Supplementary material describing detailed methods and additional figures and references

Non-ALS population controls

The whole genome sequencing (WGS) data from 238 individuals were included as non-ALS population controls. These included 138 samples of African genetic ancestry from the Simons Genome Diversity Project,[1] the 1000 Genomes Project,[2] the South African Human Genome Program,[3] and other samples from South Africa.[4] The remaining 100 non-African samples which comprised 50 European and 50 Asian ancestry samples from the 1000 Genomes Project[2] were specifically included for the purpose of demonstrating genetic admixture of the SAC samples (see supplementary figure 2).

Whole genome sequencing

DNA from 103 ALS patients was extracted from peripheral blood as previously described.[4] Forty-six of these were participants in the CReATe Consortium's *Phenotype, Genotype and Biomarker* study (clinicaltrials.gov: NCT02327845) where 30X WGS was performed at Hudson Alpha (United States) using a PCR-prepared library on the Illumina NovaSeq instrument with 2 x 150bp read length. The remaining samples were sequenced to 30X coverage at the Kinghorn Centre for Clinical Genomics (Australia) using PCR-free library preparation on the Illumina HiSeq X Ten or Novaseq instruments with 2 x 150bp read length. The control datasets were sequenced on various Illumina instruments to a read depth ≥ 30X with read lengths of 100 or 150bp.

Ancestry principal component analysis

Bi-allelic single nucleotide polymorphisms (SNPs) from the joint called case and control WGS VCF file were subjected to variant-level filtering (removing markers with >5% missing genotypes, minor allele frequency <1%, Hardy-Weinberg equilibrium p value < 0.00001 and different missing data rates between cases and controls) using PLINK v1.9 (http://pngu.mgh.harvard.edu/purcell/plink/).[5] This was followed by linkage disequilibrium (LD) based SNP pruning (considering a window of 50 SNPs and removing those with LD>0.2) and a principal component analysis (PCA) to extract the top 10 principal components (PCs) of the variance-standardized relationship matrix. To construct the ancestry

PCA plot the eigenvectors for the first two PCs were visualized using Genesis PCA and Admixture Plot viewer (figure 1).[6]

Ancestry admixture analysis

Given the admixed nature of the South African Coloured (SAC) population, we inferred ancestry proportions for each SAC individual from the four major source populations (k=4) using ADMIXTURE software[7] and representative proxy samples: EUR (10 British ancestry controls from the IPDC cohort), EAS (Vietnamese controls from the IPDC cohort), SAB (the Black African patients with ALS in this study) and KHS (1 Khoisan individual from the SGDP cohort). The output from this analysis was visualized using Genesis PCA and Admixture Plot viewer[6] (supplementary figure 2). In line with previous findings[8], we calculated the average African-ancestry contribution (SAB in red plus KHS in blue) in the 76 SAC ALS patients in this study to be 49% (IQR 23-76%). Given the genetic heterogeneity within the SAC group, we also mapped each of the likely pathogenic and pathogenic mutations described in this study to each patient to verify that our findings are a true reflection of the mutation spectrum in ALS patients with African or admixed African genetic ancestry. We observed that 10 individuals in the SAC group have no significant African genetic ancestry by admixture analysis (largely EUR and EAS admixed individuals) though none of these harboured likely pathogenic and pathogenic mutations.

ALS-associated genes

The panel of 44 genes selected in this study includes 21 ALS-associated genes screened for mutations in the CReATe Consortium's *Phenotype, Genotype and Biomarker* (PGB1) genetic testing study, 16 genes approved for genetic testing in familial ALS ±FTD by the UK Genetic Testing Network steering group,[8] and 7 candidate genes which still require replication or functional data to confirm their pathogenic role in ALS.[9] Supplementary figure 1 shows these 44 ALS-associated genes grouped together according to their associated clinical phenotype as described in the Online Mendelian Inheritance in Man (OMIM) database.[10]

Standard variant annotation and filtering

Variants in the joint called case and control WGS VCF file which passed variant quality score recalibration were decomposed (multiallelic variants split) and normalized using the vt tool.[11] The Ensembl variant effect predictor (VEP)[12] was used to annotate variants with their sequence context information such as gene and transcript level annotations using cache

version 101_GRCh38. Additional functional annotations were added using the SpliceAI[13] vep plugin and the dbNSFP version 4.1a and dbscSNV1.1 databases.[14,15] The most biologically relevant consequence per gene was selected using the vep --flag_pick_allele_gene flag which was configured to prioritize MANE transcript status (matched annotations from NCBI and EBI) (http://dev-tark.ensembl.org/web/mane_project/) followed by canonical status, APPRIS isoform annotation, transcript support level, biotype of transcript, CCDS status, consequence rank and length. Clinvar annotations[16] were added from the 28/01/2021 ClinVar release.

Variant classification according to ACMG guidelines

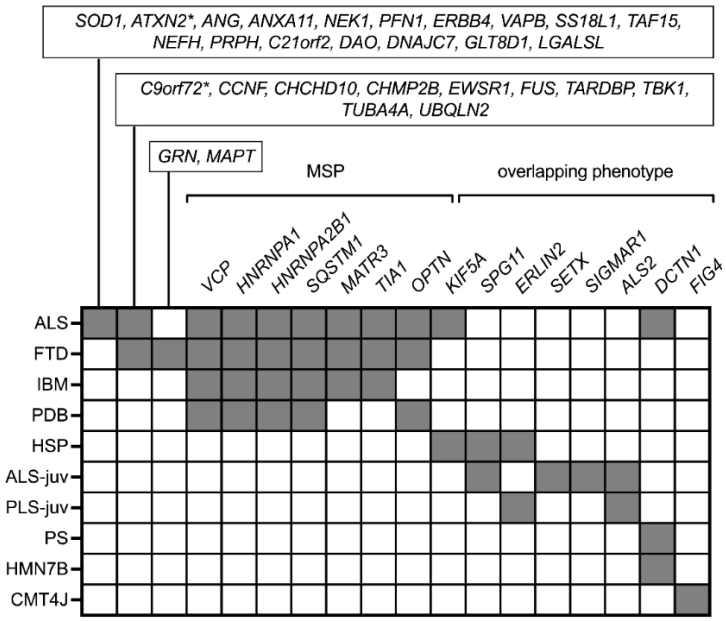
For the subset of rare variants in gnomAD (PM2) which exceeded 1 allele in the South African ALS sample (n=7), we also used MAF information from the AWI-GEN dataset[17] to interpret the rarity of these variants in African-ancestry samples (200 alleles) and assigned BS1 where the MAF exceeded 0.1% in this control dataset.

REVEL[18] was used for missense variant effect prediction as it was shown to be the top performing tool in accurately predicting the pathogenicity of missense variants and was not influenced by factors such as gene-level constraint and Mendelian inheritance pattern.[19,20] For functional effect prediction of splice region variants (including synonymous and nonsynonymous missense variants) we used the adaptive boosting (ada) and random forests (rf) ensemble scores from the dbscSNV database [14] assigning PP3 for those variants with both an ada_score and rf_score > 0.6 (predicted to alter splicing), otherwise BP4 was assigned (no predicted impact on splicing). Both BP4 and PP3 were assigned at a supporting strength level.

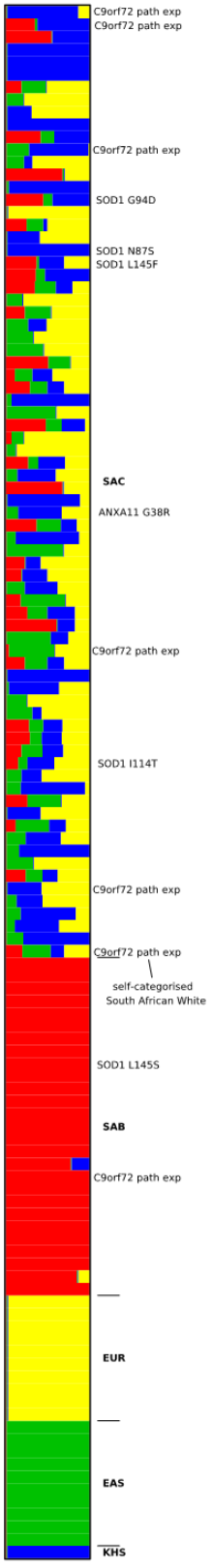
BP1 was assigned at a supporting strength level for missense variants in *NEK1* and *KIF5A* genes as truncating variants are presently the only known mechanism of variant pathogenicity for these two genes in the context of ALS.[21–25] BP3 was assigned at a supporting strength level for an in-frame insertion in the *FUS* gene which was identified in a repetitive region without a known function. PS3 was assigned with varying strength levels depending on the strength of the functional evidence demonstrating a variant's damaging effect on a gene product. PVS1 was assigned to loss of function (LOF) variants in the *NEK1* gene but downgraded to a supporting strength level as *NEK1* does not meet the criteria for a LOF disease mechanism in ALS and *NEK1* LOF variants are observed in controls.[26] PM5 was assigned at a moderate strength level for novel missense changes at an amino acid residue

where a different missense change determined to be pathogenic has been seen before. PP2 was assigned at a supporting strength level for *SOD1* missense variants as *SOD1* is a gene with low rate of benign missense variation and pathogenic missense variants in *SOD1* are a common cause of ALS. PS4 was assigned at a moderate strength level for *SOD1*, *TARDBP* and *ANXA11* missense variants where ≥ 4 case reports could be identified.

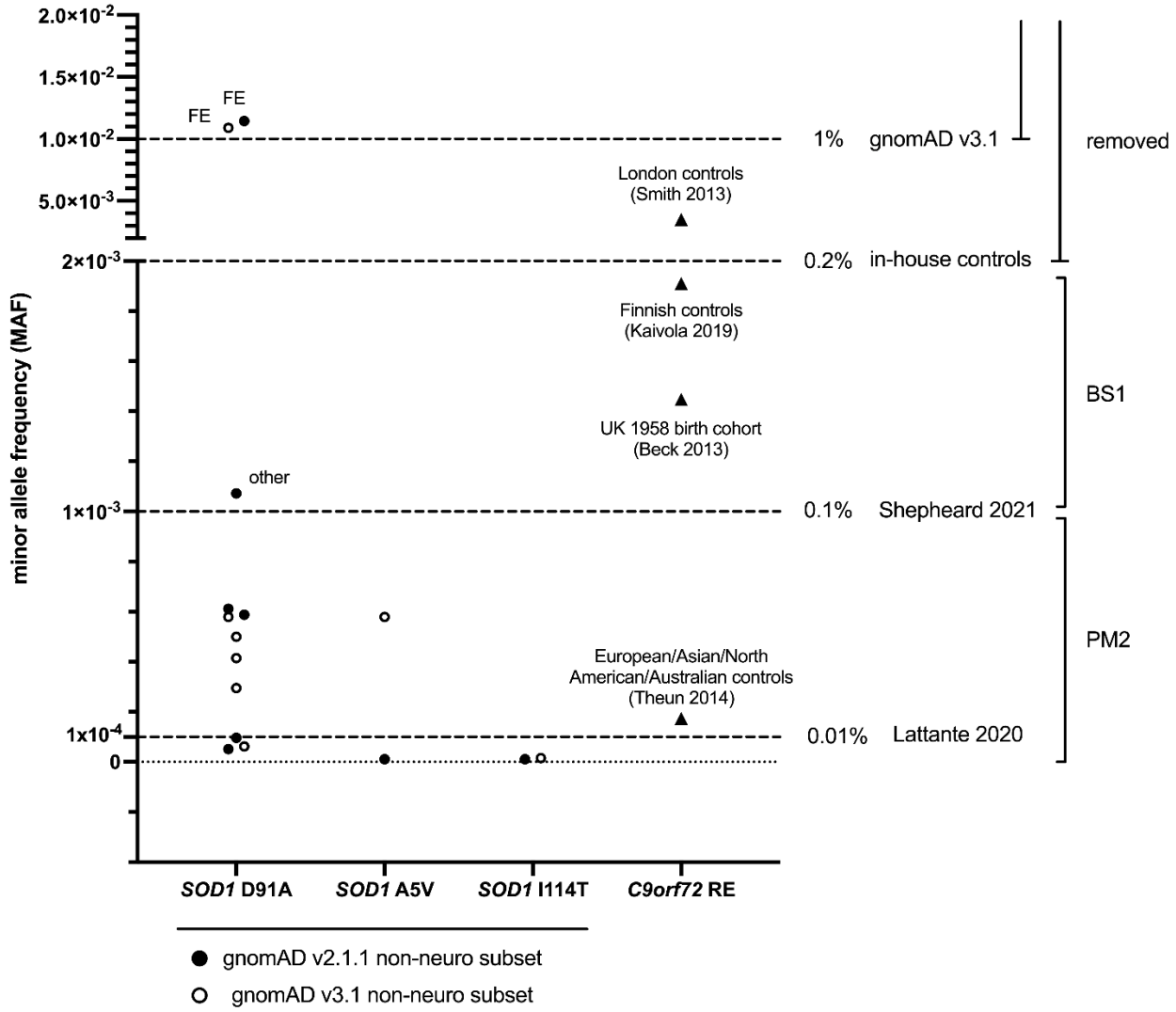
Supplementary Figures.



e figure 1. The 44 ALS genes which were curated for mutations in 103 South Africans with ALS. Short variants from WGS data were determined for all genes except **C9orf72* (screened only for the known pathogenic repeat expansion mutation) and **ATXN2* (screened only for intermediate length repeat expansions). ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia; IBM, inclusion body myopathy; MSP, multisystem proteinopathy; PDB, Paget's disease of the bone; HSP, hereditary spastic paraplegia; ALS-juv, juvenile-onset ALS; PLS-juv, juvenile-onset primary lateral sclerosis; PS, Perry syndrome; HMN7B, distal hereditary motor neuropathy type 7B; CMT4J, Charcot-Marie-Tooth Type 4J.



efigure 2. Population structure (admixture) plot showing ancestry proportions in each South African Coloured (SAC) ALS patient estimated from four source populations. European ancestry (EUR in yellow) and East Asian ancestry (EAS in green) are represented by 10 British and 10 Vietnamese ancestry controls from the 1000 Genomes Project[2] respectively. South African Black ancestry (SAB in red) is represented by the Black African patients with ALS in this study and Khoisan ancestry (KHS in blue) is represented by one Khoisan individual from the SGDP sample.[1] Individuals (represented by horizontal bars) with likely pathogenic and pathogenic mutations described in this study are indicated.



efigure 3. Approach to rare variant frequency filtering. The minor allele frequency (MAF) of established pathogenic ALS mutations in the gnomAD database (*SOD1*) and published reports (*C9orf72*) are shown. For *SOD1* variants, each data point indicates the MAF in the various population subgroups (FE=Finnish European, other=ancestry not specified, NFE=non-Finnish European, SA=South Asian, LAT=Latino, admixed American, AFR=African, African-American, ASH=Ashkenazi Jewish, EA=East Asian) in the gnomAD database (filled circles refer to gnomAD v2.1.1 exomes and genomes non-neuro subset and open circles refer to gnomAD v3.1 genomes). The *C9orf72* repeat expansion mutation carrier frequency is reported for various control cohorts from different publications. [27–31]

The STREGA reporting guidelines were used (https://www.goodreports.org/reporting-checklists/strega/)(eMethods).[32]

Additional references

1 Mallick S, Li H, Lipson M, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature 2016;**538**:201–6. doi:10.1038/nature18964

2 Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. Nature 2015;**526**:68–74. doi:10.1038/nature15393

3 Choudhury A, Ramsay M, Hazelhurst S, et al. Whole-genome sequencing for an enhanced understanding of genetic variation among South Africans. Nat Commun 2017;**8**:1–12. doi:10.1038/s41467-017-00663-9

4 Nel M, Mulder N, Europa TA, et al. Using Whole Genome Sequencing in an African Subphenotype of Myasthenia Gravis to Generate a Pathogenetic Hypothesis. Front Genet 2019;**10**. doi:10.3389/fgene.2019.00136

5 Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 2007;**81**:559–75. doi:10.1086/519795

6 Buchmann R, (Sydney Brenner Institute for Molecular Bioscience U of the W, Hazelhurst, Scott (Sydney Brenner Institute for Molecular Bioscience U of the W. Genesis Manual. 2014.

7 Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 2011;**12**:246. doi:10.1186/1471-2105-12-246

8 De Wit E, Delport W, Rugamika CE, *et al.* Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. *Hum Genet* 2010;**128**:145–53. doi:10.1007/s00439-010-0836-1

9 Shepheard SR, Parker MD, Cooper-Knock J, et al. Value of systematic genetic screening of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2021;:jnnp-2020-325014. doi:10.1136/jnnp-2020-325014

10 Gregory JM, Fagegaltier D, Phatnani H, et al. Genetics of Amyotrophic Lateral Sclerosis. Curr Genet Med Rep 2020;**8**:121–31. doi:10.1007/s40142-020-00194-8

11 Hamosh A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res 2004;**33**:D514–7. doi:10.1093/nar/gki033

12 Tan A, Abecasis GR, Kang HM. Unified representation of genetic variants. Bioinformatics 2015;**31**:2202–4. doi:10.1093/bioinformatics/btv112

13 McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;**17**:122. doi:10.1186/s13059-016-0974-4

14 Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell 2019;**176**:535-548.e24. doi:10.1016/j.cell.2018.12.015

15 Liu X, Li C, Mou C, et al. dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. Genome Med 2020;**12**:103. doi:10.1186/s13073-020-00803-9

16 Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. Nucleic Acids Res 2014;**42**:13534–44. doi:10.1093/nar/gku1206

17 Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;**44**:D862-8. doi:10.1093/nar/gkv1222

18 Ramsay M, Crowther N, Tambo E, et al. H3Africa AWI-Gen Collaborative Centre: A resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. Glob Heal Epidemiol Genomics 2016;**1**. doi:10.1017/gheg.2016.17

19 Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;**99**:877–85. doi:10.1016/j.ajhg.2016.08.016

20 Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Biol 2017;**18**:225. doi:10.1186/s13059-017-1353-5

21 Tian Y, Pesaran T, Chamberlin A, et al. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. Sci Rep 2019;**9**:12752. doi:10.1038/s41598-019-49224-8

22 Kenna KP, Van Doormaal PTC, Dekker AM, et al. NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. Nat Genet 2016;**48**:1037–42. doi:10.1038/ng.3626

23 Gratten J, Zhao Q, Benyamin B, et al. Whole-exome sequencing in amyotrophic lateral sclerosis suggests NEK1 is a risk gene in Chinese. Genome Med 2017;**9**:1–9. doi:10.1186/s13073-017-0487-0

24 Nguyen HP, Van Mossevelde S, Dillen L, et al. NEK1 genetic variability in a Belgian cohort of ALS and ALS-FTD patients. Neurobiol Aging 2018;**61**:255.e1-255.e7. doi:10.1016/j.neurobiolaging.2017.08.021

25 Naruse H, Ishiura H, Mitsui J, et al. Loss-of-function variants in NEK1 are associated with an increased risk of sporadic ALS in the Japanese population. J Hum Genet Published Online First: 2020. doi:10.1038/s10038-020-00830-9

26 Nicolas A, Kenna KP, Renton AE, et al. Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. Neuron 2018;**97**:1268-1283.e6. doi:10.1016/j.neuron.2018.02.027

27 Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat 2018;**39**:1517–24. doi:10.1002/humu.23626

28 Smith BN, Newhouse S, Shatunov A, et al. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. Eur J Hum Genet 2013;**21**:102–8. doi:10.1038/ejhg.2012.98

29 Beck J, Poulter M, Hensman D, et al. Large C9orf72 Hexanucleotide Repeat Expansions Are Seen in Multiple Neurodegenerative Syndromes and Are More Frequent Than Expected in the UK Population. Am J Hum Genet 2013;**92**:345–53. doi:10.1016/j.ajhg.2013.01.011

30 Theuns J, Verstraeten A, Sleegers K, et al. Global investigation and meta-analysis of the C9orf72 (G4C2)n repeat in Parkinson disease. Neurology 2014;**83**:1906–13. doi:10.1212/WNL.0000000000001012

31 Kaivola K, Kiviharju A, Jansson L, et al. C9orf72 hexanucleotide repeat length in older population: normal variation and effects on cognition. Neurobiol Aging 2019;**84**:242.e7-242.e12. doi:10.1016/j.neurobiolaging.2019.02.026

32 Little J, Higgens JP, Ioannidis JP, et al. STrengthening the Reporting of Genetic Association Studies (STREGA): An extension of the STROBE statement. PLOS Medicine https://doi.org/10.1371/journal.pmed.1000022.

Supplementary table 1. Sixty-two variants identified in 42 ALS genes in South Africans with ALS (n=103) classified according to ACMG evidence codes (see eMethods): likely benign (LB) variants (A, n=18), variants of uncertain significance (VUS) (B, n=38), likely pathogenic (LP) variants (C, n=1) and pathogenic (P) variants (D, n=5). HGVSp refers to the Human Genome Variation Society protein nomenclature to describe sequence variants. gnomAD v2.1.1 and v3.1 refers to frequency data (non-neuro subsets) from the gnomAD database (https://gnomad.broadinstitute.org); variants were reported as absent from gnomAD non-neuro subsets if sufficient coverage (>20x) of the variant site could be confirmed by examining the IGV read data plots in the gnomAD browser. REVEL refers to the metapredictor score described by Ioannidis et al., 2016.[1] All variants were heterozygous singletons unless indicated where hom^n refers to the number of homozygotes, het^n refers to the number of heterozygotes and $alleles^n$ refers to the total number of alleles in the ALS sample. AWI-GEN MAF refers to the frequency of the variant in a sample of 100 Black South African controls (200 alleles);[2] this information is included for variants which were rare in gnomAD but found in more than 1 ALS sample (n=7, indicated by •). In the evidence codes column, ♦ indicates variants prioritized for future research based on the available evidence (see comments column) which are referenced in table 2A. For ClinVar interpretations, B=benign, LB=likely benign, VUS=variant of uncertain significance, LP=likely pathogenic, P=pathogenic, * indicates the review status of a record on 08/09/21 where 1 star=criteria provided, single submitter/conflicting interpretations, two stars=criteria provided, multiple submitters, no conflicts and no stars=no assertion criteria provided.

A. Variants classified as likely benign according to ACMG criteria								
gene	**genomic change (hg38)**	**HGVSp (exon number)**	**variant class**	**gnomAD v2.1.1 ; v3.1 [hom^n, het^n, $alleles^n$] (AWI-GEN MAF)**	**REVEL score**	**evidence codes (• reclassified LB)**	**ClinVar interpretation (review status)**	**comments**
ALS2	chr2:201715771:C>T	R1302H (25)	missense	2.4E-3 ; 2.9E-3	0.11	BS1_strong, BP4_supp	LB (**)	mutations in the *ALS2* gene cause juvenile-onset AR ALS (ALS2) and HSP
NEK1	chr4:169438121:G>C	S909C (28)	missense	2.4E-3 ; 2.2E-3	0.13	BS1_strong, BP4_supp, BP1_supp	LB (*)	*NEK1* LOF variants more frequent in ALS cases vs controls [3–6], pathogenic role for missense *NEK1* variants in ALS not established
NEK1	chr4:169477476:G>C	R721G (25)	missense	2.7E-05 ; 1.5E-04	0.07	PM2_supp, BP4_supp, BP1_supp	VUS (*)	
NEK1	chr4:169479449:A>G	M689T (24)	missense	2.8E-3 ; 2.4E-3	0.05	BS1_strong, BP4_supp, BP1_supp	LB (*)	
NEK1	chr4:169585470:T>C	Y229C (10)	missense	1.7E-4 ; 2.2E-3	0.26	BS1_strong, BP4_supp, BP1_supp	LB/VUS (*)	
MATR3	chr5:139325651:A>G	N787S (15)	missense	2.7E-3 ; 2.2E-3 [0, 2, 2]	0.21	BS1_strong, BP4_supp	B/LB (**)	-
SQSTM1	chr5:179833236:G>A	G320E (6)	missense	9.3E-5 ; 4.9E-6 [0, 2, 2] (5.0E-03)	0.38	BS1_strong, BP4_supp•	VUS (**)	BS1 assigned as MAF>0.1% in AWI-GEN dataset
SETX	chr9:132327078:T>G	D1507A (10)	missense	5.2E-4 ; 4.6E-4 [0, 2, 2] (5.0E-03)	0.45	BS1_strong, BP4_supp•	VUS (*)	BS1 assigned as MAF>0.1% in AWI-GEN dataset

SETX	chr9:132328165:G>A	R1145W (10)	missense	1.1E-5 ; 6.7E-5 [1, 2, 4] (5.0E-03)	0.28	BS1_strong, BP4_supp•	VUS (*)	
KIF5A	chr12:57569669:G>T	S368I (11)	missense	1.1E-5 ; 0	0.23	PM2_supp, BP4_supp, BP1_supp	-	ALS-associated *KIF5A* variants are predominantly heterozygous LOF variants in the C-terminal region (residues 998-1007) [7]
SPG11	chr15:44566301:G>C	D2253E (37)	missense	4.3E-3 ; 5.0E-3	0.13	BS1_strong, BP4_supp	LB/VUS (*)	homozygous or compound heterozygous LOF *SPG11* variants segregate in juvenile ALS (ALS5) families [8], BS1 assigned for D1421N as MAF>0.1% in AWI-GEN dataset
SPG11	chr15:44596256:C>T	D1421N (25)	missense	7.5E-5 ; 8.0E-6 [0, 2, 2] (5.0E-03)	0.11	BS1_strong, BP4_supp•	VUS (*)	
SPG11	chr15:44657201:T>C	K255E (4)	missense	1.0E-3 ; 1.3E-3	0.03	BS1_strong, BP4_supp	VUS (*)	
CCNF	chr16:2448977:G>A	R406Q (11)	missense	7.7E-3 ; 5.0E-3 [0, 2, 2]	0.24	BS1_strong, BP4_supp	B (*)	-
FUS	chr16:31185096: C>CGGCGGT	G230_G231dup (6)	inframe insertion	7.9E-5; 1.3E-4 [0, 3, 3] (0)	NA	BS1_strong, BP3_supp•	-	gnomAD frequencies reported for insGGTGGT or insGGCGGC which results in equivalent amino acid change, i.e. G230_G231dup, BS1 assigned as inframe insertions and deletions are common in this stretch of 10 glycine residues (aa pos 222-231) in exon 6 of FUS protein (e.g. G231dup MAF 2.0E-3 in gnomAD)
GRN	chr17:44350237:C>A	S120Y (5)	missense	2.7E-3 ; 2.4E-3	0.15	BS1_strong, BP4_supp	B/LB (**)	-
GRN	chr17:44350271:C>G	F131L (5)	missense	1.8E-3 ; 1.4E-3	0.36	BS1_strong, BP4_supp	LB/VUS (*)	
NEFH	chr22:29489379:C>T	S580F (4)	missense	9.2E-4 ; 1.1E-3 [0, 2, 2]	0.22	BS1_strong, BP4_supp	-	
B. Variants of uncertain significance as classified according to ACMG criteria								
gene	**genomic change (hg38)**	**HGVSp (exon number)**	**variant class**	**gnomAD v2.1.1 ; v3.1 [hom^n, het^n, alleles^n]**	**REVEL score**	**evidence codes (♦ = research)**	**ClinVar interpretation (review status)**	**comments**
TARDBP	chr1:11016874:C>T	A90V (3)	missense	3.5E-4 ; 4.2E-4	0.17	PM2_supp, BP4_supp, PS3_supp, PS4_mod ♦	VUS (**)	commonly cited ALS-associated (risk factor?) variant, functional studies show altered protein function though not to the same degree as pathogenic mutant [9]
ALS2	chr2:201723407:C>T	V1183M (22)	missense	5.7E-4 ; 5.0E-4 [0, 2, 2] (0)	0.08	PM2_supp, BP4_supp•	VUS (*)	
TUBA4A	chr2:219252100:C>G	G45A (2)	missense	2.0E-4 ; 0	0.16	PM2_supp, BP4_supp	-	-

LGALSL	chr2:64456407:G>A	R106K (4)	missense	3.7E-6 ; 1.6E-5	0.10	PM2_supp, BP4_supp	-	*LGALSL* variants more frequent in ALS cases vs controls though pathogenicity of individual variants not confirmed [10,11]
DCTN1	chr2:74362700:C>G	E1187Q (30)	missense	absent	0.29	PM2_supp, BP4_supp	-	*DCTN1* E34Q located in the *DCTN1* CAP-Gly domain, found in 1 ALS case and 1 control, PS3 assigned at supporting strength level as 1 study showed altered protein function though not to the same degree as pathogenic mutant [12]
DCTN1	chr2:74363304:C>G	S1112T (28)	missense	absent	0.14	PM2_supp, BP4_supp	-	
DCTN1	chr2:74378179:C>G	E34Q (2)	missense	6.9E-5 ; 5.8E-5	0.77	PM2_supp, PP3_supp, PS3_supp♦	VUS (**)	
NEK1	chr4:169424747: GCACAGACTT ATCTACATCAC ATTTAGAGTG CTGAGAAT>G	D997AfsTer8 (31)	frameshift	absent	NA	PM2_supp, PVS1_supp♦	-	*NEK1* LOF variants more frequent in ALS cases vs controls [3–6], PVS1 evidence code downgraded to supporting strength level as *NEK1* does not meet the criteria for a LOF disease mechanism in ALS and *NEK1* LOF variants are observed in controls [13]
NEK1	chr4:169508777: CTT>C	K580RfdTer19 (20)	frameshift	absent	NA	PM2_supp, PVS1_supp♦	-	
NEK1	chr4:169587582:C>A	E195Ter (9)	stop gained	absent	NA	PM2_supp, PVS1_supp♦	-	
MATR3	chr5:139315732:A>G	H337R (6)	missense	9.0E-4 ; 6.8E-4	0.28	PM2_supp, BP4_supp	LB (*)	-
SQSTM1	chr5:179820941:C>T	A2V (1)	missense	4.3E-5 ; 2.6E-4	0.04	PM2_supp, BP4_supp	VUS (*)	PM5 assigned as *SQSTM1* P387L found in FTD cases [14] including segregation with disease in FTD family [15], found in familial PDB with incomplete penetrance [16] and functional studies indicate altered protein function [17,18]
SQSTM1	chr5:179833776:C>G	P387A (7)	missense	1.1E-5 ; 1.6E-5	0.66	PM2_supp, PP3_supp, PM5_mod♦	-	
FIG4	chr6:109716449:T>A	V57D (3)	missense	absent	0.38	PM2_supp, BP4_supp	VUS (*)	LOF variants found in ALS cases [19] but incomplete penetrance [20], pathogenic role of specific missense variants uncertain
SETX	chr9:132264856:G>C	L2473V (26)	missense	1.3E-4 ; 7.3E-5	0.28	PM2_supp, BP4_supp	VUS (*)	*SETX* harbours pathogenic heterozygous missense variants segregating in juvenile ALS (ALS4) families distributed throughout the protein [21,22]
SETX	chr9:132300688:T>C	I1830M (12)	missense	0 ; 6.8E-5 [0, 2, 2] (0)	0.17	PM2_supp, BP4_supp•	-	
SETX	chr9:132327208:G>A	P1464S (10)	missense	2.9E-5 ; 8.0E-6	0.20	PM2_supp, BP4_supp	VUS (*)	
SETX	chr9:132329314:T>A	N762Y (10)	missense	1.1E-5 ; 0	0.23	PM2_supp, BP4_supp	-	
ANXA11	chr10:80164100:G>C	S301C (10)	missense	0 ; 8.0E-6	0.63	PM2_supp, PP3_supp	-	-
PRPH	chr12:49297188:C>T	A304V (5)	missense	3.3E-5 ; 0	0.68	PM2_supporting, PP3_supp	-	limited data on *PRPH* mutations in ALS
HNRNPA1	chr12:54282856:A>G	N245D (7)	missense	absent	0.36	PM2_supporting, BP4_supp	-	-

TBK1	chr12:64466943:G>A	R134H (5)	missense	3.0E-6 ; 1.6E-5	0.73	PM2_supp, PP3_supp, PS3_supp♦	-	*TBK1* R134H identified in ALS case [10] and functional studies indicate altered protein function [23]
DAO	chr12:108887563:C>T	P103L (3)	missense (splice region)	2.8E-3 ; 1.6E-3	0.44 ada_score=0.994 rf_score=0.752	BS1_strong, PP3_supp	LB (*)	P103L found in 2 FALS patients but also in controls [24]
SPG11	chr15:44570607:A>G	M2132T (34)	missense	6.0E-5 ; 5.0E-4	0.76	PM2_supp, PP3_supp	VUS (*)	M2132T also identified in an African ancestry case with hereditary spastic paraplegia from another cohort
SPG11	chr15:44600566:C>T	R1196H (21)	missense	8.9E-6 ; 1.2E-5	0.43	PM2_supp, BP4_supp	VUS (**)	homozygous or compound heterozygous LOF *SPG11* variants segregate in juvenile ALS (ALS5) families [8]
SPG11	chr15:44621836:C>T	R848K (14)	missense	2.6E-5 ; 2.1E-4	0.08	PM2_supp, BP4_supp	-	
SPG11	chr15:44628823:T>C	K638R (10)	missense	2.3E-4 ; 2.0E-4	0.09	PM2_supp, BP4_supp	VUS (*)	
CCNF	chr16:2445550:A>T	Y341F (10)	missense	9.2E-4 ; 7.9E-4	0.18	PM2_supp, BP4_supp	-	-
CCNF	chr16:2453488:G>A	G556R (15)	missense	absent	0.06	PM2_supp, BP4_supp	-	
CCNF	chr16:2456926:C>T	P756L (17)	missense	3.0E-4 ; 5.4E-4	0.01	PM2_supp, BP4_supp	-	
FUS	chr16:31185061:C>T	R216C (6)	missense	2.9E-4 ; 1.4E-3	0.68	BS1_strong, PP3_supp	P	pathogenic ALS/FTD FUS mutations cluster in exons 3, 5-6 and 13-15 [25]; R274C has MAF 2.7E-04 in ProjectMine controls
FUS	chr16:31188345:C>T	R274C (8)	missense	3.7E-6 ; 0	0.11	PM2_supp, BP4_supp	-	
SOD1	chr21:31667309:T>A	D97E (4)	missense	6.2E-5 ; 3.1E-5	0.31	PM2_supp, BP4_supp, PP2_supp♦	-	1 report describes recessive ALS family where affected individuals carry heterozygous *SOD1* D97N and D91A mutations [26], D97V detected in 1 healthy control (73yrs, no family history of ALS) (ClinVar accession SCV000994920.1)
EWSR1	chr22:29282514:C>T	P180S (6)	missense	absent	0.57	PM2_supp, PP3_supp	-	limited data on *EWSR1* mutations in ALS
NEFH	chr22:29480518:G>A	G86S (1)	missense	2.0E-4 ; 2.0E-4	0.22	PM2_supp, BP4_supp	-	*NEFH* N390T found in 1 FTD case [27]
NEFH	chr22:29485808:A>C	N390T (3)	missense	4.0E-4 ; 1.7E-4	0.84	PM2_supp, PP3_supp♦	VUS (*)	
NEFH	chr22:29489215:G>C	K525N (4)	missense	4.2E-4 ; 1.9E-4	0.22	PM2_supp, BP4_supp	VUS (*)	
UBQLN2	chrX:56565446:C>T	P525S (1)	missense	8.2E-3 ; 6.2E-6	0.53	BS1_strong, PP3_supp♦	B/LB/VUS (*)	*UBQLN2* P525S found in 2 FTD families with incomplete penetrance [28,29], BS3/PS3 not assigned as functional studies report conflicting results [30–32]
C. Variants classified as likely pathogenic according to ACMG criteria								
ANXA11	chr10:80170859:C>T	G38R (5)	missense	3.9E-5 ; 3.8E-5	0.26	PM2_supp, BP4_supp, PS3_strong, PS4_mod ♦	P	*ANXA11* G38R found in ALS cases [33–35], PS3 assigned at moderate strength

								level for this variant as multiple functional studies indicate altered protein function [33,35–37]
D. Variants classified as pathogenic according to ACMG criteria								
gene	**genomic change (hg38)**	**HGVSp (exon number)**	**variant class**	**gnomAD v2.1.1 ; v3.1 [hom^n, het^n, alleles^n]**	**REVEL score**	**evidence codes (♦ = research)**	**ClinVar interpretation (review status)**	**comments**
SOD1	chr21:31667278:A>G	N87S (4)	missense	absent	0.85	PM2_supp, PP3_supp, PP2_supp, PS4_mod, PS3_strong	P/LP (**)	reported in heterozygous state in many ALS cases [38–43] and in homozygous state in individual with juvenile-onset ALS [44]
SOD1	chr21:31667299:G>A	G94D (4)	missense	absent	0.91	PM2_supp, PP3_supp, PP2_supp, PS4_mod, PS3_strong	P (*)	observed in several ALS cases including segregation in families
SOD1	chr21:31667359:T>C	I114T (4)	missense	1.1E-5 ; 1.6E-5	0.99	PM2_supp, PP3_supp, PP2_supp, PS4_mod, PS3_strong	P (**)	I114T most frequent *SOD1* mutation in the UK and the third most prevalent *SOD1* mutation worldwide [45]
SOD1	chr21:31668547:T>C	L145S (5)	missense	absent	0.96	PM2_supp, PP3_supp, PP2_supp, PS4_mod, PS3_strong	P (**)	observed in several ALS cases including segregation in families
SOD1	chr21:31668548:G>C	L145F (5)	missense	absent	0.92	PM2_supp, PP3_supp, PP2_supp, PS4_mod, PS3_strong	P (**)	observed in several ALS cases including segregation in families

1 Ioannidis NM, Rothstein JH, Pejaver V, *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* 2016;**99**:877–85. doi:10.1016/j.ajhg.2016.08.016

2 Ramsay M, Crowther N, Tambo E, *et al.* H3Africa AWI-Gen Collaborative Centre: A resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. *Glob Heal Epidemiol Genomics* 2016;**1**. doi:10.1017/gheg.2016.17

3 Kenna KP, Van Doormaal PTC, Dekker AM, *et al.* NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. *Nat Genet* 2016;**48**:1037–42. doi:10.1038/ng.3626

4 Gratten J, Zhao Q, Benyamin B, *et al.* Whole-exome sequencing in amyotrophic lateral sclerosis suggests NEK1 is a risk gene in Chinese. *Genome Med* 2017;**9**:1–9. doi:10.1186/s13073-017-0487-0

5 Nguyen HP, Van Mossevelde S, Dillen L, *et al.* NEK1 genetic variability in a Belgian cohort of ALS and ALS-FTD patients. *Neurobiol Aging* 2018;**61**:255.e1-255.e7. doi:10.1016/j.neurobiolaging.2017.08.021

6 Naruse H, Ishiura H, Mitsui J, *et al.* Loss-of-function variants in NEK1 are associated with an increased risk of sporadic ALS in the Japanese population. *J Hum Genet* Published Online First: 2020. doi:10.1038/s10038-020-00830-9

7 Nicolas A, Kenna KP, Renton AE, *et al.* Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* 2018;**97**:1268-1283.e6. doi:10.1016/j.neuron.2018.02.027

8 Orlacchio A, Babalini C, Borreca A, *et al.* SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain* 2010;**133**:591–8.

doi:10.1093/brain/awp325
9 Wobst HJ, Wesolowski SS, Chadchankar J, *et al.* Cytoplasmic Relocalization of TAR DNA-Binding Protein 43 Is Not Sufficient to Reproduce Cellular Pathologies Associated with ALS In vitro. *Front Mol Neurosci* 2017;**10**. doi:10.3389/fnmol.2017.00046
10 Cirulli ET, Lasseigne BN, Petrovski S, *et al.* Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* 2015;**347**:1436–41. doi:10.1126/science.aaa3650
11 Gelfman S, Dugger S, de Araujo Martins Moreno C, *et al.* A new approach for rare variation collapsing on functional protein domains implicates specific genic regions in ALS. *Genome Res* 2019;**29**:809–18. doi:10.1101/gr.243592.118
12 Stockmann M, Meyer-Ohlendorf M, Achberger K, *et al.* The dynactin p150 subunit: Cell biology studies of sequence changes found in ALS/MND and Parkinsonian Syndromes. *J Neural Transm* 2013;**120**:785–98. doi:10.1007/s00702-012-0910-z
13 Abou Tayoun AN, Pesaran T, DiStefano MT, *et al.* Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat* 2018;**39**:1517–24. doi:10.1002/humu.23626
14 van der Zee J, Van Langenhove T, Kovacs GG, *et al.* Rare mutations in SQSTM1 modify susceptibility to frontotemporal lobar degeneration. *Acta Neuropathol* 2014;**128**:397–410. doi:10.1007/s00401-014-1298-7
15 Le Ber I. SQSTM1 Mutations in French Patients With Frontotemporal Dementia or Frontotemporal Dementia With Amyotrophic Lateral Sclerosis. *JAMA Neurol* Published Online First: 16 September 2013. doi:10.1001/jamaneurol.2013.3849
16 Johnson-Pais TL, Wisdom JH, Weldon KS, *et al.* Three Novel Mutations in SQSTM1 Identified in Familial Paget's Disease of Bone. *J Bone Miner Res* 2003;**18**:1748–53. doi:10.1359/jbmr.2003.18.10.1748
17 Cavey JR, Ralston SH, Sheppard PW, *et al.* Loss of Ubiquitin Binding Is a Unifying Mechanism by Which Mutations of SQSTM1 Cause Paget's Disease of Bone. *Calcif Tissue Int* 2006;**78**:271–7. doi:10.1007/s00223-005-1299-6
18 Leach RJ, Singer FR, Ench Y, *et al.* Clinical and Cellular Phenotypes Associated With Sequestosome 1 (SQSTM1) Mutations. *J Bone Miner Res* 2006;**21**:P45–50. doi:10.1359/jbmr.06s208
19 Chow CY, Landers JE, Bergren SK, *et al.* Deleterious Variants of FIG4, a Phosphoinositide Phosphatase, in Patients with ALS. *Am J Hum Genet* 2009;**84**:85–8. doi:10.1016/j.ajhg.2008.12.010
20 Osmanovic A, Rangnau I, Kosfeld A, *et al.* FIG4 variants in central European patients with amyotrophic lateral sclerosis: a whole-exome and targeted sequencing study. *Eur J Hum Genet* 2017;**25**:324–31. doi:10.1038/ejhg.2016.186
21 Arning L, Epplen JT, Rahikkala E, *et al.* The SETX missense variation spectrum as evaluated in patients with ALS4-like motor neuron diseases. *Neurogenetics* 2013;**14**:53–61. doi:10.1007/s10048-012-0347-4
22 Tripolszki K, Török D, Goudenège D, *et al.* High-throughput sequencing revealed a novel SETX mutation in a Hungarian patient with amyotrophic lateral sclerosis. *Brain Behav* 2017;**7**:1–7. doi:10.1002/brb3.669
23 Ye J, Cheung J, Gerbino V, *et al.* Effects of ALS-associated TANK binding kinase 1 mutations on protein-protein interactions and kinase activity. *Proc Natl Acad Sci U S A* 2019;**116**:24517–26. doi:10.1073/pnas.1915732116
24 Pang SYY, Hsu JS, Teo KC, *et al.* Burden of rare variants in ALS genes influences survival in familial and sporadic ALS. *Neurobiol Aging* 2017;**58**:238.e9-238.e15. doi:10.1016/j.neurobiolaging.2017.06.007
25 Mackenzie IR, Rademakers R, Neumann M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol* 2010;**9**:995–1007. doi:10.1016/S1474-4422(10)70195-2
26 Hand CK, Mayeux-Portas V, Khoris J, *et al.* Compound heterozygous D90A and D96N SOD1 mutations in a recessive amyotrophic lateral sclerosis family. *Ann Neurol* 2001;**49**:267–71. doi:10.1002/1531-8249(20010201)49:2<267::aid-ana51>3.0.co;2-d
27 Blauwendraat C, Wilke C, Simón-Sánchez J, *et al.* The wide genetic landscape of clinical frontotemporal dementia: systematic combined sequencing of 121 consecutive subjects. *Genet Med* 2018;**20**:240–9. doi:10.1038/gim.2017.102

28 Deng H-X, Chen W, Hong S-T, *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 2011;**477**:211–5. doi:10.1038/nature10353

29 Özoğuz A, Uyan Ö, Birdal G, *et al.* The distinct genetic pattern of ALS in Turkey and novel mutations. *Neurobiol Aging* 2015;**36**:1764.e9-1764.e18. doi:10.1016/j.neurobiolaging.2014.12.032

30 Chang L, Monteiro MJ. Defective Proteasome Delivery of Polyubiquitinated Proteins by Ubiquilin-2 Proteins Containing ALS Mutations. *PLoS One* 2015;**10**:e0130162. doi:10.1371/journal.pone.0130162

31 Kim SH, Stiles SG, Feichtmeier JM, *et al.* Mutation-dependent aggregation and toxicity in a Drosophila model for UBQLN2-associated ALS. *Hum Mol Genet* 2018;**27**:322–37. doi:10.1093/hmg/ddx403

32 Dao TP, Martyniak B, Canning AJ, *et al.* ALS-Linked Mutations Affect UBQLN2 Oligomerization and Phase Separation in a Position- and Amino Acid-Dependent Manner. *Structure* 2019;**27**:937-951.e5. doi:10.1016/j.str.2019.03.012

33 Smith BN, Topp SD, Fallini C, *et al.* Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. *Sci Transl Med* 2017;**9**. doi:10.1126/scitranslmed.aad9157

34 Müller K, Brenner D, Weydt P, *et al.* Comprehensive analysis of the mutation spectrum in 301 German ALS families. *J Neurol Neurosurg Psychiatry* 2018;**89**:817–27. doi:10.1136/jnnp-2017-317611

35 Teyssou E, Muratet F, Amador MDM, *et al.* Genetic screening of ANXA11 revealed novel mutations linked to amyotrophic lateral sclerosis. *Neurobiol Aging* 2021;**99**:102.e11-102.e20. doi:10.1016/j.neurobiolaging.2020.10.015

36 Liao D, Liao Q, Huang C, *et al.* Mutations of G38R and D40G cause amyotrophic lateral sclerosis by reducing Annexin A11 protein stability. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* 2018;**43**:577–82. doi:10.11817/j.issn.1672-7347.2018.06.001

37 Nahm M, Lim SM, Kim Y-E, *et al.* ANXA11 mutations in ALS cause dysregulation of calcium homeostasis and stress granule dynamics. *Sci Transl Med* 2020;**12**. doi:10.1126/scitranslmed.aax3993

38 Radunovíc A, Leigh PN. Cu/Zn superoxide dismutase gene mutations in amyotrophic lateral sclerosis: correlation between genotype and clinical features. *J Neurol Neurosurg Psychiatry* 1996;**61**:565–72. doi:10.1136/jnnp.61.6.565

39 Millecamps S, Salachas F, Cazeneuve C, *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. *J Med Genet* 2010;**47**:554–60. doi:10.1136/jmg.2010.077180

40 Kwon M-J, Baek W, Ki C-S, *et al.* Screening of the SOD1, FUS, TARDBP, ANG, and OPTN mutations in Korean patients with familial and sporadic ALS. *Neurobiol Aging* 2012;**33**:1017.e17-23. doi:10.1016/j.neurobiolaging.2011.12.003

41 Khani M, Alavi A, Nafissi S, *et al.* Observation of c.260A > G mutation in superoxide dismutase 1 that causes p.Asn86Ser in Iranian amyotrophic lateral sclerosis patient and absence of genotype/phenotype correlation. *Iran J Neurol* 2015;**14**:152–7.

42 Hou L, Jiao B, Xiao T, *et al.* Screening of SOD1, FUS and TARDBP genes in patients with amyotrophic lateral sclerosis in central-southern China. *Sci Rep* 2016;**6**:32478. doi:10.1038/srep32478

43 Kim H-J, Oh K-W, Kwon M-J, *et al.* Identification of mutations in Korean patients with amyotrophic lateral sclerosis using multigene panel testing. *Neurobiol Aging* 2016;**37**:209.e9-209.e16. doi:10.1016/j.neurobiolaging.2015.09.012

44 Hayward C, Brock DJ, Minns RA, *et al.* Homozygosity for Asn86Ser mutation in the CuZn-superoxide dismutase gene produces a severe clinical phenotype in a juvenile onset case of familial amyotrophic lateral sclerosis. *J Med Genet* 1998;**35**:174. doi:10.1136/jmg.35.2.174

45 Andersen PM. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr Neurol Neurosci Rep* 2006;**6**:37–46. doi:10.1007/s11910-996-0008-9