Genetic Studies on Chromosome 12 in Late-Onset Alzheimer Disease

William S. Wu, MS; Peter Holmans, PhD; Fabienne Wavrant-DeVrièze; Shantia Shears; Patrick Kehoe; Richard Crook; Jeremy Booth; Nigel Williams; Jordi Pérez-Tur, PhD; Kimberly Roehl; Iain Fenton; Marie-Christine Chartier-Harlin, PhD; Simon Lovestone, MD; Julie Williams, PhD; Mike Hutton, PhD; John Hardy, PhD; Michael J. Owen, MD; Alison Goate, PhD

Context.—The only genetic locus universally accepted to be important as a risk factor for late-onset Alzheimer disease (AD) is the apolipoprotein E (*APOE*) locus on chromosome 19. However, this locus does not account for all the risk in late-onset disease, and a recent report has suggested a second locus on chromosome 12p11-12.

Objective.—To look for evidence of linkage on chromosome 12 and to test for the presence of the new locus in an independent sample of familial late-onset AD cases.

Design.—Retrospective cohort study. As part of a 20-centimorgan genome screen (approximately equal to 200 markers), we tested a series of 18 genetic markers on chromosome 12 and carried out multipoint, nonparametric lod score and exclusion analyses.

Setting.—Clinic populations in the continental United States selected from the National Institute of Mental Health AD Genetics Consortium.

Patients.—We selected samples for DNA analysis from affected sibling pairs, 497 subjects from 230 families with 2 or more affected individuals with probable or definite AD with onset ages older than 60 years (mean±SD, 75 ± 6 years). Within the families, we used the 2 probable or definitely affected individuals. In families with more than 2 such cases available, we used all of them; in families with only 2 such cases in which unaffected individuals were available, we also sampled the oldest unaffected individual and used genotype data from this unaffected individual to check for nonpaternity and genotyping errors.

Main Outcome Measure.—Presence of linkage or locus on chromosome 12.

Results.—Although linkage analyses confirmed the presence of a genetic susceptibility factor at the *APOE* locus in these families with late-onset AD, we were unable to confirm the presence of a locus close to the marker *D12S1042*. A moderate lod score (1.91) was found near *D12S98* close to the α_2-macroglobulin locus in the affected pairs in which both members did not possess an *APOE* ϵ4 allele.

Conclusions.—*APOE* remains the only locus established to be a risk factor for late-onset AD. We were unable to confirm that a locus on chromosome 12p11-12 has a major effect on risk for late-onset AD, although an effect smaller than that for *APOE* cannot be excluded.

JAMA. 1998;280:619-622

From the Departments of Psychiatry and Genetics, Washington University School of Medicine, St Louis, Mo (Messrs Wu and Booth, Mss Shears and Roehl, and Dr Goate); Neuropsychiatric Genetics Unit, University of Wales College of Medicine, Cardiff (Drs Holmans, Owen, and Williams and Messrs Kehoe, Williams, and Fenton); Mayo Clinic Jacksonville, Jacksonville, Fla (Ms Wavrant-DeVrièze, Mr Crook, and Drs Pérez-Tur, Hutton, and Hardy); INSERM, Institut Pasteur de Lille, Lille, France (Ms Wavrant-DeVrièze and Dr Chartier-Harlin); and Section of Old Age Psychiatry, Institute of Psychiatry, London, England (Dr Lovestone).

Reprints: John Hardy, PhD, Neurogenetics Laboratory, Mayo Clinic Jacksonville, 4500 San Pablo Rd, Bridsall Bldg, Jacksonville, FL 32224 (e-mail: hardy.john@mayo.edu).

LATE-ONSET Alzheimer disease (AD) shows familial clustering[1] but does not show a clear mode of inheritance. The only genetic locus universally accepted as an important risk factor for late-onset AD is the apolipoprotein E (*APOE*) locus on chromosome 19.[2] A proportion of this familial clustering is accounted for by genetic variability at the *APOE* locus. The *APOE* ϵ4 allele is positively associated with disease and the *APOE2* allele is negatively associated with disease.[3-6] However, the *APOE* locus accounts for, at most, about half of the genetic risk of developing the disease.[3,7] Thus, other genes or risk factors must account for the remaining genetic risk for developing disease.

Several strategies can be used to define these other genetic risk factors: (1) the analysis of large pedigrees with late-onset disease; however, few suitable pedigrees have been ascertained[8]; (2) the genetic analysis of population isolates with disease; again, few such isolates have been reported[9]; (3) the use of association studies between alleles of candidate genes and disease; and (4) the application of genome search strategies using large numbers of sibling pairs.

Association studies have the advantage of simplicity and speed since all that is required is the availability of a case-control series of DNA samples that can be tested for the presence of particular alleles. Furthermore, association studies can detect relatively small increases in risk associated with particular alleles. However, since the identification of the *APOE* locus, a large number of positive-association studies have been reported for AD, but none of these have been consistently confirmed. These failures to confirm associations may reflect different etiologies in different populations, type I statistical errors (a particular problem because multiple testing is always involved), or linkage disequilibrium between the tested polymorphism and the functional polymorphism (with this disequilibrium not present in all populations). In addition, association studies are sensitive to population substructure and unnoticed ethnic differences between cases and controls. Furthermore, association studies currently cannot detect new genes implicated in

Table 1.—Family Structure Comparison*

Characteristic	This Series	Duke Series[10]
No. of		
Families	230	54
Affected	643 (510)	173 (141)
Patients per family	2.8 (2.3)	4.0 (2.6)
Large family subset†	3.6 (3.1)	
Pairs		
Sib	515 (292)	234 (114)
Parent-child	97 (0)	62 (4)
Avuncular	43 (4)	31 (10)
Cousin	12 (7)	22 (13)
Age at onset, mean±SD, y	75 ± 6	73 ± 7

*Numbers in parentheses are those actually sampled. The sib pairs were comprised as follows: 199 sib pairs, 27 sib trios, and 2 sib quartets.
†There were a total of 43 families.

disease etiology; they can only be used to test existing "candidate" genes.

Genome search strategies have the disadvantage that they are considerably less sensitive than association studies and require large amounts of work in terms of sample collection, data generation, and data analysis. However, these studies have the potential to identify previously unsuspected loci. Given the uncertainty of the published disease associations, we initiated a genome search for other loci. During this process, Pericak-Vance and colleagues[10] published a report of a locus on chromosome 12, and we tried to replicate this finding. A second reason for us to start the genome search on chromosome 12 is the fact that we and others have found associations between AD and the *LRP* gene (located on chromosome 12),[11-13] which encodes a brain-expressed *APOE* receptor. In addition, the ligand for this receptor, α_2-macroglobulin, is also encoded on this chromosome.

METHODS
Family Data

We selected families from those collected by the the National Institute of Mental Health AD Genetics Consortium.[7] From within this family series, we selected 230 families based on the following criteria: at least 2 affected siblings with probable or definite AD with onset ages of more than 60 years, sampled and available for genotyping. Within the families, we genotyped the probable or definitely affected siblings. For families with more than 2 affected siblings (definite or probable) available, we used all of them; in families with only 2 affected siblings but in which unaffected individuals were available, we also sampled the oldest unaffected individual, so that the genotype data from this unaffected individual could be used to check for genotyping errors (eg, data suggesting more than 4 parental chromosomes). Comparison of the structure of our family series with that used by Pericak-Vance et al[10] is shown in Table 1,

although we restricted our analysis to sibling pairs.

Generation of Marker Data

The microsatellite polymorphisms were detected by polymerase chain reaction. The forward primers were labeled at the 5′ end with a 6-carboxyfluorescein known as 6-FAM, a tetrachlorinated analogue, or a hexachlorinated analogue (Perkin Elmer Systems, San Francisco, Calif). The total 5-μL reaction contained a 50-μg DNA genomic template, 3.5-mol/L end-labeled forward primer and unlabeled reverse primer, 0.2 mol/L each of dextroadenosine triphosphate, dextroguanosine 5′-triphosphate, dextrocytidine 5′-triphosphate, dextroribothymide 5′-triphosphate, 1 unit of Taq DNA polymerase (Promega), and 1-μL 5× buffer (7.5-mmol/L magnesium chloride, 250-mmol/L potassium chloride, 50-mmol/L Tris hydrochloride [pH 8.3]). All reaction cocktails were distributed evenly to a 96-well Falcon assay plate using a Beckman-1000 workstation (Beckman Instruments, San Francisco, Calif). Polymerase chain reactions were carried out on Hybaid OmniGene thermal cyclers (Hybaid, London, England) using the following cycling conditions: initial denaturation at 96°C for 4 minutes, followed by optimized (22-30) cycles of 94°C for 1 minute, optimized annealing temperature (54°C-60°C) for 1 minute, and 72°C for 45 seconds, and a final extension of 5 minutes at 72°C. The polymerase chain reaction products were then diluted at least 8-fold using a Beckman-1000 workstation. A measurement of 0.8 μL of diluted polymerase chain reaction product was mixed with 2.5 μL of deionized formamide, 0.4 μL of internal lane standard TAMRA-350 (Perkin Elmer Systems), and 0.5 μL of blue dye, denatured at 97°C for 5 minutes, rapidly cooled on ice, and was then electrophoresed on a 6% denaturing polyacrylamide gel and the alleles were detected on a 373–automated DNA sequencer (Perkin Elmer Systems, San Francisco, Calif) using GeneScan 2.1 (Perkin Elmer Systems). Two Centre Erute Polymorphisme Humaine individuals (133101, 133102) and 1 local DNA were loaded on every gel as controls. All alleles were initially assigned and genotyped semiautomatically using Genotyper 2.0 (Perkin Elmer Systems) without any information of phenotype. Laboratory personnel were masked to the phenotypic status of individual samples.

Statistical Analysis

Two-point and multipoint-affected sibling pair analysis was performed on the entire set of genotyped sibships (292 af-

fected pairs) using MAPMAKER/SIBS.[14] This program does not permit analysis of other affected pairings. In addition, the data set was divided into 2 subsamples, each of which was analyzed separately using MAPMAKER/SIBS.[14] The first group consisted of sibling pairs who both possessed at least 1 *APOE* ε4 allele (162 affected pairs), and the second group consisted of sibling pairs in which neither member possessed an *APOE* ε4 allele (63 affected pairs). For example, a quartet of affected siblings whose *APOE* genotypes are APOE3/APOE3, APOE3/APOE3, APOE3/APOE4, and APOE3/APOE4 would contribute 1 pair to each group.

A multipoint exclusion map also was obtained for the entire sample using MAPMAKER/SIBS. For the purposes of this analysis, the disease-susceptibility model was parameterized in terms of λ_S, the relative risk to siblings of a case.[15] A number of values of λ_S ranging from 1.2 to 2.0 were tested.

RESULTS

Two-point lod scores between markers and disease are shown in Table 2. Multipoint affected sibling pair analysis on the entire data set (292 sibling pairs), with lod scores, is shown in Figure 1. The maximum lod score obtained was 0.89 at *D12S98*, corresponding to a chromosome-wide *P* value of .21. If the multiple testing arising from the *APOE* ε4 positive and negative analyses is included, the *P* value increases to .48. These *P* values were simulated from our actual sample, using the observed marker allele frequencies and the marker map used in the analyses and thus should not be conservative. (A lod score of zero, suggesting no excess of allele sharing between affected family members, was obtained in the region in which Pericak-Vance and colleagues[10] found their highest lod scores.)

We attempted to detect possible epistatic effects between *APOE* and a susceptibility locus on chromosome 12 by splitting the data set into 2 subsets. The first portion contained sibling pairs who both possessed at least 1 *APOE* ε4 allele (162 pairs) and is denoted in Figure 1 by "*APOE* ε4 both positive." The other subset contained sibling pairs with both members *APOE* ε4 negative (63 pairs), and is denoted in Figure 1 by "*APOE* ε4 both negative." As shown in Figure 1, neither group had a significant lod score close to *D12S1042*. However, the subset of affected pairs who were both *APOE* ε4 negative had a multipoint lod of 1.91 at *D12S98*, corresponding to a chromosome-wide *P* value of approximately .09 (allowing for multiple testing). The α_2-macroglobulin gene is between *D12S98* and *D12S397* at this peak, and Blacker

Chromosome 12 in Late-Onset Alzheimer Disease—Wu et al

Table 2.—Marker Information and 2-Point Lod Scores*

Locus	Map Position	Heterozygosity	No. of Families	No. of Sib Pairs	Identity by Descent Proportion	Lod Score (Data Set)	Lod Score (Large Families)	APOE ε4 Positive	APOE ε4 Negative
S372	0	0.73	178	238	0.48	0	0.09	0	0.25
S397	17	0.87	222	286	0.54	0.59	0.84	0.19	0.69
S98	24	0.65	221	283	0.56	0.92	1.17	0.10	1.12
S358	26	0.77	218	278	0.59	1.67	1.26	0.63	1.04
S391	30	0.87	223	287	0.53	0.27	0.03	0	0.50
S310	36	0.55	221	278	0.48	0	0	0	0.61
S363	37	0.62	223	287	0.50	0	0	0	0.32
S373	40	0.69	221	285	0.48	0	0	0	0.01
S1027	45	0.70	222	281	0.51	0.01	0	0	0.58
S1057	52	0.81	223	285	0.51	0.04	0.65	0.71	0
S1042	56	0.78	221	285	0.52	0.13	0.73	0.98	0
S1292	59	0.73	225	289	0.43	0	0	0	0
S398	86	0.64	223	287	0.48	0	0.03	0.3	0
S375	98	0.74	209	273	0.51	0.01	0.23	0.24	0
S379	116	0.72	197	257	0.47	0	0	0	0
S393	131	0.66	189	251	0.54	0.28	0	0.24	0.09
S395	155	0.75	211	273	0.52	0.05	0.58	0	0.26
S1045	183	0.84	204	266	0.51	0.01	0.07	0.06	0

*Map positions obtained from the genome database. Marker heterozygosity estimated from our sample. Lod scores calculated as described in the text. Numbers of affected sib pairs are less than 292 because of data loss. Identity by descent proportion is calculated for the entire sample. Allele sharing of 0.5 would be expected by chance. Positive deviations from the chance give rise to positive lod scores. Markers S373, S1027, S1057, S1042, and S1292 are those to which Pericak-Vance and colleagues[10] reported genetic linkage. The LRP gene is located close to D12S398.

and colleagues[7] have recently reported an association between this gene and AD.[16] These data, thus, would support the notion that α_2-macroglobulin is an AD locus.

The results of the multipoint exclusion analysis performed on the entire data set are shown in Figure 2. For this analysis, the disease-susceptibility model was parameterized in terms of the relative risk to siblings of a case, λ-$_S$.[15] Assuming the frequency of APOE ε4 in the general population is 0.15, the relative risk to ε4 heterozygotes is 4, and the relative risk to ε4 homozygotes is 10^3, then λ_S for the APOE locus is approximately 1.4. As shown in Figure 2, when λ_S is 1.4, a region extending from D12S310 to 12 centimorgans distal of D12S1292 is excluded (lod score <−2.0). (This region includes D12S1042, the locus to which Pericak-Vance and colleagues[10] found their strongest evidence of linkage.)

To replicate the analysis of Pericak-Vance and colleagues,[10] we selected from our sample those families with at least 3 sampled affected individuals (see Table 1 for the structure of these "large pedigrees") and analyzed these separately. The lod scores given by each locus analyzed separately are shown in Table 2. The highest multipoint lod scores were 0.47 at D12S98 and 0.44 at D12S395. This is consistent with the results obtained from the entire sample, although the lod score is reduced, as would be expected from the reduction in sample size. In this analysis, multiply affected families are not greatly different from the others in terms of their evidence for linkage. However, our series of large families (at least 3 sampled individuals) contained half the number of sampled cousin and avuncu-
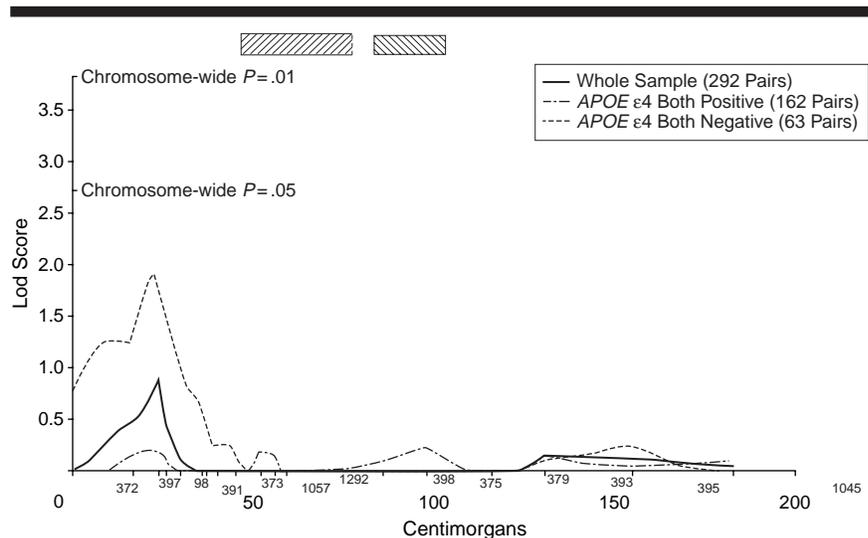

Figure 1.—Multipoint lod scores along chromosome 12 calculated using MAPMAKER/SIBS,[14] together with criteria corresponding to chromosome-wide .05 and .01 significance (based on our sample and marker map, allowing for the multiple analyses performed). The left-hand bar shows the chromosomal region where Pericak-Vance and colleagues[10] reported a lod score of more than 2 in their data set. This bar is open at the right-hand end because the region is not delimited in their publication. The right-hand bar shows the approximate chromosomal region where the LRP gene has been localized. The α_2-macroglobulin gene maps between D12S397 and D12S98.

lar pairs than the data sets in the previous report. Thus, it remains possible that some difference in ascertainment between the 2 series could account for the discrepancies in the results.

COMMENT

Our findings should not be interpreted as indicating that there is no AD-risk gene close to D12S1042, since, as with the failures to repeat the observations of genetic associations, it remains possible that there is such a locus that is important in a few families including a propor-

tion of those ascertained by Pericak-Vance and colleagues.[10] However, it is unlikely that such a locus accounts for a large proportion of cases of AD, since we were not able to detect evidence of linkage in our entire data set. However, we were able to detect evidence of linkage to a marker near the APOE locus using a dinucleotide repeat marker (D19S412). This marker gave a lod score of 1.1 on its own and 2.1 when used in a multipoint with APOE. These data, not surprisingly, resemble those reported by Blacker et al,[7] since they consist of a largely over-
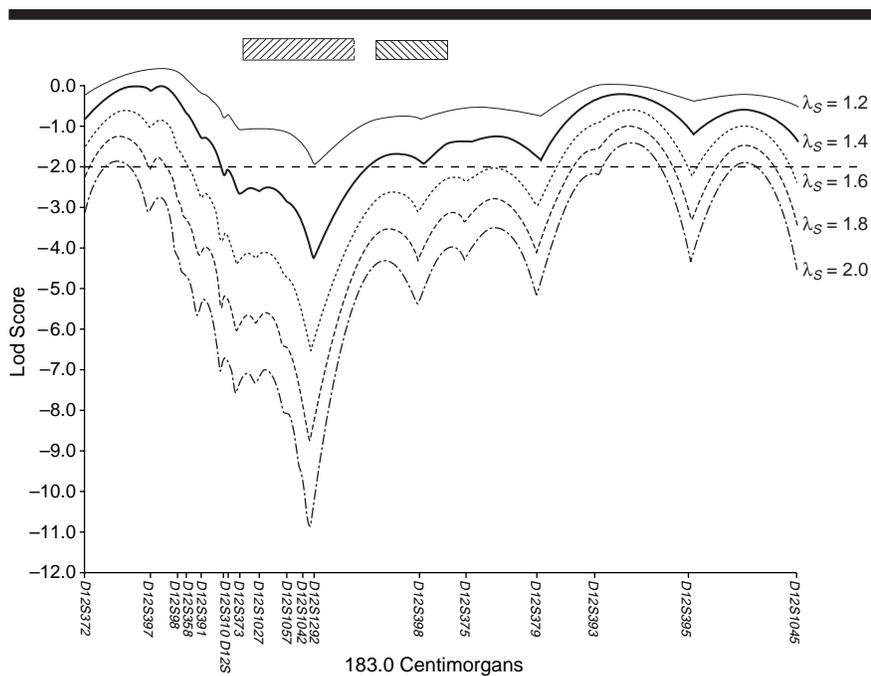
Figure 2.—Exclusion map along chromosome 12 generated using MAPMAKER/SIBS[14] at different values of λ_S.[15] The left-hand bar shows the chromosomal region where Pericak-Vance and colleagues[10] reported a lod score of more than 2 in their data set. This bar is open at the right-hand end because the region is not delimited in their publication. The right-hand bar shows the approximate chromosomal region where the *LRP* gene has been localized. The α_2-macroglobulin gene maps between *D12S397* and *D12S98*.

lapping series of samples. Furthermore, the sample we have used here was sufficiently powerful to exclude (lod < −2) a locus of similar effect size to that of *APOE* from the region in which Pericak-Vance and colleagues[10] found their most significant results. It is therefore unlikely that the difference between our results and those of Pericak-Vance and colleagues[10] is due to lack of power in our sample. We did obtain a moderate lod score ($P = .09$) in the subset of affected pairs where both members were *APOE* ε4 negative, suggesting the possible in-

volvement of a locus acting heterogeneously with *APOE* ε4. This locus is exactly at the α_2-macroglobulin locus, which Blacker and colleagues[7] have recently suggested is associated with AD.[16] This locus is more than 20 centimorgans from the region implicated by Pericak-Vance and colleagues.[10]

The *LRP* gene is close to *D12S398*, but we are not able to confirm or refute the inference from association studies,[11-15] which implicate this gene in the etiology of the disease. Our data suggest that *LRP* is not a gene of large effect. How-

ever, given the poor genetic resolution of these types of linkage studies, it remains possible that the *LRP* gene is within the "linked" region in the data set. If the α_2-macroglobulin is a predisposing locus for AD,[16] it will be important to examine the genetics of both *LRP* and its other ligands.

A serious problem in genome scans of complex disorders is the assessment of the true level of significance associated with a screen in which multiple tests are performed on a single data set. The problem is that such studies attempt to resolve 2 complex issues with a single analysis: (1) determination of the inheritance characteristics of the locus (ie, mode of inheritance, age-dependent penetrance, "phenocopy" or misdiagnosis rate, disease allele frequency, interaction with other [*APOE*] loci, etc) and (2) identification of the location of this putative gene. This is quite unlike maximum likelihood methods of analysis in simple Mendelian disorders in which the inheritance characteristics can be estimated to a high enough degree of certainty so as not to interfere with the localization. Modification of the analysis parameters or choosing a variety of linkage analyses and approaches amounts to multiple testing from a statistical perspective. Although multiple testing is a legitimate part of trying to determine the inheritance characteristics of the phenotype, it means that lod scores derived from such analyses cannot be thought of as equivalent to the lod scores in simple disorders.

References

1. Pericak-Vance MA, Haines JL. Genetic susceptibility to Alzheimer's disease. *Trends Genet*. 1995; 11:504-508.
2. Strittmatter WJ, Saunders AM, Schmechel D, et al. Apolipoprotein E: high affinity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer's disease. *Proc Natl Acad Sci U S A*. 1992;90:1971-1981.
3. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993;261:921-923.
4. Corder EH, Saunders AM, Risch N, et al. Protective effect of apolipoprotein E type 2 alleles decreases the risk of late onset Alzheimer disease. *Nat Genet*. 1994;7:180-184.
5. Chartier-Harlin M-C, Parfitt M, Legrain S, et al. Apolipoprotein E, ε4 allele as a major risk factor for sporadic early- and late-onset forms of Alzheimer's disease: analysis of the 19q13.2 chromo-

somal region. *Hum Mol Genet*. 1994;3:569-574.
6. Talbot C, Lendon CL, Craddock N, Shears S, Morris JC, Goate AM. Protection against Alzheimer's disease with ApoE ε2. *Lancet*. 1994;343:1432-1433.
7. Blacker D, Haines JL, Rodes L, et al. ApoE-4 and age at onset of Alzheimer's disease: the NIMH Genetics Initiative. *Neurology*. 1997;48:139-147.
8. Martin ES, Martin SE, Edelsohn L, Borgaonkar DS. Studies in a large family with late onset Alzheimer's disease. *Alzheimer Dis Assoc Disord*. 1997; 11:163-170.
9. Pericak-Vance MA, Haines J, Johnson CC, et al. Alzheimer's disease and apolipoprotein E in an Amish population. *Ann Neurol*. 1997;39:700-704.
10. Pericak-Vance MA, Bass MP, Yamaoka LH, et al. Complete genomic screen in late-onset familial Alzheimer's disease: evidence for a new locus on chromosome 12. *JAMA*. 1997;278:1237-1241.
11. Lendon CL, Talbot CJ, Craddock NJ, et al. Ge-

netic association studies between dementia of the Alzheimer type and three receptors for apolipoprotein E in a Caucasian population. *Neurosci Lett*. 1997;222:187-190.
12. Wavrant-DeVrièze F, Pérez-Tur J, Lambert J-C, et al. Association between the low density lipoprotein receptor-related protein (LRP) and Alzheimer's disease. *Neurosci Lett*. 1997;227:68-70.
13. Kang DE, Saitoh T, Chen X, et al. Genetic association of the low density lipoprotein receptor-related protein gene (LRP), an apolipoprotein receptor, with late onset Alzheimer's disease. *Neurology*. 1997;49:56-61.
14. Kruglyak L, Lander E. Complete multipoint analysis of qualitative and quantitative traits. *Am J Hum Genet*. 1995;57:439-454.
15. Risch N. Linkage strategies for genetically complex traits, I. *Am J Hum Genet*. 1990;46:222-228.
16. Kruglyak L. Thresholds and sample sizes. *Nat Genet*. 1996;14:132-133.