Neurobiology of Aging 123 (2023) 170–181

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging.org

Common and rare variants of *EGF* increase the genetic risk of Alzheimer's disease as revealed by targeted sequencing of growth factors in Han Chinese

Xiao Li[a,b,1], Min Xu[a,b,1], Rui Bi[a,b,d], Li-Wen Tan[c], Yong-Gang Yao[a,b,d,*], Deng-Feng Zhang[a,b,*]

[a] *Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, and KIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Disease, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China*
[b] *Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China*
[c] *Mental Health Institute of the Second Xiangya Hospital, Central South University, Changsha, China*
[d] *CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China*

ARTICLE INFO

Article history:
Received 24 April 2022
Revised 21 September 2022
Accepted 12 October 2022
Available online 3 November 2022

Keywords:
Alzheimer's disease
Growth factors
Targeted sequencing
EGF

ABSTRACT

Alzheimer's disease (AD) is the most common neurodegenerative disease with high heritability. Growth factors (GFs) might contribute to the development of AD due to their broad effects on neuronal system. We herein aimed to investigate the role of rare and common variants of GFs in genetic susceptibility of AD. We screened 23 GFs in 6324 individuals using targeted sequencing. A rare-variant-based burden test and common-variant-based single-site association analyses were performed to identify AD-associated GF genes and variants. The burden test showed an enrichment of rare missense variants ($p = 6.08 \times 10^{-4}$) in GF gene-set in AD patients. Among the GFs, *EGF* showed the strongest signal of enrichment, especially for loss-of-function variants ($p = 0.0019$). A common variant rs4698800 of *EGF* showed significant associations with AD risk ($p = 3.24 \times 10^{-5}$, OR = 1.26). The risk allele of rs4698800 was associated with an increased *EGF* expression, whereas EGF was indeed upregulated in AD brain. These findings suggested *EGF* as a novel risk gene for AD.

1. Introduction

Alzheimer's disease (AD; OMIM, 104300) is an age-related neurodegenerative disorder and the most common type of dementia (Querfurth and LaFerla, 2010; Scheltens et al., 2021). AD is characterized by a series of pathological features including deposition of extracellular β-amyloid (Aβ) plaques, intracellular neurofibrillary tangles formed by hyperphosphorylated tau, neuronal loss, profound microgliosis, and astrocytosis (Querfurth and LaFerla, 2010; De Strooper and Karran, 2016; Scheltens et al., 2021). Growth factors (GFs) are multifunctional proteins that affect the rate of cell proliferation, differentiation and migration, and also tissue remodeling and inflammation (Bottner et al., 2000; Rodrigues et al., 2010; Li et al., 2012; Seeger and Paller, 2015). There are several different kinds of GFs, including the transforming growth factor β (TGF-β) superfamily, fibroblast growth factors (FGFs), insulin derived growth factors (IGFs), epidermal growth factor (EGF) superfamily, vascular endothelial growth factor (VEGF) family, nerve growth factor (NGF) family, platelet-derived growth factor (PDGF) family, and so on. The GFs play important roles in neurogenesis, neurodegeneration (Woodbury and Ikezu, 2014), synapse formation (Schmeisser et al., 2012), axonal transport (Schindowski et al., 2008), and homeostasis of the central nervous system (CNS) (Bottner et al., 2000; Lauzon et al., 2015). The changes of GF levels in blood plasma (Mocali et al., 2004; Ray et al., 2007; Johansson et al., 2013) and brain of AD patients (Rivera et al., 2005; Mahoney et al., 2021) have been reported frequently, and they may be associated with the cognitive or clinical outcomes of AD (Ray et al., 2007; Hohman et al., 2015; Lim et al., 2016). The signaling pathways of GFs have also been demonstrated to be involved in the pathogenesis of AD (Patel et al., 2010; Turner et al., 2016). These signaling abnormalities in turn may impair synaptic plasticity and cognition in AD progression (Caraci et al., 2015; Ferreira, 2021). Due to their functional importance and the advances in recent studies, the targeting of GFs may offer consider-

* Corresponding authors at: Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, No.21, Qingsong Road, Kunming 650201, China.
E-mail addresses: yaoyg@mail.kiz.ac.cn (Y.-G. Yao), zhangdengfeng@mail.kiz.ac.cn (D.-F. Zhang).
[1] These authors contributed equally to this work.

0197-4580/$ – see front matter © 2022 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.neurobiolaging.2022.10.009

able therapeutic potential in the treatment of AD (Freiherr et al., 2013; Lauzon et al., 2015).

AD is a highly heritable disease with population heterogeneity (Lambert and Amouyel, 2007; McClellan and King, 2010). The heritability of AD ranges from 58% to 79% (Gatz et al., 2006). Apart from biomarker evaluation and functional characterizations, genetic studies have also indicated the contributions and involvement of GFs in AD. Several genetic association studies focusing on limited numbers of variants and candidate genes have identified association of some GFs with AD (Luedecking et al., 2000; Chapuis et al., 2006; Wang et al., 2012; Yang et al., 2016). However, the conclusions in these studies have been somewhat controversial, partially because of their relatively small sample size and population substructure (Chang et al., 2013; Liu et al., 2013). In recent genomic studies with large sample size of European populations or populations of European ancestry, one of the growth factors, *EGFR*, was highlighted as a potentially causal gene by gene prioritization and knockoff-based methods (He et al., 2021; Bellenguez et al., 2022). Genomic studies of AD in Chinese (Wang et al., 2018a; Zhou et al., 2018; Zhang et al., 2019; Jia et al., 2021) and Japanese and Korean populations (Miyashita et al., 2013; Shigemizu et al., 2021) have emerged in recent years, and reported several risk genes, although again the studies have been limited by relatively small sample sizes. Considering the fact that GFs have been found to be actively involved in AD, it is appropriate to perform a comprehensive analysis of GF variants in Han Chinese with AD.

In this study, we performed a 3-stage study combining targeted next-generation sequencing (NGS) and meta-analysis in 1280 AD patients and 5044 normal control subjects from the Han Chinese population. We successfully sequenced the exon and nearby untranslated regions (UTRs) of 23 GFs and examined the association of common (minor allele frequency (MAF) $\geq$ 0.01 in control samples) and rare variants (MAF $<$ 0.01) with AD. We identified a novel AD risk gene *EGF*, with both its common and rare variants being associated with AD in Han Chinese.

2. Materials and methods

2.1. Subjects

AD patients without known pathogenic variants in *APP, PSEN1* or *PSEN2* after sequencing verification were recruited from multiple provinces or municipalities in Southern and Eastern China, and were used as the stage 1 and stage 2 cohorts, respectively. All individuals with AD were diagnosed by at least 2 clinical psychiatrists following Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) and some patients have been described in our previous studies (Zhang et al., 2016; Li et al., 2017; Bi et al., 2018; Wang et al., 2016). In brief, the Southern cohort was recruited from hospitals in Sichuan, Hunan and Yunnan provinces, including 635 unrelated individuals with AD (mean age 79.7 $\pm$ 8.2 years, 40.0% male) and 1507 healthy controls (mean age 35.2 $\pm$ 15.5 years, 56.2% male). Part of the control samples from Yunnan Province has been reported in the screening stage conducting targeted NGS in our previous study (Wang et al., 2018b). The Eastern cohort including 645 cases (mean age 79.2 $\pm$ 9.1 years, 41.2% male) recruited from Shanghai and Zhejiang Province and 3537 controls from the China Metabolic Analytics Project (ChinaMAP) (Cao et al., 2020). Details about genomic DNA collection, library construction, and deep whole genome sequencing (WGS) of samples from ChinaMAP were described in the original study (Cao et al., 2020). In total, 1280 AD cases and 1507 unrelated healthy controls receiving targeted NGS and 3537 controls from the ChinaMAP (Cao et al., 2020) were analyzed in current study. Written informed consents were obtained from all participants or their guardians (for those who were unable to take care of themselves) prior to this study. This study was approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences.

2.2. Gene selection and targeted NGS

Twenty-three GFs were selected from NCBI (https://www.ncbi.nlm.nih.gov/) with verified classification from Immport (https://www.immport.org/shared/genelists) (Bhattacharya et al., 2018). Genomic DNA extraction, library construction, targeted region capture was conducted as described previously (Wang et al., 2018b). In brief, DNA was extracted from whole blood with the AxyPrep Blood Genomic DNA Miniprep Kit (Axygen Scientific). Coding region and nearby UTRs of selected genes were captured with the NimbleGen SeqCap EZ Choice Enrichment Kit (Roche NimbleGen) according to the manufacturer's instructions. DNA probes were designed using the online NimbleDesign tool. Illumina Hiseq X Ten and Novaseq 6000 Genome Analyzers were used to sequence the DNA libraries.

2.3. NGS data processing and quality control

Raw reads were trimmed and filtered with Trimmomatic (Bolger et al., 2014) and mapped to the human reference genome (GRCh37/hg19) by using the Burrows-Wheeler Aligner (Li and Durbin, 2009). Base quality score recalibration and variant calling were performed by using the Genome Analysis Toolkit (GATK v4.1) following the best practices pipeline (https://www.broadinstitute.org/gatk/guide/best-practices) (McKenna et al., 2010). Variants with quality by depth (QD) $<$ 2, Fisher's strand bias (FS) $>$ 60, strand odds ratio (SOR) $>$ 3, read position rank sum (ReadPosRankSum) $<$ -3, mapping quality (MQ) $<$ 40, or mapping quality rank sum (MQRankSum) $<$ -10 were filtered (McKenna et al., 2010; Raghavan et al., 2018). Genotypes with genotyping quality (GQ) less than 30, or heterozygous sites violated from allele balance (AB, AB $<$ 0.2 or AB $>$ 0.8) (Muyas et al., 2019) were set as missing. All variants including single nucleotide variants (SNVs) and insertions/deletions were annotated with ANNOVAR (Wang et al., 2010). The inclusion threshold for common and rare variants was $\geq$ 20$\times$ sequencing depth. Variants with genotyping rate less than 90%, or deviated from Hardy-Weinberg equilibrium (HWE $p < 1 \times 10^{-6}$) were excluded. Samples with average genotyping rate $<$ 80% were also excluded (Chang et al., 2015).

2.4. Cross validation in European populations

We used the reported data of European populations for cross validation. Briefly, the summary data from a newly published large GWAS meta-analysis (Bellenguez et al., 2022) which contains a total of 111,326 clinically diagnosed/'proxy' AD cases and 677,663 controls was used for cross-validating the associations of GF variants with AD identified in Han Chinese in this study. Summary statistics in this GWAS study are available under accession number GCST90027158 (https://www.ebi.ac.uk/gwas/).

2. 5 eQTL analysis

For the identified non-coding common variants associated with AD, we evaluated whether they may affect the expression of their corresponding gene(s). cis-eQTL datasets were extracted from peripheral blood eQTL data from the Consortium for the Architecture of Gene Expression (CAGE) (Lloyd-Jones et al., 2017) and brain eQTL data from a meta-analysis (Qi et al., 2018) of GTEx brain (GTEx Consortium et al., 2017), CommonMind Consortium (CMC) (Fromer et al., 2016) and ROSMAP (Ng et al., 2017).

2. 6 mRNA expression profiling

The genes, whose expression may be regulated by common variants as indicated by the eQTL data, were validated as to whether they were differentially expressed in brain regions or/and single cell type(s) between AD patients and controls. Bulk-tissue mRNA expression in 4 brain regions (entorhinal cortex, hippocampus, temporal cortex, and frontal cortex) were extracted from our previous study at the AlzData (http://www.alzdata.org/), which was based on a systematic integrated analysis of expression profiles of AD-affected brain tissues from 684 AD patients and 562 controls (Xu et al., 2018). Single-nucleus transcriptomes from prefrontal cortex of 48 individuals with varying degrees of AD pathology (no-pathology, early-pathology and late-pathology) were used to investigate the altered GF(s) during AD pathology. mRNA expression in 6 major brain cell types including excitatory (Ex) and inhibitory (In) neurons, astrocytes (Ast), oligodendrocytes (Oli), oligodendrocyte precursor cells (Opc), and microglia (Mic) were analyzed in the original study (Mathys et al., 2019).

2.7. Statistical analysis

The gene-set-based burden test was conducted using the optimized sequence Kernel association test (SKAT-O) (Lee et al., 2012). Rare variants within all 23 genes were put together to evaluate the association of GFs with AD risk. The association was determined based on the following categories of rare variants: loss-of-function (LoF), possibly pathogenic, and rare missense variants. Variants belonging to stop gain/loss, frameshift indels, initiation codon, and splice sites were defined as LoF. Possibly pathogenic rare missense variants were defined by using the Mendelian Clinically Applicable Pathogenicity (M-CAP) (Jagadeesh et al., 2016), which is a highly sensitive pathogenic classifier combining CADD (Amendola et al., 2015), SIFT (Ng and Henikoff, 2003), and PolyPhen-2 (Adzhubei et al., 2010) with novel features and a powerful model. Missense variants were defined as rare if they had a MAF < 0.01 in the control population. The effects of rare variants within each gene on AD were assessed both in single cohorts and the combined cohort in targeted sequencing. The gene-based burden test was also conducted by using SKAT-O (Lee et al., 2012). Rare variants defined as LoF, possibly pathogenic and missense within each gene were put together to evaluate the association of each gene with AD risk. Population frequency of each rare missense and LoF variant in East Asian population from the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/) (Karczewski et al., 2020) were retrieved for the AD-associated gene(s) identified in Han Chinese. As we did not have the detailed genotype data of each individual from the ChinaMAP (Cao et al., 2020), we performed the burden tests only using 1507 healthy samples from the Southern cohort as the control for comparison. We used a stringent Bonferroni correction to help to identify the most associated GF gene with AD.

Association between common variants and AD was measured by Fisher's exact test. A fixed effects meta-analysis combined the 2 cohorts was performed using the metafor R package (Viechtbauer, 2010). Cochran's Q-test was applied to test for residual heterogeneity. Quanto software (Gauderman, 2002) was used to evaluate the statistical power of our samples under the gene only hypothesis and log additive model. A p value < 0.05 was defined as the threshold of nominal significance. The Bonferroni correction for statistical significance was calculated on the basis of the corresponding numbers of tested variants or genes. Linkage disequilibrium (LD) plots of the identified variants in Asian and European populations were constructed by LocusZoom (http://locuszoom.sph.umich.edu/) (Pruim et al., 2010).

Genes were defined as differentially expressed genes (DEGs) in bulk-tissue with FDR < 0.05 (Xu et al., 2018). DEGs in single nucleus transcriptome met the criteria of FDR-adjusted $p < 0.01$ and absolute log2 fold change > 0.25, as described in the original research (Mathys et al., 2019).

3. Results

3.1. Variants identified by targeted NGS

The mean sequencing depth of all 23 genes was higher than 100× (Supplementary Table S1). The sequencing depth for all of the variants ranged from 26× to 187× (mean = 46). After quality control, 2431 rare variants and 171 common variants of the 23 GF genes in 6324 individuals with or without AD were included in subsequent analyses. Rare missense variants were subjected to a burden test and common variants were subjected to a single-site association analysis (Fig. 1). The Bonferroni corrected significance was defined as $p < 2.92 \times 10^{-4}$ (0.05/171) for a single common variant and $p < 2.17 \times 10^{-3}$ (0.05/23) for gene-based analysis.

We used Cochran's Q-test to evaluate potential population heterogeneity between Southern and Eastern cohorts, and found a relatively low level of heterogeneity ($I^2 < 30\%$, $p > 0.05$) for most of the common and rare variants (Supplementary Table S3 and S5). The odds ratios of these variants indicated a consistent direction of genetic effect between Southern and Eastern cohorts (Table 3**)**.

3.2. Significant enrichment of rare variants in the GF gene set in AD patients

The advantage of targeted deep sequencing is the identification of rare coding variants. We performed the SKAT-O burden test (Lee et al., 2012) by combining all rare variants of the global GF pathways. In samples from Southern China, LoF variants reached a statistical significance ($p = 0.023$). All 3 kinds of rare variants (LoF, possibly pathogenic and missense) in GFs were significantly enriched in AD patients in the combined cohort containing 1280 AD patients and 1507 controls. Rare missense variants showed the most significant enrichment ($p = 6.08 \times 10^{-4}$), while LoF ($p = 0.005$) and possibly pathogenic variants ($p = 0.004$) also contributed to the association between GFs and AD risk (Table 1). This result suggested a potentially important effect of GF rare coding variants on AD. Note that such a comparison might be biased as the control sample (1507 individuals) was from the Southern cohort.

3.3. Significant enrichment of rare variants in the EGF gene in AD patients

In order to identify the exact gene accounting for the gene-set level association, we conducted a SKAT-O burden test analysis using rare variants in each cohort (Fig. 1 and Table 2). Five genes including *EGF, GDNF, HGF, PDGFRB* and *VEGFC* passed the nominal significance of $p < 0.05$ in the combined targeting samples. The association of rare missense variants of *EGF, HGF, PDGFRB* and *VEGFC* became much stronger after a combination of the targeting samples, indicating the requirement of a larger sample size for genetic association analysis of AD in Han Chinese. Some of the GFs showed association with AD only in 1 cohort, and the association faded away (*EGFR, IGF1R* and *LTBP1*) or became weaker (*GDNF*) in the combined cohort. The inconsistent association pattern between

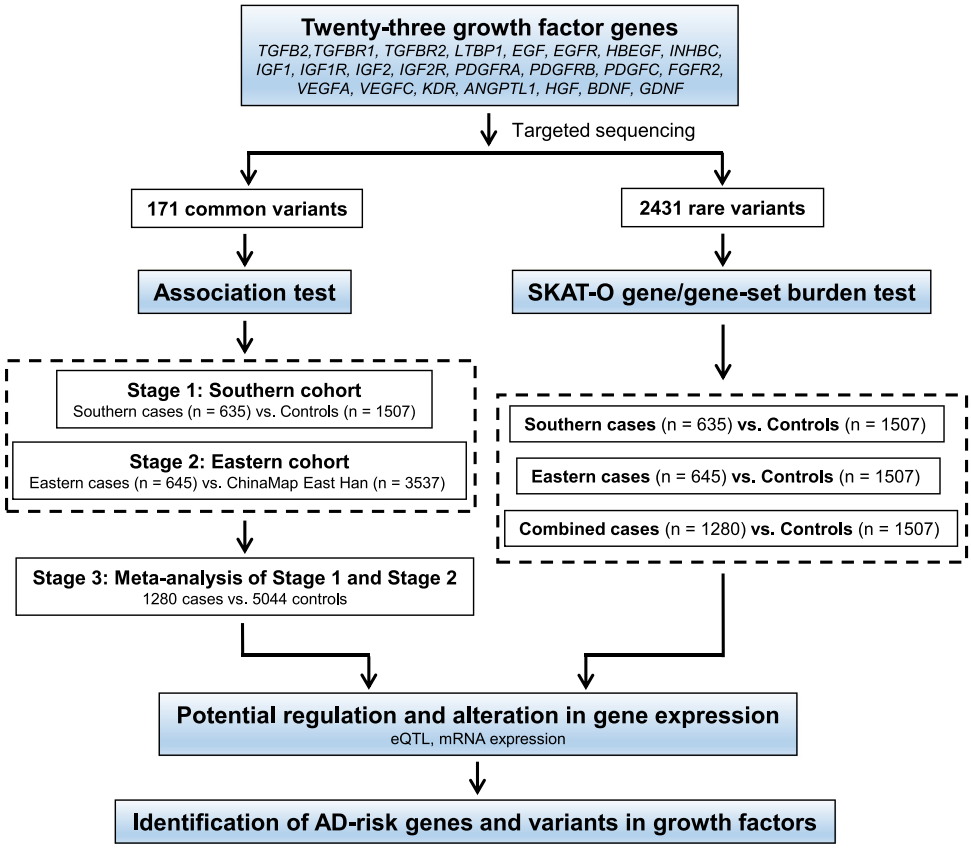


Fig. 1. The strategy for identifying AD risk genes and variants in growth factors. Common and rare variants were subjected to association analysis and burden test, respectively. A total of 3537 Eastern Han Chinese individuals with WGS data were recruited from ChinaMAP (Cao et al., 2020) and were used in the association test.

Table 1
Gene-set-based burden test showing an enrichment of rare variants in growth factors in Han Chinese with AD.

Type of rare variants	Stage 1 (Southern cohort) (635 cases / 1507 controls)	Stage 2 (Eastern cohort) (645 cases / 1507 controls)	Combined (1280 cases / 1507 controls)
LoF	0.023 (10)	0.011 (8)	0.005 (10)
Possibly pathogenic	0.158 (374)	0.007 (362)	0.004 (450)
Missense	0.203 (598)	1.46×10^{-4} (570)	6.08×10^{-4} (713)

Shown values refer to p values estimated by the gene-set-based burden test, and the numbers of variants in all 23 genes used for the test are in parentheses.
Key: LoF, loss-of-function variants.

different cohorts might be partially caused by limited sample size and/or clinical heterogeneity.

The *EGF* gene was associated with AD in both cohorts in the gene-based burden test, and the association became much stronger in the combined samples. The association between *EGF* and AD was driven by 3 LoF variants ($p = 0.0019$), 25 possibly pathogenic variants ($p = 0.010$) and 38 rare missense variants ($p = 0.013$). Only the LoF variants in *EGF* passed the threshold of statistical significance after the Bonferroni correction for the number of tested genes. The 3 LoF variants (4_110882093_C_A, p.Y379*; rs369702571, p.R394*; and rs556105355, p.R1163*) were predicted to be stop codon mutation in *EGF*. Among them, p.R1163* (T allele of rs556105355) was significantly enriched in AD patients (Fig. 2) with MAF of 0.003 ($p = 1.55 \times 10^{-3}$, OR = 8.66 in single-site association analysis) (Supplementary Table S2 and S3). The frequency of p.R1163* was 0.0004 in the ChinaMAP controls (Supplementary Table S2) and 0.002 in the population controls of East Asian from gnomAD (Supplementary Table S4), adding more support for the enrichment of this LoF variant in AD patients. This variant had a mean sequencing depth of 48× (Supplementary Table S2), providing a high confidence for base-calling, and the variant was further verified by Sanger sequencing (Supplementary Fig. S1). Taken together, *EGF* was shown to be a novel risk gene for AD as it had an enrichment of potentially harmful rare variants.

3.4. Association of a common variant rs4698800 in EGF *with AD in Han Chinese*

Besides the rare variants, we identified a total of 171 common variants in the 23 GFs. The MAF of these variants ranged from 0.01 to 0.5 in the combined control samples. We tested the association of common variants in GFs with AD risk in each cohort and the combined sample. The power to detect an odds ratio (OR) value as 1.25 for a risk allele ranged from 18.7% (MAF = 0.01) to 99.9% (MAF = 0.5). Considering an average population MAF of 0.1, the power was above 88.8% (Supplementary Fig. S2). We listed the association results for all the common variants in each independent and the combined cohort in Supplementary Table S5, and listed the variants which reached nominal significance ($p < 0.05$) in the combined cohort in Table 3. There were 9 single

Table 2
Gene-based burden test of rare variants in *EGF* and other growth factor genes in Han Chinese patients with AD.

Gene	Stage 1 (Southern cohort) (635 cases / 1507 controls)			Stage 2 (Eastern cohort) (645 cases / 1507 controls)			Combined (1280 cases / 1507 controls)		
	LoF	Possibly pathogenic	Missense	LoF	Possibly pathogenic	Missense	LoF	Possibly pathogenic	Missense
ANGPTL1	0.648 (1)	0.750 (8)	0.387 (17)	0.650 (1)	0.855 (8)	0.616 (17)	0.730 (1)	0.847 (8)	0.550 (18)
BDNF	NA	0.088 (5)	0.175 (8)	NA	0.706 (3)	0.269 (7)	NA	0.309 (6)	0.212 (10)
EGF	0.008 (3)[a]	0.046 (20)[a]	0.189 (32)	0.005 (2)[a]	0.014 (22)[a]	0.016 (31)[a]	0.0019 (3)[a]	0.010 (25)[a]	0.013 (38)[a]
EGFR	0.648 (1)	0.019 (20)[a]	0.075 (28)	0.650 (1)	0.070 (23)	0.399 (30)	0.730 (1)	0.051 (28)	0.250 (37)
FGFR2	NA	0.986 (14)	0.824 (20)	NA	0.944 (14)	0.897 (21)	NA	0.975 (17)	0.805 (24)
GDNF	NA	0.344 (4)	0.344 (4)	NA	0.003 (4)[a]	0.003 (4)[a]	NA	0.018 (5)[a]	0.018 (5)[a]
HBEGF	0.648 (1)	0.926 (2)	0.872 (8)	0.650 (1)	0.853 (3)	0.990 (8)	0.730 (1)	0.814 (3)	0.966 (9)
HGF	0.648 (1)	0.290 (15)	0.320 (18)	0.650 (1)	0.243 (15)	0.077 (19)	0.730 (1)	0.082 (19)	0.043 (24)[a]
IGF1	NA	0.670 (9)	0.604 (10)	NA	0.459 (10)	0.655 (11)	NA	0.411 (10)	0.439 (11)
IGF1R	0.753 (2)	0.965 (20)	1.000 (34)	0.755 (2)	0.627 (20)	0.037 (30)[a]	0.357 (2)	0.855 (25)	0.250 (40)
IGF2	NA	0.150 (3)	0.521 (4)	NA	0.153 (3)	0.522 (4)	NA	0.102 (4)	0.139 (5)
IGF2R	NA	0.987 (23)	0.906 (75)	NA	0.528 (26)	0.423 (78)	NA	0.750 (30)	0.548 (93)
INHBC	0.148 (1)	0.363 (6)	0.622 (9)	1.000 (0)	0.848 (5)	0.773 (9)	0.230 (1)	0.806 (7)	0.944 (11)
KDR	NA	0.112 (28)	0.182 (41)	NA	0.452 (26)	0.658 (37)	NA	0.314 (32)	0.529 (46)
LTBP1	1.000 (0)	0.373 (55)	0.515 (62)	1.000 (0)	0.144 (59)	0.007 (65)[a]	1.000 (0)	0.147 (69)	0.053 (76)
PDGFC	NA	0.666 (3)	0.661 (4)	NA	0.119 (3)	0.114 (4)	NA	0.383 (3)	0.444 (4)
PDGFRA	NA	0.084 (19)	0.089 (35)	NA	0.468 (13)	0.592 (26)	NA	0.138 (20)	0.220 (39)
PDGFRB	NA	0.489 (19)	0.301 (34)	NA	0.030 (19)[a]	0.027 (33)[a]	NA	0.215 (24)	0.013 (42)[a]
TGFB2	NA	0.753 (2)	0.914 (5)	NA	0.779 (3)	0.784 (5)	NA	0.798 (3)	0.917 (6)
TGFBR1	NA	0.525 (3)	0.523 (5)	NA	0.203 (2)	0.479 (5)	NA	0.215 (3)	0.254 (6)
TGFBR2	NA	0.487 (14)	0.364 (20)	NA	0.673 (9)	0.650 (11)	NA	0.464 (14)	0.522 (20)
VEGFA	NA	0.660 (5)	0.874 (11)	NA	0.412 (5)	0.659 (10)	NA	0.396 (7)	0.648 (13)
VEGFC	NA	0.045 (1)	0.030 (13)[a]	NA	0.648 (1)	0.063 (13)	NA	0.119 (1)	0.0024 (15)[a]

Shown values refer to p values estimated by the gene-based burden test, and the numbers of tested variants in each gene are in parentheses.
Key: NA, not available.
[a] $p < 0.05$.

Table 3
Common genetic variants of growth factors in 2 Han Chinese cohorts.

Gene	SNP	Chr:position	Alt/Ref	Stage 1 (Southern cohort) (635 cases / 1507 controls)			Stage 2 (Eastern cohort) (645 cases / 3537 controls)			Stage 3 (meta-analysis) (1280 cases / 5044 controls)		
				MAF	*p*	OR (95% CI)	MAF	*p*	OR (95% CI)	*p*	OR (95% CI)	Het *p*
EGF	rs4698800	4:110866508	T/C	0.271/0.239	0.037	1.19 (1.01-1.39)	0.248/0.200	2.93×10^{-4}	1.32 (1.13-1.53)	3.24×10^{-5}	1.26 (1.13-1.40)	0.339
EGF	rs10470911	4:110865271	G/T	0.244/0.237	0.606	1.04 (0.89-1.22)	0.236/0.200	0.004	1.23 (1.07-1.43)	0.013	1.14 (1.03-1.27)	0.119
PDGFRB	rs200684708	5:149512332	A/G	0.015/0.013	0.669	1.12 (0.61-1.99)	0.012/0.006	0.008	2.28 (1.18-4.18)	0.029	1.56 (1.05-2.34)	0.085
FGFR2	rs1613776	10:123244834	T/C	0.018/0.017	0.794	1.07 (0.61-1.81)	0.021/0.011	0.008	1.90 (1.16-3.00)	0.023	1.47 (1.05-2.06)	0.097
PDGFRB	rs246391	5:149497177	C/T	0.110/0.123	0.241	0.88 (0.70-1.09)	0.102/0.122	0.053	0.82 (0.67-1.00)	0.025	0.85 (0.73-0.98)	0.658
HGF	rs10272030	7:81350223	G/A	0.131/0.145	0.262	0.89 (0.73-1.09)	0.126/0.146	0.059	0.84 (0.70-1.01)	0.030	0.86 (0.76-0.99)	0.653
IGF2R	rs2277070	6:160445793	G/A	0.304/0.272	0.040	1.17 (1.00-1.36)	0.283/0.266	0.242	1.09 (0.95-1.25)	0.023	1.12 (1.02-1.24)	0.478
HGF	rs12536657	7:81350208	A/G	0.131/0.145	0.246	0.89 (0.73-1.08)	0.129/0.146	0.108	0.86 (0.72-1.03)	0.045	0.87 (0.77-1.00)	0.807
HGF	rs1800793	7:81346685	T/C	0.128/0.142	0.245	0.89 (0.73-1.08)	0.127/0.144	0.118	0.87 (0.72-1.04)	0.049	0.88 (0.77-1.00)	0.847

Only common variants with $p < 0.05$ in meta-analysis are listed. Meta-analysis was conducted by combining the cohorts from stage 1 and stage 2.
Cases and controls of the Southern cohort in stage 1 were recruited from Sichuan, Hunan and Yunnan Province of China, and were sequenced in this study.
Cases of Eastern cohort in stage 2 were recruited from Shanghai and Zhejiang Province of China and were sequenced in this study. The Eastern Han Chinese (n = 3537) from the ChinaMAP (Cao et al., 2020) were used as a control for comparion with cases of Eastern cohort.
Key: Alt/Ref, alternative allele/reference allele; Chr, chromosome; CI, confidence interval; Het *P*, *p* value for heterogeneity statistic. MAF, minor allele frequencies of case/control; OR, odds ratio; SNP, single nucleotide polymorphism.

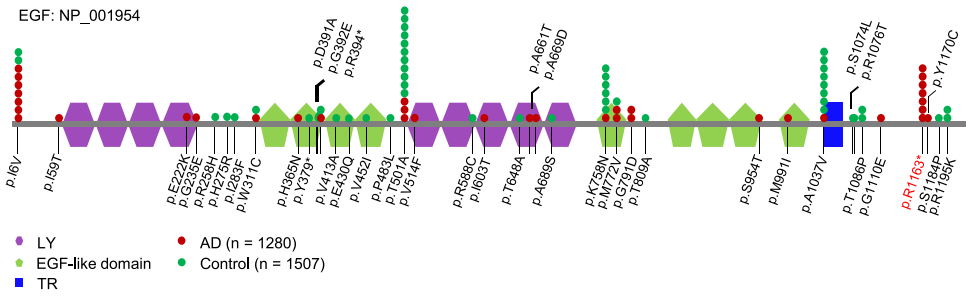


Fig. 2. Distribution of rare damaging variants of *EGF* in AD and control samples from Han Chinese population in targeted sequencing. Schematic profile of EGF was predicted by SMART (http://smart.embl-heidelberg.de/). LY, low-density lipoprotein-receptor YWTD domain; TM, transmembrane region. The AD-enriched LoF variant p.R1163* was marked in (red).

nucleotide polymorphisms (SNPs) above the nominally significant threshold in the combined cohort. Among these SNPs, 5 were nominally significant in at least 1 cohort, whereas the other 4 SNPs showed no significance in individual cohort. The OR values of each SNP indicated a consistent direction of genetic effect on disease risk between the Southern cohort and the Eastern cohort. Variant rs4698800 in *EGF* showed a nominally but consistently significant association with AD risk in both the Southern ($p = 0.037$, OR = 1.19) and the Eastern cohort ($p = 2.93 \times 10^{-4}$, OR = 1.32). This SNP remained statistically significant in the combined cohort ($p = 3.24 \times 10^{-5}$, OR = 1.26), even after the Bonferroni correction (adjusted $p = 5.54 \times 10^{-3}$). Another *EGF* variant rs10470911, which is in strong LD with rs4698800 ($r^2 > 0.8$) (Fig. 3), also showed nominally significant association with AD risk in the combined cohort ($p = 0.013$, OR = 1.14), adding more support for the association of *EGF* variants with AD (Table 3).

3.5. No association of EGF variants with AD in European population

To investigate whether the association of *EGF* variants with AD is population specific, we retrieved the association results of *EGF* SNPs from a large GWAS study (Bellenguez et al., 2022). Sixteen common variants and 7 rare variants identified in the current study were reported in this latest GWAS study. None of these variants of *EGF* was associated with AD in European population in genome-wide level (Supplementary Table S6). LD plots showed different linkage patterns of some of these SNPs with rs4698800 between Asian and European populations (Fig. 3), suggesting that population heterogeneity of *EGF* variants would affect AD susceptibility.

3.6. Regulatory effects of rs4698800 on EGF expression

Since the AD-associated common *EGF* variants were located in non-coding region, we evaluated the potentially regulatory effects of rs4698800 and rs10470911 on *EGF* expression by using the expression quantitative trait loci (eQTL) datasets of human brain (Qi et al., 2018) and whole blood (Lloyd-Jones et al., 2017). Both rs4698800 and rs10470911 showed regulatory effects on *EGF* mRNA expression in whole blood (Lloyd-Jones et al., 2017) and brain (Qi et al., 2018) (Supplementary Table S7). In particular, risk alleles rs4698800-T and rs10470911-G were associated with a higher expression of *EGF* mRNA in brain tissues than the other alleles of both SNPs (Fig. 3). We also investigated whether the index SNPs were located in regulatory elements of the *EGF* gene using the functional genomic annotations of the ENCODE data (Encode Project Consortium, 2012). The ENCODE annotations for enhancers (H3K27ac), chromatin accessibility (DNaseI hypersensitivity sites), and transcription factor binding sites (TFBSs) were retrieved from the UCSC Genome Browser (https://genome.ucsc.edu/). We found that rs4698800 was predicted to be located in a DNaseI hypersensitivity site and multiple TFBSs including FOSL2, JUND and FOS (Supplementary Fig. S3), and rs10470911 was not a regulatory variant. The regulatory effect of rs4698800 might account for the association with AD risk, and it would be rewarding to confirm this speculation by experimental assays.

3.7. Altered EGF mRNA expression in AD hippocampus

Differential expression of *EGF* mRNA between AD patients and controls were examined in 4 brain regions including entorhinal cortex, hippocampus, temporal cortex, and frontal cortex using the normalized microarray data compiled in our previous study (http://www.alzdata.org) (Xu et al., 2018). *EGF* was significantly upregulated in the hippocampus, which is one of the most vulnerable brain regions in AD patients (Mu and Gage, 2011) (Fig. 4). Also, we examined the mRNA expression pattern in 6 major brain cell types (Mathys et al., 2019). No change in the *EGF* mRNA expression pattern was observed among subjects with no-pathology, early-pathology or late-pathology (Supplementary Table S8 and S9), but the level of *EGFR* was elevated in oligodendrocytes in AD brain as compared to those without pathology (Supplementary Table S8).

4. Discussion

Many GFs, including the TGFs, IGFs, FGFs and NGFs, are naturally expressed in the brain and play important roles in the CNS, and have been suggested to influence AD progression (Schindowski et al., 2008; Woodbury and Ikezu, 2014; Lauzon et al., 2015). Recent genetic association studies have also shown an association between GF genetic variants and AD risk (He et al., 2021; Bellenguez et al., 2022). In this study, we examined the association of rare and common variants of GFs in Han Chinese cohorts containing 1280 AD patients and 5044 controls from the general population. By using the control samples from the ChinaMAP (Cao et al., 2020) for comparison, we were able to minimize the limitation of small sample size and achieve a robust association. We showed that *EGF* was a novel risk gene for AD in Han Chinese.

By using the gene-set-based burden test for rare variants in the 23 GF genes, we found that rare LoF variants in the GFs might contribute to the susceptibility of AD, as LoF variants were associated with AD in all 3 comparisons, with the most significant association in the combined samples (Table 1). The possibly pathogenic and rare missense variants were not significantly associated with AD in the Southern cohort, yet the association became significant in the combined samples. These results indicated the involvement of rare damaging variants, especially LoF of GFs in AD etiology. In the gene-based burden test, several GFs showed nominally significant enrichment of rare damaging variants in AD patients. Some GFs, including *PDGFRB, EGFR, IGF1R*, and *LTBP1* showed associations with AD only in 1 cohort (Table 2). Limited sample size and/or

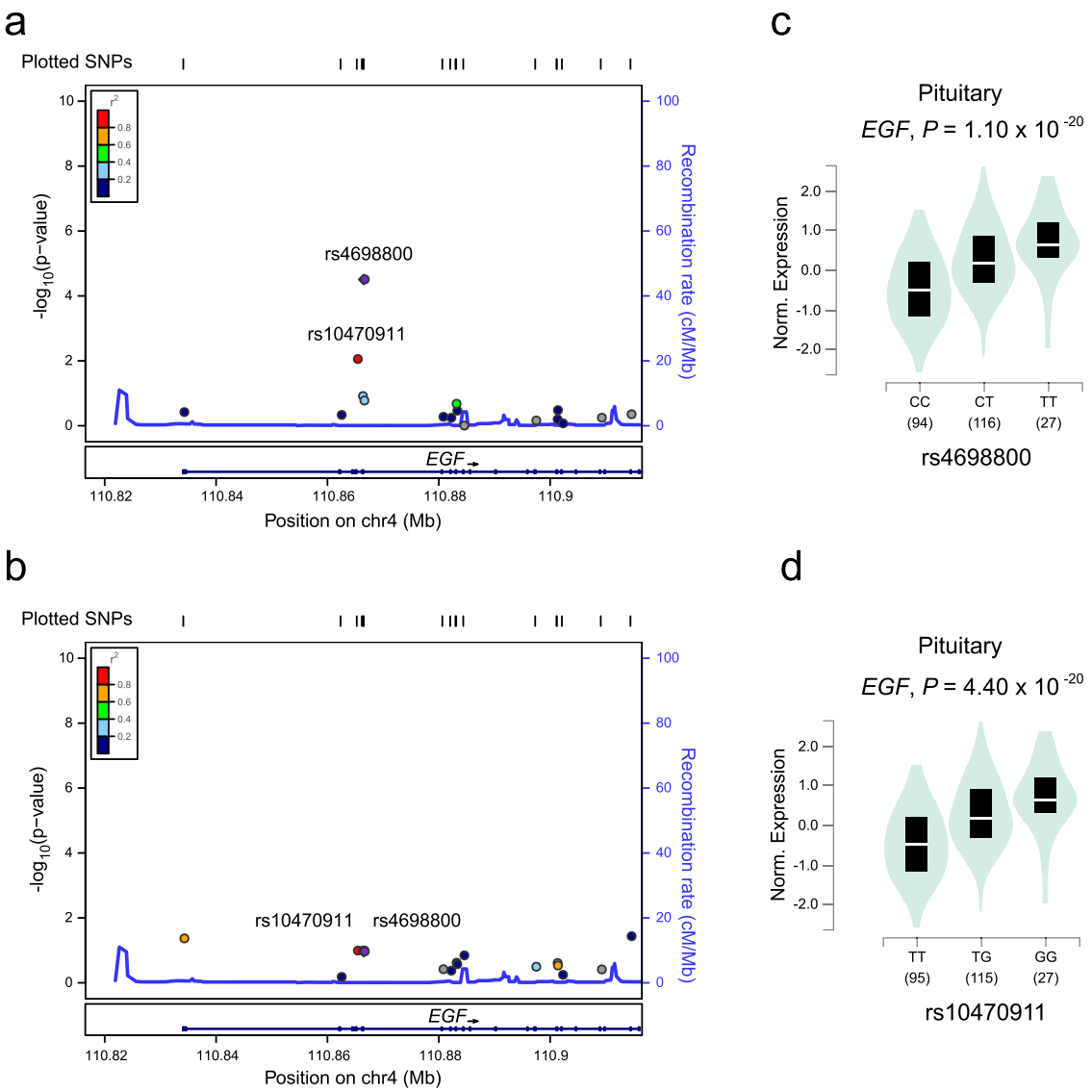


Fig. 3. Common variants of *EGF* were associated with AD risk and *EGF* mRNA expression. (a and b) LocusZoom plot of rs4698800 in Asian and European populations. *p* values of single variants of *EGF* in meta-analysis of the current study and summary statistics from the meta-analysis in GWAS study (Bellenguez et al., 2022) were used. The queried SNP rs4698800 was marked in purple. The pairwise LD patterns of other SNPs with rs4698800 were indicated by different colors. (c and d) Effects of rs4698800 and rs10470911 on *EGF* mRNA expression. Data was downloaded from GTEx portal (https://www.gtexportal.org/home).

clinical heterogeneity might contribute to this pattern, and independent samples with sufficient demographic and clinical data are needed for validating the association of these GFs with AD.

Despite of the nominal associations, we observed a robust association between *EGF* and AD, especially for LoF variants. Moreover, we also found that 2 common variants in *EGF*, out of 171 common variants of 23 GFs, were significantly associated with AD risk in Han Chinese. These findings gave further support for an active role of EGF in the development of AD. Note that the association between *EGF* variants and AD showed a population specific pattern, as it was not found in European populations and populations of European ancestry (Bellenguez et al., 2022). These findings, together with our previous observations for a population-specific pattern of association between AD and certain genes (Wang et al., 2016; Zhang et al., 2016), indicated that population heterogeneity and genetic background may affect the genetic susceptibility to human diseases. Intriguingly, the newly published large-scale GWASs highlighted the role of EGFR in AD in European population (He et al., 2021; Bellenguez et al., 2022), suggesting the robust involvement of the EGF signaling in AD development. Genetic analyses of different populations would provide convergent insights into the pathogenesis of AD.

Further characterization of the risk alleles of rs4698800 and rs10470911 using available datasets showed that these risk alleles were associated with a relatively higher *EGF* mRNA expression in the human brain (Fig. 3). This result was consistent with the observation for an increase in *EGF* mRNA expression in the hippocampus of AD patients (Fig. 4), indicating that upregulation of *EGF* may be deleterious for AD. Our results were consistent with a recent GWAS study that the protective allele of the lead variant in *EGFR* is associated with lower *EGFR* expression and that genetic downregulation of *EGFR* expression is associated with lower AD risk (Bellenguez et al., 2022). The mRNA expression of *EGFR* was elevated in oligodendrocytes in AD brain (Supplementary Table S8) (Mathys et al., 2019). All these studies indicated a higher risk of AD in individuals with higher levels of *EGF* (our study) and *EGFR* expression (Mathys et al., 2019; Bellenguez et al., 2022), and were different from the 2 studies showing a low plasma EGF level

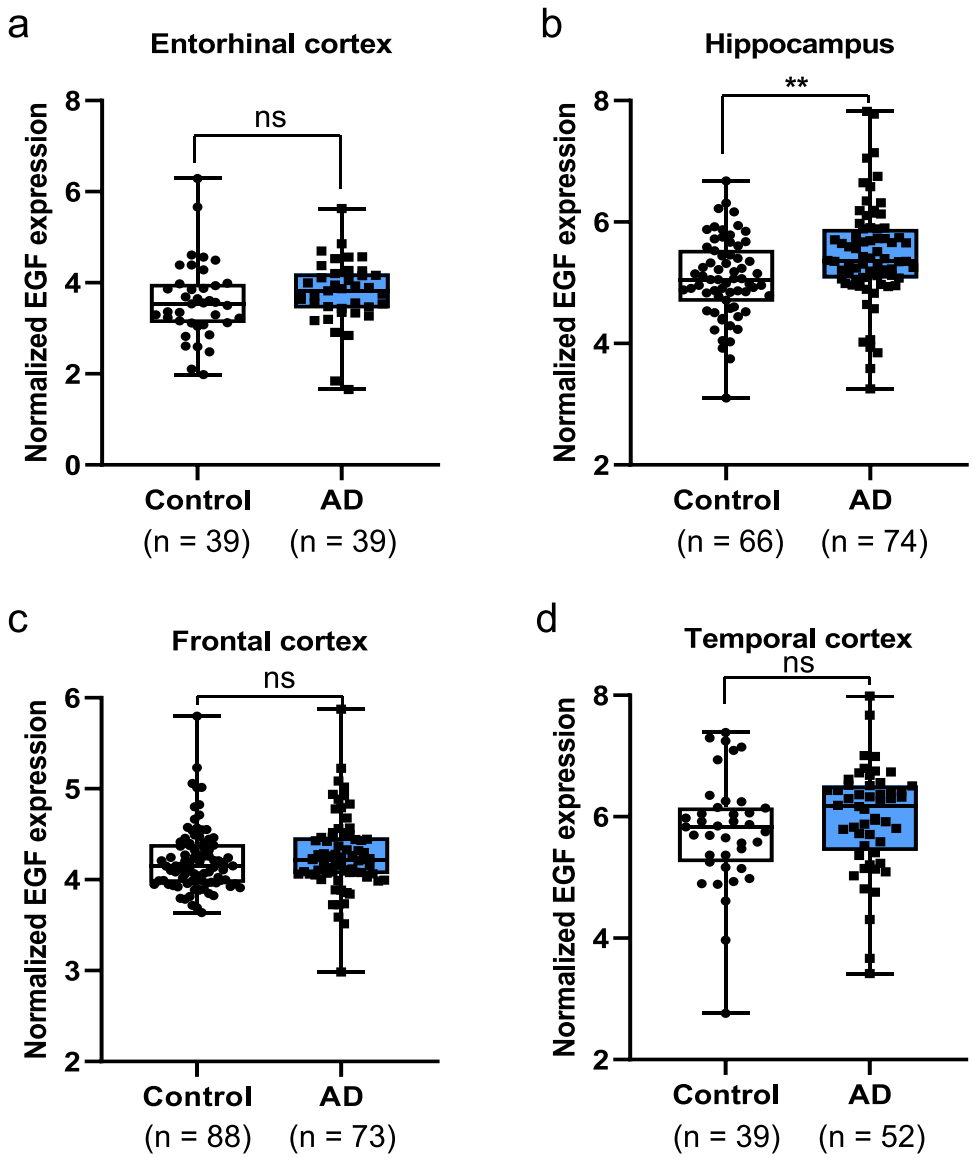


Fig. 4. Upregulation of *EGF* mRNA expression in hippocampus tissue of AD patients. The mRNA expression data of *EGF* from entorhinal cortex (a), hippocampus (b), frontal cortex (c) and temporal cortex (d) were retrieved from the AlzData (www.alzdata.org) (Xu et al., 2018). Two-tailed Student's t-test was used to test for the statistical difference between AD patients and controls. Ns, not significant; $^{**}p < 0.01$.

in AD-MCI patients (Lim et al., 2016) and a potential prevention effect of EGF on *APOE4* and Aß-induced cognitive decline in female mice aged 8 months (Thomas et al., 2016). By using single cell data from AD patients reported by Mathys et al. (2019), we found that *EGF* mRNA expression showed no apparent alteration in each of the 6 major cell types in AD brain (Supplementary Table S8 and S9). These results suggested that the effect of EGF on AD should be considered comprehensively in the context of disease stages and tissue/cell types. As we had no genotype information for these donors for single cell transcriptome analysis, we speculated that these donors may not contain the *EGF* risk alleles found in Han Chinese, because the association of *EGF* genetic variants with AD showed an apparent population-specific feature as described in the text above.

This study had several limitations. First, although our sample size had a reasonably good statistical power for association analysis, the sample size was still relatively small. The potential population substructure may further blur the association of *EGF* with AD. Second, the different sequencing depth between ChinaMAP (Cao et al., 2020) and our targeted sequencing study and the batch effect may also affect the identification of potential risk variants. Third, we did not perform any functional characterization for the potential damaging variants identified in *EGF*, nor characterize the function of EGF in the development of AD. More focused studies are warranted to confirm the association between *EGF* variants and AD and to characterize the role of this growth factor in AD.

In short, we used targeted sequencing to screen for genetic variants in 23 GFs to discern potential association with AD. Our results indicate *EGF* to be a novel risk gene for AD in Han Chinese population, with an enrichment of damaging rare missense mutations and an upregulated mRNA expression in AD patients. Validation studies with independent large samples and functional assays are needed to investigate the effect of increased EGF signaling on the development and progression of AD.

Submission declaration and verification

We declare that the work described in this manuscript has not been previously published nor being considered for publication elsewhere. All authors have been seen and given their ap-

proval for submission of the manuscript to be considered for publication in *Neurobiology of Aging*. If accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

Disclosure statement

The authors have no conflict of interest to declare.

Acknowledgements

We thank all participants in this study and Mr. Quanzhen Zheng for technical assistance. We thank Dr. Ian Logan for language editing of the manuscript. We thank Prof. Yanan Cao for sharing allele frequency data in the ChinaMAP project. This work was supported by the Ministry of Science and Technology of China [grant number 2022ZD0213500], the National Natural Science Foundation of China [grant numbers 31730037, 82022017, 31970965, 32230021, 32200489]; the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (CAS) [grant number XDB02020003]; and the Youth Innovation Promotion Association of CAS.

CRediT authorship contribution statement

Xiao Li: Investigation, Writing – original draft, Writing – review & editing, Visualization. **Min Xu:** Methodology, Writing – review & editing. **Rui Bi:** Investigation. **Li-Wen Tan:** Resources. **Yong-Gang Yao:** Conceptualization, Writing – review & editing, Supervision. **Deng-Feng Zhang:** Conceptualization, Investigation, Writing – review & editing.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neurobiolaging.2022.10.009.

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249.

Amendola, L.M., Dorschner, M.O., Robertson, P.D., Salama, J.S., Hart, R., Shirts, B.H., Murray, M.L., Tokita, M.J., Gallego, C.J., Kim, D.S., Bennett, J.T., Crosslin, D.R., Ranchalis, J., Jones, K.L., Rosenthal, E.A., Jarvik, E.R., Itsara, A., Turner, E.H., Herman, D.S., Schleit, J., Burt, A., Jamal, S.M., Abrudan, J.L., Johnson, A.D., Conlin, L.K., Dulik, M.C., Santani, A., Metterville, D.R., Kelly, M., Foreman, A.K., Lee, K., Taylor, K.D., Guo, X., Crooks, K., Kiedrowski, L.A., Raffel, L.J., Gordon, O., Machini, K., Desnick, R.J., Biesecker, L.G., Lubitz, S.A., Mulchandani, S., Cooper, G.M., Joffe, S., Richards, C.S., Yang, Y., Rotter, J.I., Rich, S.S., O'Donnell, C.J., Berg, J.S., Spinner, N.B., Evans, J.P., Fullerton, S.M., Leppig, K.A., Bennett, R.L., Bird, T., Sybert, V.P., Grady, W.M., Tabor, H.K., Kim, J.H., Bamshad, M.J., Wilfond, B., Motulsky, A.G., Scott, C.R., Pritchard, C.C., Walsh, T.D., Burke, W., Raskind, W.H., Byers, P., Hisama, F.M., Rehm, H., Nickerson, D.A., Jarvik, G.P., 2015. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. Genome Res. 25, 305–315.

Bellenguez, C., Küçükali, F., Jansen, I.E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., Holmans, P.A., Boland, A., Damotte, V., van der Lee, S.J., Costa, M.R., Kuulasmaa, T., Yang, Q., de Rojas, I., Bis, J.C., Yaqub, A., Prokic, I., Chapuis, J., Ahmad, S., Giedraitis, V., Aarsland, D., Garcia-Gonzalez, P., Abdelnour, C., Alarcón-Martín, E., Alcolea, D., Alegret, M., Alvarez, I., Álvarez, V., Armstrong, N.J., Tsolaki, A., Antúnez, C., Appollonio, I., Arcaro, M., Archetti, S., Pastor, A.A., Arosio, B., Athanasiu, L., Bailly, H., Banaj, N., Baquero, M., Barral, S., Beiser, A., Pastor, A.B., Below, J.E., Benchek, P., Benussi, L., Berr, C., Besse, C., Bessi, V., Binetti, G., Bizarro, A., Blesa, R., Boada, M., Boerwinkle, E., Borroni, B., Boschi, S., Bossu, P., Bråthen, G., Bressler, J., Bresner, C., Brodaty, H., Brookes, K.J., Brusco, L.I., Buiza-Rueda, D., Bûrger, K., Burholt, V., Bush, W.S., Calero, M., Cantwell, L.B., Chene, G., Chung, J., Cuccaro, M.L., Carracedo, Á., Cecchetti, R., Cervera-Carles, L., Charbonnier, C., Chen, H.H., Chillotti, C., Ciccone, S., Claassen, J.A.H.R., Clark, C, Conti, E., Corma-Gómez, A., Costantini, E., Custodero, C., Daian, D., Dalmasso, M.C., Daniele, A., Dardiotis, E., Dartigues, J.F., de Deyn, P.P., de Paiva Lopes, K., de Witte, L.D., Debette, S., Deckert, J., Del Ser, T., Denning, N., DeStefano, A., Dichgans, M., Diehl-Schmid, J., Diez-Fairen, M., Rossi, P.D., Djurovic, S., Duron, E., Düzel, E., Dufouil, C., Eiriksdottir, G., Engelborghs, S., Escott-Price, V., Espinosa, A., Ewers, M., Faber, K.M., Fabrizio, T., Nielsen, S.F., Fardo, D.W., Farotti, L., Fenoglio, C., Fernandez-Fuertes, M., Ferrari, R., Ferreira, C.B., Ferri, E., Fin, B., Fischer, P., Fladby, T., Fließbach, K., Fongang, B., Fornage, M., Fortea, J., Foroud, T.M., Fostinelli, S., Fox, N.C., Franco-Macías, E., Bullido, M.J., Frank–García, A., Froelich, L., Fulton-Howard, B., Galimberti, D., García-Alberca, J.M., García-González, P., Garcia-Madrona, S., Garcia-Ribas, G., Ghidoni, R., Giegling, I., Giorgio, G., Goate, A.M., Goldhardt, O., Gomez-Fonseca, D., González-Pérez, A., Graff, C., Grande, G., Green, E., Grimmer, T., Grünblatt, E., Grunin, M., Gudnason, V., Guetta-Baranes, T., Haapasalo, A., Hadjigeorgiou, G., Haines, J.L., Hamilton-Nelson, K.L., Hampel, H., Hanon, O., Hardy, J., Hartmann, A.M., Hausner, L., Harwood, J., Heilmann-Heimbach, S., Helisalmi, S., Heneka, M.T., Hernández, I., Herrmann, M.J., Hoffmann, P., Holmes, C., Holstege, H., Vilas, R.H., Hulsman, M., Humphrey, J., Biessels, G.J., Jian, X., Johansson, C., Jun, G.R., Kastumata, Y., Kauwe, J., Kehoe, P.G., Kilander, L., Ståhlbom, A.K., Kivipelto, M., Koivisto, A., Kornhuber, J., Kosmidis, M.H., Kukull, W.A., Kuksa, P.P., Kunkle, B.W., Kuzma, A.B., Lage, C., Laukka, E.J., Launer, L., Lauria, A., Lee, C.Y., Lehtisalo, J., Lerch, O., Lleó, A., Longstreth, W. Jr., Lopez, O., de Munain, A.L., Love, S., Löwemark, M., Luckcuck, L., Lunetta, K.L., Ma, Y., Macías, J, MacLeod, C.A., Maier, W., Mangialasche, F., Spallazzi, M., Marquié, M., Marshall, R., Martin, E.R., Montes, A.M., Rodríguez, C.M., Masullo, C., Mayeux, R., Mead, S., Mecocci, P., Medina, M., Meggy, A., Mehrabian, S., Mendoza, S., Menéndez–González, M., Mir, P., Moebus, S., Mol, M., Molina-Porcel, L., Montrreal, L., Morelli, L., Moreno, F., Morgan, K., Mosley, T, Nöthen, M.M., Muchnik, C., Mukherjee, S., Nacmias, B., Ngandu, T., Nicolas, G., Nordestgaard, B.G., Olaso, R., Orellana, A., Orsini, M., Ortega, G., Padovani, A., Paolo, C., Papenberg, G., Parnetti, L., Pasquier, F., Pastor, P., Peloso, G., Pérez-Cordón, A., Pérez-Tur, J., Pericard, P., Peters, O., Pijnenburg, Y.A.L., Pineda, J.A., Piñol-Ripoll, G., Pisanu, C., Polak, T., Popp, J., Posthuma, D., Priller, J., Puerta, R., Quenez, O., Quintela, I., Thomassen, J.Q., Rábano, A., Rainero, I., Rajabli, F., Ramakers, I., Real, L.M., Reinders, M.J.T., Reitz, C., Reyes-Dumeyer, D., Ridge, P., Riedel-Heller, S., Riederer, P., Roberto, N., Rodriguez-Rodriguez, E., Rongve, A., Allende, I.R., Rosende-Roca, M., Royo, J.L., Rubino, E., Rujescu, D., Sáez, M.E., Sakka, P., Saltvedt, I., Sanabria, Á., Sánchez-Arjona, M.B., Sanchez-Garcia, F., Juan, P.S., Sanchez-Valle, R., Sando, S.B., Sarnowski, C., Satizabal, C.L., Scamosci, M., Scarmeas, N., Scarpini, E., Scheltens, P., Scherbaum, N., Scherer, M., Schmid, M., Schneider, A., Schott, J.M., Selbaek, G., Seripa, D., Serrano, M., Sha, J., Shadrin, A.A., Skrobot, O., Slifer, S., Snijders, G.J.L., Soininen, H., Solfrizzi, V., Solomon, A., Song, Y., Sorbi, S., Sotolongo-Grau, O., Spalletta, G., Spottke, A., Squassina, A., Stordal, E., Tartan, J.P., Tárraga, L., Tesí, N., Thalamuthu, A., Thomas, T., Tosto, G., Traykov, L., Tremolizzo, L., Tybjaerg-Hansen, A., Uitterlinden, A., Ullgren, A., Ulstein, I., Valero, S., Valladares, O., Broeckhoven, C.V., Vance, J., Vardarajan, B.N., van der Lugt, A., Dongen, J.V., van Rooij, J., van Swieten, J., Vandenberghe, R., Verhey, F., Vidal, J.S., Vogelgsang, J., Vyhnalek, M., Wagner, M., Wallon, D., Wang, L.S., Wang, R., Weinhold, L., Wiltfang, J., Windle, G., Woods, B., Yannakoulia, M., Zare, H., Zhao, Y., Zhang, X., Zhu, C., Zulaica, M., EADB, GR@ACE, DEGESCO, EADI, GERAD, Demgene, FinnGen, ADGC, CHARGE, Farrer, L.A., Psaty, B.M., Ghanbari, M., Raj, T., Sachdev, P., Mather, K., Jessen, F., Ikram, M.A., de Mendonça, A., Hort, J., Tsolaki, M., Pericak-Vance, M.A., Amouyel, P., Williams, J., Frikke-Schmidt, R., Clarimon, J., Deleuze, J.F., Rossi, G., Seshadri, S., Andreassen, O.A., Ingelsson, M., Hiltunen, M., Sleegers, K., Schellenberg, G.D., van Duijn, C.M., Sims, R., van der Flier, W.M., Ruiz, A., Ramirez, A., Lambert, J.C., 2022. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat. Genet. 54, 412–436.

Bhattacharya, S., Dunn, P., Thomas, C.G., Smith, B., Schaefer, H., Chen, J., Hu, Z., Zalocusky, K.A., Shankar, R.D., Shen-Orr, S.S., Thomson, E., Wiser, J., Butte, A.J., 2018. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Sci. Data. 5, 180015.

Bi, R., Zhang, W., Zhang, D.F., Xu, M., Fan, Y., Hu, Q.X., Jiang, H.Y., Tan, L., Li, T., Fang, Y., Zhang, C., Yao, Y.G., 2018. Genetic association of the cytochrome c oxidase-related genes with Alzheimer's disease in Han Chinese. Neuropsychopharmacology 43, 2264–2276.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Bottner, M., Krieglstein, K., Unsicker, K., 2000. The transforming growth factor-betas: structure, signaling, and roles in nervous system development and functions. J. Neurochem. 75, 2227–2240.

Cao, Y., Li, L., Xu, M., Feng, Z., Sun, X., Lu, J., Xu, Y., Du, P., Wang, T., Hu, R., Ye, Z., Shi, L., Tang, X., Yan, L., Gao, Z., Chen, G., Zhang, Y., Chen, L., Ning, G., Bi, Y., Wang, W.ChinaMAP Consortium, 2020. The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals. Cell Res. 30, 717–731.

Caraci, F., Gulisano, W., Guida, C.A., Impellizzeri, A.A., Drago, F., Puzzo, D., Palmeri, A., 2015. A key role for TGF-beta1 in hippocampal synaptic plasticity and memory. Sci. Rep. 5, 11252.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7.

Chang, W.W., Zhang, L., Jin, Y.L., Yao, Y.S., 2013. Meta-analysis of the transforming growth factor-beta1 polymorphisms and susceptibility to Alzheimer's disease. J. Neural. Transm. (Vienna) 120, 353–360.

Chapuis, J., Tian, J., Shi, J., Bensemain, F., Cottel, D., Lendon, C., Amouyel, P., Mann, D.,

Lambert, J.C., 2006. Association study of the vascular endothelial growth factor gene with the risk of developing Alzheimer's disease. Neurobiol. Aging 27, 1212–1215.
De Strooper, B., Karran, E., 2016. The cellular phase of Alzheimer's disease. Cell 164, 603–615.
Encode Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.
Ferreira, S.T., 2021. Brain insulin, insulin-like growth factor 1 and glucagon–like peptide 1 signalling in Alzheimer's disease. J. Neuroendocrinol. 33, e12959.
Freiherr, J., Hallschmid, M., Frey 2nd, W.H., Brunner, Y.F., Chapman, C.D., Holscher, C., Craft, S., De Felice, F.G., Benedict, C., 2013. Intranasal insulin as a treatment for Alzheimer's disease: a review of basic research and clinical evidence. CNS Drugs 27, 505–514.
Fromer, M., Roussos, P., Sieberts, S.K., Johnson, J.S., Kavanagh, D.H., Perumal, T.M., Ruderfer, D.M., Oh, E.C., Topol, A., Shah, H.R., Klei, L.L., Kramer, R., Pinto, D., Gümüş, Z.H., Cicek, A.E., Dang, K.K., Browne, A., Lu, C., Xie, L., Readhead, B., Stahl, E.A., Xiao, J., Parvizi, M., Hamamsy, T., Fullard, J.F., Wang, Y.C., Mahajan, M.C., Derry, J.M., Dudley, J.T., Hemby, S.E., Logsdon, B.A., Talbot, K., Raj, T., Bennett, D.A., De Jager, P.L., Zhu, J., Zhang, B., Sullivan, P.F., Chess, A., Purcell, S.M., Shinobu, L.A., Mangravite, L.M., Toyoshiba, H., Gur, R.E., Hahn, C.G., Lewis, D.A., Haroutunian, V., Peters, M.A., Lipska, B.K., Buxbaum, J.D., Schadt, E.E., Hirai, K., Roeder, K., Brennand, K.J., Katsanis, N., Domenici, E., Devlin, B., Sklar, P., 2016. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. 19, 1442–1453.
Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske, A., Pedersen, N.L., 2006. Role of genes and environments for explaining Alzheimer disease. Arch. Gen. Psychiatry 63, 168–174.
Gauderman, W.J., 2002. Sample size requirements for matched case-control studies of gene-environment interaction. Stat. Med. 21, 35–50.
GTEx Consortium, 2017. Genetic effects on gene expression across human tissues. Nature 550, 204–213.
He, Z., Le Guen, Y., Liu, L., Lee, J., Ma, S., Yang, A.C., Liu, X., Rutledge, J., Losada, P.M., Song, B., Belloy, M.E., Butler RR 3rd, Longo, F.M., Tang, H., Mormino, E.C., Wyss–Coray, T., Greicius, M.D., Ionita-Laza, I., 2021. Genome-wide analysis of common and rare variants via multiple knockoffs at biobank scale, with an application to Alzheimer disease genetics. Am. J. Hum. Genet. 108, 2336–2353.
Hohman, T.J., Bell, S.P., Jefferson, A.L., 2015. The role of vascular endothelial growth factor in neurodegeneration and cognitive decline: exploring interactions with biomarkers of Alzheimer disease. JAMA Neurol. 72, 520–529.
Jagadeesh, K.A., Wenger, A.M., Berger, M.J., Guturu, H., Stenson, P.D., Cooper, D.N., Bernstein, J.A., Bejerano, G., 2016. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat. Genet. 48, 1581–1586.
Jia, L., Li, F., Wei, C., Zhu, M., Qu, Q., Qin, W., Tang, Y., Shen, L., Wang, Y., Shen, L., Li, H., Peng, D., Tan, L., Luo, B., Guo, Q., Tang, M., Du, Y., Zhang, J., Zhang, J., Lyu, J., Li, Y., Zhou, A., Wang, F., Chu, C., Song, H., Wu, L., Zuo, X., Han, Y., Liang, J., Wang, Q., Jin, H., Wang, W., Lu, Y., Li, F., Zhou, Y., Zhang, W., Liao, Z., Qiu, Q., Li, Y., Jiao, H., Lu, J., Jia, J., 2021. Prediction of Alzheimer's disease using multi-variants from a Chinese genome-wide association study. Brain 144, 924–937.
Johansson, P., Aberg, D., Johansson, J.O., Mattsson, N., Hansson, O., Ahren, B., Isgaard, J., Aberg, N.D., Blennow, K., Zetterberg, H., Wallin, A., Svensson, J., 2013. Serum but not cerebrospinal fluid levels of insulin-like growth factor-I (IGF-I) and IGF-binding protein-3 (IGFBP-3) are increased in Alzheimer's disease. Psychoneuroendocrinology 38, 1729–1737.
Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., Gauthier, L.D., Brand, H., Solomonson, M., Watts, N.A., Rhodes, D., Singer-Berk, M., England, E.M., Seaby, E.G., Kosmicki, J.A., Walters, R.K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J.X., Samocha, K.E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A.H., Minikel, E.V., Weisburd, B., Lek, M., Ware, J.S., Vittal, C., Armean, I.M., Bergelson, L., Cibulskis, K., Connolly, K.M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M.E., Genome Aggregation Database Consortium, Neale, B.M., Daly, M.J., MacArthur, D.G., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443.
Lambert, J.C., Amouyel, P., 2007. Genetic heterogeneity of Alzheimer's disease: complexity and advances. Psychoneuroendocrinology 32 (1), S62–S70 Suppl.
Lauzon, M.A., Daviau, A., Marcos, B., Faucheux, N., 2015. Growth factor treatment to overcome Alzheimer's dysfunctional signaling. Cell. Signal. 27, 1025–1038.
Lee, S., Emond, M.J., Bamshad, M.J., Barnes, K.C., Rieder, M.J., Nickerson, D.A., NHLBI GO Exome Sequencing Project—ESP Lung Project Team, Christiani, D.C., Wurfel, M.M., Lin, X., 2012. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am. J. Hum. Genet. 91, 224–237.
Li, G.D., Bi, R., Zhang, D.F., Xu, M., Luo, R., Wang, D., Alzheimer's Disease Neuroimaging Initiative (ADNI), Fang, Y., Li, T., Zhang, C., Yao, Y.G., 2017. Female-specific effect of the BDNF gene on Alzheimer's disease. Neurobiol. Aging 53, 192 e111-192 e119.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Li, L., Yang, M., Wang, C., Zhao, Q., Liu, J., Zhan, C., Liu, Z., Li, X., Wang, W., Yang, X., 2012. Effects of cytokines and chemokines on migration of mesenchymal stem cells following spinal cord injury. Neural. Regen. Res. 7, 1106–1112.
Lim, N.S., Swanson, C.R., Cherng, H.R., Unger, T.L., Xie, S.X., Weintraub, D., Marek, K., Stern, M.B., Siderowf, A., Investigators, P., 2016. Plasma EGF and cognitive decline in Parkinson's disease and Alzheimer's disease. Ann. Clin. Transl. Neurol. 3, 346–355.
Liu, S.Y., Zeng, F.F., Chen, Z.W., Wang, C.Y., Zhao, B., Li, K.S., 2013. Vascular endothelial growth factor gene promoter polymorphisms and Alzheimer's disease risk: a meta-analysis. CNS Neurosci. Ther. 19, 469–476.
Lloyd-Jones, L.R., Holloway, A., McRae, A., Yang, J., Small, K., Zhao, J., Zeng, B., Bakshi, A., Metspalu, A., Dermitzakis, M., Gibson, G., Spector, T., Montgomery, G., Esko, T., 2017. The genetic architecture of gene expression in peripheral blood. Am. J. Hum. Genet. 100, 228–237.
Luedecking, E.K., DeKosky, S.T., Mehdi, H., Ganguli, M., Kamboh, M.I., 2000. Analysis of genetic polymorphisms in the transforming growth factor-beta1 gene and the risk of Alzheimer's disease. Hum. Genet. 106, 565–569.
Mahoney, E.R., Dumitrescu, L., Moore, A.M., Cambronero, F.E., De Jager, P.L., Koran, M.E.I., Petyuk, V.A., Robinson, R.A.S., Goyal, S., Schneider, J.A., Bennett, D.A., Jefferson, A.L., Hohman, T.J., 2021. Brain expression of the vascular endothelial growth factor gene family in cognitive aging and alzheimer's disease. Mol. Psychiatry 26, 888–896.
Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J.Z., Menon, M., He, L., Abdurrob, F., Jiang, X., Martorell, A.J., Ransohoff, R.M., Hafler, B.P., Bennett, D.A., Kellis, M., Tsai, L.H., 2019. Single-cell transcriptomic analysis of Alzheimer's disease. Nature 570, 332–337.
McClellan, J., King, M.C., 2010. Genetic heterogeneity in human disease. Cell 141, 210–217.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.
Miyashita, A., Koike, A., Jun, G., Wang, L.S., Takahashi, S., Matsubara, E., Kawarabayashi, T., Shoji, M., Tomita, N., Arai, H., Asada, T., Harigaya, Y., Ikeda, M., Amari, M., Hanyu, H., Higuchi, S., Ikeuchi, T., Nishizawa, M., Suga, M., Kawase, Y., Akatsu, H., Kosaka, K., Yamamoto, T., Imagawa, M., Hamaguchi, T., Yamada, M., Morihara, T., Takeda, M., Takao, T., Nakata, K., Fujisawa, Y., Sasaki, K., Watanabe, K., Nakashima, K., Urakami, K., Ooya, T., Takahashi, M., Yuzuriha, T., Serikawa, K., Yoshimoto, S., Nakagawa, R., Kim, J.W., Ki, C.S., Won, H.H., Na, D.L., Seo, S.W., Mook-Jung, I., Alzheimer Disease Genetics Consortium, St George-Hyslop, P., Mayeux, R., Haines, J.L., Pericak-Vance, M.A., Yoshida, M., Nishida, N., Tokunaga, K., Yamamoto, K., Tsuji, S., Kanazawa, I., Ihara, Y., Schellenberg, G.D., Farrer, L.A., Kuwano, R., 2013. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. PLoS One 8, e58618.
Mocali, A., Cedrola, S., Della Malva, N., Bontempelli, M., Mitidieri, V.A., Bavazzano, A., Comolli, R., Paoletti, F., La Porta, C.A., 2004. Increased plasma levels of soluble CD40, together with the decrease of TGF beta 1, as possible differential markers of Alzheimer disease. Exp. Gerontol. 39, 1555–1561.
Mu, Y., Gage, F.H., 2011. Adult hippocampal neurogenesis and its role in Alzheimer's disease. Mol Neurodegener. 6, 85.
Muyas, F., Bosio, M., Puig, A., Susak, H., Domenech, L., Escaramis, G., Zapata, L., Demidov, G., Estivill, X., Rabionet, R., Ossowski, S., 2019. Allele balance bias identifies systematic genotyping errors and false disease associations. Hum. Mutat. 40, 115–126.
Ng, B., White, C.C., Klein, H.U., Sieberts, S.K., McCabe, C., Patrick, E., Xu, J., Yu, L., Gaiteri, C., Bennett, D.A., Mostafavi, S., De Jager, P.L., 2017. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat. Neurosci. 20, 1418–1426.
Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 31, 3812–3814.
Patel, N.S., Mathura, V.S., Bachmeier, C., Beaulieu-Abdelahad, D., Laporte, V., Weeks, O., Mullan, M., Paris, D., 2010. Alzheimer's beta-amyloid peptide blocks vascular endothelial growth factor mediated signaling via direct interaction with VEGFR-2. J. Neurochem. 112, 66–76.
Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., Willer, C.J., 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26, 2336–2337.
Qi, T., Wu, Y., Zeng, J., Zhang, F., Xue, A., Jiang, L., Zhu, Z., Kemper, K., Yengo, L., Zheng, Z., eQTLGen Consortium, Marioni, R.E., Montgomery, G.W., Deary, I.J., Wray, N.R., Visscher, P.M., McRae, A.F., Yang, J., 2018. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat. Commun. 9, 2282.
Querfurth, H.W., LaFerla, F.M., 2010. Alzheimer's disease. N. Engl. J. Med. 362, 329–344.
Raghavan, N.S., Brickman, A.M., Andrews, H., Manly, J.J., Schupf, N., Lantigua, R., Wolock, C.J., Kamalakaran, S., Petrovski, S., Tosto, G., Vardarajan, B.N., Goldstein, D.B., Mayeux, R., Alzheimer's Disease Sequencing Project, 2018. Whole-exome sequencing in 20,197 persons for rare variants in Alzheimer's disease. Ann. Clin. Transl. Neurol. 5, 832–842.
Ray, S., Britschgi, M., Herbert, C., Takeda-Uchimura, Y., Boxer, A., Blennow, K., Friedman, L.F., Galasko, D.R., Jutel, M., Karydas, A., Kaye, J.A., Leszek, J., Miller, B.L., Minthon, L., Quinn, J.F., Rabinovici, G.D., Robinson, W.H., Sabbagh, M.N., So, Y.T., Sparks, D.L., Tabaton, M., Tinklenberg, J., Yesavage, J.A., Tibshirani, R., Wyss-Coray, T, 2007. Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. Nat. Med. 13, 1359–1362.
Rivera, E.J., Goldin, A., Fulmer, N., Tavares, R., Wands, J.R., de la Monte, S.M., 2005.

Insulin and insulin-like growth factor expression and function deteriorate with progression of Alzheimer's disease: link to brain reductions in acetylcholine. J. Alzheimers Dis. 8, 247–268.
Rodrigues, M., Griffith, L.G., Wells, A., 2010. Growth factor regulation of proliferation and survival of multipotential stromal cells. Stem Cell Res. Ther. 1, 32.
Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chetelat, G., Teunissen, C.E., Cummings, J., van der Flier, W.M., 2021. Alzheimer's disease. Lancet 397, 1577–1590.
Schindowski, K., Belarbi, K., Buee, L., 2008. Neurotrophic factors in Alzheimer's disease: role of axonal transport. Genes Brain Behav. 7 (Suppl 1), 43–56.
Schmeisser, M.J., Baumann, B., Johannsen, S., Vindedal, G.F., Jensen, V., Hvalby, O.C., Sprengel, R., Seither, J., Maqbool, A., Magnutzki, A., Lattke, M., Oswald, F., Boeckers, T.M., Wirth, T., 2012. IkappaB kinase/nuclear factor kappaB-dependent insulin-like growth factor 2 (Igf2) expression regulates synapse formation and spine maturation via Igf2 receptor signaling. J. Neurosci. 32, 5688–5703.
Seeger, M.A., Paller, A.S., 2015. The roles of growth factors in keratinocyte migration. Adv. Wound. Care. (New Rochelle). 4, 213–224.
Shigemizu, D., Mitsumori, R., Akiyama, S., Miyashita, A., Morizono, T., Higaki, S., Asanomi, Y., Hara, N., Tamiya, G., Kinoshita, K., Ikeuchi, T., Niida, S., Ozaki, K., 2021. Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk. Transl. Psychiatry 11, 151.
Thomas, R., Zuchowska, P., Morris, A.W., Marottoli, F.M., Sunny, S., Deaton, R., Gann, P.H., Tai, L.M., 2016. Epidermal growth factor prevents APOE4 and amyloid-beta-induced cognitive and cerebrovascular deficits in female mice. Acta Neuropathol. Commun. 4, 111.
Turner, C.A., Eren-Kocak, E., Inui, E.G., Watson, S.J., Akil, H., 2016. Dysregulated fibroblast growth factor (FGF) signaling in neurological and psychiatric disorders. Semin. Cell Dev. Biol. 53, 136–143.
Viechtbauer, W., 2010. Conducting meta-Analyses in R with the metafor package. J. Stat. Softw. 36, 1–48.
Wang, B., Bao, S., Zhang, Z., Zhou, X., Wang, J., Fan, Y., Zhang, Y., Li, Y., Chen, L., Jia, Y., Li, J., Li, M., Zheng, W., Mu, N., Wang, L., Yu, Z., Wong, D.S.M., Zhang, Y., Kwan, J., Ka-Fung Mak, H., Ambalavanan, A., Zhou, S., Cai, W., Zheng, J., Huang, S., Rouleau, G.A., Yang, W., Rogaeva, E., Ma, X., St George-Hyslop, P., Chu, L.W., Song, Y.Q., 2018a. A rare variant in MLKL confers susceptibility to ApoE varepsilon4-negative Alzheimer's disease in Hong Kong Chinese population. Neurobiol. Aging 68, 160 e161-160 e167.
Wang, H.Z., Bi, R., Hu, Q.X., Xiang, Q., Zhang, C., Zhang, D.F., Zhang, W., Ma, X., Guo, W., Deng, W., Zhao, L., Ni, P., Li, M., Fang, Y., Li, T., Yao, Y.G., 2016. Validating GWAS-identified risk loci for Alzheimer's disease in Han Chinese populations. Mol. Neurobiol. 53, 379–390.
Wang, D., Fan, Y., Malhi, M., Bi, R., Wu, Y., Xu, M., Yu, X.F., Long, H., Li, Y.Y., Zhang, D.F., Yao, Y.G., 2018b. Missense variants in HIF1A and LACC1 contribute to leprosy risk in Han Chinese. Am. J. Hum. Genet. 102, 794–805.
Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164.
Wang, W., Yu, J.T., Tan, L., Liu, Q.Y., Wang, H.F., Ma, X.Y., 2012. Insulin-like growth factor 1 (IGF1) polymorphism is associated with Alzheimer's disease in Han Chinese. Neurosci. Lett. 531, 20–23.
Woodbury, M.E., Ikezu, T., 2014. Fibroblast growth factor-2 signaling in neurogenesis and neurodegeneration. J. Neuroimmune Pharmacol. 9, 92–101.
Xu, M., Zhang, D.F., Luo, R., Wu, Y., Zhou, H., Kong, L.L., Bi, R., Yao, Y.G., 2018. A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. Alzheimers Dement. 14, 215–229.
Yang, Q., Wang, E.Y., Jia, H.W., Wang, Y.P., 2016. Association between polymorphisms in transforming growth factor-beta1 and sporadic Alzheimer's disease in a Chinese population. Int. J. Neurosci. 126, 979–984.
Zhang, D.F., Fan, Y., Xu, M., Wang, G., Wang, D., Li, J., Kong, L.L., Zhou, H., Luo, R., Bi, R., Wu, Y., Li, G.D., Alzheimer's Disease Neuroimaging Initiative (ADNI), Li, M., Luo, X.J., Jiang, H.Y., Tan, L., Zhong, C., Fang, Y., Zhang, C., Sheng, N., Jiang, T., Yao, Y.G., 2019. Complement C7 is a novel risk gene for Alzheimer's disease in Han Chinese. Natl. Sci. Rev. 6, 257–274.
Zhang, D.F., Li, J., Wu, H., Cui, Y., Bi, R., Zhou, H.J., Wang, H.Z., Zhang, C., Wang, D., Alzheimer's Disease Neuroimaging Initiative (ADNI), Kong, Q.P., Li, T., Fang, Y., Jiang, T., Yao, Y.G., 2016. CFH variants affect structural and functional brain changes and genetic risk of Alzheimer's disease. Neuropsychopharmacology 41, 1034–1045.
Zhou, X., Chen, Y., Mok, K.Y., Zhao, Q., Chen, K., Chen, Y., Hardy, J., Li, Y., Fu, A.K.Y., Guo, Q., Ip, N.Y., Alzheimer's Disease Neuroimaging Initiative, 2018. Identification of genetic risk factors in the Chinese population implicates a role of immune system in Alzheimer's disease pathogenesis. Proc. Natl. Acad. Sci. U. S. A. 115, 1697–1706.

Online Supplementary Materials

This file contains 9 supplementary tables and 3 supplementary figures.

Supplementary Table S1

Sequence coverage of each gene in the targeted sequencing

Gene	Mean depth	10× coverage (%)	20× coverage (%)	30× coverage (%)	50× coverage (%)
TGFB2	147.54	100.00	100.00	95.86	90.31
TGFBR1	160.30	88.32	88.32	88.32	88.32
TGFBR2	120.41	100.00	96.34	90.93	85.07
LTBP1	145.93	98.26	93.01	92.16	89.73
EGF	165.36	100.00	100.00	100.00	99.98
EGFR	119.27	99.21	95.93	92.19	89.27
HBEGF	120.38	100.00	93.17	86.03	70.48
INHBC	125.12	100.00	100.00	100.00	100.00
IGF1	159.22	100.00	100.00	100.00	95.50
IGF1R	152.05	100.00	99.93	99.33	96.52
IGF2	105.67	100.00	92.70	85.89	73.05
IGF2R	141.68	96.63	96.13	96.13	93.97
PDGFRA	168.49	100.00	100.00	100.00	99.81
PDGFRB	107.74	100.00	100.00	99.74	88.89
PDGFC	187.67	100.00	100.00	100.00	100.00
FGFR2	158.01	97.41	96.02	95.17	92.72
VEGFA	101.54	100.00	100.00	96.23	84.38
VEGFC	151.42	99.82	99.82	87.54	81.72
KDR	159.64	100.00	100.00	100.00	94.75
ANGPTL1	245.95	100.00	100.00	100.00	100.00
HGF	162.23	100.00	100.00	100.00	100.00
BDNF	234.00	100.00	100.00	100.00	99.17
GDNF	173.11	100.00	98.74	95.41	77.50

Mean depth of a gene was calculated by an equation = (total sequence data / gene length).
N× coverage was calculated by an equation = ((the total number of nucleobases in a gene that were sequenced over N times / total gene length) × 100%).

Supplementary Table S2
Rare variants of 23 growth factor genes in Han Chinese

This table is too big and is presented as an Excel file.

Rare variants were defined by an MAF < 0.01 in the control sample. All variants listed in this table had a sequencing coverage ≥ 20×.

Supplementary Table S3
Association of rare variants of *EGF* with AD in Han Chinese

This table is too big and is presented as an Excel file.

Supplementary Table S4
Allele frequencies of *EGF* rare variants in East Asian populations from the gnomAD dataset (https://gnomad.broadinstitute.org/) (Karczewski et al., 2020)

This table is too big and is presented as an Excel file.

Supplementary Table S5
Association of common variants in 23 growth factor genes with AD in Han Chinese

This table is too big and is presented as an Excel file.

Supplementary Table S6

Association of *EGF* variants with AD in European populations

Variant ID[a]	*P*-value	Chr:position[b]	Allele[c]	Frequency[d]	OR[e]	95% CI[f]	Beta	SE[g]
rs4444903	0.869	4:109912954	A/G	0.5961	0.999	0.983-1.015	-0.0014	0.0082
rs11568849	0.421	4:109913381	A/C	0.998	0.88	0.645-1.201	-0.1277	0.1588
rs11568886	0.236	4:109941207	T/C	0.0042	1.145	0.915-1.434	0.1358	0.1146
rs10470911	0.766	4:109944115	T/G	0.662	1.003	0.986-1.02	0.0025	0.0086
rs4698755	0.810	4:109945022	A/C	0.6586	1.002	0.986-1.019	0.0021	0.0085
rs4698756	0.799	4:109945286	A/G	0.3414	0.998	0.981-1.015	-0.0022	0.0085
rs4698800	0.764	4:109945352	T/C	0.3391	0.997	0.981-1.014	-0.0026	0.0085
rs11568927	0.088	4:109959477	T/C	0.0036	1.297	0.962-1.748	0.2598	0.1523
rs11568941	0.104	4:109961824	T/G	0.0036	1.282	0.95-1.73	0.2487	0.1529
rs11568942	0.122	4:109961840	T/C	0.9955	1.129	0.968-1.316	0.1212	0.0784
rs11568943	0.440	4:109961965	A/G	0.0632	1.013	0.98-1.047	0.0129	0.0167
rs3733628	0.081	4:109963140	T/C	0.0036	1.305	0.967-1.76	0.266	0.1526
rs11568990	0.279	4:109974691	A/C	0.0041	1.163	0.885-1.528	0.1509	0.1393
rs2302135	0.111	4:109979991	A/G	0.9955	0.849	0.695-1.038	-0.1634	0.1026
rs2237051	0.338	4:109980042	A/G	0.3825	0.992	0.976-1.008	-0.008	0.0083
rs11569017	0.656	4:109980955	A/T	0.9457	1.008	0.973-1.044	0.008	0.0179
rs11569018[h]	0.249	4:109981010	A/G	0.0018	1.212	0.874-1.681	0.1924	0.1669
rs6836684[h]	0.850	4:109987933	T/C	0.0024	0.979	0.788-1.217	-0.021	0.1108
rs4698803[h]	0.610	4:109993271	A/T	0.2072	1.005	0.986-1.025	0.0052	0.0101
rs75935899[h]	0.572	4:109999791	A/G	0.002	0.925	0.705-1.213	-0.0781	0.1383
rs11568937[h]	0.564	4:109960895	T/C	0.966	0.987	0.944-1.032	-0.0131	0.0227
rs11568953[h]	0.803	4:109963240	A/G	0.9839	1.008	0.944-1.077	0.0083	0.0334
rs11568993[h]	0.004	4:109976159	T/C	0.0837	0.959	0.932-0.987	-0.042	0.0147

[a]Summary statistics from the meta-analysis (Bellenguez et al., 2022) are available through the National Human Genome Research Institute-European Bioinformatics Institute GWAS catalog under accession number GCST90027158 (https://www.ebi.ac.uk/gwas/).

[b]Chr:position, position were shown in GRCh38

[c]Effect allele / other allele

[d]Frequency of effect allele

[e]OR, odds ratio of the effect allele

[f]95% confidence interval of OR

[g]Standard error

[h]Rare variants (MAF < 0.01) in Han Chinese under study

Supplementary Table S7

eQTL effect of rs4698800 and rs10470911 on *EGF* expression

SNP	eQTL dataset	A1/A2[a]	Freq[b]	Beta	SE[c]	*P*-value
rs4698800	Brain_eMeta (Qi et al., 2018)	T/C	0.346	0.220	0.072	0.0022
rs4698800	cage_whole_blood (Lloyd-Jones et al., 2017)	T/C	0.360	0.070	0.028	0.0118
rs10470911	Brain_eMeta (Qi et al., 2018)	G/T	0.336	0.235	0.072	0.0012
rs10470911	cage_whole_blood (Lloyd-Jones et al., 2017)	G/T	0.357	0.063	0.028	0.0238

eQTL datasets include peripheral blood eQTL data from the Consortium for the Architecture of Gene Expression (CAGE) (Lloyd-Jones et al., 2017) and brain eQTL data from a meta-analysis (Qi et al., 2018) of GTEx brain (GTEx Consortium et al., 2017), CommonMind Consortium (CMC) (Fromer et al., 2016) and ROSMAP (Ng et al., 2017)

[a]Effect allele/other allele

[b]Frequency of the effect allele in the respective eQTL study

[c]SE, standard error

Supplementary Table S8
mRNA expression of 23 growth factors in six types of cells based on scRNA-seq data of human brain tissues (no-pathology vs. pathology)

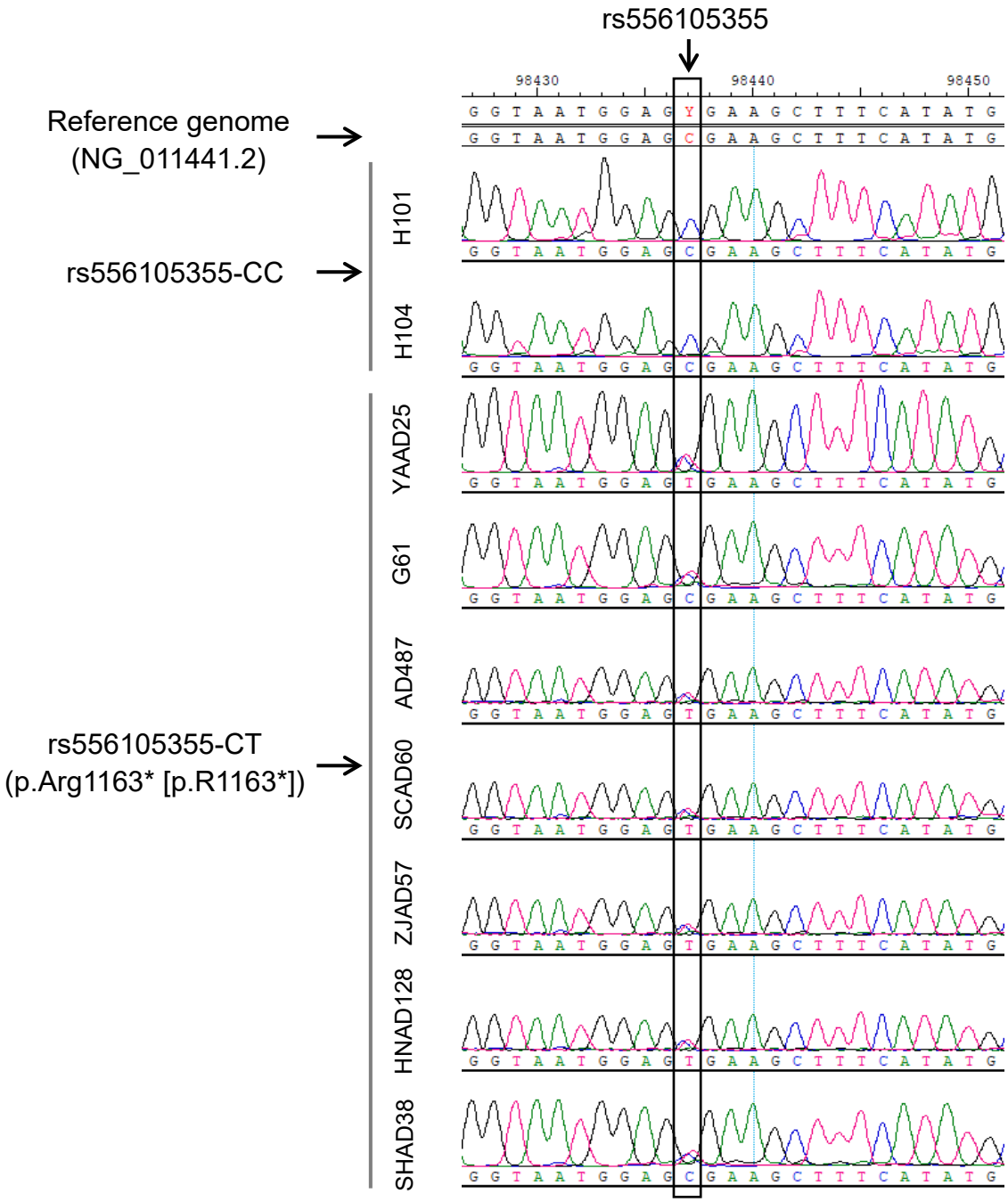
This table is too big and is presented as an Excel file.

Data in this table were retrieved from Mathys et al. (2019). We quoted notes in the Excel file from the original paper for the convenience of the reader.

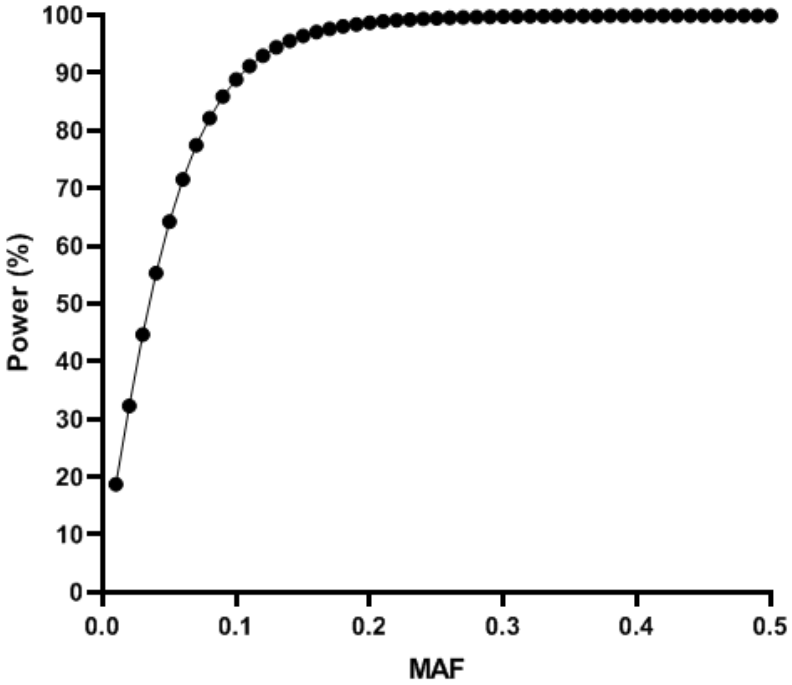
Supplementary Table S9
mRNA expression of 23 growth factors in six types of cells based on scRNA-seq data of human brain tissues (early-pathology vs. late-pathology)

This table is too big and is presented as an Excel file.

Data in this table were retrieved from Mathys et al. (2019). We quoted notes in the Excel file from the original paper for the convenience of the reader.

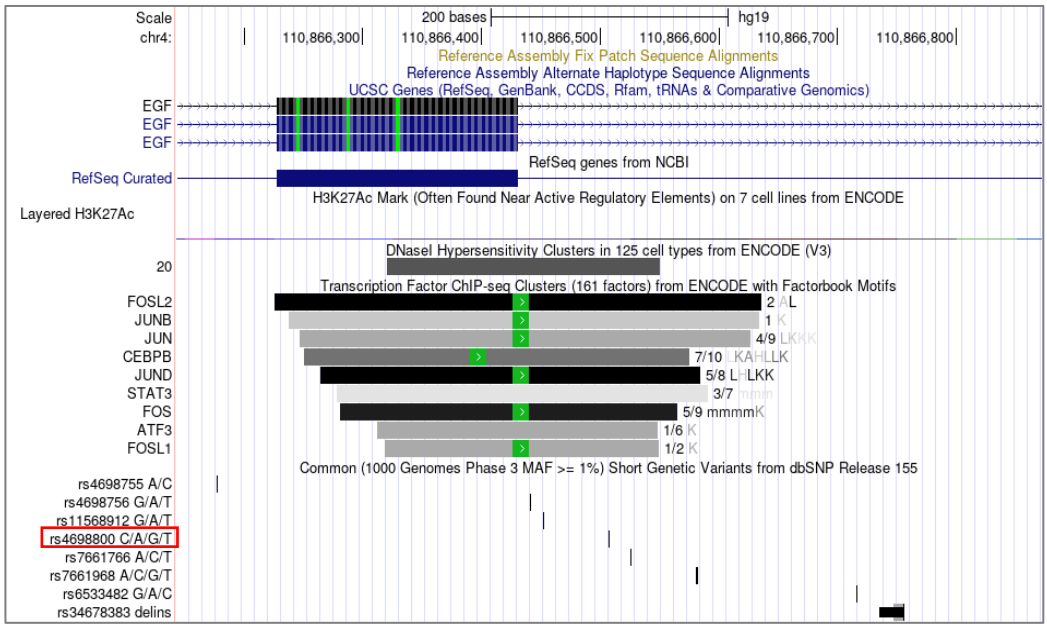


Supplementary Fig. S1. Validation of rare variant rs556105355 C>T in T allele carriers (n = 7) and non-carriers (n = 2) by Sanger sequencing. The ID of each sample was shown in the left of the corresponding sequences. We used a primer pair rs556105355-F (5'-GGCTGAGGTGGAAGGATCAC-3') / rs556105355-R (5'-CTCCATTTGGTGTGGTGGGT-3') to amplify and sequence a 506 bp fragment harboring rs556105355. *, stop codon.

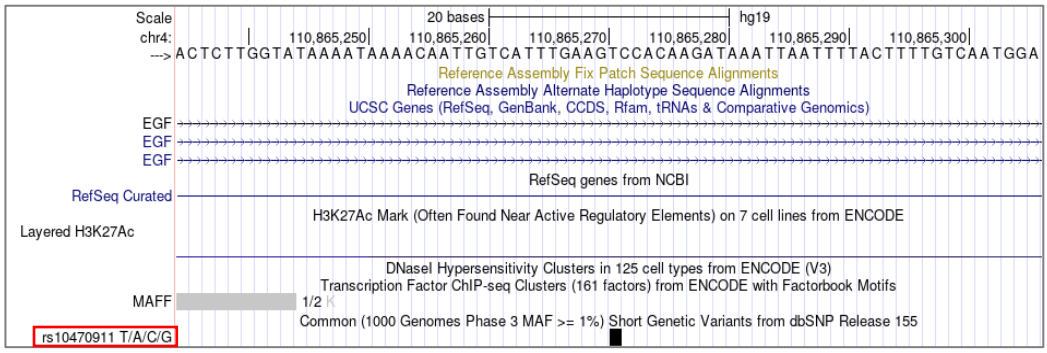


Supplementary Fig. S2. Power estimate for the case-control association analysis. Statistical power was computed under the gene only hypothesis and log additive model, with the following parameters: risk allele ranges from 0.01 to 0.5 in increments of 0.01; overall disease risk in the general population = 0.03; sample size = 1280 cases vs. 5044 controls; OR = 1.25.

a



b



Supplementary Fig. S3. Regulatory functional annotation of rs4698800 (a) and rs10470911 (b). Functional genomic annotations for enhancers (H3K27ac), chromatin accessibility (DNaseI hypersensitivity sites), and transcription factor binding sites (TFBSs) of each target variant were based on the ENCODE data retrieved from the UCSC Genome Browser (https://genome.ucsc.edu/). The target variants were marked with a red box.

References

Bellenguez C., Küçükali F., Jansen I.E., Kleineidam L., Moreno-Grau S., Amin N., Naj A.C., Campos-Martin R., Grenier-Boley B., Andrade V., Holmans P.A., et al., 2022. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat. Genet. 54, 412-436.

Fromer M., Roussos P., Sieberts S.K., Johnson J.S., Kavanagh D.H., Perumal T.M., Ruderfer D.M., Oh E.C., Topol A., Shah H.R., Klei L.L., et al., 2016. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. 19, 1442-1453.

GTEx Consortium, Laboratory Data Analysis &Coordinating Center (LDACC)-Analysis Working Group, Statistical Methods groups-Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site-NDRI, Biospecimen Collection Source Site-RPCI, et al., 2017. Genetic effects on gene expression across human tissues. Nature 550, 204-213.

Karczewski K.J., Francioli L.C., Tiao G., Cummings B.B., Alföldi J., Wang Q., Collins R.L., Laricchia K.M., Ganna A., Birnbaum D.P., Gauthier L.D., et al., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434-443.

Lloyd-Jones L.R., Holloway A., McRae A., Yang J., Small K., Zhao J., Zeng B., Bakshi A., Metspalu A., Dermitzakis M., Gibson G., Spector T., Montgomery G., Esko T., Visscher P.M.Powell J.E., 2017. The genetic architecture of gene expression in peripheral blood. Am. J. Hum. Genet. 100, 228-237.

Mathys H., Davila-Velderrain J., Peng Z., Gao F., Mohammadi S., Young J.Z., Menon M., He L., Abdurrob F., Jiang X., Martorell A.J., Ransohoff R.M., Hafler B.P., Bennett D.A., Kellis M.Tsai L.H., 2019. Single-cell transcriptomic analysis of Alzheimer's disease. Nature 570, 332-337.

Ng B., White C.C., Klein H.U., Sieberts S.K., McCabe C., Patrick E., Xu J., Yu L., Gaiteri C., Bennett D.A., Mostafavi S.De Jager P.L., 2017. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat. Neurosci. 20, 1418-1426.

Qi T., Wu Y., Zeng J., Zhang F., Xue A., Jiang L., Zhu Z., Kemper K., Yengo L., Zheng Z., et al., 2018. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat. Commun. 9, 2282.