

RESEARCH ARTICLE

Open Access



Identification of novel proteins for lacunar stroke by integrating genome-wide association data and human brain proteomes

Chengcheng Zhang¹, Fengqin Qin², Xiaojing Li¹, Xiangdong Du³ and Tao Li^{4,5*}

Abstract

Background: Previous genome-wide association studies (GWAS) have identified numerous risk genes for lacunar stroke, but it is challenging to decipher how they confer risk for the disease. We employed an integrative analytical pipeline to efficiently transform genetic associations to identify novel proteins for lacunar stroke.

Methods: We systematically integrated lacunar stroke genome-wide association study (GWAS) ($N=7338$) with human brain proteomes ($N=376$) to perform proteome-wide association studies (PWAS), Mendelian randomization (MR), and Bayesian colocalization. We also used an independent human brain proteomic dataset ($N=152$) to annotate the new genes.

Results: We found that the protein abundance of seven genes (*ICA1L*, *CAND2*, *ALDH2*, *MADD*, *MRV11*, *CSPG4*, and *PTPN11*) in the brain was associated with lacunar stroke. These seven genes were mainly expressed on the surface of glutamatergic neurons, GABAergic neurons, and astrocytes. Three genes (*ICA1L*, *CAND2*, *ALDH2*) were causal in lacunar stroke ($P < 0.05$ /proteins identified for PWAS; posterior probability of hypothesis 4 $\geq 75\%$ for Bayesian colocalization), and they were linked with lacunar stroke in confirmatory PWAS and independent MR. We also found that *ICA1L* is related to lacunar stroke at the brain transcriptome level.

Conclusions: Our present proteomic findings have identified *ICA1L*, *CAND2*, and *ALDH2* as compelling genes that may give key hints for future functional research and possible therapeutic targets for lacunar stroke.

Keywords: Lacunar stroke, Human brain proteomes, Functional genomics, *ICA1L*, *CAND2*, *ALDH2*

Background

Lacunar stroke has been recognized as a stroke subtype for over 50 years, although the etiology and whether it differs from cortical ischemic stroke are still debated [1]. Approximately 30% of patients with lacunar stroke are left dependent, and up to 25% of patients are predicted to have another stroke within 5 years [2]. The increase in

large-scale genome-wide association studies (GWAS) has greatly aided the discovery of genetic variations linked to lacunar stroke during the last decade [3]. However, deciphering the underlying biological processes responsible for the great majority of these genetic effects remains difficult, which has hampered the translation of these genetic results into novel drugs targeting these candidate genes for lacunar stroke [4].

Proteins are the most efficient biomarkers and therapeutic targets [5, 6] as they represent the major functional components of cellular and biological processes and the end products of gene expression [7]. It is critical to investigate the risk proteins in the brain disorders

*Correspondence: litaozjusc@zju.edu.cn

⁴ Department of Neurobiology, Affiliated Mental Health Center & Hangzhou Seventh People's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

[8, 9]. Previous research on lacunar stroke examined genetic, epigenetic, and transcriptome variables [10, 11], but few studies have explored brain proteins directly [12]. For example, previous studies identified an association between loci on chromosome 16q24.2 and small vessel stroke in 4203 cases and 50,728 controls [13]. In addition, a transcriptome-wide association study identified associations between the expression of six genes (*SCL25A44*, *ULK4*, *CARF*, *FAM117B*, *ICAIL*, *NBEAL1*) and lacunar stroke [14]. The current breakthrough in high-throughput proteome sequencing of complex tissues [15, 16] represents a significant step forward in the large-scale quantification of the human brain proteome. Wingo et al. developed a novel framework called proteome-wide association studies (PWAS) to combine gene and protein expression data with the results of GWAS (integrate gene expression data and GWAS results) in depression pathogenesis [12]. Ou and colleagues also revealed that particular genetic variants impact disorders by altering the quantity of brain proteins, and uncovered potentially brain-pathogenic proteins in Alzheimer's disease [17]. Thus, the causal inference of this integrated analytical approach has been empirically verified and shown to be reliable [17, 18].

Accordingly, we sought to discover novel drug targets for lacunar stroke by combining high-throughput proteomics in the brain with genetic data to determine the genomic architecture-associated protein levels. To identify potential protein biomarkers, we systematically linked protein biomarkers to lacunar stroke by taking a four-step approach. First, we used two protein quantitative trait locus (pQTL) datasets obtained from brain tissue and findings from lacunar stroke GWAS to perform a PWAS analysis. Second, we used independent Mendelian randomization (MR) analysis to verify PWAS-significant genes. Third, we used a COLOC to integrate GWAS data and brain pQTL using a Bayesian colocalization analysis to explore whether two associated signals are consistent with shared causal variant(s). Fourth, we explored the significant genes driving GWAS signals at the transcriptional level by leveraging gene expression data.

Methods and materials

Human brain protein abundance references in the discovery PWAS

The discovery PWAS data were obtained from the dorsolateral prefrontal cortex (dPFC) of postmortem brain tissues from 376 subjects recruited by the Religious Orders Study/Memory and Aging Project (ROS/MAP) [19]. For proteome sequencing, isobaric tandem mass tag peptide labeling was utilized, and peptides were assessed using liquid chromatography coupled to mass spectrometry (MS) [20]. Wingo et al. [21] used Thermo

Fisher Scientific's Proteome Discoverer suite v.2.3 and tandem MS spectra to search against the standard UniProtKB human proteome database, which has 20,338 total sequences, to assign peptide spectral matches. Genotyping was done using whole-genome sequencing or genome-wide genotyping on the Illumina OmniQuad Express or Affymetrix GeneChip 6.0 platforms. Over 8356 proteins having both proteomic and genomic data, among which 1475 protein could find significant cis associations with genetic variation.

Human brain protein abundance references in the confirmation PWAS

The confirmation PWAS data were profiled from the dPFC of postmortem brain samples from 198 participants recruited by the Banner Sun Health Research Institute (Banner) [22]. Proteomic profiling followed the same steps as the discovery proteomes, with two exceptions: only MS2 scans were collected, and MS2 spectra were compared to the UniProtKB human brain proteome database [22]. Individuals from Banner were genotyped using an Affymetrix Precision Medicine Array following the manufacturer's protocol and DNA extracted from the brain with a Qiagen GenePure kit [20]. Following quality control [12], we included 152 individuals having pQTL data in our replication analysis.

Human brain eQTL in the lacunar stroke TWAS

Transcriptomes data were profiled from postmortem brain samples donated by 452 individuals recruited by the CommonMind Consortium (CMC) [23]. These transcriptomes were profiled mainly from the dPFC. The RNA-seq data were adjusted for diagnosis, institution collecting the data, sex, disease onset age, postmortem interval (PMI), RNA integrity number (RIN), RIN^2 , clustered library batch variable, and 20 surrogate variables. The eQTL was calculated according to the formula: adjusted gene expression \sim SNP dosage + ancestry vectors + diagnosis. We retrieved the gene-level eQTL results adjusted with surrogate variable analysis. Detailed information can be found in the original study [23].

Lacunar stroke GWAS data

We used the summary association statistics from the largest GWAS of lacunar stroke by Traylor et al. [14], which included 7338 cases and 254,798 controls of European, South Asian, African, and Hispanic ancestry recruited from hospitals across the UK as part of the UK DNA Lacunar Stroke studies 1 and 2 and the International Stroke Genetics Consortium. The current study used lacunar stroke samples mostly from the magnetic resonance imaging (MRI) verified and traditional phenotypic groups. In the MRI-confirmed group, lacunar

stroke was defined as a clinical lacunar syndrome with an anatomically compatible lesion on MRI, either as (i) a high intensity region on diffusion-weighted imaging for acute infarcts or (ii) a low intensity region on fluid-attenuated inversion recovery or (iii) T1 imaging for non-acute infarcts, and the absence of other causes of stroke other than small vessel disease. In the traditional phenotyping group, lacunar stroke was also classified using the TOAST criteria, which is comprised of a clinical lacunar syndrome and the absence of other types of stroke, as well as non-lacunar infarction on CT. MRI identified 2987 patients with lacunar stroke accounting for 40.7% of all lacunar stroke patients. Meta-analysis was performed as previously described [24] by METAL tool using the fixed-effects inverse-variance weighted model [25]. Following meta-analysis, the λ_{1000} value in the transethnic analysis covering European, South Asian, African, and Hispanic was 1.005, showing no significant inflation [14]. Detailed information about the study subjects, diagnosis, genotyping, quality control, and statistical analyses was provided in the original papers [14].

Statistical analysis

Proteome-wide association studies (PWAS)

PWAS were carried out using FUSION [26]. For simplicity, we used FUSION to compute the effect of SNPs on protein abundance for proteins with significant heritability (heritability $P < 0.01$). Multiple predictive models, top1, blup, lasso, enet, and bsldmm, were adopted in the analysis [26]. Protein weights from the most predictive model were selected. Subsequently, we used FUSION to combine the genetic effect of lacunar stroke (lacunar stroke GWAS z -score) with the protein weights by calculating the linear sum of z -score \times weight for the independent SNPs at the locus to perform the PWAS of lacunar stroke.

Mendelian Randomization (MR) analysis

MR was used to verify whether lacunar stroke PWAS-significant genes (from the FUSION approach) were associated with lacunar stroke via their cis-regulated brain protein abundance. The SNPs included in the study robustly and independently ($R^2 < 0.001$) predicted exposures at a genome-wide level (5×10^{-8}). The Wald ratio calculates the log odds change in lacunar stroke risk per standard deviation change in protein biomarker in relation to the instrumenting SNP's risk allele [27]. A weighted mean of the ratio estimates weighted by the inverse variance of the ratio estimates (inverse-variance weighted approach) was employed when more than one SNP was available [28]. Complementary approaches were also used, such as weighted median, MR-Egger, simple mode, and weighted mode. To construct MR estimates,

the “TwoSampleMR” package [29] in R 4.1.02 was utilized.

Bayesian colocalization analysis

To assess the probability of the same single-nucleotide variation being responsible for both changing the lacunar stroke risk and modulating the protein levels of a gene, we used the COLOC method [26, 30]. We used the default COLOC priors of $p_1 = 10^{-4}$, $p_2 = 10^{-4}$, and $p_{12} = 10^{-5}$, where p_1 is the probability that a given variant is associated with lacunar stroke, p_2 is the probability that a given variant is a significant pQTL, and p_{12} is the probability that a given variant is both a lacunar stroke result and a pQTL. COLOC uses computed approximation Bayes factors and summary association data to generate posterior probability for the following 5 hypotheses: H_0 , No association with either GWAS or pQTL; H_1 , Association with GWAS, not with pQTL; H_2 , Association with pQTL, not with GWAS; H_3 , Association with GWAS and pQTL, two independent SNPs; and H_4 , Association with GWAS and pQTL, one shared SNP. The posterior probability (PP), represented by PP0, PP1, PP2, PP3, and PP4, quantifies support for each of the hypotheses. H_4 of 0.75 or above were chosen as strong evidence for colocalization.

Transcriptome-wide association studies (TWAS)

Using FUSION [26], which generates the linear sum of Z score weights for the independent SNPs at the locus, then the genetic influence of lacunar stroke (lacunar stroke GWAS Z score) was combined with the mRNA expression weights. The following was the fundamental procedure: firstly, FUSION computed TWAS expression weights (i.e., SNP-gene expression correlations) from the reference expression panels (i.e., CMC) [23]. To identify the best gene prediction model, FUSION did a fivefold cross-validation of each model to obtain an out-sample R^2 [26]. The imputed gene expression was then used to investigate the association with lacunar stroke [31].

Cell-type specificity analysis

Using human brain single-cell RNA sequencing (RNA-seq) data profiled from the Cell Types database (<https://portal.brain-map.org/atlas-and-data/rnaseq>), we investigated the cell type-specific expression of the risk genes. Individual layers of the cortex were dissected, and nuclei were dissociated and sorted using the neuronal marker NeuN from human brain tissues. The expression was profiled with SMART-Seq v4 or 10 \times Genomics Chromium Single Cell 3' v3 RNA-seq. CELLEX (CELL-type EXpression-specificity), a method for generating cell-type expression specificity (ES) profiles, was used to obtain gene expression specificity values [32, 33].

Results

Discovery and replication PWAS of lacunar stroke

The PWAS identified 7 genes (*ICA1L*, *CAND2*, *ALDH2*, *MADD*, *MRVI1*, *CSPG4*, and *PTPN11*) whose cis-regulated brain protein levels were associated with lacunar stroke at a false discovery rate (FDR) of $P < 0.05$ (Additional file 1: Table S1). Four genes (*ICA1L*, *CAND2*, *ALDH2*, and *MADD*) could be replicated in the independent PWAS of lacunar stroke, providing a higher confidence level (Fig. 1 and Table 1). Three of the 7 significant proteins from the discovery PWAS could not be tested in the confirmation PWAS, with 2 proteins (*CSPG4* and

PTPN11) not profiled, and *MRVI1* was profiled but did not have substantial heritability, which is likely due to the smaller sample size (Table 1).

Cell-type specificity analysis in the brain

We investigated whether the risk genes identified by PWAS were enriched in a particular brain cell type. Using human single-cell RNA-seq data from the Cell Types database (<https://portal.brain-map.org/atlas-and-data/rnaseq>), we found cell type-specific enrichment for the expression of the seven causal genes (Fig. 2). *MRVI1* and *ALDH2* were found to be more abundant in astrocytes,

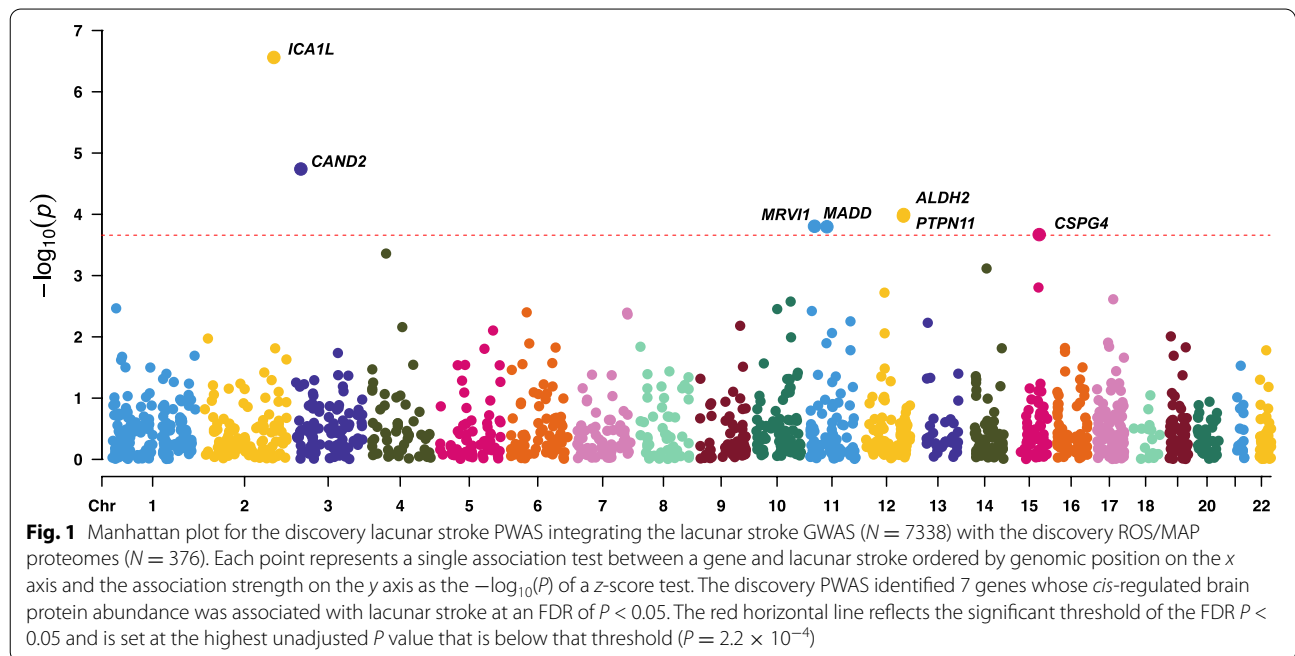
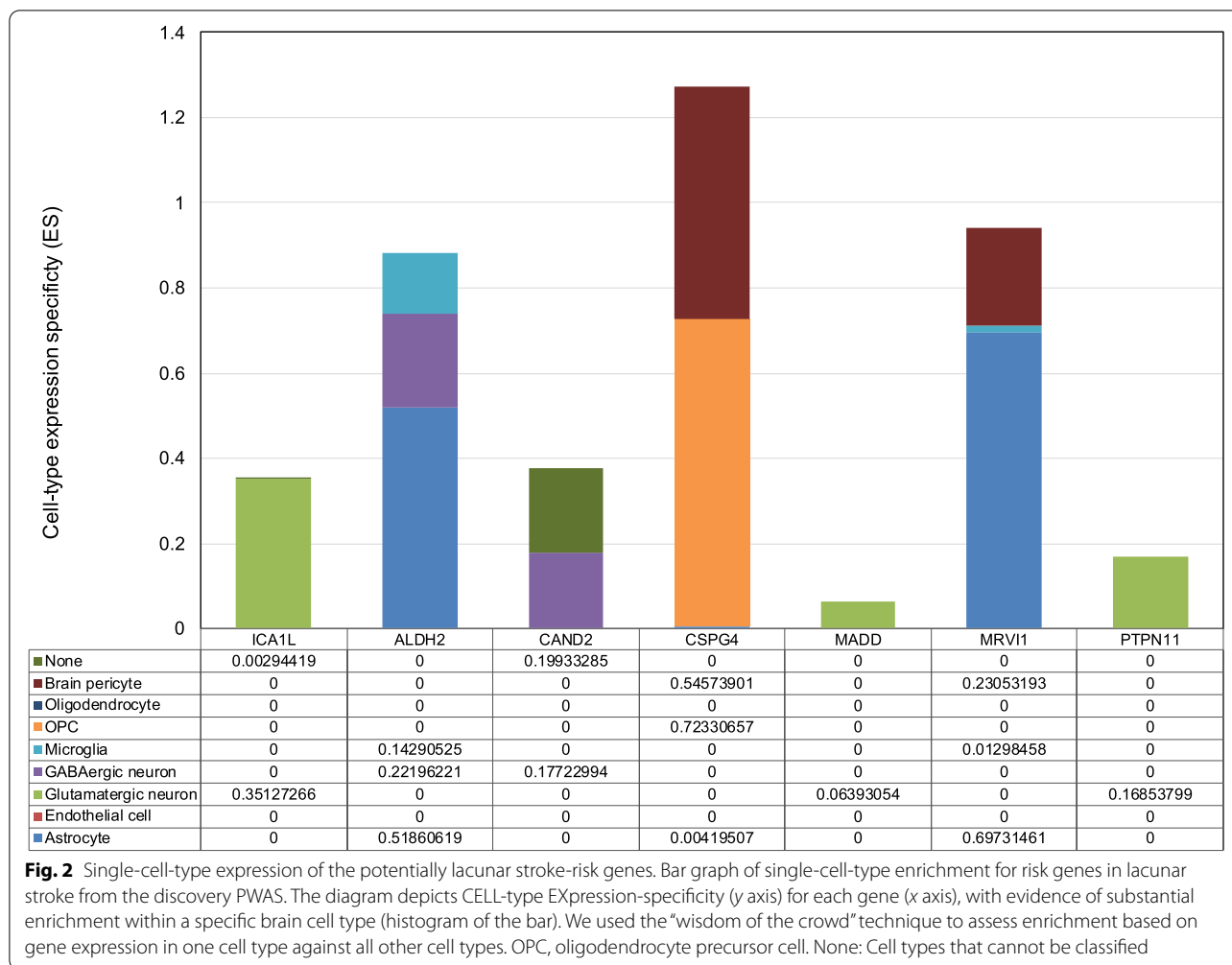


Table 1 The discovery lacunar stroke PWAS identified 7 significant genes, of which 5 were found in the confirmation PWAS, and all 4 replicated

	Gene	Chromosome	Discovery PWAS			Confirmation PWAS		Evidence for replication
			PWAS z-score	PWAS P	PWAS FDR P	PWAS z-score	PWAS P	
1	CAND2	3	-4.281	1.86E-05	1.36E-02	-4.334	1.46E-05	Yes
2	ICA1L	2	-5.136	2.81E-07	4.11E-04	-4.213	2.52E-05	Yes
3	ALDH2	12	3.871	1.09E-04	3.98E-02	3.567	3.61E-04	Yes
4	MADD	11	3.770	1.63E-04	3.98E-02	2.917	3.53E-03	Yes
5	MRVI1	11	-3.776	1.60E-04	3.98E-02	-1.680	9.30E-02	-
6	CSPG4 ^a	15	3.696	2.19E-04	4.58E-02	-	-	-
7	PTPN11 ^a	12	-3.884	1.03E-04	3.98E-02	-	-	-

This table gives the z-scores for the lacunar stroke PWAS associations with their corresponding P values and FDR-adjusted P values for all significant genes in the lacunar stroke discovery PWAS. Confirmation lacunar stroke PWAS z-scores and their corresponding unadjusted P values are provided for the significant genes in the discovery lacunar stroke PWAS

^a Protein not profiled in the confirmation proteomic dataset



whereas *ICA1L*, *PTPN11*, and *MADD* were only found in glutamatergic neurons. GABAergic neurons had higher levels of *CAND2* and *ALDH2*.

MR verify 4 genes associated with lacunar stroke using brain pQTL

Most of the analyzed proteins could only be instrumented using a single SNP; thus, MR estimates were mainly based on the Wald ratio method. We further confirmed four proteins, including *ICA1L*, *CAND2*, *ALDH2*, and *MADD*, biomarkers that revealed significant evidence of a connection in the lacunar stroke GWAS (Table 2).

Colocalization between lacunar stroke risk genes and pQTLs in the brain

Lacunar stroke PWAS associations may arise from a coincidental overlap between pQTLs and sites in linkage disequilibrium with lacunar stroke GWAS sites or from a variant associated with protein expression (the variant is a protein quantitative trait locus (pQTL)) and lacunar

stroke at the same time. Statistical colocalization analysis reported for each gene, the probability that the GWAS and pQTL share a causal variant, referred to as both hypothesis 4 (H_4) and $PP4/(PP3+PP4) \geq 0.75$. Based on a $H_4 \geq 75$ percent and $PP4/(PP3+PP4) \geq 0.75$, this analysis revealed three of the seven genes (*ICA1L*, *CAND2*, and *ALDH2*) that offered evidence of genetic colocalization (Table 2). It suggests that these three proteins play an important role in the pathophysiology of lacunar stroke.

Specificity of the lacunar stroke PWAS results

We did PWAS for other brain-related and biologic traits to understand the specificity of PWAS results for lacunar stroke, and we predicted the degree of overlap of important genes to roughly correlate to their genetic relationship. GWAS results from ischemic stroke ($N = 60,341$) [34], large-artery atherosclerotic stroke ($N = 6688$) [34], brain microbleeds ($N = 3556$) [35], neuroticism ($N = 390,278$) [36], body mass index (BMI; $N = 681,275$) [37], and waist-to-hip ratio adjusting for BMI ($N = 694,649$)

Table 2 Risk genes verified by Mendelian randomization (MR) and colocalization using brain pQTL

	Gene	Beta	SE	P value	Evidence for replication	H ₄	PP4/(PP3+PP4)	Causal variant
1	CAND2	-0.758	0.176	1.71E-05	Yes	0.94	0.99	Yes
2	ICA1L	-2.531	0.493	2.81E-07	Yes	0.99	0.99	Yes
3	ALDH2	1.157	0.314	2.28E-04	Yes	0.75	0.92	Yes
4	MADD	1.161	0.398	3.53E-03	Yes	-	-	-

This table shows the Beta, SE, and P values for the MR. MR P values and a direction of effect consistent with the discovery and replication PWAS results. For the 4 FDR-significant genes in the discovery lacunar stroke PWAS, the result of COLOC H₄, which is the Bayesian posterior probability that a genetic variant is shared by both traits (that is, the genetically regulated protein level and lacunar stroke)

[38] were combined with the discovery proteomic profiles to perform PWAS of each trait. Using FUSION, the PWAS of ischemic stroke identified 4 genes, while the PWAS of large-artery atherosclerotic stroke and brain microbleeds identified none. The PWAS of neuroticism, BMI, and WHRadjBMI, as reported by Wingo and colleagues [37], identified 72, 395, and 244 genes, respectively (FDR $P < 0.05$) (Additional file 1: Table S2-7). As expected, the lacunar stroke PWAS found that 1 in 4 (ALDH2; 25%) ischemic stroke genes overlapped with 7 lacunar stroke PWAS-significant genes, reflecting their high degree of genetic correlation. Two of 72 (2.8%) neuroticism genes (MADD and ICA1L), 4 of 395 (1%) BMI genes (MADD, ICA1L, CSPG4, and PTPN11), and 2 of 244 (0.8%) WHRadjBMI genes (CAND2 and CSPG4) overlapped with 7 lacunar stroke PWAS-significant genes. There were no overlapping genes between large-artery atherosclerotic stroke, brain microbleeds, and lacunar stroke (Fig. 3).

Examination of the potential lacunar stroke-causal proteins at the mRNA level

We combined the lacunar stroke GWAS data with human brain transcriptomes to conduct a lacunar stroke transcriptome-wide association analysis (TWAS) using FUSION. We found that the cis-regulated brain mRNA expression of the seven genes was associated with lacunar stroke (FDR $P < 0.05$) (Additional file 1: Table S8). Interestingly, we found one of the seven genes (*ICA1L*; Table 3; Additional file 1: Table S9) identified in the discovery PWAS, suggesting joint evidence from PWAS and TWAS for its role in lacunar stroke etiology.

Significance of the protein findings

To determine the importance of the 7 potentially causal genes identified from the meta-analysis of the discovery and replication PWAS analyses, we obtained the lowest P values for the SNPs within 1 Mb of each of these 7 genes using the summary statistics from the most extensive lacunar stroke GWAS ($N = 7338$) [14]. The most

significant P values were less than 5×10^{-8} in two genes (*MADD* and *ICA1L*), while the P values of SNPs in the remaining 5 genes ranged from 5.2×10^{-5} to 1.08×10^{-7} (Additional file 1: Table S10). The PWAS findings suggest that specific brain proteins likely contribute to the pathogenesis of lacunar stroke.

Discussion

In the present study, we employed a pipeline of analytical techniques investigating the functional associations between protein biomarkers in the brain and lacunar stroke risk. We identified 7 potential risk genes (*ICA1L*, *CAND2*, *ALDH2*, *MADD*, *MRVII*, *CSPG4*, and *PTPN11*) of lacunar stroke with altered protein abundances in the brain. Four (*ICA1L*, *CAND2*, *ALDH2*, *MADD*) of these 7 genes were replicated in the independent PWAS and MR validation analyses of lacunar stroke, providing a higher confidence level. Furthermore, we identified *ICA1L*, *CAND2*, and *ALDH2* from comprehensive analyses, including non-lacunar stroke brain PWAS and colocalization, and *ICA1L* was supported at the brain transcriptional level. These genes may serve as promising targets for further mechanistic and therapeutic studies.

Identifying therapeutic targets for diseases is a crucial goal of human genetics research and is particularly vital for neurovascular diseases, including lacunar stroke. Our analysis implicated genes previously investigated in lacunar stroke, such as *ICA1L* and *MADD*, as well as new candidates, including *CAND2*, *ALDH2*, *MRVII*, *CSPG4*, and *PTPN11*. Two genes (*ICA1L* and *MADD*) reported in lacunar stroke play roles at the synapse. *ICA1L* encodes a protein triggered by type IV collagen and plays a crucial role in myelination [39]. According to our lacunar stroke PWAS data, *ICA1L* has a lower abundance in the brains of lacunar stroke patients. Furthermore, we discovered that *ICA1L* was enriched in cortical glutamate neurons. Glutamate neurons are crucial components in neural development and neuropathology through their role in cell proliferation, differentiation, survival, and neural network formation. Our findings imply that decreased

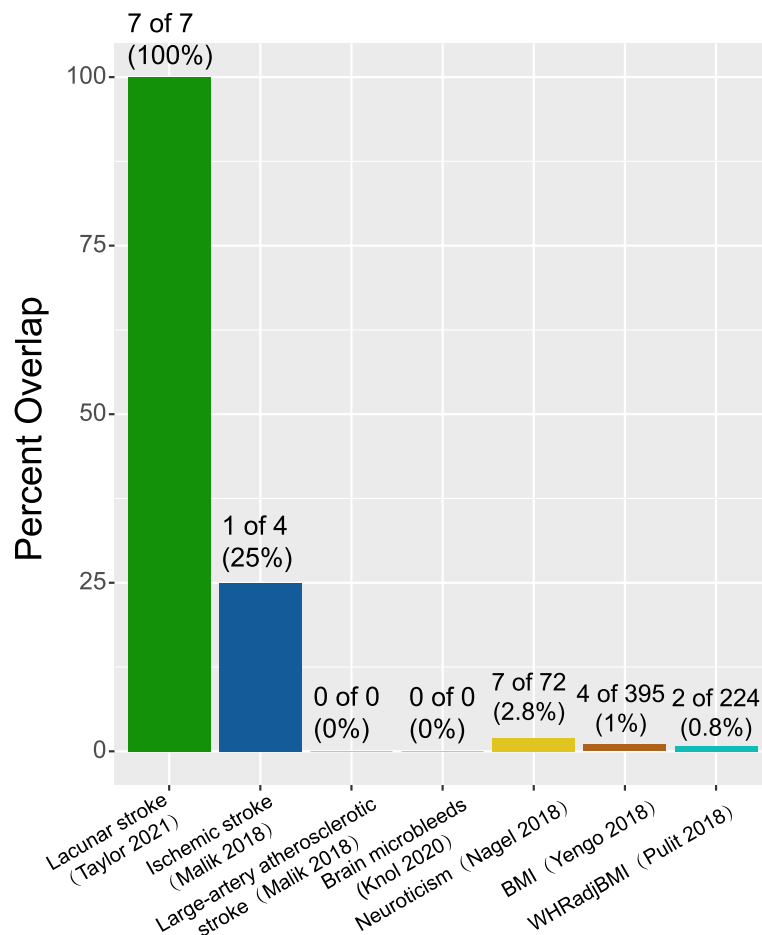


Fig. 3 Overlap of significant genes between lacunar stroke and other traits. Overlap between results of the lacunar stroke PWAS and PWAS for other traits. The PWAS used the discovery ROS/MAP proteomic dataset ($N = 376$) and GWAS summary results. The following outcomes were tested: ischemic stroke ($N = 60,341$), large-artery atherosclerotic stroke ($N = 6688$), brain microbleeds ($N = 3556$), neuroticism ($N = 390,278$), body mass index (BMI; $N = 681,275$), and waist-to-hip ratio adjusting for BMI (WHRadjBMI; $N = 694,649$). Significant genes considered for overlap are those with $FDR P < 0.05$

Table 3 Summary of the 3 lacunar stroke PWAS-significant genes with evidence for being consistent with a causal role in lacunar stroke

	Gene	Chromosome	Discovery PWAS	Confirmation PWAS	Confirmation MR	Evidence for causality COLOC	TWAS significant
1	ICA1L	2	Significant	Replicated	Replicated	Yes	Yes
2	CAND2*	3	Significant	Replicated	Replicated	Yes	N/A
3	ALDH2†	12	Significant	Replicated	Replicated	Yes	No

N/A, refers to genes that did not have significant heritability estimates to be included in the TWAS of lacunar stroke

*New gene refers to genes not within a 1-Mb window of SNPs with $P < 5.23 \times 10^{-8}$ identified in Taylor et al. lacunar stroke GWAS

ICA1L may impair excitatory synaptic signaling and contribute to the pathogenesis of lacunar stroke. *ICA1L* has also been linked to the etiology of lacunar stroke in previous transcriptome investigations [14]. Our findings

show that *MADD* is more abundant in glutamate neurons. We speculate that *MADD* is primarily involved in the transmission of apoptotic signals in neuronal signaling pathways [40], consistent with previous research

suggesting that ischemia causes excitatory glutamate toxicity [41–44].

Other notable molecular roles for the 5 novel genes in lacunar stroke include cerebral cavernous malformations, vascular inflammation, platelet adhesion, and cell apoptosis. *CAND2*, which encodes cullin-associated and neddylation-dissociated 2, plays a role in cerebral cavernous malformations [45]. Cavernous malformation is a key inducing factor in lacunar stroke and cerebral microbleeds [46]. According to our findings, *CAND2* is decreased, predominantly in GABAergic neurons in the brains of lacunar stroke patients, indicating its role in the etiology of lacunar stroke. Both *ALDH2* and *MRVII* are involved in platelet adhesion [47] and vascular inflammation [48, 49]. Previous research has linked increased blood-brain barrier permeability to an inflammatory process involving activated monocytes/macrophages in individuals with cerebral small vessel disease [50, 51]. In our study, *MRVII* was more abundant in astrocytes, which supports their roles in vascular inflammation. *CSPG4*, also known as neuron-gial antigen 2 (NG2) [52], is a protein that helps to stabilize cell-substrate connections [53–55]. Finally, we discovered a novel protein, *PTPN11*, as a new candidate for a membrane protein that suppresses cell growth and induces apoptosis [56–59]. These 7 genes are implicated in the molecular process and neuropathological changes in lacunar stroke.

Most trait-associated variants in neuropsychiatric disease are found in protein-noncoding areas of the human genome, where they have previously been linked to transcriptional levels [60–62]. As such, we applied eQTLs to understand GWAS-related transcriptional regulatory mechanisms in lacunar stroke. However, only the *ICA1L*-identified proteins exhibited changes in gene expression. There could be several reasons for this lack of agreement. First, while the exact link between eQTLs and pQTLs has yet to be discovered, the mRNA expression and protein levels of many genes are uncorrelated, owing in part to various posttranscriptional factors such as sequence characteristics implicated in protein translation and degradation [63]. Second, assay technical artifacts and differences in data analysis may impact the results significantly. While opposed to pQTL analysis [64], eQTL studies use stricter criteria to detect remote regulatory changes, resulting in a lower false-positive rate. In addition to raising thresholds, one way to improve the performance is to use strong tools like *FUSION* [26], *MR* [29], and *COLOC* [26, 30] to check findings with independent samples. To address this difficulty, however, it is essential to expand the depth and variety of multiomics sequencing at the individual level.

Clinical trials have been conducted using drug compounds targeting one of the three causal genes, including

ALDH2 (ranked as high confidence level in our findings), for alcohol dependency and parasite infection (two drugs, phase 4) [65]. Secondary analysis of those and future drugs in clinical trials would likely be helpful to prove the idea that the proteins are involved in the development of lacunar stroke.

Our study has several advantages. First, PWAS of lacunar stroke was conducted using the largest and most comprehensive human proteome and summary statistics from the most recent lacunar stroke GWAS. Second, we performed the replication PWAS using independent human brain proteome and verified the risk proteins with independent MR validation analysis. Third, based on Bayesian colocalization used to estimate the probability that two associated signals were observed at a particular site with a common causal variant, we confirmed the pathogenetic protein (*ICA1L*, *CAND2*, and *ALDH2*) of lacunar stroke. Fourth, this study analyzed both mRNA and protein levels associated with lacunar stroke utilizing both the PWAS and the TWAS. Finally, the dorsolateral prefrontal cortex in the current study was chosen because it includes the cell type most linked to lacunar stroke [14]. Furthermore, the prefrontal cortex has been proposed as a top-down control system that connects other brain areas to facilitate sophisticated cognitive functions. Prefrontal brain risk protein screening for lacunar stroke may help identify critical targets for enhanced cognitive function as well as those who are at high risk of stroke recurrence [2, 66].

The current study has several limitations. First, pQTL and eQTL mapping cannot solve all GWAS signals. At a single level, such as the protein level, the function of genes in the biological development of lacunar stroke is difficult to explain. More epigenetic investigations, based on mQTL, single-cell sequencing, and whole-genome sequencing, are needed to design tailored therapy regimens and offer a complete understanding of the molecular mechanisms implicated in lacunar stroke [67, 68]. Second, the method for detecting Slow Off-rate Modified Aptamers was limited to a subset of proteomes and did not cover the whole proteome. Third, because current proteome samples vary by ethnicity, further expansion of the scale and diversity of brain proteome data can help with more precise estimates and enable its broader applications.

Conclusions

In conclusion, we found strong evidence supporting three novel brain proteins (*ICA1L*, *CAND2*, and *ALDH2*) associated with lacunar stroke. *ICA1L* was further verified at the mRNA level. These findings offer information on the genetic and physiological processes that underpin lacunar stroke, allowing novel therapeutic targets

to be identified. Future research should take advantage of greater large-scale molecular datasets obtained from lacunar stroke-relevant tissues, which might provide unique insights into genetic and functional processes and identify potential druggable targets for new lacunar stroke treatments.

Abbreviations

DLPFC: Dorsolateral prefrontal cortex; eQTL: Expression quantitative trait locus; GWAS: Genome-wide association studies; MR: Mendelian randomization; PGC: Psychiatric Genomics Consortium; pQTL: Protein quantitative trait locus; PWAS: Proteome-wide association studies; SNPs: Single-nucleotide polymorphisms; TWAS: Transcriptome-wide association study.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-022-02408-y>.

Additional file 1: Table S1. The discovery lacunar stroke PWAS identified 7 significant genes. **Table S2.** The PWAS of ischemic stroke integrating the ischemic stroke GWAS ($N = 60,341$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S3.** The PWAS of large-artery atherosclerotic stroke integrating the large-artery atherosclerotic stroke GWAS ($N = 6,688$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S4.** The PWAS of brain microbleeds integrating the brain microbleeds GWAS ($N = 3,556$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S5.** The PWAS of neuroticism integrating the neuroticism GWAS ($N = 390,278$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S6.** The PWAS of BMI integrating the BMI GWAS ($N = 681,275$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S7.** The PWAS of WHRadjBMI integrating the WHRadjBMI GWAS ($N = 694,649$) with ROS/MAP human brain proteomic and genetic data ($N = 376$) using FUSION. **Table S8.** The TWAS of lacunar stroke integrating the lacunar stroke GWAS ($N = 7,338$) with CMC human brain transcriptome and genetic data ($N = 452$) using FUSION. **Table S9.** The lacunar stroke TWAS verified 1 significant gene. **Table S10.** SNPs located within 1 Mb of each of the 7 proteins with the lowest p -value for association with lacunar stroke.

Acknowledgements

The data available in the AD Knowledge Portal would not be possible without the participation of research volunteers and the contribution of data by collaborating researchers. We thank the participants of the ROS, MAP, Mayo, Mount Sinai Brain Bank, and Banner Sun Health Research Institute Brain and Body Donation Program for their time and participation. We thank Dr. Wingo (Division of Mental Health, Atlanta VA Medical Center, Decatur, GA, USA) for making the human brain weights public available. We thank the participants and investigators of CommonMind Consortium.

Authors' contributions

CCZ and TL designed the experiment. CCZ did the PWAS, cell-type specific analysis, and TWAS. XDD provided oversight. CCZ, FQQ, and XJL did the Bayesian colocalization analysis. CCZ and TL verified the data. All authors contributed to, reviewed, and approved the final draft of the paper. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Funding

This work was partly funded by the National Nature Science Foundation of China Key Project (T.L., 81630030 and 81920108018); National Nature Science Foundation of China Project (C.Z., 82001413); Project for Hangzhou Medical Disciplines of Excellence & Key Project for Hangzhou Medical Disciplines (T.L., 202004A11); Introduction Project of Suzhou Clinical Expert Team (X.D. and T.L.,

SZYJTD201715); China Postdoctoral Science Foundation (C.Z., 2020M673247); and Postdoctoral Foundation of West China Hospital (C.Z., 2020HXBH163).

Availability of data and materials

All data relevant to the study are included in the article or uploaded as online supplementary information. The data generated in this study will be available from the corresponding author on reasonable request. GWAS summary statistics from these analyses are available at GWAS Catalog and on the Cerebrovascular Disease Knowledge Portal.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Mental Health Center and Psychiatric Laboratory, the State Key Laboratory of Biotherapy, West China Hospital of Sichuan University, Chengdu, Sichuan, China. ²Department of Neurology, the 3rd Affiliated Hospital of Chengdu Medical College, Chengdu, Sichuan, China. ³Suzhou Psychiatric Hospital, The Affiliated Guangji Hospital of Soochow University, Suzhou, Jiangsu, China. ⁴Department of Neurobiology, Affiliated Mental Health Center & Hangzhou Seventh People's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. ⁵NHC and CAMS Key Laboratory of Medical Neurobiology, MOE Frontier Science Center for Brain Science and Brain-machine Integration, School of Brain Science and Brain Medicine, Zhejiang University, Zhejiang, Hangzhou, China.

Received: 8 January 2022 Accepted: 17 May 2022

Published online: 23 June 2022

References

1. Wardlaw JM. What causes lacunar stroke? *J Neurol Neurosurg Psychiatry*. 2005;76(5):617–9.
2. Regenhardt RW, Das AS, Lo EH, Caplan LR. Advances in understanding the pathophysiology of lacunar stroke: a review. *JAMA Neurol*. 2018;75(10):1273–81.
3. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467–84.
4. Yang C, Farias FHG. Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. *Nat Neurosci*. 2021;24(9):1302–12.
5. Sharma K, Schmitt S, Bergner CG, Tyanova S, Kannaiyan N, Manrique-Hoyos N, et al. Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci*. 2015;18(12):1819–31.
6. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet*. 2012;13(4):227–32.
7. Rolland DCM, Basrur V, Jeon Y-K, McNeil-Schwalm C, Fermin D, Conlon KP, et al. Functional proteogenomics reveals biomarkers and therapeutic targets in lymphomas. *Proc Natl Acad Sci*. 2017;114(25):6581–6.
8. Yan Z, Wang S. Proteoglycans as therapeutic targets in brain cancer. *Front Oncol*. 2020;10:1358.
9. Asoh S, Ohsawa I, Mori T, Katsura K-I, Hiraide T, Katayama Y, et al. Protection against ischemic brain injury by protein therapeutics. *Proc Natl Acad Sci*. 2002;99(26):17107–12.
10. Kuehner JN, Bruggeman EC, Wen Z, Yao B. Epigenetic regulations in neuropsychiatric disorders. *Front Genet*. 2019;10:268.
11. Imamura A, Morimoto Y, Ono S, Kurotaki N, Kanegae S, Yamamoto N, et al. Genetic and environmental factors of schizophrenia and autism

- spectrum disorder: insights from twin studies. *J Neural Transm (Vienna)*. 2020;127(11):1501–15.
12. Wingo TS, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al. Brain proteome-wide association study implicates novel proteins in depression pathogenesis. *Nat Neurosci*. 2021;24(6):810–7.
 13. Traylor M, Malik R, Nalls MA, Cotlarciuc I, Radmanesh F, Thorleifsson G, et al. Genetic variation at 16q24. 2 is associated with small vessel stroke. *Ann Neurol*. 2017;81(3):383–94.
 14. Traylor M, Persyn E, Tomppo L, Klasson S, Abedi V, Bakker MK, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol*. 2021;20(5):351–61.
 15. Wingo AP, Dammer EB, Breen MS, Logsdon BA, Duong DM, Troncosco JC, et al. Large-scale proteomic analysis of human brain identifies proteins associated with cognitive trajectory in advanced age. *Nat Commun*. 2019;10(1):1619.
 16. Timp W, Timp G. Beyond mass spectrometry, the next step in proteomics. *Sci Adv*. 2020;6(2):eaax8978.
 17. Ou Y-N, Yang Y-X, Deng Y-T, Zhang C, Hu H, Wu B-S, et al. Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. *Mol Psychiatry*. 2021;26:6065–73.
 18. Liu J, Li X, Luo X-J. Proteome-wide association study provides insights into the genetic component of protein abundance in psychiatric disorders. *Biol Psychiatr*. 2021;90(11):781–9.
 19. Wang M, Beckmann ND, Roussos P, Wang E, Zhou X, Wang Q, et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. *Sci Data*. 2018;5(1):1–16.
 20. De Jager PL, Ma Y, McCabe C, Xu J, Vardarajan BN, Felsky D, et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci Data*. 2018;5:180142.
 21. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet*. 2021;53(2):143–6.
 22. Beach TG, Adler CH, Sue LI, Serrano G, Shill HA, Walker DG, et al. Arizona study of aging and neurodegenerative disorders and brain and body donation program. *Neuropathology*. 2015;35(4):354–89.
 23. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci*. 2016;19(11):1442–53.
 24. Li M, Huang L, Grigoriou-Serbanescu M, Bergen SE, Landén M, Hultman CM, et al. Convergent lines of evidence support LRP8 as a susceptibility gene for psychosis. *Mol Neurobiol*. 2016;53(10):6608–19.
 25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
 26. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245–52.
 27. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89–98.
 28. Rasooly D, Patel CJ. Conducting a reproducible Mendelian randomization analysis using the R analytic statistical environment. *Curr Protoc Hum Genet*. 2019;101(1):e82.
 29. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.
 30. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383.
 31. Su X, Li W, Lv L, Li X, Yang J, Luo X-J, et al. Transcriptome-wide association study provides insights into the genetic component of gene expression in anxiety. *Front Genet*. 2021;12:740134.
 32. Timshel PN, Thompson JJ, Pers TH. Genetic mapping of etiologic brain cell types for obesity. *Elife*. 2020;9:e55851.
 33. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds for robust gene network inference. *Nat Methods*. 2012;9(8):796–804.
 34. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50(4):524–37.
 35. Knol MJ, Lu D, Traylor M, Adams HHH, Romero JRJ, Smith AV, et al. Association of common genetic variants with brain microbleeds: a genome-wide association study. *Neurology*. 2020;95(24):e3331–43.
 36. Nagel M, Jansen PR, Stringer S. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*. 2018;50(7):920–7.
 37. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet*. 2018;27(20):3641–9.
 38. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet*. 2019;28(1):166–74.
 39. Liu M, Xu P, Guan Z, Qian X, Dockery P, Fitzgerald U, et al. Ulk4 deficiency leads to hypomyelination in mice. *Glia*. 2018;66(1):175–90.
 40. Akiyama K, Liang YQ, Isono M, Kato N. Investigation of functional genes at homologous loci identified based on genome-wide association studies of blood lipids via high-fat diet intervention in rats using an in vivo approach. *J Atheroscler Thromb*. 2015;22(5):455–80.
 41. Centeno C, Repici M, Chatton JY, Riederer BM, Bonny C, Nicod P, et al. Role of the JNK pathway in NMDA-mediated excitotoxicity of cortical neurons. *Cell Death Differ*. 2007;14(2):240–53.
 42. Belov Kirdajova D, Kriska J, Tureckova J, Anderova M. Ischemia-triggered glutamate excitotoxicity from the perspective of glial cells. *Front Cell Neurosci*. 2020;14:51.
 43. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet*. 2009;41(1):47–55.
 44. Wang XB, Han YD, Cui JJ, Yang J, Huang ZL, et al. Associations of lipid levels susceptibility loci with coronary artery disease in Chinese population. *Lipids Health Dis*. 2015;14:80.
 45. Koskimäki J, Zhang D, Li Y, Saadat L, Moore T, Lightle R, et al. Transcriptome clarifies mechanisms of lesion genesis versus progression in models of Ccm3 cerebral cavernous malformations. *Acta Neuropathol Commun*. 2019;7(1):132.
 46. Wang Q, Meng L, Wang Z. Combined cerebral microbleeds with lacunar infarctions in familial cerebral cavernous malformations. *JAMA Neurol*. 2019;76(9):1117–8.
 47. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler G, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. *Nat Genet*. 2010;42(7):608–13.
 48. Tsai SH, Hsu LA, Tsai HY, Yeh YH, Lu CY, Chen PC, et al. Aldehyde dehydrogenase 2 protects against abdominal aortic aneurysm formation by reducing reactive oxygen species, vascular inflammation, and apoptosis of vascular smooth muscle cells. *FASEB J*. 2020;34(7):9498–511.
 49. Kessler T, Schunkert H, von Hundelshausen P. Novel approaches to fine-tune therapeutic targeting of platelets in atherosclerosis: a critical appraisal. *Thromb Haemost*. 2020;120(11):1492–504.
 50. Rouhl RP, Damoiseaux JG, Lodder J, Theunissen RO, Knottnerus IL, Staals J, et al. Vascular inflammation in cerebral small vessel disease. *Neurobiol Aging*. 2012;33(8):1800–6.
 51. Acampa M, Lazzarini PE, Manfredi C, Guideri F, Tassi R, Domenichelli C, et al. Non-stenosing carotid atherosclerosis and arterial stiffness in embolic stroke of undetermined source. *Front Neurol*. 2020;11:725.
 52. Nishiyama A, Dahlin KJ, Prince JT, Johnstone SR, Stallcup WB. The primary structure of NG2, a novel membrane-spanning proteoglycan. *J Cell Biol*. 1991;114(2):359–71.
 53. Fung K, Ramirez J, Warren HR, Aung N, Lee AM, Tzani E, et al. Genome-wide association study identifies loci for arterial stiffness index in 127,121 UK Biobank participants. *Sci Rep*. 2019;9(1):9143.
 54. Santoro C, Giugliano T, Kraemer M, Torella A, Schwitalla JC, Cirillo M, et al. Whole exome sequencing identifies MRVI1 as a susceptibility gene for moyamoya syndrome in neurofibromatosis type 1. *PLoS One*. 2018;13(7):e0200446.
 55. Rudzik R, Dziedzicko P, Rać ME, Sawczuk M, Maciejewska-Skrendo A, Safranow K, et al. Polymorphisms in GP6, PEAR1A, MRVI1, PIK3CG,

- JMJD1C, and SHH genes in patients with unstable angina. *Int J Environ Res Public Health*. 2020;17(20):7506.
56. Liu KW, Feng H, Bachoo R, Kazlauskas A, Smith EM, Symes K, et al. SHP-2/PTPN11 mediates gliomagenesis driven by PDGFRA and INK4A/ARF aberrations in mice and humans. *J Clin Investig*. 2011;121(3):905–17.
 57. Sang Y, Hou Y, Cheng R, Zheng L, Alvarez AA, Hu B, et al. Targeting PDGFR α -activated glioblastoma through specific inhibition of SHP-2-mediated signaling. *Neuro-Oncology*. 2019;21(11):1423–35.
 58. Roccograndi L, Binder ZA, Zhang L, Aceto N, Zhang Z, Bentires-Alj M, et al. SHP2 regulates proliferation and tumorigenicity of glioma stem cells. *J Neuro-Oncol*. 2017;135(3):487–96.
 59. Yang Z, Li Y, Yin F, Chan RJ. Activating PTPN11 mutants promote hematopoietic progenitor cell-cycle progression and survival. *Exp Hematol*. 2008;36(10):1285–96.
 60. Huo Y, Li S, Liu J, Li X, Luo XJ. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun*. 2019;10(1):670.
 61. Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, et al. Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nat Genet*. 2018;50(4):538–48.
 62. Kaiser J, Maibach M, Salpeter I, Hagenbuch N, de Souza VBC, Robinson MD, et al. The spinal transcriptome after cortical stroke: in search of molecular factors regulating spontaneous recovery in the spinal cord. *J Neurosci*. 2019;39(24):4714–26.
 63. Battle A, Khan Z, Wang SH, et al. Genomic variation. Impact of regulatory variation from RNA to protein. *Science*. 2015;347(6222):664–7.
 64. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. *Nat Rev Genet*. 2015;16(4):197–212.
 65. Carvalho-Silva D, Pierleoni A, Pignatelli M, Ong C, Fumis L, Karamanis N, et al. Open Targets Platform: new developments and updates two years on. *Nucleic Acids Res*. 2019;47(D1):D1056–d1065.
 66. Kwan A, Wei J, Dowling NM, Power MC, Nadareishvili Z. Cognitive impairment after lacunar stroke and the risk of recurrent stroke and death. *Cerebrovasc Dis*. 2021;50(4):383–9.
 67. Clark SJ, Lee HJ, Smallwood SA, Kelsey G, Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol*. 2016;17:72.
 68. Zhao T, Hu Y, Zang T, Wang Y. Integrate GWAS, eQTL, and mQTL Data to Identify Alzheimer's Disease-Related Genes. *Front Genet*. 2019;10:1021.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

