Research Paper: Gerotarget (Focus on Aging)

Genetic variations in *chromodomain helicase DNA-binding protein 5,* gene-environment interactions and risk of sporadic Alzheimer's disease in Chinese population

Xiao Zhu[1,7,*], Haibing Yu[1,*], Qin Xiao[2,*], Jianhao Ke[3,*], Hongmei Li[1], Zhihong Chen[1], Hongrong Ding[1], Shuilong Leng[4], Yongmei Huang[6], Jingting Zhan[6], Jinli Lei[6], Wenguo Fan[5] and Hui Luo[1,6]

[1] Key Laboratory of Medical Molecular Diagnosis, Dongguan Scientific Research Center, Guangdong Medical University, Dongguan, China

[2] Department of Blood Transfusion, Peking University Shenzhen Hospital, Shenzhen, China

[3] Tropical Crops Department, Guangdong AIB Polytechnic, Guangzhou, China

[4] Department of Human Anatomy, Guangzhou Medical University, Guangzhou, China

[5] Department of Anatomy and Physiology, Guanghua School of Stomatology, Sun Yat-sen University, Guangzhou, China

[6] Institute of Marine Medicine Research, Guangdong Medical University, Zhanjiang, China

[7] Institute of Bioinformatics, University of Georgia, Athens, GA, USA

[*] These authors have contributed equally to this work

Correspondence to: Hui Luo, **email**: luohui@gdmu.edu.cn

Wenguo Fan, **email**: fanweng@mail.sysu.edu.cn

ABSTRACT

CHD5 is an essential factor for neuronal differentiation and neurodegenerative diseases. Here, the targeted next generation sequencing and TaqMan genotyping technologies were carried out for *CHD5* gene in a two-staged case-control study in Chinese population. The genetic statistics and gene-environment interactions were analyzed to find certain risk factors of Alzheimer's disease. We found intronic rs11121295 was associated with the risk of Alzheimer's disease at both stages including combined cohorts. This risk effect presented consistently significant associations with the alcoholic subgroups at both all stages in the stratified analysis. The gene-environment interactions further supported the above findings. Our study highlighted the potential role of *CHD5* variants in conferring susceptibility to sporadic Alzheimer's disease, especially modified its risk by alcoholic intake.

INTRODUCTION

Alzheimer's disease, a progressive neurodegenerative disorder, is one of the major dementia in elderly people that usually starts slowly and worsens over time [1]. It causes a tremendous societal challenge because the popularity of Alzheimer's disease continues to rise. It affects about 6% of people over 65 years of age, and there are approximately 30 million people worldwide with Alzheimer's disease [2].

The cause of Alzheimer's disease is poorly understood. About 70% of the risk is believed to be genetic with many genes usually involved [3]. Recent studies have proved that a number of polymorphisms impact on the progression of patients with Alzheimer's disease [4, 5]. However, more than 90% of cases are scattered, as a result, only a few percentage points have clear genetic causes. One of the most frequent genetic modifications depicted is the deletion of the short arm of chromosome 1 found in around 35% of neuroblastoma [6]. Other risk factors include a history of head injuries, depression and anxiety, or hypertension.

Chromodomain helicase DNA-binding protein 5 (CHD5) located in the human chromosome 1p36.31.

CHD5 belongs to the ATP-dependent chromatin remodeling protein snf2 DNA methylase/helicase (SNF2) superfamily which is one of the nine members of the chromodomain helicase DNA-binding (CHD) family of enzymes [7, 8]. According to the latest report, the expression of CHD5 was in several brain regions and neurons [9, 10]. And then CHD5 directly regulated the targets including genes which are important for aging, Alzheimer's disease, and neuronal development. Also, intellectual impairment has been coupled with the deletion of a region of chromosome 1 near *CHD5* [11]. However, it remains to be confirmed that the specific role of CHD5

in brain development and function. Based on all of these findings, we set out to investigate whether *CHD5* single-nucleotide polymorphisms (SNPs) were associated with risk of Alzheimer's disease in a two-stage case-control study from China.

RESULTS

In the discovery study, the genotype distributions of 164 candidate variants in the case and control groups are shown in Supplementary Table 1. We found three variants perhaps were associated with Alzheimer's disease risk
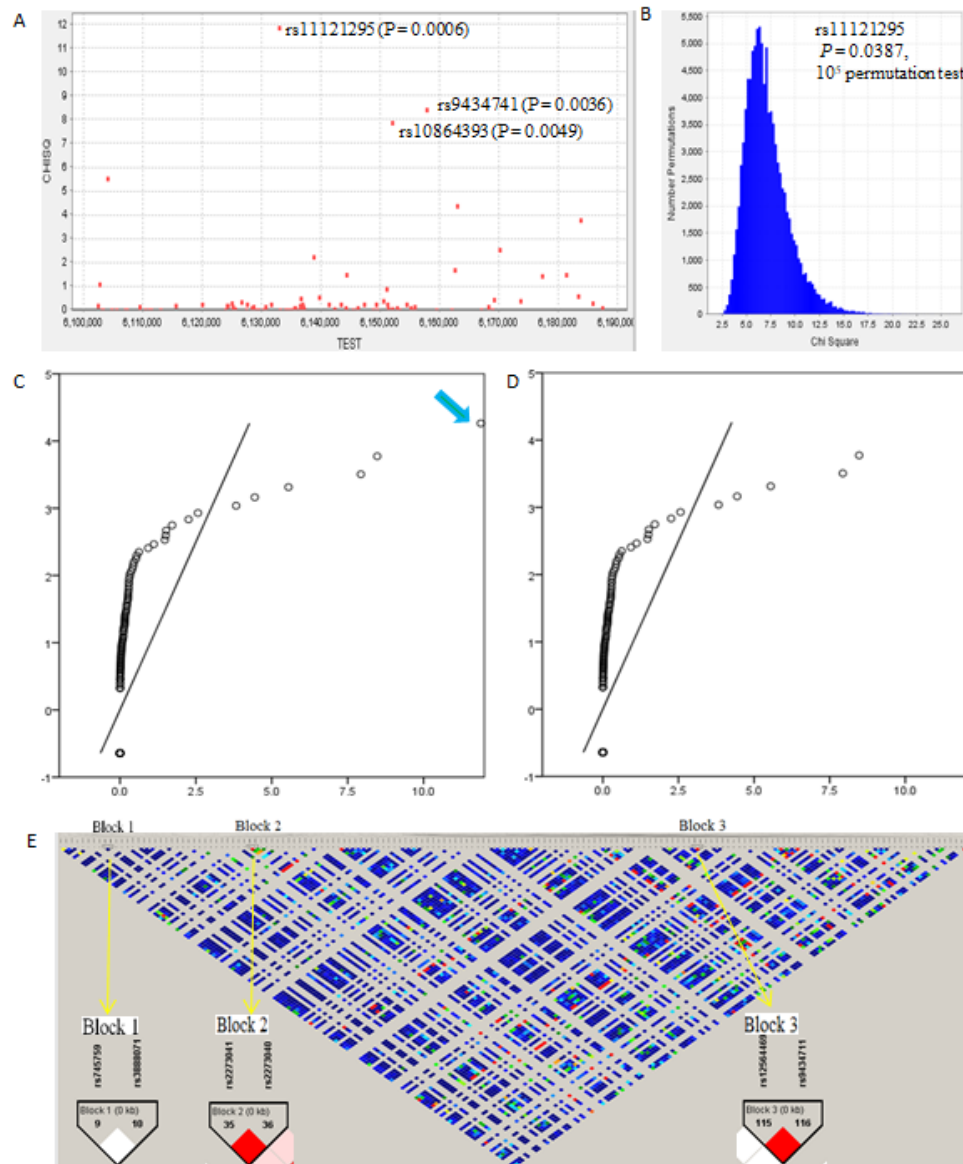
Figure 1: The association analysis of CHD5 variants with Alzheimer's disease risk and genetic mapping in the discovery study. A. Manhattan plot. The *Chi* square values were for the association of each variant with Alzheimer's disease risk, from two-sided Cochran-Armitage tests for trend. B. 10^5 permutation tests for the association analysis carried out from haploview 4.0 program (*Bioinformatics 2005;21:263-5*). C. and D. Q-Q plots for the test statistics of observed Chi-square values against expected Chi-square values (D, removing rs11121295). E. Linkage disequilibrium mapping. Block 1 includes SNP9~SNP10 (rs745759 and rs3888071). Block 2 includes SNP35~SNP36 (rs2273041 and rs2273040). Block 3 includes SNP115 and SNP116 (rs12564469 and rs9434711).

Table 1: Association results of the selected three SNPs from next generation sequencing in discovery, replication and combined studies

SNPs	Cases DD/DA/AA	Controls DD/DA/AA	RR_{Het} (95% CI)	P_{Het}	RR_{Hom} (95% CI)	P_{Hom}	P_{trend}
Discovery
rs11121295	51/99/55	89/130/36	1.33 (0.86-2.05)	.197	2.67 (1.55-4.59)	3.460×10^{-4}	5.495×10^{-4}
rs10864393	55/114/36	49/133/73	0.76 (0.48-1.21)	.249	0.44 (0.25-0.77)	.003	.004
rs9434741	8/60/137	12/112/131	0.80 (0.31-2.07)	.651	1.57 (0.62-3.96)	.337	.003
Replication
rs11121295	105/198/94	174/262/74	1.25 (0.92-1.70)	.147	2.11 (1.43-3.11)	1.603×10^{-4}	2.760×10^{-4}
rs10864393	76/225/96	94/262/154	1.06 (0.75-1.51)	.736	0.77 (0.52-1.15)	.197	.137
rs9434741	20/173/204	24/236/250	0.88 (0.47-1.64)	.687	0.98 (0.53-1.82)	.947	.605
Combined
rs11121295	156/297/149	263/392/110	1.28 (1.00-1.64)	.054	2.28 (1.67-3.13)	2.445×10^{-7}	6.463×10^{-7}
rs10864393	131/339/132	143/395/227	0.94 (0.71-1.24)	.645	0.64 (0.46-0.87)	.005	.003
rs9434741	28/233/341	36/348/381	0.86 (0.51-1.45)	.573	1.15 (0.69-1.93)	.593	.031

D,derived alleles; A, ancestral alleles; Het, heterozygous variants; Hom, homozygous variants.
P_{trend} were calculated in logistical regression models with adjustment for age, gender, smoking and drinking status.

(rs11121295, $P_{for\ alleles} = 6.00\times10^{-4}$; rs10864393, $P_{for\ alleles} = 0.0049$; and rs9434741, $P_{for\ alleles} = 0.0036$; Figure 1A). Even after 10^5 permutation tests, there was still remain significance for rs11121295 ($P = 0.0387$; Figure 1B). The observed X^2 values with the distribution of null hypothesis were deviated from the expectations at higher value of approximately 4.3 (Figure 1C). But it lost its significance after removing rs11121295 (Figure 1D).

Indeed, rs11121295 GG genotype and rs10864393 AA genotype all showed significance of risk with Alzheimer's disease in the discovery cohort ($RR_{Hom} = 2.67$, $P_{Hom} = 3.460\times10^{-4}$ for rs11121295; $RR_{Hom} = 0.44$, $P_{Hom} = 0.003$ for rs10864393; Table 1). These risk qualities increased in a dose-dependent manner as genotypes increased or decreased ($P_{trend} = 5.495\times10^{-4}$ for rs11121295, and .004 for rs10864393).

We identified three blocks with high linkage disequilibrium (Figure 1E). Block 1 includes SNP9~SNP10 (rs745759 and rs3888071). Block 2 includes SNP35~SNP36 (rs2273041 and rs2273040). Block 3 includes SNP115 and SNP116 (rs12564469 and rs9434711). The results of the haplotype-based case-control study between the Alzheimer's disease and control groups are shown in Supplementary Table 2. We found that a haplotype AG in block 3 showed a significant association with Alzheimer's disease risk ($P = 1.620\times10^{-5}$). Nevertheless, it is actually difficult to determine Alzheimer's disease risk, because the proportions of the frequencies of haplotype AG were all too small (0.055 in cases and 0.005 in controls).

According to the discovery study, SNPs rs11121295, rs10864393 and rs9434741 were selected for the replication and combined studies. Genotype frequencies distribution of rs11121295 was significantly different between the two groups not only in replication study ($RR_{Hom} = 2.11$, $P_{Hom} = 1.603\times10^{-4}$ and $P_{trend} = 2.760\times10^{-4}$) but combined study ($RR_{Hom} = 2.28$, $P_{Hom} = 2.445\times10^{-7}$ and $P_{trend} = 6.463\times10^{-7}$; Table 1).

This association of rs11121295 with the Alzheimer's disease risk was further evaluated by subgroups of age, sex, education, smoking and drinking status with replication and combined studies (Table 2). We found that those carried rs11121295 had a significantly increased risk, and this risk was more obvious not only in the older group (discovery: $P = 0.005$, and $P_{i\ (P2/P1)} = 0.011$; replication: $P = 8.700\times10^{-4}$, and $P_{i\ (P2/P1)} = 0.003$), but in the drinking group (discovery: $P = 6.941\times10^{-4}$, and $P_{i\ (P2/P1)} = 0.008$; replication: $P = 0.002$, and $P_{i\ (P2/P1)} = 0.049$) in all two case-control studies, especially their combined study ($P = 1.508\times10^{-5}$, and $P_{i\ (P2/P1)} = 8.621\times10^{-5}$ for older group; and $P = 7.164\times10^{-6}$, and $P_{i\ (P2/P1)} = 8.755\times10^{-4}$ for drinking group).

For Alzheimer's disease risk in older group, the false-positive report probability (FPRP) values of

Table 2: Stratification analysis for associations between rs11121295 and AD risk in the discovery, replication and combined studies

Variables	Discovery study				Replication study				Combined study			
	Cases/controls		OR (95% CI)	P	Cases/controls		OR (95% CI)	P	Cases/controls		OR (95% CI)	P
	GG	AA+AG			GG	AA+AG			GG	AA+AG		
	55/36	150/219			94/74	303/436			149/110	453/655		
Age (ys)												
≤60	2/33	20/183	0.56(0.12-2.49)	.435	4/68	39/366	0.55(0.19-1.60)	.266	6/101	59/549	0.55(0.23-1.31)	.174
>60	53/3	130/36	4.89(1.44-16.58)	.005	90/6	264/70	3.98(1.67-9.47)	$8.700×10-4$	143/9	394/106	4.28(2.11-8.67)	$1.508×10-5$
P_i				.011				.003				$8.621×10-5$
Gender												
Females	29/11	85/80	2.48(1.16-5.30)	.017	47/18	161/149	2.42(1.34-4.35)	.003	76/29	246/229	2.44(1.53-3.88)	$1.220×10-4$
Males	26/25	65/139	2.22(1.19-4.15)	.011	47/56	142/287	1.70(1.10-2.63)	.017	73/81	207/426	1.86(1.30-2.65)	$6.325×10-4$
P_i				.647				.176				.193
Education												
Less	34/18	108/116	2.03(1.08-3.80)	.026	42/26	145/149	1.66(0.97-2.85)	.064	76/44	253/265	1.81(1.20-2.72)	$4.207×10-3$
More	21/18	47/87	2.16(1.05-4.45)	.035	52/43	140/229	1.98(1.25-3.12)	.003	73/61	187/316	2.02(1.38-2.97)	$2943×10-4$
P_i				.743				.047				.070
Smoking												
Never	33/27	115/165	1.97(1.13-3.41)	.015	59/77	208/262	0.97(0.66-1.42)	.857	96/104	323/427	1.22(0.89-1.67)	.212
Ever+current	18/9	35/47	2.69(1.08-6.69)	.031	25/29	92/116	1.09(0.60-1.98)	.786	43/38	127/163	1.45(0.89-2.38)	.138
P_i			–	184				.917				.651
Drinking												
Never	29/26	113/168	1.66(0.93-2.96)	.086	54/55	189/297	1.54(1.02-2.34)	.041	83/81	302/465	1.58(1.13-2.21)	.008
Ever+current	26/7	37/47	4.72(1.85-12.07)	$6.941×10-4$	41/18	98/111	2.58(1.39-4.78)	.002	67/25	135/158	3.14(1.88-5.24)	$7.164×10-6$
P_i				.008				.049				$8.755×10-4$

P, *P* value for haplotype model, which obtained in logistic regression with adjustment for age, sex, smoking status and drinking status.

P_i, means P_2/P_1 or P_1/P_2.

rs11121295 GG were below 0.20 for the assigned prior probability (0.008 for the prior probability of 0.1 in the discovery study; 0.005, 0.010 for the prior probabilities of 0.1, 0.01, respectively in the replication study; and 0.002, 0.009, 0.081 for the prior probabilities of 0.1, 0.01, 0.001 respectively in the combined study). For Alzheimer's disease risk in drinking group, when the assumption of prior probability was 0.1, significant findings were noteworthy in the replication study but in the combined study (prior probability 0.011 and 0.004, respectively). Moreover, when the assumption of prior probability was 0.01, this prominent association was just found in the combined study (prior probability 0.010; Table 3).

Lastly, gene-environmental interactions were studied. We found that the age was the best one-factor model with the highest cross-validation consistency (CVC) (94/100, 93/100, 97/100) and the lowest prediction error (0.392, 0.383, 0.336) in three statistical analyses respectively (Table 4). The age plus rs11121295 was the best model for two-factors, with the highest CVC (93/100, 96/100, 97/100) and the lowest prediction error (0.323, 0.372, 0.312) in three groups respectively. Interestingly, the model with 6-factors had a maximum CVC (100/100,

99/100, 100/100) and a minimum prediction error (0.261, 0.211, 0.134) in three groups respectively, which presents a model with better prediction than one factor.

DISCUSSION

The genome-wide association studies have found some chromosome regions including rare variants that appear to affect Alzheimer's disease risk [3, 12, 13]. The potential genes including CD2AP, SlC24A4, HLA-DRB5, etc. For example, variants in the *TREM2* gene might be associated with a three to five times higher risk [14]. In our study, we reported the positive association between the rs11121295 homozygous variant and Alzheimer's disease not only in the discovery cohort but in the replication and combined studies with targeted next generation sequencing and TaqMan genotyping technologies, which has not been reported before.

The quantile-quantile (Q-Q) plots are plots of two quantiles against each other, which could examine if two data sets have roots in the same distribution. If the two sets of data have a common distribution, the points will fall on that reference line of forty-five degree angle. In

Table 3: FPRP values for associations between AD risk and rs11121295 frequencies (GG vs. AA+AG)

Variables	Statistical power[a]	Prior probability			
		.1	.01	.001	.0001
AD risk in >60 years old group					
Discovery study	.327	.008	.029	.443	.742
Replication study	.552	.005	.010	.288	.539
Combined study	.815	.002	.009	.081	.205
AD risk in drinking group					
Discovery study	.434	.254	.418	.585	.737
Replication study	.276	.011	.366	.649	.813
Combined study	.691	.004	.010	.203	.516

Table 4: MDR analysis for the prediction of AD risk with and without rs11121295

Best interaction models	Cross-validation	Average prediction error	P^a
Distcovery study			
1	94/100	0.392	<.0001
1,2	93/100	0.323	<.0001
1,2,3	97/100	0.354	<.0001
1,2,3,4	98/100	0.225	<.0001
1,2,3,4,5	90/100	0.329	<.0001
1,2,3,4,5,6[b]	100/100	0.261	<.0001
Replication study			
1	93/100	0.383	<.0001
1,2	96/100	0.372	<.0001
1,2,3	92/100	0.394	<.0001
1,2,3,4	97/100	0.373	<.0001
1,2,3,4,5	89/100	0.256	<.0001
1,2,3,4,5,6[b]	99/100	0.211	<.0001
Combined study			
1	97/100	0.336	<.0001
1,2	97/100	0.312	<.0001
1,2,3	90/100	0.375	<.0001
1,2,3,4	96/100	0.213	<.0001
1,2,3,4,5	97/100	0.168	<.0001
1,2,3,4,5,6[b]	100/100	0.134	<.0001

Labels: 1, age; 2, rs11121295; 3, drinking status; 4, education; 5, gender; 6, smoking status.
a, *P* value for 1000-fold permutation test. b, the best model with maximum cross-validation consistency and minimum prediction error rate.

Table 5: Demographics of patients with Alzheimer's disease and controls in three study cohorts

Characteristics	Discovery cohort			Replicative cohort			Combined cohort		
	Cases	Controls	P	Cases	Controls	P	Cases	Controls	P
No.of subjects	205	255		397	510		602	765	
Origin	NC	NC		SC	SC		NC+SC	NC+SC	
Females, n (%)	114 (55.61)	91 (35.69)	<.001[a]	208 (52.39)	167 (32.75)	<.001[a]	322 (53.49)	258 (33.73)	<.001[a]
Age (mean ± SD)	73.5±10.2	41.5±9.1	<.001[b]	73.0±10.2	47.2±10.7	<.001[b]	73.2±10.3	44.8±10.3	<.001[b]
Less education, n (%)	137 (66.83)	134 (52.55)	.002[a]	187 (47.10)	175 (34.31)	.003[a]	324 (53.82)	309 (40.39)	<.001[a]
missing	0	16		18	63		18	79	
Cigarette smoking, n (%)	53 (25.85)	56 (21.96)	.417[a]	117 (29.47)	145 (28.43)	.871[a]	170 (28.24)	201 (26.27)	.573[a]
missing	0	7		13	26		13	33	
Alcohol drinking, n (%)	63 (30.73)	54 (21.18)	.030[a]	139 (35.01)	129 (25.29)	.003[a]	202 (33.55)	183 (23.92)	<.001[a]
missing	0	7		15	29		15	36	
Hypertention, n (%)	26 (12.68)	0		43 (10.83)	0		69 (11.46)	0	
Stroke, n (%)	4 (1.95)	0		7 (1.76)	0		11 (1.83)	0	
Diabetes, n (%)	30 (14.63)	0		51 (12.85)	0		81 (13.46)	0	

NC, North China; SC, South China; SD, standard deviation.
[a]Pearson Chi-square test. [b]Mann-Whitney U-test.

our study, a significant change appeared after taking away rs11121295, further indicating that it is the logical risk locus.

The 1p36 is frequently deleted in neural crest tumors including neuroblastomas, raising that this region may contribute to the neurologic and developmental issues. *CHD5* gene contains 42 exons and spans more than 78 kb, it *acts on acid anhydrides in phosphorus-containing anhydrides annotated by g*ene ontology. A downstream gene CDKN2A could regulate the P53 pathway in particular, which in turn, impedes cell proliferation [9]. These suggested CHD5 plays an important role in the neurogenesis and development by activating expression of specific genes promoting neuron terminal differentiation.

Moreover, this study first explored the potential gene-environment interactions by stratification (addressed ORs) adding high-order interactions assessed by FPRP and multiple dimension reduction (MDR) analyses with five known risk factors (age, gender, education, smoking and drinking status), which suggested that older age (> 60 years), drinking, as well as rs11121295, contributed to an increased Alzheimer's disease risk at different levels. These experiment-wise results further revealed that potential gene-environment interactions seem to predispose to Alzheimer's disease. Although the FPRP can yield serious inferential errors, the FPRP was still proposed as a Bayesian prophylactic against false reports

of significant associations [15] .

In developed countries, dementia is one of the most financially costly diseases. It led to about 1.9 million deaths every year. Of them, the cause of 60% to 70% of cases is Alzheimer's disease. Its most common symptoms include language barrier, orientation disorder, mood swings, loss of self-control and behavioural issues. Alzheimer's disease most often begins in people over 60 years and older, although about 4% of cases are early-onset which begin before this. Growing evidence from previous epidemiological studies and meta-analysis has firmly established the important role of older age as a causal factor for Alzheimer's disease [16-18]. A report from the World Health Organization deemed that prevalence rates in developing regions like China are lower than that in developed regions [19]. Chan *et al.* reported Alzheimer prevalence was estimated to be 2.6% in the 65-69 age group, and 60.5% in the age 95-99 years in China in 2010 [20]. This suggested the incidence of Alzheimer's disease increases with age. Though unmatched ages with cases were used in this case-control study would lead to incorrect risk assessment, a proper statistical analysis can well control for differences between the groups. Estimates of effects can be statistically adjusted for covariates that may be different between cases and controls. Logistic regression is mostly used, the weighted one being known as the one that controls for confounders for example. This

may still be good enough for the purpose, therefore, the results are still quite valuable.

The genetic heritability of Alzheimer's disease ranges from 49% to 79% based on family studies, and about 0.1 percent of the cases are autosomal dominant inheritance, which have an onset before age 65 [21-23]. But most cases do not exhibit autosomal-dominant inheritance and are termed sporadic Alzheimer's disease, in which genetic and environmental differences may act as risk factors. In our study, drinking status with *CHD5* rs11121295 variant presented a promising association with Alzheimer's disease risk compared with non-drinking carrying that genotype, which indicated that this variant might act in response to drinking, and true associations might be detected by alcoholic stimulation. It has been shown that alcohol could modulate the effect of various cytokines, receptors or neuroimmune signaling in brain, such as midkine (MDK) [24], PPARgamma receptor [25], and HMGB1, miRNA and TLR receptors [26]. Therefore, we speculated, alcohol intake might trigger proinflammatory events through their induction of oxidative stress and extensive inflammation.

The process of Alzheimer's disease is related with tangles and plaques in the brain [2, 27]. The standard diagnosis is in view of the medical history, cognitive examinations, imaging check and blood testing to preclude other possible causes. However, evidence to support the proposals is usually not sufficient. Therefore, clinical genetic examinations may play an important role in the future.

In summary, we identified rs11121295 variant in *CHD5* gene that was highly associated with risk of developing Alzheimer's disease from Chinese descent. CHD5 is a nuclear protein which forms a nucleosome remodeling and deacetylation (NuRD) complex, and it may affect the brain neurons in the level of chromatin remodeling and gene transcription [28]. Nevertheless, we do not yet understand the mechanisms of how it is impacting on the expression of CHD5 protein in the brain. Hence, it will be much more interesting as further investigations of this gene are implemented, not only in Alzheimer's disease patients but also in healthy commons.

MATERIALS AND METHODS

Subjects

In the first step, 205 unrelated Alzheimer's disease patients and 255 healthy control people who had no history of Alzheimer's disease and other conspicuous diseases from Zibo Center Hospital in North China were included as the discovery study. Then, a replicative study including 397 Alzheimer's disease cases and 510 controls (from Guangdong Provincial People's Hospital, Guangzhou

Brain Hospital and Peking University Shenzhen Hospital in South China) was carried out. Finally, the above both were included as the combined study. At recruitment, each study participant (or his/her relative) was interviewed *via* a structured questionnaire, to obtain information on demographic characteristics, habits of cigarette smoking and alcohol drinking, as well as personal and family history of major chronic illnesses. A pack of cigarettes was defined as 20 cigarettes in China. "Ever or current smoking" were defined by valuing subjects who had smoked more than 5 packs in their whole life before the date of diagnosis for cases, or before the date of the interview for controls [29]. "Ever or current drinking" as having consumed alcoholic beverages ≥1 time/week for ≥ 6 months previously; otherwise, they were defined as non-drinkers [30]. One drink was regarded as 30 g of spirits (12.9 g of ethanol), 103 g of wine (12.3 g of ethanol), or 360 g of beer (12.6 g of ethanol) [30]. Less education implied that he only accept the primary school education or less [31]. The main features of the subjects included are summarized in Table 5. The Ethics committee of Guangdong Medical University authorized the protocol of this study. The study also adhered to tenets in the declaration of Helsinki.

Targeted sequencing, variants selection and genotyping

Genomic DNA was extracted from the whole blood leukocytes. We sequenced whole *CHD5* gene with next generation sequencing technology (Illumina Genome Analyzer) in 255 controls and 205 Alzheimer's disease samples. A targeted resequencing study was performed on the Illumina platform with pair-end 90 bp reads. Following the manufacturer's instructions, shotgun libraries were built from 5 microgram of genomic DNA, and genomic DNA diluted in Tris-EDTA buffer was sheared into about 500-bp fragments. The DNA fragments were subsequently tailed with A. Then the Illumina sequencing adaptors were ligated to the samples. Finally, the adaptor-linked fragments were enriched *via* PCRs. The prepared library was subsequently hybridized to capture probes. The captured fragments were then amplified with the following protocol: incubation at 95 °C for 5 min followed by 25 cycles of 95 °C for 15 s, 56 °C for 30 s and 72 °C for 60 s and a final extension at 72 °C for 8 min. PCR products were purified and finally sequenced with standard 2×90-bp paired-end reads on the Illumina HiSeq 2000 sequencer.

The reads were aligned to the reference genome hg19 (NCBI build 37.1) (NCBI build 37.1) [32]. Single-nucleotide variants that met any of the following criteria were then filtered: *P* for Hardy-Weinberg equilibrium < 10^{-4} [33], duplicated pair-end reads, overall depth ≤ 8×, copy number variant ≥ 2, or SNP within 10 bp of a gap. In this evaluation, we only considered the qualified SNPs,

thus yielding a 164-SNPs set, which will be used as the primary case-controls study.

We analyzed the associations of variants (allele frequencies > 1%) and Alzheimer's disease risk. Only three variants (rs11121295, rs10864393 and rs9434741) entered the next step study for their lower P values. Then, genomic DNAs from all the other subjects (510 controls and 397 cases) were genotyped by TaqMan probes in Applied Biosystems ABI 7500 Fast System (Forster City, CA) for the above selected three variants. The PCR of samples heated to 95°C for 10 min followed by 40 cycles of 92°C for 15s and 60°C for 1 min.

Statistical analysis

Chi-square test and Mann-Whitney U-test were used to assess the difference of demographics between cases and controls. The genotype distributions in controls were analyzed with Hardy-Weinberg equilibrium ($P_{HWE} > 0.01$).

Haplotype estimation and permuting association analysis were executed with Haploview program [34] for 10^5 permutations (the 'Single Markers Only' option was used), in which the subjects' phenotypes were randomly realigned. By convention if $P < 0.05$, the difference was considered statistically significant. The Q-Q plot was then performed to check the distributions of P value. We used homozygote (DD vs. AA) and heterozygote (DA vs. AA), as the models (D-derived alleles and A-ancestral alleles). Logistic regression was used to test the association, which adjusted for gender, age, education, smoking and drinking. A dose-dependent effect was assessed by the trend test of odds ratios (ORs). The variant(s) (which P for genotypes < $0.0003_{(0.05/164)}$) will be entering the next replication study. The FPRP and MDR program [35] was used to evaluate the possible high-order gene-environment interaction. The minimum average prediction error and the maximum CVC were required for the best candidate interaction model. SPSS 22.0 for windows (SPSS, Chicago, IL) and R scripts (3.0.2 Suite) were performed in the statistical analyses.

ACKNOWLEDGMENTS AND FUNDING

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

REFERENCES

1. Pistollato F, Ohayon EL, Lam A, Langley GR, Novak TJ, Pamies D, Perry G, Trushina E, Williams RS, Roher AE, Hartung T, Harnad S, Barnard N, et al. Alzheimer disease research in the 21st century: past and current failures, new perspectives and funding priorities. Oncotarget. 2016; 7: 38999-9016. https://doi.org/10.18632/oncotarget.9175.

2. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. Lancet. 2011; 377: 1019-31. doi: 10.1016/S0140-6736(10)61349-9.

3. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, et al, and Cohorts for Heart and Aging Research in Genomic Epidemiology. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013; 45: 1452-8. doi: 10.1038/ng.2802.

4. Lalli MA, Bettcher BM, Arcila ML, Garcia G, Guzman C, Madrigal L, Ramirez L, Acosta-Uribe J, Baena A, Wojta KJ, Coppola G, Fitch R, de Both MD, et al. Whole-genome sequencing suggests a chemokine gene cluster that modifies age at onset in familial Alzheimer's disease. Mol Psychiatry. 2015; 20: 1294-300. doi: 10.1038/mp.2015.131.

5. Traylor M, Adib-Samii P, Harold D, Dichgans M, Williams J, Lewis CM, Markus HS, and Alzheimer's Disease Neuroimaging Initiative, and International Stroke Genetics Consortium (ISGC), UK Young Lacunar Stroke DNA resource, and METASTROKE, and International Genomics of Alzheimer's Project (IGAP), investigators. Shared genetic contribution to Ischaemic Stroke and Alzheimer's Disease. Ann Neurol. 2016; 79:739–47. doi: 10.1002/ana.24621.

6. Bagchi A, Papazoglu C, Wu Y, Capurso D, Brodt M, Francis D, Bredel M, Vogel H, Mills AA. CHD5 is a tumor suppressor at human 1p36. Cell. 2007; 128: 459-75. doi: 10.1016/j.cell.2006.11.052.

7. Schuster EF, Stoger R. CHD5 defines a new subfamily of chromodomain-SWI2/SNF2-like helicases. Mamm Genome. 2002; 13: 117-9. doi: 10.1007/s00335-001-3042-6.

8. Li W, Mills AA. Packing for the journey: CHD5 remodels the genome. Cell Cycle. 2014; 13: 1833-4. doi: 10.4161/cc.29378.

9. Egan CM, Nyman U, Skotte J, Streubel G, Turner S, O'Connell DJ, Rraklli V, Dolan MJ, Chadderton N, Hansen K, Farrar GJ, Helin K, Holmberg J, et al. CHD5 is required for neurogenesis and has a dual role in facilitating gene expression and polycomb gene repression. Dev Cell. 2013; 26: 223-36. doi: 10.1016/j.devcel.2013.07.008.

10. Lachen-Montes M, Zelaya MV, Segura V, Fernández-Irigoyen J, Santamaría E. Progressive modulation of the

human olfactory bulb transcriptome during Alzheimer´s disease evolution: novel insights into the olfactory signaling across proteinopathies. Oncotarget. 2017; 8:69663–79. https://doi.org/10.18632/oncotarget.18193.

11. Vestin A, Mills AA. The tumor suppressor Chd5 is induced during neuronal differentiation in the developing mouse brain. Gene Expr Patterns. 2013; 13: 482-9. doi: 10.1016/j.gep.2013.09.003.

12. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet. 2007; 39: 17-23. doi: 10.1038/ng1934.

13. Vardarajan BN, Ghani M, Kahn A, Sheikh S, Sato C, Barral S, Lee JH, Cheng R, Reitz C, Lantigua R, Reyes-Dumeyer D, Medrano M, Jimenez-Velazquez IZ, et al. Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Ann Neurol. 2015; 78: 487-98. doi: 10.1002/ana.24466.

14. Carrasquillo MM, Allen M, Burgess JD, Wang X, Strickland SL, Aryal S, Siuda J, Kachadoorian ML, Medway C, Younkin CS, Nair A, Wang C, Chanana P, et al. A candidate regulatory variant at the TREM gene cluster associates with decreased Alzheimer's disease risk and increased TREML1 and TREM2 brain gene expression. Alzheimers Dement. 2017; 13: 663-73. doi: 10.1016/j.jalz.2016.10.005.

15. Lucke JF. A critique of the false-positive report probability. Genet Epidemiol. 2009; 33: 145-50. doi: 10.1002/gepi.20363.

16. Lee JH, Cheng R, Vardarajan B, Lantigua R, Reyes-Dumeyer D, Ortmann W, Graham RR, Bhangale T, Behrens TW, Medrano M, Jimenez-Velazquez IZ, Mayeux R. Genetic Modifiers of Age at Onset in Carriers of the G206A Mutation in PSEN1 With Familial Alzheimer Disease Among Caribbean Hispanics. JAMA Neurol. 2015; 72: 1043-51. doi: 10.1001/jamaneurol.2015.1424.

17. Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, Thompson WK, Besser L, Kukull WA, Holland D, Chen CH, Brewer JB, Karow DS, et al. Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. PLoS Med. 2017; 14: e1002258. doi: 10.1371/journal.pmed.1002258.

18. Carro E, Bartolome F, Bermejo-Pareja F, Villarejo-Galende A, Molina JA, Ortiz P, Calero M, Rabano A, Cantero JL, Orive G. Early diagnosis of mild cognitive impairment and Alzheimer's disease based on salivary lactoferrin. Alzheimers Dement (Amst). 2017; 8: 131-8. doi: 10.1016/j.dadm.2017.04.002.

19. World Health Organization. Neurological Disorders: Public Health Challenges. Switzerland: World Health Organization. 2006; pp. 204-07: ISBN 978-92-4-156336-9. doi: 10.1016/S0140-6736(13)60221-4.

20. Chan KY, Wang W, Wu JJ, Liu L, Theodoratou E, Car J, Middleton L, Russ TC, Deary IJ, Campbell H, Wang W, Rudan I, and Global Health Epidemiology Reference Group (GHERG). Epidemiology of Alzheimer's disease and other forms of dementia in China, 1990-2010: a systematic review and analysis. Lancet. 2013; 381: 2016-23. doi: 10.1016/S0140-6736(13)60221-4.

21. Wilson RS, Barral S, Lee JH, Leurgans SE, Foroud TM, Sweet RA, Graff-Radford N, Bird TD, Mayeux R, Bennett DA. Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. J Alzheimers Dis. 2011; 23: 249-55. doi: 10.3233/JAD-2010-101515.

22. Day GS, Musiek ES, Roe CM, Norton J, Goate AM, Cruchaga C, Cairns NJ, Morris JC. Phenotypic Similarities Between Late-Onset Autosomal Dominant and Sporadic Alzheimer Disease: A Single-Family Case-Control Study. JAMA Neurol. 2016; 73: 1125-32. doi: 10.1001/jamaneurol.2016.1236.

23. Kohli MA, Cukier HN, Hamilton-Nelson KL, Rolati S, Kunkle BW, Whitehead PL, Zuchner SL, Farrer LA, Martin ER, Beecham GW, Haines JL, Vance JM, Cuccaro ML, et al. Segregation of a rare TTC3 variant in an extended family with late-onset Alzheimer disease. Neurol Genet. 2016; 2: e41. doi: 10.1212/NXG.0000000000000041.

24. Chen H, He D, Lasek AW. Midkine in the mouse ventral tegmental area limits ethanol intake and Ccl2 gene expression. Genes Brain Behav. 2017; 16:699–708. doi: 10.1111/gbb.12384.

25. Cippitelli A, Domi E, Ubaldi M, Douglas JC, Li HW, Demopulos G, Gaitanaris G, Roberto M, Drew PD, Kane CJM, Ciccocioppo R. Protection against alcohol-induced neuronal and cognitive damage by the PPARgamma receptor agonist pioglitazone. Brain Behav Immun. 2017; 64: 320-9. doi: 10.1016/j.bbi.2017.02.001.

26. Crews FT, Lawrimore CJ, Walter TJ, Coleman LG Jr. The role of neuroimmune signaling in alcoholism. Neuropharmacology. 2017; 122: 56-73. doi: 10.1016/j.neuropharm.2017.01.031.

27. LeBlanc AC. Novel therapeutic target against Alzheimer. Oncotarget. 2017; 8: 48529-30. doi: 10.18632/oncotarget.18210.

28. Paul S, Kuo A, Schalch T, Vogel H, Joshua-Tor L, McCombie WR, Gozani O, Hammell M, Mills AA. Chd5 requires PHD-mediated histone 3 binding for tumor suppression. Cell Rep. 2013; 3: 92-102. doi: 10.1016/j.celrep.2012.12.009.

29. Tian J, Venn AJ, Blizzard L, Patton GC, Dwyer T, Gall SL. Smoking status and health-related quality of life: a longitudinal study in young adults. Qual Life Res. 2016; 25: 669-85. doi: 10.1007/s11136-015-1112-6.

30. Yuan JM, Ross RK, Gao YT, Henderson BE, Yu MC. Follow up study of moderate alcohol intake and mortality among middle aged men in Shanghai, China. BMJ. 1997; 314:18–23. doi: 10.1136/bmj.314.7073.18.

31. Quaranta D, Gainotti G, Vita MG, Lacidogna G, Scaricamazza E, Piccininni C, Marra C. Are Raw Scores

on Memory Tests Better than Age- and Education-Adjusted Scores for Predicting Progression from Amnesic Mild Cognitive Impairment to Alzheimer Disease? Curr Alzheimer Res. 2016; 13:1414–20. doi: 10.2174/1567205 013666160314145522.

32. Gu S, Fang L, Xu X. Using SOAPaligner for Short Reads Alignment. Curr Protoc Bioinformatics. 2013; 44:11.11.1-17. doi: 10.1002/0471250953.bi1111s44.

33. Cheng AY, Teo YY, Ong RT. Assessing single nucleotide variant detection and genotype calling on whole-genome sequenced individuals. Bioinformatics. 2014; 30: 1707-13. doi: 10.1093/bioinformatics/btu067.

34. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21: 263-5. doi: 10.1093/bioinformatics/bth457.

35. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics. 2003; 19:376–82. doi: 10.1093/bioinformatics/btf869.