Causal inference of genetic variants and genes in amyotrophic lateral sclerosis

Siyu Pan (pansiyu2019d@big.ac.cn)
Beijing Institute of Genomics Chinese Academy of Sciences
Xinxuan Liu
Beijing Institute of Genomics Chinese Academy of Sciences
Tianzi Liu
Beijing Institute of Genomics Chinese Academy of Sciences
Zhongming Zhao
The University of Texas Health Science Center at Houston
Yulin Dai
The University of Texas Health Science Center at Houston
Yin-Ying Wang
Beijing Institute of Genomics Chinese Academy of Sciences
Peilin Jia
Beijing Institute of Genomics Chinese Academy of Sciences
Fan Liu
Beijing Institute of Genomics Chinese Academy of Sciences https://orcid.org/0000-0001-9241-8161

Research Article

Keywords: amyotrophic lateral sclerosis, causal variants, causal genes, cell type, transcriptome-wide association analysis, colocalization analysis, summary data-based mendelian randomization analysis

Posted Date: March 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1450376/v1

Abstract

Amyotrophic lateral sclerosis (ALS) is a fatal progressive multisystem disorder with limited therapeutic options. Although genome-wide association studies (GWASs) have revealed multiple ALS susceptibility loci, unraveling causal relationships between genetic variants, genes, cell types, tissues, and risk of ALS remains challenging. Here, we reported a comprehensive post-GWAS analysis of the recent large ALS GWAS (n = 80610), including functional mapping and annotation (FUMA), transcriptome-wide association study (TWAS), colocalization (COLOC) and summary da-ta-based Mendelian Randomization analyses (SMR) in extensive multi-omics datasets. Gene property analysis highlighted inhibitory neuron 6, oligodendrocytes and GABAergic neurons (Gad1/Gad2) as functional cell types of ALS and confirmed cerebellum and cerebellar hemisphere as functional tissues of ALS. Functional annotation detected the presence of multiple deleterious variants at 3 loci (9p21.2, 12q13.3, 12q14.2) and highlighted a list of SNPs that are potentially functional. TWAS, COLOC, and SMR identified 43 genes at 24 loci, including 23 novel genes and 10 novel loci, showing significant evidence of causality. Integrating multiple lines of evidence, we further propose that rs2453555 at 9p21.2 and rs229243 at 14q12 functionally contribute to the development of ALS by regulating the expression of *C9orf72* in pituitary and *SCFD1* in skeletal muscle, respectively. Overall, our study represents the most comprehensive post GWAS study of ALS to date and provides important new insights into the etiology of ALS.

Introduction

Amyotrophic lateral sclerosis (ALS, OMIM#105400) or Lou Gehrig's disease is a progressive neurological degenerative disorder without effective treatment affecting 1 in 400 individuals worldwide [1]. More than 50% of ALS patients have cognitive and behavioral dysfunctions, causing the risk of frontotemporal dementia (FTD) [2]. With the fast pace of global aging, ALS is anticipated to reach 380,000 cases globally by 2040 [3], which will not only increase the economic and psychological burden of patients [4], but also boost annual drug development cost to ~ 2 billion dollars as estimated by the ALS Association.

The heritability of ALS has been estimated at around 0.61 (95% CI: 0.38–0.78) in a twin study [5]. Genome-wide association studies (GWASs, **Supplementary Table S1**) have identified more than 35 genetic loci associated with risk of ALS [1, 6–11], with 16 loci being identified in at least two GWASs, representing the most robust genetic associations. Identifying causal genetic variants, causal genes and cell or tissue site of action remains a challenging task as over 90% of the ALS-associated variants fall in non-coding regions with largely unknown functions [12]. Recently, 3 post-GWASs [13, 8, 14] proposed lists of genes with high probabilities of causality, providing a better understanding of the genetic basis of the pathogenesis of ALS. However, these studies had limited power and overlapping findings, which may be explained by the heterogeneity of GWAS cohorts, eQTL datasets, and methodologies used. The causal genes remain to be further identified. Importantly, the functional tissues and cell types in which ALS-associated genetic variants and genes manifest remain largely unknown. The exact causal relationships connecting genetic variants, genes, cell types, tissues, and risk of ALS are also poorly understood.

Here, we report a comprehensive functional characterization of the susceptibility loci identified in the large ALS GWAS (n = 80610) [10] using functional mapping and annotation (FUMA). Further, we systematically applied transcriptome-wide association analysis (TWAS), colocalization (COLOC) and summary data-based Mendelian Randomization analysis (SMR) to prioritize putative causal genes using 18 publicly available eQTL datasets.

Materials And Methods

Gene-based analysis, tissue specificity and cell type. We performed a gene-level enrichment analysis using the Multi-marker Analysis of GenoMic Annotation (MAGMA v1.08) [15]. MAGMA extracted each SNP *p-value* from ALS GWAS summary statistics to detect genes significantly associated with ALS by combining with LD structure. SNP projected on a specific

gene was based on its genomic location to the gene, i.e., either in the gene body or 5 kb upstream and 1.5 kb downstream of the gene. As a result, we annotated 19,297 genes. LD reference dataset was created using 1000 Genomes Project European phase 3. For other parameters, we used the default option, such as, SNP-wise mean model. Based on the number of mapped genes, multiple testing of MAGMA analysis was corrected by the Bonferroni correction method (0.05/19297). Then, to identify tissue specificity of ALS, a MAGMA gene-property analysis was performed to assess relationships between tissue-specific gene expression profiles and ALS-gene associations based on the regression model [16]. Similarly, in the cell type specificity analysis [16], we tested relationships between cell-specific gene expression profiles and ALS-gene associations based on MAGMA gene-based results and multiple testing of MAGMA analysis was corrected by the FDR correction method. Briefly, a 3-step workflow was performed with scRNA-seq datasets. In the first step, cell type analysis for each dataset was carried out, and significant cell types were retained for the next step. Cell types within the dataset tended to be correlated, especially when the cell type resolution was high, therefore a true association signal was needed to identify. The second step was systematical step-wise conditional analyses to identify independent cell types within each dataset. In the final step, cross-datasets conditional analysis was performed to examine the extent of similar association signals from significant cell types retained from the second step. Single-cell RNA sequencing (scRNA-seq) datasets, including PsychENCODE (human developmental), DropViz (mouse brain samples), were available for analysis.

Functional mapping and annotation. The bioinformatic functional analysis was performed to investigate the functional relevance for ALS through the FUMA platform v1.3.6a [17], using the following toolsets or datasets, including ANNOVAR (the version released on 2017 July)), CADD v1.4 [18], RegulomeDB v1.1 [19], eQTLGen, GTEx v8 and Hi-C data. Independent significant SNPs were defined as SNPs that had a genome-wide significance ($p < 5\times10^{-8}$) and were independent of each other (LD $r2 < 0.6$) based on the 1000 Genomes European reference panel. Lead SNPs were defined as independent significant SNPs and were independent of each other (LD $r2 < 0.1$). The LD blocks of independent significant SNPs were located close to each other (less than 250 kb), and they were merged into one genomic locus by FUMA. Each locus was represented by the most significant SNP with the smallest *p-value* in the locus. Candidate SNPs were defined as SNPs that were LD (0.6) with any of the independent significant SNPs. All candidate SNPs and independent significant SNPs were annotated. Combined Annotation–Dependent Depletion (CADD) was a method that integrated 63 annotations into a single score for every variant. A higher the score indicates the variant was the more deleterious, with a threshold of 12.37 for a variant to be considered deleterious [18]. The RegulomeDB score, a categorical score ranging from 1a to 7, assessed the evidence for the regulatory potential. The lower a variant score, the more likely it is to be a regulatory element. Mapping GWAS SNPs to genes used 3 options: (1) positional mapping. This option is based on ANNOVAR annotations by specifying the maximum distance between SNPs and genes (default 10kb) or functional consequences of SNPs on genes (e.g., CADD, RegulomeDB). (2) eQTL mapping. This option mapped SNPs to genes that may affect the expression of these genes up to 1 Mb (cis eQTL). The eQTL data, including eQTLGen (cis) and whole blood and 13 brain tissue from GTEx v8 were available for analysis. Only significant SNP-gene pairs with FDR correction ($p < 0.05$) were identified to map genes. (3) chromatin interaction mapping. This option was performed using chromatin interaction data embedded in FUMA, including PsychENCODE, enhancer-promoter (EP) correlations from FANTOM5, adult and fetal cortex from Giusti-Rodriguez et al. (https://fuma.ctglab.nl/). The significance threshold of interaction was set at FDR $< 1\times10^{-6}$, which is suggested by Schmitt et al. [20]. Circos plots of chromatin interactions and eQTLs were created for each chromosome that contains at least one risk loci.

Transcriptome-wide association study. We applied a gene mapping method, S-PrediXcan[21], using summary data to test the association between the predicted expression levels of 13 brain tissues, skeletal muscle, pituitary and whole blood from GTEx v8 and ALS separately. Considering that it may not be a single tissue associated with ALS, we used S-MultiXcan framework [22], which integrated multiple tissue panels using multivariate regression, improving the power to detect ALS-associated genes. Prediction models trained on GTEx v8 were obtained on the PredictDB website (http://predictdb.org). We chose multivariate adaptive shrinkage in R (MASHR) models, because compared to previous Elastic Net models, they are parsimonious and biologically-informed using fine-mapping variants and cross-tissues QTL patterns. In addition, to

estimate summary statistics of those variants from GTEx v8 without matching the data from ALS GWAS summary, we harmonized and imputed these missing statistics using a public GitHub repository (https://github.com/hakyimlab/summary-gwas-imputation) based on best linear unbiased predictors (BLUP) method [23]. Genes with association FDR < 0.05 were considered to be significantly associated with ALS passing the multiple testing correction. A circus plot for multiple TWASs was created by R package "circlize" [24].

In addition, we compared the identified genes or loci with the published TWASs and GWASs (**Supplementary Table S1**, *p* threshold < 5×10^{-8}), which was visualized with R package "VennDiagram" [25]. However, the latest TWAS from Megat et al. [26] was unavailable in the medrxiv and not included in the comparison.

Bayesian Colocalization analysis. We applied COLOC [27], a Bayesian colocalization method completed in R package "coloc", to assess the probability of the same variant being responsible for ALS risk and gene expression level. This approach enumerates all possible configurations, including five mutually exclusive hypotheses: (1) H_0: no causal SNP in the test region; (2) only one causal SNP for ALS GWAS; (3) only one causal eQTL for gene expression; (4) two distinct causal SNPs in the test region; (5) sharing a causal SNP within the test region. Each of these hypotheses is quantified as a posterior probability (PP), named PP0 – PP4 in the corresponding order. The prior probabilities are set by default. Default conservative priors, p_1 (1×10^{-4}), p_2 (1×10^{-4}) and p_{12} (1×10^{-5}), were used in the colocalization analysis, where p_1 and p_2 is the probability that a given SNP is associated with either of ALS and gene expression and p_{12} is the probability that a given SNP is associated with both traits.

The selection of eQTL dataset is mainly based on the result of the tissue enrichment analysis and previous known tissues of ALS. First, we assessed the posterior probability of colocalization between ALS and joint effect of Brain-eMeta eQTL data from a meta-analysis of GTEx v6p brain, the CommomMind Consortium (CMC) and Religious Orders Study and Memory and Aging Project (ROSMAP) by Meta-analysis of cis-eQTL in Correlated Samples (MeCS) [28]. Then, we extended the analysis to other specific eQTL datasets, including 13 brain tissues, skeletal muscle, pituitary and whole blood from GTEx v8, eQTLGen consortium, respectively. For three eQTL datasets, cis-association analyses were defined for SNPs within 1MB of the transcription start site or the center of the probe [29–31]. We analyzed about 2 Mb cis-region available in these eQTL datasets to confirm a sufficiently large region associated with gene expression. GTEx v8 and eQTLGen. A detailed summary of eQTL data is in **Supplementary Table S2**.

All SNPs in the ALS GWAS summary statistics that were within a region (1Mb) from the TSS and TES of a given gene were identified. These regions, where at least one SNP having a *p* value < 1×10^{-6} from GWAS that had a *p* value < 1×10^{-4} in the eQTL dataset, were included in the colocalization assessment. In addition, we assume that the posterior probability of colocalization between ALS and gene expression is driven by a single causal variant. We then applied previously suggested combination cutoffs (PP4 ≥ 0.75, PP3 + PP4 ≥ 0.9 and PP4/PP3 ≥ 3) as powerful evidence supporting a causal role for the gene to be a mediator of ALS risk. The regional association plots were generated by the R package "locuscomparer" [32].

Summary data-based Mendelian randomization analysis and HEIDI test. SMR [33] is a method performed to evaluate the association between an exposure (gene expression level) and an outcome (trait) in Mendelian randomization analysis principles by removing non-genetic confounding using a variant as an instrument variable. In a nutshell, phenotype (ALS) is the outcome (y), gene expression level is the exposure (x) and a single genetic variant, the top cis-eQTL that has a strong association with the exposure and is used as an instrumental variable (z). In the regression analysis, b_{zx} is the effect size of z on x, b_{zy} is the effect size of z on y, and b_{xy} is the effect size of x on y. The MR analysis is used to assess the effect size of gene expression on phenotype (b_{xy}), which is defined as $b_{zy}/b_{zx,}$ the latter two are estimated from GWAS summary and eQTL dataset. Like colocalization analysis, we assessed the causal effect between ALS GWAS and eQTL datasets, including eMeta (brain), 13 brain tissues, skeletal muscle, pituitary, blood from GTEx v8 and eQTLGen consortium. To decrease the false discovery rate (FDR), a Benjamini-Hochberg method correction (FDR < 0.05) was used in each SMR analysis for multiple testing.

The association signals detected by SMR did not mean that both traits were affected by the same causal variant, they were divided into three possible situations: 1) horizontal pleiotropy: exposure and outcome affected by the same variant; 2) vertical pleiotropy: causal effect between exposure and outcome; 3) linkage: the variant affected the exposure was in LD with another variant affected the outcome. We applied HEIDI built into SMR to distinguish pleiotropy from linkage, and a threshold (HEIDI > 0.01) was considered to have little evidence of heterogeneity. In this study, considering the sample size of the current datasets, we used SMR version 1.03 with the following parameters: *p-value* threshold $p < 1\times10^{-6}$ to select the top eQTL for SMR test, a threshold for the difference in SNP allele frequency between data sets was set to 1 and other parameters were default options.

Results

Study design and analysis workflow. A schematic overview of the study design is illustrated in Fig. 1. We conducted a survey of 26 GWASs [34–51, 1, 52, 10, 53–55, 11, 56] of ALS in the NHGRI-EBI GWAS Catalog (till Sep. 2021) and additionally included one most recent GWAS [57] and 3 recent post-GWASs [13, 8, 14] to summarize the current knowledge on the susceptibility loci and candidate genes of ALS (**Supplementary Table S1**, Fig. 2B). We then based our post-GWAS analyses on ALS GWAS from Nicolas et al. [10], which represents the largest-ever GWAS data for ALS so far, which incorporated data from multiple cohorts, totaling 20,806 cases and 59,804 controls of European ancestry.

Briefly, we firstly characterized a large set of possible tissues and cell types potentially functionally involved in the development of ALS by conducting a gene property analysis using FUMA. We next annotated the potential functions of a set of candidate genes using CADD score, RegulomeDB score, relative physical positioning with ALS-associated SNPs, evidence of eQTL, and chromatin interactions. We further prioritized the causality for a list of candidate genes and tissues by applying TWAS, COLOC, and SMR to 18 eQTL datasets from GTEx v8, eMeta, and eQTLGen. We finally integrated the findings from different methods and provided a list of variants and genes in corresponding tissues with high probabilities of causality.

Gene property analysis highlights a list of tissues and cell types. The platform FUMA has implemented the MAGMA method for tissue-specificity analysis using 54 tissues from GTEx. By using FUMA, we identified 13 brain regions and pituitary, in which tissue-specific gene expression profiles showed nominally significant association with ALS-gene associations ($p <$ 0.05) and two regions showed significance after Bonferroni correction: cerebellum (Bonferroni adjusted $p = 2.7\times10^{-5}$) and cerebellar hemisphere ($p = 1.3\times10^{-4}$) (**Supplementary Figure S2**). These findings are consistent with those of the most recent GWAS by van Rheenen W et al. [57], emphasizing an important role of various brain tissues and in particular cerebellum and cerebellar hemisphere in the development of ALS. In addition, we highlighted pituitary as an important tissue for the analysis of locus-specific gene expression profiles within ALS-risk loci. We thus focused on the 13 brain regions and pituitary in our subsequent analyses, and additionally included skeletal muscle, which has been implied in ALS progression [58, 59], and whole blood, which was also implied in previous studies [14].

MAGMA cell-type specificity analyses in a total of 196 cell types were performed in 8 different mouse and human single-cell RNA-seq datasets. In the human brain dataset from PsychENCODE, we identified both excitatory and inhibitory neurons as significantly (FDR < 0.05) associated with ALS (**Supplementary Figure S3A, Supplementary Table S3**), meaning that in these cells gene expression profiles are significantly associated with ALS-gene associations. Within-dataset step-wise conditional analyses identified inhibitory neuron 6 showing independent association (**Supplementary Figure S3B, Supplementary Table S4, S5**), further highlighting the likelihood of inhibitory neuron 6 being the basic functional unit of ALS. In 7 mouse datasets, within-dataset conditional analyses additionally identified oligodendrocytes and gamma-aminobutyric acidergic (GABAergic) neurons (Gad1/Gad2) as significantly associated cell types (**Supplementary Figure S3A, Supplementary Table S3-5**). To evaluate if the significantly associated cell types from distinct datasets are driven by similar genetic signals, we further conducted a cross-dataset conditional analysis for all significant cell types in all 8 datasets. This analysis highlighted that inhibitory neuron 6, oligodendrocytes and GABAergic neurons (Gad1/Gad2) are

likely driven by distinct genetic signals, while various oligodendrocytes from different datasets are likely driven by similar genetic signals (**Supplementary Figure S3C, Supplementary Table S6**).

Our findings regarding inhibitory and excitatory neurons are consistent with the very recent study of Megat et al. [26], which showed that a large part of ALS heritability lies in genes expressed in inhibitory and excitatory neurons. Our conditional analysis further pinpointed inhibitory neuron 6 as the most likely functional unit. Our findings regarding oligodendrocytes and GABAergic neurons are highly consistent with the study of Saez-Atienzar et al. [60] and partially consistent with the study of van Rheenen et al. [57], in which oligodendrocyte was not significantly enriched in ALS but significant in PD and AD proxy.

In addition, MAGMA gene-based analysis identified 7 genes showing significant association with ALS risk, and all have been identified in previous GWASs (**Supplementary Figure S1**).

Functional annotation of ALS-associated SNPs and genes. We annotated the functionality of 233 ALS-associated candidate SNPs from the 6 independent genome-wide significant loci (5q33.1 rs10463311, 9p21.2 rs3849943, 12q13.3-14.1 rs142321490, 12q14.2 rs74654358, 19p13.11 rs12973192, 21q22.3 rs75087725) in the GWAS of Nicolas et al. [10] using CADD [18], which predicts SNP deleteriousness based on an integrative annotation built from more than 60 genomic features. Of 233 candidate SNPs, the CADD analysis identified a total of 17 SNPs at 3 loci (9p21.2, 12q13.3, 12q14.2) with high scores (> 12.37) [18], suggesting a strong deleterious effect of these variants (**Supplementary Table S7**). At 9p21.2, 14 SNPs had high scores tightly surrounding *C9orf72*, with the highest score observed for rs3736319 (18.5), 83 bp upstream of *MOB3B*, and 7.6 kbp downstream of *C9orf72*, and the second highest score observed for rs10967965 (17.2), an intronic variant of *MOB3B*. At 12q14.1, two SNPs had high scores, including a UTR5 variant of *NAB2* (rs185306972) and a nonsynonymous variant of *KIF5A* (rs113247976). At 12q14.2, one synonymous variant of *TBK1* (rs41292019) had a high score.

A complementary analysis using the RegulomeDB [19] annotation further revealed three SNPs at 9p21.2 with strong evidence of regulation supported by eQTL and TF binding/DNase peak (evidence level 1f, **Supplementary Table S7**). All three were located in the intronic region of *MOB3B* and very close to *C9orf72* (< 60kbp), with one (rs10967965) also highlighted in the CADD analysis. These results provided evidence for the presence of deleterious variants with pathogenic effects and SNPs with regulatory effects in 3 ALS-associated loci and provided the list of candidate genes in these loci.

Next, positional mapping, eQTL mapping, and chromatin interaction mapping analyses were conducted in the 6 ALS-associated loci. Together, these analyses mapped to 58 genes, among which 4 genes were mapped by all three mapping methods, including 5q33.1 *TNIP1*, 9p21.2 *C9orf72*, *MOB3B*, *IFNK* (**Supplementary Table S8**). The positional mapping with a maximum distance 10 kbp mapped to 18 genes in 6 loci. The eQTL mapping identified 1,434 significant SNP-gene-tissue pairs (FDR < 0.05), mainly in brain tissues (948 pairs), which mapped to 22 expressed genes in 5 loci (**Supplementary Table S9**). The chromatin interaction mapping identified 34 genes in 5 loci with 6 genes overlapping with those from eQTL mapping, i.e., 5q33.1 *TNIP1*, 9p21.2 *C9orf72*, *MOB3B*, *IFNK*, 12q13.3 *BAZ2A*, *PRIM1*. Interestingly, different from other loci, the 9p21.2 locus clearly contained a DNA loop [61] in brain tissues (**Supplementary Figure S4**), which made parts of DNA closer together and allowed genes to be activated by regulatory elements known as enhancers. The 2 loci on chromosome 12 (12q14.1 and 12q14.2) contained more signals for both eQTL and chromatin interactions compared with the other 4 loci (**Supplementary Figure S5**). These results provided direct evidence for a list of SNPs and genes that are potentially functionally involved in the development of ALS.

Multi-tissue TWAS identified novel functional candidate genes. Based on the full GWAS summary statistics from the study of Nicolas et al. [10], we conducted a series of TWASs to test the association between gene expression levels predicted using PredictDB and ALS risk in 13 different brain tissues, pituitary, skeletal muscle, and whole blood. TWASs were carried out using S-PrediXcan, which analyzes one tissue at a time, and S-MultiXcan, which jointly analyzes all tissues.

S-PrediXcan found a total of 31 genes at 19 distinct loci showing significant (FDR < 0.05) association with ALS risk in at least one tissue. Among the 19 loci, 5 (1q23.3, 6q14.1, 16q24.1, 17p13.2, 22q13.33) are newly identified (Fig. 2B), highlighting 6 genes (*NR1I3*, *PCP4L1*, *UBE3D*, *ZDHHC7*, *MIS12*, *DENND6B*). In addition, 16 genes have not been previously suggested as functional candidates of ALS (Fig. 2A, **Supplementary Table S10**). Among the 31 genes, the most significant finding was *C9orf72* (FDR = 5.03×10^{-18} in Brain_Nucleus_accumbens_basal_ganglia), which was at orders of magnitude more significant than any other gene in any tissue (minimum FDR = 0.001). *C9orf72*, representing the most well-established gene involved in the risk of ALS, was significant not only in 11 brain regions but also in pituitary, skeletal muscle, and blood. The second most significant finding was the gene *SCFD1* at 14q12, which also showed a significant association with ALS risk in 10 brain regions, pituitary, skeletal muscle, and blood (min FDR = 0.001 in Brain_Cerebellar_Hemisphere). This gene also represented a well-established candidate gene for ALS risk. The 16 newly identified genes had similar significance levels (with FDR ranging between 0.001 and 0.05), and all were significant in up to three tissues. Among these 16 genes, 13 from 11 loci were significant in at least one brain region, and 3 from 3 different loci were significant only in non-brain tissues, i.e., skeletal muscle (12q13.3 *PIP4K2C*), blood (17q12 *DHRS11*), and pituitary (16q24.1 *ZDHHC7*).

S-MultiXcan found a total of 22 genes at 14 distinct loci, among which 6 genes at 6 distinct loci (8q13.2 *ARFGEF1*, 12q13.3 *OS9*, 12q14.1 *CTDSP2*, 12q24.31 *RP11-173P15.9*, 13q12.3 *LINC00426*, 15q25.2 *HDGFRP3*) were not identified in S-PrediXcan (Fig. 2A, **Supplementary Table S11**). Among the 6 S-MultiXcan-only loci, one gene (15q25.2 *HDGFRP3*) has been reported in a previous TWAS study [14], and the other 5 genes represent new findings. Among these 5 genes, the most significant was *ARFGEF1* (FDR = 6.4×10^{-4}), which involves in vesicular trafficking and has previously been suggested to play a role in pathogenesis in ALS based on Gene Ontology [62].

Overall, our S-PrediXcan and S-MultiXcan together identified 21 novel genes at 15 novel loci, complementing the lists of previously established functional candidate genes and ALS-associated loci.

Colocalization highlights genotype-mediated genes in corresponding tissues. GWAS associations driven by eQTLs may indicate functional mechanisms. However, few studies have investigated the colocalization with eQTLs for ALS-associated loci. We conducted a series of eQTL colocalization analyses in 13 brain tissues, pituitary, skeletal muscle, and blood utilizing a variety of eQTL datasets from GTEx v8 and eQTLGen consortium. These analyses identified a total of 9 genes at 5 loci showing significant evidence (PP4 > 0.75 & PP3 + PP4 > 0.9 & PP4/PP3 > 3, Fig. 3A, **Supplementary Table 12**) of colocalization with eQTLs in at least one tissue investigated. These included 5q33.1 (*TNIP1*, *GPX3*), 9p21.2 (*C9orf72*), 10q25.2 (*ZDHHC6*, *ACSL5*), 14q12 (*SCFD1*, *G2E3*), 14q32.12 (*TRIP11*, *RP11-529H20.6*).

The strongest signal according to the PP4/PP3 ratio was identified for rs2453555 at 9p21.2 (PP4/PP3 = 82.5), which was highly significantly associated with ALS risk (GWAS $p = 6.5\times10^{-30}$) and at the same time served as a highly significant eQTL of *C9orf72* in the pituitary gland (eQTL $p = 4.4\times10^{-12}$), strongly suggesting a causal relationship (Fig. 3B). This SNP was also a significant eQTL of *C9orf72* in spinal cord cervical, albeit at a much lower significance level (eQTL $p = 8.1\times10^{-6}$), thus having less strong evidence of colocalization (PP4/PP3 = 5.2). The association at this locus showed no evidence of colocalization with eQTLs in other tissues investigated. The SNP rs2453555 is in very high linkage disequilibrium (LD) with the most significant SNP (rs3849943, $p = 3.8\cdot10^{-30}$, $r^2 = 0.98$) in the GWAS of Nicolas et al. [10]. This result pinpoints rs2453555, which may regulate the expression of *C9orf72* in the pituitary, and consequently modifies the risk of ALS. A very recent study failed in finding colocalization signals for *C9orf72* [63]. Compared with their study, our study used the newest version of GTEx, which has an average 24% increased sample size.

The second strongest signal was observed for 10q25.2 *ZDHHC6* in the cerebellum (56.7) as well as in other 5 different brain tissues (5.1 to 19.9). The other gene (*ACSL5*) at this locus also showed significant colocalization (5.1) but at a much lower significance level than *ZDHHC6* and the signal was observed only in blood. The third signal was 5q33.1 *TNIP1* with colocalization signals in the cerebellar hemisphere (18.0) and blood (9.7) but not in other tissues. The other gene at this locus (*GPX3*) was significant in only blood (10.0). The fourth signal was 14q12 *SCFD1* in six brain tissues, skeletal muscle,

and blood at similar significance levels (PP4/PP3 ranging between 9.9–12 except in nucleus accumbens basal ganglia, where PP4/PP3 = 6.0). The other gene at this locus (*G2E3*) was detected only in skeletal muscle at a further decreased level of significance (5.5). The last signal was 14q32.12 *TRIP11* in the cerebellum, cerebellar hemisphere, pituitary, skeletal muscle, and blood with similar levels of significance (5.1–6.1). The other gene at this locus (*RP11-529H20.6*) showed a relatively weak colocalization signal in the blood (3.7). These results provided direct evidence of causality for a specific set of SNPs, genes, and corresponding tissues likely functionally involved in the development of ALS (**Supplementary Table S12**).

SMR illustrates the causal relationships between SNPs, gene expressions, and ALS risk. SMR & HEIDI is a powerful approach to test whether the effect of an SNP on a phenotype is mediated by transcription. Four studies [13, 52, 57, 64] have previously conducted SMR & HEIDI analyses for ALS. However, they were either conducted using smaller size GWAS or using a limited number of tissues. Here, based on the GWAS of Nicolas et al. (n = 80610), we additionally analyzed 9 brain tissues, pituitary, and skeletal muscle from GTEx v8.

Our SMR & HEIDI analysis identified a total of 9 genes from 6 loci with significant evidence mediating the genetic associations observed in these loci (FDR < 0.05 & HEIDI > 0.01, Table 1). These included 3q24 *PLOD2*, 9p21.2 *C9orf72*, 10q22.2 *NDST2*, 14q12 *SCFD1*, 17q12 *GGNBP2*, *MYO19*, *DYNLL2*, *ZNHIT3*, and 22q13.33 *PLXNB2*. Among these 6 loci, 3q24 and 22q13.33 have not been previously reported. At 3q24, *PLOD2* showed significant evidence mediating the association between rs149615181 and ALS risk in skeletal muscle. This gene encodes a membrane-bound homodimeric enzyme localized to the cisternae of the rough endoplasmic reticulum, which plays a critical role in the stability of intermolecular collagen crosslinks and progressive degradation of the extracellular matrix, but little is known about its potential relationship with ALS. At 22q13.33, *PLXNB2* was significant in blood. This gene is a transmembrane receptor involved in axon guidance and cell migration in response to semaphorins [65] and recently showed that it mediates the neurogenesis and neuroprotective activities of angiogenin [66], which was implicated in ALS and AD [67].

Table 1
Genes mediating the genetic associations with ALS in 6 loci from SMR & HEIDI analyses

Locus	Gene	SNP	FDR	HEIDI	Tissue	Database
3q24	PLOD2	rs149615181	0.04	0.69	**Muscle_Skeletal**	GTEx v8
9p21.2	*C9orf72*	rs2453565	2.00×10^{-5}	0.14	**Pituitary**	GTEx v8
		rs700795	0.02	0.19	**Brain_Spinal_cord_cervical_c-1**	GTEx v8
10q22.2	NDST2	rs11000785	0.05	0.07	Blood	eQTLGen
14q12	*SCFD1*	rs7144204	$3.6\cdot10^{-3}$	0.10	Blood	eQTLGen
		rs448175	0.01	0.35	Blood	GTEx v8
		rs229152	0.03	0.45	Brain_Cerebellum	GTEx v8
		rs229243	0.04	0.31	**Muscle_Skeletal**	GTEx v8
		rs2070339	0.03	0.22	Multiple brain regions	Brain-eMeta
17q12	*GGNBP2*	rs11650008	0.03	0.04	Blood	eQTLGen
	MYO19	rs7222903	0.04	0.60	Blood	eQTLGen
	DYNLL2	rs2877858	0.04	0.16	Blood	eQTLGen
	ZNHIT3	rs4796224	0.05	0.50	Blood	eQTLGen
22q13.33	PLXNB2	rs62241220	0.02	0.77	Blood	eQTLGen

All genes with FDR < 0.05 & HEIDI > 0.01 are shown. Loci that have not been identified in previous GWASs or post-GWASs are indicated in bold. Genes that have not been reported in previous SMR studies are indicated in bold. Tissues that have not been reported in previous SMR studies are indicated in bold.

Interestingly, at 9p21.2, *C9orf72* was found highly significantly mediating the association between rs2453565 and the risk of ALS in the pituitary (FDR = $2\cdot10^{-5}$), which was at orders of magnitude more significant than in other brain and non-brain tissues. This SNP is in very high LD with rs2453555 (r^2 = 0.95) identified by our colocalization analysis as described above. This result boosted the likelihood of a causal chain between rs2453555/rs2453565, expression of *C9orf72* in the pituitary and the risk of ALS. In the very recent study of van Rheenen W et al. [57], HEIDI test rejected the hypothesis that expression of *C9orf72* could mediate the association between rs2453555 and ALS risk in blood. Our finding stresses the pituitary being the correct tissue where *C9orf72* plays a functional role in the development of ALS.

At 14q12, *SCFD1* in blood (FDR = $3.6\cdot10^{-3}$ in eQTLGen), cerebellum (FDR = 0.03) and skeletal muscle (FDR = 0.04) showed significant mediatory effects on genetic association with ALS. Notably, rs229243 was detected to increase risk of ALS by modifying *SCFD1* expression in skeletal muscle. This SNP is also a significant eQTL of *SCFD1* in skeletal muscle as found in our colocalization analysis (PP4/PP3 = 9.9). Interestingly, *SCFD1* had different effect directions in blood, skeletal muscle and cerebellum (**Supplementary Table S13**). A very recent SMR analysis [64] found that rs229243 had a regulatory effect on ALS risk mediated by the expression of *SCFD1* in the blood (GTEx) and cerebellum. Our SMR and colocalization results thus further provided evidence for skeletal muscle as an additional tissue possibly of function.

At 17q12, expressions of four genes (*GGNBP2*, *MYO19*, *DYNLL2*, *ZNHIT3*) in the blood significantly mediated the genetic association in this locus (**Supplementary Table S13**). Among these four genes, *GGNBP2* was the most significant (FDR = 0.03), consistent with the finding from a previous study [52].

Integration of evidence pinpoints causal genes in corresponding tissues. Overall, our study identified a total of 43 genes at 24 loci showing significant evidence of causality (**Supplementary Table S14**). Among these 43 genes, 23 genes at 17 loci have not been functionally linked with ALS in previous studies. Among the 24 loci, 10 loci (9 from TWAS, one from SMR) have not been associated with ALS risk in previous studies.

A total of 8 genes at 6 loci were significant in at least two out of three analyses, i.e., TWAS, colocalization analysis and SMR analysis. These included 5q33.1 *TNIP1*, 9p21.2 *C9orf72*, 10q25.2 *ACSL5*, 10q22.2 *NDST2*, 14q12 *SCFD1*, 17q12 *MYO19*, *GGNBP2*, and *ZNHIT3*. All these 6 loci have been previously associated with ALS risk, and all 8 genes have been previously suggested as the functional candidates. Integrating the results from three different analyses conducted in various tissues, our study further revealed their most likely corresponding functional tissues (Table 2).

Table 2
Integration of TWAS, COLOC and SMR results.

Locus	Gene	Tissue	TWAS	COLOC (PP4/PP3)	SMR (HEIDI)	Overall evidence
5q33.1	*TNIP1*	Blood	2.2×10^{-3}	0.91 (9.70)	3.6×10^{-3} (6.7×10^{-4})	**
9p21.2	*C9orf72*	Brain_Spinal_cord_cervical_c-1	7.00×10^{-13}	0.81 (5.10)	1.5×10^{-2} (0.19)	***
	C9orf72	Pituitary	5.00×10^{-14}	0.99 (82.50)	2.00×10^{-5} (0.14)	***
10q25.2	*ACSL5*	Blood	2.7×10^{-2}	0.82 (5.10)	-	**
10q22.2	*NDST2*	Blood	3.3×10^{-2}	-	4.9×10^{-2} (0.07)	**
14q12	*SCFD1*	Brain_Cerebellum	1.7×10^{-3}	0.92 (11.81)	3.1×10^{-2} (0.45)	***
	SCFD1	Muscle_Skeletal	3.7×10^{-2}	0.91 (9.85)	3.8×10^{-2} (0.31)	***
	SCFD1	Blood	1.5×10^{-3}	0.92 (12.00)	1.1×10^{-2} (0.35)	***
	SCFD1	Brain_Anterior_cingulate_cortex_BA24	3.3×10^{-3}	0.92 (11.85)	-	**
	SCFD1	Brain_Cerebellar_Hemisphere	1.1×10^{-3}	0.92 (11.35)	-	**
	SCFD1	Brain_Cortex	2.4×10^{-3}	0.91 (10.75)	-	**
	SCFD1	Brain_Frontal_Cortex_BA9	1.3×10^{-3}	0.92 (11.80)	-	**
	SCFD1	Brain_Nucleus_accumbens_basal_ganglia	2.7×10^{-3}	0.77 (6.01)	-	**
17q12	*MYO19*	Blood	2.2×10^{-3}	-	4.0×10^{-2} (0.6)	**
	GGNBP2	Blood	1.3×10^{-2}	-	2.5×10^{-3} (0.04)	**
	ZNHIT3	Blood	3.5×10^{-2}	-	4.9×10^{-2} (0.5)	**

Blood refers to whole blood from GTEx v8 or eQTLGen depending on which is more significant. The TWAS column indicated the *p-value* (FDR). The COLOC column indicated PP4 and PP4/PP3. The SMR column indicated the *p-value* (FDR) of the SMR test and the HEIDI test. The number of asterisks in the overall evidence column represents the number of significant results from three different analyses.

The most significant finding was for 9p21.2 *C9orf72*, which was highly significant in all three analyses, and all three analyses pinpointed pituitary as the most likely functional tissue, with orders of magnitude more significant than in any

other tissues. We thus propose that in the pituitary, the expression of *C9orf72*, regulated by rs2453555, is causally associated with ALS risk.

14q12 *SCFD1* in cerebellum, skeletal muscle and blood were significant in all three analyses and multiple other brain tissues were supported by two analyses, emphasizing the multi-tissue effect of *SCFD1*.

The remaining 4 loci (5q33.1, 10q25.2, 10q22.2, and 17q12) were supported by two analyses, but all suggesting blood instead of brain tissues being the causal tissue. This finding is somehow surprising and requires experimental validations in future studies. For 17q12, three genes (*MYO19*, *GGNBP2*, and *ZNHIT3*) are functional candidates. A previous study [14] suggested *MYO19* as the most likely functional candidate of this locus, while another [52] suggested *GGNBP2*. Our analysis suggested that *GGNBP2* is less competitive with the other two as it had a more significant HEIDI ($p = 0.04$).

Discussion

This study represents the most comprehensive post-GWAS of ALS to date. Our gene property analysis highlighted inhibitory neuron 6, oligodendrocytes and GABAergic neurons (Gad1/Gad2) as functional cell types of ALS and confirmed cerebellum and cerebellar hemisphere as functional tissues of ALS. Functional annotation analysis detected the presence of multiple deleterious variants at 3 loci (9p21.2, 12q14.1, 12q14.2) and highlighted a list of SNPs that are potentially functional. TWAS, COLOC, and SMR identified 43 genes at 24 loci, including 23 novel genes and 10 novel loci, showing significant evidence of causality. Integrating multiple lines of evidence, we further propose that rs2453555 at 9p21.2 and rs229243 at 14q12 functionally contribute to the development of ALS by regulating the expression of *C9orf72* in pituitary and *SCFD1* in skeletal muscle, respectively.

Our gene property analysis suggested the cerebellum as one of the core tissues functionally involved in the development of ALS, confirming the finding in a recent GWAS [57]. The cerebellum has long been recognized as having an essential role in complete motor function and maintaining higher cognitive function levels [68]. Previous functional studies have also shown structural and functional damage to the cerebellum in ALS [69, 70].

Our gene property analysis also identified a functional relationship between a novel cell type and ALS risk, i.e., inhibitory neurons, especially inhibitory neuron 6. This finding is an important supplement of previous findings suggesting microglia [71], astrocytes [72, 60], and glutamatergic neurons [13] as the key cell types functionally involved in the development of ALS. Our finding on inhibitory neurons is highly consistent with a recent study, which showed that genetic variants associated with ALS were mostly located in genes expressed in neurons, particularly in inhibitory neurons, and ALS risk loci were significantly enriched excitatory and inhibitory neurons using single cell assay for transposase-accessible chromatin sequencing (scATAC-seq) profiles [26]. Excitatory neurons did not survive in our conditional analysis and only inhibitory neuron 6 did, suggesting a genuine involvement between this cell type and ALS. It is known that inhibitory neurons release the neurotransmitter gamma-aminobutyric acid (GABA) to regulate the initiation of excitatory neurons, ensuring our brain functions smoothly and accident-free [73, 74]. A loss of inhibitory neurons influence is an important factor leading to ALS pathogenesis [75].

Our study integrated multiple gene mapping approaches, which provided more reliable results than the studies based on a single approach. Considering the complexity of causal gene identification, i.e., a high level of redundant eQTLs [76], complex patterns of LD between ALS-associated SNPs, and a large proportion of non-coding variants, our results by integrating different mapping approaches appeared more precise and had fewer false negatives. An example of precision is *C9orf72*, which was significant in multiple analytical approaches and multiple brain tissues. Integrating these results, we found that *C9orf72* was most significant in the pituitary, which was at orders of magnitude more significant than in any other tissues, pinpointing the pituitary as the true causal tissue. An example of avoiding false negatives is *DENND6B*, which was only significant in S-PredictXcan, whereas a previous study showed that it involved in vesicle-mediated transport and Rab guanyl-nucleotide exchange factor activity, an emerging role in ALS [77].

Compared with previous studies [13, 8, 14], our study provided a comprehensive characterization of genetic loci and integrated more methods and eQTL datasets to identify causal genes that affect ALS through gene expression in the corresponding tissues. The genes identified in specific tissues could be promising ALS candidate genes for functional follow-up experiments. Integrating multiple lines of evidence, we propose that rs2453555 at 9p21.2 functionally contribute to the development of ALS by regulating the expression of *C9orf72* in the pituitary. This SNP is the top significant in both COLOC and SMR (rs2453555, rs2453565 in high LD with top rs3849943) analysis, strongly suggesting a functional role. This SNP tagging a highly pathogenic repeat expansion (GGGGCC) is also in high LD with multiple pathogenic variants at 9p21.2, although the possibility of presence of multiple causal variants at this locus cannot be excluded. The aggregation of dipeptide repeat proteins (DPRs) originating from the *C9orf72* repeat expansion could result in disorders of hormone secretion and regulation in the pituitary, followed by disruption of the hypothalamic-pituitary axis [78, 79]. Our identification of pituitary was highly consistent in our TWAS, COLOC, and SMR analysis. The failure of identification of pituitary and rs2453555 in the recent two studies [63, 57] is likely because only blood and cortex or a smaller eQTL dataset were used. Besides, the cervical spinal cord was also of interest to another tissue as a similar pattern of rs2453555. A recent neuroimaging study reported that significant cerebral white matter (WM) atrophy was detected at every cervical vertebral level of *C9orf72* hexanucleotide expansion carriers [80]. In contrast, another study showed that cervical spinal cord progressively occurs to thin in ALS patients with *C9orf72* repeat expansion [81]. This discrepancy could be due to many factors, such as sample size, imaging techniques and statistical analysis methods.

Multiple lines of evidence support that rs229243 at 14q12 functionally contributes to the development of ALS by modifying the expression of *SCFD1* in the skeletal muscle, providing a possible functional tissue. *SCFD1* is involved in vesicular transport between the endoplasmic reticulum and the Golgi [82]. Surprisingly, skeletal muscle was also found to be the relevant tissue for *SCFD1*. Although the weakening of skeletal muscle weakness was thought to be the initial hallmark of ALS, whether ALS originates in peripheral tissues (dying-back phenomenon) [83], including skeletal muscle or motor neurons (dying-forward phenomenon) [84] has been fiercely debated. Skeletal muscle was not considered pivotal to the etiology and treatment of ALS until recent years [58]. Current studies about skeletal muscle degeneration/regeneration process focused on mutant *SOD1* mouse mode to the pathology of ALS [85, 86, 59], but to our knowledge, *SCFD1* has never been investigated in molecular biology experiments or genetic models. Our findings could contribute to the understanding of skeletal muscle pathology and may provide a new therapeutic target for ALS. In addition to the above genes, we also found 23 novel functional candidate genes, among which 12 have supportive evidence from literature [87, 77, 88–101, 62, 102, 103]. For details, please see **Supplementary Table S14**.

However, our study is not without limitations. Firstly, the sample size of ALS GWAS is relatively small compared with the latest GWAS [57], a parallel to our study. The sample size in different tissues and types for eQTL tissues limited our ability to identify ALS-related genes and explore the search for causal genes in other pathologically relevant tissues. Secondly, LD structure [27] and gender differences [31] in the sample may bias findings due to unavailability of raw data. Thirdly, with the constraints of currently available data, only cis-eQTL data were included in our analysis, which may miss the actual causal genes. Fourthly, tissue pleiotropy and cell-type heterogeneity. The disease rarely works in a single tissue. Some genes could exert a causal effect on disease in specific tissues or cell types different from our reference panel, which may introduce bias and incompleteness. Therefore, identification of the possible causal tissue or cell type of each gene may be a hot topic for future development and single-cell RNA sequencing holds promise for more refined studies in the future. Finally, further biological function experiments are needed to confirm the biological role of these genes in the pathology of ALS.

Conclusions

This study established a list of causal relationships between genetic variants, candidate genes, functional cell types and tissues, and risk of ALS. The prioritized genes and tissues serve as targets of future functional and drug discovery studies.

Declarations

Data Availability

All data generated or analyzed during this study are available from published article and public databases.

Code Availability

Not applicable.

Ethics Approval

Not applicable.

Consent to Participate

Not applicable.

References

1. 10.1038/ng.3622.
2. Robberecht W, Philips T. The changing scene of amyotrophic lateral sclerosis. Nat Rev Neurosci. 2013;14(4):248–64. doi:10.1038/nrn3430.
3. Logroscino G, Piccininni M. Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference. Neuroepidemiology. 2019;52(1–2):93–103. doi:10.1159/000493386.
4. Gladman M, Zinman L. The economic impact of amyotrophic lateral sclerosis: a systematic review. Expert Rev Pharmacoecon Outcomes Res. 2015;15(3):439–50. doi:10.1586/14737167.2015.1039941.
5. Al-Chalabi A, Fang F, Hanby MF, Leigh PN, Shaw CE, Ye W, Rijsdijk F. An estimate of amyotrophic lateral sclerosis heritability using twin data. J Neurol Neurosurg Psychiatry. 2010;81(12):1324–6. doi:10.1136/jnnp.2010.207464.
6. Zhou S, Zhou Y, Qian S, Chang W, Wang L, Fan D. Amyotrophic lateral sclerosis in Beijing: Epidemiologic features and prognosis from 2010 to 2015. Brain Behav. 2018;8(11):e01131. doi:10.1002/brb3.1131.
7. van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, Lemmens R, Schelhaas HJ, Groen EJ, Huisman MH, van der Kooi AJ, de Visser M, Dahlberg C, Estrada K, Rivadeneira F, Hofman A, Zwarts MJ, van Doormaal PT, Rujescu D, Strengman E, Giegling I, Muglia P, Tomik B, Slowik A, Uitterlinden AG, Hendrich C, Waibel S, Meyer T, Ludolph AC, Glass JD, Purcell S, Cichon S, Nothen MM, Wichmann HE, Schreiber S, Vermeulen SH, Kiemeney LA, Wokke JH, Cronin S, McLaughlin RL, Hardiman O, Fumoto K, Pasterkamp RJ, Meininger V, Melki J, Leigh PN, Shaw CE, Landers JE, Al-Chalabi A, Brown RH Jr, Robberecht W, Andersen PM, Ophoff RA, van den Berg LH. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat Genet. 2009;41(10):1083–7. doi:10.1038/ng.442.
8. Xiao L, Yuan Z, Jin S, Wang T, Huang S, Zeng P. Multiple-Tissue Integrative Transcriptome-Wide Association Studies Discovered New Genes Associated With Amyotrophic Lateral Sclerosis. Front Genet. 2020;11:587243. doi:10.3389/fgene.2020.587243.

9. Project Min EALSSC. Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. Eur J Hum Genet. 2018;26(10):1537–46. doi:10.1038/s41431-018-0177-4.
10. 10.1016/j.neuron.2018.02.027.
11. Iacoangeli A, Lin T, Al Khleifat A, Jones AR, Opie-Martin S, Coleman JRI, Shatunov A, Sproviero W, Williams KL, Garton F, Restuadi R, Henders AK, Mather KA, Needham M, Mathers S, Nicholson GA, Rowe DB, Henderson R, McCombe PA, Pamphlett R, Blair IP, Schultz D, Sachdev PS, Newhouse SJ, Proitsi P, Fogh I, Ngo ST, Dobson RJB, Wray NR, Steyn FJ, Al-Chalabi A. Genome-wide Meta-analysis Finds the ACSL5-ZDHHC6 Locus Is Associated with ALS and Links Weight Loss to the Disease Genetics. Cell Rep. 2020;33(4):108323. doi:10.1016/j.celrep.2020.108323.
12. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43. doi:10.1038/nature13835.
13. Du Y, Wen Y, Guo X, Hao J, Wang W, He A, Fan Q, Li P, Liu L, Liang X, Zhang F. A Genome-wide Expression Association Analysis Identifies Genes and Pathways Associated with Amyotrophic Lateral Sclerosis. Cell Mol Neurobiol. 2018;38(3):635–9. doi:10.1007/s10571-017-0512-2.
14. Park S, Kim D, Song J, Joo JWJ. (2021) An Integrative Transcriptome-Wide Analysis of Amyotrophic Lateral Sclerosis for the Identification of Potential Genetic Markers and Drug Candidates. Int J Mol Sci 22 (6). doi:10.3390/ijms22063216.
15. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11(4):e1004219. doi:10.1371/journal.pcbi.1004219.
16. Watanabe K, Umicevic Mirkov M, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. Nat Commun. 2019;10(1):3222. doi:10.1038/s41467-019-11181-1.
17. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826. doi:10.1038/s41467-017-01261-5.
18. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310–5. doi:10.1038/ng.2892.
19. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–7. doi:10.1101/gr.137323.112.
20. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL, Ren B. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. Cell Rep. 2016;17(8):2042–59. doi:10.1016/j.celrep.2016.10.061.
21. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL, Stahl EA, Huckins LM, Consortium GT, Nicolae DL, Cox NJ, Im HK. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun. 2018;9(1):1825. doi:10.1038/s41467-018-03621-1.
22. Barbeira AN, Pividori M, Zheng J, Wheeler HE, Nicolae DL, Im HK. Integrating predicted transcriptome from multiple tissues improves association detection. PLoS Genet. 2019;15(1):e1007889. doi:10.1371/journal.pgen.1007889.
23. Pasaniuc B, Zaitlen N, Shi H, Bhatia G, Gusev A, Pickrell J, Hirschhorn J, Strachan DP, Patterson N, Price AL. Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. Bioinformatics. 2014;30(20):2906–14. doi:10.1093/bioinformatics/btu416.
24. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics. 2014;30(19):2811–2. doi:10.1093/bioinformatics/btu393.

25. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 2011;12(1):1–7. doi:10.1186/1471-2105-12-35.
26. Megat S, Mora N, Sanogo J, Catanese A, Alami NO, Freischmidt A, Mingaj X, de Calbiac H, Muratet F, Dirrig-Grosch S, Dieterle S, Van Bakel N, Müller K, Sieverding K, Weishaupt J, Andersen PM, Weber M, Neuwirth C, Margelisch M, Sommacal A, van Eijk KR, Veldink JH, consortium PMAs, Lautrette G, Couratier P, Boeckers T, Ludolph AC, Roselli F, Yilmazer-Hanke D, Millecamps S, Kabashi E, Storkebaum E, Sellier C, Dupuis L. (2021) Loss of nucleoporin Nup50 is a risk factor for amyotrophic lateral sclerosis. medRxiv:2021.2008.2023.21262299. doi:10.1101/2021.08.23.21262299.
27. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383. doi:10.1371/journal.pgen.1004383.
28. Ng B, White CC, Klein HU, Sieberts SK, McCabe C, Patrick E, Xu J, Yu L, Gaiteri C, Bennett DA, Mostafavi S, De Jager PL. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat Neurosci. 2017;20(10):1418–26. doi:10.1038/nn.4632.
29. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R, Kasela S, Pervjakova N, Alvaes I, Fave M-J, Agbessi M, Christiansen M, Jansen R, Seppälä I, Tong L, Teumer A, Schramm K, Hemani G, Verlouw J, Yaghootkar H, Sönmez R, Brown A, Kukushkina V, Kalnapenkis A, Rüeger S, Porcu E, Kronberg-Guzman J, Kettunen J, Powell J, Lee B, Zhang F, Arindrarto W, Beutner F, Brugge H, Dmitreva J, Elansary M, Fairfax B, Georges M, Heijmans B, Kähönen M, Kim Y, Knight J, Kovacs P, Krohn K, Li S, Loeffler M, Marigorta U, Mei H, Momozawa Y, Müller-Nurasyid M, Nauck M, Nivard M, Penninx B, Pritchard J, Raitakari O, Rotzchke O, Slagboom E, Stehouwer CDA, Stumvoll M, Sullivan P, Hoen PACt, Thiery J, Tönjes A, van Dongen J, van Iterson M, Veldink J, Völker U, Wijmenga C, Swertz M, Andiappan A, Montgomery G, Ripatti S, Perola M, Kutalik Z, Dermitzakis E, Bergmann S, Frayling T, van Meurs J, Prokisch H, Ahsan H, Pierce B, Lehtimäki T, Boomsma D, Psaty B, Gharib S, Awadalla P, Milani L, Ouwehand W, Downes K, Stegle O, Battle A, Yang J, Visscher P, Scholz M, Gibson G, Esko T, Franke L, Consortium B, i QTLC (2018) Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. bioRxiv. doi:10.1101/447367.
30. Qi T, Wu Y, Zeng J, Zhang F, Xue A, Jiang L, Zhu Z, Kemper K, Yengo L, Zheng Z, e QC, Marioni RE, Montgomery GW, Deary IJ, Wray NR, Visscher PM, McRae AF, Yang J. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun. 2018;9(1):2282. doi:10.1038/s41467-018-04558-1.
31. 10.1126/science.aaz1776.
32. Liu B, Gloudemans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with gene expression complicate GWAS follow-up. Nat Genet. 2019;51(5):768–9. doi:10.1038/s41588-019-0404-0.
33. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7. doi:10.1038/ng.3538.
34. Van Es MA, Van Vught PW, Blauw HM, Franke L, Saris CG, Andersen PM, Van Den Bosch L, de Jong SW, van't Slot R, Birve A. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. Lancet Neurol. 2007;6(10):869–77. doi:10.1016/S1474-4422(07)70222-3.
35. Schymick JC, Scholz SW, Fung H-C, Britton A, Arepalli S, Gibbs JR, Lombardo F, Matarin M, Kasperaviciute D, Hernandez DG. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol. 2007;6(4):322–8. doi:10.1016/S1474-4422(07)70037-6.
36. Dunckley T, Huentelman MJ, Craig DW, Pearson JV, Szelinger S, Joshipura K, Halperin RF, Stamper C, Jensen KR, Letizia D, Hesterlee SE, Pestronk A, Levine T, Bertorini T, Graves MC, Mozaffar T, Jackson CE, Bosch P, McVey A, Dick A, Barohn R, Lomen-Hoerth C, Rosenfeld J, O'Connor DT, Zhang K, Crook R, Ryberg H, Hutton M, Katz J, Simpson EP, Mitsumoto H, Bowser R, Miller RG, Appel SH, Stephan DA. Whole-Genome Analysis of Sporadic Amyotrophic Lateral Sclerosis. N Engl J Med. 2007;357(8):775–88. doi:10.1056/NEJMoa070174.

37. Van Es MA, Van Vught PW, Blauw HM, Franke L, Saris CG, Van Den Bosch L, De Jong SW, De Jong V, Baas F, Van't Slot R. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. Nat Genet. 2008;40(1):29–31. doi:10.1038/ng.2007.52.
38. Cronin S, Berger S, Ding J, Schymick JC, Washecka N, Hernandez DG, Greenway MJ, Bradley DG, Traynor BJ, Hardiman O. A genome-wide association study of sporadic ALS in a homogenous Irish population. Hum Mol Genet. 2008;17(5):768–74. doi:10.1093/hmg/ddm361.
39. van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, Lemmens R, Schelhaas HJ, Groen EJ, Huisman MH. Genome-wide association study identifies 19p13. 3 (UNC13A) and 9p21. 2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat Genet. 2009;41(10):1083–7. doi:10.1038/ng.442.
40. Landers JE, Melki J, Meininger V, Glass JD, Van Den Berg LH, Van Es MA, Sapp PC, Van Vught PW, McKenna-Yasek DM, Blauw HM. (2009) Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. Proceedings of the National Academy of Sciences 106 (22):9004–9009. doi:10.1073/pnas.0812937106.
41. Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, Johnson L, Veldink JH, van Es MA, van den Berg LH. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. Lancet Neurol. 2010;9(10):986–94. doi:10.1016/S1474-4422(10)70197-6.
42. Laaksovirta H, Peuralinna T, Schymick JC, Scholz SW, Lai S-L, Myllykangas L, Sulkava R, Jansson L, Hernandez DG, Gibbs JR. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol. 2010;9(10):978–85. doi:10.1016/S1474-4422(10)70184-8.
43. Kwee LC, Liu Y, Haynes C, Gibson JR, Stone A, Schichman SA, Kamel F, Nelson LM, Topol B, Van Den Eeden SK. A high-density genome-wide association screen of sporadic ALS in US veterans. PLoS ONE. 2012;7(3):e32768. doi:10.1371/journal.pone.0032768.
44. Deng M, Wei L, Zuo X, Tian Y, Xie F, Hu P, Zhu C, Yu F, Meng Y, Wang H. Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis. Nat Genet. 2013;45(6):697–700. doi:10.1038/ng.2627.
45. Consortium A. (2013) Age of onset of amyotrophic lateral sclerosis is modulated by a locus on 1p34. 1. Neurobiology of aging 34 (1):357. e357-357. e319. doi:10.1016/j.neurobiolaging.2012.07.017.
46. Xie T, Deng L, Mei P, Zhou Y, Wang B, Zhang J, Lin J, Wei Y, Zhang X, Xu R. (2014) A genome-wide association study combining pathway analysis for typical sporadic amyotrophic lateral sclerosis in Chinese Han populations. Neurobiology of aging 35 (7):1778. e1779-1778.
e1723
. doi:10.1016/j.neurobiolaging.2014.01.014.
47. Fogh I, Ratti A, Gellera C, Lin K, Tiloca C, Moskvina V, Corrado L, Sorarù G, Cereda C, Corti S. A genome-wide association meta-analysis identifies a novel locus at 17q11. 2 associated with sporadic amyotrophic lateral sclerosis. Hum Mol Genet. 2014;23(8):2220–31. doi:10.1093/hmg/ddt587.
48. Diekstra FP, Van Deerlin VM, Van Swieten JC, Al-Chalabi A, Ludolph AC, Weishaupt JH, Hardiman O, Landers JE, Brown RH Jr, Van Es MA. C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: a genome-wide meta-analysis. Ann Neurol. 2014;76(1):120–33. doi:10.1002/ana.24198.
49. McLaughlin RL, Kenna KP, Vajda A, Bede P, Elamin M, Cronin S, Donaghy CG, Bradley DG, Hardiman O. Second-generation Irish genome-wide association study for amyotrophic lateral sclerosis. Neurobiol Aging. 2015;36(2):1221 e1227-1213. doi:10.1016/j.neurobiolaging.2014.08.030.
50. Fogh I, Lin K, Tiloca C, Rooney J, Gellera C, Diekstra FP, Ratti A, Shatunov A, Van Es MA, Proitsi P. Association of a locus in the CAMTA1 gene with survival in patients with sporadic amyotrophic lateral sclerosis. JAMA Neurol. 2016;73(7):812–20. doi:10.1001/jamaneurol.2016.1114.

51. Chen C-J, Chen C-M, Pai T-W, Chang H-T, Hwang C-S. A genome-wide association study on amyotrophic lateral sclerosis in the Taiwanese Han population. Biomark Med. 2016;10(6):597–611. doi:10.2217/bmm.15.115.
52. Benyamin B, He J, Zhao Q, Gratten J, Garton F, Leo PJ, Liu Z, Mangelsdorf M, Al-Chalabi A, Anderson L, Butler TJ, Chen L, Chen XD, Cremin K, Deng HW, Devine M, Edson J, Fifita JA, Furlong S, Han YY, Harris J, Henders AK, Jeffree RL, Jin ZB, Li Z, Li T, Li M, Lin Y, Liu X, Marshall M, McCann EP, Mowry BJ, Ngo ST, Pamphlett R, Ran S, Reutens DC, Rowe DB, Sachdev P, Shah S, Song S, Tan LJ, Tang L, van den Berg LH, van Rheenen W, Veldink JH, Wallace RH, Wheeler L, Williams KL, Wu J, Wu X, Yang J, Yue W, Zhang ZH, Zhang D, Noakes PG, Blair IP, Henderson RD, McCombe PA, Visscher PM, Xu H, Bartlett PF, Brown MA, Wray NR, Fan D. Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis. Nat Commun. 2017;8(1):611. doi:10.1038/s41467-017-00471-1.
53. Wei L, Tian Y, Chen Y, Wei Q, Chen F, Cao B, Wu Y, Zhao B, Chen X, Xie C. (2019) Identification of TYW3/CRYZ and FGD4 as susceptibility genes for amyotrophic lateral sclerosis. Neurol Genet 5 (6). doi:10.1212/NXG.0000000000000375.
54. Dekker AM, Diekstra FP, Pulit SL, Tazelaar GH, van der Spek RA, van Rheenen W, van Eijk KR, Calvo A, Brunetti M, Van Damme P. Exome array analysis of rare and low frequency variants in amyotrophic lateral sclerosis. Sci Rep. 2019;9(1):1–8. doi:10.1038/s41598-019-42091-3.
55. Nakamura R, Misawa K, Tohnai G, Nakatochi M, Furuhashi S, Atsuta N, Hayashi N, Yokoi D, Watanabe H, Watanabe H. A multi-ethnic meta-analysis identifies novel genes, including ACSL5, associated with amyotrophic lateral sclerosis. Commun biology. 2020;3(1):1–9. doi:10.1038/s42003-020-01251-2.
56. Nakamura R, Misawa K, Tohnai G, Nakatochi M, Furuhashi S, Atsuta N, Hayashi N, Yokoi D, Watanabe H, Watanabe H, Katsuno M, Izumi Y, Kanai K, Hattori N, Morita M, Taniguchi A, Kano O, Oda M, Shibuya K, Kuwabara S, Suzuki N, Aoki M, Ohta Y, Yamashita T, Abe K, Hashimoto R, Aiba I, Okamoto K, Mizoguchi K, Hasegawa K, Okada Y, Ishihara T, Onodera O, Nakashima K, Kaji R, Kamatani Y, Ikegawa S, Momozawa Y, Kubo M, Ishida N, Minegishi N, Nagasaki M, Sobue G. A multi-ethnic meta-analysis identifies novel genes, including ACSL5, associated with amyotrophic lateral sclerosis. Commun Biol. 2020;3(1):526. doi:10.1038/s42003-020-01251-2.
57. 10.1038/s41588-021-00973-1.
58. Loeffler JP, Picchiarelli G, Dupuis L, Gonzalez De Aguilar JL. The Role of Skeletal Muscle in Amyotrophic Lateral Sclerosis. Brain Pathol. 2016;26(2):227–36. doi:10.1111/bpa.12350.
59. Badu-Mensah A, Guo X, McAleer CW, Rumsey JW, Hickman JJ. Functional skeletal muscle model derived from SOD1-mutant ALS patient iPSCs recapitulates hallmarks of disease progression. Sci Rep. 2020;10(1):14302. doi:10.1038/s41598-020-70510-3.
60. Saez-Atienzar S, Bandres-Ciga S, Langston RG, Kim JJ, Choi SW, Reynolds RH, Abramzon Y, Dewan R, Ahmed S, Landers JE, Chia R, Ryten M, Cookson MR, Nalls MA, Chiò A, Traynor BJ. Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types. Sci Adv. 2021;7(3):eabd9036. doi:10.1126/sciadv.abd9036.
61. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. Cell Rep. 2016;15(9):2038–49. doi:10.1016/j.celrep.2016.04.085.
62. Saris CG, Horvath S, van Vught PW, van Es MA, Blauw HM, Fuller TF, Langfelder P, DeYoung J, Wokke JH, Veldink JH, van den Berg LH, Ophoff RA. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. BMC Genomics. 2009;10:405. doi:10.1186/1471-2164-10-405.
63. Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, Fumis L, Hayhurst J, Buniello A, Karim MA, Wright D, Hercules A, Papa E, Fauman EB, Barrett JC, Todd JA, Ochoa D, Dunham I, Ghoussaini M. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 2021;53(11):1527–33. doi:10.1038/s41588-021-00945-5.
64. Iacoangeli A, Fogh I, Selvackadunco S, Topp SD, Shatunov A, van Rheenen W, Al-Khleifat A, Opie-Martin S, Ratti A, Calvo A, Consortium UKBE, Van Damme P, Robberecht W, Chio A, Dobson RJ, Hardiman O, Shaw CE, van den Berg LH, Andersen PM, Smith BN, Silani V, Veldink JH, Breen G, Troakes C, Al-Chalabi A, Jones AR. SCFD1 expression

quantitative trait loci in amyotrophic lateral sclerosis are differentially expressed. Brain Commun. 2021;3(4):fcab236. doi:10.1093/braincomms/fcab236.

65. Perrot V, Vazquez-Prado J, Gutkind JS. Plexin B regulates Rho through the guanine nucleotide exchange factors leukemia-associated Rho GEF (LARG) and PDZ-RhoGEF. J Biol Chem. 2002;277(45):43115–20. doi:10.1074/jbc.M206005200.
66. Yu W, Goncalves KA, Li S, Kishikawa H, Sun G, Yang H, Vanli N, Wu Y, Jiang Y, Hu MG, Friedel RH, Hu GF. Plexin-B2 Mediates Physiologic and Pathologic Functions of Angiogenin. Cell. 2017;171(4):849–64 e825. doi:10.1016/j.cell.2017.10.005.
67. Subramanian V, Crabtree B, Acharya KR. Human angiogenin is a neuroprotective factor and amyotrophic lateral sclerosis associated angiogenin variants affect neurite extension/pathfinding and survival of motor neurons. Hum Mol Genet. 2008;17(1):130–49. doi:10.1093/hmg/ddm290.
68. Koziol LF, Budding D, Andreasen N, D'Arrigo S, Bulgheroni S, Imamizu H, Ito M, Manto M, Marvel C, Parker K, Pezzulo G, Ramnani N, Riva D, Schmahmann J, Vandervert L, Yamazaki T. Consensus paper: the cerebellum's role in movement and cognition. Cerebellum. 2014;13(1):151–77. doi:10.1007/s12311-013-0511-x.
69. Qiu T, Zhang Y, Tang X, Liu X, Wang Y, Zhou C, Luo C, Zhang J. Precentral degeneration and cerebellar compensation in amyotrophic lateral sclerosis: A multimodal MRI analysis. Hum Brain Mapp. 2019;40(12):3464–74. doi:10.1002/hbm.24609.
70. 10.1136/jnnp-2014-308172.
71. Clarke BE, Patani R. The microglial component of amyotrophic lateral sclerosis. Brain. 2020;143(12):3526–39. doi:10.1093/brain/awaa309.
72. Do-Ha D, Buskila Y, Ooi L. Impairments in Motor Neurons, Interneurons and Astrocytes Contribute to Hyperexcitability in ALS: Underlying Mechanisms and Paths to Therapy. Mol Neurobiol. 2018;55(2):1410–8. doi:10.1007/s12035-017-0392-y.
73. Foerster BR, Pomper MG, Callaghan BC, Petrou M, Edden RA, Mohamed MA, Welsh RC, Carlos RC, Barker PB, Feldman EL. An imbalance between excitatory and inhibitory neurotransmitters in amyotrophic lateral sclerosis revealed by use of 3-T proton magnetic resonance spectroscopy. JAMA Neurol. 2013;70(8):1009–16. doi:10.1001/jamaneurol.2013.234.
74. Ramirez-Jarquin UN, Tapia R. Excitatory and Inhibitory Neuronal Circuits in the Spinal Cord and Their Role in the Control of Motor Neuron Function and Degeneration. ACS Chem Neurosci. 2018;9(2):211–6. doi:10.1021/acschemneuro.7b00503.
75. Turner MR, Kiernan MC. Does interneuronal dysfunction contribute to neurodegeneration in amyotrophic lateral sclerosis? Amyotroph Lateral Scler. 2012;13(3):245–50. doi:10.3109/17482968.2011.636050.
76. Barbeira AN, Bonazzola R, Gamazon ER, Liang Y, Park Y, Kim-Hellmuth S, Wang G, Jiang Z, Zhou D, Hormozdiari F, Liu B, Rao A, Hamel AR, Pividori MD, Aguet F, Group GTGW, Bastarache L, Jordan DM, Verbanck M, Do R, Consortium GT, Stephens M, Ardlie K, McCarthy M, Montgomery SB, Segre AV, Brown CD, Lappalainen T, Wen X, Im HK. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. Genome Biol. 2021;22(1):49. doi:10.1186/s13059-020-02252-4.
77. Droppelmann CA, Campos-Melo D, Volkening K, Strong MJ. The emerging role of guanine nucleotide exchange factors in ALS and other neurodegenerative diseases. Front Cell Neurosci. 2014;8:282. doi:10.3389/fncel.2014.00282.
78. Pellecchia MT, Pivonello R, Monsurro MR, Trojsi F, Longo K, Piccirillo G, Pivonello C, Rocco M, Di Somma C, Colao A, Tedeschi G, Barone P. The GH-IGF system in amyotrophic lateral sclerosis: correlations between pituitary GH secretion capacity, insulin-like growth factors and clinical features. Eur J Neurol. 2010;17(5):666–71. doi:10.1111/j.1468-1331.2009.02896.x.

79. Dedeene L, Van Schoor E, Ospitalieri S, Ronisz A, Weishaupt JH, Otto M, Ludolph AC, Scheuerle A, Vandenberghe R, Van Damme P, Poesen K, Thal DR. Dipeptide repeat protein and TDP-43 pathology along the hypothalamic-pituitary axis in C9orf72 and non-C9orf72 ALS and FTLD-TDP cases. Acta Neuropathol. 2020;140(5):777–81. doi:10.1007/s00401-020-02216-9.
80. Querin G, Bede P, El Mendili MM, Li M, Pelegrini-Issac M, Rinaldi D, Catala M, Saracino D, Salachas F, Camuzat A, Marchand-Pauvert V, Cohen-Adad J, Colliot O, Le Ber I, Pradat PF, Predict to Prevent Frontotemporal Lobar D, Amyotrophic Lateral Sclerosis Study G (2019) Presymptomatic spinal cord pathology in c9orf72 mutation carriers: A longitudinal neuroimaging study. Ann Neurol 86 (2):158–167. doi:10.1002/ana.25520.
81. van der Burgh HK, Westeneng HJ, Meier JM, van Es MA, Veldink JH, Hendrikse J, van den Heuvel MP, van den Berg LH. Cross-sectional and longitudinal assessment of the upper cervical spinal cord in motor neuron disease. Neuroimage Clin. 2019;24:101984. doi:10.1016/j.nicl.2019.101984.
82. Hou N, Yang Y, Scott IC, Lou X. The Sec domain protein Scfd1 facilitates trafficking of ECM components during chondrogenesis. Dev Biol. 2017;421(1):8–15. doi:10.1016/j.ydbio.2016.11.010.
83. Dadon-Nachum M, Melamed E, Offen D. The "dying-back" phenomenon of motor neurons in ALS. J Mol Neurosci. 2011;43(3):470–7. doi:10.1007/s12031-010-9467-1.
84. Braak H, Brettschneider J, Ludolph AC, Lee VM, Trojanowski JQ, Del Tredici K. Amyotrophic lateral sclerosis–a model of corticofugal axonal spread. Nat Rev Neurol. 2013;9(12):708–14. doi:10.1038/nrneurol.2013.221.
85. Cheng AJ, Allodi I, Chaillou T, Schlittler M, Ivarsson N, Lanner JT, Thams S, Hedlund E, Andersson DC. Intact single muscle fibres from SOD1(G93A) amyotrophic lateral sclerosis mice display preserved specific force, fatigue resistance and training-like adaptations. J Physiol. 2019;597(12):3133–46. doi:10.1113/JP277456.
86. Steyn FJ, Li R, Kirk SE, Tefera TW, Xie TY, Tracey TJ, Kelk D, Wimberger E, Garton FC, Roberts L, Chapman SE, Coombes JS, Leevy WM, Ferri A, Valle C, Rene F, Loeffler JP, McCombe PA, Henderson RD, Ngo ST. Altered skeletal muscle glucose-fatty acid flux in amyotrophic lateral sclerosis. Brain Commun. 2020;2(2):fcaa154. doi:10.1093/braincomms/fcaa154.
87. Marat AL, Dokainish H, McPherson PS. DENN domain proteins: regulators of Rab GTPases. J Biol Chem. 2011;286(16):13791–800. doi:10.1074/jbc.R110.217067.
88. Shah PR. Complex genetic approaches to neurodegenerative diseases. University of London; 2007.
89. Al-Ramahi I, Giridharan SSP, Chen YC, Patnaik S, Safren N, Hasegawa J, de Haro M, Wagner Gee AK, Titus SA, Jeong H, Clarke J, Krainc D, Zheng W, Irvine RF, Barmada S, Ferrer M, Southall N, Weisman LS, Botas J, Marugan JJ. (2017) Inhibition of PIP4Kgamma ameliorates the pathological effects of mutant huntingtin protein. Elife 6. doi:10.7554/eLife.29123.
90. Rappaport N, Twik M, Plaschkes I, Nudel R, Iny Stein T, Levitt J, Gershoni M, Morrey CP, Safran M, Lancet D. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic Acids Res. 2017;45(D1):D877–87. doi:10.1093/nar/gkw1012.
91. Farhan SMK, Howrigan DP, Abbott LE, Klim JR, Topp SD, Byrnes AE, Churchhouse C, Phatnani H, Smith BN, Rampersaud E, Wu G, Wuu J, Shatunov A, Iacoangeli A, Al Khleifat A, Mordes DA, Ghosh S, Consortium A, Consortium F, Project Min EC, Consortium CR, Eggan K, Rademakers R, McCauley JL, Schule R, Zuchner S, Benatar M, Taylor JP, Nalls M, Gotkine M, Shaw PJ, Morrison KE, Al-Chalabi A, Traynor B, Shaw CE, Goldstein DB, Harms MB, Daly MJ, Neale BM. Exome sequencing in amyotrophic lateral sclerosis implicates a novel gene, DNAJC7, encoding a heat-shock protein. Nat Neurosci. 2019;22(12):1966–74. doi:10.1038/s41593-019-0530-0.
92. Zhang Y, Yang HT, Kadash-Edmondson K, Pan Y, Pan Z, Davidson BL, Xing Y. Regional Variation of Splicing QTLs in Human Brain. Am J Hum Genet. 2020;107(2):196–210. doi:10.1016/j.ajhg.2020.06.002.
93. Da Cruz S, Cleveland DW. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. Curr Opin Neurobiol. 2011;21(6):904–19. doi:10.1016/j.conb.2011.05.029.

94. Buée. L, Bussière. T, Buée-Scherrer V, Hof DA. PR. Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. Brain Res Brain Res Rev. 2000;33(1):95–130. doi:10.1016/s0165-0173(00)00019-9.
95. Zhao G, Oztan A, Ye Y, Schwarz TL. Kinetochore Proteins Have a Post-Mitotic Function in Neurodevelopment. Dev Cell. 2019;48(6):873–82 e874. doi:10.1016/j.devcel.2019.02.003.
96. Kerkenberg N, Wachsmuth L, Zhang M, Schettler C, Ponimaskin E, Faber C, Baune BT, Zhang W, Hohoff C. Brain microstructural changes in mice persist in adulthood and are modulated by the palmitoyl acyltransferase ZDHHC7. Eur J Neurosci. 2021;54(6):5951–67. doi:10.1111/ejn.15415.
97. Montibeller L, de Belleroche J. Amyotrophic lateral sclerosis (ALS) and Alzheimer's disease (AD) are characterised by differential activation of ER stress pathways: focus on UPR target genes. Cell Stress Chaperones. 2018;23(5):897–912. doi:10.1007/s12192-018-0897-y.
98. Figueroa-Romero C, Hur J, Bender DE, Delaney CE, Cataldo MD, Smith AL, Yung R, Ruden DM, Callaghan BC, Feldman EL. Identification of epigenetically altered genes in sporadic amyotrophic lateral sclerosis. PLoS ONE. 2012;7(12):e52672. doi:10.1371/journal.pone.0052672.
99. Liu D, Liu C, Li J, Azadzoi K, Yang Y, Fei Z, Dou K, Kowall NW, Choi HP, Vieira F, Yang JH. Proteomic analysis reveals differentially regulated protein acetylation in human amyotrophic lateral sclerosis spinal cord. PLoS ONE. 2013;8(12):e80779. doi:10.1371/journal.pone.0080779.
100. Kikuchi H, Yamada T, Furuya H, Doh-ura K, Ohyagi Y, Iwaki T, Kira J. Involvement of cathepsin B in the motor neuron degeneration of amyotrophic lateral sclerosis. Acta Neuropathol. 2003;105(5):462–8. doi:10.1007/s00401-002-0667-9.
101. Hunter M, Spiller KJ, Dominique MA, Xu H, Hunter FW, Fang TC, Canter RG, Roberts CJ, Ransohoff RM, Trojanowski JQ, Lee VM. Microglial transcriptome analysis in the rNLS8 mouse model of TDP-43 proteinopathy reveals discrete expression profiles associated with neurodegenerative progression and recovery. Acta Neuropathol Commun. 2021;9(1):140. doi:10.1186/s40478-021-01239-x.
102. Diana A, Pillai R, Bongioanni P, O'Keeffe AG, Miller RG, Moore DH. Gamma aminobutyric acid (GABA) modulators for amyotrophic lateral sclerosis/motor neuron disease. Cochrane Database Syst Rev. 2017;1:CD006049. doi:10.1002/14651858.CD006049.pub2.
103. Teoh J, Subramanian N, Pero ME, Bartolini F, Amador A, Kanber A, Williams D, Petri S, Yang M, Allen AS, Beal J, Haut SR, Frankel WN. Arfgef1 haploinsufficiency in mice alters neuronal endosome composition and decreases membrane surface postsynaptic GABAA receptors. Neurobiol Dis. 2020;134:104632. doi:10.1016/j.nbd.2019.104632.

Tables

Table 1. Genes mediating the genetic associations with ALS in 6 loci from SMR & HEIDI analyses

Locus	Gene	SNP	FDR	HEIDI	Tissue	Database
3q24	***PLOD2***	rs149615181	0.04	0.69	**Muscle_Skeletal**	GTEx v8
9p21.2	*C9orf72*	rs2453565	2.00×10^{-5}	0.14	**Pituitary**	GTEx v8
		rs700795	0.02	0.19	**Brain_Spinal_cord_cervical_c-1**	GTEx v8
10q22.2	***NDST2***	rs11000785	0.05	0.07	Blood	eQTLGen
14q12	*SCFD1*	rs7144204	$3.6´10^{-3}$	0.10	Blood	eQTLGen
		rs448175	0.01	0.35	Blood	GTEx v8
		rs229152	0.03	0.45	Brain_Cerebellum	GTEx v8
		rs229243	0.04	0.31	**Muscle_Skeletal**	GTEx v8
		rs2070339	0.03	0.22	Multiple brain regions	Brain-eMeta
17q12	*GGNBP2*	rs11650008	0.03	0.04	Blood	eQTLGen
	MYO19	rs7222903	0.04	0.60	Blood	eQTLGen
	DYNLL2	rs2877858	0.04	0.16	Blood	eQTLGen
	ZNHIT3	rs4796224	0.05	0.50	Blood	eQTLGen
22q13.33	***PLXNB2***	rs62241220	0.02	0.77	Blood	eQTLGen

All genes with FDR<0.05 & HEIDI>0.01 are shown. Loci that have not been identified in previous GWASs or post-GWASs are indicated in bold. Genes that have not been reported in previous SMR studies are indicated in bold. Tissues that have not been reported in previous SMR studies are indicated in bold.

Table 2. Integration of TWAS, COLOC and SMR results.

Locus	Gene	Tissue	TWAS	COLOC (PP4/PP3)	SMR (HEIDI)	Overall evidence
5q33.1	*TNIP1*	Blood	2.2×10^{-3}	0.91 (9.70)	3.6×10^{-3} (6.7×10^{-4})	**
9p21.2	*C9orf72*	Brain_Spinal_cord_cervical_c-1	7.00×10^{-13}	0.81 (5.10)	1.5×10^{-2} (0.19)	***
	C9orf72	Pituitary	5.00×10^{-14}	0.99 (82.50)	2.00×10^{-5} (0.14)	***
10q25.2	*ACSL5*	Blood	2.7×10^{-2}	0.82 (5.10)	-	**
10q22.2	*NDST2*	Blood	3.3×10^{-2}	-	4.9×10^{-2} (0.07)	**
14q12	*SCFD1*	Brain_Cerebellum	1.7×10^{-3}	0.92 (11.81)	3.1×10^{-2} (0.45)	***
	SCFD1	Muscle_Skeletal	3.7×10^{-2}	0.91 (9.85)	3.8×10^{-2} (0.31)	***
	SCFD1	Blood	1.5×10^{-3}	0.92 (12.00)	1.1×10^{-2} (0.35)	***
	SCFD1	Brain_Anterior_cingulate_cortex_BA24	3.3×10^{-3}	0.92 (11.85)	-	**
	SCFD1	Brain_Cerebellar_Hemisphere	1.1×10^{-3}	0.92 (11.35)	-	**
	SCFD1	Brain_Cortex	2.4×10^{-3}	0.91 (10.75)	-	**
	SCFD1	Brain_Frontal_Cortex_BA9	1.3×10^{-3}	0.92 (11.80)	-	**
	SCFD1	Brain_Nucleus_accumbens_basal_ganglia	2.7×10^{-3}	0.77 (6.01)	-	**
17q12	*MYO19*	Blood	2.2×10^{-3}	-	4.0×10^{-2} (0.6)	**
	GGNBP2	Blood	1.3×10^{-2}	-	2.5×10^{-3} (0.04)	**
	ZNHIT3	Blood	3.5×10^{-2}	-	4.9×10^{-2} (0.5)	**

Blood refers to whole blood from GTEx v8 or eQTLGen depending on which is more significant. The TWAS column indicated the *p-value* (FDR). The COLOC column indicated PP4 and PP4/PP3. The SMR column indicated the *p-value* (FDR) of the SMR test and the HEIDI test. The number of asterisks in the overall evidence column represents the number of significant results from three different analyses.

Figures

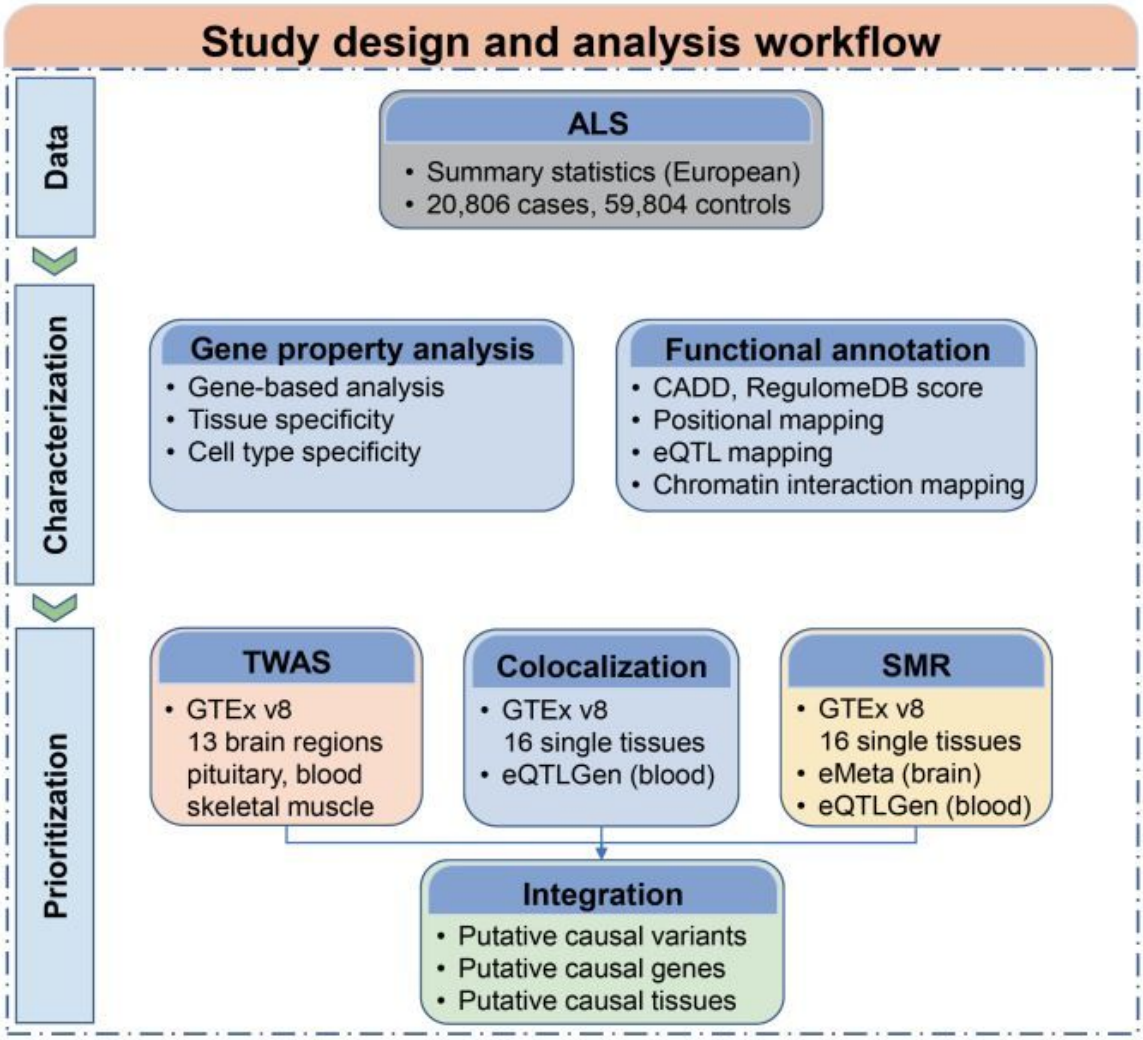

Figure 1

Schematic overview of the study design. Data, data collection and preprocessing. Characterization, Characterizing ALS risk loci, including gene property analysis and functional annotation. Prioritization, putative causal genes prioritization and summarizing the evidence.

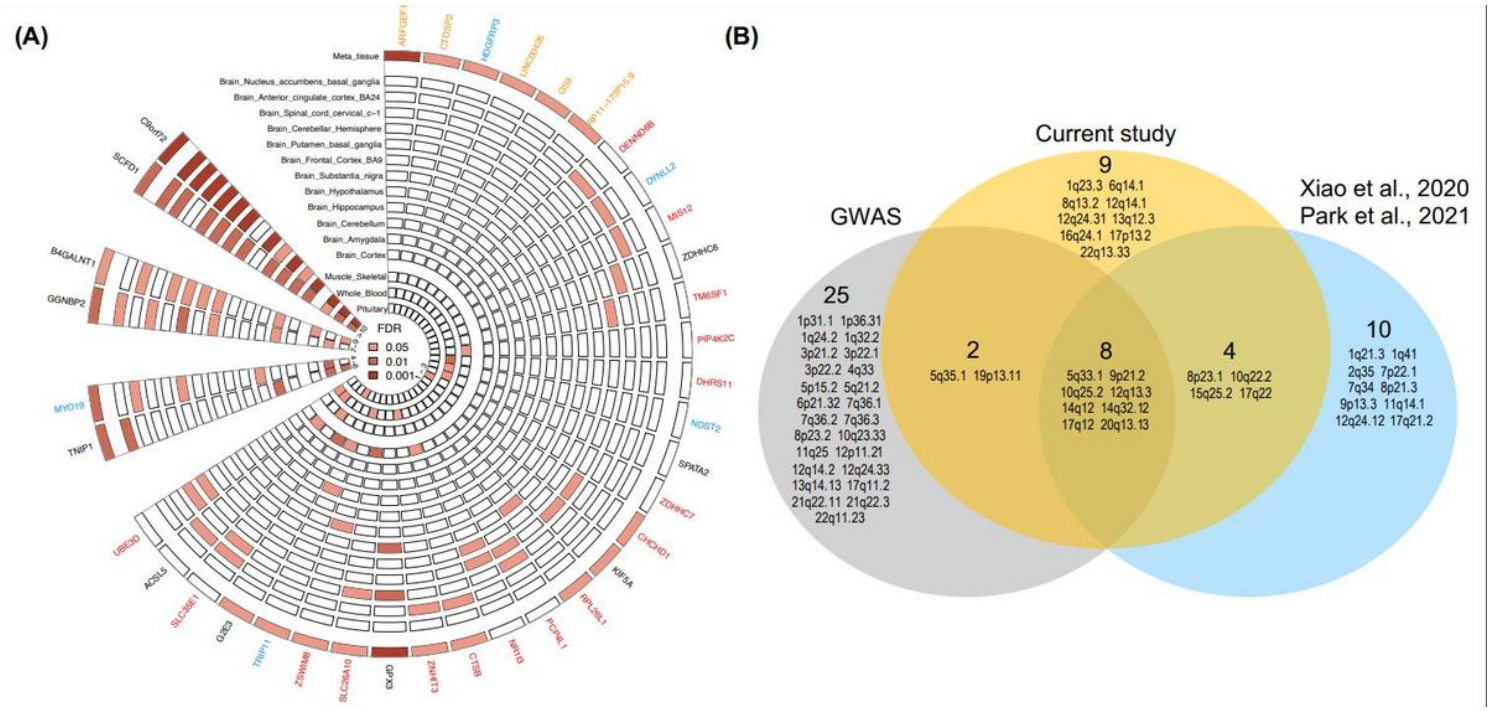

Figure 2

Identification of genes associated with ALS by TWAS analysis and comparison with other studies. (A) Genes associated with ALS were identified by S-PredictXcan and S-MetaXcan across tissues (FDR<0.05). Meta_tissue represented the joint effect of gene expression from different tissues using S-MetaXcan. The circles represented different tissues from outside to inside, where the brain tissues were clustered together, while each sector expressed a different gene. The strength of color of each cell indicated the significance of the association of genes (sectors) with ALS in different tissues (circles). Among these genes, red and orange respectively indicated genes newly discovered from S-PrediXcan and S-MultiXcan, blue indicated the 5 replicated TWAS-discovered genes recently, and black indicated genes previously reported by GWAS. Brain tissues are grouped, and grouped genes with similar TWAS patterns (the number of significant signals labeled in the innermost ring). (B) Venn diagram displayed loci identified in different studies. Gray circle represented published GWASs, blue circle represented 2 recent TWASs from Xiao et al., 2020[8] and Park et al., 2021[14], and orange circle was our current TWAS. The number represents the number of loci in each region.

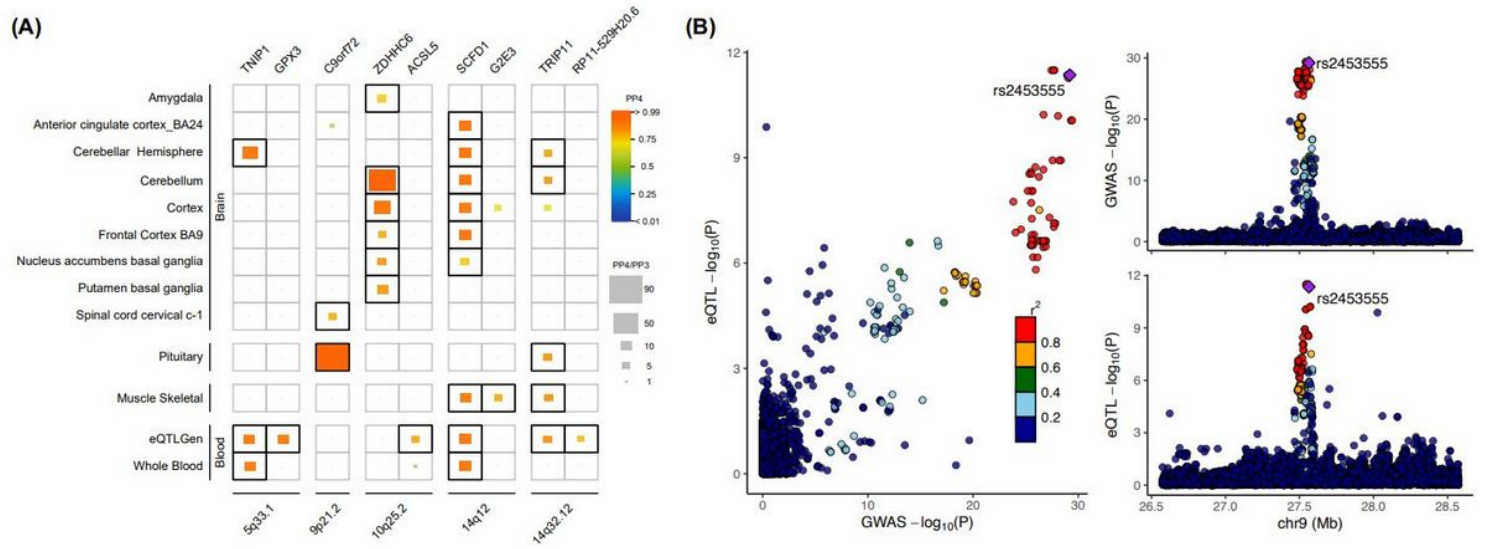

Figure 3

Colocalization of genetic ALS association and eQTL in different tissues. (A) Heatmap of significant colocalization (PP4≥0.75, PP3+PP4≥0.9 and PP4/PP3≥3) in a total of 16 tissues analyzed. The horizontal axis represented genes under different cytobands, and the vertical axis represented different tissues, where brain regions and the blood tissues were lined up together, respectively. The cell color indicated the posterior probability of colocalization with orange indicating larger values, and the size of the inside squares was proportional to the PP4/PP3 ratio. (B) Illustration of the *C9orf72* locus in pituitary (PP4≥0.99, PP3+PP4≥0.99, PP4/PP3=82.5). Each dot represented a genetic variant with the candidate causal variant, rs2453555, shown as a purple diamond. The color of other variants indicated their linkage disequilibrium (r2) based on 1000 Genomes Project European reference panel with the purple diamond from blue to red. The left panel showed - log10 *p* values for SNP associations with ALS were on the x-axes, and - log10 *p* values for associations with gene expression levels were on the y-axes. The right panel illustrated genomic positions based on GRCh37 are on the x-axes and - log10 *p* values of ALS GWAS (upper panel) and - log10 *p* values of gene expression at *C9orf72* in pituitary gland (below panel).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.pdf
- SupplementaryTables.xlsx