Genetic association study on in and around the APOE in late-onset Alzheimer disease in Japanese

Norihiro Takei a, Akinori Miyashita a, Tamao Tsukie a, Hiroyuki Arai b, Takashi Asada c, Masaki Imagawa d, Mikio Shoji e, Susumu Higuchi f, Katsuya Urayama g, Hideo Kimura h, Akiyoshi Kakitai i, Hitoshi Takahashi j, Shoji Tsui k, Ichiro Kanazawa h, Yasuo Iharal j, Shoji Odani m, Ryozo Kuwano a,⁎ and the Japanese Genetic Study Consortium for Alzheimer Disease 1

A R T I C L E  I N F O

Article history:
Received 23 October 2008
Accepted 21 January 2009
Available online 3 February 2009

Keywords:
Alzheimer disease
APOE
PVRL2
TOMM40
APOC1
SNP
Association
Linkage disequilibrium
Case-control study

A B S T R A C T

The ε4 allele of APOE is a well-characterized genetic risk factor for late-onset Alzheimer disease (LOAD). Nevertheless, using high-density single nucleotide polymorphisms (SNPs), there have only been a few studies involving genetic association and linkage disequilibrium (LD) analyses of in and around the APOE. Here, we report fine mapping of a genomic region (about 200 kb) including the APOE in Japanese using 260 SNPs (mean intermarker distance, 0.77 kb). A case-control study demonstrated that 36 of these SNPs exhibited significance after adjustment for multiple testing. These SNPs are located in a genomic region including four genes, PVRL2, TOMM40, APOE, and APOC1. Recombination rate estimation revealed that the associated region is firmly sandwiched between two recombination hotspots. Strong LD between these SNPs was observed (mean |D'| = 0.914). These data suggest that the three genes other than APOE, i.e. PVRL2, TOMM40 and APOC1, could also yield a predisposition to LOAD.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Alzheimer disease (AD) is the main cause of dementia in the elderly. Its main neuropathological features are extracellular deposition of amyloid-β protein (Aβ1-42) and intracellular formation of neurofibrillary tangles. Late-onset AD (LOAD), accounting for ~95% of AD, is thought to be a multi-factorial disease, probably caused by complicated interactions between genetic and environmental factors.

To date, only the apolipoprotein E gene (APOE) on chromosome 19q is universally recognized as a major disease susceptibility gene for LOAD [1–3]. APOE has three common alleles, APOE-ε2, APOE-ε3 and APOE-ε4. These three alleles are defined by two non-synonymous single nucleotide polymorphisms (SNPs), rs429358 (TGC→CGC, Cys112Arg) and rs7412 (GCC→TGC, Arg158Cys): APOE-ε2, T-T (Cys-Cys); APOE-ε3, T-C (Cys-Arg); and APOE-ε4, C-C (Arg-Arg). Among these alleles, the APOE-ε3 one is the most frequent (0.49–0.91), the APOE-ε2 one being the rarest (0.00–0.15), in all populations thus far investigated [4]. We previously reported that in Japanese normal controls (≥60 y.o.) the frequencies of the APOE-ε2, APOE-ε3 and
APOE-ε4 alleles are 0.05, 0.86 and 0.09, respectively [5,6]. Compared to the APOE-ε3 allele, the APOE-ε4 allele is a strong risk factor with an odds ratio (OR) of 2.0–4.0, the APOE-ε2 allele being protective as to LOAD [7,8].

Recent genome-wide association studies (GWAS) reconfirmed that an APOE linkage disequilibrium (LD) locus is strongly associated with LOAD [9–11]. Through these GWAS, it appears to be finally demonstrated that reproducible, strong association signals are only observed in and around the APOE. Up to now, it has been implicitly assumed that the genetic association in and around the APOE in LOAD thoroughly reflects the APOE-ε4 association itself, owing to strong LD with the two APOE SNPs rs429358 and rs7412 [11]. However, it was recently suggested that multiple (cis-regulatory) SNPs in and around the APOE may contribute to disease susceptibility via alteration of gene expression. Belbin et al. [12] found that the rate of the cognitive decline in AD patients is affected by a SNP (rs440446) lying in a regulatory element of intron 1 of the APOE. Interestingly, the effect of this SNP on the cognitive decline seems to be independent of the APOE-ε4 allele. Quantitative trait analysis also indicated that three cis-acting SNPs (rs449647, rs17684509 and rs7247551) of the APOE and three SNPs (rs11556505, n17664883 and rs157584) within the translocase of the outer mitochondrial membrane 40 homolog (yeast) gene (TOMM40), proximally located upstream of the APOE, are associated with the level of cerebrospinal fluid APOE in healthy non-demented subjects [13]. These data raise the possibility that other SNPs which affect a variety of quantitative traits involved in the LOAD pathogenesis might be hidden in and around the APOE. Therefore, it would be very important to conduct a genetic association study using a sufficient number of SNPs to entirely survey the genomic region comprising the APOE.

Here, we carried out a high-density SNP-based case-control association study (N = 1262 [LOAD, 547; control, 715]) using 260 SNPs for an approximately 200 kb genomic region comprising 8 genes in addition to the APOE: the Cas-Br-M (murine) ecotropic retroviral transforming sequence c gene (CBLC), the basal cell adhesion molecule (Lutheran blood group) gene (BCAM), the poliovirus receptor-related 2 (herpesvirus entry mediator B) gene (PVRL2),

Table 1
Subject information

<table>
<thead>
<tr>
<th></th>
<th>LOAD</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>547</td>
<td>715</td>
</tr>
<tr>
<td>Frequency of female</td>
<td>0.717</td>
<td>0.544</td>
</tr>
<tr>
<td>AAO/AAE (mean±SD)</td>
<td>73.0±6.9</td>
<td>73.0±6.8</td>
</tr>
<tr>
<td>APOE genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2*2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2*3</td>
<td>19</td>
<td>57</td>
</tr>
<tr>
<td>2*4</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>3*3</td>
<td>279</td>
<td>529</td>
</tr>
<tr>
<td>3*4</td>
<td>196</td>
<td>117</td>
</tr>
<tr>
<td>4*4</td>
<td>47</td>
<td>1</td>
</tr>
<tr>
<td>APOE-ε4 allele frequency</td>
<td>0.269</td>
<td>0.087</td>
</tr>
</tbody>
</table>

Fig. 1. Genomic region associated with LOAD. (A) Plot of $-\log_{10} P_{\text{unadj-allele}}$-values of 171 SNPs. Horizontal arrows indicate the transcriptional orientations of individual genes. The dotted line indicates the significant threshold of the $P_{\text{unadj-allele}}$-value = 0.0003 (Bonferroni correction). The most significant SNP, rs429358 ($P_{\text{unadj-allele}}$-value = 4.32E−34), is labeled. (B) Recombination hotspots, HS1 and HS2. Genomic positions are according to NCBI build 36.
TOMM40, the apolipoprotein C-I gene (APOC1), the apolipoprotein C-IV gene (APOC4), the apolipoprotein C-II gene (APOC2), and the cleft lip and palate associated transmembrane protein 1 gene (CLPTM1). Fine LD mapping and estimation of recombination hotspots were also performed to clarify the genetic structure in this region.

Results

Distribution of APOE genotypes

We determined the APOE genotypes of all subjects (N = 1262) enrolled in this study, by means of TaqMan<sup>a</sup> allelic discrimination and direct cycle sequencing for the two non-synonymous SNPs rs429358 and rs7412. Perfect concordance in the genotyping was obtained between the two methods. Deviation from the Hardy–Weinberg equilibrium (HWE) was assessed in controls (N = 715). The HWE P-values of SNPs rs429358 and rs7412 were 0.0358 and 0.1934, respectively. The genotypic distribution of the APOE is presented in Table 1. We compared the allelic distribution of the APOE using a 2 × 3 contingency table between LOAD patients (25; e3, 733; e4, 296) and controls (e2, 72; e3, 1232; e4, 126). As expected, a significant difference was observed: degree of freedom [df] = 2, chi-square value = 152.34, P-value = 3.06 × 3.4. We further computed an estimate of the relative risk for LOAD using a 2 × 2 contingency table. Compared to the APOE-e4 allele, the APOE-e4 showed a strong risk effect for LOAD (df = 1, chi-square value = 141.15, P-value = 1.49E – 32; OR = 3.74, 95% confidence intervals [Cls] = 2.98–4.20) and the APOE-e2 one was protective (df = 1, chi-square value = 6.42, P-value = 1.13E – 2; OR = 0.55, 95% CIs = 0.35–0.70), which are consistent with results of many preceding genetic studies on AD [8; ALZgene web site, http://www.alzforum.org/res/commgen/allgene/].

Allelic association study on in and around the APOE

We conducted high-density SNP genotyping in and around the APOE to identify novel SNPs associated with LOAD in Japanese. Two hundred and fifty-seven SNPs (Supplementary Table 1), spanning about 200 kb (49.988–50.188 Mb), were selected from a genomic region including the nine genes (Fig 1A) mentioned under Introduction. They were genotyped using a sample set, All, comprising 1262 subjects (LOAD, 547; control, 715). Among them, 171 SNPs were actually polymorphic with minor allele frequency (MAF) ≥0.01 and did not show significant deviation from the HWE (P-value ≥0.01) in the Japanese population: 8 missense mutations, 5 silent mutations, 1 in the 5' UTR and 4 in the 3' UTR, 108 in introns, and 45 in intergenic regions (Supplementary Table 1). The mean inter-marker distance ± standard deviation (SD) was 1.18 ± 1.79 kb (95% CIs, 0.91–1.45; range, 7–13,801 bp).

Table 2

Case-control study of the 36 significant SNPs

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Position (bp)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Allele&lt;sup&gt;b&lt;/sup&gt;</th>
<th>MAF&lt;sup&gt;c&lt;/sup&gt;</th>
<th>GSR&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Number of alleles</th>
<th>Number of genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs149300</td>
<td>50066010</td>
<td>A/g</td>
<td>0.493</td>
<td>98.6</td>
<td>112 278 150 206 348 150</td>
<td>502 578 760 648</td>
</tr>
<tr>
<td>rs534447</td>
<td>50067012</td>
<td>C/t</td>
<td>0.487</td>
<td>92.5</td>
<td>149 269 91 153 325 280</td>
<td>567 451 631 685</td>
</tr>
<tr>
<td>rs620976</td>
<td>50079039</td>
<td>G/a</td>
<td>0.238</td>
<td>97.9</td>
<td>142 276 70 414 267 206</td>
<td>790 372 1104 304</td>
</tr>
<tr>
<td>rs282591</td>
<td>50085130</td>
<td>C/g</td>
<td>0.273</td>
<td>98.3</td>
<td>142 276 70 414 267 206</td>
<td>790 372 1104 304</td>
</tr>
<tr>
<td>rs395908</td>
<td>50075799</td>
<td>G/a</td>
<td>0.280</td>
<td>97.9</td>
<td>142 276 70 414 267 206</td>
<td>790 372 1104 304</td>
</tr>
<tr>
<td>rs395915</td>
<td>50085179</td>
<td>G/a</td>
<td>0.273</td>
<td>98.3</td>
<td>142 276 70 414 267 206</td>
<td>790 372 1104 304</td>
</tr>
<tr>
<td>rs395920</td>
<td>50085180</td>
<td>G/a</td>
<td>0.257</td>
<td>97.9</td>
<td>142 276 70 414 267 206</td>
<td>790 372 1104 304</td>
</tr>
</tbody>
</table>

M and in represent major and minor alleles, respectively. The MAF and GSR values were calculated for the sample set All.

<sup>a</sup> Genomic positions are according to NCBI build 36.
<sup>b</sup> Major and minor alleles are indicated in upper and lower case letters, respectively.
<sup>c</sup> MAF, minor allele frequency.
<sup>d</sup> GSR, genotyping success rate.
associated with LOAD. Thirty-five of these SNPs exhibited significance: range of unadjusted allelic \( P \)-values (\( P_{\text{unadj-allele-value}} \)), 4.32E−34 to 2.00E−4 (Table 2 and Fig. 1A). These SNPs remained significant even after Bonferroni correction for multiple testing (Fig. 1A). The genomic positions of the SNPs, spanning approximately 52.3 kb, are presented in Fig. 1 and Table 2. It was found that the associated region encompasses three additional genes to the \( APOE \), i.e. \( PVRL2 \), \( TOMM40 \) and \( APOC1 \) (Fig. 1).

The 171 SNPs were further examined by stratified analysis, based on the \( APOE \) genotypes, using the chi-squared test. Three sub-sample sets, \( \text{Positive-} \varepsilon_4 \) (LOAD, 247; control, 123), \( \text{Negative-} \varepsilon_4 \) (LOAD, 296; control, 586), and \( \varepsilon_3 \times 3 \) (LOAD, 277; control, 525), were used. We did not observe any significant allelic association of these SNPs with LOAD in any sub-sample set (Supplementary Fig. 1).

Novel variants in and around \( PVRL2 \) and \( TOMM40 \)

Among the four associated genes mentioned above, \( APOE \) and \( APOC1 \) have been genetically well investigated as disease susceptibility genes for LOAD [8; ALZgene web site, http://www.alzforum.org/res/com/gen/alzgene/]. On the other hand, the genetic and biological relationships of the two remaining genes, \( PVRL2 \) and \( TOMM40 \), to LOAD have not been well evaluated so far. For these two genes, it is meaningful to explore novel DNA variants that could affect gene function and transcription. To achieve this purpose, we sequenced a total of 19 exons (\( PVRL2 \), 10 exons; and \( TOMM40 \), 9 exons), their exon–intron boundaries and an about 200 bp 5′ region upstream of these genes. Two non-synonymous variants were identified in exons 2 (GenBank accession number, NM_001042724: c.349G→A, p.Glu101Lys; designated as 19AyD0324) and 7 (GenBank accession number, NM_001042724:c.1111G→A, p.Ala355Thr; designated as 19AyD0325) within \( PVRL2 \) (Supplementary Table 2). For \( TOMM40 \), there was one variant 337 bp upstream from the translation initiation site (GenBank accession number, NC_00019.8:g.-337T→C; designated as 19AyD0326) (Supplementary Table 2). A search of the UCSC genomic database (http://genome.ucsc.edu/) demonstrated that all three variants are novel.

The variants, i.e. 19AyD0324, 19AyD0325 and 19AyD0326, were genotyped by means of the TaqMan® method using the overall sample set All. The minor alleles for the two variants 19AyD0324 (nucleotide A) and 19AyD0325 (nucleotide A) both exhibited very low frequencies: 19AyD0324, 0.00821 in LOAD and 0.0141 in LOAD.
controls; and 19AyD0325, 0.0009 in LOAD and 0.0007 in controls. No allelic association with LOAD was exhibited by these two variants: \( P_{\text{unadj-allele-values}} \) of 0.172 for 19AyD0324 and 0.853 for 19AyD0325. However, the variant 19AyD0326 was highly polymorphic with MAF of 0.268 for LOAD and 0.137 for controls, and was significantly associated with LOAD (\( P_{\text{unadj-allele-value}} \) \( \leq 1.50 \times 10^{-16} \); OR [95% CIs], 2.31 [1.89–2.86]) (Table 2).

**LD patterns and recombination hotspots**

Fine-scale LD maps in and around the APOE were constructed using LOAD (\( N = 547 \)) and control subjects (\( N = 715 \)) to further characterize the significant SNPs (Table 2). We first surveyed LD patterns across the 200-kb genomic region in and around the APOE. For this analysis, 172 SNPs, including the SNP 19AyD0326, with

---

**Fig. 3.** LD maps for the significant 36 SNPs plus the APOE SNP rs7412. \( D' \) was computed using LOAD (A, \( N = 547 \)), control (B, \( N = 715 \)), and LOAD + Control (C, \( N = 1262 \)) subjects. The two APOE SNPs, rs429358 and rs7412, are enclosed by rectangles.
These SNPs were in a relatively strong LD region in both LOAD (Fig. 3A) and control subjects (Fig. 3B). In particular, the LD between the most significant APOE SNP, rs429358, and the remaining 35 SNPs was very strong. The mean values of $D^\prime$ and $r^\prime$ were 0.91 (SD, 0.13; 95% CI, 0.87–0.95; range, 0.53–1.00) and 0.32 (SD, 0.22; 95% CI, 0.25–0.40; range, 0.01–0.88) in LOAD subjects, and 0.91 (SD, 0.14; 95% CI, 0.87–0.96; range, 0.44–1.00) and 0.22 (SD, 0.18; 95% CI, 0.16–0.28; range, 0.01–0.83) in controls, respectively.

Case-control haplotype analysis

Since combinations of multiple SNPs might increase the LOAD risk, we conducted case-control haplotype analysis, using the 36 significant SNPs plus the APOE SNP rs7412 with the sample set All. Among the 18 haplotypes estimated, three carrying the APOE-e4 allele (H3, H7 and H12) exhibited significance and caused an increased risk for LOAD (Table 3). Haplotype H14 involving the APOE-e2 allele showed a protective effect as to LOAD. We did not observe a more significant haplotype than APOE SNP rs429358 ($P_{\text{unadj}}$-allele-value = 4.32E–34) (Tables 2 and 3).

Discussion

Here we carried out a case-control association study using high-density SNPs to discover a disease susceptibility gene(s) for LOAD in and around the APOE in Japanese. It was found that 36 SNPs are strongly associated with LOAD (Fig. 1 and Table 2). These SNPs, which were in strong LD (Fig. 3), are located in a genomic region encompassing four genes, PVRL2, TOMM40, APOE and APOC1. Recombination rate estimation revealed that the associated region is firmly sandwiched between two recombination hotspots, HS1 and HS2 (Fig. 1B). These data raise the possibility that all four genes could be genetically promising candidates for a predisposition to LOAD.

Further biological studies will be helpful to determine which genes are actually involved in the LOAD pathogenesis. To our knowledge, there has only been one paper concerning the genetic association of PVRL2 with LOAD, i.e. that of Martin et al. [14]. Expression of PVRL2 has been detected in many organs [15] including brain (http://genecards.ccbzb.re.kr/cgi-bin/carddisp.pl?gene=PVRL2) and in several neuronal cell lines [16,17]. PVRL2 serves as an entry mediator for herpes simplex virus type 1 (HSV-1) [16,18,19], and its DNA has been detected in the brains of elderly normal subjects and neuronal disease patients including ones with AD [20,21]. HSV-1 is present in a high proportion of AD brains with the APOE-e4 allele in comparison with in non-AD ones [22]. Although there is no direct evidence that PVRL2 is involved in LOAD development, this gene seems to be an interesting candidate for LOAD in terms of environmental pathogens such as viruses.

Through several genetic approaches, TOMM40 has been shown to be strongly associated with LOAD in Caucasians [23-25]. TOMM40 encodes a channel protein that is essential for formation of a component (TOM complex) of the translocase of the mitochondrial outer membrane [26]. Recently, it was shown that TOMM40 is colocalized with mitochondrial attached amyloid beta (Aβ) precursor protein (APP) in the brains of AD patients, especially in ones with the APOE-e3–e4 genotype, but not in those of age-matched control subjects: APOE genotype-dependent arrest of APP transport through the TOM complex possibly causes a mitochondrial dysfunction, resulting in LOAD development [27]. Thus, TOMM40 is considered to be a candidate for the LOAD pathogenesis, which was also pointed out by Li et al. [11].

Table 3  Case-control haplotype analysis of the 36 significant SNPs plus APOE SNP rs7412

<table>
<thead>
<tr>
<th>ID</th>
<th>Haplotype</th>
<th>APOE</th>
<th>Frequencies</th>
<th>Chi-square</th>
<th>Chi-squared</th>
<th>Permutation</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All</td>
<td>LOAD</td>
<td>Control</td>
<td>P&lt;value</td>
<td>P&lt;value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>ACGGCGCTTCGTCATCCGAGATCCGGAAGA</td>
<td>c3</td>
<td>0.291</td>
<td>0.275</td>
<td>0.337</td>
<td>10.153</td>
<td>0.0014</td>
<td>0.0352</td>
</tr>
<tr>
<td>H2</td>
<td>gtaGGCTTCCTCATCGATCGCGCGATCGGCGA</td>
<td>c3</td>
<td>0.073</td>
<td>0.073</td>
<td>0.081</td>
<td>0.54</td>
<td>0.4624</td>
<td>1.0000</td>
</tr>
<tr>
<td>H3</td>
<td>gtaagaagTCACTGATCGCGCGATCGGCGGAC</td>
<td>c4</td>
<td>0.069</td>
<td>0.042</td>
<td>0.064</td>
<td>41.971</td>
<td>9.26E-11</td>
<td>1.0E–4</td>
</tr>
<tr>
<td>H4</td>
<td>gtcGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.043</td>
<td>0.035</td>
<td>0.055</td>
<td>5.307</td>
<td>0.0212</td>
<td>0.3785</td>
</tr>
<tr>
<td>H5</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.038</td>
<td>0.038</td>
<td>0.043</td>
<td>0.412</td>
<td>0.5212</td>
<td>1.0000</td>
</tr>
<tr>
<td>H6</td>
<td>gtaagaagTCATCGGATCGCGCGATCGGCGGAC</td>
<td>c4</td>
<td>0.031</td>
<td>0.054</td>
<td>0.017</td>
<td>24.842</td>
<td>6.22E-07</td>
<td>1.0E–4</td>
</tr>
<tr>
<td>H7</td>
<td>gtaagaagTCACTGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.030</td>
<td>0.028</td>
<td>0.035</td>
<td>0.789</td>
<td>0.3744</td>
<td>1.0000</td>
</tr>
<tr>
<td>H8</td>
<td>gtaagaagTCACTGATCGCGCGATCGGCGGAC</td>
<td>c4</td>
<td>0.029</td>
<td>0.030</td>
<td>0.031</td>
<td>0.052</td>
<td>0.8191</td>
<td>1.0000</td>
</tr>
<tr>
<td>H9</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.026</td>
<td>0.028</td>
<td>0.027</td>
<td>0.023</td>
<td>0.8791</td>
<td>1.0000</td>
</tr>
<tr>
<td>H10</td>
<td>gtaagaagTCACTGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.025</td>
<td>0.021</td>
<td>0.012</td>
<td>2.73</td>
<td>0.0985</td>
<td>0.9308</td>
</tr>
<tr>
<td>H11</td>
<td>gtaagaagTCACTGATCGCGCGATCGGCGGAC</td>
<td>c4</td>
<td>0.024</td>
<td>0.041</td>
<td>0.013</td>
<td>17.826</td>
<td>2.42E-05</td>
<td>0.0005</td>
</tr>
<tr>
<td>H12</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.018</td>
<td>0.017</td>
<td>0.021</td>
<td>0.649</td>
<td>0.4206</td>
<td>1.0000</td>
</tr>
<tr>
<td>H13</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.016</td>
<td>0.007</td>
<td>0.024</td>
<td>9.696</td>
<td>0.0018</td>
<td>0.0457</td>
</tr>
<tr>
<td>H14</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.014</td>
<td>0.012</td>
<td>0.017</td>
<td>1.203</td>
<td>0.2726</td>
<td>1.0000</td>
</tr>
<tr>
<td>H15</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.010</td>
<td>0.013</td>
<td>0.010</td>
<td>0.412</td>
<td>0.5211</td>
<td>1.0000</td>
</tr>
<tr>
<td>H16</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.010</td>
<td>0.014</td>
<td>0.008</td>
<td>2.517</td>
<td>0.1126</td>
<td>0.9561</td>
</tr>
<tr>
<td>H17</td>
<td>gttGGCTTCATGCGATCGCGCGATCGGCGGAC</td>
<td>c3</td>
<td>0.010</td>
<td>0.014</td>
<td>0.008</td>
<td>1.949</td>
<td>0.1627</td>
<td>0.9935</td>
</tr>
</tbody>
</table>

| Others | 0.211 | 0.164 | 0.156 |
| Sum | 1.000 | 1.000 | 1.000 |

a Major and minor alleles are indicated in upper and lower case letters, respectively. The SNP order, from left to right, follows the genomic position. The APOE SNPs, rs429358 (T/c) and rs7412 (C/t), are indicated in bold.

b 10,000 iterations.
APOC1 is involved in lipoprotein/lipid homeostasis, of which failure leads to AD [28]. Immunohistochemical staining of human hippocampal tissues revealed that APOC1 is produced in astrocytes, but not in neurons [29], and is colocalized with Aβ40 and APOE in senile plaques [30]. APOC1 inhibits Aβ1–40 aggregation and causes Aβ1–40 oligomer-induced neuronal death [30]. A study on transgenic mice demonstrated that APOC1 impairs learning and memory functions [30]. Judging from these findings, APOC1 may contribute to the LOAD etiology.

In conclusion, through a genetic association study on in and around the APOE, we found that three additional genes, PVRL2, TOMM40 and APOC1, exhibited significant association with LOAD (Table 2). By means of only genetic approaches, it is difficult to determine which gene is the most likely disease susceptibility gene for LOAD in the associated region. At the present it is reasonable to suppose that all four genes could contribute to the LOAD pathogenesis from a functional point of view of these genes.

Materials and methods

Subjects

Blood samples were collected by the Japanese Genetic Study Consortium for AD (JGSCAD), organized in 2000: the members are listed in our recent publications [5,6]. All subjects were Japanese. Probable AD cases met the criteria of the National Institute of Neurological and Communicative Disorders, and Stroke-Alzheimer’s Disease and Related Disorders (NINCDS-ADRDA) [31]. Control subjects had no signs of dementia and lived in an unassisted manner in the local community. Age at onset (AAO) is here defined as the age at which the family and/or individuals first noted cognitive problems while working or in daily activities. The Mini-Mental State Examination (MMSE) and Clinical Dementia Rating and/or Function Assessment Staging were used for evaluation of the cognitive impairment. The basic demographics of the LOAD patients and the non-demented control subjects are presented in Table 1. A total of 1262 subjects comprising 547 LOAD patients (female, 71.7%) and 715 controls (female, 54.4%), which is referred to as overall sample set in this paper, was used. The mean AAO ± SD in LOAD patients was 73.0 ± 6.9 (range, 60–93). The mean age at examination (AAE) ± SD in control subjects was 73.0 ± 6.8 (range, 60–93). There was no significant difference between AAO in LOAD patients and AAE in control subjects with the unpaired Student’s t-test (P-value, 0.965). The MMSE score in LOAD patients (mean ± SD, 16.9 ± 7.7) was significantly lower than that in the controls (mean ± SD, 27.4 ± 2.6): unpaired Student’s t-test, P-value < 0.0001.

To construct three sub-sample sets, the subjects were stratified as to the APOE-ε4 or APOE-ε3 carrier status: Negative-ε4 (LOAD, 298; control, 590), Positive-ε4 (LOAD, 249; control, 125), and ε3*3 (LOAD, 279; control, 529).

The present study was approved by the Institutional Review Board of Niigata University, and by all participating institutes. Informed consent was obtained from all controls and appropriate proxies for patients, and all samples were anonymously analyzed for genotyping.

Extraction and quantification of genomic DNA

Genomic DNA was extracted from peripheral blood with a QIAamp® DNA Blood Maxi Kit (Qiagen), and was fluorometrically quantified with Quanti-IT™ PicoGreen® reagent (Invitrogen) according to the manufacturer’s protocol.

SNPs and genotyping

SNP information was obtained from four open databases: NCBI dbSNP (Build 126; http://www.ncbi.nlm.nih.gov/SNP/), Interna-

tional HapMap Project (Rel#21a/phaseII on NCBI Build 35 assembly, dbSNP Build 125; http://www.hapmap.org/index.html), Ensemble Human (Version 45 on NCBI Build 36; http://www. ensembl.org/Homo_sapiens/), and Celera myScience (Version R27g on NCBI Build 35; http://myscience.appliedbiosystems.com/). From these databases, 333 SNPs located in a 200 kb genomic region including the APOE on chromosome 19q were initially selected in the present study. Among them, 76 SNPs were excluded since some did not satisfy the design criteria for the TaqMan® SNP genotyping assay and others were not correctly positioned in the genomic region. Finally, we used 257 SNPs for genotyping: mean inter-marker distance ± SD, 0.78 ± 1.04 kb; 95% CIs, 0.66–0.91; range, 7–6,099 bp (Supplementary Table 1).

SNP genotyping was conducted with an ABI PRISM 7900HT system using TaqMan® SNP Genotyping Assays (Applied Biosystems). To evaluate the genotyping quality of each SNP, a HWE test was performed in 715 control subjects (see also under Statistical analysis).

Sequencing

The APOE genotypes of all subjects were determined by means of sequencing with a pair of primers, as described previously [5]. In order to explore novel mutations or SNPs in and around TOMM40 and PVRL2, all exons and their exon–intron boundaries, and the 3’ upstream regions of these genes were sequenced with 18 pairs of primers in 48 individuals (24 LOAD patients and 24 control subjects). The primer sequences are available on request. Each sequence analysis was carried out by means of direct cycle sequencing with an ABI PRISM 3100 genetic analyzer and a BigDye® Terminator v3.1 kit (Applied Biosystems).

Statistical analysis

To determine the disease susceptibility SNPs for LOAD, the allelic association at each of the 172 SNPs showing polymorphism and HWE P-value ≥ 0.01 was assessed by means of a chi-squared test based on a 2 × 2 contingency table in comparison with the allele frequencies in LOAD patients and control subjects. Bonferroni correction was applied to correct for multiple testing (172 tests). A P-value of 0.0003 (= 0.05/172) was set as the significance threshold for the allelic association test. OR with 95% CIs of each SNP was computed from the 2 × 2 allelic contingency table as an estimate of the relative risk for the disease. For calculation of the P_unadj-allele-value and OR with 95% CIs, the following four sample sets were used: All, Negative-ε4, Positive-ε4 and ε3*3. For all statistical analyses mentioned above, we used SNPAlalyze software version 6.0.1 (DYNACOM, Japan; http://www.dynacom.co.jp/).

Using the Haploview version 4.0 software (http://www.broad.mit.edu/mpg/haplovie/index.php) [32], we performed a HWE test, calculation of pair-wise LD measures (|D'| and r-square), haplotype estimation based on an expectation-maximization algorithm, and case-control haplotype analysis (number of permutations, 10,000) which involved haplotypes with frequency ≥ 0.01.

The statistical significance was set at P-value < 0.05.

Recombination rate estimation

Recombination rates were calculated using software LDhat version 2.1 (http://www.stats.ox.ac.uk/~mcevean/LDhat/), which involves a Bayesian reversible-jump Markov Chain Monte Carlo (rjMCMC) scheme under the crossing-over model [33–35]. They were averaged, based on the results of 10-times computing for a randomly selected 96 samples from the entire sample set All. Every computation was run with the parameters recommended in the accompanying documentation: θ = 0.01 and rjMCMC = 1,000,000 iterations (http://www.stats. ox.ac.uk/~mcevean/LDhat/manual.pdf).
Conflict of interest statement
None declared.

Acknowledgments

We wish to express our gratitude to the patients with AD and their families, and the control individuals for their participation in this study, without whom this genetic work would have been impossible. We are very grateful to the members of JGSCAD for the collection of blood samples: the members were listed in our previous publications [5,6]. We would like to also thank K. Horigome, N. Yahata, K. Takadono, A. Kitamura, M. Hirose, Y. Satoh, A. Hirokawa, T. Hosino, S. Yanagihara, and M. Saitoh for their technical assistance, and N. J. Halewood for critical reading of the manuscript.

This study was supported by a Grant-in-Aid for Scientific Research on Priority Areas (C) – Advanced Brain Science Project – from the Ministry of Education, Culture, Sports, Science and Technology, Japan (R.K. and Y.I.), and a KAKENHI (Grant-in-Aid for Scientific Research) on Priority Areas – Comprehensive Genomics – from the Ministry of Education, Culture, Sports, Science and Technology of Japan (R.K.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2009.01.003.

References