Shared genes between Alzheimer’s disease and ischemic stroke

Chang-Juan Wei1,2 | Pan Cui1,2 | He Li1,2 | Wen-Jing Lang1,2 | Gui-You Liu3 | Xiao-Feng Ma1,2

1Department of Neurology, Tianjin Medical University General Hospital, Tianjin, China
2Tianjin Neurological Institute, Key Laboratory of Post-neurotrauma Neuro-repair and Regeneration in Central Nervous System, Ministry of Education, Tianjin, China
3School of Life Science and Technology, Harbin Institute of Technology, Harbin, China

Correspondence
Gui-You Liu, School of Life Science and Technology, Harbin Institute of Technology, Harbin, China.
Email: liuguiyou1981@163.com
and
Xiao-Feng Ma, Department of Neurology, Tianjin Medical University General Hospital, Tianjin, China.
Email: xfengma@tmu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 81300945, 81401361 and 81870954

Summary
Aims: Although converging evidence from experimental and epidemiological studies indicates Alzheimer’s disease (AD) and ischemic stroke (IS) are related, the genetic basis underlying their links is less well characterized. Traditional SNP-based genome-wide association studies (GWAS) have failed to uncover shared susceptibility variants of AD and IS. Therefore, this study was designed to investigate whether pleiotropic genes existed between AD and IS to account for their phenotypic association, although this was not reported in previous studies.

Methods: Taking advantage of large-scale GWAS summary statistics of AD (17,008 AD cases and 37,154 controls) and IS (10,307 IS cases and 19,326 controls), we performed gene-based analysis implemented in VEGAS2 and Fisher’s meta-analysis of the set of overlapped genes of nominal significance in both diseases. Subsequently, gene expression analysis in AD- or IS-associated expression datasets was conducted to explore the transcriptional alterations of pleiotropic genes identified.

Results: 16 AD-IS pleiotropic genes surpassed the cutoff for Bonferroni-corrected significance. Notably, MS4A4A and TREM2, two established AD-susceptibility genes showed remarkable alterations in the spleens and brains afflicted by IS, respectively. Among the prioritized genes identified by virtue of literature-based knowledge, most are immune-relevant genes (EPHA1, MS4A4A, UBE2L3 and TREM2), implicating crucial roles of the immune system in the pathogenesis of AD and IS.

Conclusions: The observation that AD and IS had shared disease-associated genes offered mechanistic insights into their common pathogenesis, predominantly involving the immune system. More importantly, our findings have important implications for future research directions, which are encouraged to verify the involvement of these candidates in AD and IS and interpret the exact molecular mechanisms of action.

KEYWORDS
Alzheimer’s disease, gene expression analysis, gene-based analysis, genome-wide association studies, ischemic stroke, pleiotropic genes, Versatile Gene-based Association Study-2 version 2
1 | INTRODUCTION

Alzheimer’s disease (AD) is the world’s leading cause of dementia. The hallmarks of AD are extracellular amyloid-β peptide (Aβ) accumulation and intracellular neurofibrillary tangles (NFTs), the latter of which is formed by hyperphosphorylated tau protein. Along with the notorious reputation as the second leading cause of mortality and disability worldwide,1-2 stroke is another major cause of age-related cognitive decline and dementia.3 As the most prevalent form, ischemic stroke (IS) accounts for ~85% of stroke incidents.3 Collectively, AD and IS both exert a large burden on global public healthcare and clinical practice.

Growing evidence indicates that there are links between AD and IS. Firstly, emerging epidemiologic research shows that AD is associated with considerable increased risk of IS,4,5 and vice versa.6 Secondly, neuropathological studies show that cerebrovascular lesions frequently coexist with AD pathology.7 The two mixed pathologies act synergistically in increasing the odds of clinical dementia.8 Indeed, a handful of studies have reported that brain ischemia is a non-neglectable factor driving the development of AD through dysregulated expression of AD-associated genes, such as Aβ precursor processing genes and tau protein gene.9,10 Lastly, tau protein, a core hallmark of AD, can exacerbate brain injury in experimental IS through tau-mediated iron export and excitotoxicity.11,12 Taken together, we hypothesized that there might be a shared genetic basis underlying these connections between AD and IS.

Genome-wide association studies (GWAS) have yielded new insights into the genetics of AD13,14 and IS.15,16 Shared genetic variants between AD and IS or its subtypes, have been first determined by Traylor et al.17 They tested whether established genome-wide single nucleotide polymorphisms (SNPs) for AD or IS were significantly associated with the other disease. Yet no such variants have been found. Conventional GWAS methods just focus on significant SNPs judging by overly stringent criterion (P < 5.00E−08) when exploring the genome. There is an emerging consensus that, however, complex diseases are mostly driven by the joint action of a large proportion of SNPs having modest effects well below genome-wide significance.18 Alternatively, gene-based analysis can obtain more validated associations by combining the effects of all SNPs within corresponding genes, thus expand knowledge about genetic architectures of complex diseases.

Hence, in our present study, we performed gene-based association tests to identify potential candidate genes shared between AD and IS. Next, gene expression analyses were conducted to evaluate shared genes’ expression alterations in AD and IS brains or peripheral blood versus matched controls. Furthermore, to further interpret the molecular mechanisms that underpin AD and IS, systematic dissection of individual genes’ functionalities was conducted.

2 | MATERIALS AND METHODS

2.1 | Samples

The GWAS statistics data for AD and IS analyses were from the International Genomics of Alzheimer’s Project (IGAP)13 and the METASTROKE consortium of the International Stroke Genetics Consortium,15 respectively. Both GWAS datasets are based on populations of European descent imputed to the 1000 Genomes Project (1000G) reference panel. Poorly genotyped or imputed SNPs and SNPs with a minor allele frequency (MAF) of less than 0.01 were filtered out in both datasets.

IGAP is a large-scale, two-stage GWAS (for more details, see the research by Lambert et al15). In stage 1, a meta-analysis of four published GWAS samples comprising 17,008 AD cases and 37,154 controls was conducted. After quality control, 7,055,881 SNPs were available for analysis. In stage 2, 11,632 SNPs showing moderate evidence of association (P < 1.00E−03) in stage 1 were genotyped and tested for replication in an independent sample totaling 8,572 cases and 11,312 controls. Lastly, a meta-analysis combining results from stages 1 and 2 was performed. For the present study, we used only summary data from stage 1.

The METASTROKE collaboration genotyped and imputed approximately 9 million SNPs from a meta-analysis of 12 independent GWAS comprising 10,307 IS cases and 19,326 controls. Differing from prior IS GWAS data imputed to the HapMap panel, which comprised up to 2.5 million SNPs with MAF more than 0.05, this expanded set of SNPs informed by 1000G project also included low-frequency variants (MAF 0.01-0.05) and totaled 8.3 million quality-controlled SNPs for analysis. For more detailed information, refer to the original study.15

2.2 | Statistical analyses

2.2.1 | VEGAS2 method

Using the GWAS summary data for AD and IS, we performed a gene-based association test implemented in an updated version of Versatile Gene-based Association Study-2 version 2 (VEGAS2).19 Among various methods of gene-based analysis, VEGAS2 is particularly feasible for analyzing GWAS summary statistics where individual-level genotypic and phenotypic data are unavailable. By uploading the individual SNPs’ IDs and their association P-values, VEGAS2 sums the effects of all the SNPs within a gene and corrects for linkage disequilibrium (LD) referring to 1000G reference set and thus, generates a gene-based test statistic by doing simulations from the multivariate normal distribution. The simulation approach is computationally more efficient than other methods that rely on permutations, such as PLINK, minSNP.20 The default “symmetric boundaries ±0 kb outside gene and SNPs in LD above r² = 0.8” was chosen to define gene boundaries, which meant that the effects of SNPs within a gene, also outside of the gene with r² > 0.8 with the ones in the gene, were considered to calculate gene-based P-values. This option both took account into the effects of nearby regulatory SNPs and reduced the non-specificity caused by large boundaries like ±50 kb.

To reduce the possibility of a single disease driving the cross-disease associations and uncover truly pleiotropic genes shared by AD and IS, we focused on shared genes that were nominally significant in each disease (P_{AD} < 0.05 and P_{IS} < 0.05).
### TABLE 1  Gene-based association results from gene-based meta-analyses of Alzheimer's disease (AD) and ischemic stroke (IS)

| Region       | Gene      | AD             | IS               |
|--------------|-----------|----------------|------------------|
|              |           | **P value**    | **Top-SNP**      | **Top-SNP P value** | **P combined** |
| 7q34         | ZYX       | 5.10E−04       | rs6464548        | 1.62E−06             |               |
| 7q34-q35     | EPHA1     | 1.00E−06       | rs10808026       | 1.42E−11             |               |
| 15q25.2      | EFTUD1    | 1.09E−04       | rs905450         | 2.82E−06             |               |
| 11q12.2      | MS4A4A    | 1.50E−05       | rs55777218       | 3.94E−12             |               |
| 22q11.21     | YD JC     | 2.01E−04       | rs2298428        | 7.28E−05             |               |
| 22q11.21     | UBE2L3    | 2.72E−04       | rs12168746       | 8.82E−05             |               |
| 8q22.3       | PABPC1    | 4.76E−03       | rs1693547        | 9.19E−04             |               |
| 16p12.2      | RRN3P1    | 1.56E−03       | rs12600118       | 3.92E−05             |               |
| 15q25.2      | FAM154B   | 2.56E−04       | rs28522807       | 2.86E−04             |               |
| 17q25.1      | SLC16A5   | 7.22E−04       | rs8078881        | 2.32E−04             |               |
| 12q24.13     | HECTD4    | 3.68E−02       | rs147910225      | 2.31E−04             |               |
| 5q31.3       | ANKHD1-EIF4EBP3 | 8.61E−04 | rs78028717     | 2.52E−04             |               |
| 8p23.1       | PINX1     | 1.23E−03       | rs7014168        | 3.60E−04             |               |
| 22q11.23     | SLC2A11   | 1.12E−03       | rs5760076        | 3.91E−05             |               |
| 12q24.13     | OA52      | 2.50E−04       | rs1635142        | 4.27E−05             |               |
| 5q31.3       | ANKHD1    | 1.18E−03       | rs78028717       | 2.52E−04             |               |
| 6p21.1       | TREM2     | 1.52E−03       | rs6933067        | 1.07E−03             |               |

Top-SNP, most significant SNP from each gene. Significantly associated genes whose P-values in both phenotypes were below 0.001 were shown in bold.
2.2.2 | Meta-analysis using Fisher’s method

We used Fisher’s method to combine the P-values calculated by the VEGAS2 for each gene shared by AD and IS. For a given gene, the Fisher formula for meta-analysis is:

\[ x^2 = -2 \sum_{i=1}^{k} \ln (P_i) \]

where \( P_i \) are the P-values of the genes in the \( i \)-th study and \( k \) is the total number of studies. \( x^2 \) follows a chi-square distribution with \( 2k \) degrees of freedom.\(^{21}\) The gene-based meta-analysis was carried out using the R software package.

To avoid false positive signals, we applied the stringent Bonferroni correction accounting for the number of genes and phenotypes tested, that is, the significance threshold was set at \( 0.05/2N \), where \( N \) represented the number of shared genes with nominal significance in both AD and IS.

2.2.3 | Gene expression analysis method

To explore the expression alterations of shared genes in each disease, we surveyed the expression datasets of AD and IS from the Gene Expression Omnibus (GEO) repository. Gene expression analysis was mainly restricted to brains and peripheral blood, as they are the most affected by AD- or IS-associated pathology.

Because AD-associated neuropathology shows regional specificity, expression profiles from discrete brain regions are more informative for discerning AD molecular signatures than analyses based on whole-brain expression data. Thus, we examined expression data from separate regions from postmortem human AD and control brains, including the dorsolateral prefrontal cortex (PFC) and hippocampus regions. The former (GSE44770) sampled 549 brains of 376 late-onset AD patients and 173 controls.\(^ {22}\) The latter (GSE48350) comprised 19 AD cases versus 43 matched controls.\(^ {23}\)

For IS, samples of brains and other tissues (eg spleen) from patients were generally not available. Considering that the core features of IS pathophysiology in rodents and humans are analogous, we included the expression data from peri-infarct brain areas of rats (GSE55260).\(^ {24}\) We also analyzed transcriptional data from peripheral blood collected within 24 hours of stroke onset in 39 IS patients and 24 controls (GSE16561).\(^ {25}\) Additionally, we explored transcriptional profiles from mouse spleens (GSE70841) as spleen is the major lymphoid organ involved in the inflammatory milieu secondary to brain ischemia. All gene aliases in rat or mouse were transformed to the official symbols corresponding to human genes.

If there were multiple transcripts within the same gene, the one with the smallest \( P \)-value was selected. Due to between-study heterogeneity, not all transcripts of AD-IS pleiotropic genes appeared in each dataset. The differential expression was determined by Bonferroni correction accounting for the number of shared genes present in each dataset (\( P = 0.05/n, n \leq 16 \)).

3 | RESULTS

3.1 | Gene-based testing for risk genes of AD and IS

Firstly, the VEGAS2 method was applied to individual GWAS summary data from AD and IS. For AD, 34 genes exceeded gene-wide significance (\( P < 2.35E−6 \) for 21,244 gene tests). Apart from 20 genes at 19q13.31-q13.32 harboring the well-known APOE locus, we confirmed 12 genes (CD33, ABCA7, HLA-DRB6, MS4A2/MS4A6A, EPHA1, PICALM, CR1, CLU, MTCH2/SLC39A13, and BIN1) within 10 established risk loci (CD33, ABCA7, HLA-DRB5-DRB1 region, MS4A locus, EPHA1, PICALM, CR1, CLU, CELF1, and BIN1).\(^ {13}\) Considering the other 39 genes outside the APOE locus passing a loose significance threshold (\( P < 1.00E−4 \)), 23 genes belonged to the established risk loci, with 4 more loci validated (CD2AP, PTK2B, SORL1, and SLC24A4).\(^ {13}\) The well-replicated results showed the good performance of VEGAS2 for gene-based analysis. For the full results, see Supporting Information Table S1.

For IS, no genes of gene-wide significance were recognized (\( P < 2.28E−06 \) for 21,913 gene tests), while two genes surpassed a loose significance threshold (\( P < 1.00E−04 \)): ZYX (\( P = 2.70E−05 \)) and NCR3LG1 (\( P = 7.30E−05 \)).

3.2 | Gene-based testing for shared genes between AD and IS

After overlapping 1915 AD genes and 1288 IS genes with nominal significance, 130 shared genes remained for further Fisher’s meta-analysis. We identified 17 genes reaching the significance threshold at \( P_{\text{combined}} < 1.92E−4 [0.05/(2*130)] \) (Table 1), some of which were located in adjacent positions, that is, ZYX/EPHA1 at 7q34-7q35, EFTUD1/FAM154B at 15q25.2, YDJC/UBE2L3 and SLC2A11 at 22q11, HECTD4/OAS2 at 12q24.13, and ANKHDI1-EIF4EBP3/ANKHDI1 at 5q31.3. The latter was regarded as one gene as they were likely to arise from the same SNPs signals as ANKHDI1-EIF4EBP3 is a fusion transcript of ANKHDI1 and its downstream EIF4EBP3. It’s worth noting that ZYX and EFTUD1 even reached \( P < 0.001 \) in both AD and IS datasets. For the full results, refer to Supporting Information Table S2.

3.3 | Gene expression analyses of shared genes

To validate the relevance of the 16 AD-IS genes, we evaluated their expressions in brains or peripheral blood of AD or IS cases versus controls. Importantly, MS4A4A and TREM2, two established AD-susceptibility genes, showed remarkable alterations in the spleens and brains afflicted by IS, respectively. A comparison of differentially expressed genes in each dataset was presented in Table 2.

AD-associated expression profiles from the PFC region (Supporting Information Table S3A) revealed that 8 genes were differentially expressed, with MS4A4A, UBE2L3, TREM2, and HECTD4 being the most significant, while just 3 genes’ expression
levels (HECTD4, YDJC, and PABPC1) were altered in the hippocampal regions compared with control brains (Supporting Information Table S3B). In peri-infarcted rat brains (Supporting Information Table S4A), we observed expression changes of 4 genes (TREM2, PABPC1, ZYX, and YDJC). Notably, TREM2 showed nearly 4-fold upregulation versus controls (log2-fold change = 1.99). Transcriptional changes of four genes (MS4A4A, ZYX, EPHA1 and SLC2A11) were present in the peripheral blood of IS patients (Supporting Information Table S4B). Particularly, differential expression of MS4A4A was even more evident in mouse spleens, with over 5-fold increase (log2-fold change = 2.37; Supporting Information Table S4C). It’s necessary to point out that, however, the transcriptional data from rat brains and mouse spleens were not compelling enough due to the small sample size. More robust transcriptional datasets are needed to further validate the candidate genes.

### 4 | DISCUSSION

Different from conventional SNP-based GWAS studies, our research used VEGAS2 gene-based association test to detect pleiotropic genes jointly associated with AD and IS. To avoid the problem of joint effects arising from a dominant association with one single disease, we focused on shared genes with nominal significance in both AD and IS (P < 0.05). By this criterion, 16 genes survived the stringent Bonferroni correction. Next, they were screened for differential expression in AD and IS cases versus controls. To search for supportive evidence for their relevance to AD and IS, we paid extra attention to the molecular functions of these genes.

#### 4.1 ZYX and EPHA1 gene at 7q34-7q35

Zyxin (ZYX) encodes a zinc-binding adaptor protein that translocates from focal adhesions to the nucleus to conduct signal transduction and modulate gene expression. Recently, zyxin has been identified as a novel target of Aβ metabolism in AD.\(^26\) Besides, Zyxin is a novel interacting partner of SIRT1\(^27\) that is protective against aging-associated pathologies like AD\(^28\) and IS.\(^29\) Here, Zyxin was jointly associated with AD and IS (P\(_{\text{combined}} = 2.63E^{-07}\) with P < 0.001 in each disease. Of note was that ZYX showed the most significant association (P = 2.70E-05) in the gene-based analysis of IS.

EPHA1 is an established risk locus of AD, and our gene-based analysis confirmed its gene-wide association with AD (P = 1.00E-06). Ephrin type-A receptor 1 (EPHA1) belongs to the eph receptor subfamily that is the largest family of receptor tyrosine kinases, mediating axonal guidance, synaptic plasticity, and

### TABLE 2 Expression changes of shared genes in distinct expression datasets for Alzheimer’s disease (AD) and ischemic stroke (IS)

| Gene   | AD                        | IS                        |
|--------|---------------------------|---------------------------|
|        | GSE44770                  | GSE48350                  |
| ZYX    | 2.41E-01                  | 2.13E-01                  |
| EPHA1  | 6.12E-01                  | 4.26E-02                  |
| EFTUD1 | 2.48E-01                  | 4.15E-02                  |
| MS4A4A | 1.08E-24 **               | 3.87E-03**                |
| YDJC   | –                         | 1.14E-04                  |
| UBE2L3 | 3.24E-20                  | 4.45E-01                  |
| PABPC1 | 1.90E-11                  | 2.99E-04                  |
| RRN3P1 | –                         | 2.69E-01                  |
| FAM154B| –                         | 3.14E-01                  |
| SLC16A5| 4.93E-10                  | 3.15E-01                  |
| HECTD4 | 8.74E-16                  | 5.17E-05                  |
| ANKHD1-EF4EBP3 | 4.30E-01 | 2.05E-01 |
| PINX1  | 2.46E-04                  | 3.13E-01                  |
| SLC2A11| 1.41E-11                  | 5.39E-03                  |
| OAS2   | 1.13E-01                  | 1.46E-01                  |
| TREM2  | 6.80E-20                  | 1.34E-01                  |

**Significance threshold**

| AD        | IS                        |
|-----------|---------------------------|
| GSE55260  | GSE16561                  |
| Brain     | Peripheral blood          |
|          | Spleen                    |
| 8.43E-04  | 3.38E-04                  |
| 9.48E-01  | 1.22E-03                  |
| 4.79E-02  | 5.31E-02                  |
| –         | 2.73E-05                  |
| 7.58E-03  | 2.17E-01                  |
| 7.32E-01  | 5.68E-02                  |
| 4.17E-05* | 7.40E-02                  |
| –         | 5.27E-01                  |
| 5.54E-02  | 8.80E-01                  |
| 5.13E-02  | 3.64E-02                  |
| 9.75E-02  | 4.15E-02                  |
| 2.03E-01  | 8.44E-01                  |
| –         | 2.55E-03                  |
| 1.29E-01  | 2.04E-01                  |
| 9.29E-06* | 5.37E-01                  |
| 3.85E-03  | 3.57E-03                  |
| 7.10E-01  | 5.56E-03                  |

GEO accession, GSE44772, GSE48350, GSE55260, GSE16561 and GSE70841; PFC, dorsolateral prefrontal cortex; “—”, (no data) meant that the transcripts of this gene did not exist in corresponding datasets. Bolded P-values of genes achieved Bonferroni-corrected significance, adjusted for the number of shared genes present in each expression dataset (0.05/n, n ≤ 16). *Gene expression level showing more than 2 folds upregulation or downregulation compared to controls. **Genes near the significance threshold.
cell-to-cell communication in the central nervous system (CNS). Moreover, EPHA1 can modulate leukocyte extravasation, chemotaxis, and inflammatory cell migration. Indeed, ample evidence has implicated the involvement of EPHA1, as well as MS4A4A and TREM2 listed below, in the immune module of AD. As previously described, we observed no aberrant EPHA1 expression in the PFC region of AD patients, neither in the hippocampus.

### 4.2 | MS4A4A gene at 11q12.2

MS4A4A is a strong AD candidate gene within the membrane-spanning 4-domains subfamily A (MS4A) gene cluster, an established risk locus for AD. It is predominantly expressed in immune cells, including resident microglia in the CNS. Following microglial activation, MS4A4A shows increased expression. Similarly, MS4A4A is upregulated in activated dendritic cells (DCs) and M1 macrophages, while not detected in immature DCs and M2 macrophages. Here, MS4A4A showed significantly altered expression in the PFC region of AD brains (P = 1.08E–24), yet displayed an expression alteration of suggestive evidence in the hippocampal region (P = 3.87E–03).

Notably, MS4A4A was also differentially expressed in peripheral blood of IS patients (P = 2.73E–05) and in mouse spleens after IS with over 5-fold upregulation (P = 1.57E–05; log2-fold change = 2.37). Following IS, increased incidence of infections occurs, mainly in the form of pneumonia and urinary tract infections. The underlying mechanism is insufficient antigen-presentation of monocytes/macrophages and DCs in peripheral immune organs, resulting from downregulation of MHC class II and co-stimulatory molecules and remarkable reduction of proinflammatory cytokines. Here, we speculated that the pronounced upregulation of MS4A4A in the spleen following IS might reflect a phenotypic switch of monocytes from the proinflammatory M1 phenotype to the anti-inflammatory M2 phenotype.

### 4.3 | UBE2L3, YDJC and SLC2A11 genes at 22q11

Ubiquitin-conjugating enzyme 2 L3 (UBE2L3) and YDJC are located at 22q11.21. The genetic relationship between the 22q11.21 region and multiple autoimmune diseases has been extensively elucidated. Besides, SNPs near YDJC are suggested to be a pleiotropic locus between AD and Crohn disease. Lately, UBE2L3 has been identified as a hub gene in the gene regulatory networks of AD. UBE2L3 encodes an E2 ubiquitin-conjugating enzyme. Through its action on ubiquitination in NF-κB signaling, UBE2L3 promotes NF-κB activation, thus mediates its link with numerous autoimmune diseases. Moreover, UBE2L3 modulates pro-IL-1β processing and mature IL-1β secretion, the deregulation of which pronouncedly intensifies neuronal damage in both AD and IS.

In addition, UBE2L3 directly interacts with the parkin protein, a ubiquitin-protein ligase that is protective against not only neurodegenerative diseases but also cerebral ischemia-reperfusion injury. Nonetheless, there is no conclusive evidence to date demonstrating a causative link between UBE2L3 and AD or IS. The function of YDJC remains largely obscure. Solute-carrier 2A family member 11 (SLC2A11) encodes GLUT11, a fructose-transporting protein that might participate in fructose consumption in the CNS.

### 4.4 | PABPC1 gene at 8q22.3

Poly(A) binding protein cytoplasmic 1 (PABPC1), one type of RNA-binding proteins, is a central component of cytoplasmic stress granules comprising proteins and mRNAs stalled at the translation initiation step. Pathological stress granules play crucial roles in neurodegenerative disorders, also in brain ischemia. Moreover, abnormal cytoplasmic inclusions of PABPC1 have been observed in human ALS spinal cord neurons. In the present work, PABPC1 was differentially expressed in both the PFC (P = 1.90E–11) and hippocampal regions (P = 2.99E–04) of AD. Evidently in IS brains, its expression was significantly enhanced (log2-fold change = 1.19) compared to control subjects (Supporting Information Table S4A).

### 4.5 | HECTD4 and OAS2 gene at 12q24

Mounting GWAS studies have demonstrated the pleiotropic effects of 12q24 locus on type 1 diabetes, celiac disease, coronary artery disease and a number of cardiovascular risk factors, including hypertension, cholesterol levels, and glycemia. Importantly, the 12q24 region has been suggested to be a risk locus of IS. Recently, SNPs near HECTD4 are shown to be associated with memory performance. HECT domain E3 ubiquitin-protein ligase 4 (HECTD4) is a E3 ubiquitin-protein ligase. HECT-type E3s can function with UBE2L3 discussed above in the ubiquitin system. Here, HECTD4 was differentially expressed in both AD’s PFC (P = 8.74E–16) and hippocampal (P = 1.73E–03) regions compared to controls. Regrettably, the transcripts of HECTD4 were not detected in any expression datasets of IS.

### 4.6 | PINX1 gene at 8p23.1

PIN2 (TERF1) interacting telomerase inhibitor 1 (PINX1) protein is a potent telomerase inhibitor and a microtubule-binding protein essential for chromosome segregation in mitosis. In addition to its biological significance in various cancers, PINX1 gene is associated with subclinical cardiovascular events like carotid intima media thickness, blood lipids, and involved in AD as a potential interactor of Ap.

### 4.7 | TREM2 gene at 6p21.1

Triggering receptor expressed on myeloid cells-2 (TREM2) is highly expressed on microglia as an innate immune receptor involved in phagocytosis, clearance of damaged neurons, and inhibition of the microglial proinflammatory response. Mutation of rare variants in TREM2 confers a substantial increase in AD risk, which has been
experimentally proved. TREM2 has been indicated to be upregulated and participate in ischemic brain damage by modulating microglial phenotypes despite conflicting findings. Here, TREM2 was shared by AD and IS with individual P = 0.0015 and P = 0.0095, respectively. Since rare variants (MAF < 0.01) were excluded from the GWAS panels of both AD and IS, we assumed that the true joint association signal of TREM2 with AD and IS might be stronger than we observed. TREM2 transcriptional changes were observed in the PFC region (P = 6.80E−20) of AD, and in rat peri-infarcted brains (P = 9.29E−06). Expectedly, the latter showed nearly 4-fold increased expression for TREM2 (log2-fold change = 1.99; Supporting Information Table S4A).

4.8 | Others

Though we found that the following genes (ie EFTUD1/FAM154B at 15q25.2, RRN3P1, SLCO1A5, and ANKHD1‐EIF4EBP3) were common association signals for AD and IS in terms of bioinformatics, less is known about their biological roles due to lack of overwhelming evidence related to AD or IS.

Based on the concise discussion about their biological significance, partial pleiotropic genes underlying AD and IS were prioritized, including ZXY, EPHA1, MS4A4A, UBE2L3, PABPC1, HECTD4, PNX1, and TREM2. Consistent with our previous findings from the pathway-oriented perspective, we once again highlighted the critical roles of neuroinflammation in the development and progression of AD and IS, since half of them (EPHA1, MS4A4A, UBE2L3 and TREM2) were engaged in immune signaling.

Although gene‐based tests increase the power to detect disease‐associated genes harboring multiple associated variants, they do have limitations. Firstly, the VEGAS test is prone to underestimating effects of low‐frequency SNPs correlated with few SNPs in LD blocks, but may unable to distinguish the truly casual genes from several adjacent ones colocalizing in one significant locus. Secondly, genes revealed by positional proximity to significant variants are not necessarily the casual ones for disease pathogenesis. In complex diseases, significant variants are mostly located in intronic/intergenic areas, presumably regulating gene expression, including acting on distant genes. Next, we leveraged the GEO dataset to estimate the expression alteration of candidate genes in disease‐related tissues, the reliability of which largely relied on the raw data, for instance, the size of tissue samples in the original studies. Further powerful approaches and sophisticated functional interpretation analyses are warranted to prioritize causal genes. Moreover, being pathologically and genetically heterogeneous, IS has different etiological subtypes (ie large vessel disease, cardioembolic stroke, and small vessel disease, undetermined and other).

Here, we just surveyed the genetic link between AD and overall IS.

5 | CONCLUSIONS

We presented a gene‐based strategy that corroborates shared candidate genes between AD and IS, with gene expression analysis ensued, which provided a typical example of how genetic studies could add to biological understanding of cross‐trait etiology. Literature mining supported the potential association of partial novel candidate genes with both AD and IS. Our findings yielded mechanistic insights into the common pathogenesis underlying AD and IS, predominantly involving the immune system, and might suggest common intervention targets. More importantly, our findings should encourage more studies to verify the involvement of these candidates in AD and IS and interpret the exact molecular mechanisms of action.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (81300945 to G.Y.L.; and 81401361 and 81870954 to X‐F. M.). We’re greatly appreciative of the kind offer of GWAS summary statistics from the IGAP and METASTROKE collaboration.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Kassebaum NJ, Arora M, Barber RM, et al. Global, regional, and national disability‐adjusted life‐years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990‐2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. 2016;388(10053):1603‐1658.
2. Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all‐cause mortality, and cause‐specific mortality for 249 causes of death, 1980‐2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. 2016;388(10053):1459‐1544.
3. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet. 2011;377(9778):1693‐1702.
4. Tolppanen AM, Lavinainen P, Solomon A, Kivipelto M, Soininen H, Hartikainen S. Incidence of stroke in people with Alzheimer disease: a national register‐based approach. Neurology. 2013;80(4):353‐358.
5. Chi NF, Chien LN, Ku HL, Hu CJ, Chiou HY. Alzheimer disease and risk of stroke: a population‐based cohort study. Neurology. 2013;80(8):705‐711.
6. Gamaldo A, Moghekar A, Kilada S, Resnick SM, Zonderman AB, O’Brien R. Effect of a clinical stroke on the risk of dementia in a prospective cohort. Neurology. 2006;67(8):1363‐1369.
7. Dong S, Maniar S, Manole MD, Sun D. Cerebral hypoperfusion and other shared brain pathologies in ischemic stroke and Alzheimer’s disease. Transl Stroke Res. 2018;9(3):238‐250.
8. Gorenlick PB, Scuteri A, Black SE, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American heart association/american stroke association. Stroke. 2011;42(9):2672‐2713.
9. Piutta R, Jablonski M, Ufamek‐Kuzior M, et al. Sporadic Alzheimer’s disease begins as episodes of brain ischaemia and ischmically dysregulated Alzheimer’s disease genes. Mol Neurobiol. 2013;48(3):500‐515.
10. Salminen A, Kauppinen A, Kaarniranta K. Hypoxia/ischemia active processing of Amyloid Precursor Protein: impact of vascular dysfunction in the pathogenesis of Alzheimer’s disease. J Neurochem. 2017;140(4):536‐549.
11. Bi M, Gladbach A, van Eersel J, et al. Tau exacerbates excitotoxic brain damage in an animal model of stroke. *Nat Commun*. 2017;8(1):473.

12. Tuozzex, Lei P, Jackman KA, et al. Tau-mediated iron export prevents ferroptotic damage after ischemic stroke. *Mol Psychiatry*. 2017;22(11):1520-1530.

13. Lambert JC, Ibrahim-Baass BA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45(12):1452-1458.

14. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019;51(3):404-413.

15. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology*. 2016;86(13):1217-1226.

16. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50(5):524-537.

17. Traylor M, Adib-Samii P, Harold D, et al. Shared genetic contribution to Ischaemic Stroke and Alzheimer's Disease. *Ann Neurol*. 2016;79(5):739-747.

18. Boyle EA, Li Y, Pritchard JK. An expanded view of complex traits: from polygenic to omnicogenic. *Cell*. 2017;169(7):1177-1186.

19. Mishra A, Macgregor S. VEGAS2: Software for more flexible gene-based testing. *Twin Res Hum Genet*. 2015;18(1):86-91.

20. Liu JZ, Mcrae AF, Nyholt DR, et al. A versatile genome-based test for genome-wide association studies. *Am J Hum Genet*. 2010;87(1):139-145.

21. Begum F, Ghosh D, Tseng GC, Feingold E. Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Res*. 2012;40(9):3777-3784.

22. Zhang B, Gaiteri C, Bodea LG, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*. 2013;153(3):707-720.

23. Berchold NC, Coleman PD, Cribs DH, Rogers J, Gillen DL, Cotman CW. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging*. 2013;34(6):1653-1661.

24. Buga AM, Margaritescu C, Scholz CJ, Rudu E, Zelenak C, Popa-Wagner A. Transcriptomics of post-stroke angiogenesis in the aged brain. *Front Aging Neurosci*. 2014;6:44.

25. Barr TI, Conley Y, Ding J, et al. Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. *Neurology*. 2010;75(11):1009-1014.

26. Lanni C, Necchi D, Pinto A, et al. Zyxin is a novel target for beta-amyloid peptide: characterization of its role in Alzheimer's pathogenesis. *J Neurochem*. 2013;125(5):790-799.

27. Fujita Y, Yamaguchi A, Hata K, Endo M, Yamaguchi N, Yamashita T. Zyxin is a novel interacting partner for SIRT1. *BMC Cell Biol*. 2009;10:6.

28. Herskovits AZ, Guarente L. SIRT1 in neurodevelopment and brain senescence. *Neuron*. 2014;81(3):229-236.

29. Han J-W, Zheng H-F, Cui Y, et al. Genome-wide association study in the Rotterdam Study. *Alzheimers Dement*. 2018;14(7):848-857.

30. Karch CM, Cruchaga C, Goate A. Alzheimer’s Disease Genetics: From the bench to the clinic. *Neuron*. 2014;83(1):11-26.

31. Karch CM, Jeng AT, Nowotny P, Cady J, Cruchaga C, Goate AM. Expression of Novel Alzheimer’s disease risk genes in control and Alzheimer’s disease brains. *PloS ONE*. 2012;7(11):e50976.

32. Allen M, Zou F, Chai HS, et al. Novel late-onset Alzheimer disease loci variants associate with brain gene expression. *Neurology*. 2012;79(3):221-228.

33. Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD3D and CD2AP are associated with Alzheimer's disease. *Nat Genet*. 2011;43(5):429-435.

34. Ahmad S, Bannister C, van der Lee SJ, et al. Disentangling the biological pathways involved in early features of Alzheimer’s disease in the Rotterdam Study. *Alzheimers Dement*. 2018;14(7):848-857.

35. Heneka MT, Golenbock DT, Latz E. Innate immunity in Alzheimer’s disease. *Nat Immunol*. 2015;16(3):229-236.

36. Fujita Y, Yamaguchi A, Hata K, Endo M, Yamaguchi N, Yamashita T. Zyxin is a novel target for beta-amyloid peptide: characterization of its role in Alzheimer's disease. *Neurobiol Aging*. 2013;34(12):2879-2888.

37. Mishra A, Macgregor S. VEGAS2: Software for more flexible gene-based testing. *Twin Res Hum Genet*. 2015;18(1):86-91.

38. Karch CM, Jeng AT, Nowotny P, Cady J, Cruchaga C, Goate AM. Expression of Novel Alzheimer’s disease risk genes in control and Alzheimer’s disease brains. *PloS ONE*. 2012;7(11):e50976.

39. Allen M, Zou F, Chai HS, et al. Novel late-onset Alzheimer disease loci variants associate with brain gene expression. *Neurology*. 2012;79(3):221-228.

40. Koller J, Bissel S, Wiley C, Stauffer M, Murdoch G. Differential microglial expression of new Alzheimer’s disease-associated genes MS4A4A and MS4A6A. *Alzheimers Dement*. 2012;8(4):P253-P253.

41. Sanjal R, Polvak MJ, Zuccollo J, et al. MS4AAA: a novel cell surface marker for M2 macrophages and plasma cells. *Immunol Cell Biol*. 2017;95(7):611-619.

42. Rowveder G, Ellekjaer H, Salvesen O, Naalsund E, Indredavik B. Functional outcome after common poststroke complications occurring in the first 90 days. *Stroke*. 2015;46(1):65-70.

43. Meisel C, Schwab JM, Prass K, Meisel A, Dirmag U. Central nervous system injury-induced immune deficiency syndrome. *Nat Rev Neurosci*. 2005;6(10):775-786.

44. Chamorro A, Meisel A, Planas AM, Urra X, van de Beek D, Veltkamp R. The immunology of acute stroke. *Nat Rev Neurol*. 2012;8(7):401-410.

45. Han J-W, Zheng H-F, Cui Y, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet*. 2009;41(11):1234-1237.

46. Zhernakova A, Stahl EA, Trynka G, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PloS Genet*. 2011;7(2):e1002004.

47. Ellinghaus D, Ellinghaus E, Nair R, et al. Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Hum Genet*. 2012;90(4):636-647.

48. Yokoyama JS, Wang Y, Schork AJ, et al. Association between genetic traits for immune-mediated diseases and alzheimer disease. *JAMA Neurol*. 2016;73(6):691-697.

49. Kawalia SB, Raschka T, Naz M, de Matos SR, Senger P, Hofmann-Apitius M. Analytical strategy to prioritize Alzheimer’s disease candidate genes in gene regulatory networks using public expression data. *J Alzheimers Dis*. 2017;59(4):1237-1254.

50. Fu B, Li S, Wang L, Berman MA, Dord ME. The ubiquitin conjugating enzyme UBE2L3 regulates TNFα-induced linear ubiquitination. *Cell Res*. 2014;24(3):376-379.

51. Lewis MJ, Vyse S, Shields AM, et al. UBE2L3 polymorphism amplifies NF-kappaB activation and promotes plasma cell development, linking linear ubiquitination to multiple autoimmune diseases. *Am J Hum Genet*. 2015;96(2):221-234.

52. Lewis MJ, Vyse S, Shields A, et al. Effect of UBE2L3 genotype on regulation of the linear ubiquitin chain assembly complex in systemic lupus erythematosus. *Lancet*. 2015;385(Suppl 1):S9.

53. Eldridge M, Sanchez-Garrido J, Hoben GF, Goddard PJ, Shenoy AR. The Atypical Ubiquitin E2 Conjugase UBE2L3 is an Indirect
55. Gelderblom M, Sobey CG, Kleinschnitz C, Magnus T. Danger signals in stroke. *Ageing Res Rev*. 2015;24(Pt A):77-82.
56. Heneka MT, Kummer MP, Stutz A, et al. NLRL3 is activated in Alzheimer's disease and contributes to pathology in APP/PS1 mice. *Nature*. 2013;493(7434):674-678.
57. Shimura H, Hattori N, Kubo S-I, et al. Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet*. 2000;25(3):302-305.
58. Burns MP, Zhang L, Rebeck GW, Querfurth HW, Moussa CE. Parkin promotes intracellular Abeta1-42 clearance. *Hum Mol Genet*. 2009;18(17):3206-3216.
59. Zheng Q, Huang T, Zhang L, et al. Dysregulation of ubiquitin-proteasome system in neurodegenerative diseases. *Front Aging Neurosci*. 2016;8:303.

60. Ayuso MI, Martinez-Alonso E, Regidor I, Alcazar A. Stress granule and the genesis of neurodegenerative disease. *Front Aging Neurosci*. 2012;4:79.
61. Bag J, Bhattacharjee RB. Multiple levels of post-transcriptional regulation of antigen presentation are affected in tumour cells. *Oxid Med Cell Longev*. 2012;2012:927861.
62. Dewey CM, Cenk B, Sephton CF, Johnson BA, Herz J, Yu G. Deregulation of neuronal activity after brain ischemia is independent of eukaryotic translation initiation factor 2-α phosphorylation and is correlated with a decrease in eIF4B and eIF4E proteins. *J Biol Chem*. 2016;291(53):27252-27264.
63. Fittschen M, Lastres-Becker I, Halbach Mv, et al. Genetic ablation of ataxin-2 increases several global translation factors in their transcript abundance but decreases translation rate. *Neurogenetics*. 2015;16(3):181-192.
64. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-678.
65. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet*. 2007;39(7):857-864.
66. Wu X, Wang L, Chen S, et al. New susceptibility loci identified by genome-wide association study of type 1 diabetes. *Nat Genet*. 2009;41(4):395-402.
67. Kato N, Takeuchi F, Tabara Y, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet*. 2011;43(6):531-538.
68. Soranzo N, Spector TD, Mangino M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet*. 2009;41(11):1182-1190.
69. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet*. 2007;39(7):857-864.
70. Hunt KA, Zhernakova A, Smith AV, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2012;44(8):890-894.
71. Lu X, Wang L, Chen S, et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet*. 2012;44(8):890-894.
72. Kato N, Takeuchi F, Tabara Y, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet*. 2011;43(6):531-538.
73. Soranzo N, Spector TD, Mangino M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet*. 2009;41(11):1182-1190.
95. Ulland TK, Song WM, Huang SC, et al. TREM2 maintains microglial metabolic fitness in Alzheimer's disease. Cell. 2017;170(4):649-663. e613.

96. Sieber MW, Jaenisch N, Brehm M, et al. Attenuated inflammatory response in triggering receptor expressed on myeloid cells 2 (TREM2) knock-out mice following stroke. PLoS ONE. 2013;8(1):e52982.

97. Kawabori M, Kacimi R, Kauppinen T, et al. Triggering receptor expressed on myeloid cells 2 (TREM2) deficiency attenuates phagocytic activities of microglia and exacerbates ischemic damage in experimental stroke. J Neurosci. 2015;35(8):3384-3396.

98. Zhai Q, Li F, Chen X, et al. Triggering receptor expressed on myeloid cells 2, a novel regulator of immunocyte phenotypes, Confers Neuroprotection by Relieving Neuroinflammation. Anesthesiology. 2017;127(1):98-110.

99. Kawabori M, Hokari M, Zheng Z, et al. Triggering receptor expressed on myeloid cells-2 correlates to hypothermic neuroprotection in ischemic stroke. Ther Hypothermia Temp Manag. 2013;3(4):189-198.

100. Wu R, Li X, Xu P, et al. TREM2 protects against cerebral ischemia/reperfusion injury. Mol Brain. 2017;10(1):20.

101. Cui P, Ma X, Li H, Lang W, Hao J. Shared biological pathways between Alzheimer’s disease and ischemic stroke. Front Neurosci. 2018;12(605).

102. Huang H, Chanda P, Alonso A, Bader JS, Arking DE. Gene-based tests of association. PLoS Genet. 2011;7(7):e1002177.

103. Markus HS, Bevan S. Mechanisms and treatment of ischaemic stroke—insights from genetic associations. Nat Rev Neurol. 2014;10(12):723-730.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Wei C-J, Cui P, Li H, Lang W-J, Liu G-Y, Ma X-F. Shared genes between Alzheimer’s disease and ischemic stroke. CNS Neurosci Ther. 2019;25:855–864. https://doi.org/10.1111/cns.13117