



Original research

# Systematic druggable genome-wide Mendelian randomisation identifies therapeutic targets for Alzheimer's disease

Wei-Ming Su,<sup>1,2</sup> Xiao-Jing Gu,<sup>3</sup> Meng Dou,<sup>4</sup> Qing-Qing Duan,<sup>1,2</sup> Zheng Jiang,<sup>1,2</sup> Kang-Fu Yin,<sup>1,2</sup> Wei-Chen Cai,<sup>1,2</sup> Bei Cao,<sup>1,2</sup> Yi Wang,<sup>5</sup> Yong-Ping Chen

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jnnp-2023-331142>).

<sup>1</sup>Department of Neurology, West China Hospital, Sichuan University, Chengdu, Sichuan, China

<sup>2</sup>Institute of Brain Science and Brain-inspired Technology, West China Hospital, Sichuan University, Chengdu, Sichuan, China

<sup>3</sup>Department of Mental Health Center, West China Hospital, Sichuan University, Chengdu, China

<sup>4</sup>Chengdu Computer Application Institute, Chinese Academy of Sciences, Chengdu, China

<sup>5</sup>Department of Pathophysiology, West China College of Basic medical sciences & Forensic Medicine, Sichuan University, Chengdu, China

## Correspondence to

Professor Yong-Ping Chen, Department of Neurology, West China Hospital, Sichuan University, Chengdu, Sichuan, China; [yongping.chen@wchscu.cn](mailto:yongping.chen@wchscu.cn)

W-MS and X-JG contributed equally.

Received 20 January 2023  
Accepted 5 June 2023  
Published Online First 22 June 2023

## ABSTRACT

**Background** Alzheimer's disease (AD) is the leading cause of dementia. Currently, there are no effective disease-modifying treatments for AD. Mendelian randomisation (MR) has been widely used to repurpose licensed drugs and discover novel therapeutic targets. Thus, we aimed to identify novel therapeutic targets for AD and analyse their pathophysiological mechanisms and potential side effects.

**Methods** A two-sample MR integrating the identified druggable genes was performed to estimate the causal effects of blood and brain druggable expression quantitative trait loci (eQTLs) on AD. A repeat study was conducted using different blood and brain eQTL data sources to validate the identified genes. Using AD markers with available genome-wide association studies data, we evaluated the causal relationship between established AD markers to explore possible mechanisms. Finally, the potential side effects of the druggable genes for AD treatment were assessed using a genome-wide MR.

**Results** Overall, 5883 unique druggable genes were aggregated; 33 unique potential druggable genes for AD were identified in at least one dataset (brain or blood), and 5 were validated in a different dataset. Among them, three prior druggable genes (epoxide hydrolase 2 (*EPHX2*), *SERPINB1* and *SIGLEC11*) reached significant levels in both blood and brain tissues. *EPHX2* may mediate the pathogenesis of AD by affecting the entire hippocampal volume. Further genome-wide MR analysis revealed no potential side effects of treatments targeting *EPHX2*, *SERPINB1* or *SIGLEC11*.

**Conclusions** This study provides genetic evidence supporting the potential therapeutic benefits of targeting the three druggable genes for AD treatment, which will be useful for prioritising AD drug development.

## BACKGROUND

Alzheimer's disease (AD) is the most common neurodegenerative disease and is the leading cause of cognitive impairment and dementia.<sup>1</sup> An estimated 6.5 million Americans aged  $\geq 65$  years are living with AD.<sup>2</sup> This number might reach 13.8 million by 2060.<sup>2</sup> As an age-dependent disease, AD is a threat to public health because of the rapid ageing of the world population. Over the past three decades, an increasing trend in the burden of AD has been observed globally.<sup>3</sup> The rapid increase in the prevalence and societal and economic burden

## WHAT IS ALREADY KNOWN ON THIS TOPIC

→ The rapid increase in the prevalence and societal and economic burden of Alzheimer's disease (AD) has prompted an urgent need for efficient intervention and treatment for AD. Currently, there are no effective disease-modifying treatments for AD.

## WHAT THIS STUDY ADDS

→ This study provides genetic evidence supporting the potential therapeutic benefits of targeting the three druggable genes (epoxide hydrolase 2 (*EPHX2*), *SERPINB1* and *SIGLEC11*) for AD treatment and analyses their pathophysiological mechanisms and potential side effects.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

→ Our study makes a significant contribution to the literature because we provide genetic evidence supporting the potential therapeutic benefits of targeting the three druggable genes for AD treatment, which will be useful for prioritising AD drug development.

of AD has prompted an urgent need for efficient intervention and treatment for AD.<sup>4</sup>

However, AD treatment remains challenging. First, only a few drugs are available for AD.<sup>5</sup> Second, most approved drugs exert their therapeutic effects by targeting only a few specific molecular targets. For example, almost all antidementia drugs exert their therapeutic effects by inhibiting cholinesterase enzyme or glutaminergic pathway.<sup>6</sup> Third, in addition to their beneficial therapeutic effects, antidementia drugs have considerable side effects, including nausea, vomiting, loss of appetite, headache and constipation.<sup>7</sup> Importantly, these antidementia drugs cannot cure or delay AD progression; they only provide symptomatic treatment.

Over the past few decades, great efforts have been made to explore disease-modifying treatments for AD; however, little progress has been achieved. Most agents targeting amyloid beta (A $\beta$ ), aggregated tau, or neurofibrillary tangles may delay disease progression, but there is no substantial evidence suggesting their clinical benefits.<sup>8</sup> Therapeutic stasis is primarily attributed to unknown AD pathophysiology. Fortunately, large-scale



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Su W-M, Gu X-J, Dou M, et al. *J Neurol Neurosurg Psychiatry* 2023;94:954–961.

human genetic studies provide an opportunity for novel drug development for many complex diseases because drug targets supported by genetic evidence have a greater chance of success in drug discovery pipelines.<sup>9,10</sup> In other words, ‘druggable’ genes encoding proteins or gene expression can provide powerful clues to inform drug targets.<sup>11</sup> Over the past few years, numerous large-scale genome-wide association studies (GWASs) have identified many single-nucleotide polymorphisms (SNPs) associated with AD risk. However, GWAS cannot clearly and directly provide clues regarding causal genes and drug targets because many identified SNPs are located in non-coding or intergene regions.

Mendelian randomisation (MR) is an approach used to assess the causality between a modifiable exposure or risk factor and a clinically relevant outcome.<sup>12</sup> MR analysis has been widely used to repurpose licensed drugs<sup>13</sup> and discover novel therapeutic targets<sup>14,15</sup> by integrating summary data from disease GWASs and expression quantitative trait locus (eQTL) studies. The expression levels of a gene may be considered as a kind of life-long exposure, and eQTL located in the genomic region of the druggable genes are always considered as proxies.<sup>16,17</sup>

Therefore, we performed a systematic druggable genome-wide MR to identify therapeutic targets for AD. First, a two-sample

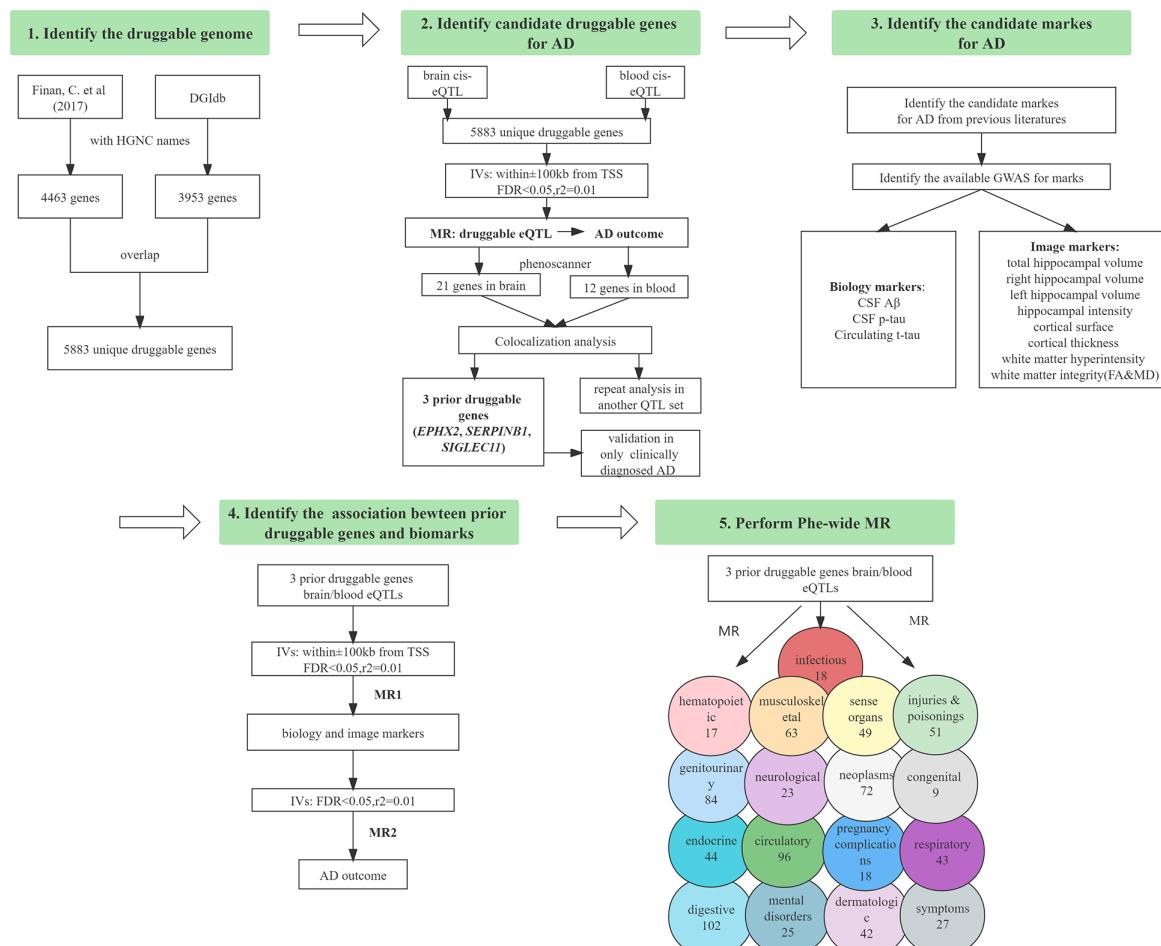
MR analysis was performed to estimate the causal effects of blood and brain druggable eQTLs on AD. Second, we performed colocalisation analyses to verify the robustness of the expression instrumental variables (IVs) and conducted a repeat study using different blood and brain eQTL data sources to validate the identified genes. Third, we evaluated the causal relationships between the identified genes and established AD markers to explore the possible mechanisms by which these genes are involved in AD pathogenesis. Finally, we assessed the potential adverse effects of the identified druggable genes for AD treatment using a phenotype-wide MR.

## MATERIALS AND METHODS

The overall study design is illustrated in figure 1. Further details of the methods and materials used are provided as follows.

### Identification of druggable genes

Druggable genes were obtained from the Drug–Gene Interaction Database (DGIdb V.4.2.0, <https://www.dgidb.org/>)<sup>18</sup> and a recent review on the ‘druggability’ of genes.<sup>11</sup> The DGIdb provides information on drug–gene interactions and druggable genes



**Figure 1** Overview of this study. First, we identified druggable genes. Second, a two-sample MR was conducted to estimate the causal effects of blood and brain druggable eQTLs on AD. We then performed colocalisation analyses to verify the robustness of the expressions' IVs, and a repeat and validation study was conducted. Third, we aimed to identify AD biomarkers with GWAS data available. Subsequently, we evaluated the causal relationship of prior expressions on AD biomarkers to explore possible mechanisms. Finally, we assessed the potential side effects of targeting the prior druggable gene products for AD treatment via a phenotype-wide MR. AD, Alzheimer's disease; CSF, cerebrospinal fluid; DGIdb, Drug–Gene Interaction Database; eQTL, expression quantitative trait locus; FA, fractional anisotropy; FDR, false discovery rate; GWAS, genome-wide association study; IV, instrumental variable; MD, mean diffusivity; MR, Mendelian randomisation; TSS, transcriptional start site.

from publications, databases and other web-based sources.<sup>18</sup> We downloaded the ‘Categories Data’ (released in February 2022), including all genes in the druggable categories in the DGIdb, from all sources mapped to Entrez genes. We also obtained a list of druggable genes from a review by Finan *et al.*<sup>11</sup>

### eQTL datasets

The discovery blood eQTL dataset was obtained from eQTLGen (<https://eqtlgen.org/>), where the cis-eQTLs of 16 987 genes were obtained from 31 684 blood samples of healthy European-ancestry individuals. Fully significant cis-eQTL results (false discovery rate (FDR) <0.05) and allele frequency information were obtained. The discovery brain eQTL data were obtained from the PsychENCODE consortia (<http://resource.psychen-code.org>), which included 1387 prefrontal cortex and primarily European samples. We downloaded all significant eQTLs (FDR <0.05) for genes with expression of >0.1 fragments per kilobase per million mapped fragments in at least 10 samples and all SNP information. Further details on the data are presented in online supplemental table S1 and in the original publications.<sup>19 20</sup>

In addition, we used whole blood and brain frontal lobe cis-eQTLs from the Genotype-Tissue Expression project (GTEx) V8 (<https://gtexportal.org/home/datasets>) to validate our findings.<sup>21</sup> Descriptions of donor enrolment, consent process, biospecimen procurement, sample fixation and other information are presented in online supplemental table S1 and on the website (<https://www.gtexportal.org>).

### AD GWAS dataset

The outcome data used in this study were obtained from the most recent and largest AD GWAS, which included 111 326 clinically diagnosed or ‘proxy’ AD cases and 677 663 controls.<sup>22</sup> It comprises various European GWAS consortia already working on AD and a new dataset collated from 15 European countries. The subgroups included in this GWAS are presented in online supplemental table S2 and more information is available in the original publication.<sup>22</sup> Moreover, a GWAS meta-analysis of only clinically diagnosed AD cases (stage I: 21 982 cases, 41 944 controls)<sup>23</sup> was used to validate our findings (online supplemental table S2).

### eQTL MR analysis

MR analyses were performed using the R package TwoSampleMR V.0.5.6.<sup>24</sup> The eQTLs of the drug genome were selected as the exposure data. To generate IVs, we selected SNPs with FDR of <0.05 and within ±100 kb from each gene’s transcriptional start site (TSS).<sup>19</sup> SNPs in each eQTL were then clumped at  $r^2 < 0.01$  using European samples from the 1000 Genomes Project.<sup>24 25</sup> The R package PhenoScanner V.1.0 was used to find IVs-related phenotypes. Furthermore, we removed SNPs that were directly related to the outcome data (AD) and AD-associated traits (eg, dementia and family history of AD).<sup>26</sup> The outcome data were then loaded and harmonised using the in-built functions. The Wald ratio method was used to compute the MR estimates for each SNP. Where more than one SNP was available, a weighted mean of the ratio estimates, weighted by the inverse variance of the ratio estimates (inverse-variance weighted, IVW), was used. We assessed whether the MR–Egger intercept significantly deviated from 0 to test horizontal pleiotropy when the number of SNPs is no less than three. Moreover, Cochran’s Q methods were used to test for heterogeneity between Wald ratios.<sup>27</sup> FDR-corrected p values were calculated, and FDR of <0.05

was considered significant. In the replication studies, nominal p values of <0.05 were considered statistically significant.

### Colocalisation analysis

Occasionally, an SNP is located in two or more gene regions. In this case, if an SNP contained eQTL information for two or more different genes, its effect on the disease (here, AD) would be mixed by different genes, and colocalisation analysis was used to confirm that AD and the eQTLs may share causal genetic variants. Briefly, for significant MR results in the discovery, we performed a colocalisation analysis for AD risk and the SNP within ±100 kb from each gene’s TSS in eQTL via the R package COLOC V.5.2.0<sup>28</sup> with  $P_1 = 1 \times 10^{-4}$ ,  $P_2 = 1 \times 10^{-4}$  and  $P_{12} = 1 \times 10^{-5}$ . The probability that a given SNP is linked to AD is denoted as  $P_1$ ; the probability that a given SNP is a significant eQTL is denoted as  $P_2$ ; and the probability that a given SNP is an outcome of both AD and eQTL is denoted as  $P_{12}$ . The COLOC package was used to test five hypotheses. The posterior probability (PP) was used to quantify support for all the hypotheses, and they were identified as  $PPH_0$ – $PPH_4$ :  $PPH_0$ , no association with either trait;  $PPH_1$ , association with the expression of the gene but not the AD risk;  $PPH_2$ , association with the AD risk but not the expression of the gene;  $PPH_3$ , association with the AD risk and expression of the gene, with distinct causal variants; and  $PPH_4$ , association with the AD risk and expression of the gene, with a shared causal variant.<sup>28</sup> We restricted our analysis to genes reaching  $PPH_3 + PPH_4$  of ≥0.8 owing to limited power in the colocalisation analysis.<sup>28</sup>

### Markers associated with AD

To determine whether these three AD-causal genes (epoxide hydrolase 2 (*EPHX2*), *SIGLEC11* and *SERPINB1*) which we identified in our study via MR analysis were involved in the hypothetical pathological mechanism of AD, we first performed another set of two-sample MR with the same MR parameters between the three identified genes’ eQTL (*EPHX2*, *SIGLEC11* and *SERPINB1*) in the brain and blood and the hypothetical biomarkers to investigate the effect of gene expression on AD biomarkers. We then performed the second step using the markers’ GWASs as exposures and the AD GWAS as the outcome to explore the mediating role. The biochemical and neuroimaging markers of AD were obtained from previous reviews<sup>1 29 30</sup> (online supplemental table S3).

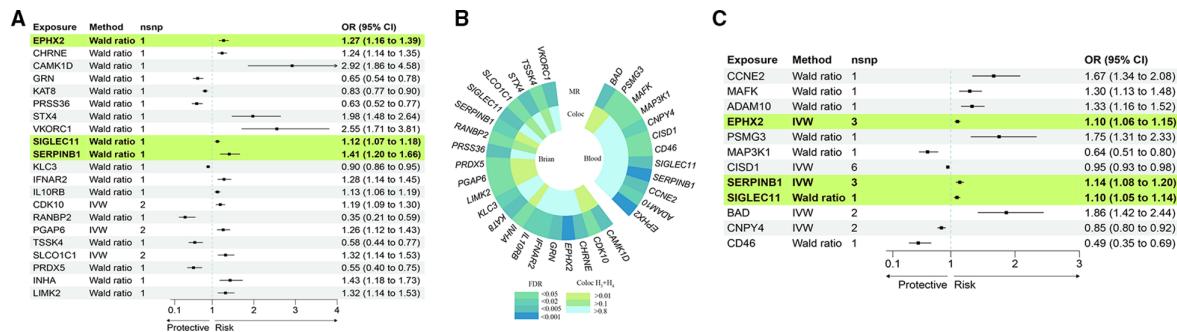
### Phenome-wide MR

To study the potential side effects of these three prior druggable genes, we used gene expression as exposure and summary statistics of diseases in the UK Biobank cohort ( $n \leq 408\,961$ ) as outcomes to perform phenome-wide MR. Disease GWASs from the UK Biobank were performed using the Scalable and Accurate Implementation of Generalised Mixed Model (SAIGE V.0.29) method to account for unbalanced case–control ratios.<sup>31</sup> Owing to statistical power, we chose 783 traits (diseases) with more than 500 cases for phenome-MR analyses. Summary statistics of disease-associated SNPs were downloaded from the SAIGE GWAS (<https://www.leelabsg.org/resources>).<sup>31</sup> Further details are provided in the publication.<sup>31</sup> We then used the same parameters to perform MR analysis using the prior druggable gene eQTL in the blood via the IVW method. The causal effects are considered statistically significant at FDR<0.05.

## RESULTS

### Druggable genome

Using the data available in DGIdb v4.2.0, we identified 3953 genes as potentially druggable (online supplemental table S4).



**Figure 2** MR and colocalisation analysis results in the discovery phase. (A) Forest plot for MR results between brain eQTL and AD. (B) FDR and value of colocalisation analysis for significant MR result genes. (C) Forest plot for the MR result between blood eQTL and AD. The fluorescent background and bold font (A,C) indicate that these genes are significant in both tissues. The left ring (B) represents genes in the brain, and the right ring represents genes in the blood. AD, Alzheimer's disease; eQTL, expression quantitative trait locus; FDR, false discovery rate; MR, Mendelian randomisation.

In addition, we extracted 4463 genes with druggability from a review<sup>11</sup> (online supplemental table S5). Finally, 5883 unique druggable genes with Human Genome Organisation Gene Nomenclature Committee names were aggregated from the two sources mentioned above for further analysis (online supplemental table S6).

### Candidate druggable genes for AD

We used eQTLs from the brain tissue and blood to intersect with druggable genes to obtain druggable eQTLs. All primitive significant brain eQTLs ( $FDR < 0.05$ ) included 14 653 gene symbols after removing SNPs beyond  $\pm 100$  kb from the TSS. When overlapping with the druggable genome, we obtained the brain druggable eQTL as the exposure, which included 4422 gene symbols. After the SNPs were clumped, 3021 genes remained to run the MR. After removing the SNPs directly related to AD, 21 candidate druggable genes for AD were identified following FDR adjustment ( $FDR < 0.05$ ) using the Wald ratio or the IVW method (figure 2). The IVs for significant expression, PhenoScanner outcome, and full results of the MR are presented in online supplemental table S7–9.

Regarding blood eQTL, by running criteria similar to the brain eQTL, blood druggable eQTL exposure, including 2860 gene symbols, was obtained to perform MR. As a result, we found 12 potential drug targets (figure 2 and online supplemental table S10). In summary, 33 unique potential druggable genes for AD were identified in at least one dataset (brain and blood). Notably, three prior druggable genes (*EPHX2*, *SERPINB1* and *SIGLEC11*) reached significant levels in the blood and brain tissues. They all showed risk influence on AD (in the brain, *EPHX2* OR 1.27, *SERPINB1* OR 1.41 and *SIGLEC11* OR 1.12; in the blood, *EPHX* OR 1.10, *SERPINB1* OR 1.14 and *SIGLEC11* OR 1.10, respectively).

Subsequently, we performed a colocalisation analysis to further determine the probability that SNPs associated with AD and eQTL shared causal genetic variants using the SNP within  $\pm 100$  kb from the TSS of the 33 potential druggable genes. The results suggested that most identified genes and AD likely share a causal variant within the region (figure 2 and online supplemental table S11). For three prior genes, *EPHX2* and *SIGLEC11* produced significant results in both datasets ( $PPH_3 + PPH_4 > 0.8$ ), and *SERPINB1* also reached significant levels in the blood ( $PPH_3 + PPH_4 = 0.98$ ) but not in the brain ( $PPH_3 + PPH_4 = 0.42$ ) in the further colocalisation analysis. When only the GWAS of clinically diagnosed AD was used, *SIGLEC11* and *SERPINB1* passed FDR of  $< 0.05$  in the brain eQTL, and the FDR of *SIGLEC11*

in the blood was very close to 0.05 (online supplemental table S12).

Additionally, during our replication, MR using whole blood and brain frontal lobe cis-eQTL from GTEx V.8,<sup>21</sup> *SIGLEC11*, *SERPINB1*, *LIMK2*, *CD46* and *ADAM10* showed suggestive significance ( $p < 0.05$ ) in either the blood or brain (online supplemental table S13 and S14).

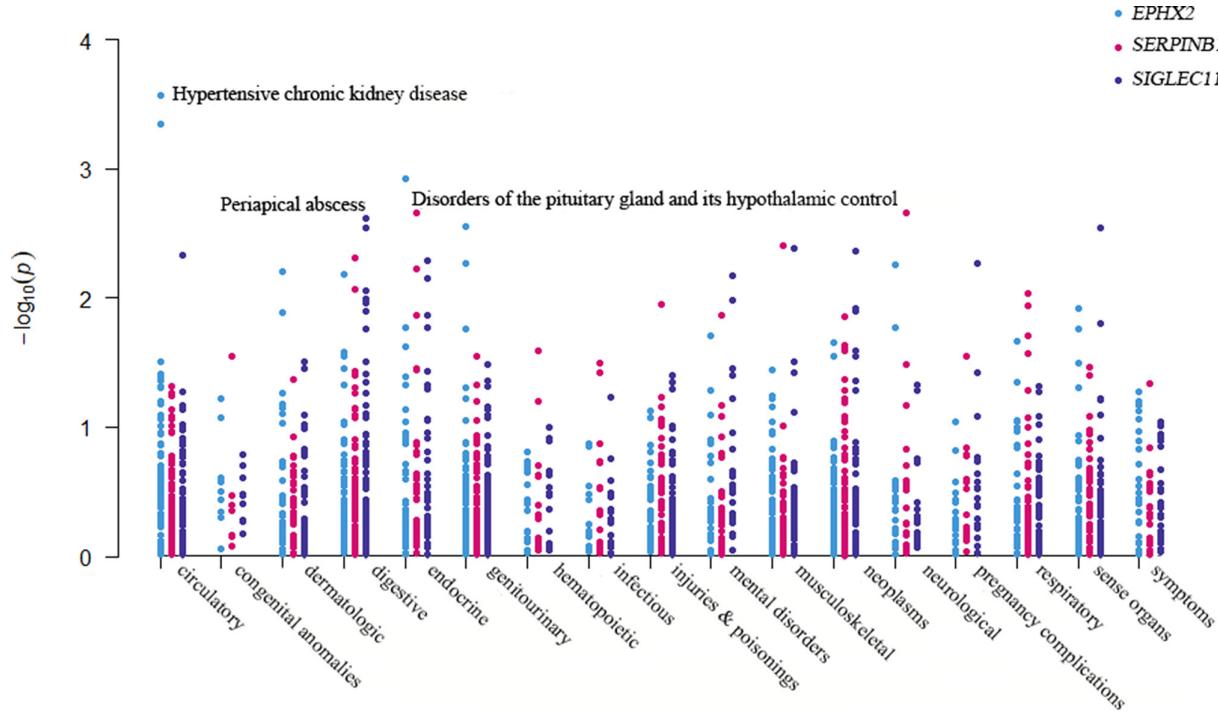
### Association between AD druggable genes and markers

Previous reviews were screened to identify the biochemical and neuroimaging markers of AD.<sup>1,29</sup> Owing to the restricted GWAS data on AD markers, we only found the following markers with sufficient data: cerebrospinal fluid (CSF)  $\text{A}\beta$ ,<sup>32</sup> CSF p-tau,<sup>32</sup> circulating t-tau,<sup>33</sup> total/right/left hippocampal volume,<sup>34,35</sup> cortical surface and thickness,<sup>36</sup> white matter hyperintensities<sup>35</sup> and white matter integrity (fractional anisotropy and mean diffusivity).<sup>37</sup>

We found that blood and brain *EPHX2* may be associated with the whole hippocampal volume (unadjusted  $p$  values = 0.019 and 0.0088, respectively; online supplemental table S15). However, no association was identified between three prior druggable genes (*EPHX2*, *SERPINB1* and *SIGLEC11*) and the aforementioned AD markers. In addition, using the marker GWAS as an exposure and AD as the outcome, we observed that only the whole brain surface affected the AD outcome (unadjusted  $p$  value = 0.00869, online supplemental table S16).

### Phenome-wide MR analysis of AD prior druggable genes

Because most drugs function through blood circulation, we assessed whether the three AD-associated expressions in the blood had beneficial or deleterious effects on other indications. Therefore, we performed a broader MR screening of 783 non-AD diseases or traits in the UK Biobank (SAIGE V.0.29, online supplemental table S17).<sup>31</sup> Overall, no significant association was identified ( $FDR < 0.05$ ) using the IVW method (figure 3), although we observed some trends. Higher blood *EPHX2* levels may potentially benefit hypertensive chronic kidney disease (OR 0.73,  $p = 0.00027$ ) (figure 3 and online supplemental table S18). In addition, high levels of *SERPINB1* caused pituitary gland disorders and hypothalamic control (OR 1.67,  $p = 0.0023$ ) (figure 3 and online supplemental table S19). Blood *SIGLEC11* levels were most likely associated with periapical abscesses (OR 0.74,  $p = 0.0025$ ) (figure 3 and online supplemental table S20). No other diseases may be linked to these druggable genes that



**Figure 3** Manhattan plot for genome-wide MR results of blood EPHX2, SERPINB1 and SIGLEC11. Note: ordinate representation of the p value in genome-wide MR results. A dot represents a disease trait, and different colours represent the MR result of different expressions. EPHX2, epoxide hydrolase 2; MR, Mendelian randomisation.

passed a significant FDR (0.05), and the summary results are presented in online supplemental table S18–20.

## DISCUSSION

The development of novel therapeutic agents for AD is extremely challenging. A major cause for this dilemma is the unknown pathophysiology of AD. In this study, based on blood and brain druggable eQTLs, we found 33 druggable gene expressions that may influence AD outcomes, 5 of which were also found in a repeated study. Indeed, our study provides robust evidence that three genes (*EPHX2*, *SERPINB1* and *SIGLEC11*) are causally associated with AD, and all show evidence of genetic colocalisation with AD outcomes. Furthermore, we found associations between *EPHX2* and whole hippocampal volume, which may provide potential clues to the pathological mechanisms involved in AD. In addition, genome-wide MR highlighted additional beneficial indications of therapeutics targeting the three AD-associated genes and indicated a few potential safety concerns.

*EPHX2* encodes a member of the epoxide hydrolase family, which is found in the cytosol and peroxisomes. It binds to specific epoxides and converts them to the corresponding dihydrodiols.<sup>38</sup> Soluble epoxide hydrolase (sEH) expression is elevated in the hippocampus of AD mice, and sEH levels are also higher in the brain of patients with AD compared with healthy individuals.<sup>39</sup> In pharmacological studies, the genetic deletion of *EPHX2* or the downregulation of sEH level reduces A $\beta$  deposition in the brain of AD model mice, prevents astrocyte reactivation and alleviates cognitive deficits.<sup>39,40</sup> Furthermore, a proteome-wide association study integrating a smaller AD GWAS with the discovery of the Religious Orders Study, Memory and Ageing Project brain proteomes<sup>41</sup> and a non-druggable genomic MR study using protein quantitative trait locus data<sup>42</sup> reported that *EPHX2* might affect AD. Consistent

with these results, we provide druggable genetic evidence for the directionally consistent effects of *EPHX2* in the blood or brain tissue on AD outcomes. We also observed that higher levels of *EPHX2* may be associated with lower hippocampal volume. More importantly, through genome-wide MR analysis, we observed no significant side effects associated with *EPHX2* expression (online supplemental table S18). sEH expression is elevated in the AD hippocampus<sup>39,40</sup>; whether the pathological accumulation of sEH accelerates hippocampal atrophy and leads to AD requires further research. PF-750, a drug that inhibits fatty acid amide hydrolase, weakly inhibits sEH.<sup>43</sup> However, this compound has not yet reached phase I clinical trials (preclinical/research compound). Another drug, AR9281, may also interact with *EPHX2* without a clear direction yet.<sup>44</sup> Overall, our results indicated that *EPHX2* is a promising target for treating AD by affecting hippocampal volume. However, PF-750 or compounds that target *EPHX2* deserve further investigation.

Serpin family B member 1 (*SERPINB1*) encodes a member of the serpin family of proteinase inhibitors, maintaining homeostasis by neutralising overexpressed proteinase activity through their function as suicide substrates.<sup>45</sup> A locus associated with A $\beta$ 42 within *SERPINB1* on chromosome 6p25 has been reported to be associated with AD risk in a GWAS study.<sup>46</sup> A previous study indicated that the intronic *SERPINB1* variant rs316341 affects the expression of *SERPINB1* in various tissues, including the hippocampus, suggesting that *SERPINB1* may influence AD through an A $\beta$ -associated mechanism.<sup>46</sup> In addition, *SERPINB1* is upregulated in patients with AD,<sup>47</sup> and higher levels of *SERPINB1* expression in the prefrontal cortex are associated with higher levels of amyloidosis.<sup>48</sup> Our MR study also revealed that higher *SERPINB1* levels increased the risk of AD. However, we found no association between *SERPINB1* and AD biomarkers using MR in this study. Because of the unavailability

of a large proportion of AD pathological marker GWAS data, negative results are not credible. More basic experiments and larger GWAS data are needed to determine the relationship between *SERPINB1* and A $\beta$  or other AD biomarkers. *SERPINB1* is the only druggable gene under the tier 3 B term (encoding secreted or extracellular proteins, proteins with more distant similarity to approved drug targets and members of key druggable gene families not already included in tier 1 or 2) according to Finan *et al.*<sup>11</sup> Currently, there are no drugs with *SERPINB1* as a clear target. Heparin from bovines is predicted to interact with *SERPINB1* ([www.novoseek.com](http://www.novoseek.com)); however, its exact role is unclear. Our genome-wide MR analysis showed that the association of the drug with blood *SERPINB1* produced no side effects in other systems. Hence, drug development targeting *SERPINB1* may be promising, and further research is needed to elucidate the mechanism of action of *SERPINB1* on the risk of AD.

*SIGLEC11*, another robust AD-associated gene identified in the present study, encodes a member of the sialic acid-binding immunoglobulin-like lectin family. It is mainly expressed in peripheral blood leucocytes, macrophages and microglia and binds to an  $\alpha$ -2-8-linked sialic acid.<sup>49</sup> It is a risk factor for AD.<sup>50</sup> *SIGLEC11* may be associated with ceramide-related inflammation and anti-inflammatory pathways in AD.<sup>50</sup> In addition to our findings, higher levels of *SIGLEC11* may trigger an inflammatory response that can produce neurotoxic effects or promote the deposition of pathological proteins in AD, although the exact mechanism is unknown. Further investigation of *SIGLEC11* in the context of AD is required.

In addition, the remaining expression products could provide further insights into anti-AD drug discovery. When comparing the MR results using eQTLs derived from the two cohorts, *SIGLEC11*, *SERPINB1*, *LIMK2*, *CD46* and *ADAM10* were repeated in at least one tissue sample. *LIMK2* regulates cortical development by regulating neural progenitor cell proliferation and migration.<sup>51</sup> Consistent with our findings, a previous study identified *CD46* as significantly dysregulated in late-onset AD.<sup>52</sup> Moreover, *ADAM10* is an important enzyme in AD pathology and may function in AD via astrocytes.<sup>53</sup> 2,4-Diacetylphloroglucinol reduces A $\beta$  production and secretion by regulating *ADAM10* and its intracellular trafficking in cellular and animal models of AD.<sup>54</sup> Furthermore, the development of *ADAM10* endocytosis inhibitors for AD treatment yielded promising preclinical results.<sup>55</sup> Therefore, multiple studies support the drug targets identified in our study.

Since our primary objective was to provide genetic evidence to improve success rates in clinical trials for AD drug discovery, we decided to reduce the number of false positives. Therefore, we integrated two-stage MR (discovery and validation) in blood and brain tissues and used colocalisation analysis to identify robust druggable genes. We removed SNPs directly related to AD and AD-related diseases to prevent strengthening the MR findings using PhenoScanner. Finally, the three robust druggable genes were further analysed for AD biomarkers and genome-wide MR. The strengths of our study include the following. Previously, there have been a few MR studies on QTLs and AD outcome<sup>42,56</sup>; however, we used the largest and latest AD GWAS,<sup>22</sup> and our MR study is the first to study druggable genes in the context of AD and AD biomarkers. There were few significant results between markers and AD when running MR; however, we cannot exclude the association between markers and AD.

Our study also has some limitations. First, the number of IVs in MR for eQTL is limited, and most are no more than three SNPs, which limits the level of confidence in the MR results. Because

of the number of IVs, we only ran heterogeneity and pleiotropy analyses for MR between AD markers and AD outcomes but not for the MR-involved eQTLs. Second, because the data were locked, many AD markers were not included in our study, such as neurofibrillary tangles, inflammatory plaques, soluble *TREM2* and positron emission tomography images, leading to potential associations being missed. Third, the findings were not repeated at the protein level, considering that many genes did not have IVs after filtering the p value when we conducted a preliminary study. In addition, our initial findings were based on AD GWAS, including clinical and proxy cases. MR with proxy cases yields counterintuitive findings.<sup>57</sup> Nonetheless, *SIGLEC11* and *SERPINB1* reached adjusted significance only in clinically diagnosed AD GWAS. This may be due to insufficient sample size and the fact that *EPHX2* was not repeated. Furthermore, it is difficult to determine the most appropriate tissue for AD discovery. The genes that reached significant levels in the blood and brain tissues may have stronger evidence. Another key limitation is that MR analysis cannot fully recapitulate clinical trials. MR analysis differs from clinical trials, which typically investigate comparably high drug doses over a short time.<sup>14</sup> Additionally, the GWASs conducted in the UK Biobank were restricted to Caucasians, which might restrict the applicability of our findings to other groups and races. Moreover, although all participants in the UK Biobank were of European ancestry, geographical clustering was not adjusted in the original GWASs. It was reported that the health outcomes appear geographically structured and that coincident structure in health outcomes and genotype data can yield biased associations in UK Biobank.<sup>58</sup> Therefore, our results might be biased by population heterogeneity in the UK Biobank.<sup>58</sup> Future studies with participants from the same region were worthy to be performed. Furthermore, we only explored the side effects of *EPHX2*, *SIGLEC11* and *SERPINB1* eQTL in diseases from the UK Biobank. However, the effects of drugs on their targets are quite extensive, and many off-target effects cannot be explored using MR. Further exploration is required through subsequent basic and clinical trials to gain a more comprehensive understanding. Finally, there were not as many approved drugs as expected, making it difficult to study new uses of the old drugs; however, we found three robust druggable genes.

## CONCLUSION

This study found three robust druggable genes (*EPHX2*, *SERPINB1* and *SIGLEC11*) and 30 candidate druggable genes for AD. In addition, AD biomarkers and genome-wide MR analyses of *EPHX2*, *SERPINB1* and *SIGLEC11* were conducted. In summary, we provide genetic evidence supporting the potential therapeutic benefits of three druggable genes for AD, which will be useful for prioritising AD drug development.

**Acknowledgements** The authors thank the patients with AD in the included studies and all the researchers.

**Contributors** W-MS: conception, methodology, analysis and writing of the manuscript; J-XG: methodology, analysis and validation; MD, Q-QD, ZJ, K-FY and W-CC: validation; BC and YW: validation and supervision of the project; Y-PC: wrote, revised, discussed the final edition and responsible for the overall content; all authors read and approved the final version of the manuscript.

**Funding** This study was supported by the National Key Research and Development Program of China (grant number 2022YFC2703101 to YPC), the National Natural Science Fund of Sichuan (grant number 2022NSFSC0749 to BC), the Science and Technology Bureau Fund of Sichuan Province (grant number 2023YFS0269 to YPC), the National Natural Science Fund of China (grant number 81971188 to YPC) and the 1·3·5 Project for Disciplines of Excellence, Clinical Research Fund, West China Hospital, Sichuan University (grant number 2023HXFH032 to YPC).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. Data analyzed in this study were a re-analysis of existing data, which are openly available at locations cited in the main text.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Yong-Ping Chen <http://orcid.org/0000-0002-5603-4192>

#### REFERENCES

- 1 Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer's disease. *The Lancet* 2021;397:1577–90.
- 2 2022 Alzheimer's disease facts and figures. *Alzheimers Dement* 2022;18:700–89.
- 3 Collaborators GBD. Global, regional, and national burden of Alzheimer's disease and other Dementias, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 2019;18:88–106.
- 4 Schwam EM, Abu-Shakra S, del Valle M, et al. Health economics and the value of therapy in Alzheimer's disease. *Alzheimers Dement* 2007;3:143–51.
- 5 Passeri E, Elkhoury K, Morsink M, et al. Alzheimer's disease: treatment strategies and their limitations. *Int J Mol Sci* 2022;23.
- 6 Atri A. Current and future treatments in Alzheimer's disease. *Semin Neurol* 2019;39:227–40.
- 7 Briggs R, Kennelly SP, O'Neill D. Drug treatments in Alzheimer's disease. *Clin Med (Lond)* 2016;16:247–53.
- 8 Kuo YC, Rajesh R. Challenges in the treatment of Alzheimer's disease: recent progress and treatment strategies of pharmaceuticals targeting notable pathological factors. *Expert Rev Neurother* 2019;19:623–52.
- 9 Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet* 2015;47:856–60.
- 10 King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet* 2019;15.
- 11 Finan C, Gaulton A, Kruger FA, et al. The Druggable genome and support for target identification and validation in drug development. *Sci Transl Med* 2017;9:383.
- 12 Sekula P, Del Greco M F, Pattaro C, et al. Mendelian randomization as an approach to assess causality using observational data. *JASN* 2016;27:3253–65.
- 13 Swerdlow DL, Holmes MV, Kuchenbaecker KB, et al. The Interleukin-6 receptor as a target for prevention of coronary heart disease: a Mendelian Randomisation analysis. *Lancet* 2012;379:1214–24.
- 14 Storm CS, Kia DA, Almramhi MM, et al. Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the Druggable genome. *Nat Commun* 2021;12:7342.
- 15 Gaziano L, Giambartolomei C, Pereira AC, et al. Actionable Druggable genome-wide Mendelian randomization identifies Repurposing opportunities for COVID-19. *Nat Med* 2021;27:668–76.
- 16 Schmidt AF, Finan C, Gordillo-Marañón M, et al. Genetic drug target validation using Mendelian Randomisation. *Nat Commun* 2020;11:3255.
- 17 Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait Gene targets. *Nat Genet* 2016;48:481–7.
- 18 Freshour SL, Kiwala S, Cotto KC, et al. Integration of the drug-Gene interaction database (Dgidb 4.0) with open Crowdsource efforts. *Nucleic Acids Res* 2021;49:D1144–51.
- 19 Vösa U, Claringbould A, Westra H-J, et al. Large-scale Cis- and Trans-eQTL analyses identify thousands of genetic Loci and Polygenic scores that regulate blood gene expression. *Nat Genet* 2021;53:1300–10.
- 20 Wang D, Liu S, Warrell J, et al. Comprehensive functional Genomic resource and integrative model for the human brain. *Science* 2018;362.
- 21 Lonsdale J, Thomas J, Salvatore M, et al. The genotype-tissue expression (Gtex) project. *Nat Genet* 2013;45:580–5.
- 22 Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related Dementias. *Nat Genet* 2022;54:412–36.
- 23 Alzheimer Disease Genetics Consortium (ADGC), The European Alzheimer's Disease Initiative (EADI), Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk Loci and Implicates Abeta, Tau, immunity and lipid processing. *Nat Genet* 2019;51:414–30.
- 24 Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human Phenome. *Elife* 2018;7.
- 25 Genomes Project C, Abecasis GR, Auton A. An integrated map of genetic variation from 1,092 human Genomes. *Nature* 2012;491:56–65.
- 26 Staley JR, Blackshaw J, Kamat MA, et al. Phenoscanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;32:3207–9.
- 27 Greco M FD, Minelli C, Sheehan NA, et al. Detecting Pleiotropy in Mendelian Randomisation studies with summary data and a continuous outcome. *Stat Med* 2015;34:2926–40.
- 28 Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for Colocalisation between pairs of genetic Association studies using summary Statistics. *PLOS Genet* 2014;10.
- 29 Aerqin Q, Wang Z-T, Wu K-M, et al. Omics-based biomarkers discovery for Alzheimer's disease. *Cell Mol Life Sci* 2022;79:585.
- 30 Hu H-Y, Ou Y-N, Shen X-N, et al. White matter Hyperintensities and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 36 prospective studies. *Neurosci Biobehav Rev* 2021;120:16–27.
- 31 Zhou W, Nielsen JB, Fritzsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic Association studies. *Nat Genet* 2018;50:1335–41.
- 32 Jansen IE, van der Lee SJ, Gomez-Fonseca D, et al. Genome-wide meta-analysis for Alzheimer's disease cerebrospinal fluid biomarkers. *Acta Neuropathol* 2022;144:821–42.
- 33 Sarnowski C, Ghanbari M, Bis JC, et al. Meta-analysis of genome-wide Association studies identifies ancestry-specific associations underlying circulating total Tau levels. *Commun Biol* 2022;5.
- 34 van der Meer D, Rokicki J, Kaufmann T, et al. Brain scans from 21,297 individuals reveal the genetic architecture of hippocampal Subfield volumes. *Mol Psychiatry* 2020;25:3053–65.
- 35 Smith SM, Douaud G, Chen W, et al. An expanded set of genome-wide Association studies of brain imaging phenotypes in UK Biobank. *Nat Neurosci* 2021;24:737–45.
- 36 Grasby KL, Jahanshad N, Painter JN, et al. The genetic architecture of the human cerebral cortex. *Science* 2020;367.
- 37 Persyn E, Hanscombe KB, Howson JMM, et al. Genome-wide Association study of MRI markers of cerebral small vessel disease in 42,310 participants. *Nat Commun* 2020;11:2175.
- 38 Sandberg M, Hassett C, Adman ET, et al. Identification and functional characterization of human soluble Epoxide Hydrolase genetic Polymorphisms. *J Biol Chem* 2000;275:28873–81.
- 39 Grinán-Ferré C, Codony S, Pujol E, et al. Pharmacological inhibition of soluble Epoxide Hydrolase as a new therapy for Alzheimer's disease. *Neurotherapeutics* 2020;17:1825–35.
- 40 Chen W, Wang M, Zhu M, et al. 14,15-Epoxyeicosatrienoic acid Alleviates pathology in a mouse model of Alzheimer's disease. *J Neurosci* 2020;40:8188–203.
- 41 Wingo AP, Liu Y, Gerasimov ES, et al. Integrating human brain Proteomes with genome-wide Association data Implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet* 2021;53:143–6.
- 42 Ou Y-N, Yang Y-X, Deng Y-T, 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.
- 43 Kodani SD, Wan D, Wagner KM, et al. Design and potency of dual soluble Epoxide Hydrolase/fatty acid amide Hydrolase inhibitors. *ACS Omega* 2018;3:14076–86.
- 44 Berman HM, Westbrook J, Feng Z, et al. The protein data Bank. *Nucleic Acids Res* 2000;28:235–42.
- 45 Choi YJ, Kim S, Choi Y, et al. Serpinb1-mediated Checkpoint of inflammatory caspase activation. *Nat Immunol* 2019;20:276–87.
- 46 Deming Y, Li Z, Kapoor M, et al. Genome-wide Association study identifies four novel Loci associated with Alzheimer's Endophenotypes and disease modifiers. *Acta Neuropathol* 2017;133:839–56.
- 47 Zattoni M, Mearelli M, Vanni S, et al. Serpin signatures in Prion and Alzheimer's diseases. *Mol Neurobiol* 2022;59:3778–99.
- 48 Deming Y, Dumitrescu L, Barnes LL, et al. Sex-specific genetic predictors of Alzheimer's disease biomarkers. *Acta Neuropathol* 2018;136:857–72.
- 49 Angata T, Kerr SC, Greaves DR, et al. Cloning and characterization of human Siglec-11. A recently evolved signaling molecule that can interact with SHP-1 and SHP-2 and is expressed by tissue Macrophages, including brain Microglia. *J Biol Chem* 2002;277:24466–74.

- 50 Mizuno S, Ogishima S, Kitatani K, et al. Network analysis of a comprehensive knowledge repository reveals a dual role for Ceramide in Alzheimer's disease. *PLoS ONE* 2016;11.
- 51 Mao R, Deng R, Wei Y, et al. Limk1 and Limk2 regulate cortical development through affecting neural progenitor cell proliferation and migration. *Mol Brain* 2019;12:67.
- 52 Nho K, Nudelman K, Allen M, et al. Genome-wide Transcriptome analysis identifies novel Dysregulated genes implicated in Alzheimer's pathology. *Alzheimers Dement* 2020;16:1213–23.
- 53 Elsworth RJ, Hill EJ, Dunleavy C, et al. The role of Adam10 in Astrocytes: implications for Alzheimer's disease. *Front Aging Neurosci* 2022;14.
- 54 Jang B-G, Choi B, Kim S, et al. 2,4-Diacetylphloroglucinol reduces beta-Amyloid production and secretion by regulating Adam10 and intracellular trafficking in cellular and animal models of Alzheimer's disease. *Cells* 2022;11.
- 55 Musardo S, Therin S, Pelucchi S, et al. The development of Adam10 Endocytosis inhibitors for the treatment of Alzheimer's disease. *Mol Ther* 2022;30:2474–90.
- 56 Jin C, Lee B, Shen L, et al. Integrating multi-Omics summary data using a Mendelian randomization framework. *Briefings in Bioinformatics* 2022;23.
- 57 Liu H, Hu Y, Zhang Y, et al. Mendelian randomization highlights significant difference and genetic heterogeneity in clinically diagnosed Alzheimer's disease GWAS and self-report proxy phenotype GWAX. *Alzheimers Res Ther* 2022;14:17.
- 58 Haworth S, Mitchell R, Corbin L, et al. Apparent latent structure within the UK Biobank sample has implications for Epidemiological analysis. *Nat Commun* 2019;10.