Neurogenetics (2004) 5:201–208
DOI 10.1007/s10048-004-0189-9

ORIGINAL ARTICLE

Kristin K. Nicodemus · Judith E. Stenger · Donald E. Schmechel · Kathleen A. Welsh-Bohmer · Ann M. Saunders · Allen D. Roses · John R. Gilbert · Jeffery M. Vance · Jonathan L. Haines · Margaret A. Pericak-Vance · Eden R. Martin

Comprehensive association analysis of *APOE* regulatory region polymorphisms in Alzheimer disease

Received: 16 April 2004 / Accepted: 15 July 2004 / Published online: 29 September 2004

Abstract Several recent case-control studies have examined the association between single nucleotide polymorphisms (SNPs) in the promoter region of the apolipoprotein E gene (*APOE*) and risk of Alzheimer disease (AD), with conflicting results. We assessed the relation between five *APOE* region SNPs and risk of AD in both case-control and family-based analyses. We observed a statistically significant association with the +5361T allele in the overall case-control analysis (*P* value=0.04) after adjusting for the known effect of the *APOE-4* allele. Further analysis revealed this association to be limited to carriers of the *APOE-4* allele. Age-stratified analyses in the patients with age at onset of 80 years or greater and age-matched controls showed that the −219T allele (*P* value=0.009) and the +113C allele (*P* value=0.03) are associated with increased risk of AD. Despite these findings, haplotype and family-based association analyses were unable to provide evidence that the *APOE* region SNPs influenced risk of AD independent of the *APOE-4* allele. In addition to risk, we tested for association between the SNPs and age at onset of AD, but found no association in the case-control or family based analyses. In conclusion, SNPs +5361, or a SNP in strong linkage disequilibrium, may confer some additional risk for developing AD beyond the risk due to *APOE-4*; however, the effect independent of *APOE-4* is likely to be small.

Keywords Alzheimer disease · *APOE* · Single nucleotide polymorphisms · Haplotype · Age at onset

K. K. Nicodemus · J. E. Stenger · J. R. Gilbert · J. M. Vance · M. A. Pericak-Vance
Department of Medicine and Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, USA

K. K. Nicodemus
Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

D. E. Schmechel · K. A. Welsh-Bohmer
Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina, USA

D. E. Schmechel · K. A. Welsh-Bohmer
Joseph and Kathleen Bryan Alzheimer's Disease Research Center, North Carolina, USA

A. M. Saunders · A. D. Roses
GlaxoSmithKline Research and Development, Research Triangle Park, North Carolina, USA

J. L. Haines
Program in Human Genetics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

E. R. Martin (✉)
Department of Medicine and Center for Human Genetics, Duke University Medical Center, DUMC Box 3445, Durham, NC 27710, USA
e-mail: emartin@chg.duhs.duke.edu
Tel.: +1-919-6840601
Fax: +1-919-6840915

Introduction

Apolipoprotein E (APOE) is a well established genetic risk factor for late-onset Alzheimer disease (AD), with the *APOE-4* allele increasing risk and the *APOE-2* allele decreasing risk relative to the most-common *APOE-3* allele [1, 2]. However, the presence of the *APOE-4* allele is neither necessary nor sufficient to cause AD, indicating that additional genetic or non-genetic factors influencing AD risk are yet to be identified. Several other polymorphisms within the promoter region of *APOE* and the APOE-coding sequence have been described [3], and there has been great interest in determining if polymorphisms with potential regulatory function confer additional risk for AD beyond the risk associated with *APOE-4*.

Studies of transcriptional activity have demonstrated that two single nucleotide polymorphisms (SNPs) in the *APOE* promoter, –491A/T and –427C/T, up-regulate transcription of the *APOE* gene, possibly influencing the

risk of AD [4, 5]. Several studies have assessed the effect on AD of these SNPs and an additional promoter SNP, –219G/T, in association studies [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. Results have been inconsistent partly because of the difficulty separating the effect of these SNPs, if any, from the effect of the classical *APOE* alleles. Determining an independent effect is complicated by the strong linkage disequilibrium (LD) between these SNPs and the classical *APOE* alleles.

We sought to assess the association between five SNPs in the *APOE* region: three in the promoter region (–491A/T, –427C/T, and –219G/T), one within the APOE-coding region (+113C/G), and one that is 862 bp from the 3′ end of the *APOE* gene coding sequence mapping to the tail end of a putative regulatory region (+5361C/T) [3]. We tested for association with AD in both a large case-control sample and large family sample. This approach provides the power of a case-control test complemented by the protection against false-positive results due to population stratification offered by the family based analysis. In addition to testing for effects on risk, we examined the effect of these SNPs on age at onset in both familial and singleton AD cases.

Materials and methods

Study populations

Case-control sample

Our case-control sample consists of 293 unrelated cases with probable or definite AD (singleton cases) and 298 unrelated controls (mostly spouses of AD patients) ascertained through The Joseph and Kathleen Bryan Alzheimer's Disease Research Center (ADRC) and the Center for Human Genetics (CHG) at Duke University. Clinical diagnosis of AD was based on consensus criteria [18]. The reported age at onset of AD patients was defined as the age at which the caregiver, family, and/or individual first noted cognitive problems (most often short-term memory loss and more rarely other problems such as dysphasia or disorientation to time or place followed closely by memory change) sufficient to interfere with independent daily activities. Controls had no obvious signs of cognitive or neurological impairment when enrolled in the study as determined by personal interview by clinical personnel of Duke CHG or ADRC. All individuals included in this study are Caucasian. This is a unique resource, and these *APOE* polymorphisms have not been tested previously in this sample.

Family sample

Our family sample includes three sets of families: the Collaborative Alzheimer Project families (ascertained through The Joseph and Kathleen Bryan ADRC and the CHG at Duke University, and the University of California at Los Angeles Neuro-psychiatric Institute), families ascertained by the National Institutes of Mental Health (NIMH) Alzheimer Disease Genetics Initiative, and families from the Indiana University Alzheimer's Disease Center's National Cell Repository (IU). All of the families (n =248) include at least one family member probably or definitely affected with AD [18], and at least one unaffected family member sampled. Criteria for AD diagnosis and screening of unaffected relatives were the same as described above, except for families ascertained by IU who classified unaffected individuals on the basis of self-report. A total of 592 affected and 452 unaffected family members were genotyped in this study. All individuals were Caucasian.

SNPs and genotyping

A map of the *APOE* region and SNPs studied is shown in Fig. 1. The *APOE* gene structure, transcript information, and protein positions were extracted from the public Ensembl server (http://www.ensembl.org). SNP positions were verified by using primers and/or surrounding sequence data to query the NIH human genome

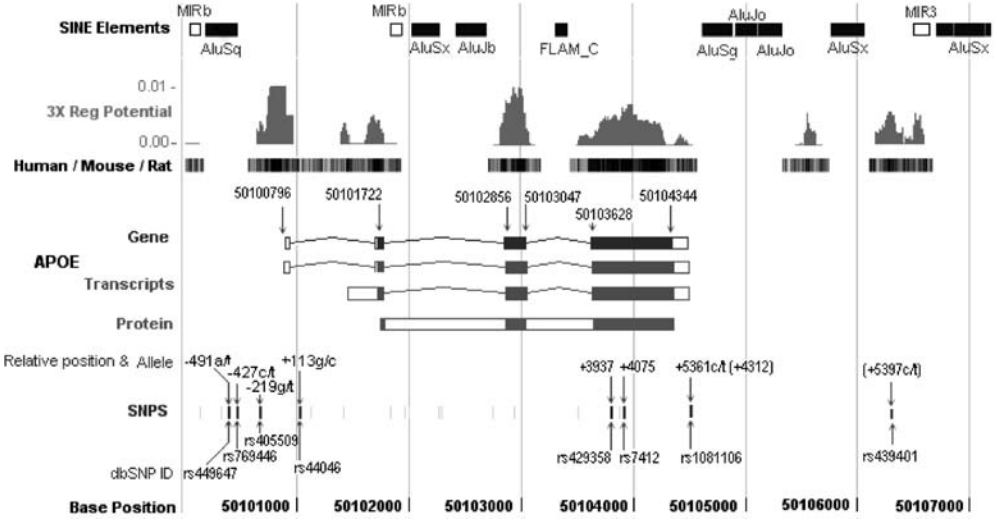

Fig. 1 Gene structure and relevant features of *APOE* and surrounding sequences. Features shown include: short interspersed repetitive elements (SINE), gene transcript and protein information, potential regulatory regions, regions of conservation, and locations of single nucleotide polymorphisms (SNPs) studied. The relevant SNPs are labeled with the dbSNP rs number, as well as the designation given by Nickerson et al. [3], and in parentheses are the actual positions, using the original nomenclature in which the positions are numbered with their position relative to the start of transcription. SNPs +3937 C/T and +4075 C/T are the classical *APOE* alleles

Table 1 Primers used for identification of *APOE* region single nucleotide polymorphisms (SNP). All primers are oriented 5′ to 3′

SNP	Primer	
–491 A/T	Forward	CACGCC TGGCTAACTTTTGT
	Reverse	CACAGTGGGCGAATCACTTA
–427 C/T	Forward	TGACCTTAAGTGATTCGCCC
	Reverse	CTAGGGGGC TGGACAGAAGT
–219 G/T	Forward	CTCCACATTCCCCTTCCAC
	Reverse	AGTCCCCAGGAAGGGAGGA
+113 G/C	Forward	GCTCAGGGGCCT CTAGAAAG
	Reverse	CTCCTCCTCTCCCCAAGC
+5361 C/T	Forward	CCACCTTGGCCTCCTGAGTA
	Reverse	GCAACATATTGAGACCTT GTCTCTACA

sequence assembly build No. 34 at UCSC using BLAT [19]. NCBI dbSNP was queried extensively to identify the official designation and sequence information for +5361C/T (which lays 1,049 bp closer to the start of transcription than expected). Using the UCSC genome browser bioinformatics tools, human and rodent DNA (*Mus musculus* and *Rattus norvegicus*) were aligned to determine conserved sequences and their regulatory potentials, based largely on phylogenetic footprinting [20]. Sequences that score high as potential regulatory regions correlate well with sequences highly conserved in mammals.

Using the Ensembl and UCSC and NCBI genome browsers as well as NCBI's HomoloGene we also looked for synteny between more distantly related organisms but found no orthologous genomic regions in any sequenced vertebrate outside of mammals. The apolipoprotein domain (pfam01442) is conserved in several avian and teleost species, but we are not aware of any phenotypic model for AD in these organisms.

Genomic DNA was obtained from the repositories (NIMH, IU) or isolated from whole blood samples by the Duke CHG DNA Banking Core using Puregene (Gentra Systems, Minneapolis, Minn., USA). The *APOE* promoter polymorphisms were genotyped using the 5′ nuclease allelic discrimination Taqman assay in a 384-well format on the ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, Calif., USA) as described by Oliveira et al. [21]. The following probes were developed (all are oriented 5′ to 3′): SNP –491A/T (rs449647), ccaggctggtctcaa**A/T**ctcctgacc; SNP –427C/T (rs769446), caggcgtgagc**C**accgccc and attacaggcgtgagc**T**accgccc; SNP –219G/T (rs405509), agggtgtctg**G/T**attactgggcga; SNP 113G/C (rs44046), tgggaa **G/C** ccctggcctcca; and SNP 5361C/T (rs1081106), ccagctttt**C/T**attattatt. Primers used are shown in Table 1 and were developed utilizing the publicly available MITPrimer3 software [22]. The classical *APOE* alleles [corresponding to allele combinations at SNP +3937 (rs429358) and SNP +4075 (rs7412)] were genotyped as previously reported [2].

Statistical methods

Descriptive analyses

We tested for LD and deviations from Hardy-Weinberg equilibrium (HWE) using the program Genetic Data Analysis (GDA) [23]. We report exact tests with 3,200 replications for LD and HWE for overall cases (unrelated cases plus one case from each family) and unrelated controls separately. The LD measure D' was calculated using the GOLD program [24].

Case-control analyses

To improve statistical power, we selected at random one affected individual from each family in the family sample to pool with the unrelated cases for analysis (total number of cases=547, proportion of unrelated singleton cases=54%). Unaffected family members were not included in the case-control analysis since their genotypes are not independent of the family cases. Case-control analyses for single alleles and for genotypes were conducted using logistic regression from the SAS program (SAS Institute, Cary, N.C., USA, version 8.1). Exact logistic regression was used for comparisons with sparse data. Models tested included a genotype-based model where both genotypes (e.g., AA and Aa) were evaluated versus a reference genotype homozygous for the more-frequent allele (aa); and an allele-based model that assumes an additive effect on the log scale for the alleles (e.g., having no A alleles=0, having one A allele=1, having two A alleles=2). The genotype- and allele-based models led to similar conclusions for all comparisons, thus we present results only from the allele-based models.

To control for possible confounding due to sex and age, we included sex and age at examination (AAE) as covariates in the regression analysis. In addition, because we were interested in the independent effect of the *APOE* region SNPs, association analyses in the overall sample were conducted including two terms in the model to control for the effect of *APOE-4* (*APOE 4/4*, and *APOE 4/2* plus *APOE 4/3* combined, versus the reference group of individuals with no *APOE-4* alleles). The presence or absence of *APOE-2* also was included in all models; however, the parameter estimate was not significant and did not substantially change other parameter estimates, thus the term was dropped from the final analyses.

To assess the association between AD and SNP haplotypes, we used the program haplo.score, which implements the test proposed by Schaid et al. [25].

To test for association between the *APOE* region SNPs and age at onset (AAO) of AD, AAO for the cases was considered as a dependent variable within a case-only analysis. To test allelic and genotype associations, linear regression was performed with AAO as the dependent variable, with independent variables coded as above and using the same covariates, using SAS version 8.1. To test for association between AAO and SNP haplotypes the program haplo.score was used [25].

Family based association analyses

Analysis of family data was conducted using the pedigree disequilibrium test (PDT) for single-locus tests [26, 27]. In addition, we used an extension of the PDT, the genotype-PDT (GenoPDT), to assess association between genotypes and risk of AD in the family data [28]. To evaluate association between SNP haplotypes we used the haplotype family based association test (HBAT), an extension of the FBAT program [29]. Global and single-haplotype *P* values for the haplotypic test were based on permutation tests with 10,000 replicates.

To assess the effect of the *APOE* region SNPs on AAO in the families, we used the variance components method implemented in the QTDT package [30]. Although we tested both the Fulker model for sibpairs [31] and the orthogonal model for complete pedigrees, *P* values are reported for the orthogonal model tests only. *P* values for the orthogonal model tests did not differ significantly from the Fulker model *P* values. All *P* values reported are based on 1000 Monte Carlo Markov Chain permutations. For all analyses a *P* value of 0.05 or smaller was considered statistically significant.

Results

Descriptive statistics

In the overall case-control set, *t* -tests revealed a significant difference between mean AAE in cases and controls (mean for cases=74.3, SD±8.41, mean for controls=68.0, SD±8.86). In addition, the proportion of females was higher (65.8%) in the case sample than the control sample (52.0%) (*P* value<0.001). The proportion of females was higher in the unrelated singleton cases (37.6%) than in

Table 2 Pairwise linkage disequilibrium (top of diagonal—exact test *P* values; below diagonal—D' values) in the overall case sample (*n* =547) and unrelated controls (*n* =298)

Overall cases						
	–491AT	–427CT	–219GT	+113CG	*APOE*	+5361CT
–491AT		0.10	0.02	<0.00001	<0.00001	0.46
–427CT	0.663		0.12	0.42	0.007	0.74
–219GT	0.244	0.112		<0.00001	<0.00001	<0.00001
+113CG	0.187	0.359	0.979		<0.00001	0.02
APOE	0.361	0.110	0.502	0.408		<0.00001
+5361CT	0.600	0.675	0.984	1.00	0.702	
Unrelated controls						
	–491AT	–427CT	–219GT	+113CG	*APOE*	+5361CT
–491AT		0.02	0.13	0.0006	<0.00001	0.03
–427CT	0.998		0.39	0.10	<0.00001	0.14
–219GT	0.256	0.157		<0.00001	<0.00001	<0.00001
+113CG	0.309	0.123	0.981		<0.00001	<0.00001
APOE	0.304	0.276	0.363	0.295		0.37
+5361CT	0.998	0.995	0.999	0.999	0.384	

Table 3 *APOE* genotypic SNP percentages in overall cases and controls and by *APOE* genotype subgroups

SNP	Genotype	Overall		*APOE-4*-positive		*APOE-4*-negative	
		Cases	Controls	Cases	Controls	Cases	Controls
		(*n* =547)	(*n* =298)	(*n* =376)	(*n* =78)	(*n* =171)	(*n* =220)
–491A/T	A/A	77.33	66.78	85.90	80.77	58.48	61.82
	A/T	20.84	28.19	14.10	16.67	35.67	32.27
	T/T	1.83	5.03	0.0	2.56	5.85	5.91
–427C/T	C/C	0.73	1.34	0.80	1.28	0.58	1.36
	C/T	12.07	17.45	11.70	15.38	12.87	18.18
	T/T	87.20	81.21	87.50	83.33	86.55	80.45
–219G/T	G/G	18.10	28.19	10.11	10.26	35.67	34.55
	G/T	48.08	48.32	49.73	62.82	44.44	43.18
	T/T	33.82	23.49	40.16	26.92	19.88	22.27
+113C/G	C/C	5.67	16.78	0.53	0.0	16.96	22.73
	C/G	36.56	38.93	31.91	32.05	46.78	41.36
	G/G	57.77	44.30	67.55	67.95	36.26	35.91
+5361C/T	C/C	0.0	0.34	0.0	0.0	0.0	0.45
	C/T	9.51	19.13	5.59	12.82	18.13	21.36
	T/T	90.49	80.54	94.41	87.18	81.87	78.18

cases from the family sample (26.1%) (*P* value=0.01) and the mean age at onset was slightly lower in unrelated singleton cases (67) versus familial cases (68) (*t* -test *P* value=0.01). However, *APOE* region SNP genotype frequencies were not significantly different for any SNP between unrelated singleton cases and cases from the family sample (data not shown).

There was no significant evidence for deviation from HWE for all SNPs except the –219G/T SNP (*P* =0.008) in the unrelated controls. Review of genotypes for this marker did not reveal obvious genotyping errors, and 200 of 202 duplicated samples had exact matches of genotype calls for this marker (less than 1% error). Furthermore, –219G/T SNP did not deviate from HWE in unaffected family members or the overall case sample for this SNP. Thus it is unlikely that the deviation in unrelated controls represents substantive genotyping error or HW disequilibrium in the population. There was significant LD between the SNPs and classical *APOE* alleles across the entire region (Table 2). In addition, several SNPs were in strong LD with one another.

Case-control analyses

Table 3 shows genotype frequencies in overall cases (unrelated singleton cases plus one case from each family) and unrelated controls in *APOE-4*-positive and *APOE-4*-negative subgroups. Genotype frequencies were similar to those previously reported [6].

Results from significant tests for association between SNPs and AD in cases and controls are shown in Table 4. For comparison, the odds ratios (ORs) and associated 95% confidence intervals (CIs) in the overall sample for having one *APOE-4* allele and having two *APOE-4* alleles were OR=4.34, CI=3.19–6.44 and OR=20.78, CI=9.25–46.7, respectively. We began by testing for association between the SNPs and disease in the overall sample adjusting for the effect of the *APOE-4* allele. Only SNP +5361C/T was found to be associated in the overall sample after controlling for the effect of *APOE-4* status (Table 4), with the +5361T allele positively associated with risk of AD (*P* value=0.04).

As a further examination, we tested the SNPs for association in groups of cases and controls stratified by *APOE* genotype: *APOE-4/3* and *APOE-3/3* (Table 4).

Table 4 Case-control association analyses: overall and stratified by *APOE* genotype. All results are adjusted for sex and age at examination. Only significant results are shown (*OR* odds ratio, *CI* confidence interval)

Group	Sample size	SNP	Positively associated allele	OR	95% CI	*P* value*
Overall (adjusted for *APOE-4*)	547 cases (9.5% C allele carriers, *n* =52) 298 controls (19.5% C allele carriers, *n* =57)	+5361	T	1.63	(1.03–2.60)	0.04
Group	**Sample size**	**SNP**	**Positively associated allele**	**OR**	**95% CI**	***P* value****
Stratified by *APOE* genotype		+5361	T	3.21	(1.17–8.76)	0.02
APOE-4/3	260 cases (7.7% C allele carriers, *n* =20) 64 controls (15.6% C allele carriers, *n* =10)					
APOE-3/3	148 cases 178 controls	None significant				

*Asymptotic normal *P* value
**Exact test *P* value

Table 5 Case-control association analyses stratified in cases with age at onset (AAO) ≥80 years and controls with age at examination (AAE) >80 years. All results are adjusted for sex and *APOE* status and are based on exact logistic regression. Only significant results shown

Group	Sample size	SNP	Positively associated allele	OR	95% CI	*P* value
Age 80 years or greater	55 cases 27 controls	−219 [a] +113 [b]	T C	3.51 3.21	(1.31–10.95) (1.12–10.67)	0.009 0.03

[a] Case/control frequencies (numbers) for −219: G/G 16.4 (9)/55.6 (15), G/T 56.4 (31)/29.6 (8), 27.3 (15)/14.8 (4)
[b] Case/control frequencies (numbers) for +113: G/G 43.6 (24)/59.3 (16), G/C 47.3 (26)/25.9 (7), C/C 9.1 (5)/14.8 (4)

There were too few controls with *APOE-4/4* genotype and too few individuals carrying *APOE-2* to allow for meaningful inference in these strata. Significant evidence for association was found only in the *APOE-4/3* stratum, with SNP +5361T again showing positive association with risk (exact *P* value=0.02). Although it did not reach our nominal significance level, SNP +113C was close to significant (exact *P* value=0.06) in this stratum. Since we observed significant results in *APOE-4/3* individuals and not in individuals with *APOE-3/3*, we also tested for interaction between the *APOE* SNPs and the *APOE-4* allele; however, no significant interaction was detected.

Because others have shown that the promoter SNPs have effects limited to cases with very late-onset AD [6, 12], we also tested for association in cases with AAO≥80 years and age-matched controls with AAE≥80 years. Of the 55 cases with AOO≥≥80 years, 43 (78.2%) were unrelated singleton cases. We found alleles +219 T and +113 C to be significantly positively associated with AD (exact *P* values=0.009 and 0.03, respectively), after adjusting for the effect of *APOE-4* (Table 5). SNP +5361C/T, which was significant in the overall case-control analysis, did not reach our nominal significance level of 0.05 in the age ≥80 years subgroup; however, the *P* value of 0.06 was close to significant and the trend was the same with +5361T being more frequent in cases than controls. Although we also tested for association in other age groups (70–79, 60–69, and <59), the AAO ≥80 years group was the only age group that showed significant results for any markers.

Table 6 Overall case-control and family based haplotype association analyses. Only significant results are shown

Analysis set	Haplotype[a]	Estimated frequency	Direction of association[b]	*P* value
Case-control				
	A-T-T-G-4-T	0.219	+	<0.001
	A-T-G-G-4-T	0.073	+	<0.001
	A-T-G-G-3-T	0.230	−	0.035
	A-T-T-C-3-T	0.169	−	<0.0001
	T-T-T-C-3-T	0.071	−	<0.0001
	A-T-G-G-3-C	0.062	−	<0.0001
	T-T-G-G-2-T	0.031	−	0.004
Family based				
	A-T-T-G-4-T	0.258	+	<0.00001
	A-T-G-G-4-T	0.073	+	0.0002
	A-T-T-C-3-T	0.168	−	0.0002

[a] Haplotypes indicated by SNP alleles with markers in the order: −491A/T, −427C/T, −219G/T, +113C/G, *APOE* alleles 2/3/4, +5361C/T
[b] Direction is + if haplotype is positively associated with AD and − if haplotype is negatively associated with AD

In the overall dataset, haplotype analysis of all five SNPs and the classical *APOE* alleles simultaneously revealed significant associations between several haplotypes and AD (Table 6). However, the only haplotypes that were positively associated with disease contained *APOE-4*. Thus, the haplotype analysis revealed no evidence of positive association with AD risk beyond the effect of *APOE-4*. Further analyses at the five *APOE* SNPs in samples stratified by genotype at the classical *APOE* locus (*APOE-4/3*and *APOE-3/3*) showed no dif-

ferences in haplotype frequencies between cases and controls.

To explore whether these polymorphisms have any effect on AAO of AD, we tested for association in the overall case sample (unrelated cases and one randomly selected case from each family). We found no evidence of association with any of the *APOE* region SNP alleles, genotypes or haplotypes and AAO of AD in the overall case sample or the sample stratified by *APOE* genotype.

Family based analysis

For comparison, the *P* values for the classical *APOE* alleles in the overall family group using the PDT were $P = 6\times10^{-7}$ for *APOE-2*, $P = 1\times10^{-9}$ for *APOE-3*, and $P = 2\times10^{-13}$ for *APOE-4*, with an overall global test *P* value of 6×10^{-17}.

Because of the strong LD between the *APOE* region SNPs and the classical *APOE* alleles, we conducted analysis using a conditional PDT test. This test is conducted by considering only family members with the same genotype at the *APOE* locus to control for possible confounding by the classical *APOE* alleles. The conditional PDT conducted conditioning on *APOE-3/3* genotype and *APOE-4/3* genotype yielded no significant results. This suggests that no detectable independent effect of the *APOE* region SNPs exists in these families. Conditioning on other genotypes was not conducted because of small numbers of families in these strata.

Haplotype association analyses, considering all five SNPs and the *APOE* locus, were conducted in the overall family sample (Table 6). Global tests of expected haplotype frequencies transmitted to affected offspring versus observed values revealed a significant association between the six-marker haplotype and AD, as expected (*P* value <0.0001). Similar to the case-control analyses, in the overall family group, only haplotypes containing the *APOE-4* allele were positively associated with AD. Analyses in *APOE-4/3* and *3/3* genotype strata showed no association with AD. Examination of the haplotype analyses reveals similar frequencies and patterns of association in family and case-control analyses.

Association between AAO of AD and *APOE* region SNPs was assessed in families using a conditional QTDT test, with sex as a covariate. As in our PDT analyses, we considered only family members with the same *APOE* genotype. No SNPs showed any association with age at AD onset.

Discussion

We have conducted a comprehensive analysis of five *APOE* polymorphisms in both case-control and family based samples. One SNP, SNP +5361C/T, showed evidence of association when we accounted for the *APOE-4* allele in the case-control analysis. Interestingly, this polymorphism showed no evidence of association in individuals without an *APOE-4* allele. This was demonstrated by the lack of association among cases and controls carrying an *APOE-3/3* genotype, and further supported by the haplotype analyses, which found positive association only for haplotypes carrying *APOE-4*. This suggests the possibility of interaction between the SNPs and *APOE-4*, so that the SNPs only influence risk if an individual carries *APOE-4*; however, statistical tests for interaction were not significant.

Despite the association with SNP +5361 in the case-control analysis, we were not able to demonstrate an effect independent of *APOE-4* in our family-based analysis. There are several possible explanations for this discrepancy. One possibility is that this polymorphism really does have a modest effect on AD risk, but we had sufficient power to detect the association only in the case-control analysis. Our case-control analysis would be expected to be more powerful than the family based analysis, not only because unrelated controls may provide more power [32], but also because we pooled independent family cases and unrelated singleton cases to practically double the size of the case sample. Notably, the results were not significant when only unrelated singleton cases were used.

Another possible explanation for the discrepancy between family and case-control analysis results is that there really is no independent effect of these *APOE* polymorphisms, and the significant case-control result is a random false-positive result. Given the number of different analyses that we conducted (case-control and family based over multiple markers and data stratifications), the most significant *P* value of 0.009 is not overwhelmingly significant and may simply have resulted by chance. Although many of our tests are correlated making an adjusted significance level difficult to calculate, we have conducted more than five independent tests (considering our AAO stratifications alone), thus even a liberal correction for the number of independent tests would make this result non-significant. Replication will be required to provide further support that these are not random false-positive findings.

A third possibility is that the case-control result is false positive due to population stratification. Examination of the allele and haplotype frequencies between singleton cases and cases from families used in the case-control analyses did not reveal significant differences, thus combining family and singleton cases in the case-control tests is unlikely to have lead to excessive false-positive results. However, though we have been careful to sample a well-matched set of unrelated cases and controls, composed entirely of Caucasians ascertained at the Duke ADRC and CHG, we cannot completely rule out hidden population stratification as a factor.

To our knowledge, the +5361 SNP has not been studied for association with AD. This SNP lies in a downstream regulatory region, at the end of a potential enhancer region identified by using the 2 and 3-Way Regulatory Potential tools [20], and although it has no known functional significance the SNP may affect AD

risk by influencing gene expression. Alternatively, our significant finding of association with this SNP may be due to LD between this SNP and an untested polymorphism, in the regulatory region or gene itself, that is responsible for the increase in risk. Further studies of this polymorphism and others in LD with it will be required to determine the underlying biological mechanism.

The other four SNPs that we tested have been examined in several studies. Bullido and Valdivieso [9] found association with +113C/G in AD cases and controls, and reported an increase in risk conferred by the +113 C allele and the +113-C/G genotype in *APOE-4*-negative individuals. Our analyses for +113C/G were not statistically significant, although the tests had *P* values of 0.06 in *APOE-4/3* individuals and the AAO≥80 years group, with the +113 C allele more frequent in cases than controls in both analyses. Other studies have reported association with AD independent of *APOE-4* with each of the promoter SNPs that we studied (-491A/T, –427C/T, and –219G/T) [6, 7, 10, 14, 33, 34], but results have been inconsistent. We found no association with any of these polymorphisms and AD when we accounted for the known effect of the classical *APOE* alleles.

Our study has the advantage of having both a large set of unrelated cases and controls and a large, well-characterized family sample, thus our analyses should have comparable power to the other studies. To explore our statistical power to detect allele frequency differences between cases and controls, we estimated power at each marker. We considered power for the additive logistic regression model for SNP alleles adjusting for the number of *APOE-4* alleles. The correlation between *APOE-4* and SNP allele covariates in the model was specified by estimating the amount of LD. These power calculations show that we should have had 80% power to detect a minimum allele frequency difference between cases and controls ranging from 0.044 to 0.088 for the different SNPs. If the SNPs have more subtle effects leading to smaller allele frequency differences we may have not been able to detect the effect. However, these power calculations suggest that it is unlikely that these polymorphisms have a substantial effect in our overall sample.

Perhaps most interesting was our finding of association in the very late-onset AD cases (AAO ≥80 years). Since risk of AD due to *APOE-4* is strongest in individuals with AAO between 60 and 70 years [1], it is this very late onset AD that must be explained by other factors. Our findings are consistent with those of Lambert et al. [6], who reported increased risk, independent of *APOE* genotype, conferred by the –219 T/T genotype was limited to individuals with AAO greater than 80 years. Heijmans et al. [12] reported significant positive association of the –219 T allele with risk of dementia in a cohort study of Caucasian individuals aged 85 years and older with *APOE-3/3* genotype, showing the –219G/T polymorphism confers excess risk of dementia independent of *APOE-4* in this older sample. We also found an association with AD and +113C/G in the AAO≥80 years subgroup, but not in individuals with earlier onset. This could indicate a direct effect of the +113C/G polymorphism on very late-onset AD, but also could result from the strong LD between this SNP and –219G/T.

In conclusion, our failure to replicate previous findings of association with *APOE* promoter polymorphisms in our large case-control and family samples suggests that these polymorphisms do not have a strong effect on risk of general late-onset AD in American Caucasians. We did, however, find modest evidence of association of the +5361 SNP after adjusting for *APOE* status. Although it is likely that the effect on risk independent of the classical *APOE* alleles is small, this result suggests that it may be worthwhile for future studies to examine this SNP and other polymorphisms in upstream regulatory regions.

Acknowledgements We thank all the families whose participation made this project possible. This research was supported by grants from the National Institutes of Health: R01 AG20135, R01 NS31153, R01 AG19757, R01 AG021547, and grants from the Alzheimer Association, including a Zenith award ZEN-01–2935. We also gratefully acknowledge the personnel at the Duke Center for Human Genetics who played an important part in this research.

References

1. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261:921–923
2. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan DR, Alberts MJ (1993) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology 43:1467–1472
3. Nickerson DA, Taylor SL, Fullerton SM, Weiss KM, Clark AG, Stengard JH, Salomaa V, Boerwinkle E, Sing CF (2000) Sequence diversity and large-scale typing of SNPs in the human apolipoprotein E gene. Genome Res 10:1532–1545
4. Artiga MJ, Bullido MJ, Sastre I, Recuero M, García MA, Aldudo J, Vázquez J, Valdivieso F (1998) Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. FEBS Lett 421:105–108
5. Artiga MJ, Bullido MJ, Frank A, Sastre I, Recuero M, García MA, Lendon CL, Han SW, Morris JC, Vázquez J, Goate A, Valdivieso F (1998) Risk for Alzheimer's disease correlates with transcriptional activity of the *APOE* gene. Hum Mol Genet 7:1887–1892
6. Lambert JC, Araria-Goumidi L, Myllykangas L, Ellis C, Wang JC, Bullido MJ, Harris JM, Artiga MJ, Hernandez D, Kwon JM, Frigard B, Petersen RC, Cumming AM, Pasquier F, Sastre I, Tienari PJ, Frank A, Sulkava R, Morris JC, St Clair D, Mann DM, Wavrant-DeVrieze F, Ezquerra-Trabalon M, Amouyel P, Hardy J, Haltia M, Valdivieso F, Goate AM, Perez-Tur J, Lendon CL, Chartier-Harlin MC (2002) Contribution of *APOE* promoter polymorphisms to Alzheimer's disease risk. Neurology 59:59–66
7. Bullido MJ, Artiga MJ, Recuero M, Sastre I, Garcia MA, Aldudo J, Lendon C, Han SW, Morris JC, Frank A, Vazquez J, Goate A, Valdivieso F (1998) A polymorphism in the regulatory region of *APOE* associated with risk for Alzheimer's dementia. Nat Genet 18:69–71
8. Casadei VM, Ferri C, Veglia F, Gavazzi A, Salani G, Cattaneo M, Sorbi S, Annoni G, Licastro F, Mariani C, Franceschi M, Grimaldi LM (1999) APOE-491 promoter polymorphism is a risk factor for late-onset Alzheimer's disease. Neurology 53:1888–1889

9. Bullido MJ, Valdivieso F (2000) Apolipoprotein E gene promoter polymorphisms in Alzheimer's disease. Microsc Res Tech 50:261–267
10. Wang JC, Kwon JM, Shah P, Morris JC, Goate A (2000) Effect of *APOE* genotype and promoter polymorphism on risk of Alzheimer's disease. Neurology 55:1644–1649
11. Zurutuza L, Verpillat P, Raux G, Hannequin D, Puel M, Belliard S, Michon A, Pothin Y, Camuzat A, Penet C, Martin C, Brice A, Campion D, Clerget-Darpoux F, Frebourg T (2000) *APOE* promoter polymorphisms do not confer independent risk for Alzheimer's disease in a French population. Eur J Hum Genet 8:713–716
12. Heijmans BT, Slagboom PE, Gussekloo J, Droog S, Lagaay AM, Kluft C, Knook DL, Westendorp RG (2002) Association of *APOE* epsilon2/epsilon3/epsilon4 and promoter gene variants with dementia but not cardiovascular mortality in old age. Am J Med Genet 107:201–208
13. Halimi G, Duplan L, Bideau C, Iniesta D, Berthezene P, Oddoze C, Verdier JM, Michel B, Berge-Lefranc JL (2000) Association of *APOE* promoter but not A2 M polymorphisms with risk of developing Alzheimer's disease. Neuroreport 11:3599–3601
14. Alvarez-Arcaya A, Combarros O, Llorca J, Sanchez-Guerra M, Berciano J, Fernandez-Luna JL (2001) The –491 TT apolipoprotein E promoter polymorphism is associated with reduced risk for sporadic Alzheimer's disease. Neurosci Lett 304:204–208
15. Zill P, Engel R, Hampel H, Behrens S, Burger K, Padberg F, Stubner S, Moller HJ, Ackenheil M, Bondy B (2001) Polymorphisms in the apolipoprotein E (*APOE*) gene in gerontopsychiatric patients. Eur Arch Psychiatry Clin Neurosci 251:24–28
16. Beyer K, Lao JI, Gomez M, Riutort N, Latorre P, Mate JL, Ariza A (2002) Identification of a protective allele against Alzheimer disease in the *APOE* gene promoter. Neuroreport 13:1403–1405
17. Roks G, Cruts M, Houwing-Duistermaat JJ, Dermaut B, Serneels S, Havekes LM, Hofman A, Breteler MM, Van Broeckhoven C, Van Duijn CM (2002) Effect of the APOE-491A/T promoter polymorphism on apolipoprotein E levels and risk of Alzheimer disease: The Rotterdam Study. Am J Med Genet 114:570–573
18. McKhann G, Drachman D, Folstein M (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34:939–944
19. Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12:656–664
20. Elnitski L, Hardison RC, Li J, Yang S, Kolbe D, Eswara P, O'Connor MJ, Schwartz S, Miller W, Chiaromonte F (2003) Distinguishing regulatory DNA from neutral sites. Genome Res 13:64–72
21. Oliveira SA, Scott WK, Nance MA, Watts RL, Hubble JP, Koller WC, Lyons KE, Pahwa R, Stern MB, Hiner BC, Jankovic J, Ondo WG, Allen FH, Jr., Scott BL, Goetz CG, Small GW, Mastaglia FL, Stajich JM, Zhang F, Booze MW, Reaves JA, Middleton LT, Haines JL, Pericak-Vance MA, Vance JM, Martin ER (2003) Association study of Parkin gene polymorphisms with idiopathic Parkinson disease. Arch Neurol 60:975–980
22. Rozen S, Skaletsky HJ. Primer 3 (1998) Whitehead Institute
23. Zaykin D, Zhivotovsky L, Weir BS (1995) Exact tests for association between alleles at arbitrary numbers of loci. Genetica 96:169–178
24. Abecasis GR, Cookson WO (2000) GOLD—graphical overview of linkage disequilibrium. BioInformatics 16:182–183
25. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 70:425–434
26. Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 67:146–154
27. Martin ER, Bass MP, Kaplan NL (2001) Correcting for a potential bias in the pedigree disequilibrium test. Am J Hum Genet 68:1065–1067
28. Martin ER, Bass MP, Gilbert JR, Pericak-Vance MA, Hauser ER (2003) Genotype-based association test for general pedigrees: the genotype-PDT. Genet Epidemiol 25:203–213
29. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM (2004) Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. Genet Epidemiol 26:61–69
30. Abecasis GR, Cardon LR, Cookson WO (2000) A general test of association for quantitative traits in nuclear families. Am J Hum Genet 66:279–292
31. Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999) Combined linkage and association sib-pair analysis for quantitative traits. Am J Hum Genet 64:259–267
32. Risch N, Teng J (1998) The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases: DNA pooling. Genome Res 8:1273–1288
33. Lambert JC, Pasquier F, Cottel D, Frigard B, Amouyel P, Chartier-Harlin MC (1998) A new polymorphism in the *APOE* promoter associated with risk of developing Alzheimer's disease. Hum Mol Genet 7:533–540
34. Town T, Paris D, Fallin D, Duara R, Barker W, Gold M, Crawford F, Mullan M (1998) The –491A/T apolipoprotein E promoter polymorphism association with Alzheimer's disease: independent risk and linkage disequilibrium with the known *APOE* polymorphism. Neurosci Lett 252:95–98