

1 Research Article

2 **Title**

3 **Association of rare *APOE* missense variants with Alzheimer's disease in the Japanese**
4 **population**

5

6 **Authors**

7 Akinori Miyashita^{a,†,*}, Ai Obinata^{a,†}, Norikazu Hara^{a,†}, Risa Mitsumori^{b,†}, Daita Kaneda^c,
8 Yoshio Hashizume^c, Terunori Sano^d, Masaki Takao^d, Ramil Gabdul Khaev^e, Mari Tada^f,
9 Akiyoshi Kakita^e, Akira Arakawa^g, Maho Morishima^g, Shigeo Murayama^{g,h}, Yuko Saito^g,
10 Hiroyuki Hatsuta^{i,j}, Tomoyasu Matsubara^g, Akio Akagi^k, Yuichi Riku^k, Hiroaki Miyahara^k, Jun
11 Sone^k, Mari Yoshida^k, Haruyasu Yamaguchi^l, Tamao Tsukie^a, Mai Hasegawa^a, Kensaku
12 Kasuga^a, Masataka Kikuchi^a, Ryozo Kuwano^m, Takeshi Iwatsubo^{n,o,p}, Japanese Alzheimer's
13 Disease Neuroimaging Initiative^p, Shumpei Niida^q, Kouichi Ozaki^{b,r}, Takeshi Ikeuchi^{a,*}

14

15 **Affiliations**

- 16 a) Department of Molecular Genetics, Brain Research Institute, Niigata University, Niigata,
17 Japan
- 18 b) Medical Genome Center, Research Institute, National Center for Geriatrics and
19 Gerontology, Obu, Japan
- 20 c) Department of Neuropathology, Fukushima Hospital, Toyohashi, Japan
- 21 d) Department of Clinical Laboratory and Internal Medicine, National Center of Neurology
22 and Psychiatry (NCNP), National Center Hospital, Kodaira, Japan
- 23 e) Department of Pathology, Brain Research Institute, Niigata University, Niigata, Japan
- 24 f) Department of Pathology Neuroscience, Brain Research Institute, Niigata University,
25 Niigata, Japan
- 26 g) Department of Neuropathology (the Brain Bank for Aging Research), Tokyo Metropolitan
27 Institute for Geriatrics and Gerontology, Itabashi-ku, Japan
- 28 h) Brain Bank for Neurodevelopmental, Neurological and Psychiatric Disorders, United
29 Graduate School of Child Development, Osaka University, Osaka, Japan
- 30 i) Department of Neurology, Osaka Metropolitan University Graduate School of Medicine,
31 Osaka, Japan
- 32 j) Hatsuta Neurology Clinic, Osaka, Japan

33 k) Department of Neuropathology, Institute for Medical Science of Aging, Aichi Medical

34 University, Nagakute, Japan

35 l) Professor Emeritus, Gunma University, Maebashi, Japan

36 m) Social Welfare Corporation Asahigawaso, Asahigawaso Research Institute, Okayama,

37 Japan

38 n) Department of Neuropathology, Graduate School of Medicine, The University of Tokyo,

39 Bunkyo-ku, Japan

40 o) National Institute of Neuroscience, National Center of Neurology and Psychiatry, Kodaira,

41 Japan

42 p) Japanese Alzheimer's Disease Neuroimaging Initiative

43 q) Research Institute, National Center for Geriatrics and Gerontology, Obu, Japan

44 r) RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

45

46 †, Contributed equally to this work

47 *, Correspondence to: Akinori Miyashita (Ph.D.) and Takeshi Ikeuchi (M.D., Ph.D.),

48 Department of Molecular Genetics, Brain Research Institute, Niigata University, 1-757

49 Asahimachi, Chuo-ku, Niigata 951-8585, Japan; E-mail, [A.M.] [miyashi2020@bri.niigata-](mailto:miyashi2020@bri.niigata-u.ac.jp)
50 [u.ac.jp](mailto:miyashi2020@bri.niigata-u.ac.jp), and [T.I.] ikeuchi@bri.niigata-u.ac.jp.

51

52 **ABSTRACT**

53 **Background:**

54 *APOE* is a major susceptibility gene for Alzheimer's disease (AD). Recent studies in Europe
55 and the US have identified rare missense variants (RMVs) in *APOE* which are significantly
56 associated with AD. However, little is known regarding *APOE* RMVs in East Asians, including
57 the Japanese, and their association with AD and lipid metabolism.

58 **Objective:**

59 To identify *APOE* RMVs in the Japanese population and investigate their association with AD
60 and lipid metabolism, including low-density lipoprotein cholesterol levels.

61 **Subjects:**

62 *APOE* RMVs were explored in the NIG (2,589 subjects) and ToMMo (3,307 subjects) cohorts.
63 A case-control study was performed involving 6,471 AD and 20,270 control subjects.

64 **Methods:**

65 Sanger sequencing or whole-exome sequencing was performed on NIG subjects. We used
66 genotype data from ToMMo. *APOE* RMV frequencies in the Japanese were compared with
67 various ethnic populations. Associations among *APOE* RMV genotypes, AD, and lipoproteins
68 were examined.

69 **Results:**

70 Fourteen RMVs were identified (minor allele frequency 0.02 – 0.73%), 10 of which were
71 specifically found in East Asians. Five RMVs previously reported overseas were not detected
72 in Japanese individuals. Two RMVs (rs140808909 and rs190853081) exhibiting complete
73 linkage disequilibrium were associated with AD protective effects: $P_{\text{Bonferroni}} = 3.63\text{E-}02$, OR
74 (95% CI) = 0.70 (0.54–0.92). No significant differences in cholesterol levels were observed
75 between RMV carriers and non-carriers.

76 **Conclusions:**

77 *APOE* RMVs were identified in the Japanese population, with two showing potential protective
78 effects against AD. Further studies involving larger cohorts are required to confirm our findings
79 and investigate the role of *APOE* RMVs in AD and lipid metabolism.

80

81 **Keywords:**

82 Alzheimer's disease, *APOE*, rare missense variants, Japanese, lipid, cholesterol

83

84 **Running title:**

85 *APOE* RMVs in Japanese

86

87 **Introduction**

88 Alzheimer’s disease (AD) is the leading cause of dementia in elderly individuals. AD is a
89 multifactorial disease resulting from the interaction between environmental factors and
90 individual genetic susceptibility. A large twin epidemiological study in Sweden estimated the
91 heritability of AD to be 60–80%. ¹ Multiple loci are suggested to contribute to the onset of AD.
92 ² This indicates that genetic factors play a significant role in the development and progression
93 of AD. Large-scale genome-wide association studies (GWAS) have identified 90 susceptibility
94 genes/loci for AD. ³ Recent studies have focused on applying polygenic risk score analysis ⁴⁻⁶
95 and polygenic resilience score analysis ⁷, which integrate multiple AD susceptibility variants to
96 evaluate individual risks, with potential clinical applications being explored.

97 Among the AD susceptibility genes/loci identified to date, *APOE* is the most powerful
98 and explanatory gene across various populations. ⁸⁻¹¹ A search for “Alzheimer” on the GWAS
99 Catalog website ¹² reveals that the genomic region containing *APOE* shows a remarkably strong
100 association and is a “multiple hit locus” with numerous variants associated with AD. Recently,
101 rare missense variants (RMVs) of *APOE* have been reported to be associated with AD, shedding
102 light on new links between *APOE* and AD. For example, RMVs such as p.L46P (p.L28P;
103 rs769452 [Pittsburgh]) ¹³, p.R154S (p.R136S; rs121918393 [Christchurch]) ¹⁴, p.R163C

104 (p.R145C; rs769455) ¹⁵, p.V254E (p.V236E; rs199768005 [Jacksonville]) ¹⁶, and p.R269G
105 (p.R251G; rs267606661) ¹⁷ have been reported in Western populations, showing resistance to
106 amyloid- β (A β), phosphorylated tau protein, or the *APOE-e4* allele. Rasmussen *et al.* ¹⁸
107 identified 20 *APOE* RMVs in an analysis involving more than 100,000 participants and
108 demonstrated their association with plasma *APOE* levels and AD onset.

109 However, there have been few reports regarding the identification and significance of
110 *APOE* RMVs in the Japanese population. Therefore, this study was the first attempt at
111 identifying *APOE* RMVs in the Japanese population. Here, we examined the genetic association
112 between the identified RMVs and AD. Moreover, we investigated whether there were
113 differences in peripheral blood lipoprotein levels between carriers and non-carriers of these
114 RMVs.

115

116 **Subjects And Methods**

117 ***Subjects for identifying APOE RMVs***

118 We explored *APOE* RMVs in two Japanese cohorts (**Table 1**). The first cohort, referred to as
119 the NIG, consisted of 2,589 participants (1,150 females and 1,439 males) from medical and

120 research institutions across Japan. This cohort was primarily established for research-level
121 genetic diagnoses, including AD and various other neurological disorders. Of these, 826 cases
122 were based on clinical diagnoses, and 1,763 were based on pathological diagnoses. Additionally,
123 508 of the 2,589 NIG subjects were derived from the Japanese Alzheimer's Disease
124 Neuroimaging Initiative (J-ADNI).¹⁹ Sanger sequencing (SS), whole-exome sequencing
125 (WES), and a combination of both (SS + WES) were performed on 839, 1,509, and 241 subjects,
126 respectively. The second cohort, referred to as ToMMo, consisted of 3,307 community-dwelling
127 healthy subjects (1,810 females and 1,497 males) recruited by the Tohoku Medical Megabank
128 Organization.²⁰ In compliance with the data usage policy (Research ID: 2019-0074), individual
129 genotype data for 176 variants within a genomic region of approximately 7.2 kb, including
130 *APOE* (positioned at 45,407,254 – 45,414,449 bp in GRCh37/hg19), were obtained. The total
131 number of subjects from both cohorts was 5,896 (2,960 females and 2,936 males), with
132 participants aged ≥ 20 years (mean: 64.1 years, median: 66.0 years, range: 20–111 years).

133 This study was approved and conducted according to the ethical guidelines of the
134 Ethics Committee of Niigata University and all participating medical and research institutions
135 (approval number of the Genetic Ethics Review Committee, Niigata University: G2017-0012,
136 G2018-0034, 2018-0409, and G2022-0009).

137

138 ***SS for APOE***

139 SS was performed in 1,080 subjects from the NIG cohort (839 SS; 241 SS + WES) (**Table 1**).

140 Genomic DNA was extracted from peripheral blood or autopsied brain tissues and quantified

141 using the NanoDrop spectrophotometer (Thermo Fisher Scientific). DNA quality was assessed

142 by calculating the DNA integrity number (Agilent Technologies). The primers used for

143 sequencing the entire coding region of *APOE* is provided in **Supplementary Table 1**.

144

145 ***WES***

146 WES was performed on 1,750 participants from the NIG (1,509 for WES and 241 for SS +

147 WES) (**Table 1**). Genomic DNA was used to prepare an exome library using the SureSelect

148 Human All Exon V6 Kit (Agilent Technologies). The library was sequenced on an NovaSeq

149 6000 platform (Illumina) in 151-cycle paired-end mode. Sequence reads were processed with

150 fastp (v.0.19.5) using the default settings for quality control and adapter trimming. The cleaned

151 reads were then mapped to the human reference genome (hg38) using BWA-MEM (v.0.7.15-

152 r1140) with default parameters. Subsequent analyses, including read processing, variant calling,

153 and filtration, were performed following GATK4 Best Practices. The resulting variant call sets
154 were annotated using snpEff (v.4.3t) and stored in an SQLite database using GEMINI (v.0.20.1).

155

156 ***Determination of six common APOE genotypes***

157 We determined six common *APOE* genotypes ($e2*2$, $e2*3$, $e2*4$, $e3*3$, $e3*4$, and $e4*4$) in all
158 5,896 subjects from the NIG and ToMMo cohorts (**Table 1**) based on nucleotide combinations
159 (haplotypes) of the common missense variants rs429358 (T>C: p.C130R [p.C112R]) and
160 rs7412 (C>T: p.R176C [p.R158C]) in *APOE*. For the NIG cohort, these genotypes were
161 determined using both the SS and TaqMan methods. Information regarding the TaqMan assays
162 used (TaqMan assay IDs: 19AyD0001 for rs429358 and 19AyD0002 for rs7412) is provided in
163 **Supplementary Table 1**. For the ToMMo cohort, genotypes were determined based on digital
164 data (genotype: 0/0, 0/1, and 1/1). The results are presented in **Table 1**. The risk effects of *APOE*
165 genotypes and alleles in Japanese patients with AD are shown in **Supplementary Figure 1**.
166 *APOE* genotype and allele information reported by Asanomi *et al.* ²¹ was used in this analysis.

167

168 ***Case-control study***

169 We conducted a case-control study to examine whether the identified *APOE* RMVs were
170 genetically associated with AD. Information concerning the subjects used in this analysis is
171 provided in **Table 3**, including sex, age (age at testing/death), and *APOE* genotype. Genomic
172 variants for each subject were comprehensively genotyped using an Asian Screening Array
173 (ASA) (Illumina). The number of samples analyzed by the ASA was 6,471 patients with AD
174 and 20,270 control subjects.

175 None of the 14 *APOE* RMVs identified in the NIG and ToMMo cohorts were included
176 in the ASA. Therefore, genotypes were inferred through imputation for each subject. We
177 performed SNP imputation using Minimac4 with the Japanese reference panel, which was
178 constructed from 1,000 Genomes Project Phase 3 data (1KGP3 [May 2013, n = 2,503]) and
179 3,181 Japanese whole-genome sequences from the National Center for Geriatrics and
180 Gerontology (NCGG) biobank database. Variants with an INFO score ≥ 0.3 were included in
181 the association analysis. Among the 14 *APOE* RMVs, genotype data were obtained for 4
182 variants (rs121918392 [e5], rs762461580, rs140808909 [e7], and rs190853081 [e7]) (**Table 4**).

183

184 ***TA cloning to verify RMV alleles***

185 In one case (Subject ID #35: female, age range at death 90-99, $\epsilon 3^*3$) with three *APOE* RMVs
186 (all heterozygous: rs267606663-G/a, rs140808909 [e7]-G/a, rs190853081 [e7]-G/a)
187 (**Supplementary Table 5**), we investigated whether these RMVs were located on the same
188 chromosome. All the RMVs were located in exon 4 of *APOE*. Genomic DNA from the subject
189 was used as a template for PCR, followed by TA cloning (TaKaRa Mighty TA-cloning Kit). The
190 sequences of the primer pair used for PCR were as follows: forward primer 5 '-
191 TGCCGATGACCTGCAGAA-3', and the reverse primer was 5'-
192 GCTGCATGTCTTCCACCAG-3'.

193

194 ***Lipoprotein data collection***

195 Measurement data for total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C),
196 high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG) were available for the J-
197 ADNI subjects included in the NIG (508 cases: baseline) and ToMMo (TC: 234 subjects; non-
198 TC: 1,630 subjects). We examined the quantitative differences between *APOE* RMV carriers
199 and non-carriers, if any.

200

201 **Statistical analysis**

202 Differences in sex distribution and *APOE* genotypic and allelic distributions between cases
203 (AD) and controls were assessed using Fisher's exact test based on 2×2 or 2×3 contingency
204 tables (**Table 3, Supplementary Table 4**). Additionally, unadjusted odds ratios (OR) and 95%
205 confidence intervals (95% CI) were calculated (**Table 3, Supplementary Table 4**). For age,
206 unpaired Student's *t*-test with Welch's correction was used (**Table 3**).

207 Logistic regression analysis was conducted on *APOE* RMVs in the case-control study,
208 with sex, age, and principal components as covariates to calculate *P*-values and ORs (95% CI)
209 (**Table 4**). The Bonferroni method was applied to adjust for multiple comparisons and
210 corrections were made for the four *APOE* RMVs in the ASA analysis.

211 A meta-analysis was conducted to integrate and assess the association between *APOE*
212 genotypes/alleles and AD derived from the two cohorts as described in a study by Asanomi *et*
213 *al.* ²¹: the NCGG and the Japanese Genetic Study Consortium for Alzheimer's Disease
214 (JGSCAD) ²⁹ (**Supplementary Figure 1**). Both fixed-effects and random-effects models were
215 assumed, and Chi-square *P*-values, pooled ORs, and 95% CIs were calculated. Cochran's Q
216 statistic and I² statistic were also computed to evaluate heterogeneity between the two datasets.

217 Comparisons of TC, LDL-C, HDL-C, TG, and the LDL-C/HDL-C (LH) ratio between
218 the two groups (*APOE* RMV carriers vs. non-carriers) were conducted using the non-parametric
219 Mann-Whitney *U*-test (**Supplementary Figure 4**).

220 The series of genetic statistical analyses related to the case-control study (**Table 4,**
221 **Figure 2**) were conducted using PLINK (v.1.90; <https://www.cog-genomics.org/plink/>).³⁰
222 Meta-analysis (**Supplementary Figure 1**) was conducted using StatsDirect (v.4.0.4;
223 <https://www.statsdirect.com>). Other statistical analyses (**Tables 1 and 3, Supplementary Table**
224 **4, Figure 3, Supplementary Figure 4**) were performed using GraphPad Prism software
225 (v.10.3.1; <https://www.graphpad.com>). *P*-values less than 0.05 were considered statistically
226 significant.

227

228 **Databases**

229 Databases and analysis websites used in this study are listed below in alphabetical order.

- 230 • Alzforum MUTATIONS APOE: <https://www.alzforum.org/mutations/apoe>
- 231 • ClinVar: <https://www.ncbi.nlm.nih.gov/clinvar/>
- 232 • dbSNP: <https://www.ncbi.nlm.nih.gov/snp/>

- 233 • Ensemble Human (GRCh38.p14): https://asia.ensembl.org/Homo_sapiens/Info/Index
- 234 • gnomAD (v4.1.0): <https://gnomad.broadinstitute.org>
- 235 • GWAS Catalog: <https://www.ebi.ac.uk/gwas/home>
- 236 • jMorp: <https://jmorp.megabank.tohoku.ac.jp>
- 237 • LDpair: <https://ldlink.nih.gov/?tab=ldpair>
- 238 • LDpop: <https://ldlink.nih.gov/?tab=ldpop>
- 239 • UCSC genome: <https://genome.ucsc.edu>

240

241 **Results**

242 ***Identification of APOE RMVs in the Japanese population***

243 Using 2,589 subjects from NIG (**Table 1**), we identified seven *APOE* RMVs (**Table 2, Figure**
244 **1**): rs1969812567 (p.A7P), rs121918392 (p.E21K), rs762461580 (p.R50H), rs370594287
245 (p.Q64H), rs267606663 (p.R242Q), rs140808909 (p.E262K), and rs190853081 (p.E263K). All
246 carriers were heterozygous, and no homozygotes were identified. A summary of the clinical and
247 pathological diagnostic information of these patients is provided in **Supplementary Table 5**.
248 Among these seven RMVs, the two variants with the highest alternative allele frequencies were

249 rs140808909 (p.E262K) and rs190853081 (p.E263K), both with a minor allele frequency
250 (MAF) of 0.00657 (**Table 2**). These two variants were in complete linkage disequilibrium (LD)
251 (LD coefficient $R^2 = 1$), indicating that they were fully co-inherited (**Supplementary Figure**
252 **2**). Both variants were previously identified in Japanese patients with dyslipidemia and
253 atherosclerosis, defined as the *APOE-e7* allele. ²²⁻²⁴ rs121918392 (p.E21K) was defined as the
254 *APOE-e5* allele ^{22, 25-27} and was the third-most frequent variant after the *APOE-e7*-defining
255 rs140808909 (p.E262K) and rs190853081 (p.E263K), with an MAF of 0.00348 (**Table 2**).
256 rs121918392 (p.E21K) was also associated with dyslipidemia and atherosclerosis. ^{22, 25-27}

257 In an analysis of 3,307 subjects from the ToMMo cohort (**Table 1**), we identified 14
258 *APOE* RMVs (**Table 2, Figure 1**). Seven variants were identified in both the NIG and ToMMo
259 cohorts. Additionally, seven variants were found in the ToMMo cohort but not in the NIG
260 cohort: rs201672011 (p.E31K), rs748703149 (p.G145S), rs954186991 (p.A156T),
261 rs1339651557 (p.R185H), rs1969871875 (p.E186K), rs1385100920 (p.G206D), and
262 rs767339630 (p.R278H). All carriers were heterozygous, and no homozygotes were identified.
263 For two variants in complete LD, rs140808909 (p.E262K) and rs190853081 (p.E263K), we
264 identified one homozygous individual: sex, male; age range at entry, 60-69 years; *APOE*
265 genotype, *e3*3*; and body mass index, 23.0. This individual had LDL-C, HDL-C, and TG levels

266 of 96, 49, and 67, respectively, all within normal ranges (TC data were not available).

267 Five previously reported variants, rs769452 (p.L46P: Pittsburgh) ¹³, rs121918393
268 (p.R154S: Christchurch) ¹⁴, rs769455 (p.R163C) ¹⁵, rs199768005 (p.V254E: Jacksonville) ¹⁶,
269 and rs267606661 (p.R269G) ¹⁷ (**Figure 1**), were not detected in either NIG or ToMMo cohorts
270 (**Table 2**). Furthermore, these variants were not observed in the Japanese multi-omics database
271 jMorp (54 KJPN, 38 KJPN, and 14 KJPN) ²⁸ (**Supplementary Table 3**). These variants are
272 predominantly observed in non-East Asian populations.

273

274 ***Relationship between identified APOE RMVs and common APOE alleles, e2, e3, and e4***

275 Clarifying the common *APOE* alleles (*e2*, *e3*, and *e4*) on which each identified *APOE* RMV is
276 located is crucial for understanding any potential genetic relationship between AD and *APOE*.
277 Therefore, for each *APOE* RMV, we categorized subjects from the NIG and ToMMo cohorts
278 (**Table 1**) according to the common *APOE* genotypes: *e2***2*, *e2***3*, *e2***4*, *e3***3*, *e3***4*, and *e4***4*
279 (**Supplementary Table 2**). The most common *APOE* genotype among *APOE* RMV carriers
280 was *e3***3*. Several individuals were also identified with *e3***4* and *e2***3*, and one individual with
281 *e4***4*. None of the *APOE* RMV carriers were found with *e2***2* or *e2***4*. The variant rs121918392

282 (p.E21K), which defines the *e5* allele, and two variants rs140808909 (p.E262K) and
283 rs190853081 (p.E263K), which are in complete LD and define the *e7* allele, appeared to co-
284 segregate with the *e3* allele (**Supplementary Table 2**). The rs767339630 (p. R278H) variant
285 was found in a subject with *e4*4*, suggesting that this variant is likely to co-segregate with the
286 *e4* allele (**Supplementary Table 2**).

287

288 ***Ethnicity dependent frequencies of APOE RMVs***

289 To investigate whether the 14 *APOE* RMVs identified in the NIG and ToMMo cohorts exist in
290 other populations and to assess their frequencies, we explored public databases
291 (**Supplementary Table 3**). First, we examined the jMorp database.²⁸ Alternative alleles were
292 observed in all 14 variants. For the two variants in complete LD, rs140808909 (p.E262K) and
293 rs190853081 (p.E263K), several homozygous individuals were identified: twelve in 54 KJPN,
294 six in 38 KJPN, and three in 14 KJPN. This result was consistent with our analysis, which
295 showed that these two *APOE* RMVs had the highest alternative allele frequencies. The other 12
296 variants were observed only in heterozygous individuals.

297 Four variants, rs201672011 (p.E31K), rs762461580 (p.R50H), rs370594287 (p.Q64H),

298 and rs748703149 (p.G145S), were also found in populations other than East Asian ones
299 **(Supplementary Table 3)**. In contrast, six variants, rs121918392 (p.E21K), rs954186991
300 (p.A156T), rs1339651557 (p.R185H), rs267606663 (p.R242Q), rs140808909 (p.E262K), and
301 rs190853081 (p.E263K), were highly ethnicity-dependent and observed only in the East Asian
302 population **(Supplementary Table 3)**. Information regarding frequencies of the four variants,
303 rs1969812567 (p.A7P), rs1969871875 (p.E186K), rs1385100920 (p.G206D), and rs767339630
304 (p.R278H), was not available for other populations **(Supplementary Table 3)**.

305

306 ***Effect of APOE RMVs on AD***

307 To investigate whether the 14 identified RMVs **(Table 2, Figure 1)** were genetically associated
308 with AD, we conducted a case-control study involving 6,471 AD cases and 20,270 controls
309 **(Table 3)**. Of the 14 RMVs, examined, there were four variants for which genotype data were
310 obtained by imputation. Two variants, rs140808909 (p.E262K) and rs190853081 (p.E263K),
311 that define the *e7* allele, showed a statistically significant association with AD **(Table 4)**. Since
312 these two RMVs are in complete LD, they exhibited the same *P*-values and OR (95% CI):
313 $P_{\text{Unadjusted}} = 8.95\text{E-}03$, $P_{\text{Bonferroni}} = 3.58\text{E-}02$, and OR (95% CI) = 0.70 (0.54 – 0.92) **(Table 4,**

314 **Figure 2).** An OR < 1.00 suggests that the *e7* allele may have a protective effect against AD.

315

316 ***Association between APOE genotypes/alleles and AD in APOE RMV carriers***

317 By focusing on carriers of the four RMVs analyzed in the case-control study (**Table 4**), we
318 investigated whether *APOE* genotypes and alleles were associated with AD. Because no carriers
319 of *APOE* RMVs with the *e2*2*, *e2*4*, or *e4*4* genotypes were identified, they were excluded
320 from analysis. Additionally, owing to the low frequency of the *e2*3* genotype and *e2* allele,
321 they were excluded from analysis. As a result, the comparison of genotypes focused on *e3*3*
322 and *e3*4*, and we compared *e3* and *e4* for alleles. Results of the analyses are presented in **Figure**
323 **3** and **Supplementary Table 4**. For reference, the OR (95% CI) including both carriers and
324 non-carriers of *APOE* RMVs, denoted as “All,” were also provided, derived from the data in
325 **Table 3**. The OR (95% CI) for both *APOE* genotypes and alleles exhibited similar trends.
326 Specifically, for the three RMVs defining the *e5* (rs121918392) and *e7* (rs140808909 and
327 rs190853081) alleles, there was a tendency for reducing the risk of the *e4* allele. For example,
328 in a comparison between the *e3*3* and *e3*4* genotypes, the OR (95% CI) for All was 3.04 (2.85-
329 3.24), whereas it was 2.20 (1.01-4.95) for *e5* (rs121918392). In contrast, for rs762461580, there

330 was a slight increase in risk associated with the *e4* allele. For example, in the same genotype
331 comparison, the OR (95% CI) for All was 3.04 (2.85-3.24), but was 3.78 (1.05-12.54) for
332 rs762461580.

333

334 ***Comparison of Lipoprotein Levels Between APOE RMV Carriers and Non-Carriers***

335 We compared blood lipoprotein levels between carriers and non-carriers of *APOE* RMVs in J-
336 ADNI participants (508 subjects) from the NIG cohort and in participants from the ToMMo
337 cohort (TC: 234 subjects; and non-TC: 1,630 subjects). No statistically significant differences
338 were observed between *APOE* RMV carriers and non-carriers in terms of TC, LDL-C, HDL-C,
339 TG, or LH ratios (**Supplementary Figure 4**).

340

341 **Discussion**

342 In this study, we identified 14 *APOE* RMVs among 5,896 Japanese individuals (**Table 2, Figure**
343 **1**). We found that some *APOE* RMVs were specific to East Asians, including Japanese
344 (**Supplementary Table 3**). Notably, variants previously reported in Western populations
345 (Christchurch¹⁴ and Jacksonville variants¹⁶) were completely absent in the Japanese population

346 **(Table 2, Supplementary Table 3)**. These findings suggested that RMVs have distinct ethnic
347 and regional specificities. To investigate their genetic association with AD, we conducted a
348 case-control study involving 26,741 individuals **(Table 3)**. Two *APOE* RMVs, rs140808909
349 and rs190853081, which were in complete LD **(Supplementary Figure 2)** and defined as the
350 *e7* allele ²²⁻²⁴, showed a statistically significant association with AD, with a protective effect
351 **(Table 4, Figure 2)**. To the best of our knowledge, this association between *APOE* RMV and
352 AD is the first of its kind in East Asians. Future studies with larger sample sizes are strongly
353 recommended to validate these findings, as reproducibility is a critical focus. Considering the
354 genetic diversity and ethnic specificity of the genomic variants ³¹, detailed analyses of each
355 ethnic group are important.

356 In this exploratory study, we identified 14 *APOE* RMVs **(Table 2, Figure 1)**, with the
357 *e7* allele ²²⁻²⁴ being the most intriguing because of its potential protective effects against AD.
358 The two RMVs that define the *e7* allele replace two glutamate (E) residues with lysine (K) at
359 the entry point of the lipid-binding domain on the C-terminal side of APOE **(Figure 1)**. This
360 substitution reduces the binding affinity of *e7* to LDL receptors and increases its affinity for
361 heparin. ³² No previous studies have reported a genetic association between *e7* and AD.
362 However, Youn *et al.* ³³ suggested that the *e7* allele may be involved in cognitive impairment,

363 based on a study of 344 amnesic patients and 345 controls in a Korean population. They
364 identified three heterozygous carriers of the *e7* allele, all of whom belonged to the amnesic
365 group. Two of these individuals were clinically diagnosed with AD and small vessel disease,
366 and the remaining one was diagnosed with subjective cognitive impairment. They also
367 predicted the three-dimensional protein structure of *e7*, showing that its C-terminal structure
368 resembled that of *e4*, despite having an *e3* background. They concluded that *e7* might behave
369 like *e4* and potentially exert adverse effects on cognitive function. While our study suggests a
370 protective effect of the *e7* allele against AD, the relationship between cognitive function and
371 the *e7* allele is an important area for future investigation.

372 We compared peripheral blood cholesterol levels (TC, LDL-C, HDL-C, TG, and LH
373 ratios) between carriers and non-carriers of the identified *APOE* RMVs. Contrary to
374 expectations, no significant differences were observed in cholesterol levels (**Supplementary**
375 **Figure 4**). Owing to the rarity of *APOE* RMVs, all carriers were grouped into a single category
376 for analysis in this study. Different *APOE* RMVs may have varying effects on peripheral blood
377 cholesterol levels; ideally, each *APOE* RMV should be analyzed separately. For instance, a
378 study concerning the *e5* and *e7* alleles revealed a trend toward higher TC (> 220 mg/dL) and
379 TG (> 150 mg/dL) levels compared with controls.³⁴ Considering the low population frequency

380 of *APOE* RMVs (< 1%), comparative analyses with larger sample sizes are warranted.

381 This study had certain limitations. First, the number of identified RMVs (**Table 2,**
382 **Figure 1**) was lower than expected, and further exploratory studies using a larger sample of
383 Japanese individuals are needed. Specifically, if the focus is on late-onset AD, exploratory
384 studies of RMVs using samples from older individuals, particularly those aged 65 and above,
385 would be valuable. Second, we were unable to verify the reproducibility of the genetic
386 association between AD and the *e7* allele (rs140808909 and rs190853081) in a larger
387 independent sample of Japanese individuals, which warrants further investigation. Third,
388 because this study was based on cross-sectional data, the longitudinal relationship between
389 RMVs and the risk of developing AD remains unclear. Finally, functional assays of RMVs have
390 not yet been conducted, leaving the impact of RMVs on AD pathology and the underlying
391 molecular mechanisms unclear.

392 *APOE* plays a crucial role in the pathophysiology of AD, particularly in relation to A β
393 and tau. ³⁵⁻³⁶ Previous research has largely focused on the major missense alleles, *e2*, *e3* and *e4*,
394 of *APOE*. In addition to the major *APOE* missense variants, research on *APOE* RMVs,
395 including the Christchurch variant (rs121918393: p.R154S [p.R136S]) ^{14, 37-41}, is expected to
396 expand, further deepening our understanding of the role of *APOE* in AD pathophysiology.

397 Additional studies are required to elucidate the functional significance of the 14 *APOE* RMVs

398 identified in this study.

399

400 **Author Contributions**

401 AM and Tikeuchi: Study design, data analysis, interpretation of data, and manuscript

402 preparation and revision. AO, NH, TT, MH, KK, and MK: Nucleic acid extraction, quality

403 control, sequencing, genotyping analyses, data interpretation, and manuscript revision. RM, SN,

404 and KO: case-control studies using gene chips, data interpretation, and manuscript revision. DK,

405 YH, TS, MTakao, RG, MTada, AK, AArakawa, MM, SM, YS, HH, TM, AAkagi, YR, HM, JS,

406 MY, and HY provided the autopsied brains and revised the manuscript. RK and TIwatsubo:

407 Data interpretation and manuscript revision. All authors have read and approved the final

408 version of the manuscript.

409

410 **ACKNOWLEDGMENTS**

411 We wish to thank all of the participants for their involvement in this study. We also gratefully

412 acknowledge the ToMMo cohort (**Table 1**) for providing individual genomic variant data on

413 *APOE* and its surrounding regions as well as various datasets from the subjects (research ID:
414 2019-0074). In the preparation of this paper, an initial rough English proofreading was
415 performed using ChatGPT-4o, followed by proofreading by a native speaker of English.

416

417 **FUNDING**

418 This study was supported by: (1) JSPS Grant-in-Aid for Scientific Research (KAKENHI) (C),
419 21K07271 (AM); (2) JSPS Grant-in-Aid for Scientific Research (KAKENHI) (B), 21H03537
420 (AM); (3) AMED, JP21dk0207055 (AM); (4) JSPS Grant-in-Aid for Scientific Research
421 (KAKENHI), 22H04923 (DK); (5) AMED, JP21wm0425019 (MT); (6) Intramural fund from
422 NCNP, Grant Number 6-8 (MT); (7) JSPS Grant-in-Aid for Scientific Research (KAKENHI)
423 (C), 18K06506 (MT); (8) JSPS Grant-in-Aid for Scientific Research (KAKENHI) (C),
424 21K06417 (MT); (9) JSPS Grant-in-Aid for Transformative Research Areas, 22H04923 (MT);
425 (10) the Research Committee of Prion Disease and Slow Virus Infection, the Ministry of Health,
426 Labour and Welfare Health and Labour Sciences Research Grants, Japan (MT); (11) AMED,
427 JP24wm0425019 (AK) ; (12) AMED, JP21wm0425019 (YS, SM); (13) JSPS Grant-in-Aid for
428 Scientific Research (KAKENHI), 22H04923 (CoBiA) (YS); (14) JSPS Grant-in-Aid for

429 Scientific Research (KAKENHI) (C), 22K07550 (YS); (15) MHLW Research on rare and
430 intractable diseases Program Grant Number JPMH23FC1008 (YS); (16) AMED,
431 JP21wm0425019 (MY); (17) AMED, JP18kk0205009 (SN); (18) AMED, JP21dk0207045
432 (SN); (19) AMED, JP22dk0207060 (SN); (20) AMED, JP21dk0207045 (KO); (21) AMED,
433 JP22dk0207060 (KO); (22) Research Funding for Longevity Sciences from NCGG (24-15)
434 (KO); (23) a grant from the Japanese Ministry of Health, Labour, and Welfare for Research on
435 Dementia (KO); (24) AMED, JP23wm0525019 (MK); (25) AMED, JP23dk0207060 (MK);
436 (26) JSPS Grant-in-Aid for Scientific Research (KAKENHI) (B), 23K27515 (TIwatsubo); and
437 (27) AMED, JP22dk0207060 (TIkeuchi)

438

439 **CONFLICTS OF INTEREST**

440 The authors declare no conflicts of interest.

441

442 **Supplementary Material**

443 Supplementary figures and tables are available in the electronic version of this article.

444

445 **Data Availability**

446 The data supporting the findings of this study are generally available upon request from the
447 corresponding authors (AM and Tikeuchi). However, the WES data used in this study are not
448 publicly available due to privacy or ethical restrictions.

449

450 **References**

- 451 [1] Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining
452 Alzheimer disease. *Arch Gen Psychiatry* 2006; 63: 168-174.
- 453 [2] Daw EW, Payami H, Nemens EJ, et al. The number of trait loci in late-onset Alzheimer
454 disease. *Am J Hum Genet* 2000; 66: 196-204.
- 455 [3] Lambert JC, Ramirez A, Grenier-Boley B, et al. Step by step: towards a better
456 understanding of the genetic architecture of Alzheimer's disease. *Mol Psychiatry* 2023; 28:
457 2716-2727.
- 458 [4] Escott-Price V, Sims R, Bannister C, et al. Common polygenic variation enhances risk
459 prediction for Alzheimer's disease. *Brain* 2015; 138: 3673-3684.
- 460 [5] Stocker H, Möllers T, Perna L, et al. The genetic risk of Alzheimer's disease beyond APOE

- 461 $\epsilon 4$: systematic review of Alzheimer's genetic risk scores. *Transl Psychiatry* 2018; 8: 166.
- 462 [6] Kikuchi M, Miyashita A, Hara N. Alzheimer's Disease Neuroimaging Initiative; Japanese
463 Alzheimer's Disease Neuroimaging Initiative Polygenic effects on the risk of Alzheimer's
464 disease in the Japanese population. *Alzheimers Res Ther* 2024; **16**: 45.
- 465 [7] Hou J, Hess JL, Armstrong N, et al. Polygenic resilience scores capture protective genetic
466 effects for Alzheimer's disease. *Transl Psychiatry* 2022; 12: 296.
- 467 [8] Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the
468 association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis.
469 APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* 1997; 278: 1349-1356.
- 470 [9] Bertram L, McQueen MB, Mullin K, et al. Systematic meta-analyses of Alzheimer disease
471 genetic association studies: the AlzGene database. *Nat Genet* 2007; 39: 17-23.
- 472 [10] Belloy ME, Andrews SJ, Le Guen Y, et al. APOE Genotype and Alzheimer Disease Risk
473 Across Age, Sex, and Population Ancestry. *JAMA Neurol* 2023; 80: 1284-1294.
- 474 [11] Ridge PG, Hoyt KB, Boehme K, et al. Assessment of the genetic variance of late-onset
475 Alzheimer's disease. *Neurobiol Aging* 2016; 41: 200.e13-200.e20.
- 476 [12] Sollis E, Mosaku A, Abid A, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and
477 deposition resource. *Nucleic Acids Res* 2023; 51: D977-D985.

- 478 [13] Kamboh MI, Aston CE, Perez-Tur J, et al. A novel mutation in the apolipoprotein E gene
479 (APOE*4 Pittsburgh) is associated with the risk of late-onset Alzheimer's disease.
480 *Neurosci Lett* 1999; 263: 129-132.
- 481 [14] Arboleda-Velasquez JF, Lopera F, O'Hare M, et al. Resistance to autosomal dominant
482 Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. *Nat Med* 2019;
483 25: 1680-1683.
- 484 [15] Le Guen Y, Raulin AC, Logue MW, et al. Association of African Ancestry-Specific APOE
485 Missense Variant R145C With Risk of Alzheimer Disease. *JAMA* 2023; 329: 551-560.
- 486 [16] Medway CW, Abdul-Hay S, Mims T, et al. ApoE variant p.V236E is associated with
487 markedly reduced risk of Alzheimer's disease. *Mol Neurodegener* 2014; 9: 11.
- 488 [17] Le Guen Y, Belloy ME, Grenier-Boley Bet al. Association of Rare APOE Missense
489 Variants V236E and R251G With Risk of Alzheimer Disease. *JAMA Neurol* 2022; 79: 652-
490 663.
- 491 [18] Rasmussen KL, Tybjaerg-Hansen A, Nordestgaard BG, et al. APOE and dementia -
492 resequencing and genotyping in 105,597 individuals. *Alzheimers Dement* 2020; 16: 1624-
493 1637.
- 494 [19] Iwatsubo T, Iwata A, Suzuki K, et al. Japanese and North American Alzheimer's Disease

- 495 Neuroimaging Initiative studies: Harmonization for international trials. *Alzheimers*
496 *Dement* 2018; 14: 1077-1087.
- 497 [20] Tadaka S, Katsuoka F, Ueki M, et al. 3.5KJPNv2: an allele frequency panel of 3552
498 Japanese individuals including the X chromosome. *Hum Genome Var* 2019; 6: 28.
- 499 [21] Asanomi Y, Shigemizu D, Miyashita A, et al. A rare functional variant of SHARPIN
500 attenuates the inflammatory response and associates with increased risk of late-onset
501 Alzheimer's disease. *Mol Med* 2019; 25: 20.
- 502 [22] Yamamura T, Yamamoto A, Sumiyoshi T, et al. New mutants of apolipoprotein E
503 associated with atherosclerotic diseases but not to type III hyperlipoproteinemia. *J Clin*
504 *Invest* 1984; 74: 1229-1237.
- 505 [23] Maeda H, Nakamura H, Kobori S, et al. Identification of human apolipoprotein E variant
506 gene: apolipoprotein E7 (Glu244,245----Lys244,245). *J Biochem* 1989; 105: 51-54.
- 507 [24] Tajima S, Yamamura T, Menju M, et al. Analysis of apolipoprotein E7 (apolipoprotein E-
508 Suita) gene from a patient with hyperlipoproteinemia. *J Biochem* 1989; 105: 249-253.
- 509 [25] Yamamura T, Yamamoto A, Hiramori K, et al. A new isoform of apolipoprotein E--apo E-
510 5--associated with hyperlipidemia and atherosclerosis. *Atherosclerosis* 1984; 50: 159-172.
- 511 [26] Tajima S, Yamamura T, Yamamoto A. Analysis of apolipoprotein E5 gene from a patient

- 512 with hyperlipoproteinemia. *J Biochem* 1988; 104: 48-52.
- 513 [27] Maeda H, Nakamura H, Kobori S, et al. Molecular cloning of a human apolipoprotein E
514 variant: E5 (Glu3----Lys3). *J Biochem* 1989; 105: 491-493.
- 515 [28] Tadaka S, Kawashima J, Hishinuma E, et al. jMorp: Japanese Multi-Omics Reference
516 Panel update report 2023. *Nucleic Acids Res* 2024; 52: D622-D632.
- 517 [29] Kuwano R, Miyashita A, Arai H, et al. Dynamin-binding protein gene on chromosome 10q
518 is associated with late-onset Alzheimer's disease. *Hum Mol Genet* 2006; 15: 2170-2182.
- 519 [30] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association
520 and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559-575.
- 521 [31] Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and
522 population history from 929 diverse genomes. *Science* 2020; 367: eaay5012.
- 523 [32] Yamamura T, Dong LM, Yamamoto A. Characterization of apolipoprotein E7 (Glu(244)--
524 >Lys, Glu(245)--->Lys), a mutant apolipoprotein E associated with hyperlipidemia and
525 atherosclerosis. *J Lipid Res* 1999; 40: 253-259.
- 526 [33] Youn YC, Lim YK, Han SH, et al. Apolipoprotein ε 7 allele in memory complaints: insights
527 through protein structure prediction. *Clin Interv Aging* 2017; 12: 1095-1102.
- 528 [34] Matsunaga A, Sasaki J, Moriyama K, et al. Population frequency of apolipoprotein E5

- 529 (Glu3-->Lys) and E7 (Glu244-->Lys, Glu245-->Lys) variants in western Japan. *Clin Genet*
- 530 1995; 48: 93-99.
- 531 [35] Yamazaki Y, Zhao N, Caulfield TR, et al. Apolipoprotein E and Alzheimer disease:
- 532 pathobiology and targeting strategies. *Nat Rev Neurol* 2019; 15: 501-518.
- 533 [36] Belloy ME, Napolioni V, Greicius MD. A Quarter Century of APOE and Alzheimer's
- 534 Disease: Progress to Date and the Path Forward. *Neuron* 2019; 101: 820-838.
- 535 [37] Wardell MR, Brennan SO, Janus ED, et al. Apolipoprotein E2-Christchurch (136 Arg----
- 536 Ser). New variant of human apolipoprotein E in a patient with type III
- 537 hyperlipoproteinemia. *J Clin Invest* 1987; 80: 483-490.
- 538 [38] Sepulveda-Falla D, Sanchez JS, Almeida MC, et al. Distinct tau neuropathology and
- 539 cellular profiles of an APOE3 Christchurch homozygote protected against autosomal
- 540 dominant Alzheimer's dementia. *Acta Neuropathol* 2022; 144: 589-601.
- 541 [39] Nelson MR, Liu P, Agrawal A, et al. The APOE-R136S mutation protects against APOE4-
- 542 driven Tau pathology, neurodegeneration and neuroinflammation. *Nat Neurosci* 2023; 26:
- 543 2104-2121.
- 544 [40] Chen Y, Song S, Parhizkar S, et al. APOE3ch alters microglial response and suppresses
- 545 A β -induced tau seeding and spread. *Cell* 2024; 187: 428-445.e20.

- 546 [41] Murakami R, Watanabe H, Hashimoto H, et al. Inhibitory Roles of Apolipoprotein E
547 Christchurch Astrocytes in Curbing Tau Propagation Using Human Pluripotent Stem Cell-
548 Derived Models. *J Neurosci* 2024; 44: e1709232024.

549 **Tables**

550 **Table 1. The information on subjects used in the exploration of *APOE* RMVs.**

551 **Table 2. *APOE* RMVs identified.**

552 **Table 3. The information on subjects used in the case-control study.**

553 **Table 4. Case-control study on *APOE* RMVs identified.**

554 **Supplementary Table 1. SS primers for all *APOE* exons and TaqMan SNP genotyping
555 assays for rs429358 and rs7412.**

556 **Supplementary Table 2. *APOE* RMV carriers expanded based on six common *APOE*
557 genotypes.**

558 **Supplementary Table 3. Details of 14 RMVs identified in this study and 5 RMVs
559 previously reported elsewhere.**

560 **Supplementary Table 4. Case-control study on *APOE* RMVs identified.**

561 **Supplementary Table 5. Clinical and neuropathological diagnoses of *APOE* RMV carriers
562 in the NIG cohort.**

563

564

565 **Figure Legends**

566 **Figure 1. *APOE* RMVs identified in this study.**

567 The 14 *APOE* RMVs identified in this study have been added to the Alzforum MUTATIONS
568 *APOE* diagram (<https://www.alzforum.org/mutations/apoe>). RMVs identified in both ToMMo
569 and NIG cohorts are marked in red, whereas those identified only in ToMMo are marked in
570 magenta.

571

572 **Figure 2. Case-control study on *APOE* RMVs identified.**

573 Of the 14 *APOE* RMVs identified in this study, four imputed variants were included in the
574 analysis. Two RMVs, rs140808909 and rs190853081, which were in complete LD and were
575 defined as the *e7* allele, remained statistically significant after multiple comparison corrections.
576 The OR was < 1.0 , suggesting a protective effect against AD. Detailed numerical data are
577 presented in **Table 4**.

578

579 **Figure 3. Case-control study on *APOE* genotypes and alleles in carriers with *APOE* RMVs**

580 **identified.**

581 For the four *APOE* RMVs analyzed in the case-control study (**Table 4, Figure 2**), we conducted

582 a case-control analysis focusing on subjects carrying alternative alleles, examining *APOE*

583 genotypes (**A**) and alleles (**B**). A forest plot showing ORs and 95% CIs is provided. For

584 genotypes (**A**), the OR (95% CI) of $e3*4$ relative to $e3*3$ was calculated. For alleles (**B**), the

585 OR (95% CI) of $e4$ relative to $e3$ was calculated. Bold text indicates that the 95% CI of the OR

586 does not include 1.00. “All” indicates the OR and its 95% CI for all samples (6,471 AD cases

587 and 20,270 controls) as shown in **Table 3**. Ref, reference.

588

589 **Supplementary Figure 1. Meta-analysis of *APOE* genotypes and alleles in NCGG and**

590 **JGSCAD.**

591 A meta-analysis was conducted using *APOE* genotype data from two Japanese cohorts, the

592 NCGG (Table S2) and JGSCAD (Table S4), as reported by Asanomi *et al.* ²¹ The pooled ORs

593 and its 95% CI were calculated using both fixed effects (Fixed) and random effects (Random)

594 models, and are presented in a forest plot. For genotype analysis (**A**), the OR and 95% CI for

595 five genotypes ($e2*2$, $e2*3$, $e2*4$, $e3*4$, $e4*4$) were calculated in comparison to $e3*3$. For allele
596 analysis (**B**), the OR and 95% CI for two alleles ($e2$ and $e4$) were calculated in comparison to
597 $e3$. The OR and 95% CI are highlighted in bold if the 95% CI does not include 1.00. Ref,
598 reference.

599

600 **Supplementary Figure 2. Pairwise LD measures between two *APOE* RMVs, rs140808909**
601 **and rs190853081, determining the $e7$ allele.**

602 Using the LDpair Tool on the LDlink website (<https://ldlink.nih.gov/?tab=ldpair>), we analyzed
603 the haplotypes and their frequencies, as well as the strength of LD (D' and R^2) between the two
604 RMVs, rs140808909 and rs190853081, in the East Asian (EAS) population (504 subjects). Sub-
605 analyses were conducted separately for the Japanese in Tokyo, Japan (JPT, 104 subjects) and
606 the Southern Han Chinese (CHS, 105 subjects) populations. Three A-A haplotypes were
607 observed in the EAS population, one in JPT, and two in CHS. Notably, these two RMVs were
608 not detected in populations outside the EAS database.

609

610 **Supplementary Figure 3. Sequence verification in an individual carrying three *APOE***

611 **RMVs: p.Arg242Gln (rs267606663-G/a), p.Glu262Lys (rs140808909-G/a), and**
612 **p.Glu263Lys (rs190853081-G/a).**

613 Subject ID #35 (*APOE* genotype, $e3^*3$; **Supplementary Table 5**) harbored three RMVs. The
614 results of SS of the genomic DNA are shown in **A**. Three RMVs (rs267606663, rs140808909,
615 and rs190853081) were detected, all of which were G/A heterozygotes. PCR products generated
616 using genomic DNA as a template were cloned into a TA vector and SS of the plasmid DNA
617 was performed. The results are presented in Section **B**. It was found that the alternative allele
618 “a” of rs267606663 (G/a) does not co-segregate with the alternative alleles “a” of the other two
619 RMVs, rs140808909 (G/a) and rs190853081 (G/a), and is located independently on a different
620 chromosome. Asterisks indicate nucleotide substitutions; black represents the reference allele,
621 and red represents the alternative allele.

622

623 **Supplementary Figure 4. A comparison of TC, LDL-C, HDL-C, and TG levels and L/H**
624 **ratios between *APOE* RMV carriers and non-carriers in the J-ADNI and ToMMo cohorts.**

625 Each measurement item is represented as a box-and-whisker plot (whiskers, 5th to 95th
626 percentiles; +, mean) divided into J-ADNI and ToMMo groups and further subdivided into

627 *APOE* RMV carriers and non-carriers. Numbers in parentheses indicate the sample size for each
628 group. The light blue vertical lines indicate reference values for each measurement item: **A**, TC
629 150–219 mg/dL; **B**, LDL-C 70–139 mg/dL; **C**, HDL-C 40–96 mg/dL [females 40–96 mg/dL,
630 and males 40–86 mg/dL]; **D**, TG 50–149 mg/dL; and **E**, LH ratio ≤ 1.5 . No significant
631 differences were observed in any of these comparisons. P_{MW} , P -value of Mann-Whitney U -test.

Table 1. The information on subjects used in the exploration of *APOE* RMVs

Cohort	Group	No. of subjects			Sex		Age (AAE/AAD)			Special mention	
		All	CL	NP	F	M	Mean	Median	SD		Range: Min-Max
NIG	SS	839	0	839	348	491	76.8	79.0	13.1	20 - 103	Including 508 J-ADNI subjects: AD, 140; MCI, 221; and CN, 147
	WES	1,509	826	683	723	786	74.1	75.0	10.9	24 - 111	
	SS+WES	241	0	241	79	162	70.9	71.0	12.8	21 - 103	
ToMMo	3.5KJPNv2	3,307	3,307	0	1,810	1,497	55.9	61.0	14.1	20 - 90	
NIG+ToMMo	-	5,896	4,133	1,763	2,960	2,936	64.1	66.0	16.2	20 - 111	
Cohort	<i>APOE</i> Genotype						Allele				
	<i>e2*2</i>	<i>e2*3</i>	<i>e2*4</i>	<i>e3*3</i>	<i>e3*4</i>	<i>e4*4</i>	<i>e2</i>	<i>e3</i>	<i>e4</i>		
NIG	6 (0.23%)	155 (5.99%)	27 (1.04%)	1,685 (65.08%)	621 (23.99%)	95 (3.67%)	194 (3.75%)	4,146 (80.07%)	838 (16.18%)		
ToMMo	6 (0.18%)	251 (7.59%)	30 (0.91%)	2,428 (73.42%)	564 (17.05%)	28 (0.85%)	293 (4.43%)	5,671 (85.74%)	650 (9.83%)		
NIG+ToMMo	12 (0.20%)	406 (6.89%)	57 (0.97%)	4,113 (69.76%)	1,185 (20.10%)	123 (2.09%)	487 (4.13%)	9,817 (83.25%)	1,488 (12.62%)		

Frequency (%) of each *APOE* genotype/allele in parentheses
AAD, age at death for neuropathologically-characterized subjects
AAE, age at entry/examination for clinically-verified subjects
AD, Alzheimer's disease
All, CL+NP
CL, clinically-verified subjects
CN, cognitively normal
F, female
J-ADNI, Japanese Alzheimer's Disease Neuroimaging Initiative
M, male
Max, maximum
MCI, mild cognitive impairment
Min, minimum
NIG, Niigata University
NP, neuropathologically-characterized subjects
SD, standard deviation
SS, Sanger sequencing
ToMMo, Tohoku Medical Megabank Organization
WES, whole exome sequencing

Table 2. *APOE* RMVs identified

dbSNP (rs ID)	Genomic position (GRCh38.p14)	Allele		Nucleotide (Triplet)	a.a.	Substitution		NIG						ToMMo						Special mention
		Ref (0)	Alt (1)			a.a. position	Genotype			Allele			Genotype			Allele				
							Immature (317 a.a.) [†]	Mature (299 a.a.) [‡]	0/0	0/1	1/1	0	1	MAF	0/0	0/1	1/1	0	1	
RMVs identified in this study																				
rs1969812567 [§]	44,906,643	G	c	GCG > cCG	Ala(A) > Pro(P)	7	Cleaved	2,588	1	0	5,177	1	1.93E-04	3,304	2	0	6,610	2	3.02E-04	
rs121918392	44,907,777	G	a	GAG > aAG	Glu(E) > Lys(K)	21		2,571	18	0	5,160	18	3.48E-03	3,284	23	0	6,591	23	3.48E-03	e5
rs201672011	44,907,807	G	a	GAG > aAG	Glu(E) > Lys(K)	31		2,589	0	0	5,178	0	0.00E+00	3,306	1	0	6,613	1	1.51E-04	
rs762461580	44,907,865	G	a	CGC > CaC	Arg(R) > His(H)	50		2,586	3	0	5,175	3	5.79E-04	3,304	3	0	6,611	3	4.54E-04	
rs370594287	44,907,908	G	c	CAG > CAc	Gln(Q) > His(H)	64		2,587	2	0	5,176	2	3.86E-04	3,306	1	0	6,613	1	1.51E-04	
rs748703149	44,908,729	G	a	GGC > aGC	Gly(G) > Ser(S)	145		2,589	0	0	5,178	0	0.00E+00	3,305	2	0	6,612	2	3.02E-04	
rs954186991	44,908,762	G	a	GCC > aCC	Ala(A) > Thr(T)	156		2,589	0	0	5,178	0	0.00E+00	3,306	1	0	6,613	1	1.51E-04	
rs1339651557	44,908,850	G	a	CGC > CaC	Arg(R) > His(H)	185		2,589	0	0	5,178	0	0.00E+00	3,306	1	0	6,613	1	1.51E-04	
rs1969871875	44,908,852	G	a	GAG > aAG	Glu(E) > Lys(K)	186		2,589	0	0	5,178	0	0.00E+00	3,305	2	0	6,612	2	3.02E-04	
rs1385100920	44,908,913	G	a	GGC > GaC	Gly(G) > Asp(D)	206		2,589	0	0	5,178	0	0.00E+00	3,306	1	0	6,613	1	1.51E-04	
rs267606663 [¶]	44,909,021	G	a	CGG > CaG	Arg(R) > Gln(Q)	242		2,586	3	0	5,175	3	5.79E-04	3,299	7	0	6,605	7	1.06E-03	
rs140808909	44,909,080	G	a	GAG > aAG	Glu(E) > Lys(K)	262		2,555	34	0	5,144	34	6.57E-03	3,260	46	1	6,566	48	7.26E-03	e7
rs190853081	44,909,083	G	a	GAG > aAG	Glu(E) > Lys(K)	263		2,555	34	0	5,144	34	6.57E-03	3,260	46	1	6,566	48	7.26E-03	e7
rs767339630	44,909,129	G	a	CGC > CaC	Arg(R) > His(H)	278		2,589	0	0	5,178	0	0.00E+00	3,305	2	0	6,612	2	3.02E-04	
RMVs previously reported elsewhere																				
rs769452	44,907,853	T	c	CTG > CcG	Leu(L) > Pro(P)	46		2,589	0	0	5,178	0	0.00E+00	NA	NA	NA	NC	NC	NC	Pittsburgh
rs121918393	44,908,756	C	a	CGC > aGC	Arg(R) > Ser(S)	154		2,589	0	0	5,178	0	0.00E+00	NA	NA	NA	NC	NC	NC	Christchurch
rs769455	44,908,783	C	t	CGT > iGT	Arg(R) > Cys(C)	163		2,589	0	0	5,178	0	0.00E+00	NA	NA	NA	NC	NC	NC	
rs199768005	44,909,057	T	a	GTG > GaG	Val(V) > Glu(E)	254		2,589	0	0	5,178	0	0.00E+00	NA	NA	NA	NC	NC	NC	Jacksonville
rs267606661	44,909,101	C	g	CGC > gGC	Arg(R) > Gly(G)	269		2,589	0	0	5,178	0	0.00E+00	NA	NA	NA	NC	NC	NC	

[§], one subject without genotypes for two SNPs, rs1969812567 and rs267606663, was excluded in ToMMo

[†], number containing a signal peptide consisting of 18 amino acids on the N-terminal side

[‡], number excluding the signal peptide consisting of 18 amino acids on the N-terminal side

a.a., amino acid

Alt, alternative (assigned the number one, 1)

MAF, minor allele (= alt. allele 1) frequency

NA, not available

NC, not calculated

Ref, reference (assigned the number zero, 0)

Table 3. The information on subjects used in the case-control study

Genotyping platform	Group	No. of subjects	Sex		Age (AAE/AAD)			APOE								
			F	M _{Ref}	Mean	Median	SD	Genotype			Allele					
								<i>e2</i> *2	<i>e2</i> *3	<i>e2</i> *4	<i>e3</i> *3 _{Ref}	<i>e3</i> *4	<i>e4</i> *4	<i>e2</i>	<i>e3</i> _{Ref}	<i>e4</i>
ASA (imputed)	Cases	6,471	4,159	2,312	76.9	81.0	7.0	4	272	62	3,320	2,365	448	342	9,277	3,323
								(0.06%)	(4.20%)	(0.96%)	(51.31%)	(36.55%)	(6.92%)	(2.64%)	(71.68%)	(25.68%)
	Controls	20,270	10,539	9,731	69.6	74.0	11.7	52	1,633	207	14,727	3,452	199	1,944	34,539	4,057
								(0.26%)	(8.06%)	(1.02%)	(72.65%)	(17.03%)	(0.98%)	(4.80%)	(85.20%)	(10.01%)
	Sum	26,741	14,698	12,043	71.4	54.0	11.2	56	1,905	269	18,047	5,817	647	2,286	43,816	7,380
								(0.21%)	(7.12%)	(1.01%)	(67.49%)	(21.75%)	(2.42%)	(4.27%)	(81.93%)	(13.80%)
			<i>P</i> _{FE}	<0.0001	<i>P</i> _t	<0.0001	<i>P</i> _{FE}	0.0248	<0.0001	0.0571	-	<0.0001	<0.0001	<0.0001	-	<0.0001
			OR	1.66			OR	0.34	0.74	1.33	1.00	3.04	9.99	0.65	1.00	3.05
			95% CI	1.57 - 1.76			95% CI	0.13 - 0.90	0.65 - 0.84	0.99 - 1.77	-	2.85 - 3.24	8.42 - 11.83	0.58 - 0.74	-	2.90 - 3.21

Frequency (%) of each *APOE* genotype/allele in parentheses.

P < 0.05 is shown in bold.

AAD, age at death for neuropathologically-characterized subjects

AAE, age at entry/examination for clinically-verified subjects

ASA, asian screening array (Illumina)

CI, confidence interval

F, female

M, male

max, maximum

min, minimum

OR, odds ratio (crude)

*P*_{FE}, *P* value for Fisher's exact test

*P*_{MW}, *P* value calculated by Mann-Whitney test

*P*_t, unpaired *t*-test with Welch's correction

Ref, reference

SD, standard deviation

Table 4. Case-control study on *APOE* RMVs identified

dbSNP (rs ID)	Allele		No. of subjects		No. of RMV carriers		MAF		<i>P</i>		OR	95% CI	Special mention
	Ref	Alt	Cases	Controls	Cases	Controls	Cases	Controls	Unadjusted	Bonferroni [†]			
<u>ASA: imputed</u>													
rs121918392	G	a	6,413	19,999	57	194	4.41E-03	4.80E-03	5.63E-01	1.00E+00	0.92	0.68 - 1.23	<i>e5</i>
rs762461580	G	a	6,443	20,121	27	72	2.09E-03	1.81E-03	5.27E-01	1.00E+00	1.15	0.74 - 1.78	
rs140808909	G	a	6,402	19,893	68	300	5.26E-03	7.45E-03	8.95E-03	3.58E-02	0.70	0.54 - 0.92	<i>e7</i>
rs190853081	G	a	6,402	19,893	68	300	5.26E-03	7.45E-03	8.95E-03	3.58E-02	0.70	0.54 - 0.92	<i>e7</i>

The OR and its 95% CI are illustrated visually in **Figure 2**.

Alt, alternative

ASA, asian screening array (Illumina)

CI, confidence interval

MAF, minor allele frequency

OR, odds ratio (adjusted)

Ref, reference

RMV, rare missense variant

[†], Bonferroni-corrected *P* value, calculated by setting the number of tests to four

Fig. 1. *APOE* rare missense variants identified in this study

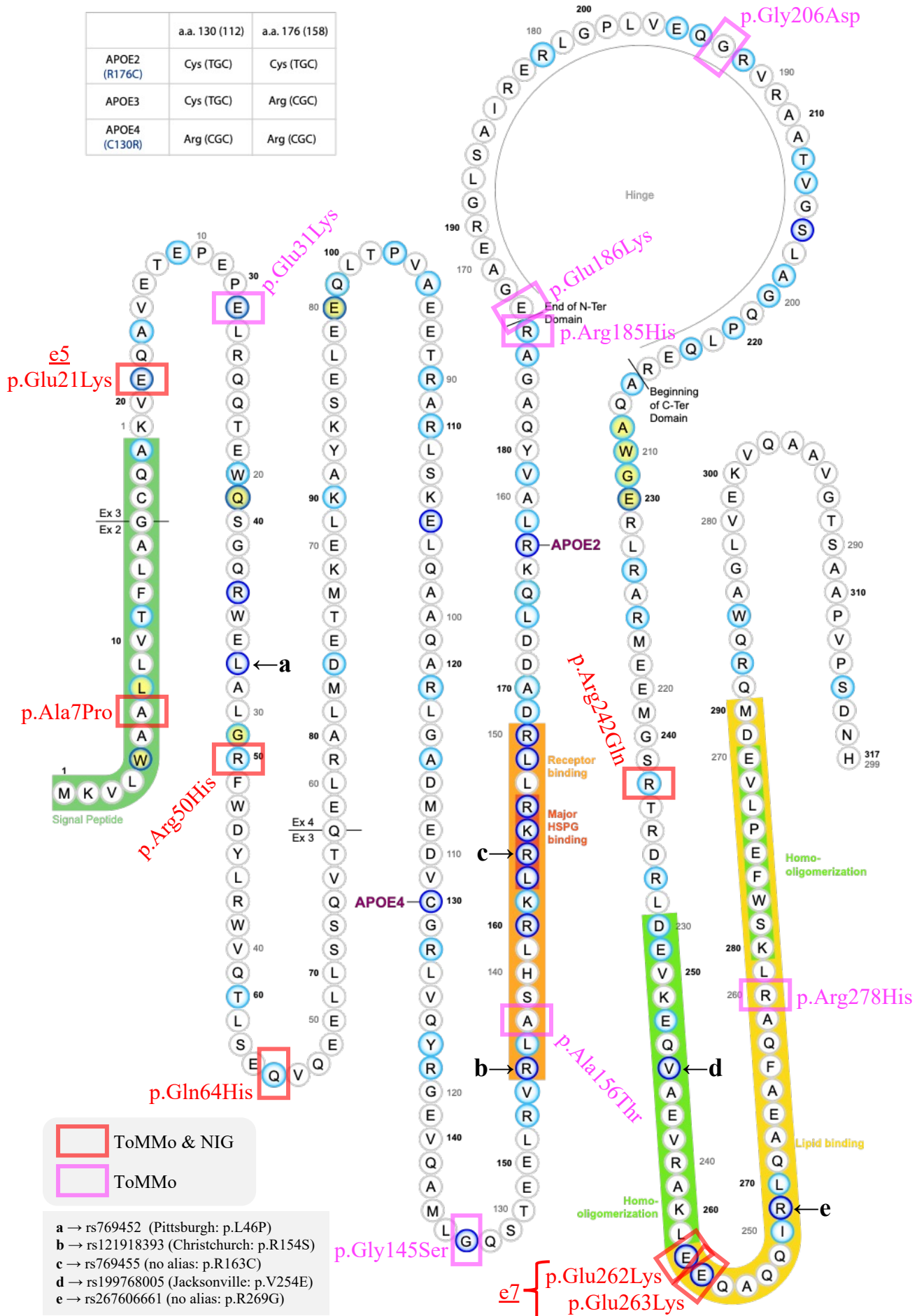


Fig. 2. Case-control study on *APOE* RMVs identified

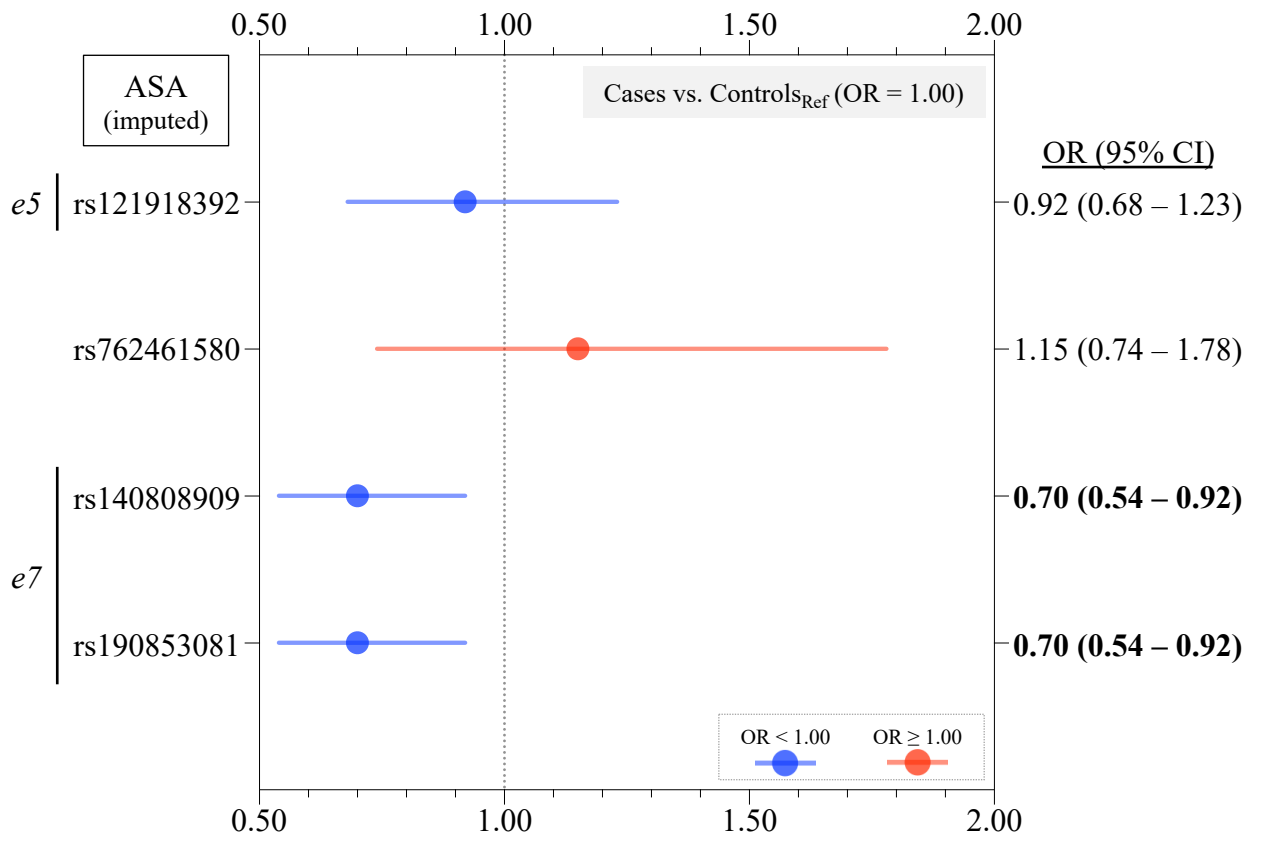


Fig. 3. Case-control study on *APOE* genotypes and alleles in carriers with *APOE* RMVs identified

