Title: Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways

Authors: Elizabeth T. Cirulli[1]†, Brittany N. Lasseigne[2]†, Slavé Petrovski[3], Peter C. Sapp[4], Patrick A. Dion[5], Claire S. Leblond[5], Julien Couthouis[6], Yi-Fan Lu[3], Quanli Wang[3], Brian J. Krueger[3], Zhong Ren[3], Jonathan Keebler[7], Yujun Han[7], Shawn E. Levy[2], Braden E. Boone[2], Jack R. Wimbish[2], Lindsay W. Jones[2], Angela L. Jones[2], John P. Carulli[8], Aaron Day-Williams[8], John F. Staropoli[8], Winnie W. Xin[9], Alessandra Chesi[6], Alya R. Raphael[6], Diane McKenna-Yasek[4], Janet Cady[10], J.M.B. Vianney de Jong[11], Kevin P. Kenna[12], Bradley N. Smith[13], Simon Topp[13], Jack Miller[13], Athina Gkazi[13], FALS Sequencing Consortium[14], Ammar Al-Chalabi[13], Leonard H. van den Berg[15], Jan Veldink[15], Vincenzo Silani[16], Nicola Ticozzi[16], Christopher E. Shaw[13], Robert H. Baloh[17], Stanley Appel[18], Ericka Simpson[18], Clotilde Lagier-Tourenne[19], Stefan M. Pulst[20], Summer Gibson[20], John Q. Trojanowski[21], Lauren Elman[22], Leo McCluskey[22], Murray Grossman[23], Neil A. Shneider[24], Wendy K. Chung[25], John M. Ravits[26], Jonathan D. Glass[27], Katherine B. Sims[9], Vivianna M. Van Deerlin[21], Tom Maniatis[28], Sebastian D. Hayes[8,29], Alban Ordureau[29], Sharan Swarup[29], John Landers[4], Frank Baas[11], Andrew S. Allen[30], Richard S. Bedlack[31], J. Wade Harper[29], Aaron D. Gitler[6], Guy A. Rouleau[5], Robert Brown[4], Matthew B. Harms[10], Gregory M. Cooper[2], Tim Harris[8]*, Richard M. Myers[2]‡, David B. Goldstein[1]‡

Affiliations:

[1]Center for Applied Genomics and Precision Medicine, Duke University School of Medicine, Durham, NC 27708, USA.

[2]HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA.

[3]Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA.

[4]Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01655, USA.

[5]Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal Qc H3A 2B4, Canada.

[6]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA.

[7]Duke University School of Medicine, Durham, NC 27708, USA.

[8]Biogen Idec, Cambridge, MA 02142, USA.

[9]Neurogenetics DNA Diagnostic Laboratory, Center for Human Genetics Research, Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA.

[10]Neurology, Washington University School of Medicine, St. Louis, MO 63110, USA.

[11]Department of Genome Analysis, Academic Medical Center. Meibergdreef 9, 1105AZ Amsterdam, The Netherlands.

[12]Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Republic of Ireland.

[13]Department of Basic and Clinical Neuroscience, King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, SE5 8AF

[14]The full author list is included at the end of the manuscript.

[15]Department of Neurology, Brain Center Rudolf Magnus, University Medical Centre Utrecht, 3508 GA Utrecht, The Netherlands.

[16]Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan 20149, Italy; Department of Pathophysiology and Transplantation, 'Dino Ferrari' Center - Università degli Studi di Milano, Milan 20122 Italy.

[17]Cedars Sinai Medical Center, Los Angeles, CA 90048, USA.

[18]Houston Methodist Hospital, Houston, TX 77030, USA; Weill Cornell Medical College of Cornell University, New York NY 10065, USA.

[19]Ludwig Institute for Cancer Research and Department of Neurosciences, University of California, San Diego La Jolla, CA 92093, USA.

[20]Department of Neurology, University of Utah School of Medicine; Salt Lake City, UT

84112, USA.

[21]Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA.

[22]Department of Neurology, Penn ALS Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA.

[23]Department of Neurology, Penn Frontotemporal Degeneration Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA.

[24]Department of Neurology, Center for Motor Neuron Biology and Disease, Columbia University, New York, NY 10032, USA.

[25]Department of Pediatrics and Medicine, Columbia University, New York, NY 10032, USA.

[26]Department of Neurosciences, University of California, San Diego La Jolla, CA 92093, USA.

[27]Department of Neurology, Emory University, Atlanta, GA 30322, USA.

[28]Department of Biochemistry & Molecular Biophysics, Columbia University, New York, NY 10027, USA.

[29]Department of Cell Biology, Harvard Medical School, Boston, MA 02481, USA.

[30]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC 27708, USA.

[31]Duke ALS Clinic and Durham VA Medical Center, Durham, NC 27708, USA.

†These authors contributed equally to this work

‡These authors contributed equally to this work

*Correspondence to: tim.harris@biogenidec.com

Abstract: Amyotrophic lateral sclerosis (ALS) is a devastating neurological disease with no effective treatments. Current evidence suggests ALS may have lower genetic heterogeneity than other complex neurological diseases, indicating that moderate-scale sequencing studies may identify new genes contributing to predisposition for ALS. We performed whole exome sequencing of 2,874 ALS patients and compared to 6,405 controls. A number of known ALS genes were associated, and TANK-Binding Kinase 1 (*TBK1)* was identified as an ALS gene.TBK1, a member of the non-canonical IκB kinase family, is known to phosphorylate the ALS gene optineurin (OPTN), a key autophagy component that we find plays an important role in sporadic ALS. These findings identify groups of ALS patients with defined molecular etiologies and suggest avenues for therapeutic intervention.

One Sentence Summary: The exomes of 2,874 patients with ALS and 6,405 controls reveal an ALS predisposition-associated gene.

Main Text: Amyotrophic lateral sclerosis (ALS) is a fatal, progressive neurodegenerative disease characterized by loss of motor neuron function for which there is no effective treatment or definitive diagnostic test (most cases are diagnosed clinically) (*1*). Approximately 10% of ALS cases are familial and inherited in an autosomal dominant, autosomal recessive, or X-linked mode, while the remainder are apparently sporadic(*2, 3*). Dozens of genes collectively explain a majority of familial cases but only a minority (about 10%) of sporadic cases(*2, 3*) (Table 1).

Protein aggregates are a common feature of ALS pathology. These aggregates often include proteins encoded by genes that cause ALS when mutated, including those encoding SOD1, TARDBP, and FUS(*4*). Multiple genes (e.g. *C9orf72, GRN, VCP, UBQLN2, OPTN, NIPA1, SQSTM1*) in addition to *TARDBP* harbor variants pathogenic for TARDBP proteinopathy manifesting as ALS. This pathological TARDBP is part of a common pathway linked to neurodegeneration caused by diverse genetic abnormalities(*5*). While murine models of ALS are limited, silencing certain ALS genes can cause regression of the disease phenotypes and clearance of the protein aggregates(*6*).

Identifying ALS genes

To identify genetic variants associated with ALS, we sequenced the exomes of 2,874 patients with ALS and 6,405 controls. We ran a standard collapsing analysis where the gene was the unit of analysis, and we coded individuals based on the presence or absence of "qualifying" variants in each sequenced gene, where qualifying was defined based on one of six different genetic models (*7*). A total of 17,248 genes had more than one case or control sample with a genetic variant meeting the inclusion criteria for at least one of the genetic models tested (Figs. 1, S1, S2). After correcting for multiple tests, the known ALS gene *SOD1* ($p=7.23 \times 10^{-8}$; dominant coding model) was found to have an experiment-wide significant enrichment of rare variants in ALS cases as compared to controls, with qualifying variants in 0.870% of cases and 0.078% of controls. The genes *HLA-B*, *ZNF729*, *SIRPA*, and *TP53* were found to have a significant enrichment of variants in controls; however, these associations appear to be due to

sequencing differences and to subsets of the controls having been ascertained on the basis of relevant phenotypes.

Based on their associations with ALS in a preliminary discovery phase analysis utilizing 2,843 cases and 4,310 controls, we chose 51 genes (Table S4) for analysis in a further 1,318 cases and 2,371 controls (sequenced using either whole exome or custom capture)(*7*). This analysis definitively identified TANK-Binding Kinase 1 (*TBK1)* as an ALS gene with a discovery association $p=1.13x10^{-5}$, a replication $p=5.78x10^{-7}$, and a combined $p=3.63x10^{-11}$ (dominant not benign model). In the combined dataset, dominant not benign variants in this gene were found in 1.097% of cases and 0.194% of controls, with loss-of-function (LoF) variants occurring in 0.382% of cases and 0.034% of controls.

Analysis of clinical features

We also performed gene-based collapsing analyses to identify genes associated with patients' age of onset, site of onset, and survival time. No genes showed genome-wide significant association with any of these features. When applying multiple-test correction only on known ALS predisposition genes and *TBK1*, we found that D-amino acid oxidase (*DAO)* was significantly associated with survival times, with variant carriers showing shorter survival times ($p=5.5x10^{-7}$, dominant coding model). In mice, DAO is required for the clearance of D-serine. Indeed, D-serine levels are increased in SOD1 mutant mice and spinal cords from humans with familial (FALS) or sporadic ALS (SALS)(*8, 9*). Known FALS mutations seem to reduce DAO activity, leading to neurotoxicity(*10*).

ALS patients with mutations in more than one known ALS gene appear to have a younger age of onset(*11*). While we did not replicate this finding in our dataset, our lack of sequence data for known *C9orf72* carriers, by far the most common ALS variant, as well as our lack of information about *ATXN2* expansions inhibits our ability to adequately assess such an association.

Associations with other ALS genes

Although *SOD1* was the only already known ALS gene to attain a genome-wide significant association in our data, many other known ALS genes showed strong association. For example, rare coding variants in *TARDBP* occurred in 0.661% of the ALS cases and 0.094% of controls in our study, ranking it second to *SOD1* genome-wide (discovery dominant coding model, $p=2.97x10^{-6}$; Fig. 1). The strongest association results for *TARDBP* were found in the model considering any coding or splice variant, while the more restrictive model considering only damaging non-synonymous and LoF variants showed a weaker signal. Moreover, we observed a clear concentration of variants in the 3' protein-coding portion of the gene in the cases as compared with the controls (Fig. 2), consistent with previous studies of *TARDBP* in ALS(*3*).

In *OPTN,* we observed rare damaging variants in 0.620% of ALS cases and 0.228% of controls (combined dominant no benign model; p=0.002). The greatest enrichment was for LoF variants, which occur in 0.334% of cases and 0.114% of controls (combined dominant LoF model, p=0.013). Whereas the initial studies of *OPTN* in ALS found a role in only a few families with a recessive genetic model, subsequent studies identified dominant mutations. Here, dominant acting variants appeared to make a substantial contribution to sporadic disease.

Finally, we also observed a modest excess of qualifying variants in *VCP* (discovery dominant coding model; p=0.022) and of LoF variants in *SPG11* (combined dominant LoF model; p=0.017). The former was driven by variants near the cell division protein 48 domain 2 region, where variants were found in 71% of case variants as compared to 25% of control variants (Fig. 2). Similar to *OPTN*, *SPG11* has previously been reported as a cause of recessive juvenile ALS, but based on our data appears it could play a broader role because these cases did not have early onset(*12*).

We did not identify even a nominal association with other previously reported ALS genes in our dataset, including the recently reported *TUBA4A, MATR3, GLE1,* and *CHCHD10* (Table 1)*(13-17)*. A fraction of our samples were genetically screened for some

of the known genes and had positive cases removed prior to sequencing, which may partially explain the lack of signal (*7*). Additionally, comparison with genes implicated in a recent assessment of the role of 169 previously reported and candidate ALS genes in 242 sporadic ALS cases and 129 controls showed no overlap beyond signals for *SOD1* and *SPG11* (*18*). Some of these previously studied genes are mutated so rarely that even the sample size presented here is not sufficient to detect causal variant enrichment, while others simply show comparable proportions of rare variant among cases and controls. Finally, certain genes did not show associations owing to the nature of the causal variation: most known pathogenic variants in *ATXN2* and *C9orf72* are repeats that cannot be identified in our sequence data.

TBK1, autophagy and neuroinflammation

Previous studies have implicated both *OPTN* (optineurin)(*19*) and *SQSTM1* (p62)(*20*) in ALS. The current study implicates the *TBK1* gene and suggests that *OPTN* is a more important disease gene than previously recognized. These genes play important and interconnected roles in both autophagy and inflammation, emerging areas of interest in ALS research (Fig. 3)(*21-23*). Mutations in SOD1, TARDBP and FUS result in the formation of protein aggregates that stain with anti-SQSTM1 and OPTN antibodies(*24*). These aggregates are thought to lead to a cargo-specific subtype of autophagy involved in the degradation of ubiquitinated proteins through the lysosome(*25*). The SQSTM1 and OPTN proteins function as cargo receptors, recruiting ubiquitinated proteins to the autophagosome via their LC3-interaction region (LIR) motifs. TBK1 phosphorylates both OPTN and SQSTM1 (*26, 27*) and enhances the binding of OPTN to LC3, thereby facilitating the autophagic turnover of ubiquitylated infectious bacteria, a specific cargo of the OPTN adaptor(*28, 29*). Considering that TBK1 co-localizes with OPTN and SQSTM1 in autophagosomes, it is possible that all three proteins associate with protein aggregates in ALS(*28*). Indeed, TBK1 appears to play a role in degradation of protein aggregates by autophagy(*30*). Additionally, OPTN also functions in the autophagic turnover of damaged mitochondria via the PARKIN ubiquitin ligase pathway (*31*). Finally,

VCP, another gene with mutations that cause ALS, also binds to ubiquitinated protein aggregates. Thus, OPTN, SQSTM1, VCP and TBK1 may be critical components of the aggresome pathway for protein degradation. It seems that defects in this pathway can be selective for motor neuron death, in some cases apparently sparing other neuronal cell types.

In addition to their roles in autophagy, OPTN, SQSTM1 and TBK1 all function in the NF-κB pathway (Fig. 3)(*23, 32*). For example, IκB Kinases (IκKα and IκKβ) phosphorylate the IκK-related kinases that include TBK1, which in turn phosphorylate the IκB kinases, suppressing their activity in a negative autoregulatory feedback loop(*33*). TBK1 also phosphorylates and activates the transcription factor IRF3(*34-36*), and both NF-κB and IRF3 stimulate transcription of many inflammatory genes, including interferon-β (*37*). Thus, pathogenic variants in *OPTN*, *SQSTM1*, or *TBK1* would be expected to lead to defects in autophagy and in key innate immunity signaling pathways, leading to defects in the activation of cytokines and downstream genes.

The simple observation of enrichment of qualifying variants in patients shows that some of the variants we have identified influence risk of disease. We cannot determine, however, the extent to which they may interact with any others variants or other risk factors in determining risk. We therefore focus on estimating the proportion of patients in which variants in the relevant genes either "cause or contribute" to disease by subtracting the proportion of controls with qualifying variants in a gene from the proportion of cases with such variants. While we saw no enrichment of case variants in *SQSTM1*, variants in *OPTN* and *TBK1* were estimated to explain or contribute to 1.30% of cases in our dataset when taken together (combined data), suggesting an important subgroup of patients that may have a common biological etiology. No individual ALS cases had qualifying variants in more than one of these three genes.

The case variants found in *OPTN* and *TBK1* were largely heterozygous and LoF, suggesting that a reduction in trafficking of cargo through the autophagosomal pathway or disruption of autophagosomal maturation may promote disease. While the most

obvious enrichment of case variants in *TBK1* was seen for LoF, there was also a signal for non-synonymous variants, which were found in 1.026% of cases and 0.365% of controls (combined data). If any of these non-synonymous variants are selective LoF for specific TBK1 functions as opposed to complete LoF variants, they may help elucidate which TBK1 function is most relevant to disease. We did not observe any clear concentration of qualifying variants in any part of the *TBK1* gene (Fig. 4).

NEK1 associates with ALS2 and VAPB

Although no additional genes showed sufficiently strong evidence to be definitively declared disease genes at this point, some of the strongly associated genes identified here may be securely implicated as sample sizes increase. One gene of particular interest is NIMA-Related Kinase 1 (*NEK1*). This gene just reached experiment-wide significance in the combined discovery and replication data sets (discovery $p=1.08x10^{-6}$; replication p=0.001, combined $p=3.20x10^{-9}$; dominant LoF model). In the combined dataset, dominant LoF variants in this gene were found in 0.835% of cases and 0.091% of controls (Fig. S3). Additional studies are needed to confirm this suggestive association. Even if LoF variants in this gene do predispose to ALS, their relatively high prevalence in our controls and in public databases indicates that such variants have quite low penetrance given that the lifetime prevalence of ALS is approximately 0.2%.

NEK1 is a widely expressed multi-functional kinase linked to multiple cellular processes, but it has not been linked to ALS. In an unbiased proteomic search for NEK1-interacting proteins in HEK293T cells, we discovered an interaction between NEK1 and two widely expressed proteins previously found to be mutated in familial ALS – the RAB guanine nucleotide exchange factor ALS2 (also called Alsin) involved in endosomal trafficking and the endoplasmic reticulum protein VAPB involved in lipid trafficking to the plasma membrane (Fig. S4A,B, Table S5) (*38*). ALS2 reciprocally associated with NEK1 in HEK293T cells, and both ALS2 and VAPB associated with NEK1 reciprocally in a mouse neuronal cell line NSC-34 (Fig. S4C-E).

Other top genes showing interesting association patterns but not obtaining genome-wide significance included *ENAH*, with variants in 0.262% of cases and 0.011% of controls (combined dataset) (discovery $p=1.83x10^{-5}$; replication p=0.133; combined $p=9.59x10^{-6}$; recessive no benign model); *CRLF3*, with variants in 0.452% of cases and 0.094% of controls (discovery p= 0.0002; dominant coding model); *DNMT3A*, with variants in 1.002% of cases and 0.456% of controls (combined dataset) (discovery p=0.0002; replication p=0.261; combined p=0.0002; dominant not benign model); and *LGALSL*, with variants in 0.382% of cases and 0.068% of controls (combined data) (discovery p=0.0002; replication p=0.356; combined p=0.0002; dominant coding model).

Conclusions:

Here, we have implicated *TBK1* as an ALS gene, providing insight into disease biology and suggesting possible directions for drug screening programs. We have also shown that *OPTN* plays a broader role in ALS than previously recognized. Both TBK1 and OPTN are involved in autophagy, with TBK1 possibly playing a crucial role in autophagosome maturation as well as the clearance of protein aggregates(*27, 30*). These observations highlight a critical role of autophagy and/or inflammation in disease predisposition. It is also noteworthy that many drugs have been developed that act on TBK1-mediated pathways owing to their role in tumor cell survival(*39*).

We also provide a large genetic dataset for ALS, which suggests other possible ALS genes and provides a substantial collection of pathogenic variants across ALS genes. The identification of *TBK1* and the expanded role for *OPTN* as ALS genes suggests that signaling systems linked with autophagy and endosome/lipid trafficking may be prominently affected pathways in ALS. Motor neurons may be particularly sensitive to processes regulated by these trafficking pathways. It remains unclear why so many ALS genes induce formation of TARDBP inclusions, the hallmark central nervous system lesions of >95% of ALS patients. Accumulating data, however, suggest that TARDBP proteinopathy associated with toxic losses or gains of TARDBP protein function is a common pathway linked to neurodegeneration caused by multiple genetic

abnormalities(*5, 40*).

Here, an exome sequencing study has successfully identified variants that definitively predispose humans to a sporadic, complex human disease. Suggestive evidence in genes that do not yet achieve significant associations strongly motivates performing even larger exome sequencing studies in ALS. There is reason for optimism that such studies will begin to fill in an increasingly complete picture of the key genes implicated in ALS, providing multiple entry points for therapeutic intervention (Fig. 3). It is also likely that whole genome sequencing (especially with longer reads) will prove of particular value in ALS, given that there are many causal variants refractory to identification using contemporary exome sequencing. Finally, we note that effective studies will depend critically on the control samples available. For example, here we used the recently released ExAC dataset of 60,500 samples to hone in on extremely rare variants in our samples (*41*). Well-characterized, publically available control sample sets will be of particular importance for further discovery of variants associated with complex traits, in particular for whole genome sequencing studies.

References and Notes:

1. L. Poppe, L. Rue, W. Robberecht, L. Van Den Bosch, Translating biological findings into new treatment strategies for amyotrophic lateral sclerosis (ALS). *Experimental neurology*, (2014); published online EpubJul 11 (10.1016/j.expneurol.2014.07.001).
2. S. Chen, P. Sayana, X. Zhang, W. Le, Genetics of amyotrophic lateral sclerosis: an update. *Molecular neurodegeneration* **8**, 28 (2013)10.1186/1750-1326-8-28).
3. A. E. Renton, A. Chio, B. J. Traynor, State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci* **17**, 17-23 (2014); published online EpubJan (10.1038/nn.3584).
4. A. Al-Chalabi, A. Jones, C. Troakes, A. King, S. Al-Sarraj, L. H. van den Berg, The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta neuropathologica* **124**, 339-352 (2012); published online EpubSep (10.1007/s00401-012-1022-4).
5. M. Neumann, Frontotemporal lobar degeneration and amyotrophic lateral sclerosis: molecular similarities and differences. *Revue neurologique* **169**, 793-798 (2013); published online EpubOct (10.1016/j.neurol.2013.07.019).
6. R. A. Smith, T. M. Miller, K. Yamanaka, B. P. Monia, T. P. Condon, G. Hung, C. S. Lobsiger, C. M. Ward, M. McAlonis-Downes, H. Wei, E. V. Wancewicz, C. F. Bennett, D. W. Cleveland, Antisense oligonucleotide therapy for neurodegenerative disease. *The Journal of clinical investigation* **116**, 2290-2296 (2006); published online EpubAug (10.1172/JCI25424).
7. Supplementary Materials and Methods.
8. J. Sasabe, T. Chiba, M. Yamada, K. Okamoto, I. Nishimoto, M. Matsuoka, S. Aiso, D-serine is a key determinant of glutamate toxicity in amyotrophic lateral sclerosis. *The EMBO journal* **26**, 4149-4159 (2007); published online EpubSep 19 (10.1038/sj.emboj.7601840).
9. M. Thompson, J. C. Marecki, S. Marinesco, V. Labrie, J. C. Roder, S. W. Barger, J. P. Crow, Paradoxical roles of serine racemase and D-serine in the G93A mSOD1 mouse model of amyotrophic lateral sclerosis. *Journal of neurochemistry* **120**, 598-610 (2012); published online EpubFeb (10.1111/j.1471-4159.2011.07601.x).
10. P. Paul, J. de Belleroche, The role of D-amino acids in amyotrophic lateral sclerosis pathogenesis: a review. *Amino acids* **43**, 1823-1831 (2012); published online EpubNov (10.1007/s00726-012-1385-9).
11. J. Cady, P. Allred, T. Bali, A. Pestronk, A. Goate, T. M. Miller, R. D. Mitra, J. Ravits, M. B. Harms, R. H. Baloh, Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. *Annals of neurology*, (2014); published online EpubNov 7 (10.1002/ana.24306).
12. H. Daoud, S. Zhou, A. Noreau, M. Sabbagh, V. Belzil, A. Dionne-Laporte, C. Tranchant, P. Dion, G. A. Rouleau, Exome sequencing reveals SPG11 mutations causing juvenile ALS. *Neurobiology of aging* **33**, 839 e835-839 (2012); published online EpubApr (10.1016/j.neurobiolaging.2011.11.012).

13. B. N. Smith, N. Ticozzi, C. Fallini, A. S. Gkazi, S. Topp, K. P. Kenna, E. L. Scotter, J. Kost, P. Keagle, J. W. Miller, D. Calini, C. Vance, E. W. Danielson, C. Troakes, C. Tiloca, S. Al-Sarraj, E. A. Lewis, A. King, C. Colombrita, V. Pensato, B. Castellotti, J. de Belleroche, F. Baas, A. L. Ten Asbroek, P. C. Sapp, D. McKenna-Yasek, R. L. McLaughlin, M. Polak, S. Asress, J. Esteban-Perez, J. L. Munoz-Blanco, M. Simpson, S. Consortium, W. van Rheenen, F. P. Diekstra, G. Lauria, S. Duga, S. Corti, C. Cereda, L. Corrado, G. Soraru, K. E. Morrison, K. L. Williams, G. A. Nicholson, I. P. Blair, P. A. Dion, C. S. Leblond, G. A. Rouleau, O. Hardiman, J. H. Veldink, L. H. van den Berg, A. Al-Chalabi, H. Pall, P. J. Shaw, M. R. Turner, K. Talbot, F. Taroni, A. Garcia-Redondo, Z. Wu, J. D. Glass, C. Gellera, A. Ratti, R. H. Brown, Jr., V. Silani, C. E. Shaw, J. E. Landers, Exome-wide Rare Variant Analysis Identifies TUBA4A Mutations Associated with Familial ALS. *Neuron* **84**, 324-331 (2014); published online EpubOct 22 (10.1016/j.neuron.2014.09.027).
14. J. O. Johnson, E. P. Pioro, A. Boehringer, R. Chia, H. Feit, A. E. Renton, H. A. Pliner, Y. Abramzon, G. Marangi, B. J. Winborn, J. R. Gibbs, M. A. Nalls, S. Morgan, M. Shoai, J. Hardy, A. Pittman, R. W. Orrell, A. Malaspina, K. C. Sidle, P. Fratta, M. B. Harms, R. H. Baloh, A. Pestronk, C. C. Weihl, E. Rogaeva, L. Zinman, V. E. Drory, G. Borghero, G. Mora, A. Calvo, J. D. Rothstein, I. Consortium, C. Drepper, M. Sendtner, A. B. Singleton, J. P. Taylor, M. R. Cookson, G. Restagno, M. Sabatelli, R. Bowser, A. Chio, B. J. Traynor, Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci* **17**, 664-666 (2014); published online EpubMay (10.1038/nn.3688).
15. H. M. Kaneb, A. W. Folkmann, V. V. Belzil, L. E. Jao, C. S. Leblond, S. L. Girard, H. Daoud, A. Noreau, D. Rochefort, P. Hince, A. Szuto, A. Levert, S. Vidal, C. Andre-Guimont, W. Camu, J. P. Bouchard, N. Dupre, G. A. Rouleau, S. R. Wente, P. A. Dion, Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Hum Mol Genet*, (2014); published online EpubOct 24 (10.1093/hmg/ddu545).
16. A. Chaussenot, I. Le Ber, S. Ait-El-Mkadem, A. Camuzat, A. de Septenville, S. Bannwarth, E. C. Genin, V. Serre, G. Auge, F. T. D. The French research network on, A. L. S. Ftd, A. Brice, J. Pouget, V. Paquis-Flucklinger, Screening of CHCHD10 in a French cohort confirms the involvement of this gene in frontotemporal dementia with amyotrophic lateral sclerosis patients. *Neurobiol Aging*, (2014); published online EpubJul 24 (10.1016/j.neurobiolaging.2014.07.022).
17. S. Bannwarth, S. Ait-El-Mkadem, A. Chaussenot, E. C. Genin, S. Lacas-Gervais, K. Fragaki, L. Berg-Alonso, Y. Kageyama, V. Serre, D. G. Moore, A. Verschueren, C. Rouzier, I. Le Ber, G. Auge, C. Cochaud, F. Lespinasse, K. N'Guyen, A. de Septenville, A. Brice, P. Yu-Wai-Man, H. Sesaki, J. Pouget, V. Paquis-Flucklinger, A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain* **137**, 2329-2345 (2014); published online EpubAug (10.1093/brain/awu138).
18. J. Couthouis, A. R. Raphael, R. Daneshjou, A. D. Gitler, Targeted exon capture and sequencing in sporadic amyotrophic lateral sclerosis. *PLoS genetics* **10**,

e1004704 (2014); published online EpubOct (10.1371/journal.pgen.1004704).
19. H. Maruyama, H. Morino, H. Ito, Y. Izumi, H. Kato, Y. Watanabe, Y. Kinoshita, M. Kamada, H. Nodera, H. Suzuki, O. Komure, S. Matsuura, K. Kobatake, N. Morimoto, K. Abe, N. Suzuki, M. Aoki, A. Kawata, T. Hirai, T. Kato, K. Ogasawara, A. Hirano, T. Takumi, H. Kusaka, K. Hagiwara, R. Kaji, H. Kawakami, Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* **465**, 223-226 (2010); published online EpubMay 13 (10.1038/nature08971).
20. S. L. Rea, V. Majcher, M. S. Searle, R. Layfield, SQSTM1 mutations--bridging Paget disease of bone and ALS/FTLD. *Experimental cell research* **325**, 27-37 (2014); published online EpubJul 1 (10.1016/j.yexcr.2014.01.020).
21. H. Maruyama, H. Kawakami, Optineurin and amyotrophic lateral sclerosis. *Geriatrics & gerontology international* **13**, 528-532 (2013); published online EpubJul (10.1111/ggi.12022).
22. M. Thomas, J. Alegre-Abarrategui, R. Wade-Martins, RNA dysfunction and aggrephagy at the centre of an amyotrophic lateral sclerosis/frontotemporal dementia disease continuum. *Brain* **136**, 1345-1360 (2013); published online EpubMay (10.1093/brain/awt030).
23. D. Kachaner, P. Genin, E. Laplantine, R. Weil, Toward an integrative view of Optineurin functions. *Cell cycle* **11**, 2808-2818 (2012); published online EpubAug 1 (10.4161/cc.20946).
24. B. A. Keller, K. Volkening, C. A. Droppelmann, L. C. Ang, R. Rademakers, M. J. Strong, Co-aggregation of RNA binding proteins in ALS spinal motor neurons: evidence of a common pathogenic mechanism. *Acta neuropathologica* **124**, 733-747 (2012); published online EpubNov (10.1007/s00401-012-1035-z).
25. E. L. Scotter, C. Vance, A. L. Nishimura, Y. B. Lee, H. J. Chen, H. Urwin, V. Sardone, J. C. Mitchell, B. Rogelj, D. C. Rubinsztein, C. E. Shaw, Differential roles of the ubiquitin proteasome system and autophagy in the clearance of soluble and aggregated TDP-43 species. *Journal of cell science* **127**, 1263-1278 (2014); published online EpubMar 15 (10.1242/jcs.140087).
26. S. Morton, L. Hesson, M. Peggie, P. Cohen, Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. *FEBS letters* **582**, 997-1002 (2008); published online EpubMar 19 (10.1016/j.febslet.2008.02.047).
27. M. Pilli, J. Arko-Mensah, M. Ponpuak, E. Roberts, S. Master, M. A. Mandell, N. Dupont, W. Ornatowski, S. Jiang, S. B. Bradfute, J. A. Bruun, T. E. Hansen, T. Johansen, V. Deretic, TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. *Immunity* **37**, 223-234 (2012); published online EpubAug 24 (10.1016/j.immuni.2012.04.015).
28. P. Wild, H. Farhan, D. G. McEwan, S. Wagner, V. V. Rogov, N. R. Brady, B. Richter, J. Korac, O. Waidmann, C. Choudhary, V. Dotsch, D. Bumann, I. Dikic, Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science* **333**, 228-233 (2011); published online EpubJul 8 (10.1126/science.1205405).

29. C. E. Gleason, A. Ordureau, R. Gourlay, J. S. Arthur, P. Cohen, Polyubiquitin binding to optineurin is required for optimal activation of TANK-binding kinase 1 and production of interferon beta. *The Journal of biological chemistry* **286**, 35663-35674 (2011); published online EpubOct 14 (10.1074/jbc.M111.267567).
30. J. Korac, V. Schaeffer, I. Kovacevic, A. M. Clement, B. Jungblut, C. Behl, J. Terzic, I. Dikic, Ubiquitin-independent function of optineurin in autophagic clearance of protein aggregates. *Journal of cell science* **126**, 580-592 (2013); published online EpubJan 15 (10.1242/jcs.114926).
31. Y. C. Wong, E. L. Holzbaur, Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. *Proc Natl Acad Sci U S A* **111**, E4439-4448 (2014); published online EpubOct 21 (10.1073/pnas.1405752111).
32. M. Komatsu, S. Kageyama, Y. Ichimura, p62/SQSTM1/A170: physiology and pathology. *Pharmacol Res* **66**, 457-462 (2012); published online EpubDec (10.1016/j.phrs.2012.07.004).
33. K. Clark, M. Peggie, L. Plater, R. J. Sorcek, E. R. Young, J. B. Madwed, J. Hough, E. G. McIver, P. Cohen, Novel cross-talk within the IKK family controls innate immunity. *The Biochemical journal* **434**, 93-104 (2011); published online EpubFeb 15 (10.1042/BJ20101701).
34. S. Sharma, B. R. tenOever, N. Grandvaux, G. P. Zhou, R. Lin, J. Hiscott, Triggering the interferon antiviral response through an IKK-related pathway. *Science* **300**, 1148-1151 (2003); published online EpubMay 16 (10.1126/science.1081315).
35. K. A. Fitzgerald, S. M. McWhirter, K. L. Faia, D. C. Rowe, E. Latz, D. T. Golenbock, A. J. Coyle, S. M. Liao, T. Maniatis, IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. *Nature immunology* **4**, 491-496 (2003); published online EpubMay (10.1038/ni921).
36. H. Hemmi, O. Takeuchi, S. Sato, M. Yamamoto, T. Kaisho, H. Sanjo, T. Kawai, K. Hoshino, K. Takeda, S. Akira, The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection. *The Journal of experimental medicine* **199**, 1641-1650 (2004); published online EpubJun 21 (10.1084/jem.20040520).
37. M. G. Wathelet, C. H. Lin, B. S. Parekh, L. V. Ronco, P. M. Howley, T. Maniatis, Virus infection induces the assembly of coordinately activated transcription factors on the IFN-beta enhancer in vivo. *Mol Cell* **1**, 507-518 (1998); published online EpubMar (
38. Y. Yang, A. Hentati, H. X. Deng, O. Dabbagh, T. Sasaki, M. Hirano, W. Y. Hung, K. Ouahchi, J. Yan, A. C. Azim, N. Cole, G. Gascon, A. Yagmour, M. Ben-Hamida, M. Pericak-Vance, F. Hentati, T. Siddique, The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. *Nature genetics* **29**, 160-165 (2001); published online EpubOct (10.1038/ng1001-160).
39. J. Li, J. Huang, J. H. Jeong, S. J. Park, R. Wei, J. Peng, Z. Luo, Y. T. Chen, Y. Feng, J. L. Luo, Selective TBK1/IKKi dual inhibitors with anticancer potency.

International journal of cancer. Journal international du cancer **134**, 1972-1980 (2014); published online EpubApr 15 (10.1002/ijc.28507).
40. E. B. Lee, V. M. Lee, J. Q. Trojanowski, Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nature reviews. Neuroscience* **13**, 38-50 (2012); published online EpubJan (10.1038/nrn3121).
41. ExomeAggregationConsortium(ExAC). (URL: http://exac.broadinstitute.org).
42. P. Van Damme, W. Robberecht, Clinical implications of recent breakthroughs in amyotrophic lateral sclerosis. *Current opinion in neurology* **26**, 466-472 (2013); published online EpubOct (10.1097/WCO.0b013e328364c063).
43. C. Munch, A. Rolfs, T. Meyer, Heterozygous S44L missense change of the spastin gene in amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* **9**, 251-253 (2008); published online EpubAug (10.1080/17482960801900172).
44. T. Meyer, A. Schwan, J. S. Dullinger, J. Brocke, K. T. Hoffmann, C. H. Nolte, A. Hopt, U. Kopp, P. Andersen, J. T. Epplen, P. Linke, Early-onset ALS with long-term survival associated with spastin gene mutation. *Neurology* **65**, 141-143 (2005); published online EpubJul 12 (10.1212/01.wnl.0000167130.31618.0a).
45. A. Chesi, B. T. Staahl, A. Jovicic, J. Couthouis, M. Fasolino, A. R. Raphael, T. Yamazaki, L. Elias, M. Polak, C. Kelly, K. L. Williams, J. A. Fifita, N. J. Maragakis, G. A. Nicholson, O. D. King, R. Reed, G. R. Crabtree, I. P. Blair, J. D. Glass, A. D. Gitler, Exome sequencing to identify de novo mutations in sporadic ALS trios. *Nature neuroscience* **16**, 851-855 (2013); published online EpubJul (10.1038/nn.3412).
46. A. Manichaikul, J. C. Mychaleckyj, S. S. Rich, K. Daly, M. Sale, W. M. Chen, Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-2873 (2010); published online EpubNov 15 (10.1093/bioinformatics/btq559).
47. G. Jun, M. Flickinger, K. N. Hetrick, J. M. Romm, K. F. Doheny, G. R. Abecasis, M. Boehnke, H. M. Kang, Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *American journal of human genetics* **91**, 839-848 (2012); published online EpubNov 2 (10.1016/j.ajhg.2012.09.004).
48. C. Genomes Project, G. R. Abecasis, A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin, R. E. Handsaker, H. M. Kang, G. T. Marth, G. A. McVean, An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012); published online EpubNov 1 (10.1038/nature11632).
49. I. A. Adzhubei, S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, S. R. Sunyaev, A method and server for predicting damaging missense mutations. *Nature methods* **7**, 248-249 (2010); published online EpubApr (10.1038/nmeth0410-248).
50. S. Petrovski, Q. Wang, E. L. Heinzen, A. S. Allen, D. B. Goldstein, Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS genetics* **9**, e1003709 (2013)10.1371/journal.pgen.1003709).

51. ExomeVariantServer. (Seattle, WA, NHLBI GO Exome Sequencing Project (ESP), URL: http://evs.gs.washington.edu/EVS/).
52. R. CoreTeam, A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* **URL http://www.R-project.org/**, (2013).
53. C. Behrends, M. E. Sowa, S. P. Gygi, J. W. Harper, Network organization of the human autophagy system. *Nature* **466**, 68-76 (2010); published online EpubJul 1 (10.1038/nature09204).
54. M. E. Sowa, E. J. Bennett, S. P. Gygi, J. W. Harper, Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389-403 (2009); published online EpubJul 23 (10.1016/j.cell.2009.04.042).
55. E. L. Huttlin, M. P. Jedrychowski, J. E. Elias, T. Goswami, R. Rad, S. A. Beausoleil, J. Villen, W. Haas, M. E. Sowa, S. P. Gygi, A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174-1189 (2010); published online EpubDec 23 (10.1016/j.cell.2010.12.001).

Acknowledgements:

Please see the Supplementary Materials.

FALS Sequencing Consortium: Peter C. Sapp[1], Claire S. Leblond[2], Diane McKenna-Yasek[1], Kevin P. Kenna[3], Bradley N. Smith[4], Simon Topp[4], Jack Miller[4], Athina Gkazi[4], Ammar Al-Chalabi[4], Leonard H. van den Berg[5], Jan Veldink[5], Vincenzo Silani[6], Nicola Ticozzi[6], John Landers[1], Frank Baas[7], Christopher E. Shaw[4,] Jonathan D. Glass[8], Guy A. Rouleau[9], Robert Brown[1]; other consortium members can be found in the supplementary materials.

[1]Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01655, USA.

[2]Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal Qc H3A 2B4, Canada.

[3]Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Republic of Ireland.

[4]Department of Basic and Clinical Neuroscience, King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, SE5 8AF

[5]Department of Neurology, Brain Center Rudolf Magnus, University Medical Centre Utrecht, 3508 GA Utrecht, The Netherlands.

[6]Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan 20149, Italy; Department of Pathophysiology and Transplantation, 'Dino Ferrari' Center - Università degli Studi di Milano, Milan 20122 Italy.

[7]Department of Genome Analysis, Academic Medical Center. Meibergdreef 9, 1105AZ Amsterdam, The Netherlands.

[8]Department of Neurology, Emory University, Atlanta, GA 30322, USA.

[9]Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal Qc H3A 2B4, Canada.

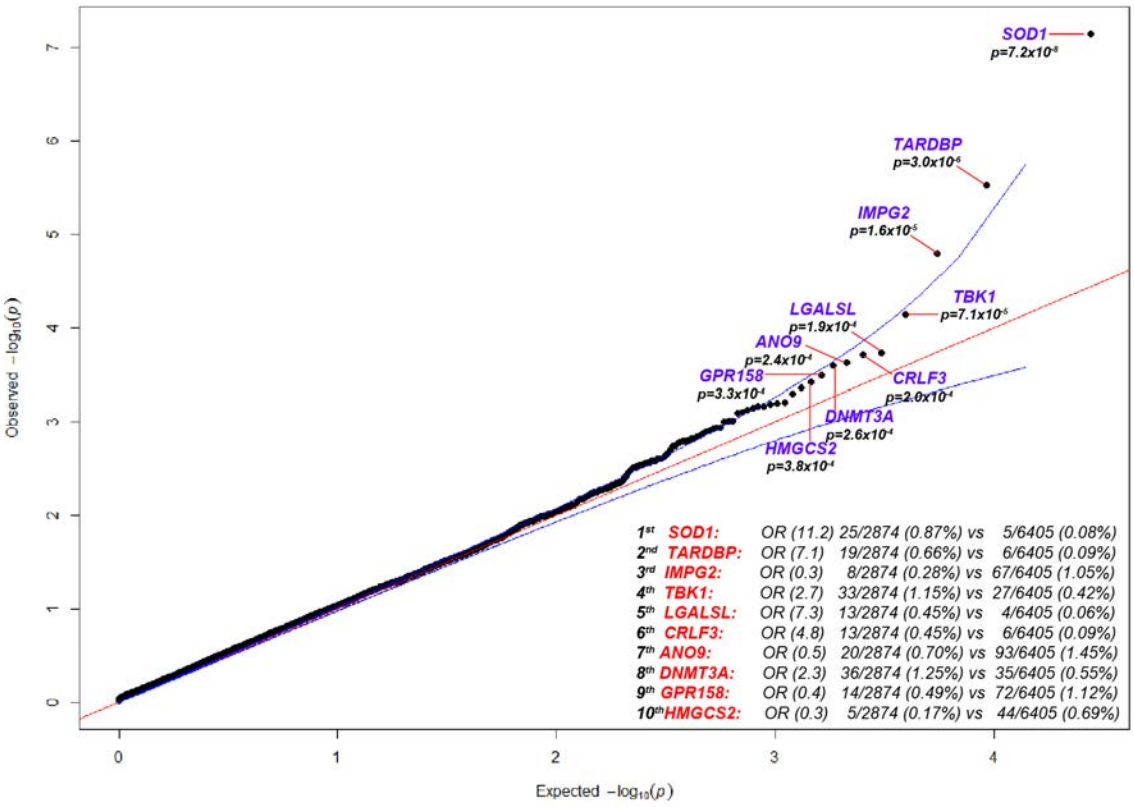

Fig. 1. QQ plot of discovery results for dominant coding model

The results for the analysis of 2,874 case and 6,405 control exomes are shown. 16,491 covered genes passed QC with more than one case or control carrier for this test. The genes with the top 10 associations are labeled. The genomic inflation factor, lambda (λ), is 1.061. The association with *SOD1* passed correction for multiple tests.

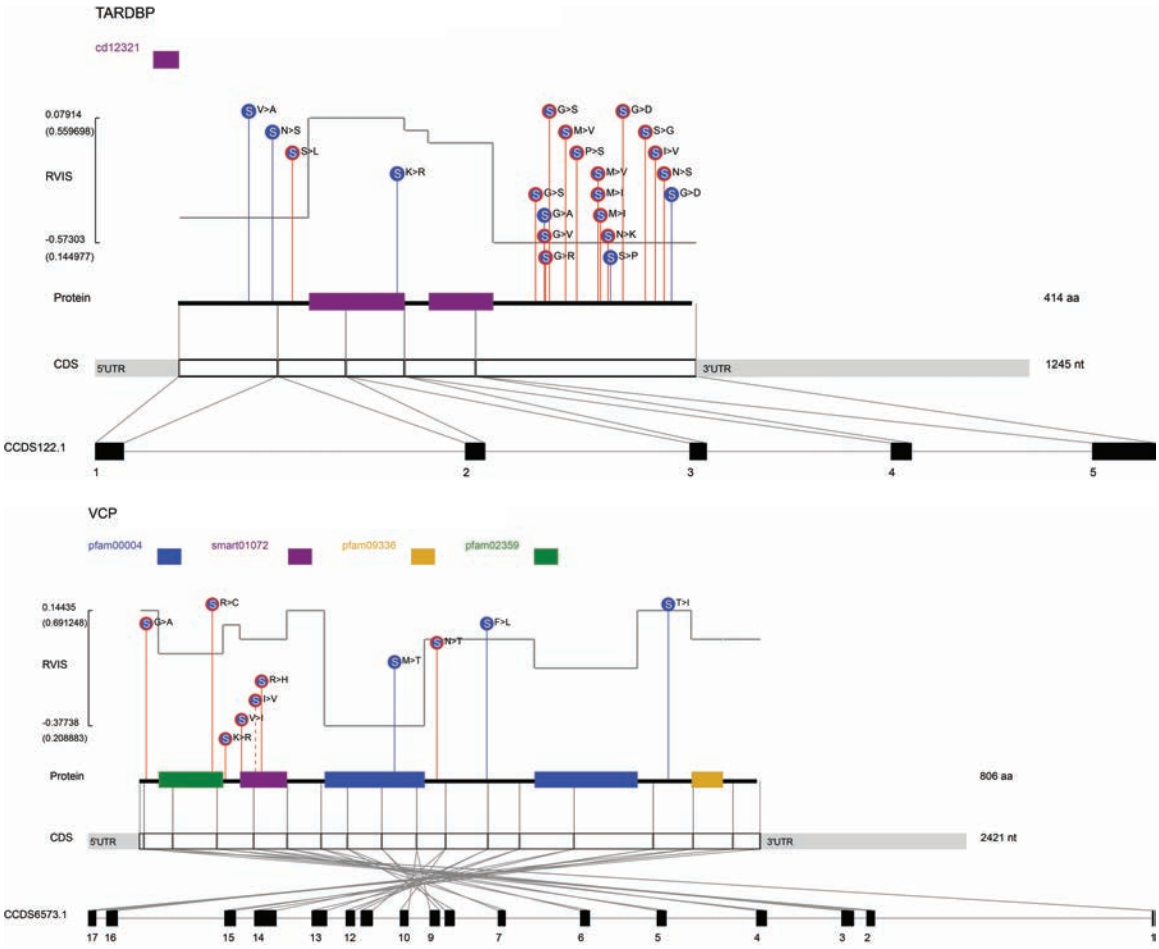

Fig. 2 Variants in *TARDBP* and *VCP*

Dominant coding variants are shown in *TARDBP* and *VCP* (discovery dataset). Case variants are enriched at the 3' end of the gene in TARDBP and near the cell division protein 48 domain 2 region in *VCP*. LoF variants are filled in red, non-synonymous variants are filled in blue, and splice variants are filled in purple. Case variants are shown with red lines, control variants are shown with blue lines, and variants found in both cases and controls are shown with dashed lines.

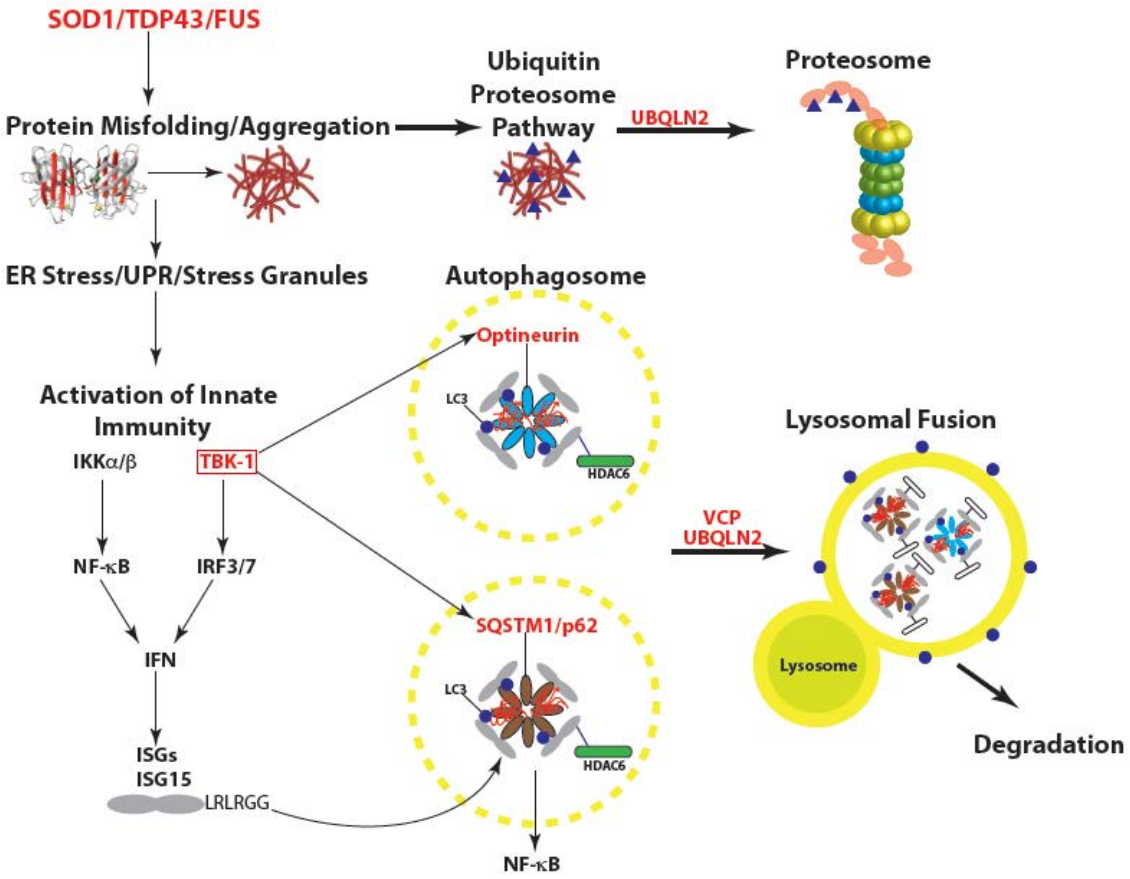

Fig. 3. Cellular processes implicated in ALS

Many of the genes implicated in ALS encode proteins that form aggregates in patients with ALS (SOD1, TARDBP and FUS) or that play key roles in innate immunity and autophagy (UBQNL2, SQSTM1, OPTN, VCP). The genes genetically implicated in ALS are indicated in red. The figure illustrates where in these key pathways the known proteins are likely to act and illustrates the role of TBK1 (boxed).

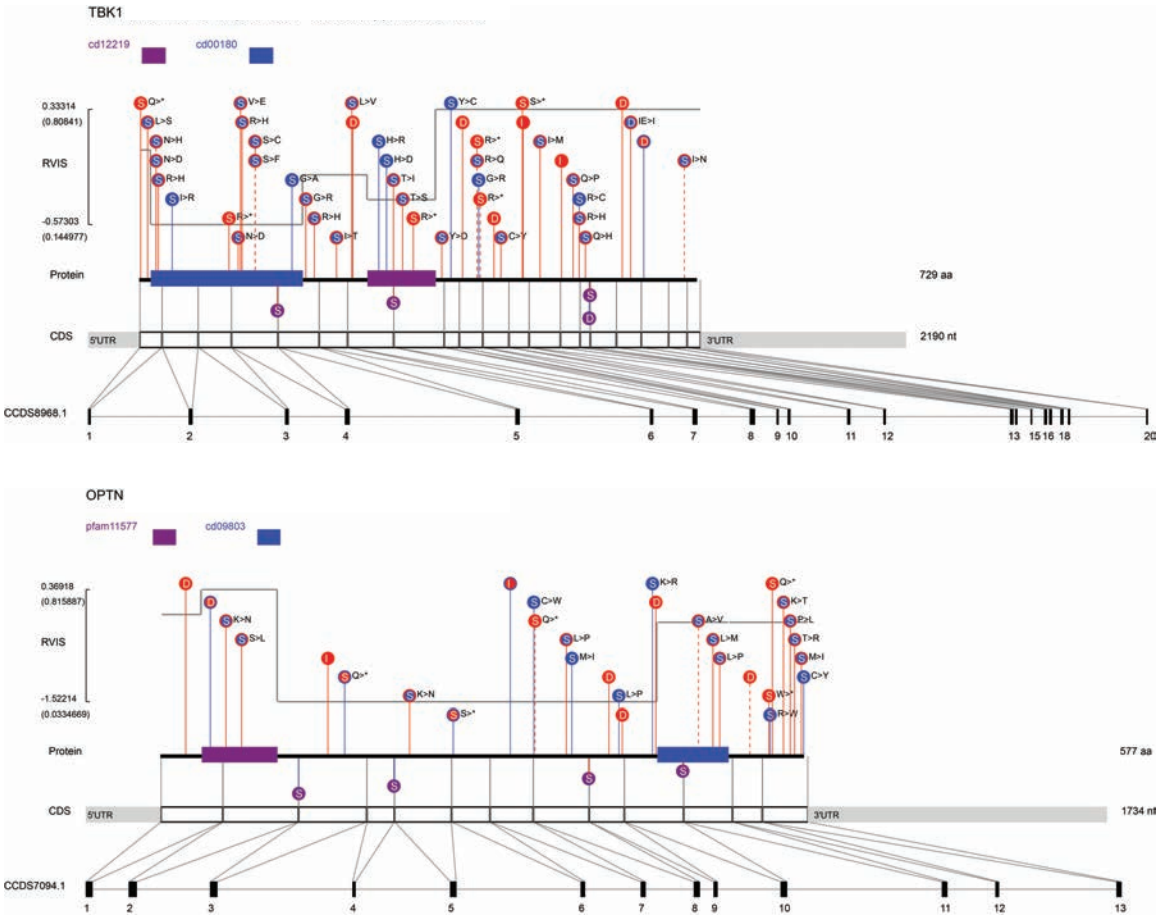

Fig. 4 Variants in *TBK1* and *OPTN*

Dominant not benign variants are shown in *TBK1* and *OPTN* (combined datasets). LoF variants are filled in red, non-synonymous variants are filled in blue, and splice variants are filled in purple. Case variants are shown with red lines, control variants are shown with blue lines, and variants found in both cases and controls are shown with dashed lines.