Variation at APOE and STH loci and Alzheimer's disease.

Lingjun Zuo
Christopher van Dyck
Xingguang Luo
Henry R. Kranzler
Bao-Zhu Yang

See next page for additional authors

Follow this and additional works at: http://digitalcommons.uconn.edu/som_articles

Recommended Citation
http://digitalcommons.uconn.edu/som_articles/13

This Article is brought to you for free and open access by the School of Medicine at DigitalCommons@UConn. It has been accepted for inclusion in SoM Articles by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.
Variation at APOE and STH loci and Alzheimer's disease


Lingjun Zuo (Lingjun.Zuo@yale.edu)
Christopher H van Dyck (christopher.vandyck@yale.edu)
Xingguang Luo (Xingguang.Luo@yale.edu)
Henry R Kranzler (kranzler@psychiatry.uchc.edu)
Bao-zhu Yang (Bao-zhu.yang@yale.edu)
Joel Gelernter (Joel.Gelernter@yale.edu)

**ISSN** 1744-9081
**Article type** Research
**Submission date** 15 March 2006
**Acceptance date** 7 April 2006
**Publication date** 7 April 2006
**Article URL** [http://www.behavioralandbrainfunctions.com/content/2/1/13](http://www.behavioralandbrainfunctions.com/content/2/1/13)

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *Behavioral and Brain Functions* are listed in PubMed and archived at PubMed Central. For information about publishing your research in *Behavioral and Brain Functions* or any BioMed Central journal, go to [http://www.behavioralandbrainfunctions.com/info/instructions/](http://www.behavioralandbrainfunctions.com/info/instructions/)

For information about other BioMed Central publications go to [http://www.biomedcentral.com/](http://www.biomedcentral.com/)

© 2006 Zuo et al., licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Variation at APOE and STH loci and Alzheimer’s disease

Lingjun Zuo,1,2 Christopher H. van Dyck,3 Xingguang Luo,1,2 Henry R. Kranzler,4
Bao-zhu Yang,1,2 Joel Gelernter*1,2

1. Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
2. VA Connecticut Healthcare System, West Haven Campus, CT, USA
3. Alzheimer’s Disease Research Unit and Cognitive Disorders Clinic, Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
4. Department of Psychiatry, University of Connecticut School of Medicine, Farmington, CT, USA.

Running title: APOE and STH loci and Alzheimer’s disease

Key words: Alzheimer’s disease, APOE gene, STH gene, dose effect, positive predictive value, positive likelihood ratio

Lingjun Zuo: Lingjun.Zuo@yale.edu
Christopher H. van Dyck: christopher.vandyck@yale.edu
Xingguang Luo: Xingguang.Luo@yale.edu
Henry R. Kranzler: kranzler@psychiatry.uchc.edu
Bao-zhu Yang: Bao-zhu.Yang@yale.edu
Joel Gelernter: Joel.Gelernter@yale.edu

*Corresponding author and reprints: Joel Gelernter, MD; Yale University School of Medicine; VA Psychiatry 116A2; 950 Campbell Avenue; West Haven, CT 06516. Tel: 203-932-5711 ext 3590; Fax: 203-937-3897.
Abstract

Background: The apolipoprotein E (APOE) and tau proteins play important roles in the pathological development of Alzheimer’s disease (AD). Many studies have shown an association between the APOE gene and AD. Association between AD and the newly discovered saitothin (STH) gene, nested within the intron of the tau gene, has been reported. The present study aimed to elucidate the association between APOE and AD, and between STH and AD in our sample.

Methods: The functional polymorphisms, rs429358 and rs7412, in the APOE gene (which together define the ε2, ε3, and ε4 alleles), and the Q7R SNP in the STH gene, were genotyped in 369 patients with AD and 289 healthy European-Americans. The associations between these two genes and AD were analyzed in a case-control design.

Results: Consistent with previously reported results, the frequencies of the APOE ε4 allele, ε4/ε4 genotype and ε3/ε4 genotype were significantly higher in AD cases than controls; the ε4/ε4 genotype frequency was significantly higher in early-onset AD (EOAD) than late-onset AD (LOAD); the frequencies of the ε2 allele, ε3 allele, ε3/ε3 genotype and ε2/ε3 genotype were significantly lower in AD cases than controls. Positive likelihood ratios (LRs+) of APOE alleles and genotypes increased in a linear trend with the number of ε4 alleles and decreased in a linear trend with the number of ε2 or ε3 alleles. There was no significant difference in the STH allele and genotype frequency distributions between AD cases and controls.

Conclusions: This study confirmed that the ε4 allele is a dose-response risk factor for AD and the ε4/ε4 genotype was associated with a significantly earlier age of onset. Moreover, we found that the ε2 allele was a dose-response protective factor for AD and the ε3 allele exerted a weaker dose-response protective effect for risk of AD compared with ε2. In a clinical setting, APOE genotyping could offer additional biological evidence of whether a subject may develop AD, but it is not robust enough to serve as an independent screening or predictive test in the diagnosis of AD. STH variation was not significantly associated with AD in our sample.

Key Words: Alzheimer’s disease, APOE gene, STH gene, dose effect, positive predictive value, positive likelihood ratio
Background

Alzheimer’s disease (AD) is the most common cause of dementia. It is a primary neurodegenerative cerebral disease in the elderly, characterized by two major histopathologic changes in the brain, i.e., extracellular amyloid plaques and intracellular neurofibrillary tangles [1,2].

Apolipoprotein E (APOE) is one of the major cholesterol transport proteins. It exists in three major isoforms, APOE2, APOE3 and APOE4. The three APOE isoforms differ in the 112th and 158th residues of their primary structures (Figure 1); these differences are classified as SNPs rs429358 and rs7412, respectively. The APOE3 protein has higher receptor affinity than the variant types APOE2 and APOE4. Substitution of the basic amino acid Arg158 in APOE3 by the neutral amino acid Cys158 in APOE2 results in the receptor affinity of APOE2 being reduced to 2% of that of APOE3 [3]. In the central nervous system, APOE mediates the uptake and redistribution of cholesterol, and different APOE isoforms modify cholesterol homeostasis by preferentially associating with specific lipoprotein particles [4]. The role of APOE in modifying cholesterol homeostasis in the brain may contribute to the relationship between APOE and AD. Furthermore, APOE exists inside the amyloid plaque, where it can bind to β-amyloid (Aβ), which is a major component of the plaque [5]. Studies have shown that APOE interacts with Aβ to form a stable complex, altering the deposition of Aβ and affecting Aβ-induced neurotoxicity [6].

Moreover, APOE may be involved in Alzheimer’s disease through a tau pathway. Studies have indicated that tau plays an important role in the physiopathology of Alzheimer’s disease and that an extended haplotype (H1), covering the entire tau gene, including a 238 bp insertion in intron 9, is associated with AD [2, 7-12], although these observations have not always been confirmed by other studies. APOE2 and APOE3 can bind to tau and prevent tau from being hyperphosphorylated. Although APOE4 also binds to tau, it cannot prevent tau from hyperphosphorylation, but destabilizes tau. The hyperphosphorylated tau can decrease tau’s affinity for microtubules and
severely disrupt microtubule stability, which has been postulated to be an important step in the formation of the paired helical filament (PHF) involved in neuronal degeneration. This may be part of the mechanism of APOE’s important role in the etiology of AD.

APOE2, APOE3 and APOE4 are encoded by the ε2, ε3 and ε4 alleles, respectively [13]. The roles of these alleles in modulating risk for AD have been widely studied. (a) The ε4 allele contributes to the risk for AD across most populations [e.g., 14, 15]. (b) AD patients with the ε4 allele have an earlier age-of-onset than those without the ε4 allele [e.g., 16, 17]. (c) The ε4 allele has a significant dose effect on the risk for AD [e.g., 18-20]. (d) The ε2 allele may protect individuals from being affected with AD [e.g., 14, 15].

Saitohin (STH), an intronless gene, has been shown to be nested in the intron between exons 9 and 10 of the tau gene, 2.5 kb downstream from exon 9. This region is functionally critical due to the splicing of exon 10. The special location of the STH gene has prompted investigation into its possible role in AD and other neurodegenerative disorders. The A224G polymorphism in the STH gene, which causes a glutamine (Q) to arginine (R) substitution at residue 7 (Q7R), is in linkage disequilibrium with the extended tau H1/H2 haplotype [21, 22]. That is, the STH Q allele is associated with tau haplotype H1, and the STH R allele is associated with haplotype H2. An initial study by Conrad et al. [23] demonstrated that the Q7R polymorphism in the STH gene was associated with risk for AD. The STH gene has also been associated with autosomal dominant frontotemporal dementia (FTD), progressive supranuclear palsy (PSP) and Pick’s disease [21, 23-25]. Nevertheless, these findings remain controversial [22, 25-29].

The purpose of the present study was to elucidate the associations between the variants at APOE and STH loci and AD in our samples, and to explore the gene-dose effects and evaluate the implications of variation at the APOE gene in the diagnosis of AD.

Methods and Materials

Subjects
The sample included 658 European-Americans, including 369 patients with AD and 289 healthy controls. The diagnosis of AD was based on criteria of the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer’s disease and Related Disorders Association (NINCDS-ADRDA) [30]. The AD cases were divided into an early-onset (EOAD) group and a late-onset (LOAD) group based on an age-of-onset of 70 years [27, 29]. Each subject was evaluated for an approximate date of AD onset, based on careful review of medical records and detailed interviews with one or more primary caregivers. The date of onset was operationally defined as the date at which the “earliest definite AD symptom” appeared. The mean age of AD patients was 73.6±8.4 years (range: 51.8 to 92.7); the mean age-of-onset was 69.3±8.3 years (range: 44.6 to 86.7; 3 subjects unknown); 143 were male, 226 were female; 180 had positive family history (FH+), 184 had negative family history (FH–), and 5 had unknown family history. Family history of AD was assessed using the Alzheimer Dementia Risk Questionnaire (ADRQ) [31] and the Dementia Questionnaire (DQ) [32]. Family history was considered positive if at least one first-degree relative met criteria for primary degenerative dementia. No cases suggestive of autosomal dominant transmission were identified.

There were two sets of control subjects who were differentiated based on the method of ascertainment. The first set of healthy controls (n=185) was recruited through advertisements in the community. They were screened using the Structured Clinical Interview for DSM-III-R (SCID), the Computerized Diagnostic Interview Schedule for DSM-III-R (C-DIS-R), the Schedule for Affective Disorders and Schizophrenia (SADS) [33], or an unstructured interview, to exclude major Axis I disorders, including substance dependence, psychotic disorders, mood disorders, anxiety disorders and dementia. Their mean age was 28.1±9.1 years (range: 18.0 to 52.0); 81 were male and 104 were female. The second set of healthy controls (n=104) was recruited primarily from among spouses of AD patients. Their mean age was 63.3±16.3 years (range: 21.1 to 87.5); 49 were male and 55 were female. They were evaluated as being in generally good medical health for their age on the basis of a comprehensive
evaluation that included medical history, physical and neurological examinations, serum chemistries, thyroid function studies, complete blood count, B12, folate, VDRL, urinalysis, electrocardiogram, and brain MRI or CT. The second set of controls was also required to have no significant evidence of cognitive impairment, as indicated by a Mini-Mental State Examination (MMSE) [34] score > 27. Subjects were recruited at Yale University School of Medicine, the University of Connecticut Health Center, or the VA Connecticut Healthcare System, West Haven Campus. Informed consent was obtained from all the patients and the controls. This study was performed after approval by the appropriate Institutional Review Boards (IRBs).

Genotyping
Genomic DNA was extracted from peripheral blood by standard methods. The region flanking the two target markers within exon 5 of APOE, rs429358 and rs7412, was amplified by a single polymerase chain reaction (PCR) using the following primers [35]: APOE-A: 5'-CGGGCACGGCTGTCCAAGGAG-3' and APOE-C: 5'-CACGCGGCCCTGTTCCAAGGAG-3' (g is mismatched). PCR was performed in a final volume of 10 µl with 1× PC2 buffer (Ab Peptides, Inc., St. Louis, MO), 1M betaine, 0.5 units of KlenTaq polymerase (Ab Peptides, Inc., St. Louis, MO) and 10 ng DNA. PCR conditions were set as follows: 95°C for 5 min; 35 cycles of 95°C for 30 s, 64°C for 30 s, and 72°C for 30 s. The genotypes were analyzed on 5% metaphor agarose gel after digestion with HhaI (New England Biolabs Inc., Beverly, MA). The size of the PCR product was 303 bp, within which there are eight constant HhaI cleavage sites (GCG|C) and two variant HhaI cleavage sites (see Figure 1).

The region flanking the Q7R marker in the STH gene was amplified by PCR using the primers from the initial study by Conrad et al. [23]. PCR was performed in a final volume of 15 µl with 1× PC2 buffer, 1M betaine, 0.75 units of KlenTaq polymerase and 25 ng DNA. PCR conditions were set as follows: 95°C for 5 min; 30 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s. The genotypes were analyzed on 3% metaphor agarose gel after digestion with Hinfl
(New England Biolabs Inc., Beverly, MA). The size of the PCR product was 226 bp. The R allele (Arginine, CGA) can be cut by *HinfI* (97bp+74bp+55bp), whereas the Q allele (Glutamine, CAA) cannot (171bp+55bp).

**Statistical analysis**

The comparisons in allele and genotype frequency distributions between two groups were performed with Fisher’s exact test. Bonferroni correction was used to adjust the α level of multiple comparisons [36].

Positive predictive values (PPVs) were calculated with Bayes’ rule [37]. $P(AD)$ was the prior probability of developing AD, i.e., the prevalence of AD (see Formula). We used 15% as the estimated prevalence of AD [38]; $P(Controls)=1-P(AD)$; $P(\epsilon|AD)$ was allele or genotype frequency in AD cases, and $P(\epsilon|Controls)$ was allele or genotype frequency in controls. Both $P(\epsilon|AD)$ and $P(\epsilon|Controls)$ were estimated from the present study; $\hat{P}(AD|\epsilon)$ was the posterior probability of developing AD given a certain allele or genotype.

Positive likelihood ratios (LRs+) were calculated by dividing the allele or genotype frequencies in AD cases by those in controls [39]. For example, if the frequency of the $\epsilon4/\epsilon4$ genotype is 0.139 in AD cases and 0.037 in controls, then the LR+ is equal to $0.139/0.037 = 3.757$.

The dose effect of the *APOE* gene, i.e., the relationship between the risk for AD and the number of *APOE* alleles, was tested by the chi-square test for trend using the software EPISTAT [40]. The relationships between the number of *APOE* alleles and their LRs+ were tested with Spearman’s rank correlation analysis implemented in SPSS 13.0 (SPSS Inc., Chicago, IL). Gene dose effects for *APOE* were plotted using a polynomial curve-fitting plot method in S-PLUS 2000 (Mathsoft Engineering & Education, Inc., Cambridge, MA).

Age, sex, and AD family history are confounders that may cause false positive or false negative results. Thus, we used stepwise logistic regression analysis to investigate the association between the risk for AD and the number of
APOE and STH alleles, controlling for the effects of the potential confounders. In the stepwise logistic regression model, the diagnosis served as the dependent variable; the independent variables included the number of APOE ε4 alleles, the number of APOE ε2 alleles, the number of STH R alleles, the interaction between STH R allele and APOE alleles, age, sex and AD family history. This analysis was performed with SPSS 13.0 software.

Results

There was no significant difference in allele frequency distributions, genotype frequency distributions or dose effects of APOE and STH gene between our two sets of controls, so we combined the two control groups into one larger control group.

Associations of APOE alleles and genotypes with Alzheimer’s disease

The comparisons of allele and genotype frequency distributions between AD cases and controls are shown in Tables 1 and 2. The genotype frequency distributions in both AD cases and controls were in Hardy-Weinberg equilibrium (HWE).

The overall allele and genotype frequency distributions in AD cases were significantly different from those in controls. The frequencies of the ε4 allele, ε3/ε4 and ε4/ε4 genotypes were significantly higher in AD cases than in controls and the frequencies of the ε2, ε3 alleles, ε2/ε3 and ε3/ε3 genotypes were significantly lower in AD cases than in controls.

We also compared allele and genotype frequencies in AD subgroups (EOAD, LOAD, FH+ AD, FH− AD, male AD and female AD) with those in controls. The overall allele and genotype frequency distribution in each of the AD subgroups was significantly different from that in controls. Specifically, the frequencies of the ε4 allele and the ε3/ε4 genotype in each of the AD subgroups, and the ε4/ε4 genotype in EOAD, FH+ AD, FH− AD, and female AD were significantly higher than those in controls; the frequencies of the ε3 allele and the
ε3/ε3 genotype in each of the AD subgroups, and the ε2 allele in EOAD, FH+ AD and female AD were significantly lower than those in controls. The genotype frequency distributions were significantly different between EOAD and LOAD [(the ε4/ε4 genotype frequency in EOAD (0.203) was significantly higher than that in LOAD (= 0.082)]. Among these differences, the nominal difference in the frequency of the ε2 allele between cases and controls was not statistically significant after Bonferroni correction.

Stepwise logistic regression analyses showed that after adjusting for age, sex, and AD family history, the ε4 and ε2 alleles were still significantly associated with risk for AD (Pε4 = 0.014, adjusted ORε4 = 1.86, 95% CIε4: 1.13-3.05; Pε2 = 0.041, adjusted ORε2 = 0.36, 95% CIε2: 0.13-0.96).

**PPVs and LR**+ of the **APOE** gene for the diagnosis of AD

PPVs and LR+ of APOE alleles and genotypes for AD are listed in Table 3. Both PPVs and LR+ of APOE alleles and genotypes were in the following order: ε4/ε4 > ε4 > ε3/ε4 > ε3 > ε2/ε4 > ε3/ε3 > ε2 > ε2/ε3.

We also compared PPVs for different subtypes of AD. PPVs of the ε4/ε4 genotype were much higher in EOAD (49.2% vs. 28.1% for LOAD), female AD (48.4% vs. 28.9% for male AD) and FH+ AD (43.6% vs. 36.8% for FH− AD). In addition, the PPV for the ε3/ε4 genotype was higher in LOAD (31.1%) than EOAD (26.0%). Finally, the PPVs were lower for FH+ AD for the ε2 allele and the ε2/ε3 genotype (3.9%, 4.6%, respectively) than for FH− AD (8.0%, 6.1%, respectively).

**Gene dose effects of the APOE gene on the risk for AD** (Table 4 and Figure 2)

The chi-square test for trend analyses showed that there was a significant positive correlation between the number of ε4 alleles and risk for AD and a significant negative correlation between the number of ε2 or ε3 alleles and risk for AD.

Similarly, the Spearman’s rank correlation analysis showed that the number of APOE alleles was significantly correlated with LR+, which increased
linearly with the number of the ε4 alleles (correlation coefficient $r_{ε4} = 1.0$; slope $K_{ε4} = 1.602$) and decreased linearly with the number of ε2 or ε3 alleles (correlation coefficient $r_{ε2 or ε3} = -1.0$; slope $K_{ε2} = -0.543$; slope $K_{ε3} = -1.122$).

**Association of the STH gene with AD**

No significant difference in STH allele and genotype frequency distributions was found between AD cases and controls. Even after adjusting for potential confounding by the *APOE* gene, age, sex and AD family history, stepwise logistic analyses showed no association of STH alleles or genotypes with AD.

**Interactive effects between the STH gene and the APOE gene**

Using STH genotypes, we grouped all subjects into QQ, RR and QR groups. We then compared APOE allele and genotype frequency distributions in these three groups in both cases and controls. No significant difference was found for any of the comparisons (data not shown).

**Discussion**

The present study confirmed the well-established association between the *APOE* gene and AD. All three *APOE* alleles (ε2, ε3 and ε4) showed dose effects on the risk for AD, and followed a co-dominant mode of inheritance. We also examined, for the first time to our knowledge for a trait in neuropsychiatry, a mathematical measure of the predictive value of each *APOE* allele and genotype for AD diagnosis risk.

In addition to a significant association between the *APOE* gene and Alzheimer’s disease, subgroup analyses revealed an association with subtypes based on age of onset, family history, and sex. The ε4 allele, the ε4/ε4 genotype and the ε3/ε4 genotype were risk factors for AD; the ε2 allele, the ε3 allele, the ε2/ε3 genotype and the ε3/ε3 genotype were protective factors for AD. These findings are consistent with those in most previous studies [e.g., 14, 15]. Further comparisons among AD subgroups and controls showed that the ε4/ε4 genotype
frequency was significantly higher in EOAD than in LOAD and controls, suggesting that the ε4/ε4 genotype can significantly reduce the age-of-onset. This is consistent with findings in other studies [e.g., 19].

We also found that the PPV of the ε4/ε4 genotype was significantly higher in females (48.4%) than in males (28.9%). Although the ε4/ε4 genotype frequency in female AD cases was significantly higher than in female controls, we found no significant difference in males. These results suggest that the ε4/ε4 genotype is a stronger risk factor for females than for males. This is consistent with findings from other studies [e.g., 41-43]. However since sex distributions were not well matched between cases and controls, it could also reflect a stratification effect by sex.

Both the chi-square test for trend and the regression analyses revealed that the risk for AD increased significantly with the number of ε4 alleles. This is also consistent with findings from other studies [e.g., 18]. In addition, we found that the risk for AD decreased with the number of ε2 or ε3 alleles. Furthermore, the dose of APOE alleles was linearly related to LR+. These results are all compatible with those from our allelewise analyses.

This information is of importance in predicting the development of AD in early life. However, not all subjects with the ε4 allele develop AD, nor do all AD patients carry the ε4 allele. On the other hand, not all subjects are protected against AD by the ε2 and ε3 alleles. Therefore, it is important to estimate the probability that these allele carriers will develop AD. We found that the ε4/ε4 genotype had a PPV of 39.90% and an LR+ of 3.76 for AD. In other words, a subject carrying two ε4 alleles has a probability of 39.90% to develop AD. In contrast, a subject carrying one ε4 allele and one ε3 allele has a probability of 28.80% to develop AD, and a subject carrying one ε4 allele and one ε2 allele has a probability of 10.82% to develop AD. Based on the interpretation of LRs+ by Ebell [39], the presence of APOE alleles can only mildly change the risk for AD, despite a highly significant association with AD. This implies that APOE genotype testing can provide evidence on whether a subject may develop AD, but it is not sufficient as an independent screening or predictive test for the diagnosis of AD.
Additionally, we found the following order for both PPVs and LRs* of APOE alleles and genotypes with respect to the diagnosis of AD: ε4/ε4 > ε4 > ε3/ε4 > ε3 > ε2/ε4 > ε3/ε3 > ε2 > ε2/ε3 (see Table 3). This order shows that: (1) ε4/ε4 > ε4, suggesting that the risk for AD increases with the number of ε4 alleles; (2) ε4 > ε3/ε4 and ε3/ε4 > AD population prevalence, suggesting that the ε3 allele reduces the risk for AD conveyed by the ε4 allele, but the protective effect of ε3 is weaker than the risk effect of ε4; (3) ε3/ε3 < ε3, suggesting that the protection against AD increases with the number of ε3 alleles; (4) ε3 > ε2/ε3 and ε2 > ε2/ε3, suggesting that the protective effect on AD risk for a genotype containing two protective alleles is greater than that for a genotype containing only one of the protective alleles; (5) ε3 > ε2 and ε3/ε3 > ε2/ε3, suggesting that the ε2 allele is a stronger protective factor for AD than the ε3 allele, which is reflected in their positions on the Y axis in the figure depicting the dose effect (Figure 2); and (6) ε3/ε4 > AD population prevalence, but ε2/ε3 < ε3/ε3 < ε3 < AD population prevalence, suggesting that without ε4, the ε3 allele and any genotypes containing the ε3 allele cannot increase risk for AD, that is, it is ε4, not ε3, that contributes to the increased risk of AD associated with the ε3/ε4 genotype. Similarly, the PPV for ε2/ε4 < AD population prevalence (i.e., a protective effect), but ε4/ε4 > ε4 > ε3/ε4 > AD population prevalence (i.e., a risk effect), suggesting that without ε2, the ε4 allele and any genotypes containing the ε4 allele (e.g., ε4/ε4 and ε3/ε4) do not have a protective effect; it is ε2, not ε4, that results in the ε2/ε4 genotype having a lower PPV. Taken together, the order of these effects suggests that ε4 is a dose-response risk factor for developing AD, ε2 is a dose-response protective factor, and ε3 is a relatively weaker dose-response protective factor. These findings are consistent with the results of our allelewise analyses, chi-square tests for trends, and logistic regression analyses.

There has been debate about whether the presence of a “bad” allele (i.e., ε4) or of a “good” allele (ε2 or ε3), or both, contribute to the association between APOE and AD. The answer to this question is important for the development of specific therapies for AD [45]. Our results tend to show that both the “bad” allele (ε4) and the “good” alleles (ε2 and ε3) are involved in the risk for AD, consistent
with codominant inheritance. These findings are supported by the evidence from studies on the neuropathological processes involved in AD [e.g., 6].

Noting both the close interaction between the APOE and the tau proteins and the physical proximity of the Tau and STH genes, we investigated the correlation between effects of the APOE and STH gene polymorphisms. We found no significant interactive effect between these two genes either in cases or in controls. This finding was consistent with our regression analysis and the studies by Conrad et al. [23] and Peplonska et al. [22]. Thus, the APOE gene affects risk for AD through a pathway independent of the STH gene polymorphism we queried.

We also found no associations between STH alleles and AD, even after adjusting for potential confounders, including age, sex, and family history. Neither the genotype analysis nor the gene-dose analysis showed any association. Our results suggest that STH may not be a risk gene for AD. The initial positive findings by Conrad et al. [23] may be attributable to sampling bias in the context of small sample sizes (51 AD cases; 30 healthy controls). Our sample size (286 AD cases; 197 healthy controls) is much larger than theirs. Moreover, our negative findings are in good agreement with many other studies, which also have much larger sample sizes (e.g., 499 AD cases and 402 controls by Verpillat et al. [25]; 225 AD cases and 144 controls by Streffer et al. [27]; 200 AD cases and 458 controls by Clark et al. [28]; 690 AD families, 903 AD cases and 320 controls by Oliveira et al. [29]; 100 AD cases and 100 controls by Peplonska et al. [22]). Additionally, the Q allele frequency (0.867) in controls in the initial study is similar to both controls and cases in our and the other negative studies; but the Q allele frequency (0.676) in AD cases is significantly lower than those in cases and controls in most of the published studies [22, 25, 27-29]. So far, there has been only one study [24] reporting a replicated positive finding between the genotype STH RR and AD (p=0.04), but even this positive finding is only nominal and does not survive after Bonferroni correction. Therefore, we conclude that the STH gene Q7R variation does not play an important role in the pathology of AD.
**Competing interests:** The authors declare that they have no competing interests.

**Authors’ contributions:**
LZ designed the study, genotyped most of the subjects, analyzed and interpreted the data, and drafted the manuscript. CHVD collected the samples, participated in the design of the study, interpreted the data, and critically revised the manuscript. XL participated in designing the study, genotyping some subjects, analyzing and interpreting the data, and drafting the manuscript. HRK collected the samples, contributed to the interpretation, and made critical comments on the manuscript. BZY contributed to statistical analysis and data interpretation. JG obtained the funding, prepared the samples, supervised the study, interpreted the data, and made critical comments on the manuscript. All authors read and approved the final manuscript.

**Acknowledgments:**
This work was supported in part by funds from the U.S. Department of Veterans Affairs (the VA Medical Research Program, and the VA Connecticut–Massachusetts Mental Illness Research, Education and Clinical Center [MIRECC], and the VA Research Enhancement Award Program [REAP] research center), National Institute of Mental Health (NIMH) grant K02-MH01387, National Institute on Drug Abuse (NIDA) grants R01-DA12849, R01-DA12690, and K24-DA15105, National Institute on Alcohol Abuse and Alcoholism (NIAAA) grants R01-AA016015-01, R01-AA11330, K08-AA13732 and K24-AA13736, National Center for Research Resources (NCRR) grant M01-RR06192 (University of Connecticut General Clinical Research Center), and Alcoholic Beverage Medical Research Foundation (ABMRF) grant award R06932 (X Luo). Ann Marie Lacobelle provided excellent technical assistance.
References


8. Farrer M, Skipper L, Berg M, Bisceglio G, Hanson M, Hardy J, Adam A, Gwinn-Hardy K, Aasly J: The tau H1 haplotype is associated and the risk of


15. Korovaitseva GI, Shcherbatykh TV, Selezneva NV, Gavrilova SI, Golimbet VE,


35. Tsai MS, Tangalos EG, Petersen RC, Smith GE, Schaid DJ, Kokmen E, Ivnik


2000, 33:529-537.


Figure 1. *HhaI* cleavage sites within the *APOE* amplicon. \(^1\) The residue positions are updated in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP/). \(\varepsilon_2\) allele (haplotype): 130Cys+176Cys (old:112Cys+158Cys) = 91bp+83bp+others (≤33bp); \(\varepsilon_3\) allele (haplotype): 130Cys+176Arg (old:112Cys+158Arg) = 91bp+48bp+35bp+others (≤33bp); \(\varepsilon_4\) allele (haplotype): 130Arg+176Arg (old:112Arg+158Arg) = 72+48+35+19+others (≤33bp).

Figure 2. Dose effects of *APOE* gene alleles. X-axis represents the copies of alleles (\(\varepsilon_2\) allele, pink line; \(\varepsilon_3\) allele, green line; \(\varepsilon_4\) allele, red line); Y-axis represents the positive likelihood ratios of genotype frequency (in AD cases vs. in Controls).
Table 1. Distributions of allele frequencies of APOE and STH gene variations in European-American subjects

<table>
<thead>
<tr>
<th></th>
<th>APOE alleles</th>
<th>Exact p-values</th>
<th>STH alleles</th>
<th>Exact p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ε2</td>
<td>ε3</td>
<td>ε4</td>
<td>Total</td>
</tr>
<tr>
<td>AD</td>
<td>0.024</td>
<td>17</td>
<td>0.617</td>
<td>443</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EOAD</td>
<td>0.017</td>
<td>6</td>
<td>0.590</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOAD</td>
<td>0.027</td>
<td>10</td>
<td>0.644</td>
<td>237</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FH+</td>
<td>0.014</td>
<td>5</td>
<td>0.590</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FH-</td>
<td>0.030</td>
<td>11</td>
<td>0.641</td>
<td>232</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.032</td>
<td>9</td>
<td>0.635</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.018</td>
<td>8</td>
<td>0.606</td>
<td>264</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>0.061</td>
<td>30</td>
<td>0.801</td>
<td>394</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.065</td>
<td>15</td>
<td>0.800</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.057</td>
<td>15</td>
<td>0.802</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

f, frequency; n, number of alleles; N, number of individuals; 2N, number of chromosomes; AD, Alzheimer’s disease; EOAD, early-onset AD (<70yrs); LOAD, late-onset AD (>70yrs); Controls, the total controls combining two subgroups through two different diagnostic evaluations.

The numbers in last column denote: 1. Comparing overall AD with controls (α was set at 0.017 for each APOE allele comparison); 2. Comparing EOAD, LOAD, FH+ AD and FH- AD with controls respectively (α was set at 0.003 for each APOE allele comparison); 3. Comparing male AD with male controls and female AD with female controls (α was set at 0.003 for each APOE allele comparison).
Table 2. Distributions of the genotype frequencies of the APOE gene variations in European-American subjects

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
</tr>
<tr>
<td>AD</td>
<td>0.036</td>
<td>13</td>
<td>0.011</td>
<td>4</td>
<td>0.384</td>
<td>138</td>
<td>0.429</td>
<td>154</td>
<td>0.139</td>
<td>50</td>
<td>359</td>
<td>0.001</td>
</tr>
<tr>
<td>EOAD</td>
<td>0.029</td>
<td>5</td>
<td>0.006</td>
<td>1</td>
<td>0.390</td>
<td>67</td>
<td>0.372</td>
<td>64</td>
<td>0.203</td>
<td>35</td>
<td>172</td>
<td>0.004</td>
</tr>
<tr>
<td>LOAD</td>
<td>0.038</td>
<td>7</td>
<td>0.016</td>
<td>3</td>
<td>0.386</td>
<td>71</td>
<td>0.478</td>
<td>88</td>
<td>0.082</td>
<td>15</td>
<td>184</td>
<td>0.010</td>
</tr>
<tr>
<td>FH+</td>
<td>0.029</td>
<td>5</td>
<td>0.000</td>
<td>0</td>
<td>0.341</td>
<td>59</td>
<td>0.468</td>
<td>81</td>
<td>0.162</td>
<td>28</td>
<td>173</td>
<td>0.004</td>
</tr>
<tr>
<td>FH-</td>
<td>0.039</td>
<td>7</td>
<td>0.022</td>
<td>4</td>
<td>0.425</td>
<td>77</td>
<td>0.392</td>
<td>71</td>
<td>0.122</td>
<td>22</td>
<td>181</td>
<td>0.010</td>
</tr>
<tr>
<td>Male</td>
<td>0.043</td>
<td>6</td>
<td>0.021</td>
<td>3</td>
<td>0.390</td>
<td>55</td>
<td>0.447</td>
<td>63</td>
<td>0.099</td>
<td>14</td>
<td>141</td>
<td>0.053</td>
</tr>
<tr>
<td>Female</td>
<td>0.032</td>
<td>7</td>
<td>0.005</td>
<td>1</td>
<td>0.381</td>
<td>83</td>
<td>0.417</td>
<td>91</td>
<td>0.165</td>
<td>36</td>
<td>218</td>
<td>0.015</td>
</tr>
</tbody>
</table>

EOAD vs. LOAD

<table>
<thead>
<tr>
<th>Controls</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>N</td>
<td>f</td>
</tr>
<tr>
<td>0.106</td>
<td>26</td>
<td>0.016</td>
</tr>
<tr>
<td>0.113</td>
<td>13</td>
<td>0.017</td>
</tr>
<tr>
<td>0.099</td>
<td>13</td>
<td>0.015</td>
</tr>
</tbody>
</table>

f, N, AD, EOAD, LOAD, Controls and the comparing methods see Table 1. EOAD vs. LOAD, comparing EOAD and LOAD. α was set at 0.010 (or 0.002) for each genotype comparison between overall AD and controls (or between AD subgroups and controls), referring to Table 1. ε2/ε3, the genotype consists of allele ε2 and allele ε3; as analogy.
Table 3. Interpretation of PPVs and LRs\(^+\) of APOE alleles and genotypes

| allele/genotype | PPV \([ \hat{p} (AD|\epsilon) ]\) | LR\(^+\) | LR range\(^+\) | Significance for diagnosis |
|-----------------|----------------------------------|-----------|-----------------|---------------------------|
| \(\epsilon 4/\epsilon 4\) | 39.90%                           | 3.76      | 5~10            | greatly increasing risk for AD |
| \(\epsilon 4\)     | 31.50%                           | 2.60      | 2~5             | moderately increasing risk for AD |
| \(\epsilon 3/\epsilon 4\) | 28.80%                           | 2.29      | 0.5~1           | minimally decreasing risk for AD |
| \(\epsilon 3\)     | 28.80%                           | 1         | 0.1~0.2         | moderately decreasing risk for AD |
| \(\epsilon 2/\epsilon 3\) | 15.00%                           | 0.77      | <0.1            | greatly decreasing risk for AD |

PPV, estimated positive predictive value; LR\(^+\), positive likelihood ratios.

\(^{+}\)Reference to [39].
**Table 4. Distributions of the frequencies of the genotypes with ascending APOE allele number**

<table>
<thead>
<tr>
<th></th>
<th>Number of ε2 alleles</th>
<th></th>
<th></th>
<th></th>
<th>Chi-Square</th>
<th></th>
<th></th>
<th></th>
<th>Chi-Square</th>
<th></th>
<th></th>
<th></th>
<th>Chi-Square</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
<td>f</td>
<td>N</td>
</tr>
<tr>
<td>AD</td>
<td>0.953</td>
<td>342</td>
<td>0.047</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>11.32</td>
<td>0.001</td>
<td>0.150</td>
<td>54</td>
<td>0.465</td>
<td>167</td>
<td>0.384</td>
<td>138</td>
<td>43.02</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>EOAD</td>
<td>0.965</td>
<td>166</td>
<td>0.035</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>9.73</td>
<td>0.002</td>
<td>0.209</td>
<td>36</td>
<td>0.401</td>
<td>69</td>
<td>0.390</td>
<td>67</td>
<td>37.41</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LOAD</td>
<td>0.946</td>
<td>174</td>
<td>0.054</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>5.69</td>
<td>0.017</td>
<td>0.098</td>
<td>18</td>
<td>0.516</td>
<td>95</td>
<td>0.386</td>
<td>71</td>
<td>26.27</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FH+</td>
<td>0.971</td>
<td>168</td>
<td>0.029</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>11.46</td>
<td>0.001</td>
<td>0.162</td>
<td>28</td>
<td>0.497</td>
<td>86</td>
<td>0.341</td>
<td>59</td>
<td>41.06</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FH−</td>
<td>0.939</td>
<td>170</td>
<td>0.061</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>4.49</td>
<td>0.034</td>
<td>0.144</td>
<td>26</td>
<td>0.431</td>
<td>78</td>
<td>0.425</td>
<td>77</td>
<td>24.67</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Male</td>
<td>0.936</td>
<td>132</td>
<td>0.064</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>3.29</td>
<td>0.069</td>
<td>0.121</td>
<td>17</td>
<td>0.489</td>
<td>69</td>
<td>0.390</td>
<td>55</td>
<td>16.00</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>0.963</td>
<td>210</td>
<td>0.037</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>8.02</td>
<td>0.005</td>
<td>0.170</td>
<td>37</td>
<td>0.450</td>
<td>98</td>
<td>0.381</td>
<td>83</td>
<td>26.33</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Controls</td>
<td>0.878</td>
<td>216</td>
<td>0.122</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0.053</td>
<td>0.293</td>
<td>72</td>
<td>0.654</td>
<td>161</td>
<td>0.760</td>
<td>187</td>
<td>0.203</td>
<td>50</td>
<td>0.037</td>
</tr>
<tr>
<td>Male</td>
<td>0.870</td>
<td>100</td>
<td>0.130</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0.061</td>
<td>0.278</td>
<td>32</td>
<td>0.661</td>
<td>76</td>
<td>0.774</td>
<td>89</td>
<td>0.183</td>
<td>21</td>
<td>0.043</td>
</tr>
<tr>
<td>Female</td>
<td>0.885</td>
<td>116</td>
<td>0.115</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0.046</td>
<td>0.305</td>
<td>40</td>
<td>0.649</td>
<td>85</td>
<td>0.748</td>
<td>98</td>
<td>0.221</td>
<td>29</td>
<td>0.031</td>
</tr>
</tbody>
</table>

f, N, AD, EOAD, LOAD, Controls, and the comparing methods see Tables 1 and 2.

α was set at 0.017 (or 0.003) for each Chi-Square test for trend analysis in overall AD (or AD subgroups), referring to Tables 1 and 2.
<p>| Length (bp) | (|\langle 19+72=91 \rangle| | \langle 48+35=83 \rangle| |
|---|---|
| <strong>HhaI</strong> cleavage sites 0 | 33 | 49 | 68 | 140 | 158 | 206 | 241 | 248 | 259 | 272 | 303 |
| New residue positions | 118 | 124 | 130 | 176 |
| Old residue positions | 112 | 158 |
| Nucleotide substitution | T/C | T/C |
| Amino acid substitution | Cys/Arg | Cys/Arg |
| NCBI database Rs# | rs429358 | rs7412 |</p>
<table>
<thead>
<tr>
<th>Copies of alleles</th>
<th>Positive likelihood Ratios of genotype frequency (AD/Controls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0.150/0.053=2.830)</td>
</tr>
<tr>
<td></td>
<td>(0.440/0.203=2.167)</td>
</tr>
<tr>
<td></td>
<td>(0.421/0.760=0.554)</td>
</tr>
<tr>
<td></td>
<td>(0.953/0.878=1.085)</td>
</tr>
<tr>
<td></td>
<td>(0.047/0.122=0.385)</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
</tr>
</tbody>
</table>

The slopes: $K_{e1}=1.602, K_{e2}=-1.122, K_{e3}=-0.543$
Additional files provided with this submission:

Additional file 1: Cover letter.doc: 339Kb
http://www.behavioralandbrainfunctions.com/imedia/122372619999737/sup1.DOC