



Contents lists available at ScienceDirect

Genomics

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

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 Urakami^g, Hideo Kimura^h, Akiyoshi Kakitaⁱ, Hitoshi Takahashi^j, Shoji Tsuji^k, Ichiro Kanazawa^h, Yasuo Ihara^l, Shoji Odani^m, Ryoza Kuwano^{a,*} and the Japanese Genetic Study Consortium for Alzheimer Disease¹

^a Department of Molecular Genetics, Bioresource Science Branch, Center for Bioresources, Brain Research Institute, Niigata University, Niigata 951-8585, Japan

^b Department of Geriatrics and Gerontology, Institute of Development, Aging and Cancer, Tohoku University, Sendai 980-8575, Japan

^c Department of Psychiatry, Institute of Clinical Medicine, University of Tsukuba, Tsukuba 305-8575, Japan

^d Imagawa Clinic, Fukushima-ku, Osaka 553-0003, Japan

^e Department of Neurology, Neuroscience and Biophysiological Science, Hirosaki University, School of Medicine, Hirosaki 036-8562, Japan

^f Division of Clinical Research, Kurihama Alcoholism Center, National Hospital Organization, Yokosuka 239-0841, Japan

^g Division of Environment and Health Science, Department of Biological Regulation, Faculty of Medicine, Tottori University, Yonago 683-8503, Japan

^h National Center for Neurology and Psychiatry, Kodaira 187-8502, Japan

ⁱ Department of Pathological Neuroscience, Brain Research Institute, Niigata University, Niigata 951-8585, Japan

^j Department of Pathology, Brain Research Institute, Niigata University, Niigata 951-8585, Japan

^k Department of Neurology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

^l Department of Medical Life Systems, Doshisha University, Kyoto 619-0225, Japan

^m Life Science Course, Graduate School of Science and Technology, Niigata University, Niigata 950-2181, Japan

ARTICLE INFO

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

ABSTRACT

The $\epsilon 4$ 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 ($A\beta$) 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 (CGC → 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 (0.0–0.15), 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

* Corresponding author. 1-757 Asahimachi, Chuo-ku, Niigata 951-8585, Japan. Fax: +81 25 227 0793.

E-mail address: ryosun@bri.niigata-u.ac.jp (R. Kuwano).

¹ Members listed in our recent publications [5,6].

Table 1
Subject information

	LOAD	Control
No. of subjects	547	715
Frequency of female	0.717	0.544
AAO/AAE (mean \pm SD)	73.0 \pm 6.9	73.0 \pm 6.8
<i>APOE</i> genotype		
2*2	0	4
2*3	19	57
2*4	6	7
3*3	279	529
3*4	196	117
4*4	47	1
<i>APOE</i> - ϵ 4 allele frequency	0.269	0.087

APOE- ϵ 4 alleles are 0.05, 0.86 and 0.09, respectively [5,6]. Compared to the *APOE*- ϵ 3 allele, the *APOE*- ϵ 4 allele is a strong risk factor with an odds ratio (OR) of 2.0–4.0, the *APOE*- ϵ 2 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*- ϵ 4 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*- ϵ 4 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 nondemented 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*),

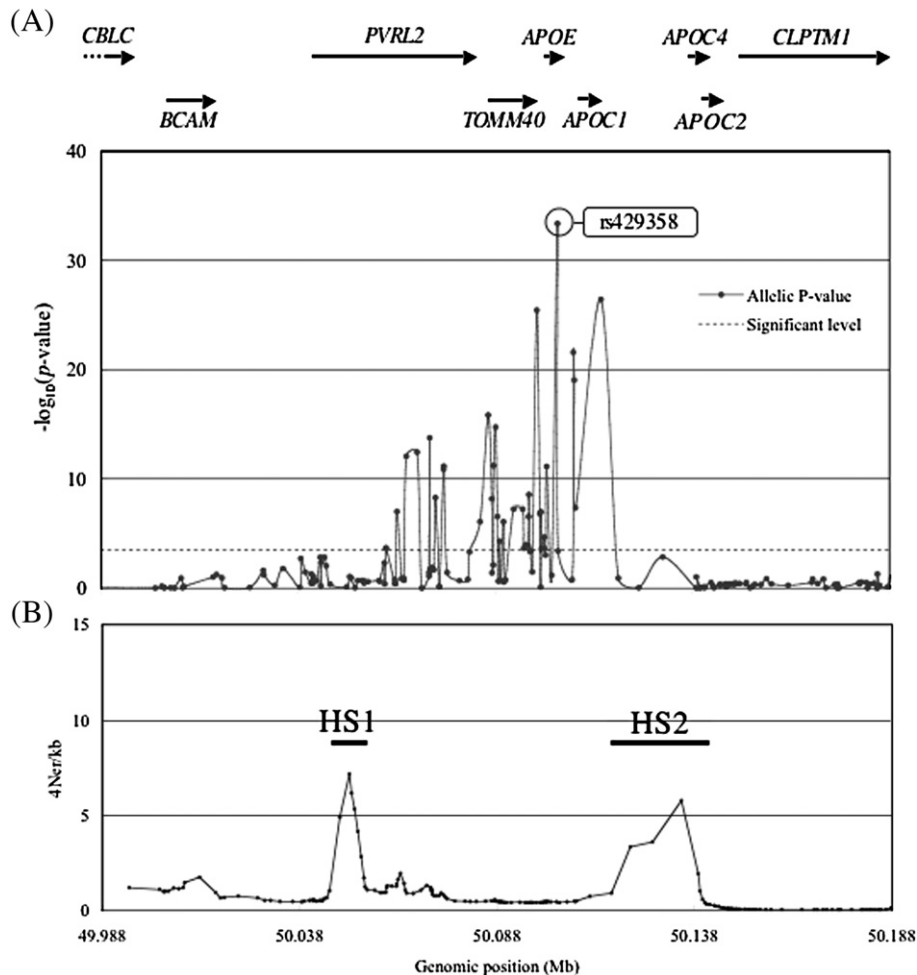


Fig. 1. Genomic region associated with LOAD. (A) Plot of $-\log_{10} P_{\text{unadj-allele}}$ -values of 171 SNPs. Horizontal arrows indicate the transcriptional orientations of individual genes. The dotted line indicates the significant threshold of the $P_{\text{unadj-allele}} = 0.0003$ (Bonferroni correction). The most significant SNP, rs429358 ($P_{\text{unadj-allele}} = 4.32E-34$), 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[®] 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 ($\epsilon 2$, 25; $\epsilon 3$, 773; $\epsilon 4$, 296) and controls ($\epsilon 2$, 72; $\epsilon 3$, 1232; $\epsilon 4$, 126). As expected, a significant difference was observed: degree of freedom [df]=2, chi-square value = 152.34, P -value = 3.06E–34. We further computed an estimate of the relative risk for LOAD using a 2×2 contingency table. Compared to the *APOE*- $\epsilon 3$ allele, the *APOE*- $\epsilon 4$ one exhibited a strong

risk effect for LOAD ($df=1$, chi-square value = 141.15, P -value = 1.49E–32; OR = 3.74, 95% confidence intervals [CIs] = 2.98–4.20) and the *APOE*- $\epsilon 2$ 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/com/gen/alzgene/>].

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 \pm standard deviation (SD) was 1.18 ± 1.79 kb (95% CIs, 0.91–1.45; range, 7–13,801 bp).

For these 171 SNPs, an allelic contingency table (2×2)-based chi-squared test was performed using the *All* sample set to identify SNPs

Table 2
Case-control study of the 36 significant SNPs

SNP ID	Position(bp) ^a	Allele ^b	MAF ^c	GSR ^d	Number of genotypes						Number of alleles				Allelic P -value	OR	95% CIs
					LOAD			Control			LOAD		Control				
					MM	Mm	mm	MM	Mm	mm	M	m	M	m			
rs419010	50060160	A/g	0.493	98.6	112	278	150	206	348	150	502	578	760	648	2.10E–04	1.35	(1.15–1.58)
rs394221	50060264	C/t	0.487	92.5	149	269	91	153	325	180	567	451	631	685	2.04E–04	0.73	(0.62–0.86)
rs4803766	50063008	G/a	0.434	98.2	136	273	130	265	328	107	545	533	858	542	9.16E–08	1.55	(1.32–1.82)
rs395908	50065579	G/a	0.280	98.2	228	240	70	414	261	26	696	380	1089	313	9.24E–13	1.90	(1.59–2.27)
rs519113	50069307	C/g	0.273	98.3	237	226	73	423	258	23	700	372	1104	304	3.77E–13	1.93	(1.61–2.31)
rs412776	50071356	G/a	0.257	97.9	245	230	65	440	237	19	720	360	1117	275	1.81E–14	2.03	(1.69–2.44)
rs3865427	50072801	C/a	0.209	98.3	299	198	43	481	206	14	796	284	1168	234	5.27E–09	1.78	(1.47–2.16)
rs3852860	50074806	T/c	0.321	96.9	191	268	77	359	292	36	650	422	1010	364	1.34E–11	1.80	(1.52–2.14)
rs3852861	50074901	T/g	0.318	98.7	196	261	80	375	294	39	653	421	1044	372	6.99E–12	1.81	(1.53–2.15)
rs6857	50084094	C/t	0.180	97.0	316	184	26	504	183	11	816	236	1191	205	7.88E–07	1.68	(1.37–2.07)
I9AyD0326	50086176	T/c	0.194	99.3	292	212	40	524	176	9	796	292	1224	194	1.50E–16	2.31	(1.89–2.83)
rs157580	50087106	A/g	0.467	98.2	192	262	85	162	351	187	646	432	675	725	6.91E–09	0.62	(0.53–0.73)
rs157581	50087554	T/c	0.329	97.1	194	251	90	363	281	47	639	431	1007	375	6.31E–12	1.81	(1.53–2.15)
rs157582	50088059	C/t	0.292	97.9	212	240	80	417	250	36	664	400	1084	322	1.87E–15	2.03	(1.70–2.42)
rs157583	50088513	G/t	0.043	96.8	465	57	7	659	34	0	987	71	1352	34	2.70E–07	2.86	(1.89–4.34)
rs1160983	50089069	G/a	0.048	98.4	511	30	0	616	82	3	1052	30	1314	88	4.68E–05	0.43	(0.28–0.65)
rs157587	50090046	A/g	0.043	97.5	467	58	6	668	36	0	984	70	1372	36	7.79E–07	2.71	(1.80–4.09)
rs283817	50092615	T/g	0.044	98.6	468	61	7	673	35	0	997	75	1381	35	5.43E–08	2.97	(1.97–4.47)
rs573199	50094959	C/t	0.038	97.8	472	52	7	675	28	0	996	66	1378	28	5.73E–08	3.26	(2.08–5.11)
rs1160985	50095252	C/t	0.266	98.7	323	192	28	346	300	57	838	248	986	414	2.12E–04	0.70	(0.59–0.85)
rs760136	50095698	A/g	0.265	97.5	323	191	27	338	294	57	837	245	970	408	1.03E–04	0.70	(0.58–0.84)
rs741780	50096271	T/c	0.269	97.9	314	193	26	340	306	57	821	245	986	420	1.31E–04	0.70	(0.58–0.84)
rs394819	50096419	G/t	0.043	98.7	473	58	7	672	35	0	1004	72	1379	35	2.76E–07	2.83	(1.87–4.27)
rs405697	50096531	G/a	0.496	98.4	98	276	167	216	348	137	472	610	780	622	2.90E–09	1.62	(1.38–1.90)
rs10119	50098513	C/t	0.229	99.1	249	225	66	509	189	13	723	357	1207	215	3.62E–26	2.77	(2.29–3.36)
rs446037	50099277	G/t	0.044	98.6	468	60	7	674	34	1	996	74	1382	36	1.45E–07	2.85	(1.90–4.28)
rs434132	50099560	C/g	0.043	99.0	474	59	7	675	34	0	1007	73	1384	34	9.64E–08	2.95	(1.95–4.47)
rs7259620	50099628	G/a	0.264	98.3	319	189	27	349	301	56	827	243	999	413	2.54E–04	0.71	(0.59–0.85)
rs449647	50100404	A/t	0.049	97.6	457	70	2	658	44	1	984	74	1360	46	2.14E–05	2.22	(1.53–3.24)
rs769446	50100468	T/c	0.036	99.4	522	22	0	647	60	4	1066	22	1354	68	2.28E–04	0.41	(0.25–0.67)
rs440446	50101007	C/g	0.426	93.9	125	270	129	261	319	81	520	528	841	481	7.67E–12	1.78	(1.51–2.09)
rs429358	50103781	T/c	0.167	100.0	298	202	47	590	124	1	798	296	1304	126	4.32E–34	3.84	(3.06–4.81)
rs7256200	50107775	G/t	0.128	98.9	348	164	27	608	101	0	860	218	1317	101	2.74E–22	3.31	(2.57–4.25)
rs483082	50108018	G/t	0.208	97.1	274	213	51	504	172	12	761	315	1180	196	9.58E–20	2.49	(2.04–3.04)
rs584007	50108318	G/a	0.439	98.2	129	279	130	248	358	95	537	539	854	548	4.44E–08	1.56	(1.33–1.84)
rs4420638	50114786	A/g	0.203	98.1	297	199	41	565	135	1	793	281	1265	137	3.78E–27	3.27	(2.62–4.09)

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

^a Genomic positions are according to NCBI build 36.

^b Major and minor alleles are indicated in upper and lower case letters, respectively.

^c MAF, minor allele frequency.

^d GSR, genotyping success rate.

associated with LOAD. Thirty-five of these SNPs exhibited significance: range of unadjusted allelic P -values ($P_{\text{unadj-allele-value}}$), $4.32E-34$ – $2.00E-4$ (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, *Positive-ε4* (LOAD, 247; control, 123), *Negative-ε4* (LOAD, 296; control, 586), and $\epsilon3^*3$ (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

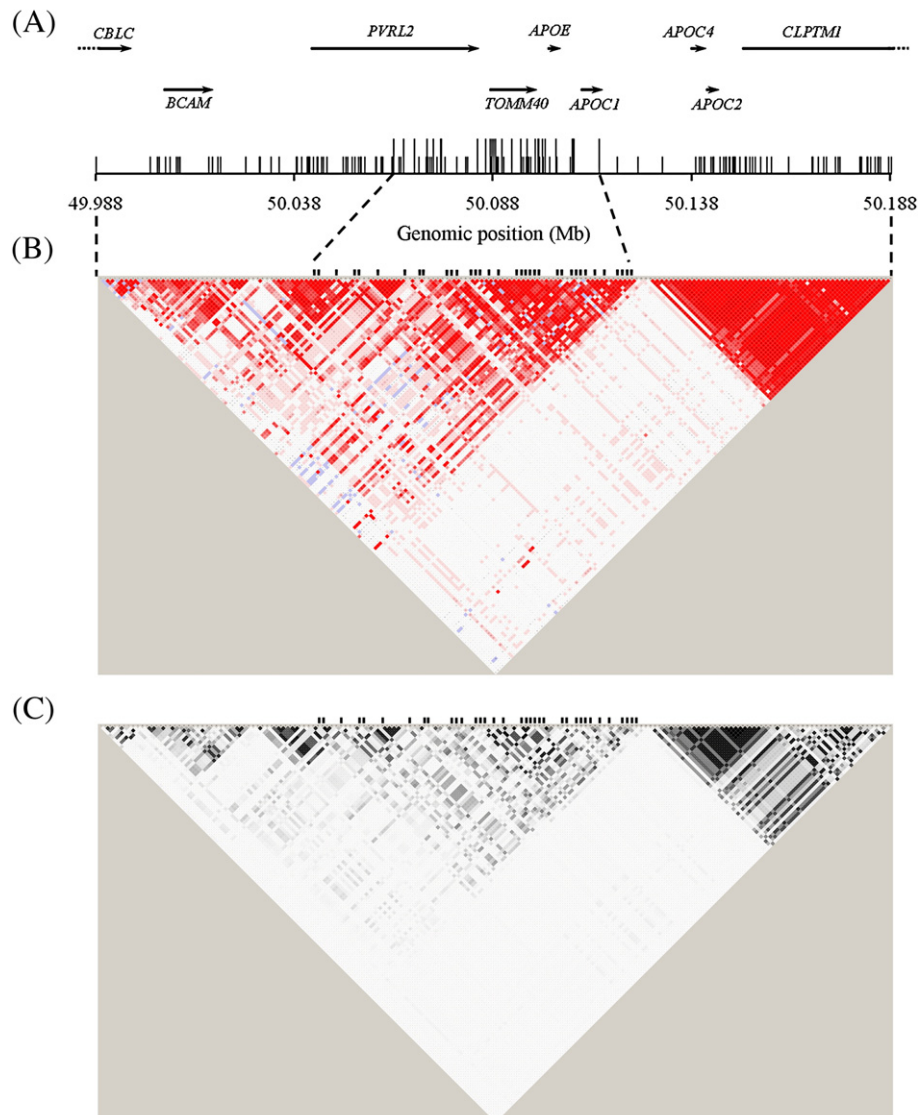


Fig. 2. LD mapping. (A) Horizontal arrows indicate the transcriptional orientations of individual genes. Vertical lines indicate the 171 SNPs, described under Results. The 35 significant SNPs are indicated by long vertical lines. $|D'|$ and r -square are indicated in (B) and (C), respectively. The thick vertical lines in (B) and (C) indicate the significant SNPs.

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, 1.50×10^{-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

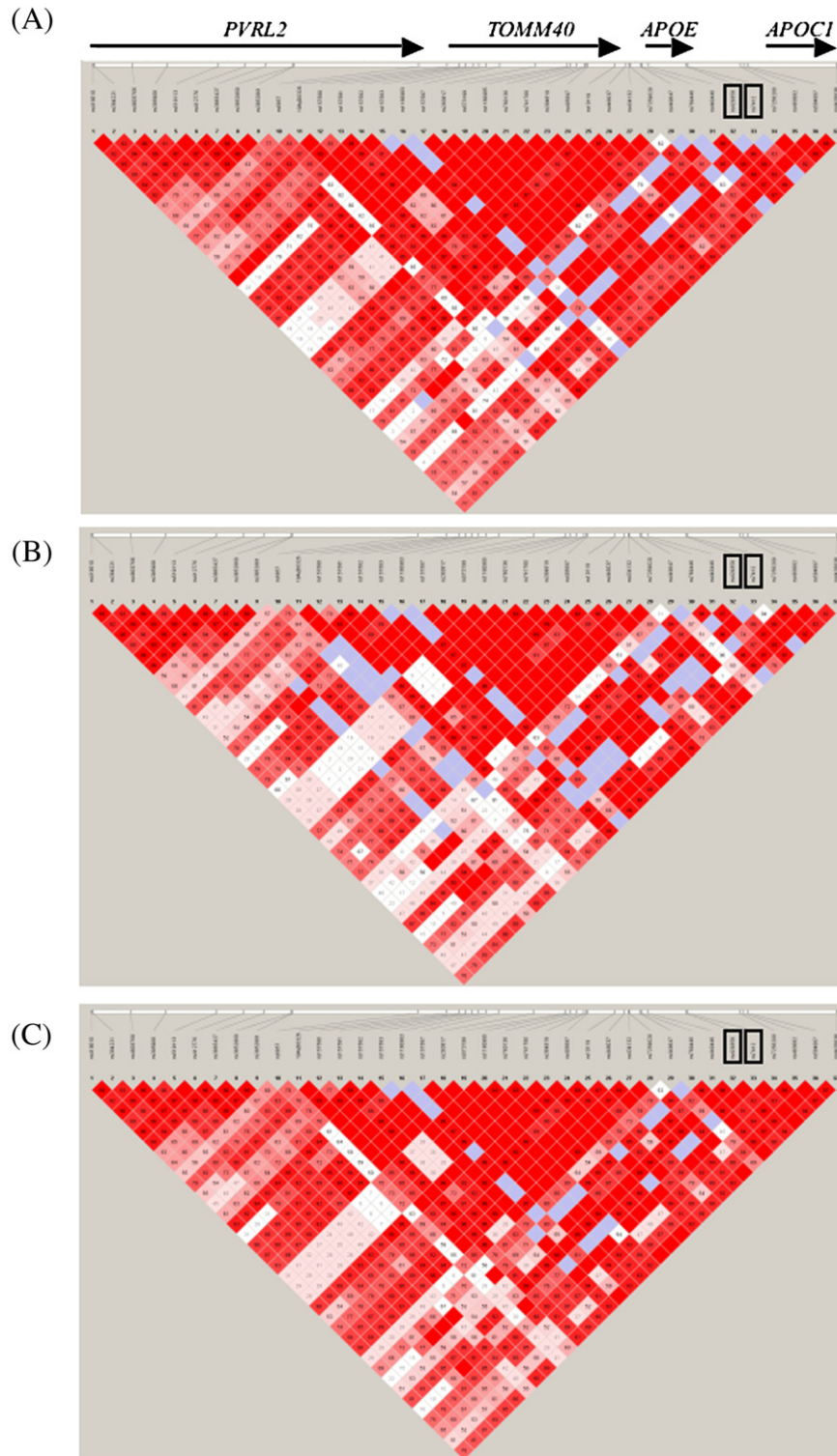


Fig. 3. LD maps for the significant 36 SNPs plus the *APOE* SNP rs7412. $|D'|$ was computed using LOAD (A, $N=547$), control (B, $N=715$), and LOAD + Control (C, $N=1262$) subjects. The two *APOE* SNPs, rs429358 and rs7412, are enclosed by rectangles.

MAF ≥ 0.01 , were used. The mean inter-marker distance \pm SD was 1.17 ± 1.78 kb (95% CIs, 0.9–1.44; range, 7–13,801 bp). We observed a highly structured single LD block spanning about 48.49 kb downstream of the *APOC1*, which encompassed three genes, i.e. *APOC4*, *APOC2* and *CLPTM1* (Fig. 2B and Supplementary Fig. 2). A solid LD pattern was not detected for the *BCAM-APOC1* region extending roughly for 117.5 kb, suggesting the possibility that recombination has frequently occurred in this region during human evolution.

In order to determine whether or not there are recombination hotspots in and around the *APOE*, we estimated the recombination rates. Two major recombination hotspots were identified in intron 1 of *PVRL2* (HS1), and an intergenic region between *APOC1* and *APOC4* (HS2) (Fig. 1B). There were no remarkable differences in the positions of these hotspots between LOAD and control subjects (data not shown). All 36 significant SNPs (Table 2) were found to be sandwiched between HS 1 and HS2, spanning approximately 54.6 kb (Fig. 1B). No definite recombination hotspots were detected in the solid LD region extending from *APOC4* to *CLPTM1* (Figs. 1B and 2B).

We carried out LD analysis using only the significant 36 SNPs. 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 $|D'|$ and r^2 were 0.91 (SD, 0.13; 95% CIs, 0.87–0.95; range, 0.53–1.00) and 0.32 (SD, 0.22; 95% CIs, 0.25–0.40; range, 0.01–0.88) in LOAD subjects, and 0.91 (SD, 0.14; 95% CIs, 0.87–0.96; range, 0.44–1.00) and 0.22 (SD, 0.18; 95% CIs, 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-ε4* allele (H3, H7 and H12) exhibited significance and caused an increased risk for LOAD (Table 3). Haplotype H14 involving the *APOE-ε2* 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.cccb.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-ε4* 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 mitochondrially attached amyloid beta ($A\beta$) precursor protein (APP) in the brains of AD patients, especially in ones with the *APOE-ε3*4* 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

ID	Haplotype ^a	APOE	Frequencies			Chi-square	Chi-squared P-value	Permutation P-value ^b	OR	95% CIs
			All	LOAD	Control					
H1	ACGGCGCTTCTgTCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.291	0.275	0.337	10.153	0.0014	0.0352	0.75	(0.62–0.89)
H2	gtaCGGCTTCTATCGGATCtgcGGCGCaATgTCGGGA	$\epsilon 3$	0.073	0.073	0.081	0.54	0.4624	1.0000	0.89	(0.65–1.21)
H3	gtaagaacgtCActGGATCCATGGtGCGATgCttGg	$\epsilon 4$	0.069	0.113	0.042	41.971	9.26E-11	<1.0E-4	2.89	(2.07–4.03)
H4	gtGGCGCTTCTgTCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.043	0.035	0.055	5.307	0.0212	0.3785	0.62	(0.41–0.94)
H5	gtaCGGCTTCTgTCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.038	0.038	0.043	0.412	0.5212	1.0000	0.87	(0.57–1.33)
H6	ACGGCGCTTCTgTCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.032	0.030	0.037	0.849	0.3569	1.0000	0.81	(0.51–1.28)
H7	gtaagaCcgCTActtGggtCATtGttGtTgcCGtGg	$\epsilon 4$	0.031	0.054	0.017	24.842	6.22E-07	<1.0E-4	3.36	(2.03–5.56)
H8	gtaagaacgtCActGGATCCATGGtGCGATCTCGGgA	$\epsilon 3$	0.030	0.028	0.035	0.789	0.3744	1.0000	0.80	(0.50–1.30)
H9	ACGGCGCTTCTATCGGATCtgcGGCGCaATgTCGGGA	$\epsilon 3$	0.029	0.030	0.031	0.052	0.8191	1.0000	0.94	(0.58–1.53)
H10	ACGGCGCTTCTATCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.026	0.028	0.027	0.023	0.8791	1.0000	1.04	(0.63–1.72)
H11	gtaagaacgtCTATCGGATCtgcGGCGCaATgTCGGGA	$\epsilon 3$	0.025	0.021	0.032	2.73	0.0985	0.9308	0.64	(0.38–1.09)
H12	gtaagaacgtCActGGATCCATGGtGCGATgCttGg	$\epsilon 4$	0.024	0.041	0.013	17.826	2.42E-05	0.0005	3.21	(1.81–5.69)
H13	gtGGCGCTTCTATCGGATCtgcGGCGCaATgTCGGGA	$\epsilon 3$	0.018	0.017	0.021	0.649	0.4206	1.0000	0.78	(0.42–1.43)
H14	ACGGCGCcgCTActGAtATctgcGGCGCaAcgTtGtGA	$\epsilon 2$	0.016	0.007	0.024	9.696	0.0018	0.0457	0.30	(0.13–0.67)
H15	ACGGCGCTTtTgTCGGATCCATgACGCGATCTCGGgA	$\epsilon 3$	0.014	0.012	0.017	1.203	0.2726	1.0000	0.67	(0.33–1.37)
H16	ACGGCGCTTtActGGATCCATGtGCGATCTCGGgA	$\epsilon 3$	0.010	0.013	0.010	0.412	0.5211	1.0000	1.29	(0.59–2.80)
H17	ACGGCGCcgCTActGAtATctgcGGCGCaATCTCGGgA	$\epsilon 3$	0.010	0.007	0.014	2.517	0.1126	0.9561	0.50	(0.21–1.20)
H18	ACGGCGCcgCTATCGGATCtgcGGCGCaATgTCGGGA	$\epsilon 3$	0.010	0.014	0.008	1.949	0.1627	0.9935	1.76	(0.79–3.92)
	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 A β and APOE in senile plaques [30]. APOC1 inhibits A β _{1–40} aggregation and causes A β _{1–40} 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 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 *All* in this paper, was used. The mean AAO \pm SD in LOAD patients was 73.0 \pm 6.9 (range, 60–93). The mean age at examination (AAE) \pm SD in control subjects was 73.0 \pm 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 \pm SD, 16.9 \pm 7.7) was significantly lower than that in the controls (mean \pm SD, 27.4 \pm 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- ϵ 4 or APOE- ϵ 3 carrier status: *Negative- ϵ 4* (LOAD, 298; control, 590), *Positive- ϵ 4* (LOAD, 249; control, 125), and *ϵ 3*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 Quant-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 \pm SD, 0.78 \pm 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 5' 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 \geq 0.01 was assessed by means of a chi-squared test based on a 2 \times 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 \times 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- ϵ 4*, *Positive- ϵ 4* and *ϵ 3*3*. For all statistical analyses mentioned above, we used SNPalyze 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/haploview/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 \geq 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/~mcvean/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: $\theta = 0.01$ and rjMCMC = 1,000,000 iterations (<http://www.stats.ox.ac.uk/~mcvean/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](https://doi.org/10.1016/j.ygeno.2009.01.003).

References

- [1] W.J. Strittmatter, et al., Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 1977–1981.
- [2] D.E. Schemmel, et al., Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of Apolipoprotein E genotype in late-onset Alzheimer disease, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 9649–9653.
- [3] E.H. Corder, et al., Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families, *Science* 261 (1993) 921–923.
- [4] R.M. Corbo, R. Scacchi, Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a 'thrifty' allele? *Ann. Hum. Genet.* 63 (1999) 301–310.
- [5] R. Kuwano, et al., Dynamin-binding protein gene on chromosome 10q is associated with late-onset Alzheimer's disease, *Hum. Mol. Genet.* 15 (2006) 2170–2182.
- [6] A. Miyashita, et al., Genetic association of CTNNA3 with late-onset Alzheimer's disease in females, *Hum. Mol. Genet.* 16 (2007) 2854–2869.
- [7] L.A. Farrer, et al., Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium, *JAMA* 278 (1997) 1349–1356.
- [8] L. Bertram, M.B. McQueen, K. Mullin, D. Blacker, R.E. Tanzi, Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database, *Nat. Genet.* 39 (2007) 17–23.
- [9] J.V. Pearson, et al., Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies, *Am. J. Hum. Genet.* 80 (2007) 126–139.
- [10] K.D. Coon, et al., A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease, *J. Clin. Psychiatry* 68 (2007) 613–618.
- [11] H. Li, et al., Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease, *Arch. Neurol.* 65 (2008) 45–53.
- [12] O. Belbin, et al., Regulatory region single nucleotide polymorphisms of the apolipoprotein E gene and the rate of cognitive decline in Alzheimer's disease, *Hum. Mol. Genet.* 16 (2007) 2199–2208.
- [13] L.M. Bekris, et al., Multiple SNPs within and surrounding the apolipoprotein E gene influence cerebrospinal fluid apolipoprotein E protein levels, *J. Alzheimer's Dis.* 13 (2008) 255–266.
- [14] E.R. Martin, et al., Analysis of association at single nucleotide polymorphisms in the APOE region, *Genomics* 63 (2000) 7–12.
- [15] F. Eberle, P. Dubreuil, M.G. Mattei, E. Devillard, M. Lopez, The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is the true homolog of the murine MPH gene, *Gene* 159 (1995) 267–272.
- [16] M.S. Warner, et al., A cell surface protein with herpesvirus entry activity (HvE) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus, *Virology* 246 (1998) 179–189.
- [17] D. Shukla, C.L. Rowe, Y. Dong, V.R. Racaniello, P.G. Spear, The murine homolog (Mph) of human herpesvirus entry protein B (HvE) mediates entry of pseudorabies virus but not herpes simplex virus types 1 and 2, *J. Virol.* 73 (1999) 4493–4497.
- [18] M. Lopez, F. Cocchi, L. Menotti, E. Avitabile, P. Dubreuil, G. Campadelli-Fiume, Nectin2alpha (PRR2alpha or HvE) and nectin2delta are low-efficiency mediators for entry of herpes simplex virus mutants carrying the Leu25Pro substitution in glycoprotein D, *J. Virol.* 74 (2000) 1267–1274.
- [19] W.M. Martinez, P.G. Spear, Structural features of nectin-2 (HvE) required for herpes simplex virus entry, *J. Virol.* 75 (2001) 11185–11195.
- [20] L. Gordon, S. McQuaid, S.L. Cosby, Detection of herpes simplex virus (types 1 and 2) and human herpesvirus 6 DNA in human brain tissue by polymerase chain reaction, *Clin. Diagn. Virol.* 6 (1996) 33–40.
- [21] W.R. Lin, M.A. Wozniak, R.J. Cooper, G.K. Wilcock, R.F. Itzhaki, Herpesviruses in brain and Alzheimer's disease, *J. Pathol.* 197 (2002) 395–402.
- [22] R.F. Itzhaki, W.R. Lin, D. Shang, G.K. Wilcock, B. Faragher, G.A. Jamieson, Herpes simplex virus type 1 in brain and risk of Alzheimer's disease, *Lancet* 349 (1997) 241–244.
- [23] E.R. Martin, et al., SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease, *Am. J. Hum. Genet.* 67 (2000) 383–394.
- [24] C.E. Yu, et al., Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: patterns of linkage disequilibrium and disease-marker association, *Genomics* 89 (2007) 655–665.
- [25] A. Grupe, et al., Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants, *Hum. Mol. Genet.* 16 (2007) 865–873.
- [26] N.J. Hoogenraad, L.A. Ward, M.T. Ryan, Import and assembly of proteins into mitochondria of mammalian cells, *Biochim. Biophys. Acta* 1592 (2002) 97–105.
- [27] L. Devi, B.M. Prabhu, D.F. Galati, N.G. Avadhani, H.K. Anandatheerthavarada, Accumulation of amyloid precursor protein in the mitochondrial import channels of human Alzheimer's disease brain is associated with mitochondrial dysfunction, *J. Neurosci.* 26 (2006) 9057–9068.
- [28] C.J. Carter, Convergence of genes implicated in Alzheimer's disease on the cerebral cholesterol shuttle: APP, cholesterol, lipoproteins, and atherosclerosis, *Neurochem. Int.* 50 (2007) 12–38.
- [29] C. Petit-Turcotte, et al., Apolipoprotein C-I expression in the brain in Alzheimer's disease, *Neurobiol. Dis.* 8 (2001) 953–963.
- [30] K. Abildaveyeva, et al., Human apolipoprotein C-I expression in mice impairs learning and memory functions, *J. Lipid Res.* 49 (2008) 856–869.
- [31] G. McKhann, et al., 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 (1984) 939–944.
- [32] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, *Bioinformatics* 21 (2005) 263–265.
- [33] G. McVean, P. Awadalla, P. Fearnhead, A coalescent-based method for detecting and estimating recombination from gene sequences, *Genetics* 160 (2002) 1231–1241.
- [34] G.A. McVean, et al., The fine-scale structure of recombination rate variation in the human genome, *Science* 304 (2004) 581–584.
- [35] A. Auton, G. McVean, Recombination rate estimation in presence of hotspots, *Genome Res.* 17 (2007) 1219–1227.