



Causal Inference of Genetic Variants and Genes in Amyotrophic Lateral Sclerosis

Siyu Pan^{1,2}, Xinxuan Liu^{1,3}, Tianzi Liu¹, Zhongming Zhao⁴, Yulin Dai⁴, Yin-Ying Wang¹, Peilin Jia^{1*} and Fan Liu^{1,3*}

¹CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation, Beijing, China, ²College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, ³School of Future Technology, University of Chinese Academy of Sciences, Beijing, China, ⁴Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States

OPEN ACCESS

Edited by:

Xiao Chang,
Children's Hospital of Philadelphia,
United States

Reviewed by:

Jie Ping,
Vanderbilt University Medical Center,
United States

Yao Yu,

University of Texas MD Anderson
Cancer Center, United States

*Correspondence:

Fan Liu
liufan@big.ac.cn
Peilin Jia
pjia@big.ac.cn

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 10 April 2022

Accepted: 16 May 2022

Published: 22 June 2022

Citation:

Pan S, Liu X, Liu T, Zhao Z, Dai Y,
Wang Y-Y, Jia P and Liu F (2022)
Causal Inference of Genetic Variants
and Genes in Amyotrophic
Lateral Sclerosis.
Front. Genet. 13:917142.
doi: 10.3389/fgene.2022.917142

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, the exact identities of causal variants, genes, cell types, tissues, and their functional roles in the development of ALS remain largely unknown. Here, we reported a comprehensive post-GWAS analysis of the recent large ALS GWAS ($n = 80,610$), including functional mapping and annotation (FUMA), transcriptome-wide association study (TWAS), colocalization (COLOC), and summary data-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 three loci (9p21.2, 12q13.3, and 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 proposed 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. Together, these results advance our understanding of the biological etiology of ALS, feed into new therapies, and provide a guide for subsequent functional experiments.

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

INTRODUCTION

Amyotrophic lateral sclerosis (ALS, OMIM#105400) is a progressive neurological degenerative disorder without effective treatment affecting 1 in 400 individuals worldwide (van Rheen et al., 2016). With the fast pace of global aging, ALS is anticipated to reach 380,000 cases globally by 2040 (Logroscino and Piccininni, 2019). The heritability of ALS has been estimated at around 0.61 (95% CI: 0.38–0.78) in a twin study (Al-Chalabi et al., 2010). Genome-wide association studies (GWASs, **Supplementary Table S1**) have identified more than 35 genetic loci associated with risk of ALS (van

Es et al., 2009; van Rheenen et al., 2016; Nicolas et al., 2018; Project MinE ALS Sequencing Consortium, 2018; Zhou et al., 2018; Iacoangeli et al., 2020; Xiao et al., 2020), 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 noncoding regions with largely unknown functions (Farh et al., 2015). Recently, three post-GWASs (Du et al., 2018; Xiao et al., 2020; Park et al., 2021) proposed lists of genes with high probabilities of causality, providing a better understanding of the genetic basis of the pathogenesis of ALS. However, the exact identities of causal variants, genes, cell types, and tissues remain largely unknown, leaving alone the complex causal relationships between them.

Here, we report a comprehensive functional characterization of the susceptibility loci identified in the large ALS GWAS ($n = 80,610$) (Nicolas et al., 2018) using functional mapping and annotation (FUMA). Furthermore, 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

ALS GWAS summary data reported by Nicolas et al. (2018) were downloaded from ALS Variant Server (<http://als.umassmed.edu>). This dataset incorporated several previous cohorts of ALS for meta-analysis, mainly including previous 20,806 cases and 59,804 controls of European ancestry. We harmonized GWAS summary statistics to the 1,000 reference genomes. After removing SNPs with low MAF (<0.01), our post-GWAS analysis included a total of 9,657,890 SNPs.

Gene-Based Analysis, Tissue Specificity, and Cell Type

We performed a gene-level enrichment analysis using the MAGMA in FUMA platform v1.3.6a (Watanabe et al., 2017). MAGMA gene-property analysis was performed to assess relationships between tissue-specific or cell-specific gene expression profiles and ALS-gene associations (Watanabe et al., 2019). In cell-specific analysis, a 3-step workflow was performed with scRNA-seq datasets, 1) cell type analysis for each dataset was carried out, and significant cell types were retained for the next step; 2) step-wise conditional analyses to identify independent cell types within each dataset; and 3) cross-datasets conditional analysis was performed to examine the extent of similar association signals from significant cell types retained from the second step. For scRNA-seq datasets, PsychENCODE and DropViz were available for analysis.

Functional Mapping and Annotation

The bioinformatic functional analysis was performed to investigate the functional relevance for ALS through FUMA (Watanabe et al., 2017), using the following toolsets or

datasets, including ANNOVAR, CADD, RegulomeDB, eQTLGen, GTEx v8, and Hi-C data. Independent significant SNPs, lead SNPs, and genomic risk loci were defined to use the default parameters in FUMA. A CADD score with a threshold of 12.37 for a variant was considered deleterious (Kircher et al., 2014). The lower a RegulomeDB score, the more likely it is to be a regulatory element. Mapping SNPs to genes used 3 options including positional mapping, eQTL mapping, and chromatin interaction mapping. We restricted to eQTLGen, whole blood, and 13 brain tissues from GTEx v8 in eQTL mapping. Only significant SNP-gene pairs with FDR correction ($p < 0.05$) were identified to map genes. Chromatin interaction mapping was restricted to PsychENCODE, enhancer-promoter (EP) correlations from FANTOM5, adult cortex, and fetal cortex. The significance threshold of interaction was set at $FDR < 1 \times 10^{-6}$.

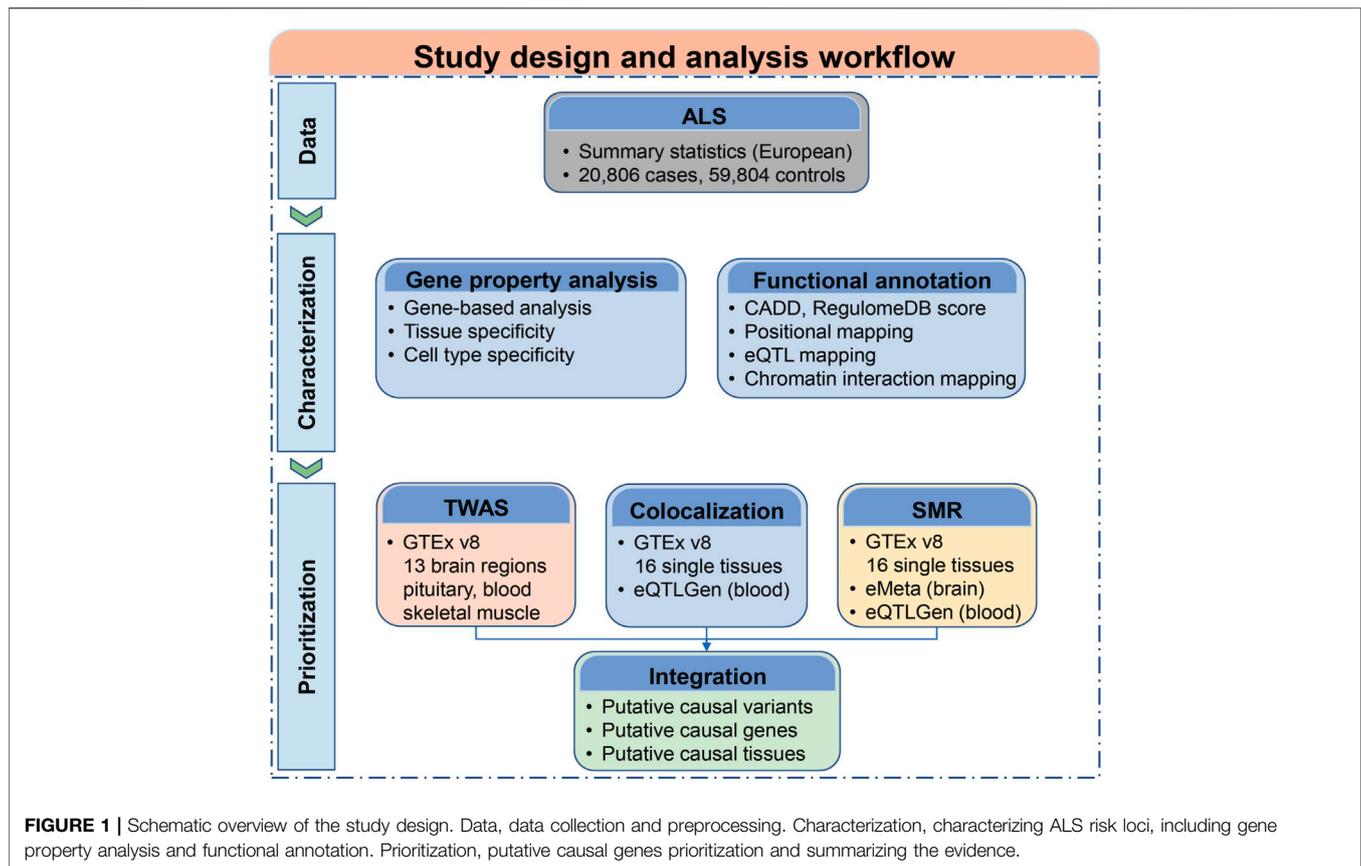
Transcriptome-Wide Association Study

We applied S-PrediXcan (Barbeira et al., 2018) to test the association between the predicted expression levels of 13 brain tissues, skeletal muscle, pituitary, and whole blood from GTEx v8 and ALS, respectively. S-MultiXcan (Barbeira et al., 2019) integrated multiple tissue panels to improve the power to detect ALS-associated genes. Prediction models trained on GTEx v8 were obtained on the PredictDB website (<http://predictdb.org>). Genes with association $FDR < 0.05$ were considered to be significantly associated with ALS. A circus plot for multiple TWASs was created by the R package “circlize” (Gu et al., 2014).

We compared the identified loci with the published TWASs and GWASs (**Supplementary Table S1**, p threshold $< 5 \times 10^{-8}$), which was visualized with the R package “VennDiagram” (Chen and Boutros, 2011). However, the latest TWAS (Megat et al., 2021) was unavailable in the medRxiv and not included in the comparison.

Colocalization Analysis

We applied COLOC (Giambartolomei et al., 2014) to assess the probability of the same variant being responsible for ALS risk and gene expression. Default conservative priors, $p_1 (1 \times 10^{-4})$, $p_2 (1 \times 10^{-4})$, and $p_{12} (1 \times 10^{-5})$ for a causal SNP in either ALS or gene expression and a shared causal SNP. We assessed the posterior probability of colocalization between ALS and Brain-eMeta (Ng et al., 2017), 13 brain tissues, skeletal muscle, pituitary, and whole blood from GTEx v8, eQTLGen, respectively. For three eQTL datasets, cis-association analyses were defined for SNPs within 1 MB of the transcription start site or the center of the probe (Qi et al., 2018; Vösa et al., 2021; Aguet et al., 2020). A detailed summary of eQTL data is in **Supplementary Table S2**. The regions in colocalization assessment had at least one SNP with a p value $< 1 \times 10^{-6}$ from GWAS that had a p value $< 1 \times 10^{-4}$ in the eQTL dataset. We applied suggested combination cutoffs ($PP4 \geq 0.75$, $PP3 + PP4 \geq 0.9$, and $PP4/PP3 \geq 3$) as powerful evidence supporting a causal role for the gene to be a mediator of ALS risk. The regional association plots were generated by R package “LocusCompareR” (Liu et al., 2019).



Summary Data-Based Mendelian Randomization Analysis and HEIDI Test

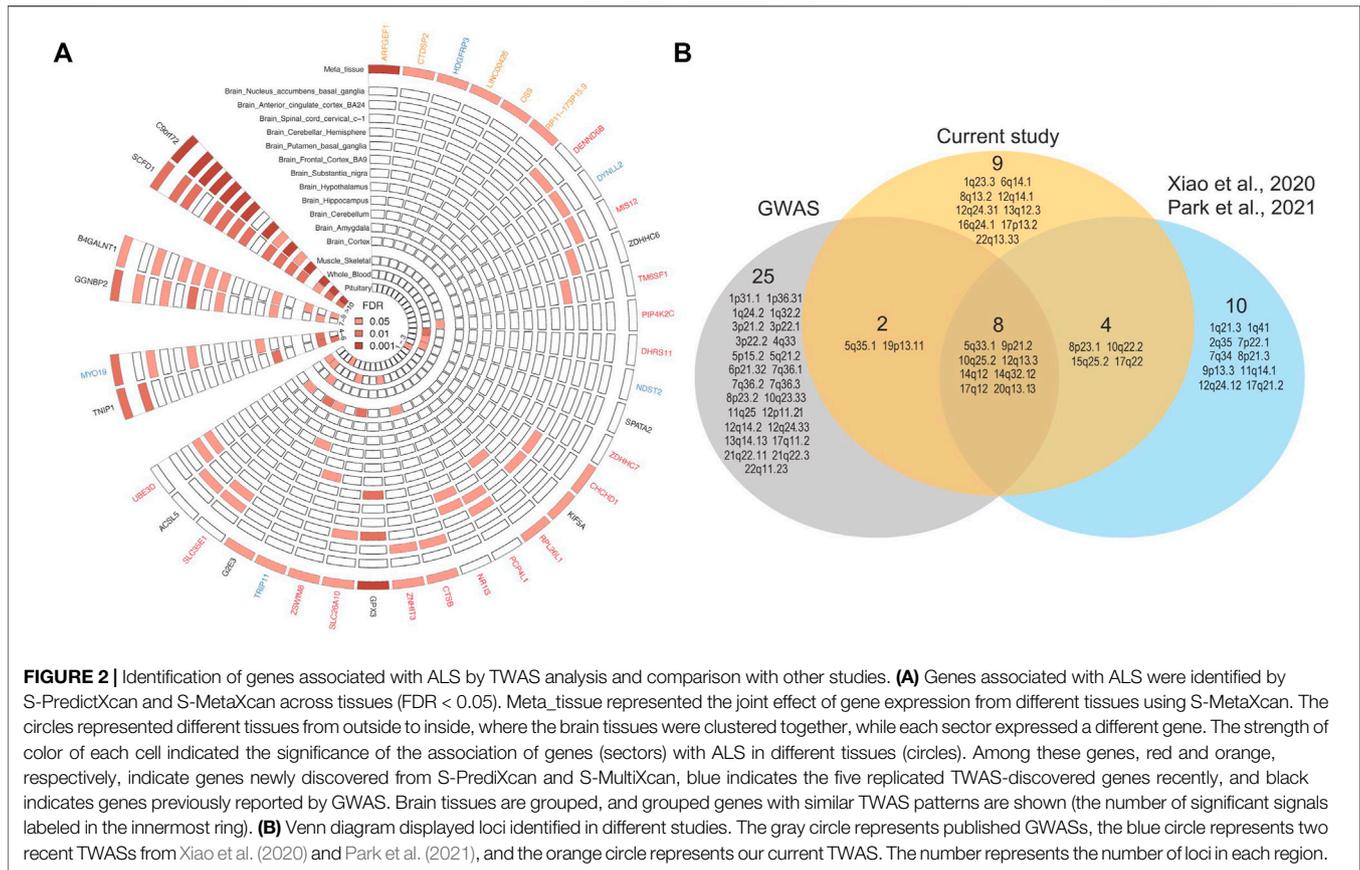
SMR (Zhu et al., 2016) was performed to evaluate the association between an exposure and an outcome in Mendelian randomization analysis principles by removing nongenetic confounding using a variant as an instrument variable. We used the same tissue as in the COOC analysis. A Benjamini–Hochberg method correction ($FDR < 0.05$) was used in each SMR analysis. HEIDI was applied to distinguish pleiotropy from linkage, and a threshold ($HEIDI > 0.01$) was considered to have little evidence of heterogeneity. We used SMR with the following parameters: *p*-value threshold $p < 1 \times 10^{-6}$ to select the top eQTL for the SMR test, a threshold for the difference in SNP allele frequency between datasets 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 **Figure 1**. We conducted a survey of 26 GWASs (Dunckley et al., 2007; Schymick et al., 2007; Van Es et al., 2007; Cronin et al., 2008; Van Es et al., 2008; Landers et al., 2009; van Es et al., 2009; Laaksovirta et al., 2010; Shatunov et al., 2010; Kwee et al., 2012; The ALSGEN Consortium, 2013; Deng et al., 2013; Diekstra

et al., 2014; Fogh et al., 2014; Xie et al., 2014; McLaughlin et al., 2015; Chen et al., 2016; Fogh et al., 2016; van Rheenen et al., 2016; Benyamin et al., 2017; Nicolas et al., 2018; Dekker et al., 2019; Wei et al., 2019; Iacoangeli et al., 2020; Nakamura et al., 2020) of ALS in the NHGRI-EBI GWAS Catalog (till September 2021) and additionally included one most recent GWAS (van Rheenen et al., 2021) and 3 recent post-GWASs (Du et al., 2018; Xiao et al., 2020; Park et al., 2021) to summarize the current knowledge on the candidate genes of ALS (**Supplementary Table S1; Figure 2B**). We then based our post-GWAS analyses on ALS GWAS from Nicolas et al. (2018), which represents the largest-ever GWAS data for ALS, totaling 20,806 cases and 59,804 controls of European ancestry. We first 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 FUMA has implemented the MAGMA tissue-specificity analysis using 54 tissues from GTEx. We identified 13 brain regions and pituitary with nominally significant association with ALS-gene associations ($p < 0.05$), and two regions showed significance after Bonferroni correction: cerebellum ($p = 2.7 \times 10^{-5}$) and cerebellar hemisphere ($p = 1.3 \times 10^{-4}$) (Supplementary Figure S2). 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 (Loeffler et al., 2016; Badu-Mensah et al., 2020), and whole blood, which was also implied in previous studies (Park et al., 2021).

MAGMA cell-type specificity analyses in a total of 196 cell types were performed in 8 different mouse and human scRNA-seq datasets. In the 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 were significantly associated with ALS-gene associations. Within-dataset step-wise conditional analyses identified inhibitory neuron 6 showing independent association (Supplementary Figure S3B; Supplementary Tables S4, S5), further highlighting the likelihood of inhibitory neuron 6 being the basic functional unit of ALS. In seven 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 Tables S3–S5). Cross-dataset conditional 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).

In addition, MAGMA identified seven genes showing significant association with ALS, 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, and 21q22.3 rs75087725) (Kircher et al., 2014). Of 233 candidate SNPs, CADD identified a total of 17 SNPs at 3 loci (9p21.2, 12q13.3, and 12q14.2) with high scores (>12.37), 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. RegulomeDB (Boyle et al., 2012) 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* (<60 kbp), 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 three ALS-associated loci and provided the list of candidate genes in these loci.

Positional mapping, eQTL mapping, and chromatin interaction mapping were mapped to 58 genes, among which 4 genes were mapped by all three mapping methods, including 5q33.1 *TNIP1*, 9p21.2 *C9orf72*, *MOB3B*, and *IFNK* (**Supplementary Table S8**). Interestingly, different from other loci, the 9p21.2 locus clearly contained a DNA loop (Fudenberg et al., 2016) 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 two loci (12q14.1 and 12q14.2) contained more signals for both eQTL and chromatin interactions compared with the other four 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

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, and 22q13.33) are newly identified (**Figure 2B**), highlighting six genes (*NR1I3*, *PCP4L1*, *UBE3D*, *ZDHHC7*, *MIS12*, and *DENND6B*). In addition, 16 genes have not been previously suggested as functional candidates for ALS (**Figure 2A; Supplementary Table S10**). Among the 31 genes, the most significant finding was *C9orf72* ($FDR = 5.03 \times 10^{-18}$ 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 the 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). 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 tissue, and 3 from 3 different loci were significant only in nonbrain tissues, that is, skeletal muscle (12q13.3 *PIP4K2C*), blood (17q12 *DHRS11*), and pituitary (16q24.1 *ZDHHC7*).

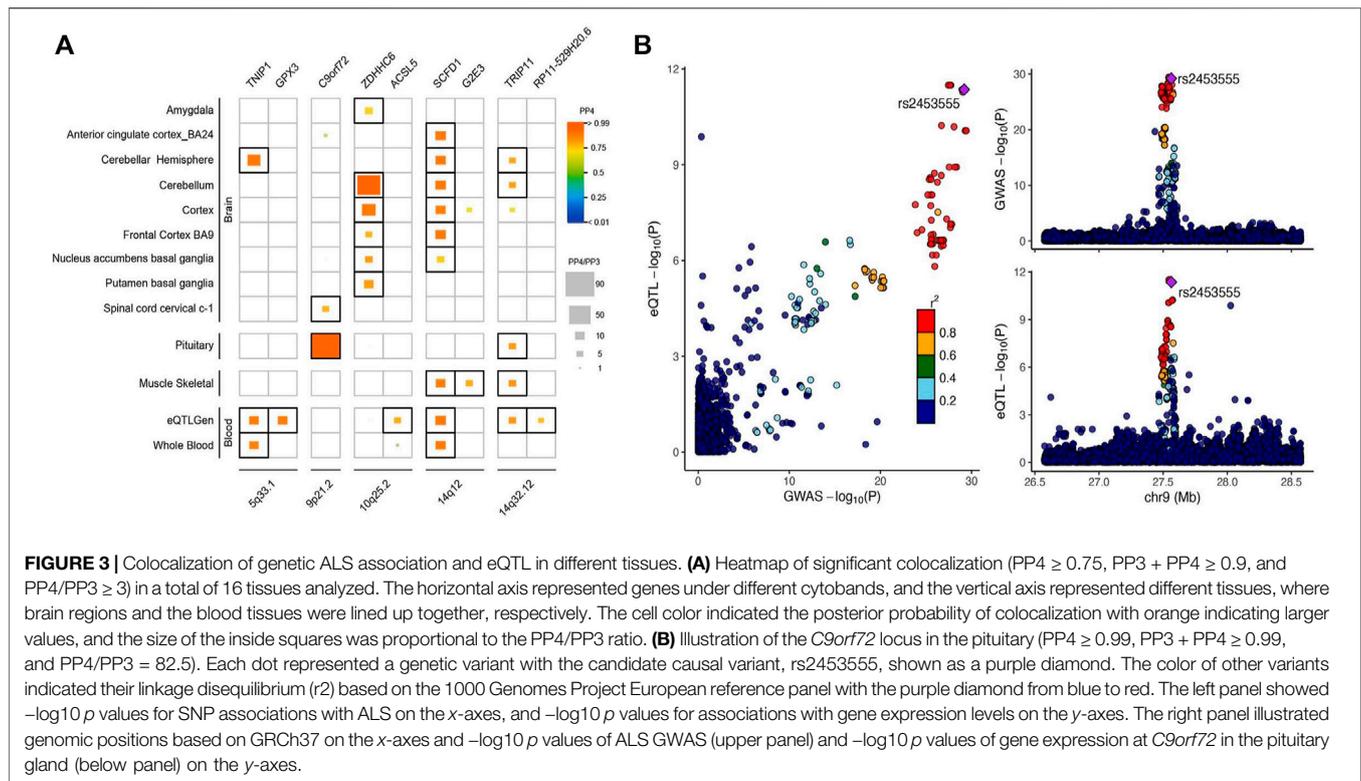
S-MultiXcan found a total of 22 genes at 14 distinct loci, among which six genes at six distinct loci were not identified in S-PrediXcan (**Figure 2A; Supplementary Table S11**), among which, five loci were novel. The most significant was *ARFGEF1* ($FDR = 6.4 \times 10^{-4}$), which involves vesicular trafficking and has previously been suggested to play a role in pathogenesis in ALS (Saris et al., 2009). 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.

We additionally conducted a TWAS using two different methods, that is, a unified test for molecular signatures (UTMOST) (Hu et al., 2019) and joint-tissue imputation (JTI) (Zhou et al., 2020). As the key gene *C9orf72* was not imputed in the UTMOST model, we focused on JTI. JTI identified a total of 18 genes (14 loci) in 16 tissues from GTEx v8 that reached the $FDR < 0.05$ significance level in the corresponding tissue. Among these, 11 genes (5q33.1 *TNIP1*, 9p21.2 *C9orf72*, 10q25.2 *ACSL5*, 12q13.3 *B4GALNT1*, 12q13.3 *PIP4K2C*, 14q12 *SCFD1*, 14q32.12 *TRIP11*, 17q12 *DHRS11*, 17q12 *ZNHIT3*, 17q12 *GGNBP2*, and 17q22 *DYNLL2*) at eight distinct loci overlapped with S-PrediXcan and S-MultiXcan (**Supplementary Table S12**). For the remaining seven genes, two genes (*SLC9A8* and *SNAI1*) were located at the locus 20q13.13, whereas S-PrediXcan identified a different gene *SPATA2*. One gene, *PLOD2* at 3q24, though failed to be identified by PrediXcan, was reported by our subsequent SMR analysis. The remaining four genes (4p16.3 *NSD2*, 5q22.1 *CAMK4*, 11p13 *AL356215.1*, and 20q11.22 *PIGU*) at four different loci have not been reported before (**Supplementary Table S12**).

Colocalization Highlights Genotype-Mediated Genes in Corresponding Tissues

We conducted a series of eQTL colocalization analyses in 13 brain tissues, pituitary, skeletal muscle, and blood from GTEx v8 and eQTLGen. These analyses identified a total of 9 genes at 5 loci showing significant evidence ($PP4 > 0.75$, $PP3 + PP4 > 0.9$, and $PP4/PP3 > 3$, **Figure 3A; Supplementary Table S13**) of colocalization with eQTLs in at least one tissue. These included 5q33.1 (*TNIP1* and *GPX3*), 9p21.2 (*C9orf72*), 10q25.2 (*ZDHHC6* and *ACSL5*), 14q12 (*SCFD1* and *G2E3*), and 14q32.12 (*TRIP11* and *RP11-529H20.6*).

The strongest signal according to the $PP4/PP3$ was identified for rs2453555 at 9p21.2 ($PP4/PP3 = 82.5$), which was highly significantly associated with ALS risk (GWAS $p = 6.5 \times 10^{-30}$), and at the same time served as a highly significant eQTL of *C9orf72* in the pituitary gland (eQTL $p = 4.4 \times 10^{-12}$), strongly suggesting a causal relationship (**Figure 3B**). This SNP was also a significant eQTL of *C9orf72* in spinal cord cervical at a much



lower significance level (eQTL $p = 8.1 \times 10^{-6}$, $PP4/PP3 = 5.2$). 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, $p = 3.8 \times 10^{-30}$, $r^2 = 0.98$) in the GWAS of Nicolas et al. (2018). 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* (Mountjoy et al., 2021). 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 five different brain tissues (5.1–19.9). 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 and 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 S13).

SMR Illustrates the Causal Relationships Between SNPs, Gene Expression, and ALS Risk

We conducted a comprehensive SMR analysis for ALS based on the GWAS of Nicolas et al. ($n = 80,610$) in 9 brain tissues, pituitary, and skeletal muscle. We identified a total of 9 genes from 6 loci with significant evidence mediating the genetic associations observed in these loci ($FDR < 0.05$ and $HEIDI > 0.01$, 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 (*PLOD2* and rs149615181, skeletal muscle) and 22q13.33 (*PLXNB2* and rs62241220, blood) have not been previously reported. Interestingly, at 9p21.2, *C9orf72* was found highly significantly mediating the association between rs2453555 and the risk of ALS in the pituitary ($FDR = 2 \times 10^{-5}$), which was at orders of magnitude more significant than in other brain and nonbrain tissues. This SNP is in very high LD with rs2453555 ($r^2 = 0.95$) identified by our colocalization analysis. This result boosted the likelihood of a causal chain between rs2453555/rs2453555, expression of *C9orf72* in the pituitary, and the risk of ALS. In the recent study of van Rheenen et al. (2021), the HEIDI test rejected the hypothesis

TABLE 1 | Genes mediating the genetic associations with ALS in six loci from SMR and HEIDI analyses.

Locus	Gene	SNP	FDR	HEIDI	Tissue	Database	
3q24	<i>PLOD2</i>	rs149615181	0.04	0.69	Muscle_Skeletal	GTEX v8	
	<i>C9orf72</i>	rs2453555	2.00×10^{-5}	0.14	Pituitary	GTEX v8	
		rs700795	0.02	0.19	Brain_Spinal_cord_cervical_c-1	GTEX v8	
10q22.2	<i>NDST2</i>	rs11000785	0.05	0.07	Blood	eQTLGen	
	<i>SCFD1</i>	rs7144204	3.6×10^{-3}	0.10	Blood	eQTLGen	
14q12		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	
		rs11650008	0.03	0.04	Blood	eQTLGen	
		<i>MYO19</i>	rs7222903	0.04	0.60	Blood	eQTLGen
		<i>DYNLL2</i>	rs2877858	0.04	0.16	Blood	eQTLGen
17q12	<i>ZNHIT3</i>	rs4796224	0.05	0.50	Blood	eQTLGen	
	<i>PLXNB2</i>	rs62241220	0.02	0.77	Blood	eQTLGen	

All genes with $FDR < 0.05$ and $HEIDI > 0.01$ 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.

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 ($FDR = 3.6 \times 10^{-3}$ 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 the 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 ($PP4/PP3 = 9.9$). A recent SMR analysis (Iacoangeli et al., 2021) found that rs229243 had a regulatory effect on ALS risk mediated by the expression of *SCFD1* in the blood 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*, and *ZNHIT3*) in the blood significantly mediated the genetic association in this locus (Supplementary Table S14). Among these four genes, *GGNBP2* was the most significant ($FDR = 0.03$), consistent with a previous study (Benyamin et al., 2017).

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 S15). 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 (nine from TWAS, one from SMR) have not been associated with ALS risk in previous studies.

A total of eight genes at six loci were significant in at least two out of three analyses. These included 5q33.1 *TNIP1*, 9p21.2 *C9orf72*, 10q25.2 *ACSL5*, 10q22.2 *NDST2*, 14q12 *SCFD1*, 17q12 *MYO19*, *GGNBP2*, and *ZNHIT3*. All these six loci have been previously associated with ALS risk, and all eight 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).

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 the 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 four 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 (Park et al., 2021) suggested *MYO19* as the most likely functional candidate of this locus, while another (Benyamin et al., 2017) 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, and 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 identified that rs2453555 at 9p21.2 and rs229243 at

TABLE 2 | Integration of TWAS, COLOC, and SMR results.

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

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 identified a functional relationship between a novel cell type and ALS risk, especially inhibitory neuron 6. This finding is an important supplement to previous findings suggesting microglia (Clarke and Patani, 2020), astrocytes (Do-Ha et al., 2018; Saez-Atienzar et al., 2021), and glutamatergic neurons Cronin et al. (2008) 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 in excitatory and inhibitory neurons using single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) profiles (Megat et al., 2021). 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 (Foerster et al., 2013; Ramirez-Jarquín and Tapia, 2018). A loss of inhibitory neuron influence is an important factor leading to ALS pathogenesis (Turner and Kiernan, 2012).

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 and 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 the presence of multiple causal variants at this locus cannot be excluded. The aggregation of dipeptide repeats 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 (Pellecchia et al., 2010; Dedeene et al., 2020). 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 (Mountjoy et al., 2021; van Rheenen et al., 2021) is likely because only blood and cortex or a smaller eQTL dataset were used. In addition, 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 (Querin et al., 2019). In contrast, another study showed that the cervical spinal cord progressively occurs to thin in ALS patients with *C9orf72* repeat expansion (van der Burgh et al., 2019). 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 (Hou et al., 2017). Surprisingly, skeletal muscle was also found to be the relevant tissue for *SCFD1*. Although the weakening of skeletal muscle was thought to be the initial hallmark of ALS, whether ALS originates in peripheral tissues (dying-back phenomenon) (Dadon-Nachum et al., 2011), including skeletal muscle or motor neurons (dying-forward phenomenon) (Braak et al., 2013) has been fiercely debated. Skeletal muscle was not considered pivotal to the etiology and treatment of ALS until recent years (Loeffler et al., 2016). Current studies about skeletal muscle degeneration/regeneration process focused on mutant *SOD1* mouse model to the pathology of ALS (Cheng et al., 2019; Badu-Mensah et al., 2020; Steyn et al., 2020), 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, we also found 23 novel functional candidate genes (Supplementary Table S15), among which 12 have supportive evidence from the literature.

However, our study is not without limitations. First, the sample size of ALS GWAS is relatively small compared with the latest GWAS (van Rheenen et al., 2021), 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. Second, LD structure (Giambartolomei et al., 2014) and gender differences (Aguet et al., 2020) in the sample may bias findings due to the unavailability of raw data. Third, with the constraints of currently available data, only cis-eQTL data were included in our analysis, which may miss the actual causal genes. Fourth are the 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.

In conclusion, we established causal relationships between genetic variants, candidate genes, functional cell types and tissues, and ALS risk. The prioritized genes and tissues serve as targets for future functional and drug discovery studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding authors.

REFERENCES

- Aguet, F., Anand, S., Ardlie, K. G., Gabriel, S., Getz, G. A., Graubert, A., et al. (2020). The GTEx Consortium Atlas of Genetic Regulatory Effects across Human Tissues. *Science* 369 (6509), 1318–1330. doi:10.1126/science.aaz1776
- Al-Chalabi, A., Fang, F., Hanby, M. F., Leigh, P. N., Shaw, C. E., Ye, W., et al. (2010). An Estimate of Amyotrophic Lateral Sclerosis Heritability Using Twin Data. *J. Neurol. Neurosurg. Psychiatry* 81 (12), 1324–1326. doi:10.1136/jnnp.2010.207464
- Badu-Mensah, A., Guo, X., McAleer, C. W., Rumsey, J. W., and Hickman, J. J. (2020). Functional Skeletal Muscle Model Derived from SOD1-Mutant ALS Patient iPSCs Recapitulates Hallmarks of Disease Progression. *Sci. Rep.* 10 (1), 14302. doi:10.1038/s41598-020-70510-3
- Barbeira, A. N., Dickinson, S. P., Dickinson, S. P., Bonazzola, R., Zheng, J., Wheeler, H. E., et al. (2018). Exploring the Phenotypic Consequences of Tissue Specific Gene Expression Variation Inferred from GWAS Summary Statistics. *Nat. Commun.* 9 (1), 1825. doi:10.1038/s41467-018-03621-1
- Barbeira, A. N., Pividori, M., Zheng, J., Wheeler, H. E., Nicolae, D. L., and Im, H. K. (2019). Integrating Predicted Transcriptome from Multiple Tissues Improves Association Detection. *PLoS Genet.* 15 (1), e1007889. doi:10.1371/journal.pgen.1007889
- Benyamin, B., He, J., Zhao, Q., Gratten, J., Garton, F., Leo, P. J., et al. (2017). Cross-ethnic Meta-Analysis Identifies Association of the GPX3-TNIP1 Locus with Amyotrophic Lateral Sclerosis. *Nat. Commun.* 8 (1), 611. doi:10.1038/s41467-017-00471-1
- Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., et al. (2012). Annotation of Functional Variation in Personal Genomes Using RegulomeDB. *Genome Res.* 22 (9), 1790–1797. doi:10.1101/gr.137323.112
- Braak, H., Brettschneider, J., Ludolph, A. C., Lee, V. M., Trojanowski, J. Q., and Tredici, K. D. (2013). Amyotrophic Lateral Sclerosis-A Model of Corticofugal Axonal Spread. *Nat. Rev. Neurol.* 9 (12), 708–714. doi:10.1038/nrneuro.2013.221
- Chen, H., and Boutros, P. C. (2011). VennDiagram: A Package for the Generation of Highly-Customizable Venn and Euler Diagrams in R. *BMC Bioinform.* 12 (1), 1–7. doi:10.1186/1471-2105-12-35
- Chen, C.-J., Chen, C.-M., Pai, T.-W., Chang, H.-T., and Hwang, C.-S. (2016). A Genome-wide Association Study on Amyotrophic Lateral Sclerosis in the Taiwanese Han Population. *Biomarkers Med.* 10 (6), 597–611. doi:10.2217/bmm.15.115
- Cheng, A. J., Allodi, I., Chaillou, T., Schlittler, M., Ivarsson, N., Lanner, J. T., et al. (2019). Intact Single Muscle Fibres from SOD1 G93A Amyotrophic Lateral Sclerosis Mice Display Preserved Specific Force, Fatigue Resistance and Training-like Adaptations. *J. Physiol.* 597 (12), 3133–3146. doi:10.1113/JP277456
- Clarke, B. E., and Patani, R. (2020). The Microglial Component of Amyotrophic Lateral Sclerosis. *Brain* 143 (12), 3526–3539. doi:10.1093/brain/awaa309
- Cronin, S., Berger, S., Ding, J., Schymick, J. C., Washecka, N., Hernandez, D. G., et al. (2008). A Genome-wide Association Study of Sporadic ALS in a Homogenous Irish Population. *Hum. Mol. Genet.* 17 (5), 768–774. doi:10.1093/hmg/ddm361
- Dadon-Nachum, M., Melamed, E., and Offen, D. (2011). The "Dying-Back" Phenomenon of Motor Neurons in ALS. *J. Mol. Neurosci.* 43 (3), 470–477. doi:10.1007/s12031-010-9467-1
- Dedeene, L., Van Schoor, E., Ospitalieri, S., Ronisz, A., Weishaupt, J. H., Otto, M., et al. (2020). Dipeptide Repeat Protein and TDP-43 Pathology along the Hypothalamic-Pituitary axis in C9orf72 and Non-C9orf72 ALS and FTLD-

AUTHOR CONTRIBUTIONS

PJ and FL conceived the study. ZZ, YD, SP and Y-YW conducted data curation. SP, XL, and TL conducted formal analysis. SP, FL, and PJ wrote the initial manuscript. FL and PJ supervised the study, All authors critically reviewed the manuscript and provided approval of the version to be published.

FUNDING

This research was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB38010400 and XDPB25), the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (KFJ-STZ-ZDTP-079 and KFJ-STZ-QYZD-2021-08-001), the Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), and the National Natural Science Foundation of China (81930056).

ACKNOWLEDGMENTS

We thank consortia such as the GTEx and eQTLGen for providing publicly available data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.917142/full#supplementary-material>

- TDP Cases. *Acta Neuropathol.* 140 (5), 777–781. doi:10.1007/s00401-020-02216-9
- Dekker, A. M., Diekstra, F. P., Pulit, S. L., Tazelaar, G. H. P., van der Spek, R. A., van Rheenen, W., et al. (2019). Exome Array Analysis of Rare and Low Frequency Variants in Amyotrophic Lateral Sclerosis. *Sci. Rep.* 9 (1), 1–8. doi:10.1038/s41598-019-42091-3
- Deng, M., Wei, L., Zuo, X., Tian, Y., Xie, F., Hu, P., et al. (2013). Genome-wide Association Analyses in Han Chinese Identify Two New Susceptibility Loci for Amyotrophic Lateral Sclerosis. *Nat. Genet.* 45 (6), 697–700. doi:10.1038/ng.2627
- Diekstra, F. P., Van Deerlin, V. M., Van Swieten, J. C., Al-Chalabi, A., Ludolph, A. C., Weishaupt, J. H., et al. (2014). C9orf72 and UNC13A are Shared Risk Loci for Amyotrophic Lateral Sclerosis and Frontotemporal Dementia: A Genome-wide Meta-analysis. *Ann. Neurol.* 76 (1), 120–133. doi:10.1002/ana.24198
- Do-Ha, D., Buskila, Y., and Ooi, L. (2018). Impairments in Motor Neurons, Interneurons and Astrocytes Contribute to Hyperexcitability in ALS: Underlying Mechanisms and Paths to Therapy. *Mol. Neurobiol.* 55 (2), 1410–1418. doi:10.1007/s12035-017-0392-y
- Du, Y., Wen, Y., Guo, X., Hao, J., Wang, W., He, A., et al. (2018). A Genome-wide Expression Association Analysis Identifies Genes and Pathways Associated with Amyotrophic Lateral Sclerosis. *Cell Mol. Neurobiol.* 38 (3), 635–639. doi:10.1007/s10571-017-0512-2
- Dunckley, T., Huentelman, M. J., Craig, D. W., Pearson, J. V., Szelinger, S., Joshipura, K., et al. (2007). Whole-Genome Analysis of Sporadic Amyotrophic Lateral Sclerosis. *N. Engl. J. Med.* 357 (8), 775–788. doi:10.1056/NEJMoa070174
- Farh, K. K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., et al. (2015). Genetic and Epigenetic Fine Mapping of Causal Autoimmune Disease Variants. *Nature* 518 (7539), 337–343. doi:10.1038/nature13835
- Foerster, B. R., Pomper, M. G., Callaghan, B. C., Petrou, M., Edden, R. A. E., Mohamed, M. A., et al. (2013). An Imbalance between Excitatory and Inhibitory Neurotransmitters in Amyotrophic Lateral Sclerosis Revealed by Use of 3-T Proton Magnetic Resonance Spectroscopy. *JAMA Neurol.* 70 (8), 1009–1016. doi:10.1001/jamaneurol.2013.234
- Fogh, I., Ratti, A., Gellera, C., Lin, K., Tiloca, C., Moskvina, V., et al. (2014). A Genome-wide Association Meta-Analysis Identifies a Novel Locus at 17q11.2 Associated with Sporadic Amyotrophic Lateral Sclerosis. *Hum. Mol. Genet.* 23 (8), 2220–2231. doi:10.1093/hmg/ddt587
- Fogh, I., Lin, K., Tiloca, C., Rooney, J., Gellera, C., Diekstra, F. P., et al. (2016). Association of a Locus in the CAMTA1 Gene with Survival in Patients with Sporadic Amyotrophic Lateral Sclerosis. *JAMA Neurol.* 73 (7), 812–820. doi:10.1001/jamaneurol.2016.1114
- Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15 (9), 2038–2049. doi:10.1016/j.celrep.2016.04.085
- Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., et al. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* 10 (5), e1004383. doi:10.1371/journal.pgen.1004383
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize Implements and Enhances Circular Visualization in R. *Bioinformatics* 30 (19), 2811–2812. doi:10.1093/bioinformatics/btu393
- Hou, N., Yang, Y., Scott, I. C., and Lou, X. (2017). The Sec Domain Protein Scfd1 Facilitates Trafficking of ECM Components during Chondrogenesis. *Dev. Biol.* 421 (1), 8–15. doi:10.1016/j.ydbio.2016.11.010
- Hu, Y., Li, M., Li, M., Lu, Q., Weng, H., Wang, J., et al. (2019). A Statistical Framework for Cross-Tissue Transcriptome-wide Association Analysis. *Nat. Genet.* 51 (3), 568–576. doi:10.1038/s41588-019-0345-7
- Iacoangeli, A., Lin, T., Al Khleifat, A., Jones, A. R., Opie-Martin, S., Coleman, J. R. I., et al. (2020). Genome-wide Meta-Analysis Finds the ACSL5-ZDHHHC6 Locus is Associated with ALS and Links Weight Loss to the Disease Genetics. *Cell Rep.* 33 (4), 108323. doi:10.1016/j.celrep.2020.108323
- Iacoangeli, A., Fogh, I., Selvakadunco, S., Topp, S. D., Shatunov, A., van Rheenen, W., et al. (2021). SCFD1 Expression Quantitative Trait Loci in Amyotrophic Lateral Sclerosis are Differentially Expressed. *Brain Commun.* 3 (4), fcab236. doi:10.1093/braincomms/fcab236
- Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M., and Shendure, J. (2014). A General Framework for Estimating the Relative Pathogenicity of Human Genetic Variants. *Nat. Genet.* 46 (3), 310–315. doi:10.1038/ng.2892
- Kwee, L. C., Liu, Y., Haynes, C., Gibson, J. R., Stone, A., Schichman, S. A., et al. (2012). A High-Density Genome-wide Association Screen of Sporadic ALS in US Veterans. *PLoS one* 7 (3), e32768. doi:10.1371/journal.pone.0032768
- Laaksovirta, H., Peuralinna, T., Schymick, J. C., Scholz, S. W., Lai, S.-L., Myllykangas, L., et al. (2010). Chromosome 9p21 in Amyotrophic Lateral Sclerosis in Finland: A Genome-wide Association Study. *Lancet Neurology* 9 (10), 978–985. doi:10.1016/S1474-4422(10)70184-8
- Landers, J. E., Melki, J., Meininger, V., Glass, J. D., Van Den Berg, L. H., Van Es, M. A., et al. (2009). Reduced Expression of the Kinesin-Associated Protein 3 (KIFAP3) Gene Increases Survival in Sporadic Amyotrophic Lateral Sclerosis. *Proc. Natl. Acad. Sci. U.S.A.* 106 (22), 9004–9009. doi:10.1073/pnas.0812937106
- Liu, B., Gloudemans, M. J., Rao, A. S., Ingelsson, E., and Montgomery, S. B. (2019). Abundant Associations with Gene Expression Complicate GWAS Follow-Up. *Nat. Genet.* 51 (5), 768–769. doi:10.1038/s41588-019-0404-0
- Loeffler, J. P., Picchiarelli, G., Dupuis, L., and Gonzalez De Aguilar, J. L. (2016). The Role of Skeletal Muscle in Amyotrophic Lateral Sclerosis. *Brain Pathol.* 26 (2), 227–236. doi:10.1111/bpa.12350
- Logrosino, G., and Piccininni, M. (2019). Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference. *Neuroepidemiology* 52 (1-2), 93–103. doi:10.1159/000493386
- McLaughlin, R. L., Kenna, K. P., Vajda, A., Bede, P., Elamin, M., Cronin, S., et al. (2015). A Second-Generation Irish Genome-wide Association Study for Amyotrophic Lateral Sclerosis. *Neurobiol. Aging* 36 (2), e7–1221. doi:10.1016/j.neurobiolaging.2014.08.030
- Megat, S., Mora, N., Sanogo, J., Catanese, A., Alami, N. O., Freischmidt, A., et al. (2021). Loss of Nucleoporin Nup50 is a Risk Factor for Amyotrophic Lateral Sclerosis. *medRxiv*. doi:10.1101/2021.08.23.21262299 (Online ahead of print)
- Mountjoy, E., Schmidt, E. M., Carmona, M., Schwartzentruber, J., Peat, G., Miranda, A., et al. (2021). An Open Approach to Systematically Prioritize Causal Variants and Genes at All Published Human GWAS Trait-Associated Loci. *Nat. Genet.* 53 (11), 1527–1533. doi:10.1038/s41588-021-00945-5
- Nakamura, R., Misawa, K., Tohrai, G., Nakatomi, M., Furuhashi, S., Atsuta, N., et al. (2020). A Multi-Ethnic Meta-Analysis Identifies Novel Genes, Including ACSL5, Associated with Amyotrophic Lateral Sclerosis. *Commun. Biol.* 3 (1), 526. doi:10.1038/s42003-020-01251-2
- Ng, B., White, C. C., Klein, H.-U., Sieberts, S. K., McCabe, C., Patrick, E., et al. (2017). An xQTL Map Integrates the Genetic Architecture of the Human Brain’s Transcriptome and Epigenome. *Nat. Neurosci.* 20 (10), 1418–1426. doi:10.1038/nn.4632
- Nicolas, A., Kenna, K. P., Renton, A. E., Ticozzi, N., Faghri, F., Chia, R., et al. (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* 97 (6), 1268. doi:10.1016/j.neuron.2018.02.027
- Park, S., Kim, D., Song, J., and Joo, J. W. J. (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), 3216. doi:10.3390/ijms22063216
- Pellecchia, M. T., Pivonello, R., Monsurrò, M. R., Trojsi, F., Longo, K., Piccirillo, G., et al. (2010). The GH-IGF System in Amyotrophic Lateral Sclerosis: Correlations between Pituitary GH Secretion Capacity, Insulin-like Growth Factors and Clinical Features. *Eur. J. Neurol.* 17 (5), 666–671. doi:10.1111/j.1468-1331.2009.02896.x
- Project MinE ALS Sequencing Consortium (2018). Project MinE: Study Design and Pilot Analyses of a Large-Scale Whole-Genome Sequencing Study in Amyotrophic Lateral Sclerosis. *Eur. J. Hum. Genet.* 26 (10), 1537–1546. doi:10.1038/s41431-018-0177-4
- Qi, T., Wu, Y., Wu, Y., Zeng, J., Zhang, F., Xue, A., et al. (2018). Identifying Gene Targets for Brain-Related Traits Using Transcriptomic and Methyloomic Data from Blood. *Nat. Commun.* 9 (1), 2282. doi:10.1038/s41467-018-04558-1
- Querin, G., Bede, P., El Mendili, M. M., Li, M., Péligrini-Issac, M., Rinaldi, D., et al. (2019). Presymptomatic Spinal Cord Pathology in C9orf72 Mutation Carriers: A Longitudinal Neuroimaging Study. *Ann. Neurol.* 86 (2), 158–167. doi:10.1002/ana.25520
- Ramirez-Jarquín, U. N., and Tapia, R. (2018). Excitatory and Inhibitory Neuronal Circuits in the Spinal Cord and Their Role in the Control of Motor Neuron Function and Degeneration. *ACS Chem. Neurosci.* 9 (2), 211–216. doi:10.1021/acscchemneuro.7b00503
- Saez-Atienzar, S., Bandres-Ciga, S., Langston, R. G., Kim, J. J., Choi, S. W., Reynolds, R. H., et al. (2021). Genetic Analysis of Amyotrophic Lateral

- Sclerosis Identifies Contributing Pathways and Cell Types. *Sci. Adv.* 7 (3), eabd9036. doi:10.1126/sciadv.abd9036
- Saris, C. G., Horvath, S., van Vught, P. W., van Es, M. A., Blauw, H. M., Fuller, T. F., et al. (2009). Weighted Gene Co-expression Network Analysis of the Peripheral Blood from Amyotrophic Lateral Sclerosis Patients. *BMC Genomics* 10, 405. doi:10.1186/1471-2164-10-405
- Schmick, J. C., Scholz, S. W., Fung, H.-C., Britton, A., Arepalli, S., Gibbs, J. R., et al. (2007). Genome-wide Genotyping in Amyotrophic Lateral Sclerosis and Neurologically Normal Controls: First Stage Analysis and Public Release of Data. *Lancet Neurol.* 6 (4), 322–328. doi:10.1016/S1474-4422(07)70037-6
- Shatunov, A., Mok, K., Newhouse, S., Weale, M. E., Smith, B., Vance, C., et al. (2010). Chromosome 9p21 in Sporadic Amyotrophic Lateral Sclerosis in the UK and Seven Other Countries: A Genome-wide Association Study. *Lancet Neurol.* 9 (10), 986–994. doi:10.1016/S1474-4422(10)70197-6
- Steyn, F. J., Li, R., Kirk, S. E., Tefera, T. W., Xie, T. Y., Tracey, T. J., et al. (2020). Altered Skeletal Muscle Glucose-Fatty Acid Flux in Amyotrophic Lateral Sclerosis. *Brain Commun.* 2 (2), fcaa154. doi:10.1093/braincomms/fcaa154
- The ALSGEN Consortium (2013). Age of Onset of Amyotrophic Lateral Sclerosis Is Modulated by a Locus on 1p34.1. *Neurobiol. Aging* 34 (1), e7–e357. doi:10.1016/j.neurobiolaging.2012.07.017
- Turner, M. R., and Kiernan, M. C. (2012). Does Interneuronal Dysfunction Contribute to Neurodegeneration in Amyotrophic Lateral Sclerosis? *Amyotroph. Lateral Scler.* 13 (3), 245–250. doi:10.3109/17482968.2011.636050
- van der Burgh, H. K., Westeneng, H.-J., Meier, J. M., van Es, M. A., Veldink, J. H., Hendrikse, J., et al. (2019). Cross-sectional and Longitudinal Assessment of the Upper Cervical Spinal Cord in Motor Neuron Disease. *NeuroImage Clin.* 24, 101984. doi:10.1016/j.nicl.2019.101984
- Van Es, M. A., Van Vught, P. W., Blauw, H. M., Franke, L., Saris, C. G., Andersen, P. M., et al. (2007). ITPR2 as a Susceptibility Gene in Sporadic Amyotrophic Lateral Sclerosis: a Genome-wide Association Study. *Lancet Neurol.* 6 (10), 869–877. doi:10.1016/S1474-4422(07)70222-3
- Van Es, M. A., Van Vught, P. W., Blauw, H. M., Franke, L., Saris, C. G., Van Den Bosch, L., et al. (2008). Genetic Variation in DPP6 is Associated with Susceptibility to Amyotrophic Lateral Sclerosis. *Nat. Genet.* 40 (1), 29–31. doi:10.1038/ng.2007.52
- van Es, M. A., Veldink, J. H., Saris, C. G., Blauw, H. M., van Vught, P. W., Birve, A., et al. (2009). Genome-wide Association Study Identifies 19p13.3 (UNC13A) and 9p21.2 as Susceptibility Loci for Sporadic Amyotrophic Lateral Sclerosis. *Nat. Genet.* 41 (10), 1083–1087. doi:10.1038/ng.442
- van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., et al. (2016). Genome-wide Association Analyses Identify New Risk Variants and the Genetic Architecture of Amyotrophic Lateral Sclerosis. *Nat. Genet.* 48 (9), 1043–1048. doi:10.1038/ng.3622
- van Rheenen, W., van der Spek, R. A. A., Bakker, M. K., van Vugt, J. J. F. A., Hop, P. J., Zwamborn, R. A. J., et al. (2021). Common and Rare Variant Association Analyses in Amyotrophic Lateral Sclerosis Identify 15 Risk Loci with Distinct Genetic Architectures and Neuron-specific Biology. *Nat. Genet.* 53 (12), 1636–1648. doi:10.1038/s41588-021-00973-1
- Võsa, U., Claringbould, A., Westra, H.-J., Bonder, M. J., Deelen, P., Zeng, B., et al. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* 53, 1300–1310. doi:10.1038/s41588-021-00913-z
- Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional Mapping and Annotation of Genetic Associations with FUMA. *Nat. Commun.* 8 (1), 1826. doi:10.1038/s41467-017-01261-5
- Watanabe, K., Umičević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P., and Posthuma, D. (2019). Genetic Mapping of Cell Type Specificity for Complex Traits. *Nat. Commun.* 10 (1), 3222. doi:10.1038/s41467-019-11181-1
- Wei, L., Tian, Y., Chen, Y., Wei, Q., Chen, F., Cao, B., et al. (2019). Identification of TYW3/CRYZ and FGD4 as Susceptibility Genes for Amyotrophic Lateral Sclerosis. *Neurol. Genet.* 5 (6), e375. doi:10.1212/NXG.0000000000000375
- Xiao, L., Yuan, Z., Jin, S., Wang, T., Huang, S., and Zeng, P. (2020). Multiple-Tissue Integrative Transcriptome-wide Association Studies Discovered New Genes Associated with Amyotrophic Lateral Sclerosis. *Front. Genet.* 11, 587243. doi:10.3389/fgene.2020.587243
- Xie, T., Deng, L., Mei, P., Zhou, Y., Wang, B., Zhang, J., et al. (2014). A Genome-wide Association Study Combining Pathway Analysis for Typical Sporadic Amyotrophic Lateral Sclerosis in Chinese Han Populations. *Neurobiol. Aging* 35 (7), e9–e1779. doi:10.1016/j.neurobiolaging.2014.01.014
- Zhou, S., Zhou, Y., Qian, S., Chang, W., Wang, L., and Fan, D. (2018). Amyotrophic Lateral Sclerosis in Beijing: Epidemiologic Features and Prognosis from 2010 to 2015. *Brain Behav.* 8 (11), e01131. doi:10.1002/brb3.1131
- Zhou, D., Jiang, Y., Zhong, X., Cox, N. J., Liu, C., and Gamazon, E. R. (2020). A Unified Framework for Joint-Tissue Transcriptome-wide Association and Mendelian Randomization Analysis. *Nat. Genet.* 52 (11), 1239–1246. doi:10.1038/s41588-020-0706-2
- Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., et al. (2016). Integration of Summary Data from GWAS and eQTL Studies Predicts Complex Trait Gene Targets. *Nat. Genet.* 48 (5), 481–487. doi:10.1038/ng.3538

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer YY declared a shared parent affiliation with the authors ZZ and YD to the handling editor at the time of the review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Pan, Liu, Liu, Zhao, Dai, Wang, Jia and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

ALS	amyotrophic lateral sclerosis	MYO19	myosin XIX
LD	linkage disequilibrium	DHRS11	dehydrogenase/reductase 11
eQTL	expression quantitative trait loci	GGNBP2	gametogenetin binding protein 2
FTD	frontotemporal dementia	SPATA2	spermatogenesis-associated 2
AD	Alzheimer's disease	ZSWIM8	zinc finger SWIM-type containing 8
PD	Parkinson's disease	ACSL5	acyl-CoA synthetase long chain family member 5
FUMA	functional mapping and annotation	UBE3D	ubiquitin protein ligase E3D
MAGMA	Multi-marker Analysis of GenoMic Annotation	NDST2	N-deacetylase and N-sulfotransferase 2
CADD	Combined Annotation-Dependent Depletion	ZNHIT3	zinc finger HIT-type containing 3
GWAS	genome-wide association study	POLD2	DNA polymerase delta 2, accessory subunit
TWAS	transcriptome-wide association analysis	DYNLL2	dynein light chain LC8-Type 2
MASHR	multivariate Adaptive Shrinkage in R	KIF3B	kinesin family member 3B
COLOC	colocalization	ARFGEF1	ADP ribosylation factor guanine nucleotide exchange factor 1
SMR	summary data-based Mendelian randomization analysis	GPX3	glutathione peroxidase 3
HEIDI	heterogeneity in dependent instruments	RPL26L1	ribosomal protein L26-like one
FDR	false discovery rate	CTDSP2	CTD small phosphatase 2
GTE_x	Genotype-Tissue Expression	TRIP11	thyroid hormone receptor interactor 11
scATAC-seq	single-cell assay for transposase-accessible chromatin sequencing	CHCHD1	coiled-coil-helix-coiled-coil-helix domain containing one
scRNA-seq	single-cell RNA sequencing	HDGFRP3	hepatoma-derived growth factor-related protein 3
OR	odds ratio	CTSB	cathepsin B
GABA_A	gamma-aminobutyric acid type A	SLC26A10	solute carrier family 26 member 10
C9orf72	C9orf72-SMCR8 complex subunit	LINC00426	long intergenic non-protein coding RNA 426
SCFD1	Sec1 family domain containing 1	MOB3B	MOB kinase activator 3B
TNIP1	TNFAIP3-interacting protein 1	ZDHHC6	zinc finger DHHC-type palmitoyltransferase 6
G2E3	G2/M-phase-specific E3 ubiquitin protein ligase	ATXN3	ataxin 3
B4GALNT1	beta-1,4-N-acetyl-galactosaminyltransferase 1	UNC13A	Unc-13 homolog A
ZDHHC7	zinc finger DHHC-type palmitoyltransferase 7	C21orf2	chromosome 21 open reading frame 2
		SOD1	superoxide dismutase one