Am. J. Hum. Genet. 69:895–899, 2001

Report

The Amyloid Precursor Protein Locus and Very-Late-Onset Alzheimer Disease

Jane M. Olson, Katrina A. B. Goddard, and Doreen M. Dudek

Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University, Cleveland

Although mutations in the amyloid-β precursor protein (APP) gene are known to confer high risk of Alzheimer disease (AD) to a small percentage of families in which it has early onset, convincing evidence of a major role for the APP locus in late-onset AD has not been forthcoming. In this report, we have used a covariate-based affected-sib-pair linkage method to analyze the chromosome 21 clinical and genetic data obtained on affected sibships by the National Institute of Mental Health Alzheimer Disease Genetics Initiative. The baseline model (without covariates) gave a LOD score of 0.02, which increases to 1.43 when covariates representing the additive effects of E2 and E4 are added. Larger increases in LOD scores were found when age at last examination/death (LOD score 5.54; $P = .000002$) or age at onset plus disease duration (LOD score 5.63; $P = .000006$) were included in the linkage model. We conclude that the APP locus may predispose to AD in the very elderly.

Although mutations in the amyloid-β recursor protein (APP [MIM 104760]) gene are known to confer high risk of Alzheimer disease (AD [MIM 104300]) to a small percentage of families in which it has early onset (Goate et al. 1991), convincing evidence of a major role for the APP locus in late-onset AD has not been forthcoming. Recently, using affected sib pairs (ASPs) collected as part of the National Institute of Mental Health Alzheimer Disease (NIMH AD) Genetics Initiative (see the Alzheimer Disease Genetics Initiative Data Archive web site), Kehoe et al. (1999) and Wavrant-De Vrieze et al. (1999) reported evidence that ASPs with no E4 allele at apolipoprotein E (ApoE [MIM 107741]) were more likely to show linkage to the APP region than were ASPs in whom both members had at least one E4 allele. However, the effect that these authors detected was not large. An independent genome scan of late-onset AD failed to detect linkage to chromosome 21 (Pericak-Vance et al. 1997). To date, there has been little or no investigation of possible associations between specific APP polymorphisms and late-onset AD.

Received June 19, 2001; accepted for publication July 24, 2001; electronically published August 7, 2001.

Address for correspondence and reprints: Dr. Jane M. Olson, Department of Epidemiology and Biostatistics, 2500 MetroHealth Drive, R258, Cleveland, OH 44109-1998. E-mail: olson@darwin.cwru.edu

0002-9297/2001/6904-0024$02.00

In this report, we revisit the data collected by the NIMH AD Genetics Initiative (see the Alzheimer Disease Genetics Initiative Data Archive web site) and analyze the chromosome 21 data by using a covariate-based method of ASP linkage analysis previously proposed by Goddard et al. (2001). Specifically, we examine the roles of age at onset (AAO), age at last examination or death, degree of dementia, sex, and ApoE genotype in linkage to chromosome 21. By including covariates in a linkage analysis, one allows for locus heterogeneity due to those covariates and thereby can discover linkage evidence that might otherwise be obscured. In the present report, we show that reanalysis of the NIMH AD data by a covariate-based linkage model provides convincing evidence that the APP locus plays a role in late-onset AD.

The collection, characterization, and genotyping of AD-affected sibships by the NIMH AD Genetics Initiative (see the Alzheimer Disease Genetics Initiative Data Archive web site) have been described elsewhere (Blacker et al. 1997). We restricted our analysis to 252 ASPs who had onset at age ≥ 60 years and on whom there were complete data for the following variables: AAO (considered to be the age at which the first symptoms were present), current age (considered to be the age at either the most recent examination [33%] or death [67%]), sex, ApoE genotype, and degree of clinical dementia at the time of examination (Hughes et al. 1982). The lower bound on AAO was imposed to eliminate the possible

Table 1

Descriptive Statistics of Variables and Covariates

Variable or Covariate	Value		Within-Pair Correlation
	Mean (SD)	Range	
Original variables (n = 461 sibs):			
AAO	74.77 (5.84)	61–93	
Current age	83.33 (6.72)	63–105	
Duration	8.56 (4.65)	1–28	
Dementia	2.89 (1.12)	1–5	
Sex (proportion of females)	.75		
ApoE-allele frequencies:			
E2	.024		
E3	.581		
E4	.395		
Covariates (n = 252 ASPs):[a]			
AAO	149.02 (9.97)	127–181	.37
Current age	166.60 (11.93)	134–208	.55
Duration	17.58 (7.23)	4–44	.18
Dementia	5.78 (1.70)	2–10	.10
Sex (no. of females)	1.45 (.67)	0–2	.12
E4 (no. of E4 alleles)	1.63 (1.12)	0–4	.58
E2 (no. of E2 alleles)	.08 (.34)	0–2	.25
Current age, by E4:			
No E4 alleles (n = 51)	175.12 (11.78)	154–208	
One E4 allele (n = 52)	167.50 (11.80)	143–192	
Two E4 alleles (n = 106)	165.42 (10.03)	141–190	
Three E4 alleles (n = 26)	159.19 (10.92)	134–179	
Four E4 alleles (n = 17)	156.88 (8.70)	144–169	

[a] Data are for the sum of the pairs' values before centering; covariates were centered before inclusion in linkage models.

effect that outliers might have on the linkage analysis; only four ASPs were removed for this reason. For AAO, current age, disease duration (current age minus AAO), and degree of dementia, the covariate included in the linkage analysis was the sum of the values for the two individuals. For sex, we counted the number of females in the ASP. For ApoE genotype, we considered two covariates: (1) the total number of E4 alleles in the ASP and (2) the total number of E2 alleles in the ASP.

To allow for covariate-related locus heterogeneity, we applied a covariate-based ASP LOD-score method (Goddard et al. 2001) to the genotype data for chromosome 21. The model is a one-parameter modification of the conditional logistic parameterization of the ASP LOD score, a parameterization introduced by Olson (1999). An optimal mode-of-inheritance parameter (Whittemore and Tu 1998) is specified that allows one to fit only a single additional parameter per covariate. In terms of offspring recurrence-risk ratio, conditional on K covariates, x_k, the model is parameterized as $\lambda_1(x) = \exp(\beta + \Sigma_{k=1}^{K}\gamma_k x_k)$; in terms of the recurrence-risk ratio for MZ twins, it is parameterized as $\lambda_2(x) = 3.634\lambda_1(x) - 2.634$.

This model has been implemented in the S.A.G.E. program LODPAL (see The Human Genetic Analysis Resource web site). To simplify specification of constraints on parameter estimates, all covariates are centered around their sample mean before being included within the linkage model; asymptotic distributions of the resulting likelihood-ratio tests were used to obtain P values (see Goddard et al 2001). Covariate-specific sibling recurrence-risk ratios can be obtained by the expression $\lambda_S(x) = \frac{1}{4} + \frac{1}{2}\lambda_1(x) + \frac{1}{4}\lambda_2(x)$.

In this report, we describe as "LOD" scores the likelihood-ratio statistics (LRSs) divided by 4.605 (i.e., $2\log_e 10$). Critical values for the LRSs were obtained as follows. The distribution of the LRS for the basic one-parameter model is a 50:50 mixture of a point mass at 0 and a χ^2 distribution with 1 df. Addition of K covariates gives an LRS with a distribution that is a 50:50 mixture of a χ^2 with K df and a χ^2 with $K + 1$ df. The LRS difference between nested models that differ by J covariates has a χ^2 distribution with J df. One can therefore test both the significance of the contribution of a covariate and the overall evidence for linkage.

Multipoint identity-by-descent values were computed at 2-cM intervals throughout chromosome 21, with use of all markers available from the genome scan. In this report, we describe results at the approximate location of the APP locus (2 cM proximal to D21S1435). It is necessary, when testing the significance of covariates, to compare models at the same chromosomal location; our

peak LOD scores all occurred either at or close to (i.e., ⩽4 cM from) this location.

Table 1 contains descriptive statistics of the variables used to construct covariates for the linkage analysis. In the model, the covariates were included as the sum of the values for the two individuals constituting each ASP; table 1 also includes the means and SDs for these covariates, as well as those for the original variables.

Table 2 contains the LOD scores at the APP locus. The baseline LOD score, with no covariates, has a value of only 0.02, showing no evidence of linkage. The LOD score increases to 1.13 when E4 is added. These results are consistent with what Wavrant-DeVrieze et al. (1999) found by using these data; ASPs without E4 alleles show the most evidence of linkage to the APP locus. The addition of E2 to the model is nonsignificant (i.e., a 0.30 increase in LOD score), possibly, at least in part, because E2 is relatively rare both in the population in general and in patients with AD (see table 1). However, it is instructive to consider the fitted sibling recurrence-risk ratios for this model, which show that the risk is lowest in ASPs with four E4 alleles ($\lambda_S = 0.71$), moderate in ASPs with four E3 alleles ($\lambda_S = 1.32$), and highest in ASPs with two E2 and two E3 alleles ($\lambda_S = 3.26$). No ASP in this data set had more than two E2 alleles.

The most significant covariates were those that measured age. Inclusion of AAO increases the LOD score to 2.27, and (separate) inclusion of current age increases the LOD score to 5.54; the effects of both covariates were highly significant. Current duration of disease (current age minus AAO) shows little effect when included separately but greatly increases the LOD score (to 5.63) when included along with AAO. The signs of the covariate parameters indicate that "linked" ASPs tend to be those with the latest AAO and longest disease duration, or, almost equivalently, the oldest current age.

Table 3

Recurrence-Risk Ratios, by Current Age

Current Age[a] ($x/2$)	No. of SD Units[b]	$\lambda_S(x)$
71.37	−2	.33
77.33	−1	.69
83.30	0	1.23
89.26	1	2.05
95.23	2	3.26

[a] Average, for an ASP.
[b] Above or below sample mean.

The model with current age alone is the most parsimonious and may be most useful for identification and/or ascertainment of ASPs most likely to be useful in further studies of the role of APP in late-onset AD, particularly in searches for predisposing polymorphisms. In table 3 we provide sibling recurrence-risk ratios for some values of current age; these results suggest that future searches for relevant APP polymorphisms should focus on affected individuals in their late 80s and older. In table 1 we also provide mean total current ages of ASPs, by number of E4 alleles. Current age decreases with increasing number of E4 alleles; individual patients in ASPs with no E4 alleles are, on average, ~10 years older than those in ASPs with four E4 alleles.

To ensure that our result for the APP locus was not the result of strong influence from a small number of sib pairs, we obtained the ASP-specific LOD-score contributions (a new feature of LODPAL) for the model with AAO and disease duration. Sixty-eight ASPs had LOD-score contributions >0.1, and 104 had LOD-score contributions >0.1 in absolute value; the range was −0.76 to 0.55. We remaximized the LOD-score function after eliminating the five ASPs whose contribution to the

Table 2

LOD Scores, Parameter Estimates, and Significance Levels at the APP Locus

Model	LOD Score	β	γ_1	γ_2	γ_3	γ_4	Overall P	LOD Score Difference[a]	P[a]
Baseline	.05	.0284					.3923	…	…
Plus ApoE4 (E4)	1.13	.0402	−.1131				.0483	1.08	.0257
Plus ApoE2 (E2)	.36	.0383	.3314				.3172	.31	.2322
Plus E4 plus E2	1.43	.0539	−.1089	.3778			.0618	.30	.2398
Plus sex	.19	.0303	−.0710				.4976	.14	.4220
Plus dementia	.24	.0317	.0349				.4343	.19	.3496
Plus AAO	2.27	.0580	.0192				.0032	2.22	.0014
Plus current age	5.54	.1540	.0337				.000002	5.49	.0000005
Plus duration	.64	.0370	.0136				.1576	.59	.0993
Plus AAO plus duration	5.63	.1523[b]	.0324[c]	.0361[d]			.000006	3.36	.000084
Plus AAO plus duration plus E4 plus E2	5.70	.1569	.0334	.0364	.0198	.0909	.000054	.07	.5702

[a] Between this model and nearest nested model.
[b] Standard error .0877 (computed under the assumption that ASPs are independent).
[c] Standard error .0094 (computed under the assumption that ASPs are independent).
[d] Standard error .0010 (computed under the assumption that ASPs are independent).

LOD score was >0.4 in absolute value. The resulting LOD score, 5.02, although reduced in magnitude, remains highly significant.

We have shown that the APP locus on chromosome 21 is strongly linked to AD if AAO plus disease duration or current age alone are included as covariates in the linkage model. The most parsimonious model includes only current age; the longest-lived affected individuals are most likely to have linkage to the APP region. In addition, the model with AAO plus disease duration suggests that ASPs susceptible to the detrimental effects of mutations in the APP region are those with very late onset who survive longer with the disease. It may be that this form of AD is characterized by both very late onset and slower disease progression, but, because the three variables are linearly related, we cannot statistically determine which of the two models best represents the primary features. Our results are also consistent with (*a*) Kehoe et al.'s (1999) and Wavrant-DeVrieze et al.'s (1999) findings that these ASPs have fewer ApoE4 alleles than do the rest of the sample and (*b*) Blacker et al.'s (1997) results, based on the same data, suggesting that ApoE4 exerts its maximal effect at age <70 years.

The absence of E4 alleles in these individuals may not represent an interaction at the biological level. Because these mutations in the APP region do not confer susceptibility until very late in life, most individuals with E4 alleles may have either died or at least fallen ill with E4-type AD prior to the age when these mutations become penetrant. In addition, these individuals appear to survive longer with the disease, suggesting a slower disease progression.

These results also demonstrate the usefulness of covariate-based linkage methods for both detection of linkage in the presence of locus heterogeneity and identification of which subgroups of families have linkage to which loci. In this report, we not only have established strong evidence of linkage and age-related locus heterogeneity but also have provided useful recommendations for sample selection or subselection in future research. Specifically, studies of the role that APP-region mutations play in late-onset AD should concentrate on the extremely elderly, particularly those with later AAO who have survived longer than average and, perhaps incidentally, have few E4 alleles. As have Goddard et al. (2001), we note that the ability to include covariates increases the overall probability of type I errors due to multiple analyses of the same genetic data. In this report, we have used genomewide significance levels associated with a single analysis but have focused our attention exclusively on chromosome 21, with the exception of the E4 variable, as noted below. In general, we recommend both careful prior selection of covariates and genomic regions for study and cautious interpretation of results.

It may be noted that, because disease duration adds considerably to the LOD score at the APP locus, even after AAO is taken into account, samples ascertained through new AD cases might not be as useful for studying the APP effect as are samples ascertained through a cohort of existing cases, even if only families with very late onset are collected. In cohort sampling, the effects of length-biased sampling would likely lead to oversampling of patients with APP-type symptoms, with respect to their incidence rates, but that would be desirable if one wished to "load" one's sample for families most likely to carry mutations in the APP region. It also should be noted that, in these data, disease duration is measured by subtracting AAO from age at last examination, if age at death is not available, and that it therefore represents disease duration until the time at most recent observation. On the other hand, age at death was available for approximately two-thirds of these patients. It is difficult to know precisely either how our quantitative measurement of disease duration is affected by the ascertainment scheme or how these results might change if age at death of all of these affected individuals were available.

We rescanned the entire genome, using a linkage model that includes our E4 covariate (results not shown). Other than the APP region, no genomic location—including the regions on chromosomes 6, 9, 10, and 19, previously reported on the basis of these data (Kehoe et al. 1999; Bertram et al. 2000; Collins et al. 2000; Myers et al. 2000), and the region, on chromosome 12, previously reported, by Pericak-Vance et al. (1997), for a different data set—provided even modest evidence that ASPs with fewer E4 alleles were more strongly linked than were families with more E4 alleles. On the other hand, regions on chromosomes 1p21, 2, 4, 5, 9, 14pter, 16, and 19 each showed some evidence (i.e., LOD score, for covariate effect, .74–1.94) that ASPs with more E4 alleles were more likely to be linked than were ASPs with fewer E4 alleles. These results suggest that E4 alleles may play a role in many, but not all, biological mechanisms that lead to AD.

Acknowledgments

This work was supported in part by U.S. Public Health Service grants HG01577, from the National Center for Human Genome Research, and RR03655, from the National Center for Research Resources. Some of the results in this report were obtained by S.A.G.E., which is supported by National Center for Research Resources grant RR03655. Data were collected in three projects that participated in the NIMH AD Genetic Initiative. During 1991–98, the principal investigators and coinvestigators were Marilyn S. Albert and Deborah Blacker, both of Massachusetts General Hospital, Boston (both supported by grant U01 MH46281); Susan S. Bassett, Gary A. Chase, and Marshal F. Folstein, all of Johns Hopkins University, Baltimore (all supported by grant U01 MH46290); and Rodney C. P. Go and

Lindy E. Harrall, both of University of Alabama, Birmingham (both supported by grant U01 MH46373).

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Alzheimer Disease Genetics Initiative Data Archive, http://zork.wustl.edu/nimh/ad.html (for NIMH AD Genetics Initiative)
Human Genetic Analysis Resource, The, http://darwin.cwru.edu/ (for S.A.G.E. software)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for AD [MIM 104300], APP [MIM 104760], and ApoE [MIM 107741])

References

Bertram L, Blacker D, Mullin K, Keeney D, Jones J, Basu S, Yhu S, McInnis MG, Go RCP, Vekrellis K, Selkoe DJ, Saunders AJ, Tanzi RE (2000) Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. Science 290:2302–2303

Blacker D, Haines JL, Rodes L, Terwedow H, Go RC, Harrell LE, Perry RT, Bassett SS, Chase G, Meyers D, Albert MS, Tanzi R (1997) ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. Neurology 48:139–147

Collins JS, Perry RT, Watson B Jr, Harrell LE, Acton RT, Blacker D, Albert MS, Tanzi RE, Bassett SS, McInnis MG, Campbell RD, Go RC (2000) Association of a haplotype for tumor necrosis factor in siblings with late-onset Alzheimer disease: the NIMH Alzheimer Disease Genetics Initiative. Am J Med Genet 96:823–830

Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 349: 704–706

Goddard KAB, Witte JS, Suarez BK, Catalona WJ, Olson JM (2001) Model-free linkage analysis with covariates confirms linkage of prostate cancer to chromosomes 1 and 4. Am J Hum Genet 68:1197–1206

Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL (1982) A new clinical scale for the staging of dementia. Br J Psychiatry 140:566–572

Kehoe P, Wavrant-De Vrieze F, Crook R, Wu WS, Holmans P, Fenton I, Spurlock G, Norton N, Williams H, Williams N, Lovestone S, Perez-Tur J, Hutton M, Chartier-Harlin MC, Shears S, Roehl K, Booth J, Van Voorst W, Ramic D, Williams J, Goate A, Hardy J, Owen MJ (1999) A full genome scan for late onset Alzheimer's disease. Hum Mol Genet 8: 237–245

Myers A, Holmans P, Marshall H, Kwon J, Meyer D, Ramic D, Shears S, Booth J, Wavrant-DeVrieze F, Crook R, Hamshere M, Abraham R, Tunstall N, Rice F, Carty S, Lillystone S, Kehoe P, Rudrasingham V, Jones L, Lovestone S, Perez-Tur J, Williams J, Owen MJ, Hardy J, Goate AJ (2000) Susceptibility locus for Alzheimer's disease on chromosome 10. Science 290:2304–2305

Olson JM (1999) A general conditional-logistic model for affected-relative-pair linkage studies. Am J Hum Genet 65: 1760–1769

Pericak-Vance MA, Bass MP, Yamaoka LH, Gaskell PC, Scott WK, Terwedow HA, Menold MM, Conneally PM, Small GW, Vance JM, Saunders AM, Roses AD, Haines JL (1997) Complete genome screen in late-onset familial Alzheimer disease. JAMA 278:1237–1241

Wavrant-De Vrieze, Crook R, Holmans P, Kehoe P, Owen MJ, Williams J, Roehl K, Laliiri DK, Shears S, Booth J, Wu W, Goate A, Chartier-Harlin MC, Hardy J, Perez-Tur J (1999) Genetic variability at the amyloid-β precursor protein locus may contribute to the risk of late-onset Alzheimer's disease. Neurosci Lett 269:67–70

Whittemore AS, Tu I-P (1998) Simple, robust linkage tests for affected sibs. Am J Hum Genet 62:1228–1242