HUMAN MUTATION 28(11), 1065–1073, 2007

RESEARCH ARTICLE

Effect of Heterogeneity on the Chromosome 10 Risk in Late-Onset Alzheimer Disease

Xueying Liang,[1,2] Eden R. Martin,[3,4] Nathalie Schnetz-Boutaud,[1,2] Jackie Bartlett,[1,2] Brent Anderson,[1,2] Stephan Züchner,[3–5] Harry Gwirtsman,[6] Don Schmechel,[7] Regina Carney,[3,4] John R. Gilbert,[3,4] Margaret A. Pericak-Vance,[3,4] and Jonathan L. Haines[1,2]*

[1]*Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee;* [2]*Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee;* [3]*Center for Human Genetics, Duke University, Durham, North Carolina;* [4]*Department of Medicine, Duke University, Durham, North Carolina;* [5]*Department of Psychiatry and Behavioral Science, Duke University, Durham, North Carolina;* [6]*Department of Psychiatry, Veterans Affairs (VA) Hospital Medical Center, Nashville, Tennessee;* [7]*Department of Neurology, Duke University, Durham, North Carolina*

Communicated by Christine Van Broeckhoven

With the exception of ApoE (APOE), no universally accepted genetic association has been identified with late-onset Alzheimer disease (AD). A broad region of chromosome 10 has engendered continued interest generated from both preliminary genetic linkage and candidate gene studies. To better examine this region, we combined unbiased genetic linkage with candidate gene association studies. We genotyped 36 SNPs evenly spaced across 80.2 Mb in a family-based data set containing 1,337 discordant sibling pairs in 567 multiplex families to narrow the peak region of linkage using both covariate and subset analyses. Simultaneously, we examined five functional candidate genes (*VR22, LRRTM3, PLAU, TNFRSF6,* and *IDE*) that also fell within the broad area of linkage. A total of 50 SNPs were genotyped across the genes in the family-based data set and an independent case–control data set containing 483 cases and 879 controls. Of the 50 SNPs in the five candidate genes, 22 gave nominally significant association results in at least one data set, with at least one positive SNP in each gene. SNPs rs2441718 and rs2456737 in VR22 (67.8 Mb) showed association in both family-based and case–control data sets (both P = 0.03). A two-point logarithmic odds (LOD) score of 2.69 was obtained at SNP rs1890739 (45.1 Mb, P = 0.03 in 21% of the families) when the families were ordered from low to high by ApoE LOD score using ordered subset analysis (OSA). These data continue to support a role for chromosome 10 loci in AD. However, the candidate gene and linkage analysis results did not converge, suggesting that there is more extensive heterogeneity on chromosome 10 than previously appreciated. Hum Mutat 28(11), 1065–1073, 2007. Published 2007 Wiley-Liss, Inc.†

KEY WORDS: Alzheimer disease; chromosome 10; linkage; association; VR22; alpha-catenin; LRRTM3; PLAU; TNFRSF6; IDE; APOE

INTRODUCTION

Alzheimer disease (AD; MIM# 104300) is a progressive neurodegenerative disorder of later life with a complex etiology and a strong genetic component. Mutations in three genes encoding β-amyloid precursor protein (APP; MIM# 104760), presenilin 1 (PS1; MIM# 104311), and presenilin 2 (PS2; MIM# 600759) genes cause the rare early-onset AD (EOAD). ApoE (APOE; MIM# 107741) is the only gene universally accepted as an important risk factor for late-onset AD (LOAD). However, more than 50% of Alzheimer disease cases do not carry an ApoE ε4 allele, suggesting that other genetic risk factors exist [Farrer et al., 1997; Jarvik et al., 1996; Pericak-Vance and Haines, 1995; Bennett et al., 1995].

Numerous studies have tested hundreds of functional candidate genes and many genes for AD have reported positive associations in at least one study group [Bertram et al., 2007]. However, there is no single locus that has been sufficiently replicated to be widely accepted as an AD risk gene, except ApoE. Genome screening is an alternative approach that identifies *AD* gene locations independent of any presumed knowledge of function. These have suggested that the loci for LOAD lie on several chromosomes,

Received 5 October 2006; accepted revised manuscript 30 April 2007.

*Correspondence to: Jonathan L. Haines, Ph.D., Center for Human Genetics Research, Vanderbilt University Medical Center, 519 Light Hall, Nashville,TN 37232-0700.
E-mail: jonathan@chgr.mc.vanderbilt.edu

Grant sponsor: NIH; Grant numbers: AG019757; AG021547; AG20135 and AG21886; Grant sponsor: Alzheimer Association; Grant number: IIRG-00-2006.

Current address of Eden R. Martin, Stephan Züchner, John R. Gilbert, and Margaret A. Pericak-Vance: Miami Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL 33136.

DOI 10.1002/humu.20567
Published online 15 June 2007 in Wiley InterScience (www.interscience.wiley.com).

including chromosomes 9, 10, and 12 [Pericak-Vance et al., 1997; Kehoe et al., 1999; Myers et al., 2002; Blacker et al., 2003]. Perhaps the most interesting region is chromosome 10 since two independent linkage studies identified the same chromosome region using different methods [Myers et al., 2000; Ertekin-Taner et al., 2000] and a candidate gene driven linkage analysis found a linked region at 113 cM on chromosome 10q, near the insulin-degrading enzyme (IDE) [Bertram et al., 2000]. Recently, the microsatellite marker D10S1208 (63.3 cM) was linked to AD among families with affected mothers [Bassett et al., 2002, 2005]. The multiple reports of chromosome 10 linkage point to linkage peaks scattered across an 80-Mb region and provide an impractically large area for a molecular genetic search. To examine this region in more detail, we tested several recently reported candidate genes within this broad region of linkage, VR22 (also known as CTNNA3; MIM# 607667), LRRTM3 (MIM# 610869), PLAU (MIM# 191840), TNFRSF6 (MIM# 134637), and IDE (MIM# 146680) [Bertram et al., 2000; Ertekin-Taner et al., 2003; Feuk et al., 2000; Finckh et al., 2003; Martin et al., 2005]. Simultaneously, we examined the linkage signal in more detail to see if we could narrow the peak region of linkage and converge the candidate gene and linkage data.

All the notable differences in the location of the linkage signals call attention to the potential for locus heterogeneity and highlight the need for strategies to increase homogeneity. One strategy is to examine phenotypic or genotypic subgroups, e.g., based on ApoE genotype. Another strategy is to apply new statistical techniques, such as ordered subset analysis (OSA). We performed both a linkage study at a 2-Mb spacing across this 80-Mb region and simultaneously performed an association study of five candidate genes within this region to identify a minimum candidate region on chromosome 10 and the putative underlying AD gene(s) in the region.

PATIENTS AND METHODS

Study Populations

We used a case–control sample consisting of 483 unrelated cases with probable or definite AD and 879 unrelated cognitively normal elderly controls who were either the spouses of AD patients or subjects recruited from the outpatient clinics of the participating institutions. Cases and controls were collected through the Center for Human Genetics Research (CHGR) at Vanderbilt University and the Center for Human Genetics (CHG) at Duke University. All individuals included in this study were Caucasian. Written consent was obtained from all participants in agreement with protocols approved by the institutional review board at each contributing center. AD was diagnosed according to the National Institute of Neurological and Communication Disorders and Stroke (NINCDS)-Alzheimer's Disease and Related Disorders Association (ADRDA) criteria [McKhann et al., 1984]. Controls had no obvious signs of cognitive or neurological impairment when enrolled in the study as determined by a personal interview conducted by clinical personnel at Vanderbilt CHGR and Duke CHG. All controls had results within the normal range in the Mini-Mental State Exam (MMSE) or Modified Mini-Mental State Exam (3MS). Age at onset was recorded as the age at which the first symptoms were noted by the participant or a family member. If the affected individuals were early in the disease process, we included their report of age-at-onset as part of the determination. If the disease was more advanced, we only used information as collected from multiple family members (such as spouse and children). Mean $\pm$ standard deviation (SD) of age at onset (AAO) in cases was 71.8 ± 6.0 years. The control group's mean age at examination (AAE) $\pm$ SD was 72.0 ± 6.3 years, while the case group's mean AAE $\pm$ SD was 76.5 ± 6.5 years. The AD cases were 63.3% female, while the controls were 58.5% female.

Our independent family sample had 730 pedigrees with a total of 1,337 affected discordant sibling pairs (DSP) with a total of 1,521 LOAD patients (minimum AAO $\geq$ 60 years) and 974 unaffected individuals (Table 1). Spouses of AD patients were used as controls in the case–control data set if individuals had no obvious signs of cognitive or neurological impairment when enrolled in the study as determined by a personal interview conducted by clinical personnel. All controls had results within the normal range in the MMSE or 3MS. Family data were ascertained by the following centers: the National Cell Repository for Alzheimer's Disease (NCRAD) repository at Indiana University; the Collaborative Alzheimer Project (CAP), including Duke and Vanderbilt Universities and the University of California at Los Angeles; and the National Institute of Mental Health (NIMH) repository. Criteria for AD diagnosis and screening of unaffected relatives were

TABLE 1. **Study Populations for Alzheimer Disease**

Family	Overall	NIMH	NCRAD	CAP
Family-based data set				
Total family				
Total pedigrees	**730**	**352**	**154**	**224**
Affected individuals	**1521**	**807**	**315**	**390**
Unaffected individuals	**974**	**331**	**162**	**481**
Discordant sibling pairs (DSP)	**1337**	**629**	**269**	**439**
Independent discordant sibling pairs	**674**	**283**	**129**	**262**
Pedigrees with at least one DSP	**406**	**165**	**75**	**166**
Affected relative pairs (ARP)	**188**	**66**	**26**	**96**
Pedigrees with at least one ARP	**64**	**31**	**11**	**22**
Multiplex family				
Total pedigrees	**567**	**352**	**122**	**96**
Affected sibling pairs (ASP)	**922**	**517**	**190**	**225**
Independent affected sibling pairs	**707**	**418**	**150**	**139**
Pedigrees with at least one ASP	**534**	**331**	**118**	**85**
Sporadic data set	**483**	**1**	**39**	**443**
Case				
Control	**879**	**0**	**0**	**879**

NIMH, National Institute of Mental Health repository; NCRAD, NCRAD repository at Indiana University; CAP, Collaborative Alzheimer Project.

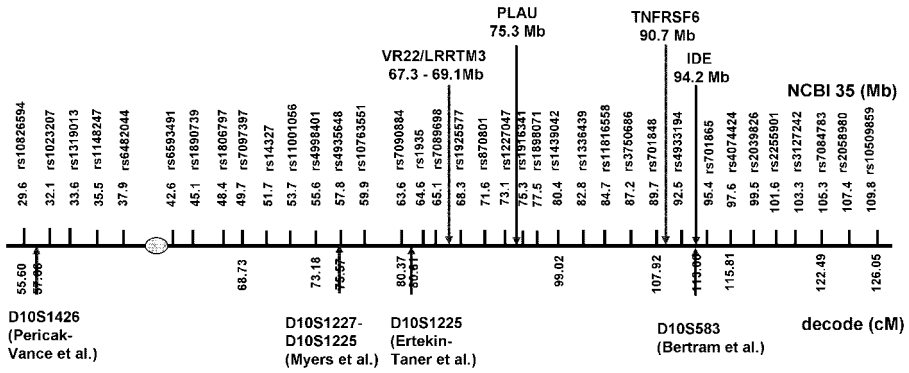

FIGURE 1. **The position of candidate genes and 36 genotyped SNPs for linkage analysis on chromosome 10. Upper panel shows the physical location (Mb) of 36 SNPs for linkage analysis according to NCBI build 35; the lower panel shows the DECODE genomic map location (cM) of linked markers shown in the literature.**

the same as described above. All participants were Caucasian Americans. Mean ± SD of AAO in affected individuals in the family-based sample was 72.9 ± 6.4 years. The mean AAE ± SD was 80.1 ± 7.1 and 69.9 ± 11.2 in affected and unaffected individuals, respectively. The AD affected group was 67.3% female, while the unaffected group was 56.2% female.

SNP Selection and Genotyping

Following informed consent, blood samples were collected from each individual. Genomic DNA was obtained from the repositories (NIMH, NCRAD) or extracted from whole blood (CAP) by use of the Puregene system (Gentra Systems, Minneapolis, MN). All samples were coded and stored at 4°C until used.

A map of the chromosome 10 region and previously reported candidate genes is shown in Fig. 1. NCBI (www.ncbi.nih.gov), Ensembl (www.ensembl.org), and Applied Biosystems websites (http://home.appliedbiosystems.com) were mined to select SNPs according to their location relative to other selected SNPs, high minor allele frequency (≥ 0.25), and availability of quality assays. We carefully selected 36 SNPs at a 2-Mb spacing for linkage analysis (Fig. 1). SNPs in candidate genes were selected using multiple criteria. First, we selected the SNPs that showed association in previous reports. Second, any known functional SNPs were included. Third, HapMap data were mined to select TagSNPs using Haploview [Barrett et al., 2005]. Finally, SNPs with high heterozygosity and available high-quality assays were used to fill any remaining large genomic gaps. SNP genotyping was performed using Assays-On-Demand or Assays-By-Design (Applied Biosystems, Foster City, CA) with the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems). Amplification was performed in a 384-well DNA Engine Tetrad 2 Peltier Thermal Cycler (MJ Research, Waltham, MA) with the following conditions: 94°C for 10 min; 92°C for 15 sec, 60°C for 1 min (50 cycles); and held at 4°C. Systematic genotyping errors were minimized by use of a system of quality control (QC) checks with duplicated samples [Rimmler et al., 1998]. Linkage disequilibrium (LD) and deviations from Hardy-Weinberg equilibrium (HWE) were assessed using Haploview [Barrett et al., 2005].

Statistical Methods

Association analyses. Case–control association analyses for single alleles and for genotypes were conducted using logistic regression (version 8.1; SAS Institute, Cary, NC). We tested three different models. Model I assumed an additive effect on the log scale for the alleles (e.g., if A is the minor allele, having no A alleles = 0, having one A allele = 1, having two A alleles = 2). Model II dichotomized genotype according to different alleles (having allele A was coded as 1, not having allele A was coded as 0). Model III categorized genotypes into three groups using the most frequent genotype as a reference genotype and other groups were compared with the reference group separately. Statistical significance was declared at $\alpha = 0.05$.

To adjust for potential confounders, we included gender and AAE as covariates in the regression analysis. In addition, we included ApoE status in the model to exclude its possible confounding effect. We considered ApoE status as three categories: those that have two ApoE ε4 alleles, those that have at least one ApoE ε4 allele, and those that have no ApoE ε4 allele. We decided to analyze the data using models with and without covariates because the data set had 7.4% missing data after we adjusted confounders.

Haplotype analyses in the case–control data set were conducted using haplo.stats [Schaid et al., 2002]. An expectation-maximization (EM) algorithm was used to generate haplotype frequency estimates under the null hypothesis. We estimated haplotype frequencies and tested association of each haplotype with a frequency of at least 5% in our case–control data set with age- and gender-adjusted score statistics except VR22. A minimum of 25 counts in the dataset for a haplotype was used as the cutoff for selecting haplotypes in VR22 because of the low haplotype frequencies in the gene. Haplotype logistic regression was modeled using a generalized linear model (GLM) algorithm including age and gender as covariates. The most frequent haplotype occurring in a similar percentage of cases and controls was selected as the baseline haplotype. To evaluate the association of subsets of alleles from the full haplotype, a sliding window of three SNPs was used and the global score statistics was reported.

Family-based association analysis was conducted using the pedigree disequilibrium test (PDT, for allelic effect) and Genotype-PDT (GenoPDT, for genotypic effect) for single-locus tests to assess association between genotypes and risk of AD in the family data [Martin et al., 2000, 2003]. The Haplotype Family Based Association Test [Horvath et al., 2001] was used for haplotype analysis in the family data set.

Linkage analyses of family data set. Two-point heterogeneity logarithmic odds (LOD) score (HLOD) analyses were computed using FASTLINK and HOMOG [Ott, 1999]. Because the mode of inheritance for AD is unknown, affected-only parametric analyses were performed using both autosomal

dominant and autosomal recessive models with disease allele frequencies of 0.001 and 0.20, respectively, to model the susceptibility allele. Greenberg et al. [1998] suggested that a critical factor in LOD-score analysis is the mode of inheritance (MOI) at the linked locus, not that of the disease or trait per se. Thus, a limited set of simple genetic models in LOD-score analysis can work well in testing for linkage. If a signal is picked up in either one of the models, there is increased power to detect true linkage using this approach.

Because it is likely that there is genetic heterogeneity in AD, we applied OSA [Hauser et al., 2004] in our linkage analysis to test for a subset of linked families. In OSA, the statistical significance of increased evidence for linkage in a subset of the data relative to evidence for linkage in the entire sample is assessed via random permutation of the order of inclusion of the families to estimate empirical P values. Since the ApoE ε4 allele is the only known risk allele for AD, we used the ApoE LOD score (at theta = 0) and the ApoE weight (the proportion of affecteds within a family who carry at least one ApoE ε4 allele) as covariates to order families from low to high and high to low. A locus on chromosome 12 has been suggested to be linked with AD [Liang et al., 2006; Mayeux et al., 2002], so the LOD score of the linked marker D12S368 was used also as a covariate to order families in OSA analysis. The same parametric models were applied to D12S368. OSA was also performed using mean age of onset as a covariate.

Multipoint LOD scores were calculated using MERLIN [Abecasis et al., 2002] and OSA multipoint LOD scores were calculated using the four covariates as described above. Conditional linkage analysis using ApoE weight as a covariate was done in Allegro [Gudbjartsson et al., 2000]. The analysis was weighted by the proportion of affecteds within a family who carry at least one ApoE ε4 allele and by proportion of affecteds within a family who do not carry ApoE ε4 allele (inverse ApoE– ε4 weighted). There was no LD between 34 out of the 36 linkage analysis SNPs and very low LD only between the other two linkage SNPs (rs1935 and rs7089698, $r^2 = 0.46$). We used the LD option of MERLIN to perform the multipoint linkage analysis to take the LD into account. An $r^2 > 0.1$ was used as the criterion to define SNPs in LD. In the OSA multipoint analysis, we excluded the SNP (rs7089698) which was in LD with the other SNP (rs1935) as it had the lower minor allele frequency. Because LD between SNPs within the five candidate genes was high, we used $r^2 > 0.1$ as the criterion in MERLIN to define SNPs in LD to be able to perform the multipoint linkage analysis in candidate genes.

RESULTS

Candidate Gene Association Analyses

We selected a number of representative SNPs in the five candidate genes, *VR22* (N = 20 SNPs), *LRRTM3* (N = 14 SNPs), *PLAU* (N = 5 SNPs), *TNFRSF6* (N = 4 SNPs), and *IDE* (N = 7 SNPs) on chromosome 10. The SNPs that showed nominal association with the disease are listed in Table 2. At least one SNP showed association in each of these five candidate genes, *VR22* (N = 7 SNPs), *LRRTM3* (N = 7 SNPs), *PLAU* (N = 4 SNPs), *TNFRSF6* (N = 3 SNPs), and *IDE* (N = 1 SNP), in either the

TABLE 2. **Candidate Genes and SNPs Showing Association in at Least One Data Set**

				Case–control[b]		Family[c]
Gene	SNP	Allele variation	Nucleotide position[a]	P	OR (95%CI)	P
VR22	rs7911820	G>T	67,534,145	0.27[d]	0.91 (0.77, 1.08)	0.03[e,*]
	rs7070570	A>G	67,534,610	0.06[f]	1.69 (0.97, 2.95)	0.03[g,*]
	rs7074454	C>T	67,534,965	0.14[h]	0.64 (0.35, 1.16)	0.01[e,*]
	rs2441718	A>G	67,806,967	0.03[i,**]	1.29 (1.02, 1.62)**	0.03[e,**]
	rs2456737	A>G	67,825,340	0.03[i,**]	1.33 (1.04, 1.72)**	0.05[g,**]
	rs7909676	A>C	68,104,803	0.33[i]	1.14 (0.88, 1.48)	0.00[g,*]
	rs3096244	C>T	69,080,192	0.15[j]	1.16 (0.95, 1.42)	0.05[e,*]
LRRTM3	rs12769870	A>G	68,347,401	0.01[f,*]	0.70 (0.53, 0.92)*	0.08[g]
	rs2394314	A>G	68,350,254	0.04[k,*]	0.72 (0.54, 0.98)*	0.14[g]
	rs10762122	C>T	68,386,380	0.02[j,*]	1.24 (1.02, 1.50)*	0.09[g]
	rs942780	A>G	68,406,547	0.48[k]	1.11 (0.83, 1.48)	0.03[g,*]
	rs1925617	G>T	68,434,823	0.17[i]	0.82 (0.61, 1.09)	0.00[g,*]
	rs1925622	A>G	68,439,644	0.48[f]	0.90 (0.66, 1.22)	0.01[g,*]
	rs1952060	C>T	68,492,940	0.24[i]	0.86 (0.67, 1.11)	0.00[g,*]
PLAU	rs1916341	A>C	75,341,168	0.20[f]	0.85 (0.66, 1.09)	0.02[e,*]
	rs2227566	C>T	75,343,737	0.35[k]	0.87 (0.65, 1.16)	0.01[e,*]
	rs2227568	C>T	75,343,885	0.11[j]	0.80 (0.61, 1.05)	0.01[g,*]
	rs4065	C>T	75,346,470	0.22[f]	0.85 (0.67, 1.10)	0.02[e,*]
TNFRSF6	rs1800682	C>T	90,739,943	0.27[a]	1.19 (0.87, 1.61)	0.03[g,*]
	rs2031612	A>G	90,756,960	0.02[d,*]	1.23 (1.04, 1.46)*	0.87[e]
	rs2296600	C>G	90,760,419	0.02[d,*]	0.82 (0.69, 0.97)*	0.26[g]
IDE	rs7099761	A>G	94,325,779	0.50[f]	0.91 (0.68, 1.20)	0.02[e,*]

[a]The map location is from NCBI dbSNP, build 35.
[b]We tested a total of six models for the case–control data set and only the most significant results are shown.
[c]We applied two methods to the family-based association test and only the most significant results are shown.
[d]Model I with genotype-only.
[e]PDT.
[f]Model III with genotype-only.
[g]Geno-PDT.
[h]Model III with both genotype and covariates.
[i]Model II with genotype-only.
[j]Model I with both genotype and covariates.
[k]Model II with both genotype and covariates.
*SNPs showing association ($P < 0.05$) in either the case–control or family-based data set, but not both.
**SNPs showing association ($P < 0.05$) in both case–control and family-based data sets.

case–control, family-based, or both data sets. The number of associated SNPs is shown in parentheses. SNPs rs2441718 and rs2456737 in VR22 showed association in both the family-based and the case–control data sets, with an odds ratio (OR) in the case–control data set of 1.29 (95% CI = 1.02–1.62; P = 0.03) and 1.33 (95% CI = 1.04–1.72; P = 0.03), respectively, when we used major allele to dichotomize genotypes. However, the association went away after we adjusted for the confounders (age, gender, and ApoE status). SNP rs2031612 in TNFRSF6 showed association in the case–control data set with an OR of 1.23 (95% CI = 1.04–1.46; P = 0.02 with the major allele A under the additive model). SNP rs10762122 in LRRTM3 also showed association in the case–control data set after we adjusted for confounders, with an OR of 1.24 (95% CI = 1.02–1.50; P = 0.02 with the major allele T). The complete sets of results are available upon request.

Using the haplo.stats program to calculate age- and gender-adjusted score statistics, two haplotypes in PLAU were associated with AD in the overall case–control data set (global P = 0.03; Table 3). Haplotype analyses on other candidate genes did not show significant results at P < 0.05 (Table 3). All five genotyped SNPs (rs1916341, rs2227564, rs2227566, rs2227568, and rs4065) in PLAU were used to define the haplotype. Haplotype GTCCC was associated with AD at P = 0.01 and haplotype GCCTC at P = 0.03 for global score statistics (Table 4). These two haplotypes remained significant if we apply false discovery rate (FDR) correction for multiple comparisons. The most frequent haplotype TCTCT was used as the referent haplotype (P = 0.43). Haplotype GTCCC was also borderline significant (P = 0.05) in the logistic regression after taking into account age and gender (data not shown). A sliding window analysis (n = 3) across five genotyped SNPs in PLAU generated three haplotypes and all three were significant (P < 0.05 for age and gender adjusted score statistics, data not show). Although we also saw a significant effect for PLAU haplotypes in the family-based data set, the association was in the opposite direction. The most common haplotype, TCTCT, was significantly associated with AD (P = 0.01) in haplotype FBAT analysis, while the haplotype GTCCC was not associated with AD (P = 0.70). Haplotype GCCTC had marginal effect (P = 0.05; Table 4).

Two-Point Linkage Analyses

A total of 36 SNPs evenly spaced across a 80.2-Mb region on chromosome 10 were genotyped in 567 multiplex families (922 affected sibling pairs). The reliability of duplicate genotyping across plates was >99% and the average genotyping efficiency was 98%. All analyses were applied to the overall data set and three stratified subsets (autopsy-confirmed, ApoE ε4+, ApoE ε4−).

Using the 36-SNP linkage panel, we conducted both FASTLINK/HOMOG two-point analyses and OSA two-point analyses using ApoE LOD score, ApoE weight, and D12S368 LOD score as covariates. In the overall data set, only SNP rs11816558 at 84.7 Mb generated a LOD > 1.0 (1.31). In the OSA analysis using the ApoE LOD score as a covariate, a two-point LOD score of 2.69 was obtained at SNP rs1890739 at 45.1 Mb (P = 0.03 in 21% of the families, low to high), rs1319013 at 33.6 Mb generated LOD = 2.50, and rs11816558 at 84.7 Mb generated LOD = 1.70 (Fig. 2). SNP rs14327 (51.7 Mb) generated a two-point LOD score of 1.94 in 13% of the families when the families were ordered from low to high using the LOD score from the linkage peak on chromosome 12 as a covariate (P = 0.009; data not shown). In the ApoE ε4− subset, SNP rs3750686 at 87.2 Mb generated a parametric two-point LOD score of 1.54. This SNP also had a two-point LOD of 2.30 in OSA using the linkage signal

TABLE 3. **Haplotypic Association of the Candidate Genes in the Case–Control Data Set**

Gene	Number of genotyped SNPs	Number of haplotypes		Global P
		Total	Frequency >5%	
VR22-A	11	109	8[b]	0.29
VR22-B	11[a]	155	3[b]	0.95
LRRTM3-A	7	30	6	0.19
LRRTM3-B	9[a]	47	5	0.90
PLAU	5	11	3	0.03*
TNFRSF6	4	9	4	0.46
IDE	7	18	5	0.84

[a]More than 10 SNPs were genotyped in the gene and they were split into two groups to be able to run in haplo.stats. The second half of the SNP set had two overlapped SNPs with the first half of the set.
[b]Haplotypes with more than 25 counts in the data set.
*P < 0.05.

TABLE 4. **Age- and Gender-Adjusted Haplotypic Association in PLAU in the Case–Control Data Set and Results of HBAT in the Family-Based Data Set**

Haplotype	SNPs[a]					Case–control				Family-based	
						Frequency					
	1	2	3	4	5	Total	Case	Control	Global P	Total frequency	P
h1	G	C	C	T	C	0.15	0.12	0.16	0.03*	0.14	0.05
h2	G	T	C	C	C	0.22	0.24	0.20	0.01*	0.23	0.70
Baseline	T	C	T	C	T	0.59	0.58	0.60	0.43	0.56	0.01*

[a]SNP1, rs1916341; SNP2, rs2227564; SNP3, rs2227566; SNP4, rs2227568; SNP5, rs4065.
*P < 0.05.

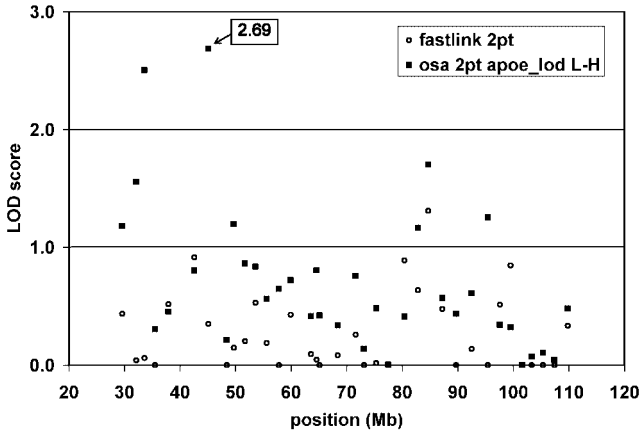

FIGURE 2. **Two-point analyses in the overall data set. Ordered subset two-point analysis in overall data set obtained a two-point LOD score of 2.69 at 45.1 Mb (P = 0.03 in 21% of the families) using the ApoE LOD score to order families from low to high.**

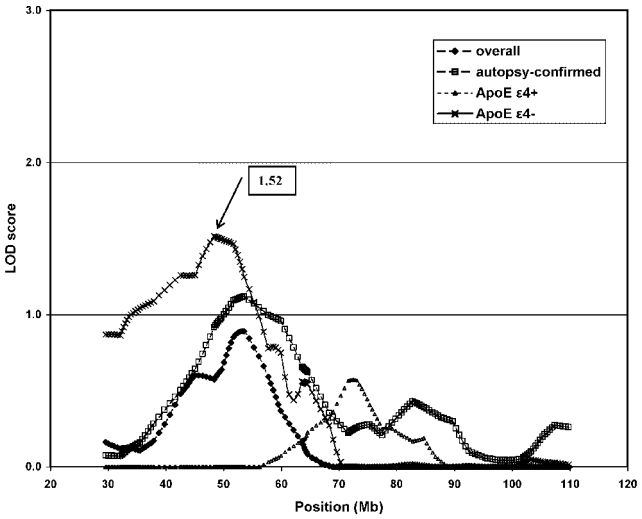

FIGURE 3. **OSA multipoint analysis (D12S368 LOD low to high). The peak LOD = 1.52 around 49 Mb between rs1806797 and rs7097397 was seen in the ApoEε4− subset when the families were ordered from low to high by the peak linkage signal on chromosome 12 (D12S368 LOD score).**

on chromosome 12 from low to high as a covariate (data not shown).

We also performed linkage analyses on the SNPs in the candidate genes, applying the same approaches as the SNP linkage panel on the overall data set. Two SNPs (rs7070570 and rs2441718) in VR22 had a LOD score >1.0 (1.30 and 1.02, respectively) in the FASTLINK/HOMOG two-point analyses. In the OSA analysis using ApoE LOD score as a covariate, a two-point LOD score of 2.30 was obtained at SNP rs2456737 in VR22 (P = 0.05 in 51% of the families, high to low). A two-point LOD score of 1.62 was obtained at SNP rs942780 in LRRTM3 (P = 0.02 in 22% of the families, families were ordered from low to high by ApoE LOD score, data not shown).

Multipoint Linkage Analyses

Multipoint analysis using the 36 SNP linkage panel on the overall data set did not show a suggestive LOD score (Fig. 3). When we applied the ApoE LOD score as a covariate to order families from low to high in OSA, the highest LOD score was 1.85 (rs10826594). The next two SNPs (rs1023207 and rs1319013) also had LOD scores higher than 1.0. SNP rs14327 (51.7 Mb) had a LOD of 1.03 using the same covariate in OSA. There was no SNP showing linkage with the disease in ApoE ε4+ subset (Fig. 3). In the ApoE ε4− subset, rs1806797 (48.3 Mb) showed a peak LOD score of 1.52 when the families were ordered from low to high by the peak linkage signal on chromosome 12 (D12S368 LOD score) [Liang et al., 2006] with P = 0.04 (Fig. 3). In the autopsy-confirmed subset, there was no effect using the number of ApoE ε4 alleles as a covariate. However, a peak LOD of 1.91 was generated at 49.7 Mb between rs7097397 and rs14327 in the OSA multipoint analysis when the ApoE LOD score was used as the covariate and the families were ranked from low to high (Fig. 4). Thus families that were not linked to ApoE or the locus on chromosome 12 had suggestive LOD scores around 49.7 Mb on chromosome 10.

We also performed multipoint analysis using SNPs in the candidate genes on the overall data set. Since there was high LD between SNPs in genes, we used $r^2 > 0.1$ to define SNPs in LD in the program MERLIN. Only one SNP (rs2456737) in VR22 had a LOD score of 1.05 (data not shown). This SNP also showed

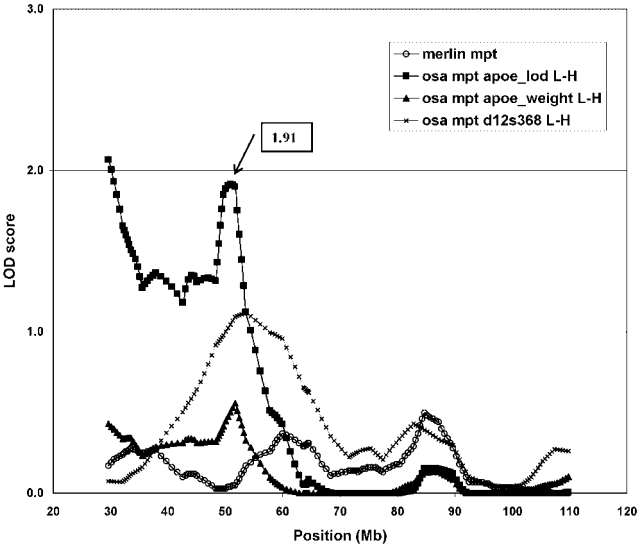

FIGURE 4. **Multipoint analyses in the autopsy-confirmed data set. Ordered subset multipoint analysis in autopsy-confirmed subset generated the peak between 40 and 60 Mb when the ApoE LOD score was used as a covariate to order families from low to high.**

association in both the case–control and the family-based data sets (P < 0.05).

Genomic Convergence of Association and Linkage Analyses

Both linkage and association studies gave us suggestive results, shown in Fig. 5. However, these results did not converge. The linkage study revealed a major peak between 40 and 60 Mb (LOD = 2.69) and a minor peak between 80 and 100 Mb (LOD = 1.70). In association analysis, VR22 (67.8 Mb), LRRTM3 (68.3 Mb), TNFRSF6 (90.7 Mb), and PLAU (75.3 Mb) showed evidence of association. VR22 showed allelic association in both the family and the case–control data sets. PLAU showed

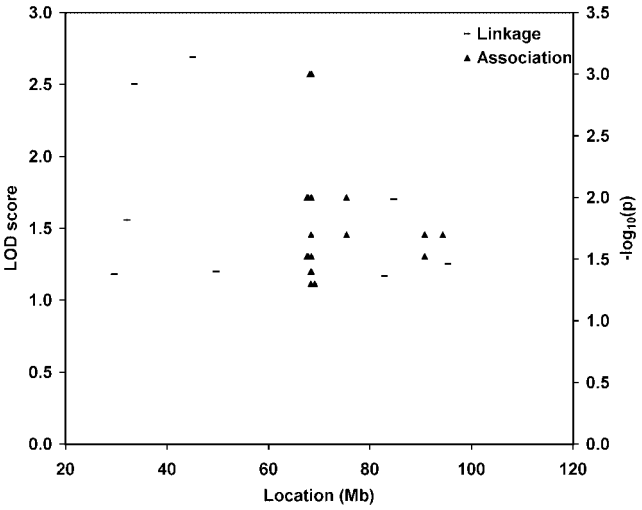

FIGURE 5. **Overall suggestive linkage and association data on chromosome 10. Triangles are $-\log_{10}$ (P) from candidate gene association study ($P < 0.05$). The dashes are two-point LOD scores (LOD > 1.0) in OSA from linkage study.**

haplotypic association in both the family and the case–control data sets. LRRTM3 demonstrated allelic association in the case–control data set after adjusting for confounders, and it survived both Bonferroni and FDR correction for multiple comparisons in the family data set. The SNPs in TNFRSF6 remained allelic association after FDR correction for multiple testing ($q = 0.2$) in the case–control data set. Only TNFRSF6 (90.7 Mb) falls near one of the observed linkage peaks (80–100 Mb).

DISCUSSION

In 2000, three groups reported significant linkage of LOAD to two regions of the long arm of chromosome 10 [Bertram et al., 2000; Ertekin-Taner et al., 2000; Myers et al., 2000]. One study [Myers et al., 2000] followed up "suggestive" [Lander and Kruglyak, 1995] linkage results on this chromosome using an affected sibling-pair method resulting in significant linkage near ~58 Mb. Using a different methodology, a second study [Ertekin-Taner et al., 2000] performed multipoint linkage analyses in five LOAD families using amyloid β (Aβ) plasma levels as a quantitative phenotype and found significant linkage very close to the same chromosomal region. The third report [Bertram et al., 2000] was a candidate gene–driven linkage analysis of 435 families using six genetic markers in a region of chromosome 10 that lies ~40 Mb distal to the one implicated by the former two studies. The candidate gene of interest encodes the insulin-degrading enzyme (*IDE*) that, together with neprilysin and other proteases, has been previously suggested to play a major role in the degradation and clearance of Aβ in brain [Selkoe, 2001]. Two of the six markers in this region showed significant linkage to AD. In addition, one marker, *D10S583*, displayed modest evidence for linkage disequilibrium with AD and one of its alleles using a diallelic test ($P = 0.04$ after correction for multiple testing). Across this broad region on chromosome 10q, there are many candidate genes. In our study, we focused on VR22 (α-catenin), LRRTM3, PLAU, TNFRSF6, and IDE. *VR22* is a large (1.8 Mb) gene located at 67.3 Mb that encodes alpha-T catenin, which is a binding partner of beta catenin. This makes *VR22* an attractive candidate gene because beta catenin interacts with presenilin 1, which has many mutations that elevate Aβ42 and cause early onset familial AD. *LRRTM3* is a nested gene in an intron of VR22, which also has plausible biological function of promoting APP processing by β-secretase (BACE1). Ertekin-Taner et al. [2003] reported two intronic SNPs in VR22 showing highly significant association ($P = 0.0001$ and 0.0006) with plasma Aβ42 in 10 extended LOAD families. Martin et al. [2005] suggested significant evidence of association between VR22 and AD in both the families and the unrelated cases and controls, and the effect was dependent on ApoE status. PLAU encodes urokinase-type plasminogen activator (uPA), which converts plasminogen to plasmin. Plasmin is involved in processing of amyloid precursor protein (APP) and degrades secreted and aggregated Aβ, a hallmark of AD. Finckh et al. [2003] found a coding SNP in PLAU showing association ($P = 0.00039$) in 347 patients with LOAD and 291 control subjects. The TNFRSF6 gene encodes FAS, a cell-surface receptor involved in apoptosis initiation. Elevated levels of FAS have been reported in the brains of AD patients [Feuk et al., 2000]. Despite these tantalizing reports, none of the associations to these genes have been widely replicated in multiple studies and thus none have been generally accepted as an *AD* risk gene.

Although we found that there was at least one SNP in each gene showing association with AD, the results were not strikingly significant, especially in consideration of the problem of multiple comparisons. We adjusted for multiple comparisons in both the family and case–control data sets by the number of genotyped SNPs in each gene. Using the overly conservative Bonferroni correction, none of the 50 genotyped SNPs was significant at global $P < 0.05$ in the case–control data set. The only SNP remaining significant after Bonferroni correction was SNP (rs1952060) in LRRTM3 in the family-based data set. Perhaps a more appropriate approach is to control for the FDR. When we applied an FDR of $q = 20\%$, three SNPs in LRRTM3 (rs1925617, rs1925622, and rs1952060) and one SNP (rs2227568) in PLAU remained significant in the family-based data set. In the case–control data set, two SNPs (rs2031612 and rs2296600) in TNFRSF6 were significant.

PLAU is the only gene showing haplotypic association with AD in both the family-based and case–control data sets. Although the associated haplotype is different between two data sets, considering the strong single marker effect (one SNP survived FDR multiple comparison correction), these results remain very interesting. More detailed examination of this gene is needed to explore its role in AD.

Martin et al. [2005] tested 11 SNPs in VR22 and LRRTM3 in an overlapping data set, and we included them in our SNP list. We extended their analyses by genotyping 12 more SNPs in VR22 and 11 more SNPs in LRRTM3. These additional SNPs were selected because they were functional SNPs or tag SNPs or filled large genomic gaps. Although the overlap with the Martin et al. [2005] data set is extensive, the current data set includes more updated clinical information. We found the same associated SNPs (rs7911820 and rs7074454 in VR22; rs1925617 in LRRTM3) at $P < 0.05$ in the family-based data set, but also identified two SNPs (rs2441718 and rs2456737 in VR22) showing association in both the family-based and case–control data sets.

The inconsistencies between our study and other groups' findings may reflect differences in the samples, study designs, or analytical techniques. They may also reflect heterogeneity within the samples or more complicated genetic mechanisms, such as the interplay of genetic and epigenetic factors. Also, we selected several SNPs in each gene based on the previous reports, spacing, and heterozygosity. But the genes have not been exhaustively

examined. There is a possibility that association could be found if we genotype all tagSNPs to cover each haplotype block and do an exhaustive test of these tagSNPs in these genes. In addition, the candidate gene selection strategy is biased by the known biological function and the function of most genes is poorly understood. This prompted us to combine linkage and candidate gene association studies to try to locate the susceptibility genes underlying AD.

We hypothesized that locus heterogeneity might be one reason for the inconsistent linkage results. To increase homogeneity, we stratified our linkage data set by autopsy confirmation and ApoE status. We also applied the OSA method that can find a more homogeneous subset by incorporating covariates into analysis.

Our data suggest that applying covariates to the analysis increases the homogeneity in the data set. Before the covariate ApoE LOD score was used in the analysis, the highest two-point LOD score was 1.31. However, we found a major linkage peak and a minor linkage peak in the overall data set after we applied the ApoE LOD score as a covariate. One was between 40 and 60 Mb with peak LOD of 2.69, at SNP rs1890739 (45.1 Mb), the other was between 80 and 100 Mb with peak LOD of 1.70, at SNP rs11816558 (84.7 Mb) (Fig. 2). Furthermore, subsetting combined with covariate analysis helped us decrease the heterogeneity and refine the linkage peak. Without applying covariates, multipoint analysis in the overall data set had no SNP with LOD > 1.0. However, in the autopsy-confirmed subset, the only peak was located between 40 and 60 Mb when we used ApoE LOD score as a covariate to order families from low to high (Fig. 4). The peak LOD was 1.91, at SNP rs7097397 (49.7 Mb). These data suggest that the families without the ApoE ε4 allele may be linked to the disease at this location. When we used D12S368, a surrogate for linkage to chromosome 12, as a covariate to order families from low to high, we also saw the peak between 40 and 60 Mb in the autopsy-confirmed and ApoE ε4− subsets. Analysis using ApoE and D12S368 LOD scores greatly increased evidence of linkage in areas showing no linkage in the overall analyses. For example, rs1890739 at 45.1 Mb reached LOD of 2.69 among 21% of the families by using the covariate of ApoE LOD score to order families from low to high in OSA, whereas the same SNP showed a LOD of 0.35 when all families were analyzed together.

Taken together, our results show evidence for linkage of LOAD to chromosome 10 (near 45.1 Mb) and this locus is independent of the chromosome 12 locus and the ApoE gene. However, it is not convergent with candidate gene association study, in which VR22 (67.8 Mb) showed allelic association in both case–control and family-based data sets (Fig. 5). Since we applied linkage and candidate gene association approaches simultaneously, we could not predict the results in advance. Our results support the role of the tested candidate genes in AD, but it is also important to study the candidate genes around the 45 Mb location. Our results suggest more complicated heterogeneity on chromosome 10 with LOAD than previously expected. Further experiments on the possible roles of the genes on chromosome 10 in AD pathology will hopefully lead to a better understanding of the etiology of AD.

ACKNOWLEDGMENTS

We thank all of the individuals who participated in this study. Funding for collection and distribution of NCRAD samples was provided, in the past, by NIH grant AG21886. This work used the core resources of the General Clinical Research Center (GCRC) (MO1 RR-00095) and the CHGR at Vanderbilt University Medical Center (VUMC), and the CHG at Duke University Medical Center (DUMC).

REFERENCES

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. 2002. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.

Bassett SS, Avramopoulos D, Fallin D. 2002. Evidence for parent of origin effect in late-onset Alzheimer disease. Am J Med Genet 114: 679–686.

Bassett SS, Kusevic I, Cristinzio C, Yassa MA, Avramopoulos D, Yousem DM, Fallin MD. 2005. Brain activation in offspring of AD cases corresponds to 10q linkage. Ann Neurol 58:142–146.

Bennett C, Crawford F, Osborne A, Diaz P, Hoyne J, Lopez R, Roques P, Duara R, Rossor M, Mullan M. 1995. Evidence that the APOE locus influences rate of disease progression in late onset familial Alzheimer's Disease but is not causative. Am J Med Genet 60:1–6.

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

Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. 2007. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 39:17–23.

Blacker D, Bertram L, Saunders AJ, Moscarillo TJ, Albert MS, Wiener H, Perry RT, Collins JS, Harrell LE, Go RC, Mahoney A, Beaty T, Fallin MD, Avramopoulos D, Chase GA, Folstein MF, McInnis MG, Bassett SS, Doheny KJ, Pugh EW, Tanzi RE. 2003. Results of a high-resolution genome screen of 437 Alzheimer's disease families. Hum Mol Genet 12:23–32.

Ertekin-Taner N, Graff-Radford N, Younkin LH, Eckman C, Baker M, Adamson J, Ronald J, Blangero J, Hutton M, Younkin SG. 2000. Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. Science 290:2303–2304.

Ertekin-Taner N, Ronald J, Asahara H, Younkin L, Hella M, Jain S, Gnida E, Younkin S, Fadale D, Ohyagi Y, Singleton A, Scanlin L, de AM, Petersen R, Graff-Radford N, Hutton M, Younkin S. 2003. Fine mapping of the alpha-T catenin gene to a quantitative trait locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. Hum Mol Genet 12: 3133–3143.

Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM, APOE and Alzheimer Disease Meta Analysis Consortium. 1997. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. JAMA 278:1349–1356.

Feuk L, Prince JA, Breen G, Emahazion T, Carothers A, St Clair D, Brookes AJ. 2000. apolipoprotein-E dependent role for the FAS receptor in early onset Alzheimer's disease: finding of a positive association for a polymorphism in the *TNFRSF6* gene. Hum Genet 107:391–396.

Finckh U, Van Hadeln K, Muller-Thomsen T, Alberici A, Binetti G, Hock C, Nitsch RM, Stoppe G, Reiss J, Gal A. 2003. Association of late-onset Alzheimer disease with a genotype of *PLAU*, the gene encoding urokinase-type plasminogen activator on chromosome 10q22.2. Neurogenetics 4:213–217.

Greenberg DA, Abreu P, Hodge SE. 1998. The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 63:870–879.

Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. 2000. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25:12–13.

Hauser ER, Watanabe RM, Duren WL, Bass MP, Langefeld CD, Boehnke M. 2004. Ordered subset analysis in genetic linkage mapping of complex traits. Genet Epidemiol 27:53–63.

Horvath S, Xu X, Laird NM. 2001. The family based association test method: strategies for studying general genotype—phenotype associations. Eur J Hum Genet 9:301–306.

Jarvik G, Larson EB, Goddard K, Schellenberg GD, Wijsman EM. 1996. Influence of apolipoprotein E genotype on the transmission of Alzheimer disease in a community-based sample. Am J Hum Genet 58:191–200. [Published erratum appears in Am J Hum Genet 1996;58:648]

Kehoe P, Wavrant-De VF, 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.

Lander ES, Kruglyak L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nature Genetics 11:241–247.

Liang X, Schnetz-Boutaud N, Kenealy SJ, Jiang L, Bartlett J, Lynch B, Gaskell PC, Gwirtsman H, McFarland L, Bembe ML, Bronson P, Gilbert JR, Martin ER, Pericak-Vance MA, Haines JL. 2006. Covariate analysis of late-onset Alzheimer disease refines the chromosome 12 locus. Mol Psychiatry 11:280–285.

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.

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.

Martin ER, Bronson PG, Li YJ, Wall N, Chung RH, Schmechel DE, Small G, Xu PT, Bartlett J, Schnetz-Boutaud N, Haines JL, Gilbert JR, Pericak-Vance MA. 2005. Interaction between the alpha-T catenin gene (*VR22*) and *APOE* in Alzheimer's disease. J Med Genet 42:787–792.

Mayeux R, Lee JH, Romas SN, Mayo D, Santana V, Williamson J, Ciappa A, Rondon HZ, Estevez P, Lantigua R, Medrano M, Torres M, Stern Y, Tycko B, Knowles JA. 2002. Chromosome-12 mapping of late-onset Alzheimer disease among Caribbean Hispanics. Am J Hum Genet 70:237–243.

McKhann G, Drachman G, 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.

Myers A, Holmans P, Marshall H, Kwon J, Meyer D, Ramic D, Shears S, Booth J, DeVrieze FW, 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 AM. 2000. Susceptibility locus for Alzheimer's disease on chromosome 10. Science 290:2304–2305.

Myers A, Wavrant De-Vrieze F, Holmans P, Hamshere M, Crook R, Compton D, Marshall H, Meyer D, Shears S, Booth J, Ramic D, Knowles H, Morris JC, Williams N, Norton N, Abraham R, Kehoe P, Williams H, Rudrasingham V, Rice F, Giles P, Tunstall N, Jones L, Lovestone S, Williams J, Owen MJ, Hardy J, Goate A. 2002. Full genome screen for Alzheimer disease: stage II analysis. Am J Med Genet 114:235–244.

Ott J. 1999. Analysis of human genetic linkage. Baltimore: Johns Hopkins University Press. 416p.

Pericak-Vance MA, Haines JL. 1995. Genetic susceptibility to Alzheimer disease. Trends Genet 11:504–508.

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 genomic screen in late-onset familial Alzheimer disease: evidence for a new locus on chromosome 12. JAMA 278:1237–1241.

Rimmler J, McDowell J, Slotterbeck BD, Haynes CS, Menold MM, Rogala A, Speer MC, Gilbert JR, Hauser ER, Vance JM, Pericak-Vance MA. 1998. Development of a data coordinating center (DCC): data quality control for complex disease studies. Am J Hum Genet 63:A240.

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.

Selkoe DJ. 2001. Clearing the brain's amyloid cobwebs. Neuron 32: 177–180.