Journal of Psychiatric Research 113 (2019) 141–147

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/jpsychires

Mutation and association analyses of dementia-causal genes in Han Chinese patients with early-onset and familial Alzheimer's disease

Guihong Wang[a,1], Deng-Feng Zhang[b,c,1], Hong-Yan Jiang[d,1], Yu Fan[b,c], Lingyan Ma[a], Zonglin Shen[d], Rui Bi[b,c], Min Xu[b,e], Liwen Tan[f], Baoci Shan[g,h,i,***], Yong-Gang Yao[b,e,i,j,*], Tao Feng[a,k,l,**]

[a] Center for Neurodegenerative Diseases, Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, 100050, China
[b] Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, China
[c] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, 650223, China
[d] Department of Psychiatry, The First Affiliated Hospital of Kunming Medical University, Kunming, 650032, China
[e] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, 650204, China
[f] Mental Health Institute of the Second Xiangya Hospital, Central South University, Changsha, 410011, China
[g] Division of Nuclear Technology and Applications, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, 100049, China
[h] Beijing Engineering Research Center of Radiographic Techniques and Equipment, Beijing, 100049, China
[i] CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, 200031, China
[j] KIZ – CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China
[k] China National Clinical Research Center for Neurological Diseases, Beijing, 100050, China
[l] Parkinson's Disease Center, Beijing Institute for Brain Disorders, Capital Medical University, Beijing, 100050, China

ARTICLE INFO

Keywords:
Alzheimer's disease
Dementia
Mutation
Association
Next-generation sequencing

ABSTRACT

Alzheimer's disease (AD) is the most common cause of dementia in the elderly. It shares clinical and pathological features with other types of dementia, such as vascular dementia (VaD), Lewy body dementia (LBD), and frontotemporal dementia (FTD). We have hypothesized that there might be an overlapping molecular mechanism and genetic basis to the different types of dementia. In this study, we analyzed the mutation pattern of dementia-causal genes in 169 Han Chinese patients with familial and early-onset AD by using whole exome sequencing or targeted resequencing. We identified 9 potentially pathogenic mutations in the AD-causal genes *APP*, *PSEN1*, *PSEN2*, and 6 mutations in a group of non-AD dementia-causal genes including the FTD-causal gene *GRN* and the VaD-causal gene *NOTCH3*. A common splice-site variant rs514492 in the FTD-causal gene *VCP* showed a positive association with AD risk ($P = 0.0003$, OR = 1.618), whereas the rare missense variant rs33949390 (p. R 1628P) in the LBD-causal gene *LRRK2* showed a protective effect on AD risk ($P = 0.0004$, OR = 0.170). The presence of putative pathogenic mutations and risk variants in these causal genes for different types of dementia in clinically diagnosed familial and early-onset AD patients suggests a need to screen for mutations of the dementia-causal genes in cases of AD to avoid misdiagnosis. These mutations also support the idea that there are overlapping pathomechanisms between AD and other forms of dementia.

1. Introduction

'Dementia' describes a group of diseases which all present with symptoms affecting cognitive, behavioral and social abilities (Leyhe et al., 2017; Love, 2005; Querfurth and LaFerla, 2010). Alzheimer's disease (AD) is the most common cause of dementia in the elderly,

* Corresponding author. Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, China.
** Corresponding author. Center for Neurodegenerative Diseases, Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, 100050, China.
*** Corresponding author. Division of Nuclear Technology and Applications, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, 100049, China.
E-mail addresses: shanbc@mail.ihep.ac.cn (B. Shan), yaoyg@mail.kiz.ac.cn (Y.-G. Yao), happyft@sina.com (T. Feng).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jpsychires.2019.03.026
Received 28 January 2019; Received in revised form 22 March 2019; Accepted 27 March 2019

accounting for 50%–70% of all cases (Alzheimer's association, 2016; Blennow et al., 2006; Jack et al., 2018; Leyhe et al., 2017; Querfurth and LaFerla, 2010). The main molecular features of AD are deposits of β-amyloid (Aβ) plaques and aggregated neurofibrillary tangles of phosphorylated tau protein (Coded by the *MAPT* gene) in the brain (Jack et al., 2018; Love, 2005; Waldemar et al., 2007). Vascular dementia (VaD) is the second most common type of dementia (up to ~20% of cases) and is the result of damage to the blood vessels (Waldemar et al., 2007; Wetterling et al., 1996). Lewy body dementia (LBD, up to ~10% dementia cases) is the third common cause of dementia after AD and VaD, and is characterized by abnormal deposits of α-synuclein in the Lewy bodies (McKeith et al., 2005). Frontotemporal dementia (FTD) is a group of diseases characterized by degeneration of the frontal and temporal lobes (Young et al., 2018). It is possible for a patient to have two types of dementia at the same time. This is known as having a mixed dementia and is usually a combination of AD and another type of dementia (Tofaris and Buckley, 2018). Some of the clinical and pathological features are shared by different types of dementia, leading to a challenge in making a clinical diagnosis of dementia (Tofaris and Buckley, 2018). What is more, considering the similar clinical and pathological phenotypes, different types of dementia might share an overlapping molecular mechanism and a common genetic basis (Tofaris and Buckley, 2018), but presenting different symptoms due to different genetic backgrounds and environmental influences.

Pathogenic mutations in dementia-causal genes have been well characterized, providing the possibility of genetic testing for the diagnosis of dementia and basic research. Three well-known genes, β-amyloid precursor protein (*APP*), Presenilin-1 (*PSEN1*) and Presenilin-2 (*PSEN2*) that involving in the production of Aβ, have been recognized to be the causal genes for early-onset familial AD (EOFAD, age at onset [AAO] < 65 years old [yr], always with a positive familial history) that is inherited in a Mendelian fashion (Campion et al., 1999; Guerreiro et al., 2013). However, the families with autosomal dominant familial AD caused by pathogenic *APP*, *PSEN1*, and *PSEN2* mutations only account for a very small proportion of dementia cases (Campion et al., 1999; Ridge et al., 2013, 2016). Whether causal genes found in other types of dementia, such as *MAPT*, *GRN*, *VCP*, *TREM2*, *SQSTM1*, *FUS*, *TARDBP*, *CHMP2B*, and *C9ORF72* for FTD (Guerreiro et al., 2015), *LRRK2*, *SNCA*, and *PINK1* for LBD (Meeus et al., 2012; Vergouw et al., 2017; Zhu et al., 2006), *NOTCH3*, *HTRA1*, and *COL4A1* for VaD (Ikram et al., 2017), and *PRNP* and *CSF1R* for other neurodegenerative diseases with dementia symptoms (Lynch et al., 2017; Sassi et al., 2018; Vergouw et al., 2017), also contribute to AD have remained to be investigated. In this study, we analyzed the mutation spectrum of these known dementia-causal genes by using the next-generation sequencing data obtained from familial and early-onset AD patients.

2. Materials and methods

2.1. Subjects

Patients with early onset AD (AAO < 65 yr) and/or a positive familial history were enrolled from Southwest and East China (N = 107, 46.7% females, age 64.6 ± 10.3 yr, AAO 56.7 ± 9.5 yr, *APOE* ε4 38.5%; South East cohort) and North China (N = 62, 54.8% females, age 58.1 ± 8.6 yr, AAO 54.7 ± 7.6 yr, *APOE* ε4 43.5%; North cohort). The majority of these patients had been described in our recent study (Zhang et al., 2018). Briefly, familial cases were defined as having at least one affected first- or second-degree relative besides the proband. For each family, we only sequenced the proband. We did not collect samples for the other patients in these families. In the South East cohort, 82 patients (including 12 familial cases) had an AAO < 65 yr and 25 individuals were late-onset familial cases (> 65 yr). In the North cohort, there were 55 patients (including 19 familial cases) with an AAO < 65 yr and the remaining cases were late-onset familial AD. All patients were diagnosed as possible or probable AD by at least two clinical psychiatrists according to the revised National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria (Jack et al., 2011; Khachaturian, 2011; McKhann et al., 1984) and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, as described in our previous studies (Bi et al., 2015; Xiang et al., 2017; Zhang et al., 2016). Patients in the North cohort also had diagnostic imaging, but this kind of test was not available for the South East cohort. Individuals clinically diagnosed as having VaD, FTD, or LBD were excluded. Sample collection complied with the Declaration of Helsinki, with written informed consents being obtained from each participant or their guardians. Exome data of 160 in-house non-dementia individuals (52 healthy individuals and 108 leprosy patients without dementia, 40.6% females, age 52.6 ± 16.5 yr; *APOE* ε4, 15% (Wang et al., 2018);) was combined with the whole genome data of Han Chinese in Beijing (CHB, N = 103) and Southern Han Chinese (CHS, N = 105) from the 1000 Genome Project phase 3 (1000 Genomes Project Consortium et al., 2015) as the control samples (N = 368) (Zhang et al., 2018). We performed principle component analysis to correct for potential population stratification using the same procedure in our previous study (Zhang et al., 2018), and found no apparent population stratification among the AD patients and the control samples (Fig. S1). Allele frequency data of 4327 East Asians from the Exome Aggregation Consortium (ExAC, accessed at 2016) (Lek et al., 2016) were also retrieved and used as a reference control. This study was approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences.

2.2. Next-generation sequencing

Genetic testing of all the patients was conducted by using next-generation sequencing technologies. The South East cohort was initially sequenced by Zhang et al. (2018) using the Nimblegen SeqCap EZ Exome Enrichment Kit v3.0 (Roche, Basel, Switzerland). The North cohort was sequenced using the IDT XGen Exome kit or customized targeted gene sequencing panel (performed by a commercial service from the PrecisionMD Company). In brief, libraries were constructed according to the manufacturer's instructions and sequenced on the Illumina HiSeq 2500 or 4000 (Illumina, San Diego, CA, USA) platform using the 150 bp paired-end module. All the sequenced reads were processed through the canonical pipeline recommended by the GATK Best Practices (Genome Analysis Toolkit, https://www.broadinstitute.org/gatk/guide/best-practices) (McKenna et al., 2010). Reads were aligned to the human genome reference assembly (build GRCh37/hg19: https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/) using the Burrows-Wheeler Aligner (Li and Durbin, 2009). Picard Tools (http://broadinstitute.github.io/picard/) were used to mark and remove duplicate reads. Variants were called by the newest GATK pipeline with a Phred-quality score > Q10. All called variants were subjected to the GATK Variant Quality Score Recalibration to filter spurious variants due to sequencing errors and mapping artifacts. As the newly identified singletons in this study might be sequencing errors, we manually checked the sequencing and mapping quality of raw reads. We also performed Sanger sequencing (Table S1) to confirm the potentially damaging variants in the dementia-causal genes that were only observed in single patients. ANNOVAR was used to annotate variants into different functional categories according to their genic location and expected effect on encoded gene products (Wang et al., 2010).

2.3. Mutation and association analyses of the dementia-causal genes

We analyzed 20 causal genes showing a Mendelian inheritance that were reported for dementia, which were captured by different sequencing strategies used in this study. The gene list contained the well-known genes *APP*, *PSEN1*, and *PSEN2* for AD (Guerreiro et al., 2013),

MAPT, *GRN*, *VCP*, *TREM2*, *SQSTM1*, *FUS*, *TARDBP*, *CHMP2B*, and *C9ORF72* for FTD (Guerreiro et al., 2015), *LRRK2*, *SNCA*, and *PINK1* for LBD (Meeus et al., 2012; Vergouw et al., 2017), *NOTCH3*, *HTRA1*, and *COL4A1* for VaD (Ikram et al., 2017), *PRNP* and *CSF1R* for other neurodegenerative diseases with dementia symptoms such as leukodystrophy (Lynch et al., 2017; Sassi et al., 2018). Nonsense, frameshift, splice-site variants, and missense variants that affect the coding region were defined as functional variants and were analyzed in this study, and the other variants of unknown significance were not discussed.

Functional variants of the above 20 dementia-causal genes in 169 AD patients, 368 controls (Zhang et al., 2018), and 4327 East Asians from the ExAC dataset (Lek et al., 2016) were retrieved and annotated. Allele frequencies of all variants in AD patients were compared with those of the combined controls ((Zhang et al., 2018) and references therein) and the population samples from the ExAC (Lek et al., 2016) by using the Fisher's exact test. A *P-value* < 0.05 was regarded as marginally significant when the AD patients were compared with the combined control or the ExAC population control. We also performed linear regression analysis comparing the cases with the combined controls, with the first three principle components as the covariates to correct for potential population stratification (Zhang et al., 2018). Note that we had no detailed information for the subjects in the ExAC dataset (Lek et al., 2016), especially for age, sex, and neurological disease assessment. There was a possibility that the ExAC samples might contain potential dementia patients, which would lead to a reduced statistical power. Therefore, the results using the ExAC data as the reference in our comparison should be interpreted with caution.

As most of the pathogenic mutations affect protein function, missense variants predicted to be damaging or deleterious by at least two of five protein-function-based algorithms (PolyPhen2 HunDiv and HunVar (Adzhubei et al., 2010), LRT (Chun and Fay, 2009), MutationTaster (Schwarz et al., 2010), and SIFT (Ng and Henikoff, 2003)) were regarded as potentially damaging mutations. The PHRED-scaled Combined Annotation-Dependent Depletion (CADD) score (Kircher et al., 2014), a method integrating diverse annotations, was also used to evaluate function potential of target variants.

We focused on two types of functional variants. The damaging mutations observed only in AD patients, but not in controls, available databases and reported studies according to a web-based search similar to that used for mtDNA variants (Bandelt et al., 2009), were defined as potentially pathogenic mutations. These case-only mutations were also classified as possibly pathogenic, probably pathogenic, and definitely pathogenic according to the classification by Guerreiro et al. (2010). Variants showing significant associations with AD by comparing AD patients with controls were regarded as functionally risk variants for AD.

Considering the limited sample size in this study, we calculated the statistic power for the association analysis using the Quanto software (version 1.2.4) (Gauderman, 2002). The power was 12.6% for allele with a minor allele frequency (MAF) of 0.01 and 58.1% for allele with a MAF of 0.5 in our samples (disease prevalence was set as 0.1) to capture an odds ratio of 2.0 under an additive model. It was thus underpowered for making a valid conclusion and validation in larger samples was essential to confirm these results. Therefore, we analyzed the AD-associated dementia-causal genes using two large datasets of populations with European ancestry. The first dataset is the International Genomics of Alzheimer's Project (IGAP) stage 1 GWAS data (Lambert et al., 2013), which contained 7,055,881 SNPs in 17,008 AD cases and 37,154 controls (Lambert et al., 2013). We downloaded the summary statistics from http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php and performed a gene-based test for the GWAS common variants using the online tool Versatile Gene-based Association Study (VEGAS2 v02, https://vegas2.qimrberghofer.edu.au/) (Liu et al., 2010; Mishra and Macgregor, 2015). The second dataset contains whole-exome sequencing data of 5815 AD cases and 4755 controls from the Alzheimer's Disease Sequencing Project (ADSP) (Bis et al., 2018). The original data were retrieved through the dbGaP (Genotypes and Phenotypes database: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000572.v7.p4) under the study accession phs000572.v7.p4, and were processed by plink/seq (https://atgu.mgh.harvard.edu/plinkseq/).

3. Results

3.1. Potentially pathogenic mutations of the dementia-causal genes in Chinese AD patients

A total of 69 functional variants (including missense, frame-error, nonsense, read-through, and splice-site) in 20 dementia-causal genes were found in 169 AD patients (Table S2, variants were listed separately for the two cohorts). All these variants had a high sequencing quality score. Most of these functional variants had a high PHRED-CADD score (Mean = 21.85 ± 7.35, range 1.04–39), suggesting they were potentially pathogenic. Among these 69 functional variants, 27 were observed only in single cases (12 patients in the South East cohort and 15 patients in the North cohort). 16 of the 27 case-only singletons were predicted to be damaging by at least two algorithms (Table S2). To exclude false positive variants calling of these 16 case-only damaging singletons, we conducted Sanger sequencing for these patients with the singletons for confirmation (Fig. S2). Among these 16 patients, two had *SQSTM1* singletons (p.E176A [South East cohort] and p.R309W [North cohort]) but were not sequenced, simply because the original DNA samples were used up for the two carriers; one patient in the South East cohort with a *PSEN2* variant (p.V88M) was confirmed to be false positive, and the remaining 13 patients with damaging singletons were confirmed to be true variants. Therefore, the validation rate (13/14) was considerably high for the singletons. We then focused on these 13 variants and the two *SQSTM1* variants (p.E176A and p.R309W) in the subsequent analyses (Table 1).

In the South East cohort, the damaging variant rs63750929 (p.G394V) of *PSEN1*, has been recorded previously in the AD mutation database (Alzheimer Disease & Frontotemporal Dementia Mutation Database: http://www.molgen.ua.ac.be/ADMutations/), and occurred in a patient with an AAO of 45 years old. The missense variant p.A379D and the splice-site variant chr1:227078976-C-T in *PSEN2* were separately found in two patients with AAO of 55 yr and 81 yr, respectively. The *PSEN2* variant p.A379D might be *possibly pathogenic* as it was observed only in a patient and this mutation has not been reported previously according to a search in the above AD mutation database. The *PSEN2* splice-site chr1:227078976-C-T seemed to be a novel variant, but its pathogenicity was unclear, as the carrier was positive for *APOE* ε4 and had a late onset age. Potentially pathogenic variants were also observed in two FTD-causal genes (*GRN* p.P21L [AAO = 49 yr] and *SQSTM1* p.E176A [AAO = 69 yr]), one LBD-causal gene (*LRRK2* p.P235A [AAO = 52 yr]), and two VaD-causal genes (*NOTCH3* p.E585A [AAO = 54 yr] and *HTRA1* p.R190H [AAO = 50 yr]) (Table 1). In the North cohort, there were 6 potentially pathogenic variants in the AD-causal genes: *APP* (p.R486W [rs201085152, AAO = 60 yr]), *PSEN1* (p.L173S [AAO = 38 yr], p.L262S [AAO = 59 yr], and p.T116I [AAO = 47 yr]) and *PSEN2* (p.N141D [AAO = 59 yr] and p.M298T [AAO = 56 yr]) (Table 1). One variant in the FTD-causal gene *SQSTM1* (p.R309W) was observed, albeit without Sanger sequencing confirmation due to no DNA available (Table 1). The two *PSEN1* variants (p.L173S and p.L262S) and one *PSEN2* variant (p.N141D) should be grouped as *definite pathogenic* (Guerreiro et al., 2010) based on the following lines of evidence. First, different mutations had been reported in the 173rd residue (p.Leu173Trp and p.Leu173Phe) and the 262nd residue (p.Leu262Val, p.Leu262Phe) of *PSEN1* according to the AD mutation database (http://www.molgen.ua.ac.be/ADMutations/). Mutations at both residues (*PSEN1* p.L173 and p.L262) were confirmed to change $A\beta_{42}$ production (Forsell et al., 1997; Kasuga et al., 2009). Second, mutations p.Asn141Tyr and p.Asn141Ile were reported in the

Table 1
Potentially damaging variants in the dementia-causal genes that were only observed in patients with early-onset or familial AD.

Gene	Chromosome: position	rs ID	Ref/alt	Protein change	Damaging	PHRED CADD	AC/AN			*P*-value	Diagnosis	AAO	Familial	APOE ε4
							AD	Controls	ExAC					
South East cohort														
PSEN1	14:73683885	rs63750929	G/T	p.G394V	4	31	1/212	0/160	NA	NA	Probable AD	45	No	–
PSEN2	1:227081771	.	C/A	p.A379D	4	27.1	1/212	0/160	NA	NA	Probable AD	55	No	–
PSEN2	1:227078976	.	C/T	Splice-site	NA	9.977	1/214	0/160	NA	NA	Probable AD	81	Yes	+
GRN	17:42426594	.	C/T	p.P21L	3	24.7	1/212	0/160	0/7866	**0.026**	Probable AD	49	No	+
SQSTM1	5:179260056	.	A/C	p.E176A	3	25.3	1/212	0/160	NA	NA	Probable AD	69	Yes	–
LRRK2	12:40634416	.	C/G	p.P235A	2	18.64	1/212	0/160	NA	NA	Mixed dementia	52	No	–
NOTCH3	19:15298002	.	T/G	p.E585A	3	27.4	1/210	0/160	NA	NA	Probable AD	54	Yes	+
HTRA1	10:124248514	.	G/A	p.R190H	3	25.2	1/212	0/160	0/7866	**0.026**	Possible AD	50	No	–
North cohort														
APP	21:27347385	rs201085152	G/A	p.R486W	4	34	1/124	NA	0/7790	**0.016**	Probable AD	60	Yes	+
PSEN1	14:73653598	.	T/C	p.L173S	4	24.1	1/124	NA	NA	NA	Possible AD	38	Yes	–
PSEN1	14:73664754	.	T/C	p.L262S	4	32	1/124	NA	NA	NA	Probable AD	59	Yes	–
PSEN1	14:73640282	rs63750730	C/T	p.T116I	4	27.1	1/124	NA	NA	NA	Probable AD	47	Yes	–
PSEN2	1:227073303	.	A/G	p.N141D	3	24.9	1/124	NA	NA	NA	Probable AD	59	No	–
PSEN2	1:227078985	.	T/C	p.M298T	2	26	1/124	NA	NA	NA	Probable AD	56	No	–
SQSTM1	5:179263447	rs539942101	C/T	p.R309W	4	26.2	1/124	NA	0/7866	**0.016**	Probable AD	NA	Yes	+

Note: Shown were mutations in AD patients, controls (Zhang et al., 2018) and ExAC dataset (Lek et al., 2016). The mutations were observed in single AD patients and predicted to be damaging by at least two of five algorithms (PolyPhen2 HunDiv and HunVar (Adzhubei et al., 2010), LRT (Chun and Fay, 2009), MutationTaster (Schwarz et al., 2010), and SIFT (Ng and Henikoff, 2003)). Thirteen singletons were confirmed to be true mutations by Sanger sequencing; two patients with different *SQSTM1* singletons (p.E176A and p.R309W) were not sequenced because no DNA samples were available; one *PSEN2* mutation (p.V88M; Table S2) was confirmed to be a sequencing error and were not listed here. The *PSEN2* splice-site chr1:227078976-C-T might not be pathogenic since the carrier was positive for *APOE* ε4 and had a late onset age of 81 years old. Significant *P* values are marked in bold. Ref/alt, reference allele and altered allele according to Genome Reference Consortium Human Build 37 (GRCh37/hg19, https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/); Damaging, number of algorithms with a prediction of damaging effect for each variant; PHRED CADD, PHRED-scaled Combined Annotation-Dependent Depletion (CADD) score (Kircher et al., 2014). AC, allele count of minor allele; AN, allele number (= chromosome number); ExAC, Exome data of 4327 East Asians from the Exome Aggregation Consortium (ExAC, accessed at 2016) (Lek et al., 2016); *P*-value, Fisher's exact test; AAO, age at onset; NA, not available or not applicable; −/+, absence/presence of APOE ε4 allele; Probands with familial history are marked with Yes, otherwise No.

same residue of *PSEN2* p.N141D and affected $A\beta_{42}$ production (Blauwendraat et al., 2016).

3.2. Association of functional variants in the dementia-causal genes with AD risk in Chinese population

Apart from the above case-only potentially pathogenic mutations, there were some variants showing a marginally significant association with AD when compared with the combined controls or population controls (Table 2), suggesting a modifying effect of these dementia genes on AD risk. In the South East cohort, we identified the AD-associated variants in *VCP* (splice-site variant, rs514492), *LRRK2* (p.R1628P and p.R1398H), and *NOTCH3* (p.R640C, p.R1050W, p.A1927T, and p.R2234C) (Table 2). Among them, *VCP* rs514492 and *LRRK2* p.R1398H were common variants. In the North cohort, variants in *VCP* (splice-site variant, rs514492), *FUS* (p.R524K), *LRRK2* (p.L153R, p.A419V, and p.R1628P), *PINK1* (p.R246*, nonsense), *CHMP2B* (p.R205W), *COL4A1* (p.Q1334H), *C9orf72* (chr9:27556798-T/A, splice-site), and *CSF1R* (p.V279M) were identified (Table 2). Among these variants, *VCP* rs514492 and *COL4A1* p.Q1334H were common and showed a marginally association with AD ($P < 0.05$; Table 2).

As our sample size was relatively small, we combined the two cohorts as one population. The common splice-site variant rs514492 in *VCP* showed a positive association with AD risk in both cohorts and the combined population ($P_{combined} = 0.0003$, OR = 1.618), indicating a robust effect of *VCP* rs514492 on AD risk. The rare *LRRK2* variant rs33949390 (p.R1628P) was associated with a reduced risk of AD in both cohorts and the combined population ($P_{combined} = 0.0004$, OR = 0.170) (Table 2). In addition, the *APOE* variant rs429358 defining the ε4 allele showed a strong association with AD risk in both cohorts and the combined population ($P_{combined} = 1.12 \times 10^{-11}$, OR = 3.513). Note that only *VCP* rs514492 and *LRRK2* rs33949390 survived the multiple testing ($P < 0.0029$, 0.05/17 variants in Table 2), together with *APOE* rs429358. However, after correction for potential population stratification by linear regression analysis, only *VCP* rs514492 showed a marginal significance ($P_{adjusted} = 0.015$). Therefore, further validation of the results in independent cohorts with larger sample size was needed.

3.3. Association of the dementia-causal genes with AD risk in populations of European ancestry

In order to cross validate the above association results in Han Chinese with AD, we analyzed the association hits in reported datasets of populations with European ancestry (Bis et al., 2018). The IGAP dataset (Lambert et al., 2013) had data available for one significant SNP rs514492 (*VCP* splice-site), but not the other significant SNPs. Unfortunately, we found no association of rs514492 with AD in the IGAP population (Lambert et al., 2013), suggesting a potential population-specific effect for this variant. We performed the gene-based test to identify potential associations of the highlighted genes in Han Chinese with AD using the IGAP dataset (Lambert et al., 2013), One gene *FUS* showed a significant association with AD at the gene-based level (gene-based test, *P-value* = 0.0048). Four genes (*CSF1R*, *C9orf72*, *COL4A1*, and *LRRK2*) had no significant gene-based *P-values* but had some suggestively associated SNPs (top SNP $P < 0.05$; Table S3).

The significant SNPs identified in Han Chinese had not been included in the ADSP dataset (Bis et al., 2018), but we observed some other suggestively significant SNPs in the above highlighted genes *NOTCH3*, *COL4A1*, *PINK1*, *LRRK2*, *CSF1R*, and *CHMP2B* in the ADSP dataset based on linear regression analyses. Most of these SNPs were intronic or synonymous variants, and only four nonsynonymous variants (*PINK1* rs61744200 [p.R501Q]; *LRRK2* rs33958906 [p.P1542S], rs78365431 [p.Q1111H], and *CSF1R* rs146406037 [p.N255I]) showed a protective effect with a suggestive significance (*P-value* < 0.05) (Table S4).

Table 2
Functional variants in the dementia-causal genes that were associated with AD susceptibility.

Gene	Chromosome: position	rs ID	Ref/alt	Protein	AD		Control		*P*-value	$P_{adjusted}$	OR	ExAC	*P*-value	OR
					AC/AN	Allele Freq	AC/AN	Allele Freq				AC/AN		
South East cohort														
APOE	chr19:45411941	rs429358	T/C	p.C130R	51/208	0.245	61/736	0.083	**3.41E-09**	1.95E-07	3.595	127/850	**0.001**	1.849
APOE	chr19:45412079	rs7412	C/T	p.R176C	9/182	0.049	62/736	0.084	0.124	0.088	0.566	30/396	0.287	0.635
VCP	chr9:35062972	rs514492	C/T	Splice-site	116/212	0.547	315/736	0.428	**0.002**	**0.010**	1.615	3666/7858	**0.021**	1.382
LRRK2	chr12:40713845	rs33949390	G/C	p.R1628P	2/212	0.009	37/736	0.050	**0.005**	0.058	0.180	139/7864	0.591	0.529
LRRK2	chr12:40702911	rs7133914	G/A	p.R1398H	27/212	0.127	46/736	0.063	**0.003**	**0.009**	2.189	825/7860	0.307	1.245
NOTCH3	chr19:15297722	.	G/A	p.R640C	2/212	0.009	0/160	0	0.508	0.999	3.812	1/7866	**0.002**	74.905
NOTCH3	chr19:15291062	rs371525707	G/A	p.R1050W	2/212	0.009	1/526	0.002	0.200	0.634	5.000	5/7856	**0.013**	14.954
NOTCH3	chr19:15276215	.	C/T	p.A1927T	2/212	0.009	0/160	0	0.508	0.999	3.812	4/7838	**0.010**	18.652
NOTCH3	chr19:15271739	rs184996545	G/A	p.R2234C	2/210	0.010	1/366	0.003	0.302	0.399	3.510	0/5022	**0.002**	NA
North cohort														
APOE	chr19:45411941	rs429358	T/C	p.C130R	29/124	0.234	61/736	0.083	**4.21E-06**	1.18E-07	3.378	127/850	**0.025**	1.738
APOE	chr19:45412079	rs7412	C/T	p.R176C	4/124	0.032	62/736	0.084	**0.044**	0.087	0.362	30/396	0.098	0.407
VCP	chr9:35062972	rs514492	C/T	Splice-site	68/124	0.548	315/736	0.428	**0.015**	**0.015**	1.623	3666/7858	0.084	1.389
FUS	chr16:31202749	rs544088874	G/A	p.R524K	1/114	0.009	NA		NA	NA	NA	1/7862	**0.028**	69.566
LRRK2	chr12:40631792	.	T/G	p.L153R	1/124	0.008	0/160	0	0.437	0.999	NA	1/7862	**0.031**	63.911
LRRK2	chr12:40646786	rs34594498	C/T	p.A419V	3/124	0.024	5/576	0.009	0.154	0.214	2.831	49/7848	**0.047**	3.946
LRRK2	chr12:40713845	rs33949390	G/C	p.R1628P	1/124	0.008	37/736	0.050	**0.032**	0.132	0.154	139/7864	0.727	0.452
PINK1	chr1:20966445	rs74315357	C/T	p.R246*	1/124	0.008	NA		NA	NA	NA	1/6282	**0.038**	51.065
CHMP2B	chr3:87302943	rs373536428	C/T	p.R205W	1/124	0.008	NA		NA	NA	NA	2/7848	**0.046**	31.894
COL4A1	chr13:110818598	rs3742207	T/G	p.Q1334H	20/124	0.161	172/736	0.234	0.080	0.056	0.631	1965/7860	**0.021**	0.577
C9orf72	chr9:27556798	.	T/A	Splice-site	1/124	0.008	NA		NA	NA	NA	2/7860	**0.046**	31.943
CSF1R	chr5:149456893	rs3829986	C/T	p.V279M	7/124	0.056	8/576	0.0139	**0.009**	0.221	4.248	260/7866	0.198	1.750
Combined														
APOE	chr19:45411941	rs429358	T/C	p.C130R	80/332	0.241	61/736	0.083	**1.12E-11**	**1.18E-07**	3.513	127/850	**3.30E-5**	1.807
VCP	chr9:35062972	rs514492	C/T	Splice-site	184/336	0.548	315/736	0.428	**0.0003**	**0.015**	1.618	3666/7858	**0.004**	1.384
LRRK2	chr12:40713845	rs33949390	G/C	p.1628RP	3/336	0.009	37/736	0.050	**0.0004**	0.132	0.170	139/7864	0.288	0.501

Note: The mutations significantly associated with AD when the allele frequency of patients was compared with that of controls (Zhang et al., 2018) or ExAC (Lek et al., 2016). Significant *P-values* are marked in bold. Ref/alt, reference allele and altered allele according to Genome Reference Consortium Human Build 37 (GRCh37/hg19, https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/); AC, allele count of minor allele; AN, allele number (= chromosome number); *P-value*, Fisher's exact test; $P_{adjusted}$, adjusted *P*-values were given by linear regression analysis of our case-control comparison, with the first three principle components as the covariates, following the same procedure in our previous study (Zhang et al., 2018); OR, Odds Ratio; ExAC, Exome data of 4327 East Asians from the Exome Aggregation Consortium (accessed at 2016) (Lek et al., 2016), Fisher's exact test was used for comparison of allele frequency in cases and ExAC data; NA, not available or not applicable.

4. Discussion

The clinical diagnosis of AD and other types of dementia has been mainly based on the medical history, neuropsychological assessment, and diagnostic imaging (Jack et al., 2011, 2018; Khachaturian, 2011; Love, 2005). As some symptoms of different types of dementia have been shown to overlap (Tofaris and Buckley, 2018), it has been hard to distinguish AD from the other types of dementia in certain cases. For example, LBD can occur alone or in combination with AD (McKeith et al., 2005); VaD is a commonly seen in AD patients and are usually diagnosed as mixed dementia (Ikram et al., 2017; Khachaturian, 2011; Love, 2005; Waldemar et al., 2007). Therefore, making a definite diagnosis has been challenging, even with postmortem evidence. Genetic testing offers a useful approach to distinguish the different types of dementia; albeit the correlation of phenotype and genotype needs further characterization. Moreover, identification of new pathogenic mutations might add insights for future mechanistic and therapeutic studies.

In this study, we analyzed 20 dementia-causal genes in clinically diagnosed early-onset or familial AD patients. We identified 9 potentially functional mutations in the three well-known AD-causal genes *APP*, *PSEN1*, *PSEN2* in 169 patients (5.3%). However, 6 of the mutations were not found in the available studies and datasets according to a web-based search (Bandelt et al., 2009). Potentially pathogenic mutations were also observed in non-AD dementia-causal genes such as *GRN*, *SQSTM1*, *LRRK2*, *NOTCH3*, and *HTRA1* (Table 1). The shared pathogenic mutations or causal genes between AD and other types of dementia could be explained by the overlapping of the pathophysiological specificity of AD with other types of dementia and/or a common genetic basis. We found that most of the potentially pathogenic mutations in the non-AD genes were found in the South East cohort. This might be partially explained by a higher frequency of mixed dementia in the cohort, especially considering the fact that most of these patients lacked imaging data during the diagnosis. Another interpretation of the data could be the misdiagnosis of the AD patients which should be assigned to the corresponding type of dementia. Assuming the latter interpretation being correct, there would be a high possibility of misdiagnoses or co-occurrence of FTD with AD and VaD with AD in clinical practice. We performed a pathway analysis by using the genes with mutations identified in the patients, but found no significant cluster of any pathway or Gene Ontology term, simply because of the small number of genes under study. Taken together, these findings would suggest a genetic screening for dementia-causal genes is necessary in the diagnosis of early-onset or familial AD. It should be mentioned that the non-coding (GGGGCC) hexanucleotide repeat expansion within the first intron of *C9orf72*, which was the major cause of FTD and ALS (Beck et al., 2013; Dols-Icardo et al., 2014), could not be detected by the whole exome sequencing used in this study. Therefore, the involvement of *C9orf72* in AD needs further study.

The association of the variant rs514492 (splice-site) of the FTD-causal gene *VCP* with AD risk seemed to be robust in this study, although the comparison should be received with caution as the two cohorts under study and the control samples were not completely matched (Table 2). Importantly, we observed a protective rare missense variant rs33949390 (p.R1628P) in the *LRRK2* gene in both cohorts ($P_{combined}$ = 0.0004, OR = 0.170) (Table 2). These observations indicated that the causal genes for other types of dementia or neuropsychiatric disorders might be susceptibility genes, if not causal, for AD too. Indeed, we recently found that risk genes for major depressive disorder might play an active role in AD (Ni et al., 2018). It will be worthwhile to confirm and investigate the involvement of the dementia-causal genes (e.g. *VCP* and *LRRK2*) in AD using independent samples with a large size.

This study had several limitations. First, the analyzed sample size was relatively small, and previous gene recognition was based on clinically diagnosed dementia patients and cognitive normal individuals, and as such we cannot rule out the presence of asymptomatic cases, or even misdiagnoses. Postmortem autopsy-based diagnosis and gene identification are needed to obtain conclusive information for any genetic testing. Second, it is unclear whether the variants observed in this study are biologically pathogenic in the development of AD although program-affiliated *in silico* prediction analyses have indicated they are potentially pathogenic. A focused functional assay, animal model study (Fan et al., 2018; Yao, 2017; Zhang et al., 2019) and a genotype-phenotype correlation analysis in pedigrees with early-onset and familial AD patients are needed for confirming the role of these mutations.

In short, the screening for mutations in 20 dementia-causal genes in 169 Han Chinese patients with early-onset or familial AD has supported the idea of there being overlapping pathomechanisms between AD and other dementias. Our results have also demonstrated the difficulty in treating AD as a single *clinical-pathological* entity (Khachaturian, 2011).

Acknowledgements

We thank Ian Logan for help with language editing and the three reviewers for their helpful comments on the early version of the manuscript. This work was supported by the National Key Research and Development Program of China (2016YFC1306501 to TF), the National Natural Science Foundation of China (31730037 to YGY, 81501464 to GW, 81560230 to HYJ, 81571226 and 81771367 to TF), the Chinese Academy of Sciences (XDB32020200 and QYZDJ-SSW-SMC005 to YGY), and the CAS “Light of West China” Program (中国科学院西部之光项目to DFZ).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpsychires.2019.03.026.

Disclosure statement

The authors declare that they have no competing interests.

References

1000 Genomes Project Consortium, et al., 2015. A global reference for human genetic variation. Nature 526 (7571), 68–74.
Adzhubei, I.A., et al., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7 (4), 248–249.
Alzheimer's association, 2016. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 12 (4), 459–509.
Bandelt, H.J., et al., 2009. Exaggerated status of "novel" and "pathogenic" mtDNA sequence variants due to inadequate database searches. Hum. Mutat. 30 (2), 191–196.
Beck, J., et al., 2013. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. Am. J. Hum. Genet. 92 (3), 345–353.
Bi, R., et al., 2015. Mitochondrial DNA haplogroup B5 confers genetic susceptibility to Alzheimer's disease in Han Chinese. Neurobiol. Aging 36 (3) 1604 e1607-1616.
Bis, J.C., et al., 2018. Whole exome sequencing study identifies novel rare and common Alzheimer's-Associated variants involved in immune response and transcriptional regulation. Mol. Psychiatr. https://doi.org/10.1038/s41380-018-0112-7.
Blauwendraat, C., et al., 2016. Pilot whole-exome sequencing of a German early-onset Alzheimer's disease cohort reveals a substantial frequency of PSEN2 variants. Neurobiol. Aging 37 208 e211-208 e217.
Blennow, K., et al., 2006. Alzheimer's disease. Lancet 368 (9533), 387–403.
Campion, D., et al., 1999. Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. Am. J. Hum. Genet. 65 (3), 664–670.
Chun, S., Fay, J.C., 2009. Identification of deleterious mutations within three human genomes. Genome Res. 19 (9), 1553–1561.
Dols-Icardo, O., et al., 2014. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. Hum. Mol. Genet. 23 (3), 749–754.
Fan, Y., et al., 2018. Does the genetic feature of the Chinese tree shrew (*Tupaia belangeri chinensis*) support its potential as a viable model for Alzheimer's disease research? J. Alzheimer's Dis. 61 (3), 1015–1028.
Forsell, C., et al., 1997. A novel pathogenic mutation (Leu262Phe) found in the presenilin 1 gene in early-onset Alzheimer's disease. Neurosci. Lett. 234 (1), 3–6.
Gauderman, W.J., 2002. Sample size requirements for matched case-control studies of

gene-environment interaction. Stat. Med. 21 (1), 35–50.
Guerreiro, R.J., et al., 2010. Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. Neurobiol. Aging 31 (5), 725–731.
Guerreiro, R., et al., 2013. SnapShot: genetics of Alzheimer's disease. Cell 155 (4) 968-968 e961.
Guerreiro, R., et al., 2015. SnapShot: genetics of ALS and FTD. Cell 160 (4) 798-798.e791.
Ikram, M., et al., 2017. Genetics of vascular dementia - review from the ICVD working group. BMC Med. 15 (1), 48.
Jack Jr., C.R., et al., 2011. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7 (3), 257–262.
Jack, C., et al., 2018. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 14 (4), 535–562.
Kasuga, K., et al., 2009. Depression and psychiatric symptoms preceding onset of dementia in a family with early-onset Alzheimer disease with a novel PSEN1 mutation. J. Neurol. 256 (8), 1351–1353.
Khachaturian, Z.S., 2011. Revised criteria for diagnosis of Alzheimer's disease: National Institute on Aging-Alzheimer's Association diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7 (3), 253–256.
Kircher, M., et al., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46 (3), 310–315.
Lambert, J.C., et al., 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genet. 45 (12), 1452–1458.
Lek, M., et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536 (7616), 285–291.
Leyhe, T., et al., 2017. A common challenge in older adults: classification, overlap, and therapy of depression and dementia. Alzheimers Dement 13 (1), 59–71.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14), 1754–1760.
Liu, J.Z., et al., 2010. A versatile gene-based test for genome-wide association studies. Am. J. Hum. Genet. 87 (1), 139–145.
Love, S., 2005. Neuropathological investigation of dementia: a guide for neurologists. J. Neurol. Neurosurg. Psychiatry 76 (Suppl. 5), v8–14.
Lynch, D., et al., 2017. Clinical and genetic characterization of leukoencephalopathies in adults. Brain 140 (5), 1204–1211.
McKeith, I.G., et al., 2005. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology 65 (12), 1863–1872.
McKenna, A., et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (9), 1297–1303.
McKhann, G., et al., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. Neurology 34 (7), 939–944.
Meeus, B., et al., 2012. The genetics of dementia with Lewy bodies: what are we missing? Arch. Neurol. 69 (9), 1113–1118.
Mishra, A., Macgregor, S., 2015. VEGAS2: software for more flexible gene-based testing. Twin Res. Hum. Genet. 18 (1), 86–91.
Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 31 (13), 3812–3814.
Ni, H., et al., 2018. The GWAS risk genes for depression may be actively involved in Alzheimer's disease. J. Alzheimer's Dis. 64 (4), 1149–1161.
Querfurth, H.W., LaFerla, F.M., 2010. Alzheimer's disease. N. Engl. J. Med. 362 (4), 329–344.
Ridge, P.G., et al., 2013. Alzheimer's disease: analyzing the missing heritability. PLoS One 8 (11) e79771.
Ridge, P.G., et al., 2016. Assessment of the genetic variance of late-onset Alzheimer's disease. Neurobiol. Aging 41 200 e213-220.
Sassi, C., et al., 2018. Mendelian adult-onset leukodystrophy genes in Alzheimer's disease: critical influence of CSF1R and NOTCH3. Neurobiol. Aging 66 179.e117-179.e129.
Schwarz, J.M., et al., 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat. Methods 7 (8), 575–576.
Tofaris, G., Buckley, N., 2018. Convergent molecular defects underpin diverse neurodegenerative diseases. J. Neurol. Neurosurg. Psychiatr. 89 (9), 962–969.
Vergouw, L.J.M., et al., 2017. An update on the genetics of dementia with Lewy bodies. Park. Relat. Disord. 43, 1–8.
Waldemar, G., et al., 2007. Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. Eur. J. Neurol. 14 (1), e1–26.
Wang, K., et al., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38 (16), e164.
Wang, D., et al., 2018. Missense variants in HIF1A and LACC1 contribute to leprosy risk in han Chinese. Am. J. Hum. Genet. 102 (5), 794–805.
Wetterling, T., et al., 1996. Comparison of different diagnostic criteria for vascular dementia (ADDTC, DSM-IV, ICD-10, NINDS-AIREN). Stroke 27 (1), 30–36.
Xiang, Q., et al., 2017. Rare genetic variants of the *Transthyretin* gene are associated with Alzheimer's disease in Han Chinese. Mol. Neurobiol. 54 (7), 5192–5200.
Yao, Y.-G., 2017. Creating animal models, why not use the Chinese tree shrew (Tupaia belangeri chinensis)?Zool. Res. 38 (3), 118–126.
Young, J., et al., 2018. Frontotemporal dementia: latest evidence and clinical implications. Ther Adv Psychopharmacol 8 (1), 33–48.
Zhang, D.F., et al., 2016. CFH variants affect structural and functional brain changes and genetic risk of Alzheimer's disease. Neuropsychopharmacology 41 (4), 1034–1045.
Zhang, D.-F., et al., 2018. Complement C7 is a novel risk gene for Alzheimer's disease in Han Chinese. *National Science Review*. https://doi.org/10.1093/nsr/nwy127.
Zhang, D.F., et al., 2019. Genetic analyses of Alzheimer's disease in China: achievements and perspectives. ACS Chem. Neurosci. 10 (2), 890–901.
Zhu, X., et al., 2006. LRRK2 in Parkinson's disease and dementia with Lewy bodies. Mol. Neurodegener. 1, 17.