Multi-dimensional scaling plot of the 44,558 genotyped samples included in analysis compared to the HapMap populations.

Figure S3. Related to Figure 1; Quartile-Quartile plot of *P*-values from the meta-analysis based on logistic regression analysis.

Figure S4. Related to Figure 3; Principal components analysis of samples included in the RVB analysis compared to the Human Diversity Panel.

Figure S5. Related to Figure 3; Quartile-Quartile plot of P values from the gene-based rare variant burden analysis of exome data.

Figure S6. Related to Figure 3; Control-control analyses.

Figure S7. Related to Fgure 3; Plot of variant call rates across the KIF5A protein-coding region in FALS versus controls analyzed by RVB testing.

Figure S8. Related to Figure 2; Principal components analysis of samples included in *KIF5A* replication cohort.

Table S1. Related to Figure 1; Demographics and baseline characteristics of patients diagnosed with ALS and control individuals included in the GWAS analysis.

Table S2. Related to Figure 1; DbGaP studies contributing to the GWAS analysis.

Table S3. Related to Figure 1; SNPs achieving genome-wide significance in the GWAS analysis.

Table S4. Related to Figure 1; Suggestive SNPs with P values less than 5.0x10-7 in the GWAS analyses.

Table S5. Related to Figure 3; DbGaP/EGA studies contributing to the RVB analysis.

Table S6. Related to Figure 2, 3; Quality control filtering of the FALS discovery and *KIF5A* replication cohorts.

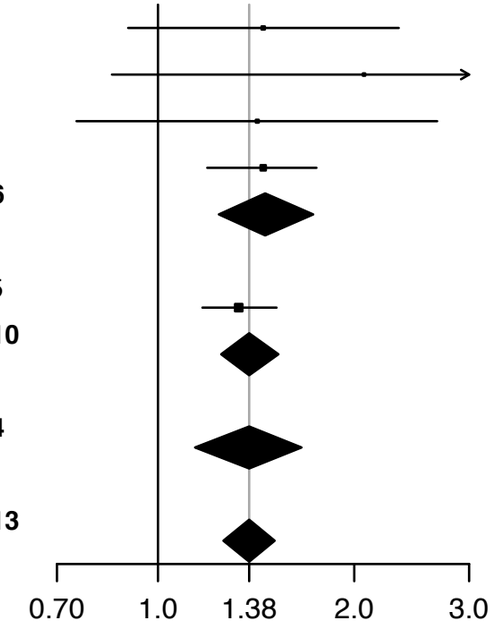
A.

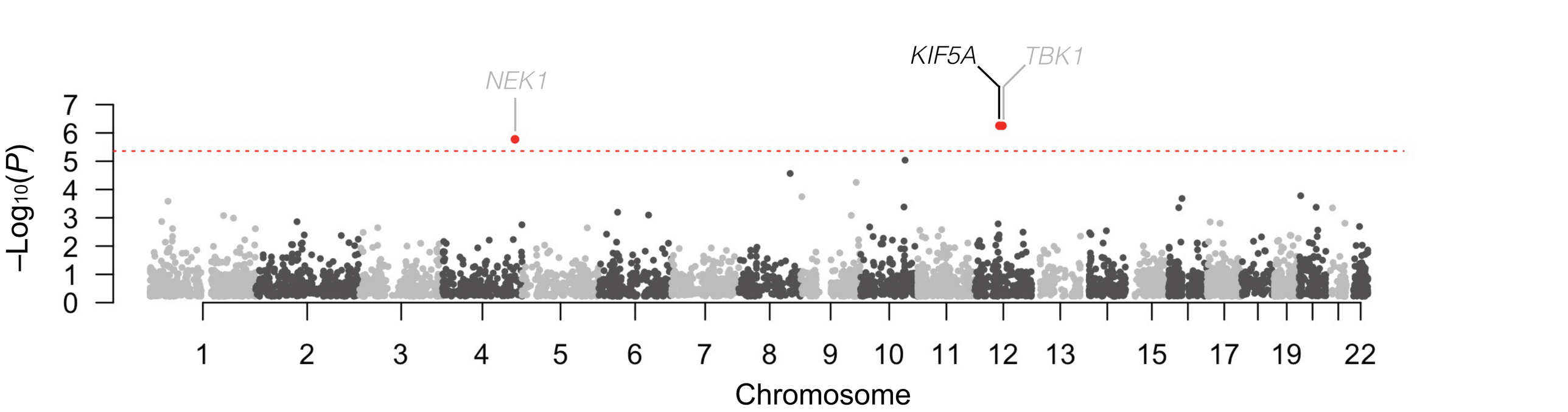
C9orf72
TNIP1
KIF5A
TBK1
UNC13A
C21orf2
$-\text{Log}_{10}(P)$
30
25
20
15
10
5
0
1
2
3
4
5
6
7
8
9
10
11
12
13
15
17
19
22
Chromosome

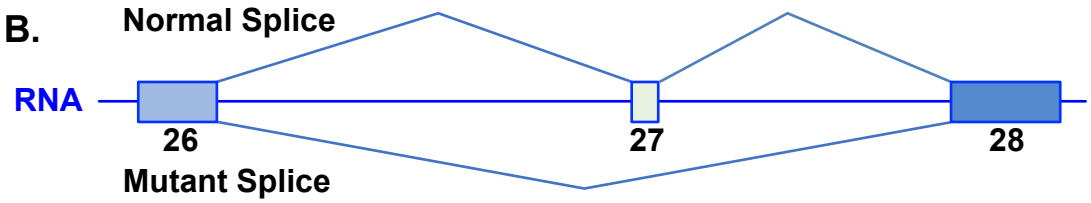
B.

rs117027576
rs118082508
rs113247976
(p.Pro986Leu)
rs116900480
rs142321490
62 cM/Mb
0 cM/Mb
$-\text{Log}_{10}(P)$
10
5
0
ATP5B -
SNORD59B -
SNORD59A -
PTGES3 -
NACA -
PRIM1 -
HSD17B6 +
SDR9C7 -
RDH16 -
GPR182 +
ZBTB39 -
TAC3 -
MYO1A -
TMEM194A -
NAB2 +
STAT6 -
LRP1 +
MIR1228 +
NXPH4 +
SHMT2 +
NDUFA4L2 -
STAC3 -
R3HDM2 -
DDIT3 -
MBD6 +
DCTN2 -
KIF5A +
PIP4K2C +
DTX3 +
ARHGEF25 +
SLC26A10 +
B4GALNT1 -
INHBC +
INHBE +
GLI1 +
ARHGAP9 -
MARS +
MARCH9 +
CYP27B1 -
METTL1 -
METTL21B +
TSFM +
AVIL -
CTDSP2 -
MIR26A2 -
LOC100506844 -
OS9 +
XRCC6BP1 +
AGAP2 -
AGAP2-AS1 +
TSPAN31 +
CDK4 -
57,000,000
57,500,000
58,000,000
58,500,000
59,000,000
Position
R2 with
rs113247976
1.00
0.75
0.50
0.25
0.00

Cohort	Cases	Controls	Case MAF	Control MAF	OR (95% CI)	P
Italy	2,853	2,143	1.28%	0.82%	1.45 (0.90 - 2.34)	
United Kingdom	449	226	3.56%	1.33%	2.07 (0.85 - 5.07)	
Belgium / France	1,150	595	1.91%	1.26%	1.42 (0.75 - 2.68)	
USA	3,777	33,365	2.01%	1.46%	1.45 (1.19 - 1.75)	
Present Study	**8,229**	**36,329**	**1.83%**	**1.42%**	**1.46 (1.24 - 1.73)**	**9.23×10^{-6}**
Van Rheenen Study	12,577	23,475	2.23%	1.75%	1.33 (1.17 - 1.52)	1.13×10^{-5}
Combined Discovery Set	**20,806**	**59,804**	**2.07%**	**1.55%**	**1.38 (1.25 - 1.53)**	**6.43×10^{-10}**
Replication	**4,159**	**18,650**	**1.78%**	**1.32%**	**1.42 (1.17 - 1.70)**	**3.82×10^{-4}**
Joint	**24,965**	**78,454**	**2.02%**	**1.49%**	**1.39 (1.33 - 1.45)**	**7.09×10^{-13}**

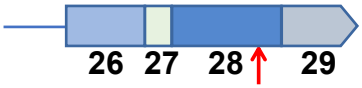



KIF5A
TBK1
NEK1
−Log10(P)
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
15
17
19
22
Chromosome

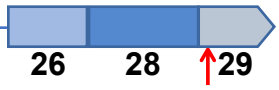


C.

Normal Splice mRNA:



Mutant Splice mRNA:



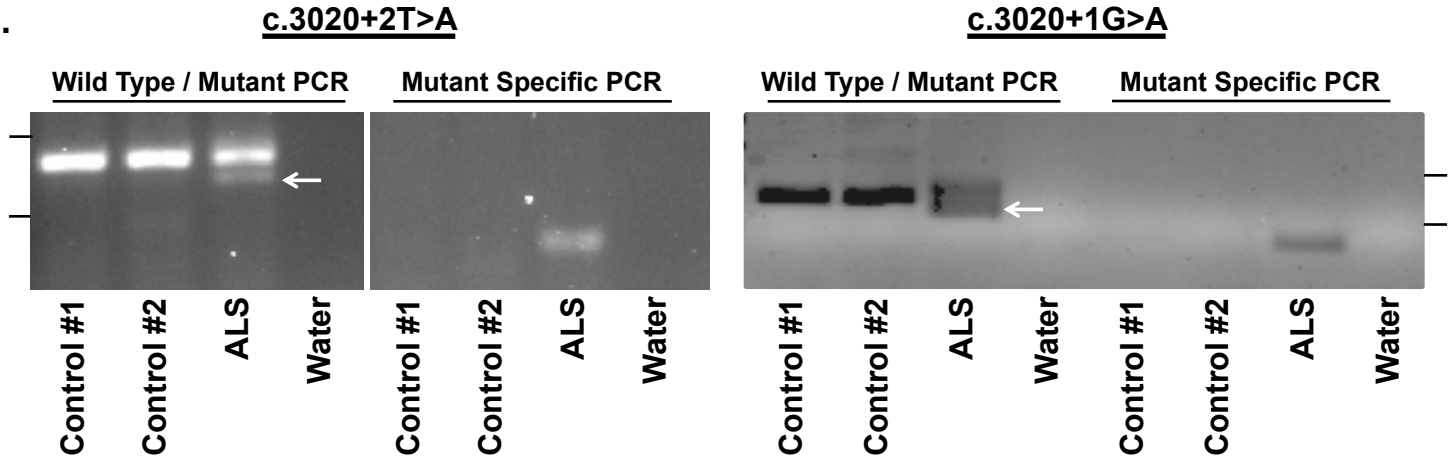
Normal Protein:

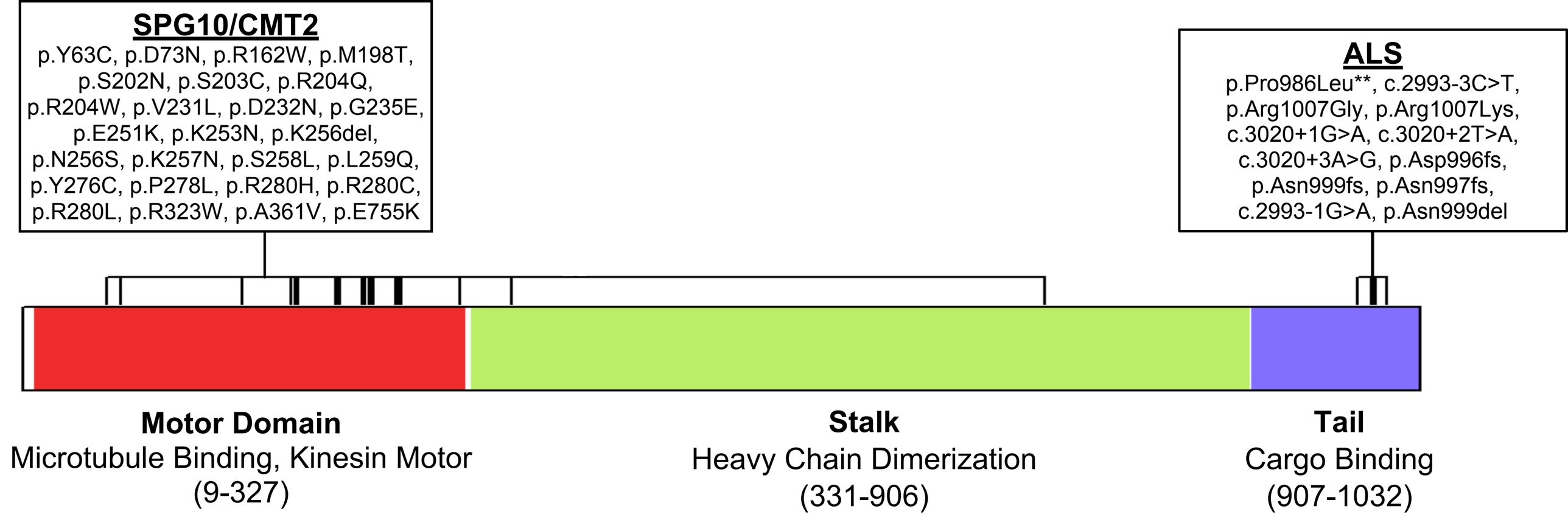
MDNGNATDINDNRSDLPCGYEAEDQAKLFPLHQETAAS*

Mutant Protein:

MDNGVTCRVAMRLRTRPSFSLSTKRQQPANLPHPRLHTCTFSF*

D.




SPG10/CMT2
p.Y63C, p.D73N, p.R162W, p.M198T, p.S202N, p.S203C, p.R204Q, p.R204W, p.V231L, p.D232N, p.G235E, p.E251K, p.K253N, p.K256del, p.N256S, p.K257N, p.S258L, p.L259Q, p.Y276C, p.P278L, p.R280H, p.R280C, p.R280L, p.R323W, p.A361V, p.E755K
ALS
p.Pro986Leu**, c.2993-3C>T, p.Arg1007Gly, p.Arg1007Lys, c.3020+1G>A, c.3020+2T>A, c.3020+3A>G, p.Asp996fs, p.Asn999fs, p.Asn997fs, c.2993-1G>A, p.Asn999del
Motor Domain
Microtubule Binding, Kinesin Motor
(9-327)
Stalk
Heavy Chain Dimerization
(331-906)
Tail
Cargo Binding
(907-1032)

SUPPLEMENTAL MATERIAL

Author Affiliations

1. Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institutes on Aging, National Institutes of Health, Porter Neuroscience Research Center, Bethesda, MD 20892, USA
2. These authors contributed equally to this work
3. Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts, USA.
4. Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
5. Ronald M. Loeb Center for Alzheimer's Disease, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
6. Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan, Italy.
7. Department of Pathophysiology and Transplantation, 'Dino Ferrari' Center - Università degli Studi di Milano, Milan 20122 Italy.
8. Molecular Genetics Section, Laboratory of Neurogenetics, National Institutes on Aging, National Institutes of Health, Porter Neuroscience Research Center, Bethesda, MD 20892, USA
9. Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA
10. Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.
11. Data Tecnica International, Glen Echo, MD, USA
12. Department of Neurology Brain Centre, Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, GA Utrecht, The Netherlands
13. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands.
14. Neurodegenerative Diseases Research Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA
15. United Kingdom Dementia Research Institute Centre, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 125 Coldharbour Lane, Camberwell, SE5 9NU London, U.K.
16. Institute of Genomic Medicine, Catholic University, Roma, Italy
17. Sobell Department of Motor Neuroscience and Movement Disorders, University College London, Institute of Neurology, London, UK
18. Genetics, Genetics and Pharmacogenomics, MRL, Merck & Co., Inc., Boston, MA 02115, USA
19. ALS Center, Salvatore Maugeri Foundation, Milan, Italy
20. 'Rita Levi Montalcini' Department of Neuroscience, University of Turin, Turin, Italy
21. ALS Center, Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy
22. Department of Neurology, Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy
23. Department of Neuroscience, St. Agostino Estense Hospital, Azienda Ospedaliero Universitaria di Modena, Modena, Italy

24. Department of Neurosciences, Ophthalmology, Genetics, Rehabilitation, Maternal and Child Health, Ospedale Policlinico San Martino, Genoa, Italy
25. Department of Medical, Surgical and Neurological Sciences, University of Siena, Siena, Italy
26. ALS Center, Salvatore Maugeri Foundation, IRCCS, Mistretta, Messina, Italy
27. ALS Clinical Research Center, University of Palermo, Palermo, Italy
28. Institute of Neurological Sciences, National Research Council, Mangone, Cosenza, Italy
29. Department of Neurology, Azienda Universitario Ospedaliera di Cagliari and University of Cagliari, Cagliari, Italy
30. Department of Clinical and Experimental Medicine, University of Messina and Nemo Sud Clinical Center for Neuromuscular Diseases, Aurora Foundation, Messina, Italy
31. Department of Basic Medical Sciences, Neurosciences and Sense Organs, University of Bari, Italy
32. Department of Medical, Surgical, Neurological, Metabolic and Aging Sciences, University of Campania "Luigi Vanvitelli", Naples, Italy
33. Il Bene Center for Immunological and Rare Neurological Diseases at Bellaria Hospital, IRCCS, Istituto delle Scienze Neurologiche, Bologna, Italy
34. Neurological Clinic, Marche Polytechnic University, Ancona, Italy
35. Department of Health Sciences, University of Eastern Piedmont, Novara, Italy
36. Department of Neurology, University of Chieti, Chieti, Italy
37. Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD, 21224, USA
38. Department of Neurology, Columbia University, New York, NY 10032 USA.
39. Institute for Genomic Medicine, Columbia University, New York, NY 10032 USA.
40. Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305 USA.
41. Bioverativ, 225 2nd Ave, Waltham, MA 02145
42. HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806 USA.
43. Center for Genomics of Neurodegenerative Diseases (CGND), New York Genome Center, New York, NY
44. Computational Biology, New York Genome Center, New York, NY
45. Gladstone Institute of Neurological Disease, San Francisco, CA, USA
46. Departments of Neurology and Physiology, University of California, San Francisco, CA, USA
47. Gladstone Institutes, San Francisco CA 94158
48. Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA
49. Broad Institute, 415 Main St, Cambridge, Massachusetts, 02142, USA
50. Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA.
51. Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA.
52. Department of Neurobiology and Behavior, Institute of Memory Impairment and Neurological Disorders, University of California, Irvine, CA 92697, USA
53. Department of Psychiatry and Human Behavior, Institute of Memory Impairment and Neurological Disorders, University of California, Irvine, CA 92697, USA.
54. The Heart Institute and Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA
55. Harvard Medical School, Department of Neurology, Massachusetts General Hospital (MGH), Boston, Massachusetts, USA
56. Neurological Clinical Research Institute (NCRI), Massachusetts General Hospital, Boston, Massachusetts, USA

57. Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA.
58. Department of Neurology, The Ohio State University Wexner Medical Center, Columbus, OH, United States
59. Department of Neurology, The Ohio State University Wexner Medical Center, Columbus, OH, United States.
60. Harvard Medical School, Department of Neurology, Massachusetts General Hospital (MGH), Boston, Massachusetts, USA.
61. Neurological Clinical Research Institute (NCRI), Massachusetts General Hospital, Boston, Massachusetts, USA.
62. Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA
63. Department of Neurology, University of Miami
64. Howard Hughes Medical Institute, Chevy Chase, MD 20815
65. Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105
66. Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN 38105
67. 3rd Neurology Unit, Motor Neuron Diseases Center, Fondazione IRCCS Istituto Neurologico 'Carlo Besta', Milan, Italy and Department of Biomedical and Clinical Sciences "Luigi Sacco", University of Milan, Italy
68. Department of Neurology, IRCCS Istituto Auxologico Italiano, Milan, Italy
69. Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, King's College London, London SE5 9RS, UK
70. Neurology Unit, IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.
71. Department of Neurosciences, University of Padova, Padova, Italy
72. Experimental Neurobiology Laboratory, 'C. Mondino' National Institute of Neurology Foundation, IRCCS, Pavia, Italy
73. ALS Center, CHU Bretonneau, Tours university, Tours, France
74. Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki, FIN-02900, Finland
75. Department of Pathology, Medicum, University of Helsinki, and HUSLAB, FIN-00014 Helsinki, Finland
76. Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki, FIN-02900, Finla
77. Dept of Neurology, Salford Royal Foundation Trust, Stott Lane, Salford, M6 8HD
78. Faculty of Human and Medical Sciences, University of Manchester, Manchester M13 9PT, UK
79. Department of Clinical Neuroscience, Institute of Neurology, University College London, London NW3 2PG, UK
80. Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, Queen Square House, London WC1N 3BG, UK
81. Centre for Neuroscience and Trauma, Blizard Institute, Queen Mary University of London, North-East London and Essex Regional Motor Neuron Disease Care Centre, London, E1 2AT, UK
82. Genomics Technology Group, Laboratory of Neurogenetics, National Institutes on Aging, National Institutes of Health, Porter Neuroscience Research Center, Bethesda, MD 20892, USA
83. Department of Neurology, Emory University, Atlanta, GA 30322, USA.
84. Division of Brain Sciences, Department of Medicine, Imperial College London, W120NN, UK

85. Department of Genome analysis and Neurogenetics, Academic Medical Centre, Amsterdam, The Netherlands
86. Department of Neurogenetics and Neurology, Academic Medical Centre, Amsterdam, The Netherlands
87. Unidad de ELA, Instituto de Investigación Hospital Gregorio Marañón de Madrid, SERMAS, Spain.
88. Computational Biology Core, Laboratory of Neurogenetics, National Institutes on Aging, National Institutes of Health, Porter Neuroscience Research Center, Bethesda, MD 20892, USA
89. Neurodegenerative Diseases Research Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA
90. Motor Neuron Disorders Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA
91. Center for Geriatric Medicine, Department of Geriatrics, Neurosciences and Orthopedics, Catholic University of Sacred Heart, Rome 00168, Italy
92. Division of Neurology, Barrow Neurological Institute, Phoenix, Arizona, USA
93. Departments of Neurology, University of Utah School of Medicine, Salt Lake City, Utah, USA
94. Department of Neuroscience, University of California San Diego, La Jolla, CA, USA
95. Mount Sinai Beth Israel Hospital, Mount Sinai School of Medicine, NYC NY
96. Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield, UK
97. Department of Neurology, Neuromuscular Center, Neurological Institute, Cleveland Clinic, Cleveland, Ohio, USA
98. Discipline of Pathology, Brain and Mind Centre, The University of Sydney, 94 Mallett St, Camperdown, NSW 2050 Australia
99. Department of Biochemistry, Penn State College of Medicine, Hershey, PA, USA
100. Department of Pathology, Penn State College of Medicine, Hershey, PA, USA
101. Research and Development Service, Veterans Affairs Boston Healthcare System, Boston, MA
102. Department of Neurology, Program in Behavioral Neuroscience, Boston University School of Medicine, Boston, MA
103. Neurology Service, VA Boston Healthcare System and Boston University Alzheimer's Disease Center, Boston MA 02130
104. Departments of Pathology and Neurology, Johns Hopkins University School of Medicine
105. Sorbonne Universités, UPMC Univ Paris 06, Inserm, CNRS, Institut du cerveau et la moelle (ICM), Assistance Publique-Hôpitaux de Paris (AP-HP) - Hôpital Pitié-Salpêtrière, Paris, France
106. Service de biochimie, CHU de Nîmes, Nîmes, France
107. Department of Neurology, Drexel University College of Medicine, Philadelphia, PA
108. Department of Neurology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA
109. Epidemiology Branch, National Institute of Environmental Health Sciences, NC 27709, USA.
110. KU Leuven – University of Leuven, Department of Neurosciences, Experimental Neurology, Leuven, Belgium
111. VIB, Center for Brain & Disease Research, Laboratory of Neurobiology, Leuven, Belgium
112. Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, California, USA
113. Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
114. Institute of Human Genetics, Technische Universität München, Munich, Germany.
115. Maurice Wohl Clinical Neuroscience Institute, King's College London, Department of Basic and Clinical Neuroscience, London, UK.

116. Institute of Physiology, Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon, Lisbon, Portugal.
117. Department of Neurosciences, Hospital de Santa Maria-CHLN, Lisbon, Portugal.
118. SURFsara, Amsterdam, the Netherlands.
119. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands
120. KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), B-3000 Leuven, Belgium
121. VIB, Center for Brain & Disease Research, Laboratory of Neurobiology, Leuven, Belgium.
122. Population Genetics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Republic of Ireland.
123. Neuromuscular Diseases Unit/ALS Clinic, Kantonsspital St. Gallen, St. Gallen, Switzerland.
124. Department of Neurology, Mayo Clinic Florida, Jacksonville, Florida 32224, USA
125. Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, USA
126. Faculty of Medicine, University of Southampton, Southampton, UK
127. Suna and Inan Kırac Foundation, Neurodegeneration Research Laboratory, Bogazici University, Istanbul, Turkey
128. ALS Unit/Neurology, Hospital San Rafael, Madrid, Spain.
129. Department of Neurology Tel-Aviv Sourasky Medical Centre, Israel
130. Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield, UK
131. Nuffield Department of Clinical Neurosciences, University of Oxford, UK
132. Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK
133. Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Republic of Ireland.
134. Centre for MND Research, Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW 2109, Australia
135. ANZAC Research Institute, Concord Hospital, University of Sydney, Sydney, NSW 2139, Australia
136. Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada
137. Unidad de ELA, Instituto de Investigación Hospital 12 de Octubre de Madrid, SERMAS, and Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER U-723), Madrid, Spain.
138. Tanz Centre for Research of Neurodegenerative Diseases, Division of Neurology, Department of Medicine, University of Toronto, Toronto, ON, M5S 3H2, Canada
139. Division of Neurology, Department of Internal Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, M4N 3M5, Canada
140. Department of Neurology, Penn State Hershey Medical Center, Hershey, PA, USA
141. Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield, UK.
142. Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA
143. Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico 'Carlo Besta', Milan 20133, Italy.
144. Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan 20149, Italy.

145. Department of Neurology, University Hospital KU Leuven (University of Leuven) and Department of Neurosciences, KU Leuven, B-3000 Leuven, Belgium
146. Centro Clinico NeMO. Institute of Neurology. Catholic University, Largo F. Vito 1. 00168 Rome. Italy
147. NEuroMuscular Omnicenter (NEMO), Serena Onlus Foundation, Milan, Italy
148. Neurology Department, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany
149. Department of Pharmacology and Clinical Neuroscience, Umeå University, Umeå SE-90185
150. ALS center, CHU Gui de Chauliac, University of Montpellier, Montpellier, France
151. Department of Neurology, Gui-de-Chauliac Hospital, Montpellier, France
152. Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, USA
153. Department of Neurology, Columbia University, New York, NY, USA
154. Department of Neurology, Emory University School of Medicine, Atlanta, GA
155. Genetics, Genetics and Pharmacogenomics, MRL, Merck & Co., Inc., West Point, PA 19486, USA
156. Neuroscience Institute of Torino, Turin 10124, Italy
157. Lead Contact

* Correspondence: john.landers@umassmed.edu; bryan.traynor@nih.gov

Consortium Members

ALS Sequencing Consortium

Andrew S. Allen[1], Stanley Appel[2], Robert H. Baloh[3], Richard S. Bedlack[4], Braden E. Boone[5], Robert Brown[6], John P. Carulli[7], Alessandra Chesi[8], Wendy K. Chung[9], Elizabeth T. Cirulli[10], Gregory M. Cooper[5], Julien Couthouis[8], Aaron G. Day-Williams[7], Patrick A. Dion[11], Summer Gibson[12], Aaron D. Gitler[8], Jonathan D. Glass[13], David B. Goldstein[14], Yujun Han[15], Matthew B. Harms[16], Tim Harris[17], Sebastian D. Hayes[18], Angela L. Jones[5], Jonathan Keebler[15], Brian J. Krueger[14], Brittany N. Lasseigne[5], Shawn E. Levy[5], Yi-Fan Lu[14], Tom Maniatis[19], Diane McKenna-Yasek[6], Timothy M. Miller[20], Richard M. Myers[5], Slavé Petrovski[14], Stefan M. Pulst[12], Alya R. Raphael[8], John M. Ravits[21], Zhong Ren[14], Guy A. Rouleau[11], Peter C. Sapp[6], Neil A. Shneider[22], Ericka Simpson[2], Katherine B. Sims[7], John F. Staropoli[7], Lindsay L. Waite[5], Quanli Wang[14], Jack R. Wimbish[5], Winnie W. Xin[23]

1. Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC 27708 USA.
2. Stanley Appel Department of Neurology, Houston Methodist, Houston,TX USA.
3. Department of Neurology, Cedars Sinai Medical Center, Los Angeles, CA 90048 USA.
4. Duke ALS Clinic and Durham VA Medical Center, Duke University School of Medicine, Durham, NC 27708 USA.
5. HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806 USA.
6. Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01655 USA.
7. Biogen Idec, Cambridge, MA 02142 USA.
8. Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305 USA.
9. Department of Pediatrics and Medicine, College of Physicians and Surgeons, Columbia University, New York NY 10032

10. Human Longevity, INC, San Diego, CA 92121 USA.
11. Department of Neurology and Neurosurgery, McGill University, Montreal, H3A 2B4 Canada
12. Department of Neurology, University of Utah School of Medicine, Salt Lake City, UT 84112 USA.
13. Department of Neurology, Emory University, Atlanta, GA 30322 USA.
14. Institute for Genomic Medicine, Columbia University, New York, NY 10032 USA.
15. Duke University School of Medicine, Duke University School of Medicine, Durham, NC 27708 USA.
16. Department of Neurology/Institute for Genomic Medicine, Columbia University, New York, NY 11032 USA.
17. Bioverativ, 225 2nd Ave, Waltham, MA 02145
18. Biogen Idec and Harvard Medical School, Cambridge, MA 02142 USA.
19. Department of Biochemistry & Molecular Biophysics, Columbia University, New York, NY 10027 USA.
20. Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110 USA.
21. Department of Neurosciences, University of California, San Diego, La Jolla CA 92093
22. Department of Neurology/Eleanor and Lou Gehrig ALS Center, Columbia University, New York, NY 10032 USA.
23. Neurogenetics DNA Diagnostic Laboratory, Center for Human Genetics Research Massachusetts General Hospital, Boston, MA 02114 USA.

Answer ALS Foundation

Julia Kaye[1,] Steven Finkbeiner[1,2,] Stacia Wyman[3,] Alexander Lenail[4,] Leandro Lima[1,] Ernest Fraenkel[4,5,] Jeffrey D Rothstein[6,] Clive N Svendsen[7,8,] Leslie M Thompson[9,10,] Jenny Van Eyk[11,] Nicholas J Maragakis[6,] James D Berry[12,13,] Jonathan D Glass[14,] Timothy M Miller[15,] Stephen J Kolb[16,17,] Robert H Baloh[18,] Merit Cudkowicz[19,20,] Emily Baxi[6]

1. Gladstone Institute of Neurological Disease, San Francisco, CA, USA
2. Departments of Neurology and Physiology, University of California, San Francisco, CA, USA
3. Gladstone Institutes, San Francisco CA 94158
4. Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA
5. Broad Institute, 415 Main St, Cambridge, Massachusetts, 02142, USA
6. Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA
7. Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA
8. Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA
9. Department of Neurobiology and Behavior, Institute of Memory Impairment and Neurological Disorders, University of California, Irvine, CA 92697, USA
10. Department of Psychiatry and Human Behavior, Institute of Memory Impairment and Neurological Disorders, University of California, Irvine, CA 92697, USA.
11. The Heart Institute and Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA
12. Harvard Medical School, Department of Neurology, Massachusetts General Hospital (MGH), Boston, Massachusetts, USA
13. Neurological Clinical Research Institute (NCRI), Massachusetts General Hospital, Boston, Massachusetts, USA

14. Department of Neurology, Emory University School of Medicine, Atlanta, GA
15. Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA
16. Department of Biological Chemistry & Pharmacology, The Ohio State University Wexner Medical Center, Columbus, OH, United States
17. Department of Neurology, The Ohio State University Wexner Medical Center, Columbus, OH, United States
18. Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, California, USA
19. Harvard Medical School, Department of Neurology, Massachusetts General Hospital (MGH), Boston, Massachusetts, USA.
20. Neurological Clinical Research Institute (NCRI), Massachusetts General Hospital, Boston, Massachusetts, USA

Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium

Michael Benatar[1], J. Paul Taylor[2,3], Gang Wu[4], Evadnie Rampersaud[4], Joanne Wuu[1], Rosa Rademakers[5], Stephan Züchner[6], Rebecca Schule[7], Jacob McCauley[6], Sumaira Hussain[1], Anne Cooley[1], Marielle Wallace[1], Christine Clayman[1], Richard Barohn[8], Jeffrey Statland[8], John Ravits[9], Andrea Swenson[10], Carlayne Jackson[11], Jaya Trivedi[12], Shaida Khan[12], Jonathan Katz[13], Liberty Jenkins[13], Ted Burns[14], Kelly Gwathmey[14], James Caress[15], Corey McMillan[16], Lauren Elman[16], Erik Pioro[17], Jeannine Heckmann[18], Yuen So[19], David Walk[20], Samuel Maiser[21], Jinghui Zhang[4]

1. Department of Neurology, University of Miami
2. Howard Hughes Medical Institute
3. Department of Cell and Molecular Biology, St. Jude Children's Research Hospital
4. Department of Computational Biology, St. Jude Children's Research Hospital
5. Department of Neuroscience, Mayo Clinic Jacksonville
6. Department of Human Genetics, University of Miami
7. University of Tubingen, Hertie Institute for Clinical Brain Research
8. Department of Neurology, University of Kansas Medical Center
9. Department of Neurosciences, University of California San Diego
10. Department of Neurology, University of Iowa
11. Department of Neurology, University of Texas Health Sciences Center San Antonio
12. Department of Neurology & Neurotherapeutics, University of Texas Southwestern Medical Center
13. Department of Neurology, California Pacific Medical Center
14. Department of Neurology, University of Virginia
15. Department of Neurology, Wake Forest University
16. Department of Neurology, University of Pennsylvania
17. Neuromuscular Center, Cleveland Clinic Ohio
18. Department of Neurology, University of Cape Town
19. Department of Neurology, Stanford University
20. Department of Neurology, University of Minnesota
21. Hennepin County Medical Center

French ALS Consortium
William Camu[1], Kevin Mouzat[2], Serge Lumbroso[2], Philippe Corcia[3], Vincent Meininger[3], Gérard Besson[4], Emmeline Lagrange[5], Pierre Clavelou[6], Nathalie Guy[6], Philippe Couratier[7], Patrick Vourch[8], Véronique Danel[9], Emilien Bernard[10], Gwendal Lemasson[11]

1. ALS center, CHU Gui de Chauliac, University of Montpellier, Montpellier, France
2. Service de biochimie, CHU de Nîmes, Nîmes, France
3. ALS Center, CHU Bretonneau, Tours university, Tours, France
4. Service de Neurologie, Hôpital Michalon, La Tronche, France
5. Service d'explorations du système nerveux, Hôpital Michalon, La Tronche, France
6. ALS center, CHU Gabriel Montpied, Clermont-Ferrand, France
7. ALS center, CHU Dupuytren, Limoges, France
8. Département de Biochimie, CHU Bretonneau, Tours, France
9. ALS center, CHRU, Lille, France,
10. ALS center, CHU Wertheimer, Lyon, France
11. ALS center, CHU Pellegrin, Bordeaux, France

Genomic Translation for ALS Care (GTAC) Consortium
Matthew B. Harms[1], David B. Goldstein[2], Neil A. Shneider[3], Stephen Goutman[4], Zachary Simmons[5], Timothy M. Miller[6], Siddharthan Chandran[7], Suvankar Pal[7], George Manousakis[8], Stanley H. Appel[9], Ericka Simpson[9], Leo Wang[10], Robert H Baloh[11], Summer Gibson[12], Richard Bedlack[13], David Lacomis[14], Dhruv Sareen[15], Alexander Sherman[16], Lucie Bruijn[17], Michelle Penny[18]

1. Department of Neurology/Institute of Genomic Medicine/Elanor and Lou Gehrig ALS Center, Columbia University, New York, New York 10032, USA
2. Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA
3. Department of Neurology/Elanor and Lou Gehrig ALS Center, Columbia University, New York, NY 10032, USA
4. Department of Neurology, University of Michigan, Ann Arbor, Michigan 48109, USA
5. Department of Neurology, The Pennsylvania State University College of Medicine, Hershey, Hershey, PA 17033, USA
6. Department of Neurology, Washington University in St. Louis, St. Louis, MO, USA
7. Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, Scotland, UK
8. Department of Neurology, University of Minnesota, Minneapolis, MN, USA
9. Stanley Appel Department of Neurology, Houston Methodist, Houston, TX, USA
10. Department of Neurology, University of Washington, Seattle, WA, USA
11. Department of Neurology, Cedars Sinai Medical Center, Los Angeles, CA, USA
12. Department of Neurology, University of Utah, Salt Lake City, UT, USA
13. Duke ALS Clinic and Durham VA Medical Center, Duke University School of Medicine, Durham, NC 27708, USA
14. Department of Neurology and Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
15. Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA
16. Neurological Clinical Research Institute, Massachusetts General Hospital, Boston, MA, USA

17. Amyotrophic Lateral Sclerosis Association, , Washington DC, , USA
18. Computational Biology, Biogen, Cambridge, MA, USA

ITALSGEN Consortium

Francesco O. Logullo[1], Isabella Simone[2], Giancarlo Logroscino[2,3], Fabrizio Salvi[4], Ilaria Bartolomei[4], Giuseppe Borghero[5], Maria Rita Murru[5], Emanuela Costantino[5], Carla Pani[5], Roberta Puddu[5], Carla Caredda[5], Valeria Piras[5], Stefania Tranquilli[5], Stefania Cuccu[5], Daniela Corongiu[5], Maurizio Melis[5], Antonio Milia[5], Francesco Marrosu[5], Maria Giovanna Marrosu[5], Gianluca Floris[5], Antonino Cannas[5], Stefania Tranquilli[5], Margherita Capasso[6], Claudia Caponnetto[7], Gianluigi Mancardi[7], Paola Origone[7], Paola Mandich[7], Francesca L. Conforti[8], Sebastiano Cavallaro[8], Gabriele Mora[9], Kalliopi Marinou[9], Riccardo Sideri[9], Silvana Penco[10,33], Lorena Mosca[10], Christian Lunetta[11], Giuseppe Lauria Pinter[12], Massimo Corbo[13], Nilo Riva[14], Paola Carrera[14], Paolo Volanti[15], Jessica Mandrioli[16], Nicola Fini[16], Antonio Fasano[16], Lucio Tremolizzo[17], Alessandro Arosio[17], Carlo Ferrarese[17], Francesca Trojsi[18], Gioacchino Tedeschi[19], Maria Rosaria Monsurrò[19], Giovanni Piccirillo[19], Cinzia Femiano[19], Anna Ticca[20], Enzo Ortu[21], Vincenzo La Bella[22], Rossella Spataro[22], Tiziana Colletti[22], Mario Sabatelli[23], Marcella Zollino[23], Amelia Conte[23], Marco Luigetti[23], Serena Lattante[23], Giuseppe Marangi[23], Marialuisa Santarelli[24], Antonio Petrucci[25], Maura Pugliatti[26], Angelo Pirisi[26], Leslie D. Parish[26], Patrizia Occhineri[26], Fabio Giannini[27], Stefania Battistini[27], Claudia Ricci[27], Michele Benigni[27], Tea B. Cau[28], Daniela Loi[28], Andrea Calvo[29], Cristina Moglia[29], Maura Brunetti[30], Marco Barberis[29], Gabriella Restagno[30], Federico Casale[29], Giuseppe Marrali[29], Giuseppe Fuda[29], Irene Ossola[29], Stefania Cammarosano[29], Antonio Canosa[29], Antonio Ilardi[29], Umberto Manera[29], Maurizio Grassano[29], Raffaella Tanel[31], Fabrizio Pisano[32]

1. Neurological Clinic, Marche Polytechnic University, Ancona, Italy
2. Department of Basic Medical Sciences, Neurosciences and Sense Organs, University of Bari, Bari, Italy
3. Department of Clinical and Research Neurology, “Pia Fondazione Cardinal G. Panico” Hospital, Tricase (LE), Bari, Italy
4. Center for Diagnosis and Cure of Rare Diseases, Department of Neurology, IRCCS Institute of Neurological Sciences, Bologna, Italy
5. Department of Neurology, Azienda Universitario Ospedaliera di Cagliari and University of Cagliari, Cagliari, Italy
6. Department of Neurology, University of Chieti, Chieti, Italy
7. Department of Neurosciences, Ophthalmology, Genetics, Rehabilitation, Maternal and Child Health, Ospedale Policlinico San Martino, Genoa, Italy
8. Institute of Neurological Sciences, National Research Council, Mangone, Cosenza, Italy
9. Department of Neurological Rehabilitation, Fondazione Salvatore Maugeri, IRCCS, Istituto Scientifico di Milano, Milano, Italy
10. Department of Laboratory Medicine, Medical Genetics, Niguarda Ca' Granda Hospital, Milan, Italy
11. NeuroMuscular Omnicenter, Serena Onlus Foundation, Milan, Italy
12. Neuroalgology and Headache Unit, IRCCS Fondazione Istituto Neurologico "Carlo Besta”, Milano, Italy
13. Department of Neurorehabilitation Sciences (P.T., M.C.), Casa Cura Policlinico, Milan, Italy
14. Department of Neurology and Institute of Experimental Neurology (INSPE), IRCCS San Raffaele Scientific Institute, Milan, Italy

15. Neurorehabilitation Unit/ALS Center, Scientific Clinical Institutes (ICS) Maugeri, IRCCS, Mistretta, Messina, Italy
16. Department of Neuroscience, S. Agostino-Estense Hospital, University of Modena and Reggio Emilia, Modena, Italy
17. Neurology Unit, School of Medicine and Surgery and NeuroMI, University of Milano-Bicocca, Monza, Italy
18. Department of Medical, Surgical, Neurological, Metabolic and Aging Sciences, University of Campania "Luigi Vanvitelli", Naples, Italy
19. Department of Medical, Surgical Neurological Metabolic and Aging Sciences, Second University of Naples, Naples, Italy
20. Department of Neurology, Azienda Ospedaliera San Francesco, Nuoro, Italy.
21. Division of Neurology, "A. Segni" Hospital, Ozieri, Italy
22. ALS Clinical Research Center, Bio. Ne. C., University of Palermo, Palermo, Italy.
23. Centro Clinico NEMO-Roma. Neurological Institute, Catholic University and I.C.O.M.M. Association for ALS Research, Rome, Italy
24. Department of Medicine, Azienda Complesso Ospedaliero, San Filippo Neri, Rome, Italy
25. Neurology Department, San Camillo Hospital, Rome, Italy
26. Department of Biomedical and Surgical Sciences, Section of Neurological, Psychiatric and Psychological Sciences, University of Ferrara, Ferrara, Italy
27. Department of Medical, Surgical and Neurological Sciences, University of Siena, Siena, Italy
28. Azienda Sanitaria Locale n. 2, Olbia-Tempio, Olbia, Italy
29. Rita Levi Montalcini Department of Neuroscience, Amyotrophic Lateral Sclerosis Center, University of Turin, Turin, Italy; Azienda Ospedaliero Universitaria Città della Salute e della Scienza, Turin, Italy
30. Molecular Genetics Unit, Department of Clinical Pathology, A.S.O. O.I.R.M.-S. Anna, 10126 Turin, Italy
31. Department of Neurology, Santa Chiara Hospital, Trento, Italy
32. Department of Neurological Rehabilitation, "Salvatore Maugeri" Clinical-Scientific Institute, Istituto di Ricovero e Cura a Carattere Scientifico, Veruno, Italy
33. Deceased

NYGC ALS Consortium

Hemali Phatnani[1], Justin Kwan[2], Dhruv Sareen[3], James R. Broach[4], Zachary Simmons[5], Ximena Arcila-Londono[6], Edward B. Lee[7], Vivianna M. Van Deerlin[7], Neil A. Shneider[8], Ernest Fraenkel[9], Lyle W. Ostrow[10], Frank Baas[11], Noah Zaitlen[12], James D. Berry[13], Andrea Malaspina[14], Pietro Fratta[15], Gregory A. Cox[16], Leslie M. Thompson[17], Steve Finkbeiner[18], Efthimios Dardiotis[19], Timothy M. Miller[20], Siddharthan Chandran[21], Suvankar Pal[21], Eran Hornstein[22], Daniel J. MacGowan[23], Terry Heiman-Patterson[24], Molly G. Hammell[25], Nikolaos. A. Patsopoulos[26,27], Joshua Dubnau[28], Avindra Nath[29]

1. Center for Genomics of Neurodegenerative Diseases (CGND), New York Genome Center, New York, NY
2. Department of Neurology, University of Maryland School of Medicine, University of Maryland ALS Clinic, Baltimore, MD"

3. Cedars-Sinai Department of Biomedical Sciences, Board of Governors Regenerative Medicine Institute and Brain Program, Cedars-Sinai Medical Center, and Department of Medicine, University of California, Los Angeles, CA
4. Department of Biochemistry and Molecular Biology, Penn State Institute for Personalized Medicine, The Pennsylvania State University, Hershey, PA
5. Department of Neurology, The Pennsylvania State University, Hershey, PA
6. Department of Neurology, Henry Ford Health System, Detroit, MI
7. Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
8. Department of Neurology, Center for Motor Neuron Biology and Disease, Institute for Genomic Medicine, Columbia University, New York, NY
9. Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA
10. Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD
11. Department of Neurogenetics, Academic Medical Centre, Amsterdam and Leiden University Medical Center, Leiden, The Netherlands
12. Department of Medicine, Lung Biology Center, University of California, San Francisco, CA
13. ALS Multidisciplinary Clinic, Neuromuscular Division, Department of Neurology, Harvard Medical School, and Neurological Clinical Research Institute, Massachusetts General Hospital, Boston, MA
14. Centre for Neuroscience and Trauma, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, and Department of Neurology, Basildon University Hospital, Basildon, United Kingdom
15. Institute of Neurology, National Hospital for Neurology and Neurosurgery, University College London, London, United Kingdom
16. The Jackson Laboratory, Bar Harbor, ME
17. Department of Psychiatry & Human Behavior, Department of Biological Chemistry, School of Medicine, and Department of Neurobiology and Behavior, School of Biological Sciences, University California, Irvine, CA
18. Taube/Koret Center for Neurodegenerative Disease Research, Roddenberry Center for Stem Cell Biology and Medicine, Gladstone Institute
19. Department of Neurology & Sensory Organs, University of Thessaly, Thessaly, Greece
20. Department of Neurology,Washington University in St. Louis, St. Louis, MO
21. Centre for Clinical Brain Sciences, Anne Rowling Regenerative Neurology Clinic, Euan MacDonald Centre for Motor Neurone Disease Research, University of Edinburgh, Edinburgh, United Kingdom
22. Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel
23. Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY
24. Center for Neurodegenerative Disorders, Department of Neurology, the Lewis Katz School of Medicine, Temple University, Philadelphia, PA
25. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Computer Science and Systems Biology Program, Ann Romney Center for Neurological Diseases, Department of Neurology and Division of Genetics in Department of Medicine, Brigham and Women's Hospital, Boston, MA
27. Program in Medical and Population Genetics, Broad Institute, Cambridge, MA
28. Department of Anesthesiology, Stony Brook University, Stony Brook, NY
29. Section of Infections of the Nervous System, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD

Project MinE ALS Sequencing Consortium
Ahmad Al Kheifat[1], Ammar Al-Chalabi[1], Peter Andersen[2], A. Nazli Basak[3], Ian P Blair[4], Adriano Chio[5], Jonathan Cooper-Knock[6], Philippe Corcia[7], Philippe Couratier[8], Mamede de Carvalho[9], Annelot Dekker[10], Vivian Drory[11], Alberto Garcia Redondo[12], Marc Gotkine[13], Orla Hardiman[14,15], Winston Hide[6,16], Alfredo Iacoangeli[17], Glass Jonathan[18], Kevin Kenna[19], Matthew Kiernan[20], Maarten Kooyman[21], John Landers[19], Russell McLaughlin[22], Bas Middelkoop[10], Jonathan Mill[23,1], Miguel Mitne Neto[24], Mattieu Moisse[25,26,27], Jesus Mora Pardina[28], Karen Morrison[29], Stephen Newhouse[17,30], Susana Pinto[9], Sara Pulit[10], Wim Robberecht[25,26,27], Aleksey Shatunov[1], Pamela Shaw[6], Chris Shaw[1], Vincenzo Silani[31], William Sproviero[1], Gijs Tazelaar[10], Nicola Ticozzi[31], Philip van Damme[25,26,27], Leonard van den Berg[10], Rick van der Spek[10], Kristel van Eijk[10], Michael van Es[10], Wouter van Rheenen[10], Joke van Vugt[10], Jan Veldink[10], Markus Weber[32], Kelly L Williams[33], Mayana Zatz[24]

1. Maurice Wohl Clinical Neuroscience Institute, King's College London, Department of Basic and Clinical Neuroscience, London, UK.
2. Umeå University, Umeå, Sweden
3. Suna and Inan Kırac Foundation, Neurodegeneration Research Laboratory, Bogazici University, Istanbul, Turkey
4. Centre for MND Research, Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW 2109, Australia
5. Universita degli Studi dei Torino, Turin, Italy
6. Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK
7. Université François-Rabelais, Tours
8. Université de Limoges, France
9. Instituto de Medicina Molecular, University of Lisbon
10. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands
11. Tel-Aviv Medical Center, Tel-Aviv, Israel
12. Hospital Carlos III, Madrid, Spain
13. Hadassah University Hospital, Jerusalem, Israel
14. Academic Unit of Neurology, Trinity College Dublin, Trinity Biomedical Sciences Institute, Dublin, Republic of Ireland
15. Department of Neurology, Beaumont Hospital, Dublin, Republic of Ireland.
16. Biostatistics Department, Harvard School of Public Health, Boston, Massachusetts, USA.
17. Department of Biostatistics, IoPPN, King's College London, London, US.
18. Emory University, Atlanta, USA
19. Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA
20. Macquarie University, Sydney, Australia
21. SURFsara, Amsterdam, the Netherlands.
22. Population Genetics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Republic of Ireland.
23. University of Exeter Medical School, Exeter University, St Luke's Campus, Magdalen Street, Exeter EX1 2LU, UK
24. Universidade de São Paulo, Brazil
25. KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), B-3000 Leuven, Belgium

26. VIB, Vesalius Research Center, Laboratory of Neurobiology, Leuven, Belgium
27. University Hospitals Leuven, Department of Neurology, Leuven, Belgium
28. Hospital San Rafael, Madrid, Spain
29. Faculty of Medicine, University of Southampton, Southampton, UK
30. Biomedical Research Centre for Mental Health, IoPPN, King's College London, London, UK
31. IRCCS Instituto Auxologico Italiano, Milan, Italy
32. Kantonspittal St Gallen, St Gallen, Switzerland
33. Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW 2109, Australia

SLAGEN Consortium

Vincenzo Silani[1,2], Nicola Ticozzi[1,2], Cinzia Gellera[3], Antonia Ratti[1,2], Franco Taroni[3], Giuseppe Lauria[4], Federico Verde[1], Isabella Fogh[1,5], Cinzia Tiloca[1], Giacomo P. Comi[2,6], Gianni Sorarù[7], Cristina Cereda[8], Sandra D'Alfonso[9], Lucia Corrado[9], Fabiola De Marchi[10], Stefania Corti[2,6], Mauro Ceroni[8,11], Letizia Mazzini[10], Gabriele Siciliano[12], Massimiliano Filosto[13], Maurizio Inghilleri[14], Silvia Peverelli[1], Claudia Colombrita[12], Barbara Poletti[1], Luca Maderna[1], Roberto Del Bo[2,6], Stella Gagliardi[8], Giorgia Querin[7], Cinzia Bertolin[7], Viviana Pensato[3], Barbara Castellotti[3]

1. Department of Neurology, IRCCS Istituto Auxologico Italiano, Milan, Italy
2. Department of Pathophysiology and Transplantation, 'Dino Ferrari' Center, Università degli Studi di Milano, Milan, Italy
3. Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico 'Carlo Besta', Milan, Italy.
4. 3rd Neurology Unit, Motor Neuron Diseases Center, Fondazione IRCCS Istituto Neurologico 'Carlo Besta', Milan, Italy.
5. Maurice Wohl Clinical Neuroscience Institute, King's College London, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, London, UK.
6. Neurology Unit, IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.
7. Department of Neurosciences, University of Padova, Padova, Italy
8. Experimental Neurobiology Laboratory, 'C. Mondino' National Institute of Neurology Foundation, IRCCS, Pavia, Italy
9. Department of Health Sciences, University of Eastern Piedmont, Novara, Italy
10. Department of Neurology, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy
11. Department of Neurological Sciences, University of Pavia, Pavia, Italy
12. Neurology Unit, Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
13. Center for Neuromuscular Diseases and Neuropathies, Unit of Neurology ASST "Spedali Civili", University of Brescia, Brescia, Italy
14. Rare Neuromuscular Diseases Centre, Department of Neurology and Psychiatry, Sapienza University, Rome, Italy

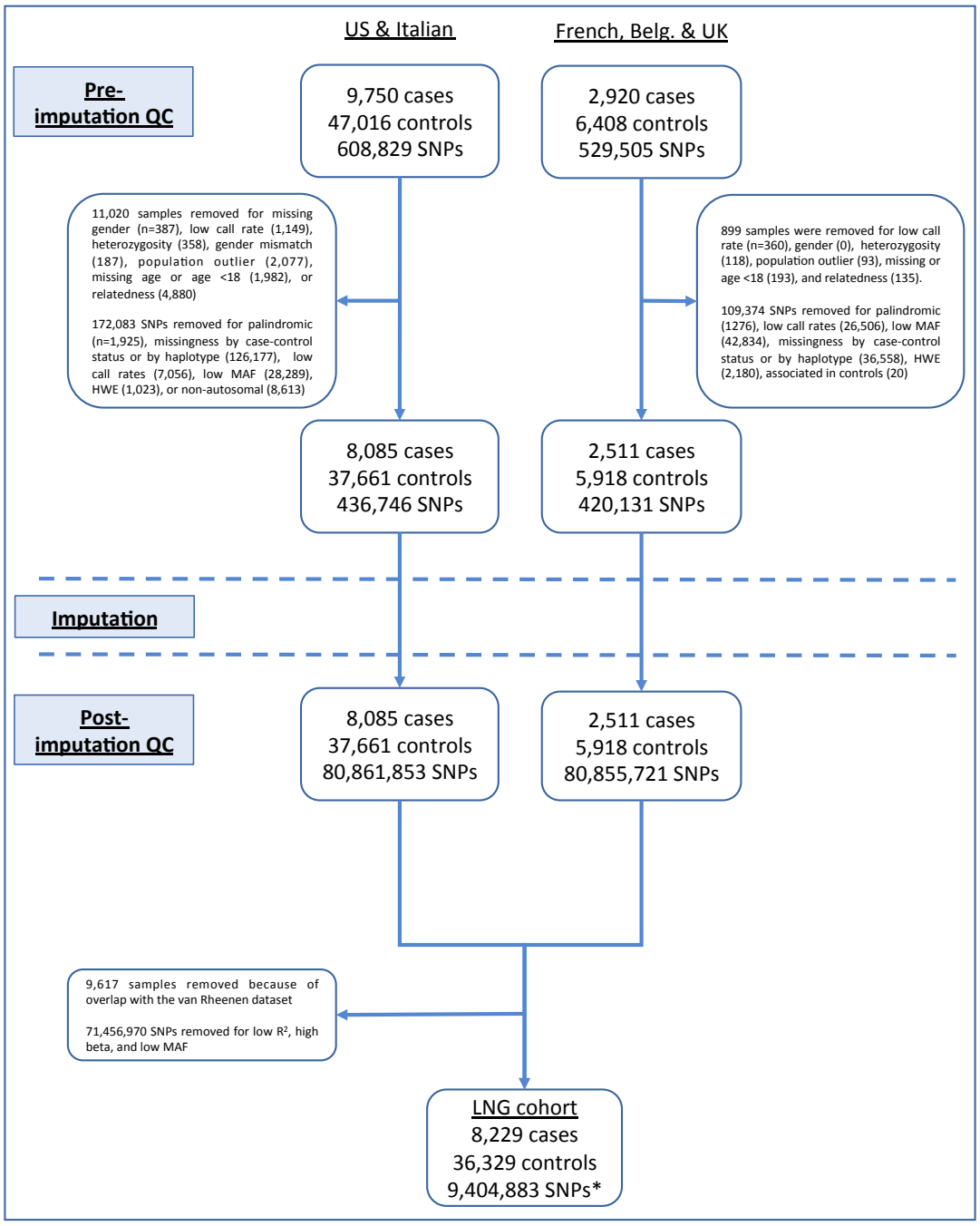


Figure S1. Related to Figure 1; Workflow showing the quality control procedures applied to the present study. *increased to 10,031,630 when merged with the Van Rheenen et al dataset; Belg., Belgium; SNP, single nucleotide polymorphism; MAF, minor allele frequency, HWE, Hardy-Weinberg equilibrium; R^2, R-square value representing imputation precision; LNG, Laboratory of Neurogenetics.

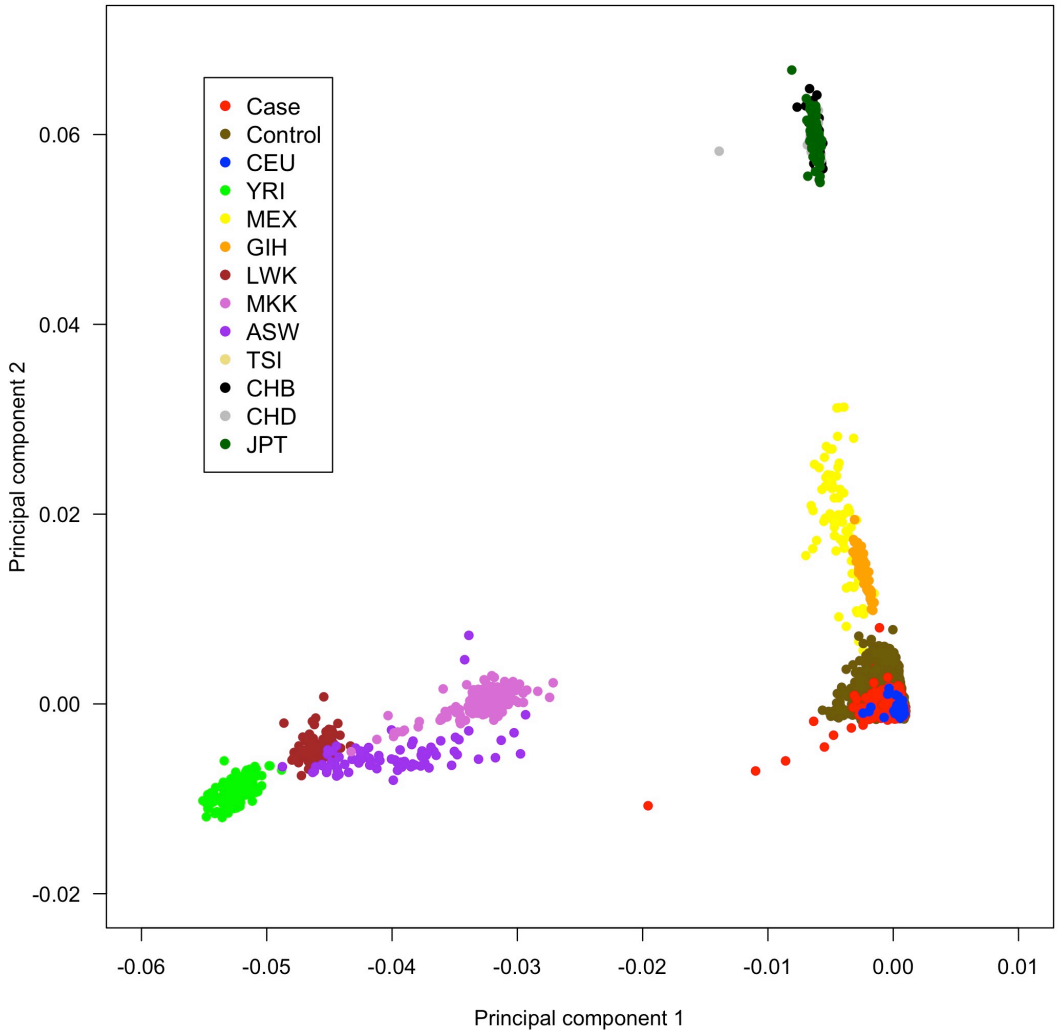


Figure S2. Related to Figure 1; Multi-dimensional scaling plot of the 44,558 genotyped samples included in analysis compared to the HapMap populations.

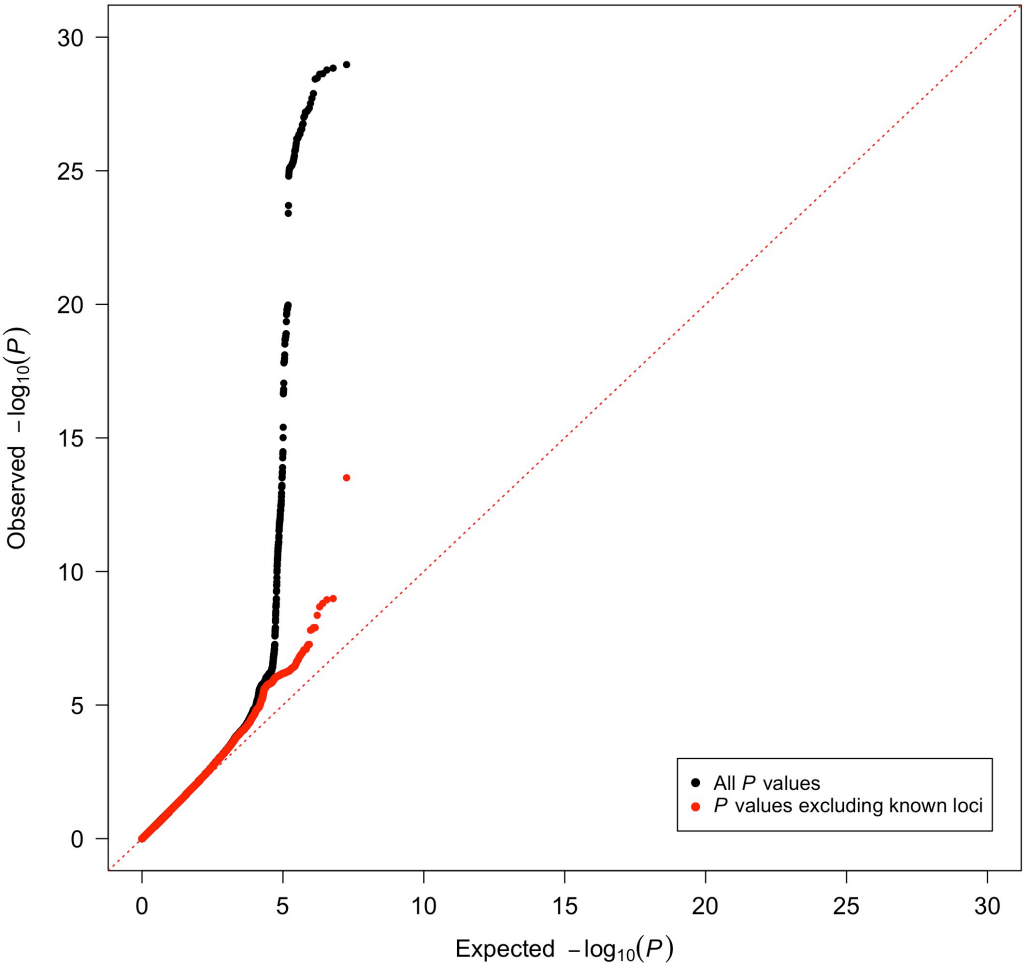


Figure S3. Related to Figure 1; Quartile-Quartile plot of *P*-values from the meta-analysis based on logistic regression analysis. The black curve represents all SNPs, and the red curve represent SNPs after excluding variants within +/- 500 kilobases of the *C9orf72* and the *UNC13A* loci. Raw genome inflation factor (λ) was 1.042 and adjusted λ scaled to 1,000 cases and 1,000 controls was 1.001 based on the entire SNP dataset.

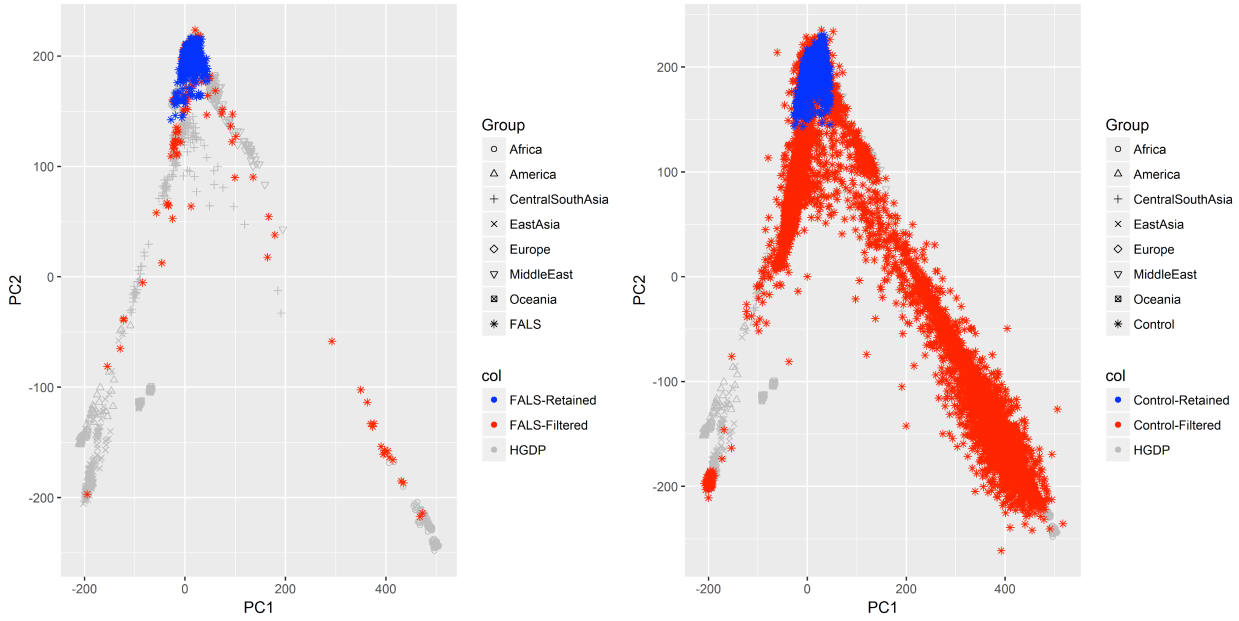


Figure S4. Related to Figure 3; Principal components analysis of samples included in the RVB analysis compared to the Human Diversity Panel. Ancestry filtering of the FALS discovery cohort was performed as follows: LASER was used to generate PCA coordinates for samples from the Human genome diversity panel (HGDP). Samples from the FALS discovery cohort were then mapped to this reference co-ordinate space. The discovery cohort was restricted to cases and controls occurring within 3 standard deviations of the mean for European HGDP samples along principal components 1-4.

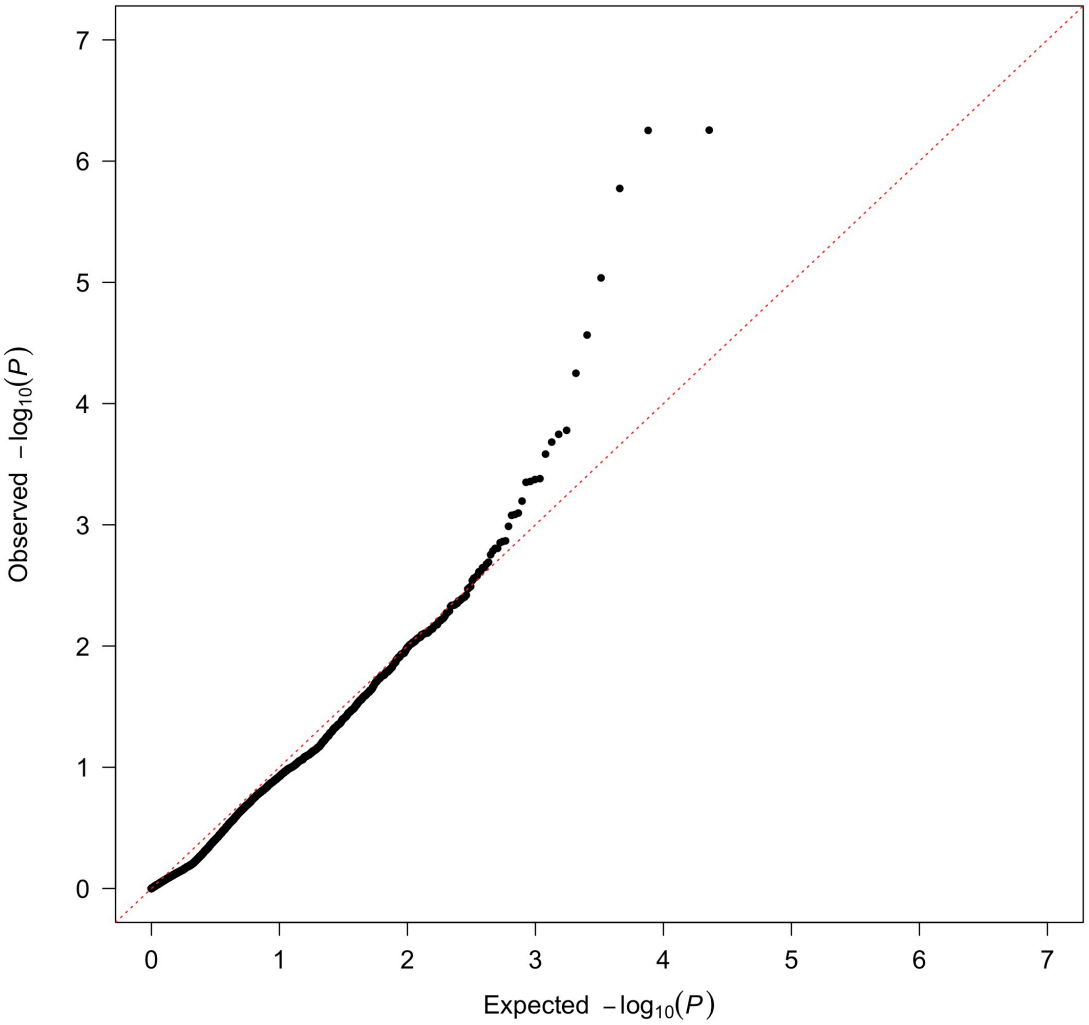


Figure S5. Related to Figure 3; Quartile-Quartile plot of P values from the gene-based rare variant burden analysis of exome data. The genomic inflation factor ($\lambda = 0.93$) was calculated based on the entire gene dataset.

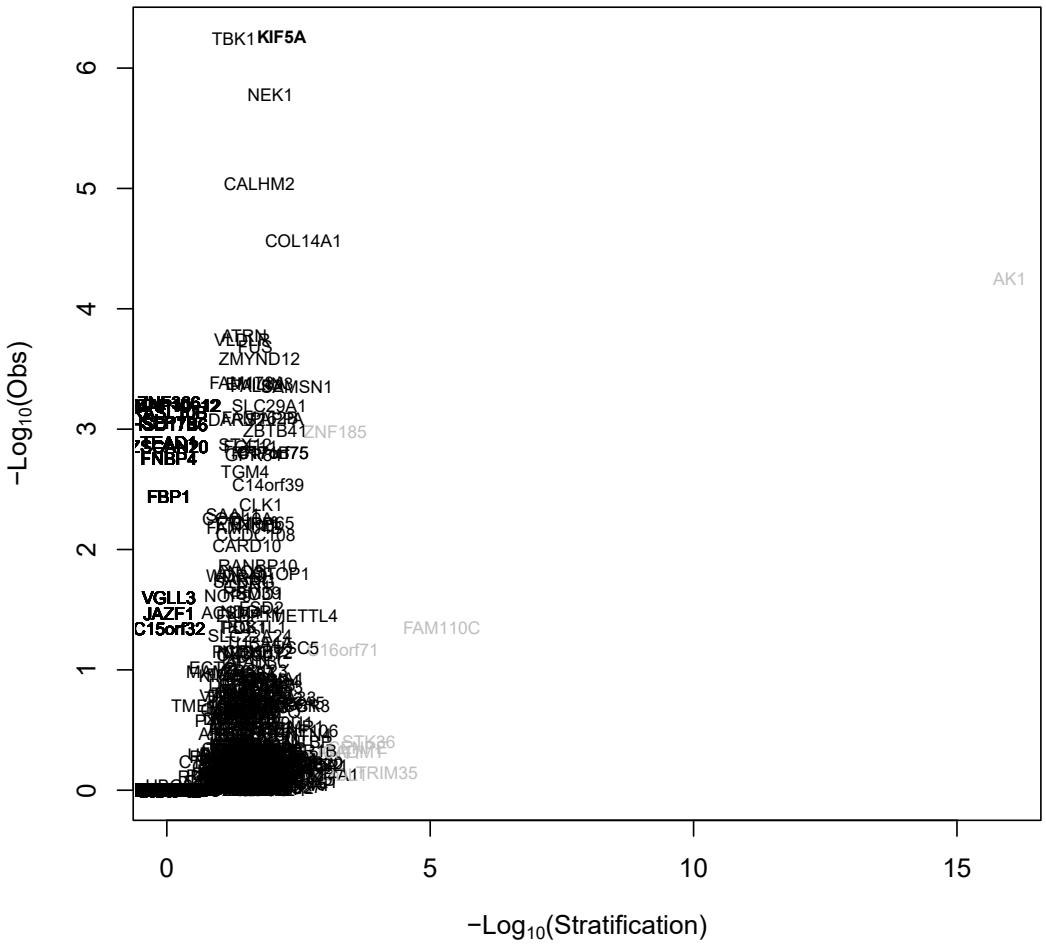


Figure S6. Related to Figure 3; Control-control analyses. *P* values from RVB analysis of FALS cases versus controls (y-axis) are plotted against minimum *P* values from RVB analyses of candidate batch effects (x-axis). To assess the potential impact of batch effects, the sample cohort was divided into 28 pseudo case-control groups based on the sequencing center or associated dbGaP project. Loci showing possible association with non-ALS related batch effects are coloured light grey. No evidence of confounder bias was observed for *KIF5A* or previously reported ALS genes.

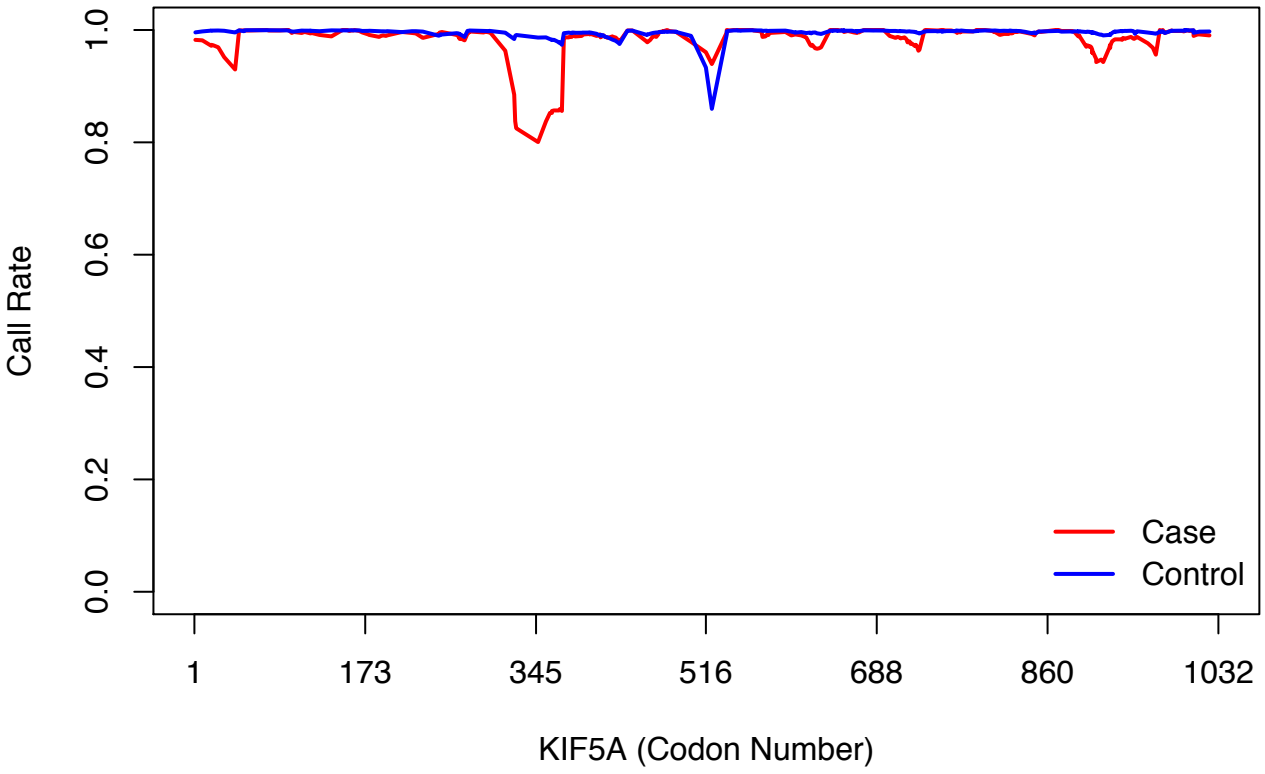


Figure S7. Related to Fgure 3; Plot of variant call rates across the KIF5A protein-coding region in FALS versus controls analyzed by RVB testing.

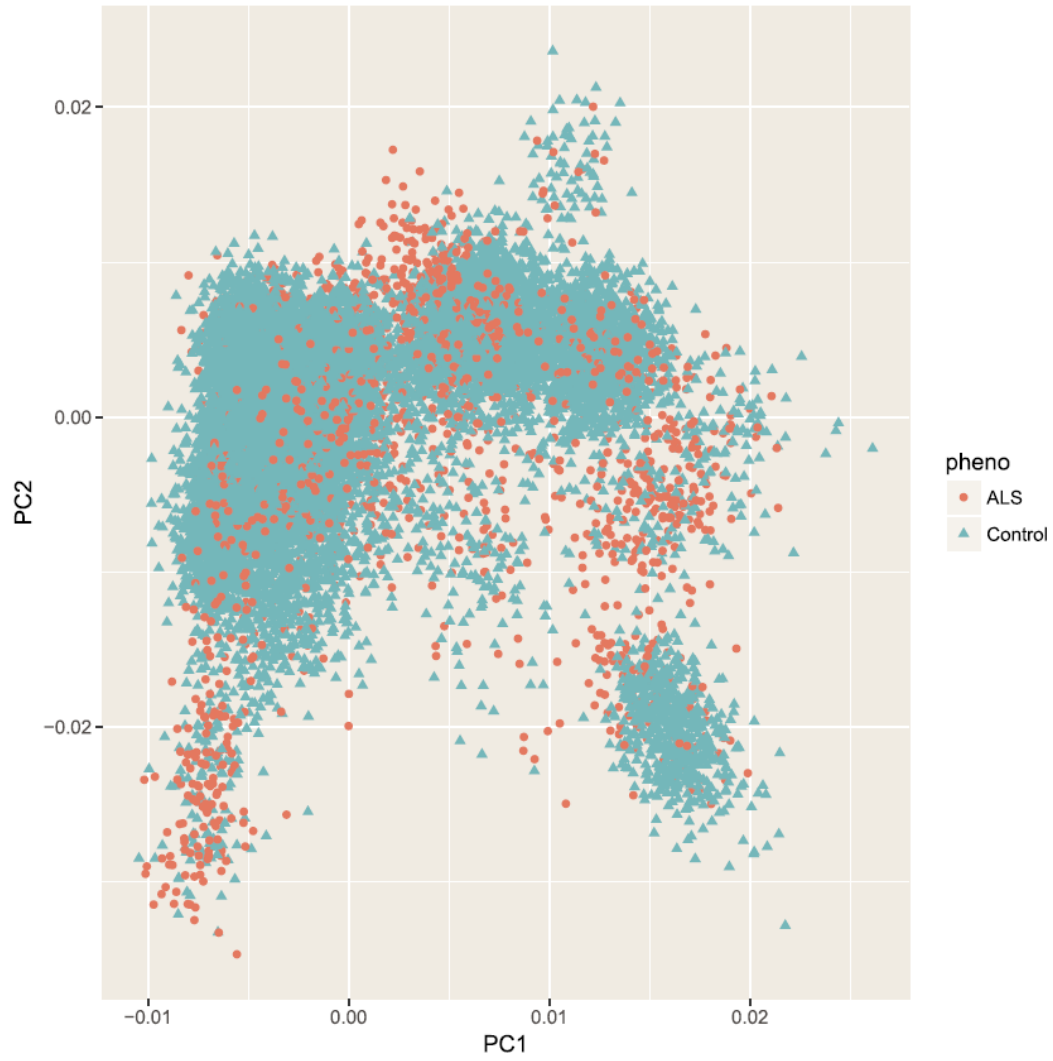


Figure S8. Related to Figure 2; Principal components analysis of samples included in *KIF5A* replication cohort.

Table S1. Related to Figure 1; Demographics and baseline characteristics of patients diagnosed with ALS and control individuals included in the GWAS analysis.

	US		Italian		UK		French & Belgian		Total cohort	
	cases	controls	cases	controls	cases	controls	cases	controls	cases	controls
Sample number	3,777	33,365	2,853	2,143	449	226	1,150	595	8,229	36,329
Female (%)	1,515 (40.1)	23,870 (71.5)	1,239 (43.4)	896 (41.8)	193 (43.0)	109 (48.2)	486 (42.3)	422 (70.9)	3,433 (41.7)	25,297 (69.6)
Age (SD)	58.1 (12.3)	64.2 (13.3)	61.8 (11.8)	50.6 (17.4)	60.3 (12.8)	57.0 (0.0)	60.5 (12.6)	66.9 (16.8)	59.8 (12.3)	63.4 (13.9)
Bulbar-onset* (%)	963 (25.5)	-	741 (26.0)	-	141 (31.4)	-	357 (31)	-	2,202 (26.8)	-
Family history† (%)	458 (12.1)	-	248 (8.7)	-	54 (12.0)	-	195 (17.0)	-	955 (11.6)	-

SD, standard deviation. *Data not available for site of symptom onset for 199 patients. †Data not available for familial history of 154 patients.

Table S2. Related to Figure 1; DbGaP studies contributing to the GWAS analysis.

Accession Number	Study	Sample number	Females (%)	Average age (SD)	Genotyping platform	Ascertainment criteria
phs000001	NEI Age-Related Eye Disease Study (AREDS)	1,644	959 (58.3)	68.2 (4.8)	HumanOmni2.5	Population controls
phs000007	Framingham Cohort	1,298	718 (55.3)	75.7 (8.6)	HumanOmni5	Population controls
phs000187	High Density SNP Association Analysis of Melanoma	1,027	414 (40.3)	51.3 (12.6)	HumanOmniExpress	Population controls
phs000196	CIDR: The NeuroGenetics Research Consortium Parkinson's Disease Study	10	6 (60)	74.3 (18.6)	HumanOmni1	Population controls
phs000292	GENEVA Genetics of Early Onset Stroke (GEOS) Study	89	0 (0)	41.5 (6.4)	HumanOmni1	Population controls
phs000304	Genes and Blood Clotting Study (GABC)	403	259 (64.3)	21.6 (3.3)	HumanOmni1	Population controls
phs000315	Woman's Health Initiative (WHI GARNET)	4,206	4206 (100)	65.7 (6.9)	HumanOmni1	Population controls
phs000368	Polycystic Ovary Syndrome Genetics (POLYGEN)	2,974	2973 (100)	46.8 (15.2)	HumanOmniExpress	Population controls
phs000372	Alzheimer's Disease Genetics Consortium Genome Wide Association Study	533	335 (62.9)	75.8 (9)	HumanOmniExpress	Population controls
phs000394	Autopsy-Confirmed Parkinson Disease GWAS Consortium (APDGC)	299	152 (50.8)	82.1 (12.6)	HumanOmni1	Population controls
phs000397	NIA Long Life Family Study (LLFS)	1,804	957 (53)	65.9 (12.3)	HumanOmni2.5	Population controls
phs000404	The Genetic Architecture of Smoking and Smoking Cessation	81	50 (61.7)	36.6 (5.9)	HumanOmni2.5	Population controls
phs000421	A Genome-Wide Association Study of Fuchs' Endothelial Corneal Dystrophy	497	294 (59.2)	70.4 (10.2)	HumanOmni2.5	Population controls
phs000428	Health and Retirement Study (HRS)	9,394	5437 (57.9)	68.4 (9.4)	HumanOmni2.5	Population controls
phs000615	NINDS Stroke Genetics Network (SiGN)	743	416 (56)	56 (16.1)	HumanOmni5	Population controls
phs000675	GWAS on Selected WHI Hormone Trial European Americans	5,626	5626 (100)	68 (5.9)	HumanOmni1	Population controls
phs000801	NCI Non-Hodgkin Lymphoma GWAS	1,544	790 (51.2)	58.4 (11.6)	HumanOmniExpress	Population controls
phs000869	Barrett's and Esophageal Adenocarcinoma Genetic Susceptibility Study (BEAGESS)	1,174	271 (23.1)	61.3 (10.9)	HumanOmni1	Population controls

Table S3. Related to Figure 1; SNPs achieving genome-wide significance in the GWAS analysis.

SNP Information				Present Study (8,229 Cases / 36,329 Controls)			Van Rheenen *et al.* (12,577 Cases / 23,475 Controls)			Combined Discovery Set (20,806 Cases / 59,804 Controls)			
SNP	**Chr**	**Position**	**Gene**	**Beta [SE]**	**OR [95% CI]**	***P***	**Beta [SE]**	**OR [95% CI]**	***P***	**I^2**	**Beta [SE]**	**OR [95% CI]**	***P***
Novel Loci													
rs117027576	12	57,316,603	*KIF5A*	0.373 [0.096]	1.45 [1.20-1.76]	$1.1x10^{-4}$	0.286 [0.070]	1.33 [1.16-1.53]	$4.3x10^{-5}$	25.6	0.316 [0.057]	1.37 [1.23-1.54]	$2.3x10^{-8}$
rs118082508	12	57,318,819	*KIF5A*	0.374 [0.096]	1.45 [1.20-1.76]	$1.0x10^{-4}$	0.288 [0.070]	1.33 [1.16-1.53]	$3.8x10^{-5}$	25.8	0.317 [0.051]	1.37 [1.23-1.54]	$2.0x10^{-8}$
rs113247976*	12	57,975,700	*KIF5A*	0.381 [0.086]	1.46 [1.23-1.74]	$9.2x10^{-6}$	0.288 [0.066]	1.33 [1.17-1.52]	$1.1x10^{-5}$	0.0	0.322 [0.052]	1.38 [1.24-1.53]	$6.4x10^{-10}$
rs116900480	12	58,656,105	*KIF5A*	0.354 [0.083]	1.42 [1.21-1.68]	$1.9x10^{-5}$	0.294 [0.065]	1.34 [1.18-1.53]	$7.1x10^{-6}$	0.0	0.317 [0.051]	1.37 [1.24-1.52]	$6.6x10^{-10}$
rs142321490	12	58,676,132	*KIF5A*	0.357 [0.082]	1.43 [1.21-1.68]	$1.5x10^{-5}$	0.292 [0.066]	1.34 [1.18-1.53]	$8.0x10^{-6}$	0.0	0.317 [0.056]	1.37 [1.24-1.52]	$6.1x10^{-10}$
Previously Published Loci													
rs10463311	5	150,410,835	*TNIP1*	-0.065 [0.024]	0.94 [0.89-0.98]	$7.8x10^{-3}$	-0.100 [0.020]	0.91 [0.87-0.94]	$8.5x10^{-7}$	0.0	-0.085 [0.016]	0.92 [0.89-0.95]	$4.0x10^{-8}$
rs3849943	9	27,543,382	*C9orf72*	-0.17 [0.024]	0.84 [0.80-0.88]	$1.4x10^{-12}$	-0.181 [0.020]	0.83 [0.80-0.87]	$4.0x10^{-19}$	0.0	-0.176 [0.016]	0.84 [0.81-0.86]	$3.8x10^{-30}$
rs74654358	12	64,881,967	*TBK1*	0.182 [0.058]	1.20 [1.07-1.34]	$1.6x10^{-3}$	0.206 [0.042]	1.23 [1.13-1.34]	$7.7x10^{-7}$	0.0	0.198 [0.034]	1.22 [1.14-1.30]	$4.7x10^{-9}$
rs12973192	19	17,753,239	*UNC13A*	-0.149 [0.026]	0.86 [0.82-0.91]	$1.3x10^{-8}$	-0.106 [0.019]	0.9 [0.87-0.93]	$2.4x10^{-8}$	38.6	-0.121 [0.015]	0.89 [0.86-0.91]	$3.9x10^{-15}$
rs75087725	21	45,753,117	*C21orf2*	0.687 [0.162]	1.99 [1.44-2.75]	$2.2x10^{-5}$	0.479 [0.074]	1.61 [1.39-1.87]	$8.7x10^{-11}$	31.1	0.515 [0.067]	1.67 [1.46-1.91]	$1.8x10^{-14}$

Position is based on Human Genome Assembly build 37. Nearest gene or previously published gene names are included. Chr, chromosome; SE, standard error; OR, odds ratio; 95% CI, 95% confidence interval; *, rs113247976 represents the p.Pro986Leu variant in *KIF5A* (NM_004984.2).

Table S4. Related to Figure 1; Suggestive SNPs with *P* values less than 5.0x10-7 in the GWAS analyses.

SNP Information				Present Study (8,229 Cases / 36,329 Controls)				Van Rheenen *et al.* (12,577 Cases / 23,475 Controls)				Combined Discovery Set (20,806 Cases / 59,804 Controls)			
SNP	**Chr**	**Position**	**Gene**	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***	**Case MAF**	**Control MAF**	**OR [95% CI]**	***P***
rs17070492	8	2,420,855	*LOC101927815*	10.01%	9.76%	1.10 [1.02-1.18]	$1.3x10^{-2}$	9.17%	10.09%	1.16 [1.09-1.23]	$1.3x10^{-6}$	9.50%	9.89%	1.13 [1.08-1.19]	$1.0 x10^{-7}$
rs10139154	14	31,147,498	*SCFD1*	34.10%	31.30%	1.07 [1.03-1.12]	$2.1x10^{-3}$	33.76%	31.17%	1.08 [1.04-1.12]	$1.9x10^{-5}$	33.90%	31.25%	1.08 [1.05-1.11]	$1.4 x10^{-7}$
rs10143310	14	92,540,381	*ATXN3*	24.85%	24.36%	1.09 [1.04-1.015]	$3.3x10^{-4}$	24.04%	22.95%	1.08 [1.04-1.13]	$2.6x10^{-4}$	24.36%	23.81%	1.09 [1.05-1.12]	$3.2 x10^{-7}$
rs9901522	17	14,673,934	*PMP22*	7.08%	6.31%	1.16 [1.06-1.26]	$5.2x10^{-4}$	6.87%	5.97%	1.16 [1.08-1.24]	$4.6x10^{-5}$	6.95%	6.18%	1.16 [1.10-1.22]	$8.6 x10^{-8}$

Table S5. Related to Figure 3; DbGaP/EGA studies contributing to the RVB analysis.

Accession Number	Study	Sample number	Females (%)
phs000179	Genetic Epidemiology of COPD (COPDGene)	2	100%
phs000254	NHLBI GO-ESP: Lung Cohorts Exome Sequencing Project (Cystic Fibrosis)	238	49.6%
phs000281	NHLBI GO-ESP: Women's Health Initiative Exome Sequencing Project (WHI) - WHISP	1904	100%
phs000290	NHLBI GO-ESP: Lung Cohorts Exome Sequencing Project (Pulmonary Arterial Hypertension)	73	82.2%
phs000291	NHLBI GO-ESP: Lung Cohorts Exome Sequencing Project (Lung Health Study of COPD)	332	37%
phs000296	NHLBI GO-ESP: Lung Cohorts Exome Sequencing Project (COPDGene)	285	52.6%
phs000307	NHLBI Framingham Heart Study Allelic Spectrum Project	1317	51.6%
phs000347	NHLBI GO-ESP: Family Studies (Aortic Disease)	29	34.5%
phs000354	NHLBI GO-ESP Family Studies: Pulmonary Arterial Hypertension	9	88.9%
phs000362	NHLBI GO-ESP: Family Studies: (Familial Atrial Fibrillation)	12	16.7%
phs000398	NHLBI GO-ESP: Heart Cohorts Exome Sequencing Project (ARIC)	800	54.6%
phs000400	NHLBI GO-ESP: Heart Cohorts Exome Sequencing Project (CHS)	186	28%
phs000401	NHLBI GO-ESP: Heart Cohorts Exome Sequencing Project (FHS)	348	36.8%
phs000402	NHLBI GO-ESP: Heart Cohorts Exome Sequencing Project (JHS)	296	58.8%
phs000403	NHLBI GO-ESP: Heart Cohorts Exome Sequencing Project (MESA)	259	45.2%
phs000422	NHLBI GO-ESP: Lung Cohorts Exome Sequencing Project (Asthma)	189	65.1%
phs000498	Jackson Heart Study Allelic Spectrum Project	1629	63.8%
phs000518	NHLBI GO-ESP Family Studies: Idiopathic Bronchiectasis	24	70.8%
phs000572	Alzheimer's Disease Sequencing Project (ADSP)	4655	58.8%
phs000632	NHLBI GO-ESP: Family Studies (Hematological Cancers)	19	36.8%
phs000651	Building on GWAS: the U.S. CHARGE consortium - Sequencing (CHARGE-S): FHS	550	61.5%
phs000667	Building on GWAS for NHLBI-Diseases: The U.S. CHARGE Consortium - Sequencing (CHARGE-S): CHS	1209	52.9%
phs000668	Building on GWAS: the U.S. CHARGE consortium - Sequencing (CHARGE-S): ARIC	5497	58.5%
phs000744	Yale Center for Mendelian Genomics (Y CMG)	1944	44.7%
phs000806	MIGen_ExS: Ottawa Heart Study	1966	33.1%
phs000814	MIGen_ExS: Italian Atherosclerosis Thrombosis and Vascular Biology	3591	11.3%
phs000908	Identification of Rare Variants in PD through Whole Exome Sequencing	105	66.7%
phs000917	MIGen_ExS: PROMIS	7298	17.9%
phs001000	MIGen_ExS: U. of Leicester	1081	0%
phs001101	MIGen_ExS: MDC	1075	44.7%
EGAO00000000079	UK10K	4062	65%
phs000101	NIH Exome Sequencing of Familial Amyotrophic Lateral Sclerosis Project	201	45%

Table S6. Related to Figure 2, 3; Quality control filtering of the FALS discovery and *KIF5A* replication cohorts.

FALS discovery cohort

Cohort	Cases	Controls
Initial Sample Set	1,463	41,410
Post HGDP Continental Ancestry Filter	1,397	24,563
Post Call Rate Filter	1,331	20,789
Post Heterozygosity Filter	1,319	20,664
Post Relatedness Filter	1,138	19,494

rs113247976 replication cohort (FALS discovery + ALS WXS/WGS replication cohort)

Cohort	Cases	Controls
Initial Sample Set	12,180*	21,533**
Post Call Rate Filter	11,916	21,050
Post Heterozygosity Filter	11,721	21,028
Post Ancestry Filter (PCA)	11,373	21,009
Post Relatedness & GWAS Checksum Filter	4,160	18,650

* All 1,138 FALS passing QC in FALS discovery cohort + 11,042 additional ALS WXS/WGS cases

** All 19,494 controls passing QC in FALS discovery cohort + 2,039 additional WXS/WGS controls

LOF screen (ALS WXS/WGS replication cohort)

Cohort	Cases	Controls
Initial Sample Set	11,042*	2,039**
Post Call Rate Filter	10,741	2,039
Post Heterozygosity Filter	10,549	2,026
Post Ancestry Filter (PCA)	10,201	2,008
Post Relatedness	9,046	1,955

* 11,042 additional ALS WXS/WGS cases not included in FALS discovery cohort

** 2,039 additional WXS/WGS controls not included in FALS discovery cohort

See Experimental Procedures for further details on filtering parameters.

Table S7. Related to Figure 3; RVB analysis according to mutation type across KIF5A and within gene sub-domains.

Analysis	FALS	Control	OR (95% CI)	*P*
Missense - Full CDS	9 (0.79%)	80 (0.41%)	1.93 (0.915-3.60)	$8.09x10^{-2}$
Missense - Motor Domain	3 (0.26%)	18 (0.09%)	3.27 (0.86-9.25)	$7.74x10^{-2}$
Missense - Microtubule Binding Domain	2 (0.18%)	8 (0.04%)	5.07 (0.95-18.52)	$5.57x10^{-2}$
Missense - Coiled-Coil Domain	3 (0.26%)	55 (0.28%)	1.01 (0.28-2.60)	$9.83x10^{-1}$
Missense - C-Terminal Domain	3 (0.26%)	7 (0.04%)	7.23 (1.74-24.55)	$9.41x10^{-3}$
Loss of Function	6 (0.53%)	3 (0.02%)	32.07 (9.05-135.27)	$5.55x10^{-7}$
Loss of Function (including frameshifts)	8 (0.70%)	3 (0.02%)	41.16 (12.61-167.57)	$3.77x10^{-9}$

FALS, familial ALS; OR, odds ratio; 95% CI, 95% confidence interval; CDS, coding sequence

Table S8. Related to Figure 3; Clinical information of probands and relatives carrying *KIF5A* LOF variants.

Position	Variant	Relation to Proband	DNA Available	Exon	cDNA	Description	Gender	Age of Onset (years)	Site of Onset	Survival (months)	Alive
57,975,729	GA>A	Proband	Y	26	c.2987delA	p.Asp996fs	M	45	n/a	n/a	n/a
57,976,382	C>T	Proband	Y	27	c.2993-3C>T	5' Splice Junction	M	29	L	>264	Y
57,976,382	C>T	Sister	Y	27	c.2993-3C>T	5' Splice Junction	F	52	L	84	N
57,976,382	C>T	Brother	Y	27	c.2993-3C>T	5' Splice Junction	M	18	L	324	N
		Brother	N				M	n/a	L	n/a	N
57,975,731	CA>C	Sporadic	Y	26	c.2989delA	p.Asn997fs	F	50	L	>96	Y
57,976,384	G>A	Sporadic	N	27	c.2993-1G>A	5' Splice Junction	n/a	52	B	n/a	n/a
57,976,385	GA>G	Proband	Y	27	c.2996delA	p.Asn999fs	M	42	L	>12	Y
		Brother	N				M	38	n/a	24	N
57,976,411	A>G	Proband	Y	27	c.3019A>G	p.Arg1007Gly	F	53	L	45	N
57,976,412	G>A	Proband	Y	27	c.3020G>A	p.Arg1007Lys	M	50	L	>108	Y
57,976,412	G>A	Proband	Y	27	c.3020G>A	p.Arg1007Lys	F	50	n/a	>240	Y
57,976,413	G>A	Proband	Y	27	c.3020+1G>A	3' Splice Junction	M	45	B	>220	Y
		Parent	N				n/a	47	n/a	156	N
		Uncle/Aunt	N				n/a	57	n/a	144	N
		Uncle/Aunt	N				n/a	55	n/a	121	N
		Uncle/Aunt	N				n/a	46	n/a	24	N
57,976,414	T>A	Proband	Y	27	c.3020+2T>A	3' Splice Junction	M	46	B	124	N
57,976,414	T>A	Brother	Y	27	c.3020+2T>A	3' Splice Junction	M	48	L	117	N
		Mother	N				F	35	L	144	N
57,976,415	A>G	Proband	Y	27	c.3020+3A>G	3' Splice Junction	M	50	B	54	N

All mutations were heterozygous; Genomic coordinates are based on Human Genome Assembly build 37; Protein change is based on transcript NM_004984.3; n/a, not applicable or not availabl

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
TRIzol Reagent	Thermo Fisher Scientific	Cat# 15596026
RNA-to-cDNA Kit	Applied Biosystems	Cat# 4368814
RNase Inhibitor	Applied Biosystems	Cat# 4374966
OneTaq Hot Start DNA Polymerase	New England BioLabs	Cat# M0481S
Critical Commercial Assays		

HumanOmniExpress-24 DNA Analysis Kit	Illumina	Cat# WG-312-3007
TruSeq Exome Enrichment Kit	Illumina	Cat# FC-121-1096
TruSeq PE Cluster Kit	Illumina	Cat# PE-401-3001
TruSeq SBS Kit	Illumina	Cat# FC-401-3001
TrueSeq DNA PCR-free Kit	Illumina	Cat# 20000902
HiSeq X Ten Reagent Kit	Illumina	Cat# FC-501-2501
Deposited Data		
Experimental Models: Cell Lines		
Experimental Models: Organisms/Strains		
Oligonucleotides		

F1 primer sequence: CAGTGGAGCCACATCTTCTG	Operon Technologies	NA
R1 primer sequence: TCTCTTGGTGGAGAGGGAAA	Operon Technologies	NA
F2 primer sequence: CCAACATGGACAATGGAGTGA	Operon Technologies	NA
Recombinant DNA		
Software and Algorithms		
ASSEDA	Tompson et al., 2007	http://www.cytognomix.com/?post_type=duka&p=2670
BWA	Wellcome Trust Sanger Institute	http://bio-bwa.sourceforge.net
GenABEL	The GenABEL Project	http://www.genabel.org
GATK	Broad Institute	https://software.broadinstitute.org/gatk/
id_geno_checksum	Broad Institute	https://personal.broadinstitute.org/sripke/share_links/checksums_download/
KING	Manichaikul et al., 2010	http://people.virginia.edu/~wc9c/KING/
LASER	University of Michigan	http://csg.sph.umich.edu/chaolong/LASER/

Mach2dat	Marchini and Howie, 2010	https://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output
METAL	Willer et al., 2010	http://csg.sph.umich.edu/abecasis/metal/index.html
Minimac3	Das et al., 2016	https://genome.sph.umich.edu/wiki/Minimac3
PLINK	Chang et al., 2015	http://zzz.bwh.harvard.edu/plink/
R	R Core Team	https://www.r-project.org
SHAPEIT	Delaneau et al., 2013	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html
dbNSFP	Liu et al., 2013	http://varianttools.sourceforge.net/Annotation/DbNSFP
Other		

Highlights

- Loss of function mutations in *KIF5A* are a cause of amyotrophic lateral sclerosis
- ALS-associated *KIF5A* mutations are distinct from HSP and CMT mutations in *KIF5A*
- Identification of *KIF5A* highlights role of cytoskeleton in ALS pathogenesis

eTOC

Using large-scale genome-wide association study and exome sequencing, we identified *KIF5A* as a novel gene associated with ALS. Our data broaden the phenotype resulting from mutations in *KIF5A* and highlight the importance of cytoskeletal defects in the pathogenesis of ALS.