Mol Neurobiol (2018) 55:1217–1226
DOI 10.1007/s12035-017-0397-6

The *Arc* Gene Confers Genetic Susceptibility to Alzheimer's Disease in Han Chinese

Rui Bi[1] · Li-Li Kong[1,2] · Min Xu[1,3] · Guo-Dong Li[1,3] · Deng-Feng Zhang[1] · Alzheimer's Disease Neuroimaging Initiative · Tao Li[4] · Yiru Fang[5] · Chen Zhang[5] · Buchang Zhang[2] · Yong-Gang Yao[1,6]

Received: 12 September 2016 / Accepted: 10 January 2017 / Published online: 20 January 2017

Abstract Alzheimer's disease (AD) is the most common form of dementia. The deposition of β-amyloid (Aβ) plaques in the brain was considered one of the main neuropathological hallmarks of AD. As the loss of synapses always occurs during AD progression, AD has been gradually regarded as a "synaptopathy." The activity-regulated cytoskeleton-associated protein (Arc) was recently identified as a key factor for AD due to its active roles in synaptic plasticity, learning, memory, and Aβ generation. However, there is little evidence to support the association of the *Arc* gene with AD. In this study, a two-stage case-control study of 1471 Han Chinese was conducted to investigate the genetic association between the *Arc* gene and AD. Variant rs10097505 in the 3′ UTR region was significantly associated with AD. The whole exons of the *Arc* gene were also screened in 99 AD patients with a high heritability (familial and/or onset age <55 years old). One missense variant (c.20G>A, p.T7I) was identified in two AD patients but was absent in the controls from the general populations. Both rs10097505 and c.20G>A were predicted to be potentially pathogenic. Further luciferase assay, data mining, and integrative analyses revealed that the AD-risk genotype AA of rs10097505 was associated with an increased *Arc* mRNA expression and an elevated Aβ level. Our results indicated that the *Arc* gene would confer susceptibility to AD in Han Chinese, probably through changing the protein structure or affecting the *Arc*

Rui Bi and Li-Li Kong contributed equally to this work.

For the Alzheimer's Disease Neuroimaging Initiative: data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Electronic supplementary material The online version of this article (doi:10.1007/s12035-017-0397-6) contains supplementary material, which is available to authorized users.

✉ Yong-Gang Yao
yaoyg@mail.kiz.ac.cn

[1] Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China

[2] Institute of Health Sciences, Anhui University, Hefei, Anhui 230601, China

[3] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China

[4] The Mental Health Center and Psychiatric Laboratory, West China Hospital, Sichuan University, Chengdu, Sichuan 610064, China

[5] Division of Mood Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China

[6] CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

expression in brain tissues, which would finally contribute to the pathogenesis and development of AD.

Keywords Alzheimer's disease · *Arc* gene · Genetic susceptibility · Han Chinese

Introduction

Alzheimer's disease (AD) is the most common form of neurodegeneration and dementia, which mainly leads to neuronal loss and severe memory deficit [1]. It is clinically characterized by memory loss and an inability to consolidate new memories [1]. Previous studies on AD pathogenesis have identified the accumulation and deposition of extracellular β-amyloid (Aβ) and neurofibrillary tangles in the brain tissue of the AD patients, which were considered the main neuropathological hallmarks of AD [2, 3]. Hitherto, most of the identified AD-risk genes were involved in the pathways relating to the regulation of Aβ generation and clearance [4, 5]. The abnormal accumulation of Aβ in the brain will lead to its self-assembly into neurotoxic Aβ oligomers, which would result in synaptic dysfunction and cause neurodegeneration [1, 6]. Reduction of Aβ plaques by using the antibody aducanumab had a slowing of clinical decline in patients with AD [7].

As the loss or damage of synapses always occurs during AD progression, AD is gradually regarded as a "synaptopathy" [8]. Activity-regulated cytoskeleton-associated protein (Arc) is a member of the immediate early gene family, which is activated following the induction of the long-term potentiation (LTP) and plays important roles in synaptic plasticity, learning, and memory [9]. Accumulating evidence supported the functional importance of Arc in memory consolidation [8–14]. The maintenance of LTP and consolidation of long-term memory were found to be impaired when Arc expression was inhibited in rat hippocampus, while the short-term memory was not affected by Arc inhibition [14, 15]. Furthermore, the synaptic binding of Aβ was found to be accompanied by induction of *Arc* expression [16], and Arc was involved in the generation of Aβ through interacting with presenilin 1 [17]. Considering the fact that accumulation of Aβ plaques in the brain and memory loss are the most typical features of AD, we hypothesize that genetic variants in the *Arc* gene may confer genetic risk to AD. However, there is little evidence to support the genetic association between the *Arc* gene and AD. A recent study in the Swedish population has identified several AD-risk variants in the *Arc* gene [18], but there is still no study for this gene in Han Chinese.

In this study, we screened variants in the *Arc* gene in 713 AD patients and 758 controls from Han Chinese populations and reanalyzed publically available genetic, expression, and AD endo-phenotype data to further clarify the potential involvement of *Arc* in AD.

Materials and Methods

Subjects

A two-stage case-control study of 1471 subjects of Han Chinese origin was performed in this study. All these subjects have been analyzed for the *APOE* and other AD-related genes in our recent studies [19–22]. In brief, the discovery stage was composed of 332 AD cases and 334 normal individuals without dementia, which was collected by the Mental Health Center of West China Hospital from Southwest China; the validating stage consisted of 381 AD cases and 424 normal controls without dementia that were collected by the Shanghai Mental Health Center and Tongde Hospital of Zhejiang Province from East China. In addition, a total of 52 healthy individuals and 99 AD patients with a high heritability (familial and/or early onset of AD at an age <55 years old; these patients were from different families and were genetically unrelated) were selected to sequence the entire exons of the *Arc* gene. Sample collection was complied with the declaration of Helsinki, with informed consent from all the subjects. The AD patients were diagnosed by at least two clinical psychiatrists. Clinical diagnoses were determined according to the Nation Institute of Neurological and Communicative Disorders Association criteria [23] and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. This study was approved by the institutional review board of the Kunming Institute of Zoology (KIZ), Chinese Academy of Sciences (CAS).

Genotyping and Sequencing of the *Arc* Gene

Genomic DNA was extracted from peripheral blood by using AxyPrep Blood Genomic DNA Miniprep Kit (Axygen, USA). The linkage disequilibrium pattern of the −10~+10-kb region of the *Arc* gene were analyzed by using Haploview software to search for SNPs that are capable of tagging more SNPs in the gene region [24]. Two SNPs (rs13260813, downstream of *Arc*; rs10097505, in the 3′UTR of *Arc*; Supplementary Table S1), with minor allele frequency (MAF) >10% in the Han Chinese in Beijing (CHB) population in the International HapMap Project database (HapMap, http://hapmap.ncbi.nlm.nih.gov/ [25]), were selected for genotyping analysis. A P value <0.025 (0.05/2) was considered statistically significant according to the Bonferroni correction. The SNaPshot assay was performed to genotype rs13260813 and rs10097505, following the procedure described in our previous studies [19, 20]. Briefly, multiplex PCR was performed by using two pairs of primers (Supplementary Table S1). After cleaning up with SAP and *Exo*I (TaKaRa Biotechnology Co. Ltd., Dalian, China), the multiplex PCR products were subjected to a single-base extension reaction with respective extension primers

(Supplementary Table S1). A mixture of 4-μL purified products and 6-μL ddH_2O was analyzed by the 3730xl DNA analyzer (Applied Biosystems) at the Kunming Biodiversity Large-Apparatus Regional Center, KIZ. The GeneMarker software (http://www.softgenetics.com/GeneMarker.html) was used to read the genotyping results.

A total of 99 AD patients with a high heritability (familial and/or early onset age <55 years old) from the sample collection in this study and 52 non-dementia healthy controls were selected to sequence the whole exons and nearby regions of the *Arc* gene by using the paired-end reads on the HiSeq 4000 sequencer (Illumina), following the same procedure in our previous studies [26, 27]. High-quality sequence reads were aligned with the reference genome hg19 by Burrows-Wheeler Aligner (http://bio-bwa.sourceforge.net/) [28] and Samtools (http://samtools.sourceforge.net/) [29, 30]. The Genome Analysis Toolkit (https://www.broadinstitute.org/gatk/) [31] was applied for SNP calling. The identified rare variants were validated by using Sanger sequencing.

Analysis of the *Arc* Genetic Variants

Genetic variants of the *Arc* gene were searched in Chinese samples in 1000 Genomes phase 3 (http://www.1000genomes.org) [32, 33] and in Asian samples in the Exome Aggregation Consortium (http://exac.broadinstitute.org) [34]. The potential function of each variant was estimated by using the Combined Annotation Dependent Depletion (CADD) database [35]. The association between genotype and gene expression was analyzed by using the available brain eQTL database BRAINEAC (http://www.braineac.org/) [36]. The Arc protein structure with or without mutation was predicted by using the RaptorX (http://raptorx.uchicago.edu/) [37].

The Arc protein sequences of 11 vertebrate species were retrieved from GenBank (http://www.ncbi.nlm.nih.gov), including human (*Homo sapiens*, NP_056008.1), chimpanzee (*Pan troglodytes*, XP_016815455.1), crab-eating macaque (*Macaca fascicularis*, XP_005564239), pig-tailed macaque (*Macaca nemestrina*, XP_011750727.1), sheep (*Ovis aries*, XP_004023625.2), cattle (*Bos taurus*, NP_001193336.1), wild yak (*Bos mutus*, XP_005912021.1), domestic guinea pig (*Cavia porcellus*, XP_005005257.1), house mouse (*Mus musculus*, NP_ 001263613.1), Norway rat (*Rattus norvegicus*, NP_062234.1), and chicken (*Gallus gallus*, NP_989763.1). Sequence alignment was performed by using the Clustal W method in MegAlign of DNASTAR Lasergene 7.1 (DNAS Inc., Madison, WI, USA), to show the evolutionary conservation of the Arc variant(s).

Power test was performed by using the Quanto software [38]. The deviation from the Hardy-Weinberg equilibrium (HWE) was calculated by using the PLINK software [39]. The allele and genotype frequencies of all SNPs in the AD patients and controls were analyzed by using two-tailed Fisher's exact test.

Data Mining of *Arc* mRNA Expression

The *Arc* gene expression level in AD was analyzed by data mining the expression data in Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/sites/GDSbrowser). Alterations of *Arc* messenger RNA (mRNA) level in different stages of AD were investigated by reanalyzing the GSE29652 dataset [40]. We further collected the expression data of entorhinal cortex from GSE26927 [41], GSE26972 [42], GSE48350 [43], and GSE5281 [44] and merged these expression data by a cross-platform normalization method Combat in R package inSilicoMerging after log2 transformation [45]. The normalized data could be accessed through our unpublished database available at www.alzdata.org. The difference between the two groups was quantified by using two-tailed Student's *t* test with the GraphPad Prism software (GraphPad Software, La Jolla, CA, USA).

Dual Luciferase Reporter Assay

The 3′UTR of the *Arc* gene (~590 bp) was amplified from AD patients with different alleles and was inserted into the psiCHECK2 luciferase reporter vector within the *Not*I and *Sgf*I restriction sites. Luciferase reporter assay was performed in two different cell lines (human glioma U251 cells and 293T cells) to investigate whether the expression pattern was consistent. Cells were introduced from the Kunming Cell Bank, KIZ, CAS, and were cultured in DMEM medium (293T cells) or 1640 medium (U251 cells), all supplemented with 10% FBS (Invitrogen, USA) at 37 °C in 5% CO_2. The constructs were transfected into cells by using the Lipofectamine 2000 transfection system (Thermo, USA). The transfected 293T cells and U251cells were harvested in 100 μL passive lysis buffer (Promega, USA) at 24 and 48 h, respectively. Luciferase activities were measured by using the Dual Luciferase Reporter Assay Kit (Promega) on an InfiniteM1000 Pro multimode microplate reader (Tecan, Switzerland). All experiments were conducted in triplicates. Two-tailed Student's *t* test was performed to quantify the statistical difference between the two groups by using GraphPad Prism software (GraphPad Software, La Jolla, CA, USA). A P value <0.05 was considered statistically significant.

Analyzing the Effect of rs10097505 on AD Endo-phenotypes

In order to further investigate the role of Arc in AD pathogenesis, we analyzed the genetic, clinical, and biomarker data of 180 patients with AD, 363 patients with mild cognitive impairment (MCI), and 214 controls. Data used in the

preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) [46]. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI was to test whether serial magnetic resonance imaging, positron emission tomography, related biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The association between different genotypes of rs10097505 and AD endo-phenotypes was analyzed by using linear regression analysis with PLINK software [39].

Results

SNP rs10097505 of the *Arc* Gene Was Significantly Associated with AD in Han Chinese

Based on the above-described criteria of SNP selection, we were able to obtain two SNPs for the *Arc* gene and genotyped them in a two-stage case-control study of 1471 Han Chinese subjects (Table 1; Supplementary Fig. S1). Both SNPs were in HWE for the controls (Table 1). The MAF of rs13260813 and rs10097505 were 35% and 38% in our control sample, respectively. The power to detect the odds ratio (OR) value as 1.5 for risk allele would be expected to be 95.5% (for rs13260813) and 94.7% (for rs10097505), with a false-positive rate being controlled as 0.05.

In the stage I study of the cohort from Southwest China, the frequencies of allele and genotype of rs13260813 were not associated with genetic risk of AD, while the allele frequency of SNP rs10097505 showed a marginally significant association with AD ($P = 0.07$; Table 1). We performed a validation study for rs10097505 in an independent Han Chinese population from East China (stage II study). The association of rs10097505 with AD could be validated in this cohort (allele, $P = 0.021$; genotype, $P = 0.012$) and remained to be significant after the Bonferroni correction for multiple testing. When the two cohorts were pooled together to increase the statistical power, we found that the association between rs10097505 and AD was much more significant in the combined population ($P = 0.003$, OR (95% CI) = 1.25 (1.08–1.45); Table 2). Considering the fact that the potential population stratification of Han regional samples [47] might affect this result, the data of CHB ($N = 103$) and Southern Han Chinese ($N = 105$) were retrieved from the 1000 Genomes phase 3 [32, 33] and combined with the Han Chinese controls in this study. We could observe a robust association of rs10097505 with AD risk in the pooled Han Chinese (Table 2). Accordantly, prediction of potential pathogenicity by using the CADD database [35] showed that rs10097505 had a high possibility to be pathogenic, while rs13260813 was probably to be benign (Table 1).

Exon Sequencing of the *Arc* Gene Identified One Potentially Pathogenic Variant

In the 99 AD patients that were selected for sequencing the exon regions of the *Arc* gene, a total of eight variants were identified (Table 3). Among these variants, only one rare missense variant (c.20G>A, p.T7I) was identified in 2 of the 99 AD patients (Fig. 1a). This variant was absent in the 353 controls from the Chinese general populations and was significantly associated with AD risk ($P = 0.04$; Table 3). We searched this variant in an unpublished whole exome dataset and identified one carrier of c.20G>A in a total of 2473 non-dementia Han Chinese individuals from the general populations (Dr. Jirong Long, personal communication). Fisher's exact test indicated that allele frequency of this variant was significantly different in AD patients versus the non-dementia Han Chinese population ($P = 0.004$).

Evolutionary conservation analysis of 11 vertebrate species showed that this non-synonymous variant occurred in a highly conserved residue of the Arc protein (Fig. 1b). Structure

Table 1 Allele and genotype frequencies of *Arc* SNPs in 332 AD cases and 334 controls from Southwest China and 381 AD cases and 424 controls from East China

SNP ID	Allele		*P* value[a]	Genotype		*P* value[a]	HWE	PHRED[b]
	Case	Control		Case	Control			
Stage I	332 AD cases versus 334 controls							
rs13260813	230/434	225/443	0.713	43/144/145	33/159/142	0.353	0.271	1.294
rs10097505	281/381	251/417	0.070	63/155/113	45/161/128	0.133	0.643	11.44
Stage II	381 AD cases versus 424 controls							
rs10097505	303/459	290/558	0.021	66/171/144	43/204/177	0.012	0.195	

[a] *P* value and Fisher's exact test

[b] The PHRED-like scaled CADD score: a score greater than ten indicates that the variant belongs to the top 10% deleterious substitutions in human genome [35]

Table 2 Allele and genotype frequencies of rs10097505 in 713 AD cases and 966 controls from Han Chinese population

Allele/genotype	AD[a]	CN[b]	*P* value[c]	OR (95%CI)	CN + CHB + CHS[d]	*P* value[c]	OR (95%CI)
AA	129	88	0.0005	1.69 (1.26–2.26)	124	0.003	1.50 (1.15–1.97)
AG	326	365	0.37	0.91 (0.74–1.12)	471	0.24	0.89 (0.73–1.08)
GG	257	305	0.11	0.84 (0.68–1.04)	371	0.36	0.91 (0.74–1.11)
A	584	541	0.003	1.25 (1.08–1.45)	719	0.03	1.17 (1.02–1.35)
G	840	975			1213		

[a] AD, 713 Han Chinese patients with AD in this study

[b] CN, 758 Han Chinese normal individuals in this study

[c] *P* value and Fisher's exact test

[d] CN + CHB + CHS, 966 normal controls (758 normal Han Chinese (CN) in this study, 103 Han Chinese from Beijing (CHB), and 105 Han Chinese from South China (CHS) in the 1000 Genomes phase 3 (http://www.1000genomes.org) [32, 33]; we assumed the CHB and CHS samples to be normal)

prediction of the domain containing the 1–140 residues by using Raptor X [37] revealed that variant p.T7I could change the tertiary structure of the domain (Fig. 1c). Pathogenicity assessment of c.20G>A by using the CADD database [35] indicated that this variant belongs to the top 10% most deleterious mutations in human genome (Table 3).

The *Arc* mRNA Expression Level Was Significantly Decreased in AD

Reanalysis of the gene expression profile of GSE29652 [40], which includes expression data of temporal cortex astrocytes representing different Braak stages of AD, showed that the *Arc* mRNA level was obviously downregulated during the development of AD (Fig. 2a). We further normalized and combined different expression datasets (GSE26927 [41], GSE26972 [42], GSE48350 [43], and GSE5281 [44]) of entorhinal cortex, which was the most affected region in AD. We found that the *Arc* mRNA level was remarkably downregulated in entorhinal cortex of AD patients when compared with the normal controls. All these data suggested that altered *Arc* expression in the brain might be involved in the pathogenesis of AD.

rs10097505 Was Associated with the *Arc* Expression Level

Based on the finding that the *Arc* mRNA level was significantly decreased in AD brain, we investigated whether the AD-risk allele/genotype of the *Arc* gene would account for the altered *Arc* expression. Through data mining of the brain

Table 3 Variants in the exons of the *Arc* gene in 99 AD patients and 353 controls from the Chinese general populations

Location	SNP ID	Function	AD[a]	CN[b]	*P* value[c]	Chinese CN[d]	*P* value[c]	ExAC [e]	*P* value[c]	PHRED[f]
c.*1542C>T	rs188213819	UTR-3	3/193	1/103	1.000	7/699	0.460	–[g]	–[g]	12.47
c.858C>G	rs28686812	Synonymous	73/125	47/57	0.175	245/461	0.613	3121/4609	0.341	0.031
c.747G>A	rs140337555	Synonymous	3/195	0/104	0.554	0/706	0.010	42/7822	0.098	14.64
c.489C>T	rs2234911	Synonymous	114/74	57/47	0.385	460/246	0.266	870/480	0.331	0.234
c.20G>A	–[g]	p.T7I	2/176	0/104	0.533	0/706	0.040	–[g]	–[g]	16.19
c.-17G>A	rs201628756	UTR-5	3/167	0/104	0.291	5/701	0.189	20/1348	0.735	9.295
c.-46T>C	rs28403912	UTR-5	129/33	79/11	0.120	590/102	0.093	271/81	0.568	6.189
c.-49C>G	rs28641083	UTR-5	128/32	79/11	0.162	590/102	0.117	213/61	0.629	5.285

[a] AD, 99 Han Chinese AD patients with a high heritability (familial and/or onset age <55 years old)

[b] CN, 52 non-dementia healthy controls

[c] *P* value and Fisher's exact test

[d] Chinese CN, 353 controls from Chinese general populations (52 non-dementia healthy controls in this study; 103 Han Chinese from Beijing, 105 Han Chinese from South China, and 93 Dai from Xishuangbanna, China in the 1000 Genomes phase 3 (http://www.1000genomes.org) [32, 33])

[e] Four thousand three hundred twenty-seven East Asians from the Exome Aggregation Consortium (ExAC: http://exac.broadinstitute.org) [34]

[f] The PHRED-like scaled CADD score: a score greater than ten indicates that the variant belongs to the top 10% most deleterious substitutions in human genome [35]

[g] No data available

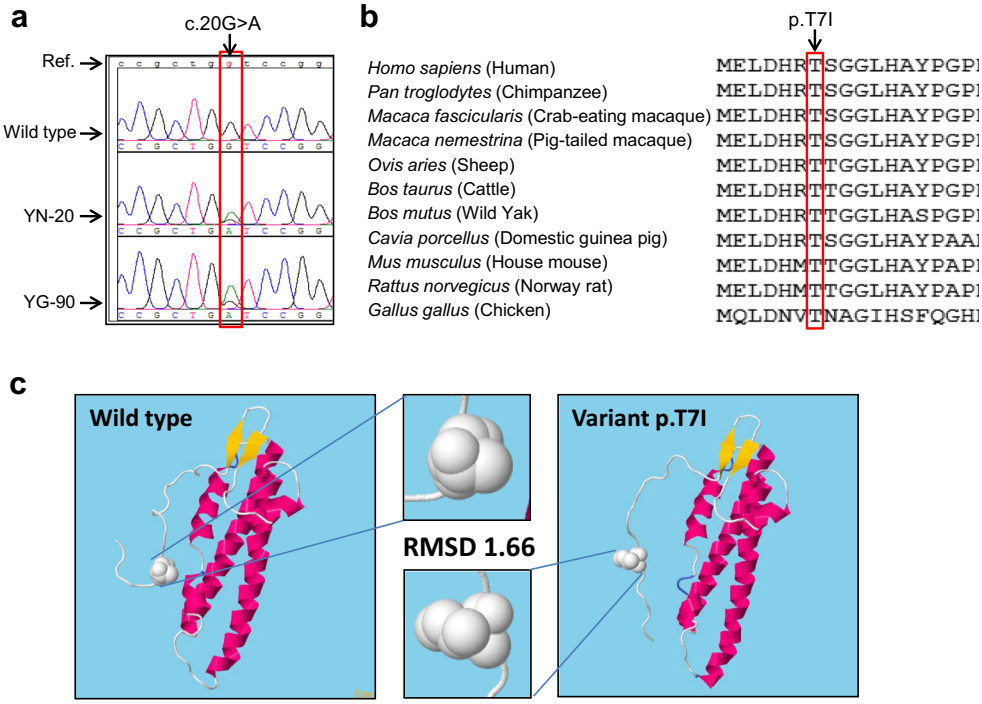


Fig. 1 Evolutionary conservation analysis and structure modeling of the Arc protein. **a** Sequencing electrophoregrams of variant c.20G>A (p.T7I), which was identified in two AD patients (YN-20 and YG-90) but not in healthy controls (wild type). **b** Evolutionary analysis of the 7th residue in the Arc protein. The Arc protein sequences of 11 vertebrate species were retrieved from GenBank. **c** Structure modeling of the 1–140 residues of wild-type Arc and variant p.T7I. The modeling was performed by using RaptorX [37]. *RMSD* - root mean square deviation, which means the average distance between the atoms of the wild-type and mutant Arc proteins

eQTL data from BRAINEAC database (http://www.braineac.org/) [36], we found that the AA genotype of the AD-risk SNP rs10097505 was significantly associated with increased hippocampal *Arc* mRNA expression level ($P = 0.033$; Fig. 3a).

In order to validate the expression result, we performed luciferase reporter assay with the 3′UTR sequence containing different alleles of rs10097505 in U251 and 293T cells. Consistent with the expression pattern, allele A of rs10097505 had a higher luciferase activity than allele G in both U251 and 293T cells (Fig. 3b). Collectively, these results suggested that rs10097505 may be related to AD risk through affecting the *Arc* gene expression.

rs10097505 Was Associated with Cognitive Ability and AD Biomarkers

Data mining of the ADNI database [46] was performed to investigate whether the *Arc* SNP rs10097505 is associated with altered cognitive ability and AD biomarkers. As shown in Fig. 4, the AD-risk genotype AA of rs10097505 was linked

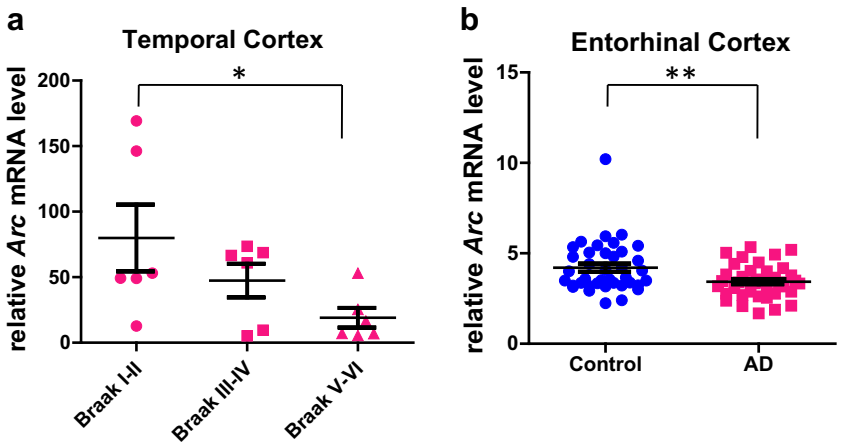


Fig. 2 The *Arc* mRNA level was significantly decreased in AD brain tissues. **a** The mRNA level of *Arc* was significantly decreased during AD progression. The *Arc* expression profile in temporal cortex astrocytes of patients at different AD Braak stages was retrieved from GSE29652 dataset [40]. **b** The *Arc* mRNA level was significantly decreased in entorhinal cortex of AD patients. Four datasets (GSE26927 [41], GSE26972 [42], GSE48350 [43], and GSE5281 [44]) in the GEO database were retrieved and normalized. The difference between the two groups was quantified by using two-tailed Student's *t* test. *P value <0.05; **P value <0.01

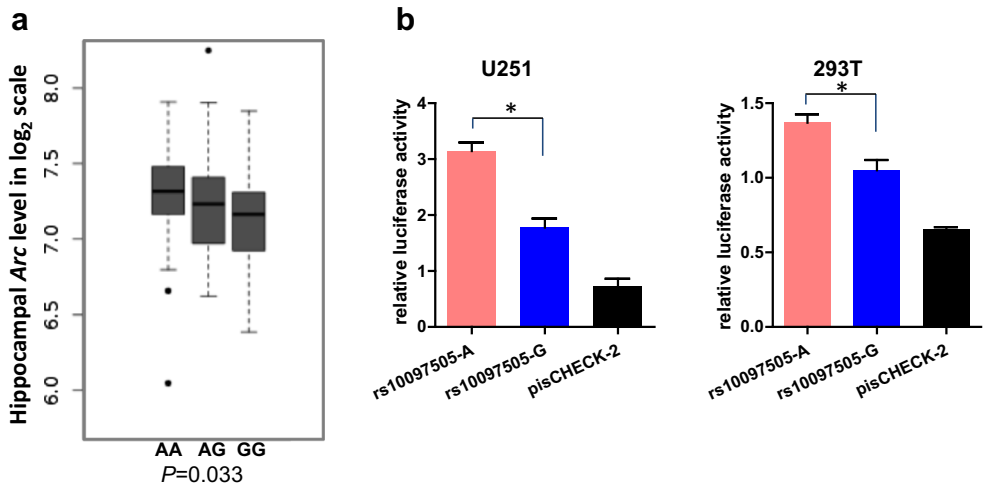


Fig. 3 Effects of rs10097505 allele and genotype on *Arc* expression. **a** The AA genotype of rs10097505 was significantly associated with increased hippocampal *Arc* expression level. The data of rs10097505 was retrieved from the brain eQTL database BRAINEAC [36]. **b** Alleles of rs10097505 had different regulatory effects. The 3′UTR sequences with different allele of rs10097505 were cloned in to the psiCHECK2 luciferase reporter vector and were transfected into U251 and 293T cells. Cells were harvested at 24 h (293T cell) and 48 h (U251 cell) after transfection and were analyzed for luciferase activities. *P value <0.05, two-tailed Student's t test

to a relatively lower score of Minimum Mental State Examination (MMSE) in the control group, while in the MCI and AD groups, no difference among different genotypes was observed. The total Aβ and Aβ1–42 levels were higher in the cerebrospinal fluid (CSF) of normal subjects with AA genotype, while the levels were similar among different genotypes in the MCI and AD groups, suggesting that carriers with AA genotype of rs10097505 had a tendency to generate more Aβ, which would be distributed in the CSF of the normal controls but would be deposited in the cortex in the MCI and AD patients (Fig. 4).

Discussion

The Arc was regarded as a key regulator of synaptic plasticity [48]. Many studies have revealed the important roles of Arc in neuropsychiatric diseases, such as AD [8, 17], schizophrenia [49, 50], and depression [51], whereas only a few studies had identified a genetic association between the *Arc* gene and these diseases [18, 49, 50]. A recent study in Norwegian populations have found that common variants in the *Arc* gene were not associated with cognitive ability [52], and genome-wide association study (GWAS) in European and American populations also identified no association between the *Arc* gene and AD [53] (Supplementary Table S2). However, these previous negative results were all based on western European populations or populations of European origin [52, 53] that had a significantly different genetic structure compared with Asian population. Therefore, whether the *Arc* gene was genetically associated with AD in Han Chinese population is still unknown. In this study, we found that the AA genotype of a common SNP in the *Arc* 3′UTR (rs10097505) was significantly associated with the genetic risk of AD in Han Chinese population (P = 0.003, OR (95% CI) = 1.25 (1.08–1.45); Tables 1 and 2), and the effect direction of rs10097505-AA was consistent with a previous study in Swedish population (P = 0.09, OR (95% CI) = 1.23 (0.97–1.54)) [18]. Further screening of rare variants in the exon region of the *Arc* gene in AD patients with a high heritability risk identified a missense mutation (c.20G>A, p.T7I) showing a higher frequency in AD patients compared with the controls. This variant c.20G>A (p.T7I) changed a highly conserved residue and would affect the predicted protein structure of Arc (Fig. 1). Web server-based prediction by using the CADD database [35] revealed that variants rs10097505 and c.20G>A in the *Arc* gene had a high possibility to be pathogenic (Tables 1 and 3). All these results suggested that both common and rare variants in the *Arc* gene might be significantly associated with the genetic risk of AD in Han Chinese.

To further characterize the underpinning of association between rs10097505 and AD, we performed data-mining and integrative analyses of the available datasets [40–44] and performed a luciferase assay. The *Arc* mRNA level was found to be significantly reduced in AD (Fig. 2), but the AD-risk genotype AA of rs10097505 was significantly associated with increased hippocampal *Arc* mRNA level (Fig. 3). The effect of allele A of rs10097505 on *Arc* expression could be demonstrated by the higher luciferase activity compared with that of allele G (Fig. 3). These results provided a seeming paradox: the *Arc* mRNA level was significantly decreased in AD patients whereas the AD-risk allele in the *Arc* gene significantly increased *Arc* expression level. It is of note that previous studies concerning the association between *Arc* mRNA expression

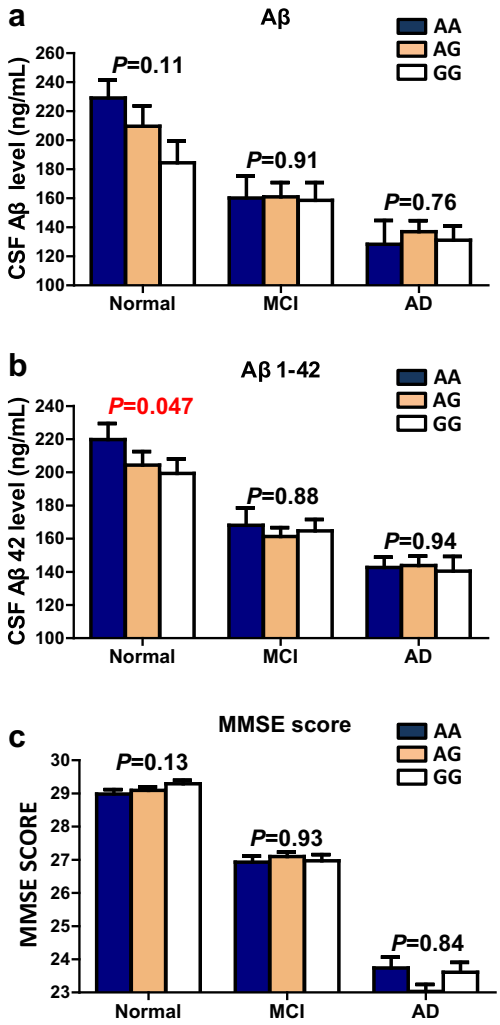


Fig. 4 SNP rs10097505 was associated with AD endo-phenotypes. The ADNI data set [46], which includes genetic, clinical, and biomarker data of 180 patients with AD, 363 patients with mild cognitive impairment (MCI), and 214 normal controls (normal), was reanalyzed. The associations between rs10097505 and the total Aβ (**a**) or Aβ1–42 (**b**) levels in the cerebrospinal fluid (CSF), and between rs10097505 and the score of Minimum Mental State Examination (MMSE) **(c)**, were analyzed by using linear regression analysis with PLINK software [39]. A *P* value <0.05 was considered statistically significant and was marked on the figure

level and AD had not reached a consistent conclusion: both increased [8, 17] and reduced levels [54] of *Arc* had been reported in AD patients or in animal models. A most recent study reported that plasma level of Arc protein was significantly higher in children with autism compared with the controls, whereas the Arc protein level was negatively correlated with the severity of autism [55]. What is the underlying reason for this discrepancy? Previous studies have identified a feedback regulation mechanism of Arc on AMPAR endocytosis, which was induced upon neuronal activation and would dampen the over excitation of neuronal network [8, 13, 56]. The homeostatic Arc expression was indispensable for maintaining the negative feedback mechanism of neuronal network and for memory consolidation [8, 12, 13, 56]. The Arc knockout mice were unable to form long-term memory [12], while Arc overexpression could increase the activity-dependent generation of Aβ [17]. Therefore, combining with the results of these previous studies and the current finding, we proposed that the disruption of homeostatic *Arc* expression, either upregulation or downregulation, would contribute to the pathogenesis of AD. The upregulation effect of the AD-related risk allele of rs10097505 on *Arc* expression might contribute to the increased generation of Aβ and conferred the genetic susceptibility to AD.

The association between rs10097505 and AD could be further demonstrated by analyzing the association of this SNP with endo-phenotypes of the AD patients and normal controls from the ADNI data set [46]. Subjects with the risk genotype AA of rs10097505 showed a relatively lower MMSE score in normal subjects, but this effect was not evident in the MCI and AD groups that already had an impaired cognitive ability (Fig. 4). Previous studies showed that the reduced CSF Aβ1–42 level could be regarded as a biomarker for AD, which would reflect the deposition of Aβ plaques in the cortex of AD patients [57–59]. As there might be no Aβ deposition in the cortex of the normal controls, the CSF Aβ level could be considered an indicator for total Aβ production in the normal controls. On this point, the regulatory effect of risk allele of rs10097505 on the CSF levels of Aβ and Aβ1–42 could be best demonstrated in normal subjects. Concordant with speculation, we observed a higher level of CSF Aβ in normal subjects with the risk allele A of rs10097505, supporting the upregulation effect of this allele on *Arc* expression (Fig. 4) and the fact that overexpression of Arc increased Aβ production [17].

In conclusion, we found that variant rs10097505 in the *Arc* gene could confer genetic susceptibility to AD in Han Chinese, probably through affecting *Arc* expression level in the brain tissues and further influencing the Aβ generation. Rare missense variant (c.20G>A, p.T7I) in the *Arc* gene might also play a role in the development of AD, although the exact mechanism awaits further study. Future independent validation studies and essential functional assays are necessary to solidify the current conclusion and to characterize the putative role of the *Arc* gene in AD.

Acknowledgements We thank all participants in this study. We thank Miss Hui-Zhen Wang and Qiu-Xiang Hu for technical assistance and Dr. Jirong Long for searching the Arc variant p.T7I in her unpublished dataset. This study was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020003) and the Bureau of Frontier Sciences and Education, Chinese Academy of Sciences (QYZDJ-SSW-SMC005). Data collection and sharing for this project was funded by the ADNI (National Institutes of Health grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and

Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

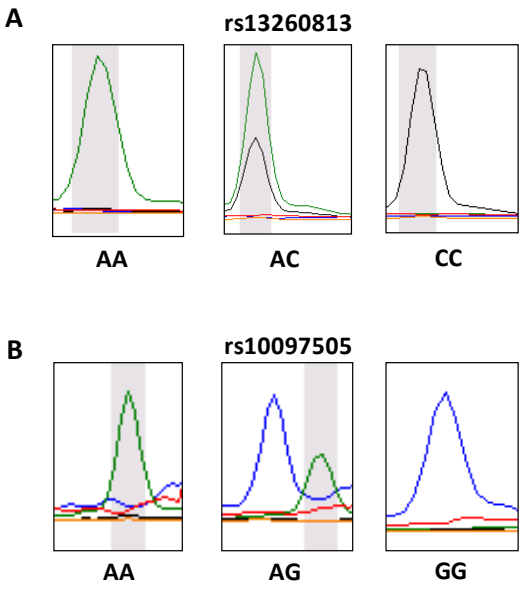
Compliance with Ethical Standards

Conflict in Interests The authors declare that they have no competing interests..

References

1. Querfurth HW, LaFerla FM (2010) Alzheimer's disease. N Engl J Med 362(4):329–344
2. Hardy JA, Higgins GA (1992) Alzheimer's disease: the amyloid cascade hypothesis. Science 256(5054):184–185
3. Hardy J, Selkoe DJ (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 297(5580):353–356
4. Bertram L, Lill CM, Tanzi RE (2010) The genetics of Alzheimer disease: back to the future. Neuron 68(2):270–281
5. Karch CM, Goate AM (2015) Alzheimer's disease risk genes and mechanisms of disease pathogenesis. Biol Psychiatry 77(1):43–51
6. De Strooper B, Karran E (2016) The cellular phase of Alzheimer's disease. Cell 164(4):603–615
7. Sevigny J, Chiao P, Bussiere T, Weinreb PH, Williams L, Maier M et al (2016) The antibody aducanumab reduces Abeta plaques in Alzheimer's disease. Nature 537(7618):50–56
8. Kerrigan TL, Randall AD (2013) A new player in the "synaptopathy" of Alzheimer's disease-arc/arg 3.1. Front Neurol 4:9
9. Bramham CR, Worley PF, Moore MJ, Guzowski JF (2008) The immediate early gene arc/arg3.1: regulation, mechanisms, and function. J Neurosci 28(46):11760–11767
10. Lyford GL, Yamagata K, Kaufmann WE, Barnes CA, Sanders LK, Copeland NG et al (1995) Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. Neuron 14(2):433–445
11. Ying SW, Futter M, Rosenblum K, Webber MJ, Hunt SP, Bliss TV et al (2002) Brain-derived neurotrophic factor induces long-term potentiation in intact adult hippocampus: requirement for ERK activation coupled to CREB and upregulation of arc synthesis. J Neurosci 22(5):1532–1540
12. Plath N, Ohana O, Dammermann B, Errington ML, Schmitz D, Gross C et al (2006) Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron 52(3):437–444
13. Waung MW, Pfeiffer BE, Nosyreva ED, Ronesi JA, Huber KM (2008) Rapid translation of arc/Arg3.1 selectively mediates mGluR-dependent LTD through persistent increases in AMPAR endocytosis rate. Neuron 59(1):84–97
14. Bramham CR, Alme MN, Bittins M, Kuipers SD, Nair RR, Pai B et al (2010) The Arc of synaptic memory. Exp Brain Res 200(2):125–140
15. Guzowski JF, Lyford GL, Stevenson GD, Houston FP, McGaugh JL, Worley PF et al (2000) Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. J Neurosci 20(11):3993–4001
16. Lacor PN, Buniel MC, Chang L, Fernandez SJ, Gong Y, Viola KL et al (2004) Synaptic targeting by Alzheimer's-related amyloid beta oligomers. J Neurosci 24(45):10191–10200
17. Wu J, Petralia RS, Kurushima H, Patel H, Jung MY, Volk L et al (2011) Arc/Arg3.1 regulates an endosomal pathway essential for activity-dependent beta-amyloid generation. Cell 147(3):615–628
18. Landgren S, von Otter M, Palmer MS, Zetterstrom C, Nilsson S, Skoog I et al (2012) A novel ARC gene polymorphism is associated with reduced risk of Alzheimer's disease. J Neural Transm (Vienna) 119(7):833–842
19. Bi R, Zhao L, Zhang C, Lu W, Feng JQ, Wang Y et al (2014) No association of the LRRK2 genetic variants with Alzheimer's disease in Han Chinese individuals. Neurobiol Aging 35(2):444. e5–9
20. Wang HZ, Bi R, Hu QX, Xiang Q, Zhang C, Zhang DF et al (2016) Validating GWAS-identified risk loci for Alzheimer's disease in Han Chinese populations. Mol Neurobiol 53(1):379–390
21. Wang HZ, Bi R, Zhang DF, Li GD, Ma XH, Fang Y et al (2016) Neprilysin confers genetic susceptibility to Alzheimer's disease in Han Chinese. Mol Neurobiol 53(7):4883–4892
22. Zhang DF, Li J, Wu H, Cui Y, Bi R, Zhou HJ et al (2016) CFH variants affect structural and functional brain changes and genetic risk of Alzheimer's disease. Neuropsychopharmacology 41(4):1034–1045
23. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (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
24. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263–265
25. International HapMap Consortium (2003) The International HapMap Project. Nature 426(6968):789–796
26. Zhang DF, Fan Y, Wang D, Bi R, Zhang C, Fang Y et al (2015) PLD3 in Alzheimer's disease: a modest effect as revealed by updated association and expression analyses. Mol Neurobiol 53(6):4034–4045
27. Wang D, Li GD, Zhang DF, Xu L, Li XA, Yu XF et al (2016) Genetic variants of the MAVS, MITA and MFN2 genes are not associated with leprosy in Han Chinese from Southwest China. Infect Genet Evol 45:105–110
28. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760
29. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27(21):2987–2993
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079

31. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43(5):491–498
32. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO et al (2015) A global reference for human genetic variation. Nature 526(7571):68–74
33. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE et al (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491(7422):56–65
34. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536(7616):285–291
35. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46(3):310–315
36. Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R et al (2014) Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci 17(10):1418–1428
37. Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H et al (2012) Template-based protein structure modeling using the RaptorX web server. Nat Protoc 7(8):1511–1522
38. Gauderman WJ (2002) Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 21(1): 35–50
39. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575
40. Simpson JE, Ince PG, Shaw PJ, Heath PR, Raman R, Garwood CJ et al (2011) Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype. Neurobiol Aging 32(10):1795–1807
41. Durrenberger PF, Fernando FS, Magliozzi R, Kashefi SN, Bonnert TP, Ferrer I et al (2012) Selection of novel reference genes for use in the human central nervous system: a BrainNet Europe study. Acta Neuropathol 124(6):893–903
42. Berson A, Barbash S, Shaltiel G, Goll Y, Hanin G, Greenberg DS et al (2012) Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. EMBO Mol Med 4(8):730–742
43. Berchtold NC, Cribbs DH, Coleman PD, Rogers J, Head E, Kim R et al (2008) Gene expression changes in the course of normal brain aging are sexually dimorphic. Proc Natl Acad Sci U S A 105(40): 15605–15610
44. Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Walker DG et al (2007) Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol Genomics 28(3):311–322
45. Taminau J, Meganck S, Lazar C, Steenhoff D, Coletta A, Molter C et al (2012) Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. BMC Bioinformatics 13:335
46. Weiner MW, Aisen PS, Jack CR Jr, Jagust WJ, Trojanowski JQ, Shaw L et al (2010) The Alzheimer's disease neuroimaging initiative: progress report and future plans. Alzheimers Dement 6(3): 202–211. e7
47. Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70(3):635–651
48. Shepherd JD, Bear MF (2011) New views of Arc, a master regulator of synaptic plasticity. Nat Neurosci 14(3):279–284
49. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P et al (2014) De novo mutations in schizophrenia implicate synaptic networks. Nature 506(7487):179–184
50. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P et al (2014) A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506(7487):185–190
51. Elizalde N, Pastor PM, Garcia-Garcia AL, Serres F, Venzala E, Huarte J et al (2010) Regulation of markers of synaptic function in mouse models of depression: chronic mild stress and decreased expression of VGLUT1. J Neurochem 114(5):1302–1314
52. Myrum C, Giddaluru S, Jacobsen K, Espeseth T, Nyberg L, Lundervold AJ et al (2015) Common variants in the ARC gene are not associated with cognitive abilities. Brain Behav 5(10): e00376
53. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez 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
54. Wegenast-Braun BM, Fulgencio Maisch A, Eicke D, Radde R, Herzig MC, Staufenbiel M et al (2009) Independent effects of intra- and extracellular Abeta on learning-related gene expression. Am J Pathol 175(1):271–282
55. Alhowikan AM (2016) Activity-regulated cytoskeleton-associated protein dysfunction may contribute to memory disorder and earlier detection of autism spectrum disorders. Med Princ Pract 25(4):350–354
56. Shepherd JD, Rumbaugh G, Wu J, Chowdhury S, Plath N, Kuhl D et al (2006) Arc/Arg3.1 mediates homeostatic synaptic scaling of AMPA receptors. Neuron 52(3):475–484
57. Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L (2006) Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. Lancet Neurol 5(3):228–234
58. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC et al (2009) Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Ann Neurol 65(4):403–413
59. Blennow K, Dubois B, Fagan AM, Lewczuk P, de Leon MJ, Hampel H (2015) Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. Alzheimers Dement 11(1):58–69



Supplementary Figure S1. SNaPshot profiles of different genotypes of rs13260813 (A) and rs10097505 (B).

Supplementary Materials

Supplementary Table S1. Primer information for genotyping SNPs in the *Arc* gene by using the SNaPshot assay

SNP ID	Allele	Position	Location	Primer(5'-3')[a]
rs13260813	A/C	143689397	Downstream	F: GACCTGTGTCTGCAGCTCC
				R: TGTGCACCTTACCCCAGA
				E:TCGATCGATCGATCGCCAGCCCAAGCACACAGAAGAAACA
rs10097505	G/A	143691186	3'-UTR	F: TTGTAGCAGAATGAGGAAGCT
				R: ATCACCCTCCCTCCGTCT
				E:ATCGATCGATCGATCGATCGGGGTGTTCTGTGATGGCCTGAGAGT

[a]F: forward primer; R: reverse primer; E: extension primer.

Supplementary Table S2. Common variants in the *Arc* gene (spanning the -10 kb ~ +10 kb region of the *Arc* gene) from the IGAP data set.

Chr	Location	SNP	Allele	Beta	SE	*P*-value
8	143691393	rs7002963	T/C	-0.030	0.018	0.092
8	143691545	rs60800647	G/C	0.055	0.027	0.039
8	143691838	rs7465272	A/T	-0.032	0.019	0.096
8	143692087	rs13252981	A/C	0.001	0.016	0.951
8	143692146	rs145258183	T/C	-0.046	0.070	0.513
8	143692326	rs36044474	C/G	0.003	0.016	0.876
8	143692395	rs13260813	C/A	-0.004	0.029	0.891
8	143692781	rs77080405	T/C	0.037	0.019	0.057
8	143693384	rs76315763	C/G	0.041	0.019	0.033
8	143693411	rs35900184	T/C	-0.029	0.019	0.122
8	143694184	rs10097505	A/G	-0.004	0.016	0.817
8	143694775	rs28686812	C/G	0.003	0.016	0.845
8	143695144	rs2234911	A/G	-0.006	0.016	0.698
8	143695678	rs28403912	A/G	-0.030	0.019	0.111
8	143695681	rs28641083	G/C	-0.031	0.019	0.109
8	143695772	rs72614036	T/C	-0.003	0.030	0.933
8	143695825	rs28420666	C/G	-0.029	0.019	0.141
8	143696041	rs28427138	A/G	0.004	0.016	0.820
8	143696080	rs28734952	A/G	0.004	0.016	0.829
8	143696234	rs28489307	C/G	0.004	0.016	0.787
8	143696832	rs28625055	A/G	0.041	0.020	0.036
8	143697426	rs28600329	C/A	0.005	0.016	0.773

Note - IGAP data source:
http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php [1]

Reference

1. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, Destefano AL, et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45 (12):1452-1458.