Causal Inference of Genetic Variants and Genes in Amyotrophic Lateral Sclerosis

Siyu Pan1,2, Xinxuan Liu1,3, Tianzi Liu1, Zhongming Zhao4, Yulin Dai4, Yin-Ying Wang1, Peilin Jia1,4 and Fan Liu1,3*

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

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 Rijswijck 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 × 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 “circize” (Gu et al., 2014).

We compared the identified loci with the published TWASs and GWASs (Supplementary Table S1, p threshold < 5 × 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, p1 (1 × 10^-5), p2 (1 × 10^-4), and p12 (1 × 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; Võsa 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 × 10^-6 from GWAS that had a p value < 1 × 10^-4 in the eQTL dataset. We applied 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 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 (Loeflfler 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). 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 IFNKB (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 × 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 DHR511), 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 ARFGGEF1 (FDR = 6.4 × 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 B4GALT1, 12q13.3 PIP4K2C, 14q12 SCFD1, 14q21.12 TRIP11, 17q12 DHR511, 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 SNAIL1) 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 × 10^{-30}), and at the same time served as a highly significant eQTL of C9orf72 in the pituitary gland (eQTL p = 4.4 × 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 \((e\text{QTL} \ p = 8.1 \times 10^{-6}, \text{PP4/PP3} = 5.2)\). 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 \((\text{PP4/PP3} \text{ ranging between 9.9 and 12 except in nucleus accumbens basal ganglia, where PP4/PP3 = 6.0})\). 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 \((\text{FDR} < 0.05 \text{ and HEIDI} > 0.01, \text{Table 1})\). 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 \((\text{PLOD2 and rs149615181, skeletal muscle})\) and 22q13.33 \((\text{PLXNB2 and rs22241220, blood})\) have not been previously reported. Interestingly, at 9p21.2, C9orf72 was found highly significantly mediating the association between rs2453565 and the risk of ALS in the pituitary \((\text{FDR} = 2 \times 10^{-3})\), 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/rs2453565, 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.
that expression of C9orf72 could mediate the association between rs2453555 and ALS risk in blood. Our findings stress 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 × 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 associated with 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.

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

| Locus | Gene | SNP | FDR | HEIDI | Tissue | Database |
|-------|------|-----|-----|-------|--------|----------|
| 3q24  | PLOD2| rs149615181 | 0.04 | 0.69 | Muscle_Skeletal | GTEx v8 |
| 9p21.2| C9orf72| rs2453555 | 2.00 × 10^{-5} | 0.14 | Pituitary | GTEx v8 |
| 10q22.2| NDST2| rs1100785 | 0.02 | 0.19 | Brain_Spinal_cord_cervical_c-1 | GTEx v8 |
| 14q12 | SCFD1| rs7144204 | 3.6 × 10^{-3} | 0.10 | Blood | eQTLGen |
| 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| rs82241220 | 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.
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 (Foerster et al., 2013; Ramirez-Jarquin and Tapia, 2018). A loss of neurons, ensuring our brain functions smoothly and accident-free 

17q12 functionally contributes to the development of ALS by regulating the expression of 9p21.2 in pituitary and 14q12 in skeletal muscle, respectively.

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

| Locus     | Gene                  | Tissue                      | TWAS     | COLOC (PP4/PP3) | SMR (HEIDI) | Overall evidence |
|-----------|-----------------------|-----------------------------|----------|-----------------|-------------|------------------|
| 5q33.1    | TNNP1                 | Blood                       | 2.2 x 10^-3 | 0.91 (9.70)     | 3.6 x 10^-3 (6.7 x 10^-6) | ++              |
| 9p21.2    | C9orf72               | Brain_Spinal_cord_cervical_c-1 | 7.00 x 10^-3 | 0.81 (5.10)     | 1.5 x 10^-3 (0.19)  | +++             |
| 9p21.2    | C9orf72               | Pituitary                   | 5.00 x 10^-4 | 0.99 (82.50)    | 2.00 x 10^-5 (0.14) | +++             |
| 10q25.2   | ACSL5                 | Blood                       | 2.7 x 10^-2  | 0.82 (5.10)     |              | ++              |
| 10q22.2   | NOS1                  | Blood                       | 3.3 x 10^-2  |                | 4.9 x 10^-2 (0.07) | ++              |
| 14q12     | SCFD1                 | Brain_Cerebellum            | 1.7 x 10^-3  | 0.92 (11.81)    | 3.1 x 10^-2 (0.45) | +++             |
| 14q12     | SCFD1                 | Muscle_Skeletal             | 3.7 x 10^-2  | 0.91 (8.85)     | 3.8 x 10^-2 (0.31) | +++             |
| 14q12     | SCFD1                 | Blood                       | 1.5 x 10^-3  | 0.92 (12.00)    | 1.1 x 10^-3 (0.35) | +++             |
| 14q12     | SCFD1                 | Brain_Anterior_cingulate_cortex_BA24 | 3.3 x 10^-3 | 0.92 (11.85)    | -             | ++              |
| 14q12     | SCFD1                 | Brain_Cerebellar_Hemisphere | 1.1 x 10^-3  | 0.92 (11.35)    | -             | ++              |
| 14q12     | SCFD1                 | Brain_Cortex               | 2.4 x 10^-3  | 0.91 (10.75)    | -             | ++              |
| 14q12     | SCFD1                 | Brain_Frontal_Cortex_BA9    | 1.3 x 10^-3  | 0.92 (11.80)    | -             | ++              |
| 14q12     | SCFD1                 | Brain_Nucleus_accumbens_basal_ganglia | 2.7 x 10^-3 | 0.77 (0.01)     | -             | ++              |
| 17q12     | MYO19                 | Blood                       | 2.2 x 10^-3  |                | 4.0 x 10^-2 (0.8)  | ++              |
| 17q12     | GNBTP2                | Blood                       | 1.3 x 10^-2  |                | 2.5 x 10^-2 (0.04) | ++              |
| 17q12     | ZNHit3                | Blood                       | 3.5 x 10^-2  |                | 4.9 x 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 and 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.
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 biological role of these genes in the pathology of ALS. In conclusion, we established causal relationships between genetic variants, candidate genes, cellular function 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., Pivodori, 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 GDX3-TNPI1 Locus with Amyotrophic Lateral Sclerosis. Nat. Commun. 8 (1), 611. doi:10.1038/s41467-017-00741-1

Boyle, A. P., Hong, E. L., Haritamah, M., Cheng, Y., Schaumb, 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

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 XDBP25), the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (KFJ-GSTS-ZDTP-079 and KFJ-STS-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

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/nrneurol.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 Bioform. 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/ bm.15.115

Cheng, A. I., Allodi, L., Chailou, 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/ JPH27456

Clarke, B. E., and Patani, R. (2020). The Microsomal Component of Amyotrophic Lateral Sclerosis. Brain 143 (12), 3526–3539. doi:10.1093/brain/awaa309

Cronin, S., Berger, S., Ding, J., Schmick, 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/ddn361

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. (2011). The “Dying-Back” Phenomenon of Motor Neurons in ALS. J. Neurol. Neurosurg. Psychiatry 82 (9), 1093–1099. doi:10.1136/jnnp.2010.232074

J. Neurol. Neurosurg. Psychiatry 81 (12), 1324–1326. doi:10.1136/jnnp.2010.207464

Pan et al. Genetic Causal Inference in ALS
Sclerosis Identifies Contributing Pathways and Cell Types. Sci. Adv. 7 (3), eab9036. doi:10.1126/sciadv.ab9036
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
Schymick, J. C., Scholtz, 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:10.3109/17482966.2011.636050
van der Burgh, H. K., Westeneng, H.-J., Meier, J. M., van Es, M. A., Veldink, J. H., Hendriksje, 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
vandeRheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Dickstra, P. P., Pult, S. L., et al. (2016). Genome-wide 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 Vught, 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
vOsa, 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 Bochove, A., and Posthumus, 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., Umlicevic, Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P., and Posthumus, 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/NGX.0000000000100375
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, Z., 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
LD linkage disequilibrium
eQTL expression quantitative trait loci
FTD frontotemporal dementia
AD Alzheimer’s disease
PD Parkinson’s disease
FUMA functional mapping and annotation
MAGMA Multi-marker Analysis of GenoMic Annotation
GWAS genome-wide association study
TWAS transcriptome-wide association analysis
MASHR multivariate Adaptive Shrinkage in R
COLOC colocalization
SMR summary data-based Mendelian randomization analysis
HEIDI heterogeneity in dependent instruments
FDR false discovery rate
GTEx Genotype-Tissue Expression
scATAC-seq single-cell assay for transposase-accessible chromatin sequencing
scRNA-seq single-cell RNA sequencing
GABA_A gamma-aminobutyric acid type A
OR odds ratio
C9orf72 C9orf72-SMCR8 complex subunit
SCFD1 Sec1 family domain containing 1
TNIP1 TNFAIP3-interacting protein 1
G2E3 G2/M-phase-specific E3 ubiquitin protein ligase
B4GALNT1 beta-1,4-N-acetyl-galactosaminyltransferase 1
ZDHHC7 zinc finger DHHC-type palmitoyltransferase 7
MYO19 myosin XIX
DHR511 dehydrogenase/reductase 11
GGNBP2 gametogenetin binding protein 2
SPATA2 spermatogenesis-associated 2
ZSWIM8 zinc finger SWIM-type containing 8
ACSL5 acyl-CoA synthetase long chain family member 5
UBE3D ubiquitin protein ligase E3D
NDST2 N-deacetylase and N-sulfotransferase 2
ZNHI3 zinc finger HIT-type containing 3
POLD2 DNA polymerase delta 2, accessory subunit
DYNLL2 dynein light chain LC8-Type 2
KIF3B kinesin family member 3B
ARFGEF1 ADP ribosylation factor guanine nucleotide exchange factor 1
GPX3 glutathione peroxidase 3
RPL26L1 ribosomal protein L26-like one
CTDSP2 CTD small phosphatase 2
TRIP11 thyroid hormone receptor interactor 11
CHCHD1 coiled-coil-helix-coiled-coil-helix domain containing one
HDGFRP3 hepatoma-derived growth factor-related protein 3
CTSB cathepsin B
SLC26A10 solute carrier family 26 member 10
LINC00426 long intergenic non-protein coding RNA 426
MOB3B MOB kinase activator 3B
ZDHHC6 zinc finger DHHC-type palmitoyltransferase 6
ATXN3 ataxin 3
UNC13A Unc-13 homolog A
C21orf2 chromosome 21 open reading frame 2
SOD1 superoxide dismutase one