Genome scan of age-at-onset in the NIMH Alzheimer disease sample uncovers multiple loci, along with evidence of both genetic and sample heterogeneity

Yoonha Choi[1], **Elizabeth E. Marchani**[2], **Thomas D. Bird**[2,3,4], **Ellen J. Steinbart**[3,4], **Deborah Blacker**[5], and **Ellen M. Wijsman**[1,2,6,*]

[1]Department of Biostatistics, University of Washington, Seattle, WA

[2]Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA

[3]Geriatric Research Education and Clinical Center, Veterans Affairs Puget Sound Health Care System, Seattle Division, Seattle, WA

[4]Department of Neurology, University of Washington, Seattle, WA

[5]Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School; Dept of Epidemiology, Harvard School of Public Health; Boston, MA

[6]Department of Genome Sciences, University of Washington, Seattle, WA

Abstract

Alzheimer's disease (AD) is a common neurodegenerative disorder of late life with a complex genetic basis. Although several genes are known to play a role in rare early-onset AD, only the *APOE* gene is known to have a high contribution to risk of the common late-onset form of the disease (LOAD, onset > 60 years). *APOE* genotypes vary in their AD risk as well as age-at-onset distributions, and it is likely that other loci will similarly affect AD age-at-onset. Here we present the first analysis of age-at-onset in the NIMH LOAD sample that allows for both a multilocus trait model and genetic heterogeneity among the contributing sites, while at the same time accommodating age censoring, effects of known genetic covariates, and full pedigree and marker information. The results provide evidence for genomic regions not previously implicated in this data set, including regions on chromosomes 7q, 15, and 19p. They also affirm evidence for loci on chromosomes 1q, 6p, 9q, 11, and, of course, the *APOE* locus on 19q, all of which have been reported previously in the same sample. The analyses failed to find evidence for linkage to chromosome 10 with inclusion of unaffected subjects and extended pedigrees. Several regions implicated in these analyses in the NIMH sample have been previously reported in genome scans of other AD samples. These results, therefore, provide independent confirmation of AD loci in family-based samples on chromosomes 1q, 7q, 19p, and suggest that further efforts towards identifying the underlying causal loci are warranted.

Keywords

MCMC; oligogenic; Bayesian; dementia; linkage analysis

---

[*]correspondence to Ellen M. Wijsman, Department of Medicine, Division of Medical Genetics, Box 357720, University of Washington, Seattle, WA 98195-7720. (206) 543-8987. wijsman@u.washington.edu.

INTRODUCTION

Alzheimer's disease (AD, MIM 104300) is a fatal and devastating neurodegenerative disorder. Late onset AD (LOAD), defined by the onset of symptoms after age 60 years, is the most prevalent form of the disease, affecting 5–10% of individuals ≥ 65 years of age and 10–45% of individuals > 85 years of age [Canadian Study of Health and Aging Working Group 1994; Evans et al., 1989; Fratiglioni et al., 1999]. Because of the late onset, delay of onset, even in the absence of prevention or cure, would provide substantial benefit to the patients and to society. It is therefore important to identify risk factors that influence age-at-onset of AD, as wellas disease risk.

There is ample evidence for a genetic basis to both risk [Akesson 1969; Bergem and Lannfelt 1997; Breitner et al., 1986; Lautenschlager et al., 1996; Meyer and Breitner 1998; Mohs et al., 1987] and age-at-onset of AD [Daw et al., 1999; Daw et al., 2000]. The existence of a genetic basis to AD was confirmed by identification of causal mutations resulting in early-onset AD in the amyloid precursor protein (*APP*) [Goate et al., 1991], presenilin 1 (*PSEN1*) [Sherrington et al., 1995] and presenilin 2 (*PSEN2*) [Levy-Lahad et al., 1995] genes, and by identification of *APOE* as the only well-established susceptibility gene for LOAD [Corder et al., 1993]. *APOE* also affects age-at-onset, with onset age decreasing with number of *ε4* alleles and increasing with number of *ε2* alleles [Corder et al., 1994; Corder et al., 1993]. Only 10–20% of the genetic variance for LOAD risk [Bennett et al., 1995; Slooter et al., 1998] and variance in age-at-onset [Daw et al., 1999; Daw et al., 2000] can be attributed to *APOE*, suggesting that additional LOAD loci remain to be discovered.

Genome scans using a variety of samples and designs have searched for additional contributing loci. Several regions have been implicated across multiple pedigree-based genome scans based on different samples [Bertram et al., 2005; Blacker et al., 2003; Farrer et al., 2003; Hahs et al., 2006; Lee et al., 2006; Li et al., 2006; Pericak-Vance et al., 2000; Rademakers et al., 2005; Scott et al., 2003; Wijsman et al., 2004]. Genomewide association studies (GWAS) have also recently been used, starting with moderate [Coon et al., 2007] and more recently large samples [Harold et al., 2009; Lambert et al., 2009; Seshadri et al., 2010] of typically unrelated subjects, although related subjects have also been used [Wijsman et al., 2011]. These studies have implicated additional loci, several of which give consistent results across studies of subjects of European descent, although not always in non-European samples [Bertram et al., 2007]. However, for these loci contribution to AD risk is much smaller than has been estimated for *APOE*. Other than *APOE*, which is detectable in both family-based linkage analyses [Holmans et al., 2005; Myers et al., 2002; Pericak-Vance et al., 2000; Sillen et al., 2008] and population-based association analyses [Abraham et al., 2008; Beecham et al., 2009; Coon et al., 2007; Grupe et al., 2007; Harold et al., 2009; Lambert et al., 2009; Wijsman et al., 2011], there is also only modest overlap between results from pedigree-based vs. GWAS designs. However, a meta-analysis of linkage-based genome scans identified regions on chromosomes 1q, 7p and 8p with strong evidence of linkage across studies, and with additional regions showing suggestive results across studies, including 6p and 19p [Butler et al., 2009].

The complex genetic basis of LOAD may explain the lack of strong consensus in regions implicated across studies. In addition to genetic heterogeneity and multilocus contributions to the phenotype, there is likely to be heterogeneity in exposure and response to environmental risk factors. Even *APOE* is likely to contribute differently to AD risk across samples because allele frequencies vary considerably among populations; e.g., there is a gradient of the high-risk *ε4* allele frequency ranging from a low of 5–9% in Mediterranean populations to over 20% in Scandinavian populations [Corbo and Scacchi 1999]. Some explicit allowance for heterogeneity has been introduced through use of stratification on

earlier vs. later onset of AD [Blacker et al., 2003], through use of a 2-locus risk model [Curtis et al., 2001], and through use of an oligogenic model with age-at-onset [Daw et al., 1999; Wijsman et al., 2005]. However, despite this complexity, most previous genome scan analyses of LOAD have been carried out under models that have limited accommodation for genetic heterogeneity, using methods that either ignore, or do not explicitly acknowledge heterogeneity.

A small number of studies have focused on age-at-onset of LOAD rather than disease risk. The well-established LOAD locus, *APOE*, appears to affect risk of AD through genotype-dependent age-at-onset [Breitner et al., 1998; Corder et al., 1994; Jarvik et al., 1995]. Analysis of age-at-onset in pedigree-based samples is complicated by limited analytic options to accommodate a complex mode-of-inheritance, censoring of age information in unaffected subjects, and applicability to general pedigrees. Previous searches for loci affecting age-at-onset in LOAD samples have made compromises, including ignoring unaffected subjects [Holmans et al., 2005], ignoring their age censoring [Lee et al., 2008], or ignoring pedigree information when adjusting for the known affects of *APOE* [Dickson et al., 2008].

Here we present the first analysis of age-at-onset in the NIMH LOAD sample that allows for both a multilocus trait model and genetic heterogeneity among the contributing sites within the sample, while at the same time accommodating age censoring and effects of known genetic covariates. This large sample represents the combined data from three different recruitment sites that differ with respect to both genetic ancestries of the local populations and overall recruitment procedures. Both of these issues are likely to cause heterogeneity in the sample. Our analyses detect differing underlying multilocus models across sites, supported by subsequent genome scans, providing evidence of this heterogeneity. Our results both affirm some regions identified previously in this sample, as well as identify additional regions with evidence for AD age-at-onset loci. Some of the new regions that we identify also have been reported in other samples, and therefore serve as confirmation of relevant AD loci.

METHODS

Data

We used the publicly-available National Institute of Mental Health (NIMH) Genetic Initiative Alzheimer disease sample (https://www.nimhgenetics.org/available_data/alzheimers_disease/). We used the data as provided, and we restricted all analyses to European Americans as defined by the "race=white" indicator. Subject collection and evaluation followed a standard protocol at three clinical sites (Centers 50–52: C50, C51, C52) as described elsewhere [Blacker et al., 1997]. The three sites differed in their geographic locations and points of contact for systematic recruitment efforts: recruiting through a memory disorders clinic, a private nonprofit geriatric outpatient clinic and nursing home, and through private hospital outpatient clinics. The three centers also differed in terms of the fraction of subjects recruited through systematic efforts as compared to non-systematic efforts such as referral from a clinician or Alzheimer's disease center, with 19%, 59%, and 69% of the sample representing subjects recruited by non-systematic procedures in C50, C51, and C52, respectively.

For the purpose of analysis, we defined affectation status as "affected" for individuals with definite, probable or possible AD based on NINCDS/ADRDA criteria [Blacker et al., 1997; McKhann et al., 1984], and as "unaffected" for individuals without dementia, or presumed to have no dementia. Among the affected subjects, autopsy documentation was available for

33%, 26%, and 35% in centers 50, 51 and 52, respectively. For other diagnosis groups, affectation status was considered to be unknown (missing). Age information used in the analysis was the age at which the first symptoms of AD were reported for affected individuals and the latest age at which an individual was observed for unaffected individuals, with all individuals with both age and diagnosis information used for analysis.

Genotyping is described elsewhere [Blacker et al., 2003; Blacker et al., 1997]. The markers consisted of 362 multiallelic markers spaced at ~10cM intervals. We restrict our presentation here to the genotypes as originally provided, since alternative options for compressing rare alleles into composite alleles for analysis did not affect the results (not shown). Marker map positions were obtained from the Rutgers map [Kong et al., 2004] and were converted to map positions based on the Haldane map function because of assumptions inherent in the underlying analysis methods.

Statistical Analysis

We used age-at-onset as the phenotype of interest, modeled as a quantitative trait [Daw et al., 1999; Heath 1997]. Age-at-onset was treated as censored for unaffected individuals and as observed (non-censored) for affected individuals. If either affection status or age information was missing, age-at-onset was treated as missing. We performed all analyses both jointly for the data from all three clinical sites, as well as separately for the three sites in order to investigate possible center effects. In addition, there were two versions of joint analysis of the complete sample: (1) assuming every individual is drawn from the same genetic population with the same underlying model parameters (All_1), and (2) assuming the three centers may differ in their genetic populations through differing allele frequencies of underlying trait loci (All_3).

The analysis approach used here differs in two important ways from alternatives such as variance-component methods used for analysis of censored age-at-onset data [Dickson et al., 2008]. First, original ages, rather than residuals from a pre-analysis, are used, and the censoring of age data in unaffected subjects is explicitly taken into account in the model. This, in turn, allows for more efficient use of *APOE* genotype information as a covariate [Wijsman and Yu 2004] because unobserved *APOE* genotypes can be probabilistically inferred from the existing data. Second, by allowing different allele frequencies for underlying quantitative trait loci (QTL), this approach can account for both the familial relatedness in pedigrees as well as effects such as center effects due to either ascertainment or population differences.

MCMC method

Here, we describe briefly the Bayesian oligogenic MCMC analyses used for most analyses and implemented in the program Loki version 2.4.7 [Daw et al., 1999; Heath 1997]. Age-at-onset, *Y*, is modeled as a censored quantitative trait. Multilocus trait genotype-specific means are assumed to be normally distributed with a common within-genotype residual variance, after adjusting for covariates. The model is:

$$Y = \mu + X\beta + \sum_{i=1}^{k} Q_i \alpha_i + e$$

where μ is an overall baseline, *X* is the incidence matrix for covariate effects, β is the vector of covariate effects, Q_i is the incidence matrix for the effect of diallelic QTL *i*, α_i is the vector of additive and dominant effects for QTL *i*, and *e* is the normally distributed residual

environmental effect. The model allows for multiple QTLs with the number of trait loci, *k*, considered as a random variable, not as a constant.

A Bayesian reversible-jump MCMC sampler [Green 1995] was used to estimate posterior distributions of model parameters conditional on the data and prior distributions on parameters. The parameters of interest are covariate effects, the number of QTLs contributing to the trait, the allele frequency and genotype effects for each QTL, and for linkage analysis, the location of each QTL. The reported allele frequency, p_A, is the frequency of allele A, and homozygote (BB) and heterozygote (AB) effects are μ_{BB}-μ_{AA} and μ_{AB}-μ_{AA}, respectively, where μ_{ij} is the mean genotypic value for genotype *ij*. For each QTL, the allele labels A and B were arbitrarily assigned such that the genotype BB effect was non-negative.

Segregation Analysis

Oligogenic segregation analysis was performed to determine the complexity of the underlying mode of inheritance. We identified models for QTLs associated with age-at-onset for the complete sample (All_1 and All_3), and also for each of three centers (C50, C51, C52) separately. Segregation analysis describes the underlying QTL model(s) without use of markers or trait locus positions, with the complexity of the oligogenic trait model serving as a useful indicator of the genetic basis of the trait for further linkage analysis [Igo et al., 2006]. All segregation analyses were performed both without and with *APOE* as a covariate, and were based on 100,000 MCMC iterations, with every 10th iteration used for computation of posterior distributions. We used a uniform distribution between 0 and 1 for QTL allele frequencies. The prior distribution on the number of QTLs was assumed to be Poisson with mean 4. The trait locus effects were assumed to have a normally distributed prior with mean 0 and variance 512, which is about four times the phenotypic variance [Wijsman and Yu 2004].

To investigate whether certain QTL models are derived by a few individuals with the most extreme phenotypic values, we performed additional segregation analyses that excluded unaffected individuals with age < 46 or > 91 years. The values of 46 and 91 are based on minimum and maximum of ages of individuals in C52, which had a narrower range of age data in unaffected subjects than did the other two centers, and also had a less complex posterior distribution in the QTL model space. QTL models were compared first, without the young unaffected individuals (age < 46); second, without the old unaffected individuals (age > 91); and third, without both the young and old unaffected individuals for C50, C51 and All_3.

Linkage Analysis

We performed linkage analysis for All_1, C50, C51, C52, and All_3 as in the segregation analysis, carrying out analysis without and with including *APOE* both as a covariate and as a marker. A sex-averaged meiotic map was used with the assumption of a complete genome length of 3,000cM, with map positions based on the Haldane map function. All linkage analyses were based on 500,000 MCMC iterations, with every 10th iteration used for inference. We used a uniform distribution for the prior distribution on QTL locations. The prior distributions on other parameters were the same as the segregation analysis. Since we used a Bayesian approach, which does not result in conventional LOD scores or P-values, we report log (base 10) of the Bayes' factor (BF) for every 2 cM bin, logBF, which is the ratio of posterior to prior odds of linkage. We use the maximum logBF, $log(BF_{max})$ in regions with high values as a test of linkage.

Calibration of logBF as supporting evidence of linkage is approximate. However, log10 BF>1.5 corresponds to strong evidence in favor of the alternative hypothesis of linkage [Kass and Raftery 1995]. In addition, in our experience, most cases with log BF>1.5 appear to be equivalent to p<0.001, calibrated either by reference to a different calibrated linkage statistic, or by carrying out computationally expensive simulations [Igo and Wijsman 2008].

Chromosome 10

We performed additional linkage analyses on chromosome 10 using affected sibling pairs (ASPs) with onset ages ≥ 65. Previous studies of part [Holmans et al., 2005; Myers et al., 2000] or all of this sample [Myers et al., 2002] reported strong evidence for linkage to Chromosome 10 based on allele sharing in ASPs, but we did not find such evidence using MCMC linkage analysis of age-at-onset in the complete sample, which also included additional relatives (see results). This analysis was, therefore, performed to determine whether the discrepancy was explained by differences in the sample included in the analysis or by difference in the analytic approach. To make our analysis comparable with the previous results, we used 304 ASPs in 303 pedigrees (1289 total individuals) and ignored unaffected siblings.

We carried out standard non-parametric linkage analysis with the Kong and Cox [Kong and Cox 1997] linear model using Merlin [Abecasis et al., 2002], and report multipoint maximum LOD scores. The Kong and Cox LOD score is maximized on a single parameter, δ, representing the degree of allele sharing among affected individuals. Under the null hypothesis, $\delta=0$, and higher δ corresponds to the alternative of excess allele sharing. For comparison, MCMC linkage analysis was performed on the same sample (N=1289), and the BF was computed with and without adjustment for *APOE* effects.

Center-Specific Validation

Center-specific analyses identified evidence of linkage to chromosomes 1 and 7 that appeared to be consistent with linkage signals in a recent genome scan for modifiers of age-at-onset among early-onset AD families with a known *PSEN2* mutation [Marchani et al., 2010]. To determine whether these signals were centered over the same regions, we performed a joint analysis including both the large early-onset (EO) family from this previous analysis with the greatest contribution to the published linkage signals, and either the individual NIMH centers supporting linkage to these regions, or all NIMH centers. In these analyses, we allowed allele frequencies to vary between each center and the early-onset family. We merged the marker maps from the two studies, and different markers with identical positions based on a Haldane map function were jittered to be 0.1cM apart to avoid analytical constraints. Analyses included *APOE* as a covariate, and were based upon 2,000,000 MCMC iterations after a 1,000 iteration burn-in, saving every 50th iteration for posterior parameter estimation.

RESULTS

Descriptive analysis

There was evidence for heterogeneity among the samples from the three NIMH centers. The centers differed in terms of family size and phenotype distributions, as well as in the *APOE* allele frequencies (Table 1), which is well known to have a significant effect on age-at-onset [Corder et al., 1993]. Center 50 had both the largest pedigrees and the highest missing data rate. Although mean ages were similar among the centers for both affected and unaffected individuals, there were fewer extreme high and low values in Center 52 than in the other centers: Centers 50 and 51 included a few young individuals with age below 40, including one affected and eight unaffected individuals for Center 50, and six unaffected individuals

for Center 51; Center 52 had no such individuals. Similarly, *APOE* frequencies varied considerably, with Center 51 representing the most extreme values among the three centers for both the *ε3* and *ε4* alleles.

Segregation Analysis

We found three main QTL models affecting age-at-onset of AD, and defined them based on their homozygote (BB) and heterozygote (AB) effects (Figure 1, Table 2). Two of the three models were obtained when all three centers were analyzed, both jointly as one population (All$_1$) (Figure 1A, B), and when all three centers were permitted center-specific QTL allele frequencies (All$_3$) (Figure 1N, O). Model 1 is slightly over-dominant, in which the trait values for the AB genotype were lower than those for either homozygote. The mean genotype effects for the AB and BB genotypes of Model 1 in analysis All$_3$ were −7.53 and 34.37 years without *APOE* adjustment and −6.68 and 33.22 years with *APOE* adjustment, respectively. In Model 2, the heterozygote effect was higher than those of either homozygote except in the analysis of C50 alone: the mean age-at-onset for the AB and BB genotypes in analysis All$_3$ were 9.93 and 3.13 years greater than that of the AA genotype without *APOE* adjustment, and 7.89 and 2.48 years greater with *APOE* adjustment, respectively. Model 2 was the least well-defined of the three main models, and may represent multiple smaller models that improve overall model fit, thus leading to overall larger confidence intervals upon model extraction. Model 3 is close to a dominant model, with similar effects of the AB and BB genotypes: the mean effects of the AB and BB genotypes in analysis All$_3$ were 22.43 and 22.65 years without *APOE* adjustment and 20.21 and 21.11 years with *APOE* adjustment, respectively.

Segregation analyses performed separately for each center gave different results, suggesting the existence of among-center heterogeneity. The three centers contribute to different QTL models, so the model spaces shifted slightly depending on which center was used in the analysis. In the analysis of C50 alone, all three models were identified. The models were similar for analysis both with and without *APOE* adjustment (Figure 1D, F), but with smaller parameter standard deviations with *APOE* adjustment. In the analysis of C51 alone, Models 2 and 3 had greater overlap (Figure 1H, I): with *APOE* adjustment, there was no model equivalent to Model 2 and again, parameter standard deviation were reduced relative to the analysis without *APOE* adjustment. For C52, the QTL model spaces were the most different from the other two centers (Figure 1K, L). Notably, there was no model equivalent to Model 1, the positions of Models 2 and 3 were also quite different from those of other centers, and there was less evidence for clarification of the QTL models with adjustment for *APOE*. These differences in the QTL models suggested that center effects should not be ignored in the linkage analysis.

The three QTL models were clearly separated in analysis of the full sample that allowed for different allele frequencies among the three centers (All$_3$) (Figure 1N, O). In Model 1, allele frequencies of allele A for C50 and C51 were similar to each other (Table 2), but C52 had a much higher A allele frequency (rare minor allele frequency), which is consistent with the absence of a QTL model equivalent to Model 1 in the analysis of C52 only. C52 had a different allele frequency for Model 2 as well. For Model 3, allele frequencies were similar for all of three centers.

Overall, adjustment for *APOE* did not substantially alter QTL models, but did improve the precision of their estimates. The mean parameter values based on the posterior parameter distributions obtained for QTLs with *APOE* adjustment had lower standard deviations than without *APOE* adjustment. In addition, the number of QTLs in each model decreased under all analysis models that included an *APOE* adjustment (Figure 1, column 3).

Additional segregation analyses showed that Model 1 is driven by the oldest unaffected individuals (age > 91 years). This model disappeared or weakened when unaffected old individuals (N=6 in C50, N=9 in C51) were excluded (Supplementary Figure 1). For example, with *APOE* adjustment, removal of the oldest unaffected subjects lead to the disappearance of model 1 for C51, and a reduction of about 10 years for the homozygote effect for All$_3$ and C50. The other QTL models (2 and 3) were essentially unaffected by exclusion of the outliers, and excluding young unaffected individuals similarly had little effect on QTL models (not shown). One possible reason for the absence of Model 1 for C52 in the segregation analysis may be that C52 includes no extremely old, unaffected individuals.

Linkage Analysis

Our genomewide linkage scan found several regions with strong to moderate evidence of linkage with age-at-onset of LOAD. Figure 2 presents the results of all chromosomes across analysis conditions, including the all-centers analyses, All$_1$ and All$_3$. In the analysis of the complete data set, we detected two strong linkage signals on chromosome 6 and 19 (Figure 2A, E). We identified moderate signals on chromosome 1 (Figure 2B, D) and 7 (Figure 2C, D) both supported by at least two centers. We also identified two strong signals on chromosomes 9 and 15 each, in only one specific center (Figure 2C, D). As discussed further below, results without or with including *APOE* as a covariate and a marker generally differed, but in most cases, the differences were modest.

Chromosome 19—As expected, strong evidence of linkage was obtained near the *APOE* locus on 19q when *APOE* was not included as either a marker or covariate (Figure 3A). This was particularly striking when the complete sample was included in the analysis (logBF$_{max}$=1.62 and 1.50 for All$_1$ and All$_3$ respectively, at 71–73cM). The analysis of each of C51 and C52 gave moderate evidence for linkage to the location of *APOE*, but with the position of the maximum BF located at slightly different positions (logBF$_{max}$=1.05, 71cM for C51; logBF$_{max}$=0.85, 97cM for C52). However, the evidence for linkage to *APOE* was weak when the analysis was performed for C50 alone (logBF$_{max}$=−0.28, 71cM). Consistent with an interpretation of *APOE* as the cause of the strong evidence for linkage at ~70cM, evidence for linkage to 19q dropped considerably when *APOE* was included as a covariate and a marker (Figure 3B). For all the analyses except that of C50 alone, which had little evidence for linkage to this region, the logBF$_{max}$ near *APOE* was decreased with *APOE* adjustment (logBF$_{max}$=0.58 and 0.78 for All$_1$ and All$_3$ respectively, at 71–73cM; logBF$_{max}$=0.61, 71cM for C51; logBF$_{max}$= 0.28, 97cM for C52).

We also detected moderate evidence for linkage to 19p. In the analysis without *APOE* adjustment, there was only weak evidence for linkage to 19p. This signal strengthened considerably in the analysis that included *APOE* both as a covariate and as a marker for analysis of All$_1$ (logBF$_{max}$=0.99, 21cM for All$_1$) with weaker but similar results for All$_3$ and C52 (Figure 3B).

To investigate the effect of *APOE* adjustment on QTL models, we compared the QTL models near *APOE* for analysis with and without adjustment for *APOE*. Figure 3C and D shows the QTL models accepted in the MCMC analysis of All$_3$ with locations in the *APOE* region from 55cM to 95cM on chromosome 19 in the analysis. Without adjustment for *APOE*, we found a QTL model with the mean trait effects of μ_{AB} -μ_{AA} = −8 and μ_{BB} -μ_{AA} = 30, which is equivalent to Model 1 in the segregation analysis. However this QTL model disappeared when *APOE* was included both as a covariate and as a marker, suggesting that Model 1 represents the effect of *APOE*. In addition, adjustment for *APOE* resulted in merging of two pairs of models that represent Model 2 and Model 3. This is what would be

expected from the analysis model, which models genotype effects as deviation from a baseline overall mean: in the absence of an *APOE* adjustment, different *APOE* genotypes would be represented in the baseline as sampled in the MCMC process, leading to multiple similar models with an effect, as seen here.

Chromosome 6—Other than chromosome 19q, chromosome 6p exhibited the strongest genome wide evidence of linkage in the analysis of the complete data (Figure 4, $logBF_{max}$=1.63 and 1.23 for All_1 and All_3 respectively, at 71cM). Two of the three centers, C51 and C52, individually also provided individual support for linkage to this region both with and without adjustment for *APOE*, with the third center providing support for linkage in the absence of *APOE* adjustment. To the extent that the BF provides localization, results from the different centers are all consistent with a QTL at the same location, since the logBF maximizes at the same position in analysis of All_1 and All_3 as well as the data from individual centers. These results are similar to those obtained for chromosome 19q near *APOE*, which was also supported most strongly by C51 and C52. This signal on chromosome 6 decreased considerably when *APOE* was included as a covariate and a marker in the linkage analysis (Figure 4B vs. E).

Chromosome 7—There was moderate support for linkage to chromosome 7q across centers, with strong support in one family. Without *APOE* adjustment, one moderate signal was detected in the analysis of C51 (Figure 4C, $logBF_{max}$ =1.19, 167cM). This signal became much stronger with *APOE* adjustment ($logBF_{max}$= 1.67, 167cM for C51) with supporting evidence at the same region in the analysis of C52 (Figure 4F). In the analysis of All_1 and All_3, we found linkage signals at ~167cM, but these were not as strong as in the analysis of C51. The logBF was essentially coincident in the analysis of all vs. only individuals centers, consistent with a single contributing locus in the separate samples. As for chromosome 19, when the analysis was performed for C50 only, the evidence for linkage was very weak on the entire chromosome without or with *APOE* adjustment.

Chromosome 1—There was moderately strong evidence for linkage to chromosome 1 (Figure 4A, D). This evidence was strongest for the individual center C52 ($logBF_{max}$=1.08, 169cM) without *APOE* adjustment (Figure 4A). There was supporting evidence for linkage from C50 ($logBF_{max}$=0.68, 187cM), but this signal dissipated when *APOE* was included as a covariate (Figure 4D). In analyses of C51, with or without *APOE* adjustment, there was no evidence of linkage to chromosome 1.

Chromosomes 9 and 15—Regions with evidence for linkage on chromosome 9 and 15 were obtained in single centers only (Figure 2). For both chromosomes, results were insensitive to *APOE* adjustment, although in both cases *APOE* adjustment weakened signals slightly. Although there was weak evidence for linkage at ~133 cM on chromosome 9 in the analysis of All_3, strong evidence for linkage was obtained for this region for C52, $logBF_{max}$= 1.69 and 1.44 without and with *APOE* adjustment, respectively (Supplementary Figure 4A). Similarly strong evidence for linkage at ~41cM on chromosome 15 was detected for C51 ($logBF_{max}$= 1.34 and 1.11 without and with *APOE* adjustment, respectively, Supplementary Figure 5A). Presumably, the weaker evidence of linkage to both these regions in the complete data (All_3) is explained by the lack of support from the other two centers.

Similar to the model space observed on chromosome 19, the QTL models on chromosome 9, and to a lesser extent on chromosome 15, were affected by *APOE* adjustment. On chromosome 9 there were parallel distributions of models evident in the absence of adjustment for *APOE* (Supplementary Figure 4B). These models combined into a single model after adjustment for *APOE* (Supplementary Figure 4C). Similar, but weaker, evidence

for QTL models was obtained for chromosome 15, and again parallel models merged when *APOE* was included as a covariate (Supplementary Figure 5B, C). The most probable explanation for this merging of models with *APOE* adjustment is again the structure of the model as genotypic effects relative to a baseline, with multiple possible *APOE* baseline genotypes.

Chromosome 10—Evidence for linkage of age-at-onset to chromosome 10 was very weak. MCMC and non-parametric linkage analysis of 303 ASPs suggests that evidence for linkage to chromosome 10 is specific to use of ASPs, not to use of the more extended families and unaffected subjects. In MCMC linkage analysis of the ASPs, we found very weak support for linkage to chromosome 10 with a maximum logBF=0.45 (BF=2.81) obtained at 109 cM, and centered over the same position as achieved a maximum LOD score. When all subjects were included, evidence for linkage dropped considerably: in the analysis of All$_3$, the evidence for linkage to the region ~109 cM dropped to log BF=−0.52 (BF = 0.3). In the non-parametric linkage analysis of 303 ASPs, we found suggestive evidence of linkage (LOD =1.57, 107.88 cM), which is similar to the MCMC linkage analysis of 303 ASPs, and similar to the analysis presented originally on the complete sample [Blacker et al., 2003] including non-European Americans. This result suggests that absence of evidence for linkage of age-at-onset to Chromosome 10 is explained by use of all European Americans in our analysis, instead of only ASPs with onset ages ≥ 65.

Center-Specific Validation

Evidence of linkage of age-at-onset to chromosomes 1 and 7 strengthened with the inclusion of an early onset AD family with a known *PSEN2* mutation (Table 3), with evidence of age-at-onset modifier loci on these chromosomes [Marchani et al., 2010]. The stronger linkage signal was detected at the same region of chromosome 7 for the analysis of C51 (logBF$_{max}$=1.43 and 1.91 without and with *APOE* adjustment, respectively at 167cM). On chromosome 1, the linkage signal in the analysis of C52 was also strengthened without (logBF$_{max}$=1.10, 170cM) and with (logBF$_{max}$ = 1.21, 195cM) *APOE* adjustment, respectively, but the regions were slightly shifted compared to those identified in the analysis of NIMH data alone, especially when *APOE* was included.

DISCUSSION

Our results provide evidence for multiple loci affecting age-at-onset of AD in the complete European-American NIMH sample, with heterogeneity across data collection centers. For this data set, evidence obtained here for linkage to the regions identified on chromosomes 7q, 15, and 19p is new. Evidence for loci on chromosomes 1q, 6p, 9q, 11, and 19q has been reported previously with an AD risk model and/or an age-at-onset model [Blacker et al., 2003; Dickson et al., 2008; Holmans et al., 2005]. The approach to handling age-at-onset information used here for the first time allowed determination of the center(s) that contribute to individual linkage signals. This demonstrated that evidence for linkage of age-at-onset to chromosomes 6p, 7q, 11 and 19q was supported by more than one center, while evidence for linkage to chromosomes 1q, 9q, 15, and 19p was obtained primarily within individual centers. These results also demonstrate the advantage of analysis of this sample with a more sophisticated and complex model than has been previously used.

Differences between center-specific and global results support heterogeneity across centers. This is supported by differences in both sample composition and in results from the genome scans. In the multi-center NIMH sample, sample heterogeneity that contributes to genetic heterogeneity may be the consequence of variation in sample ascertainment as well as geographically correlated genetic ancestry differences across centers. Even though inclusion

criteria were similar across the collection sites, de-facto ascertainment differences are suggested by differences in the resulting sample characteristics, including pedigree sizes and age ranges. This could inadvertently have an affect on the resulting genetic architecture of the samples. Evidence for differences in genetic background is suggested by differences in *APOE* allele frequencies among centers, as well as by differences in the QTL models obtained by oligogenic segregation analysis. Such genetic heterogeneity may be inevitable in any sample collected from geographically distinct regions. For example, two of the centers that contributed subjects to the NIMH sample used here also contributed subjects to the NIA LOAD sample, with a recent analysis identifying substantially different fractions of subjects with southern-European and Ashkenazi Jewish ancestral origins between those two centers [Wijsman et al., 2011]. While heterogeneity may be impossible to eliminate in large samples, knowledge of which data components contribute to evidence of linkage in any specific genomic region is useful for further studies leading to gene identification.

Other than the *APOE* region, the regions implicated in this sample differ from those identified by association-based approaches [Bertram et al., 2007]. This is not surprising, since factors that affect power to detect trait loci differ for linkage-based vs. population-based designs. For example, the strength of evidence for linkage near *APOE* in the individual NIMH centers is inversely related to the *APOE ε4* allele frequency in the centers, as would be expected by the known deleterious effect of high allele frequencies on power to detect evidence of linkage [Risch and Merikangas 1996]. In contrast, association-based approaches do poorly in the presence of rare alleles – the situation that favors linkage-based designs. The linkage-based design is therefore more likely than the population-based design to lead to identification of loci with rare risk alleles that have large differences in genotype-specific risks. Identification of such loci is also likely to translate into easier subsequent mechanistic studies than would use of genotypes with the very small differences in risks currently being identified in population-based studies. Possible approaches for identifying loci in regions with evidence of linkage require dense genotyping in the pedigrees with evidence for linkage, as discussed elsewhere [Blangero 2009; Blangero et al., 2009; Wijsman et al., 2010]. The complementary strengths of different designs support the importance of tackling the genetic basis of AD with multiple approaches, since it is unlikely that any single approach will identify all of the loci underlying genetic variance in age-at-onset or disease risk.

The results obtained here provide replication and additional support for several previously implicated regions from other AD samples. The strong evidence for linkage on chromosome 7q reported here overlaps a previously reported linkage signal in the analysis of age-at-onset modifiers of early onset AD [Marchani et al., 2010], age-at-onset in a Caribbean Hispanic sample [Lee et al., 2008], and AD status in a Dutch sample [Rademakers et al., 2005]. A candidate gene in this region is NOS3 [Bertram et al., 2007]. Similarly, evidence for linkage of age-of-onset of AD to chromosome 1, previously also reported in the NIMH sample in analysis of AD status [Blacker et al., 2003; Kehoe et al., 1999] and AD age-at-onset [Dickson et al., 2008], also has been reported in the analysis of age-at-onset in a Caribbean Hispanic sample [Lee et al., 2008], as an age-at-onset modifier in early onset AD [Marchani et al., 2010], and in analysis of AD status in other independent samples [Hiltunen et al., 2001; Liu et al., 2007]. This region on chromosome 1 is notable for the location of nicastrin, which is a component of γ-secretase along with presenilin 1 and additional subunits [Kimberly et al., 2003], and is also one of the three regions with strongest support in a meta-analysis of genome scans for AD-risk susceptibility [Butler et al., 2009]. Regions identified on chromosomes 6 and 15 contain candidate genes VEGF and ADAM10, respectively [Bertram et al., 2007]. The region on chromosome 19p identified here for the first time was previously reported in two different samples in an analysis of dementia risk [Hahs et al., 2006] and of age-at-onset [Wijsman et al., 2004], with similar strengthening of evidence

with adjustment for *APOE* as seen here. This region also was highlighted in meta-analysis [Butler et al., 2009], and contains several candidate genes, including LDLR [Bertram et al., 2007]. The region on chromosome 19q with strong linkage evidence is, of course, expected because it contains *APOE*, the strongest known LOAD genetic risk factor [Corder et al., 1993]. In the absence of adjustment for *APOE*, the variation among with respect to the position of the maximum BF for analysis of the data from individual centers may simply reflect the combination of a diallelic QTL model used for analysis of an effect governed by the triallelic *APOE* locus [Rosenthal and Wijsman 2010]. For the 19q region, this hypothesis is supported by the strong reduction in evidence for linkage to this region upon adjustment for *APOE*. However, for other regions with some variability in the location of the maximum BF, such as on chromosome 1, more data would be needed to begin to distinguish between hypotheses of multiallelic QTLs, multiple loci, or simple statistical variability in location estimates.

Identification of several regions that appear to replicate results in other, independent, samples is encouraging, as is the evidence that such regions can be detected with either age-at-onset or disease risk as an endpoint phenotype. The genomic regions in the NIMH sample that may be most amenable to further investigation are such regions that coincide with those in genome scans of other samples. These regions may contain relatively rare variants of interest, since this architecture best explains the replication across only a subset of samples and is more readily detected in a family-based than population-based design. Fortunately, rapidly developing technologies are increasingly making it possible to search directly for such variants, rather than having to depend on indirect association with tag-SNPs. These newer technologies, coupled with the existence of multiple samples that support particular regions, are likely to facilitate identification of the causal sites in future studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nature Genetics. 2002; 30:97–101. [PubMed: 11731797]

Abraham R, Moskvina V, Sims R, Hollingworth P, Morgan A, Georgieva L, Dowzell K, Cichon S, Hillmer AM, O'Donovan JC, Williams J, Owen JJ, Kirov G. A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. BMC Medical Genomics. 2008; 1:44. [PubMed: 18823527]

Akesson HO. A Population Study of Senile and Arteriosclerotic Psychoses. Human Heredity. 1969; 19(5):546. [PubMed: 5365894]

Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, Pericak-Vance MA. Genome-wide Association Study Implicates a Chromosome 12 Risk Locus for Late-Onset Alzheimer Disease. American Journal of Human Genetics. 2009; 84(1):35–43. [PubMed: 19118814]

Bennett CL, Crawford F, Osborne A, Diaz P, Hoyne J, Lopez R, Roques P, Duara R, Rossor M, Mullan MJ. Evidence that the APOE locus influences rate of disease progression in late-onset

familial Alzheimer's-disease but is not causative. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 1995; 50:1–6.

Bergem ALM, Lannfelt L. Apolipoprotein E type epsilon 4 allele, heritability and age at onset in twins with Alzheimer's disease. Clinical Genetics. 1997; 52(5):408–413. [PubMed: 9520135]

Bertram L, Hiltunen M, Parkinson M, Ingelsson M, Lange C, Ramasamy K, Mullin K, Menon R, Sampson AJ, Hsiao MY, Elliott KJ, Velicelebi G, Moscarillo T, Hyman BT, Wagner SL, Becker KD, Blacker D, Tanzi RE. Family-based association between Alzheimer's disease and variants in UBQLN1. New England Journal of Medicine. 2005; 352(9):884–894. [PubMed: 15745979]

Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nature Genetics. 2007; 39(1):17–23. [PubMed: 17192785]

Blacker D, Bertram L, Saunders AJ, Moscarillo TJ, Albert MS, Wiener H, Perry RT, Collins JS, Harrell LE, Go RCP, Mahoney A, Beaty T, Fallin MD, Avramopoulos D, Chase GA, Folstein MF, McInnis MG, Bassett SS, Doheny KJ, Pugh EW, Tanzi RE. Results of a high-resolution genome screen of 437 Alzheimer's Disease families. Human Molecular Genetics. 2003; 12(1):23–32. [PubMed: 12490529]

Blacker D, Haines JL, Rodes L, Terwedow H, Go RC, Harrell LE, Perry RT, Bassett SS, Chase G, Meyers D, Albert MS, Tanzi R. ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. Neurology. 1997; 48(1):139–47. [PubMed: 9008509]

Blangero J. Update to Blangero et al.'s "Quantitative Trait Nucleotide Analysis Using Bayesian Model Selection" (2005): From QTL Localization to Functional Variant Identification. Human Biology. 2009; 81(5–6):849–852. [PubMed: 20504201]

Blangero J, Goring HHH, Kent JW, Williams JT, Peterson CP, Almasy L, Dyer TD. Quantitative Trait Nucleotide Analysis Using Bayesian Model Selection. Human Biology. 2009; 81(5–6):829–847. [PubMed: 20504200]

Breitner JCS, Folstein MF, Murphy EA. Familial Aggregation in Alzheimer Dementia.1. a Model for the Age-Dependent Expression of an Autosomal Dominant Gene. Journal of Psychiatric Research. 1986; 20(1):31–43. [PubMed: 3712289]

Breitner JCS, Jarvik GP, Plassman BL, Saunders AM, Welsh KA. Risk of Alzheimer disease with the epsilon 4 allele for apolipoprotein E in a population-based study of men aged 62–73 years. Alzheimer Disease & Associated Disorders. 1998; 12:40–44. [PubMed: 9539409]

Butler AW, Ng MYM, Hamshere ML, Forabosco P, Wroe R, Al-Chalabi A, Lewis CM, Powell JF. Meta-analysis of linkage studies for Alzheimer's disease-A web resource. Neurobiology of Aging. 2009; 30(7):1037–1047. [PubMed: 19362756]

Canadian Study of Health and Aging Working Group. Canadian study of health and aging: study methods and prevalence of dementia. Canadian Medical Association Journal. 1994; 150(6):899–913. [PubMed: 8131123]

Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Lince DH, Zismann VL, Beach TG, Leung D, Bryden L, Halperin RF, Marlowe L, Kaleem M, Walker DG, Ravid R, Heward CB, Rogers J, Papassotiropoulos A, Reiman EM, Hardy J, Stephan DA. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. Journal of Clinical Psychiatry. 2007; 68(4):613–618. [PubMed: 17474819]

Corbo RM, Scacchi R. Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a 'thrifty' allele? Annals of Human Genetics. 1999; 63:301–310. [PubMed: 10738542]

Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, Rimmler JB, Locke PA, Conneally PM, Schmader KE, Small GW, Roses AD, Haines JL, Pericak Vance MA. Protective Effect of Apolipoprotein-E Type-2 Allele for Late-Onset Alzheimer-Disease. Nature Genetics. 1994; 7(2):180–184. [PubMed: 7920638]

Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993; 261(5123):921–3. [PubMed: 8346443]

Curtis D, North BV, Sham PC. A novel method of two-locus linkage analysis applied to a genome scan for late onset Alzheimer's disease. Annals of Human Genetics. 2001; 65:473–482. [PubMed: 11806855]

Daw EW, Heath SC, Wijsman EM. Multipoint oligogenic analysis of age-at-onset data with applications to Alzheimer's disease pedigrees. American Journal of Human Genetics. 1999; 64(3): 839–851. [PubMed: 10053019]

Daw EW, Payami H, Nemens EJ, Nochlin D, Bird TD, Schellenberg GD, Wijsman EM. The number of trait loci in late-onset Alzheimer disease. American Journal of Human Genetics. 2000; 66:196–204. [PubMed: 10631151]

Dickson MR, Li J, Wiener HW, Perry RT, Blacker D, Bassett SS, Go RCP. A genomic scan for age at onset of Alzheimer's disease in 437 families from the NIMH genetic initiative. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2008; 147B(6):784–792.

Evans DA, Funkenstein HH, Albert MS, Scherr PA, Cook NR, Chown MJ, Hebert LE, Hennekens CH, Taylor JO. Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported. JAMA. 1989; 262(18):2551–6. [PubMed: 2810583]

Farrer LA, Bowirrat A, Friedland RP, Waraska K, Korczyn AD, Baldwin CT. Identification of multiple loci for Alzheimer disease in a consanguineous Israeli-Arab community. Human Molecular Genetics. 2003; 12(4):415–22. [PubMed: 12566388]

Fratiglioni L, De Ronchi D, Aguero-Torres H. Worldwide prevalence and incidence of dementia. Drugs & Aging. 1999; 15(5):365–375. [PubMed: 10600044]

Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, Mant R, Newton P, Rooke K, Roques P, Talbot C, Pericak-Vance M, Roses A, Williamson R, Rossor M, Owen M, Hardy J. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991; 349:704–706. [PubMed: 1671712]

Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika. 1995; 82:711–732.

Grupe A, Abraham R, Li Y, Rowland C, Hollingworth P, Morgan A, Jehu L, Segurado R, Stone D, Schadt E, Karnoub M, Nowotny P, Tacey K, Catanese J, Sninsky J, Brayne C, Rubinsztein D, Gill M, Lawlor B, Lovestone S, Holmans P, O'Donovan M, Morris JC, Thal L, Goate A, Owen MJ, Williams J. Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. Human Molecular Genetics. 2007; 16(8):865–873. [PubMed: 17317784]

Hahs DW, McCauley JL, Crunk AE, McFarland LL, Gaskell PC, Jiang L, Slifer SH, Vance JM, Scott WK, Welsh-Bohmer KA, Johnson SR, Jackson CE, Pericak-Vance MA, Haines JL. A genome-wide linkage analysis of dementia in the Amish. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2006; 141B(2):160–166.

Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nature Genetics. 2009; 62:1088–1093. [PubMed: 19734902]

Heath S. Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. American Journal of Human Genetics. 1997; 61:748–760. [PubMed: 9326339]

Hiltunen M, Mannermaa A, Thompson D, Easton D, Pirskanen M, Helisalmi S, Koivisto AM, Lehtovirta M, Ryynanen M, Soininen H. Genome-wide linkage disequilibrium mapping of late-onset Alzheimer's disease in Finland. Neurology. 2001; 57(9):1663–1668. [PubMed: 11706108]

Holmans P, Hamshere M, Hollingworth P, Rice F, Tunstall N, Jones S, Moore P, DeVrieze FW, Myers A, Crook R, Compton D, Marshall H, Meyer D, Shears S, Booth J, Ramic D, Williams N, Norton N, Abraham R, Kehoe P, Williams H, Rudrasingham V, O'Donovan M, Jones L, Hardy J, Goate

A, Lovestone S, Owen M, Williams J. Genome screen for loci influencing age at onset and rate of decline in late onset Alzheimer's disease. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2005; 135B(1):24–32.

Igo RP, Chapman NH, Wijsman EM. Segregation analysis of a complex quantitative trait: Approaches for identifying influential data points. Human Heredity. 2006; 61:80–86. [PubMed: 16679774]

Igo RP, Wijsman EM. Empirical significance values for linkage analysis: Trait simulation using posterior model distributions from MCMC oligogenic segregation analysis. Genetic Epidemiology. 2008; 32(2):119–131. [PubMed: 17849492]

Jarvik GP, Wijsman EM, Kukull WA, Schellenberg GD, Yu C, Larson EB. Interactions of apolipoprotein E genotype, total cholesterol level, age, and sex in prediction of Alzheimer's disease: a case-control study. Neurology. 1995; 45:1092–1096. [PubMed: 7783869]

Kass RE, Raftery AE. Bayes Factors. Journal of the American Statistical Association. 1995; 90(430): 773–795.

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. A full genome scan for late onset Alzheimer's disease. Human Molecular Genetics. 1999; 8(2):237–245. [PubMed: 9931331]

Kimberly WT, LaVoie MJ, Ostaszewski BL, Ye WJ, Wolfe MS, Selkoe DJ. gamma-secretase is a membrane protein complex comprised of presenilin, nicastrin, aph-1, and pen-2. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100(11):6382–6387. [PubMed: 12740439]

Kong A, Cox NJ. Allele-sharing models: scores and accurate linkage tests. American Journal of Human Genetics. 1997; 61:1179–1188. [PubMed: 9345087]

Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. American Journal of Human Genetics. 2004; 75:1143–1148. [PubMed: 15486828]

Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, Letenneur L, Bettens K, Berr C, Pasquier F, Fievet N, Barberger-Gateau P, Engelborghs S, De Deyn P, Mateo I, Franck A, Helisalmi S, Porcellini E, Hanon O, de Pancorbo MM, Lendon C, Dufouil C, Jaillard C, Leveillard T, Alvarez V, Bosco P, Mancuso M, Panza F, Nacmias B, Bossu P, Piccardi P, Annoni G, Seripa D, Galimberti D, Hannequin D, Licastro F, Soininen H, Ritchie K, Blanche H, Dartigues JF, Tzourio C, Gut I, Van Broeckhoven C, Alperovitch A, Lathrop M, Amouyel P. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nature Genetics. 2009; 41:1094–1099. [PubMed: 19734903]

Lautenschlager NT, Cupples LA, Rao VS, Auerbach SA, Becker R, Burke J, Chui H, Duara R, Foley EJ, Glatt SL, Green RC, Jones R, Karlinsky H, Kukull WA, Kurz A, Larson EB, Martelli K, Sadovnick AD, Volicer L, Waring SC, Growdon JH, Farrer LA. Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: What is in store for the oldest old? Neurology. 1996; 46(3):641–650. [PubMed: 8618660]

Lee JH, Barral S, Cheng R, Chacon I, Santana V, Williamson J, Lantigua R, Medrano M, Jimenez-Velazquez IZ, Stern Y, Tycko B, Rogaeva E, Wakutani Y, Kawarai T, St George-Hyslop P, Mayeux R. Age-at-onset linkage analysis in Caribbean Hispanics with familial late-onset Alzheimer's disease. Neurogenetics. 2008; 9(1):51–60. [PubMed: 17940814]

Lee JH, Cheng R, Santana V, Williamson J, Lantigua R, Medrano M, Arriaga A, Stern Y, Tycko B, Rogaeva E, Wakutani Y, Kawarai T, St George-Hyslop P, Mayeux R. Expanded genomewide scan implicates a novel locus at 3q28 among Caribbean Hispanics with familial Alzheimer disease. Archives of Neurology. 2006; 63(11):1591–1598. [PubMed: 17101828]

Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K, Crowley AC, Fu YH, Guenette SY, Galas D, Nemens E, Wijsman EM, Bird TD, Schellenberg GD, Tanzi RE. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science. 1995; 269:973–977. [PubMed: 7638622]

Li Y, Grupe A, Rowland C, Nowotny P, Kauwe JS, Smemo S, Hinrichs A, Tacey K, Toombs TA, Kwok S, Catanese J, White TJ, Maxwell TJ, Hollingworth P, Abraham R, Rubinsztein DC, Brayne

C, Wavrant-De Vrieze F, Hardy J, O'Donovan M, Lovestone S, Morris JC, Thal LJ, Owen M, Williams J, Goate A. DAPK1 variants are associated with Alzheimer's disease and allele-specific expression. Human Molecular Genetics. 2006; 15(17):2560–8. [PubMed: 16847012]

Liu F, Arias-Vasquez A, Sleegers K, Aulchenko YS, Kayser M, Sanchez-Juan P, Feng BJ, Bertoli-Avella AM, van Swieten J, Axenovich TI, Heutink P, van Broeckhoven C, Oostra BA, van Duijn CM. A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. American Journal of Human Genetics. 2007; 81(1):17–31. [PubMed: 17564960]

Marchani EE, Bird TD, Steinbart EJ, Rosenthal E, Yu CE, Schellenberg GD, Wijsman EM. Evidence for three loci modifying age-at-onset of Alzheimer's disease in early-onset PSEN2 families. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2010; 153B:1031–1041.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34(7):939–944. [PubMed: 6610841]

Meyer JM, Breitner JCS. Multiple threshold model for the onset of Alzheimer's disease in the NAS-NRC twin panel. American Journal of Medical Genetics. 1998; 81(1):92–97. [PubMed: 9514594]

Mohs RC, Breitner JCS, Silverman JM, Davis KL. Alzheimers-Disease - Morbid Risk among 1st-Degree Relatives Approximates 50-Percent by 90 Years of Age. Archives of General Psychiatry. 1987; 44(5):405–408. [PubMed: 3579492]

Myers A, De-Vrieze FW, 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. Full Genome Screen for Alzheimer Disease: Stage II Analysis. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2002; 114:235–244.

Myers A, Holmans P, Marshall H, Kwon J, Meyer D, Ramic D, Shears S, Booth J, DeVrieze F, Crook R, Hamshere M, Abraham R, Tunstall N, Rice F, Carty S, Lillystone S, Kehoe P, Rudrasingham V, Jones L, Lovestone S, Pérez-Tur J, Williams J, Owen M, Hardy J, Goate A. Susceptibility locus for Alzheimer's disease on chromosome 10. Science. 2000; 290:2304. [PubMed: 11125144]

Pericak-Vance MA, Grubber JM, Bailey LR, Hedges D, West S, Santoro L, Kemmerer B, Hall JL, Saunders AM, Roses AD, Small GW, Scott WK, Conneally PM, Vance JM, Haines JL. Identification of Novel Genes in Late-Onset Alzheimer's Disease. Experimental Gerontology. 2000; 35:1343–1352. [PubMed: 11113612]

Rademakers R, Cruts M, Sleegers K, Dermaut B, Theuns J, Aulchenko Y, Weckx S, De Pooter T, Van den Broeck M, Corsmit E, De Rijk P, Del-Favero J, van Swieten J, van Duijn CM, Van Broeckhoven C. Linkage and association studies identify a novel locus for Alzheimer disease at 7q36 in a Dutch population-based sample. American Journal of Human Genetics. 2005; 77(4):643–52. [PubMed: 16175510]

Risch N, Merikangas KR. The future of genetic studies of complex human diseases. Science. 1996; 273:1516–1517. [PubMed: 8801636]

Rosenthal EA, Wijsman EM. Joint linkage and segregation analysis under multiallelic trait inheritance: Simplifying interpretations for complex traits. Genetic Epidemiology. 2010; 34:344–353. [PubMed: 20091797]

Scott WK, Hauser ER, Schmechel DE, Welsh-Bohmer KA, Small GW, Roses AD, Saunders AM, Gilbert JR, Vance JM, Haines JL, Pericak-Vance MA. Ordered-subsets linkage analysis detects novel Alzheimer disease loci on chromosomes 2q34 and 15q22. American Journal of Human Genetics. 2003; 73(5):1041–1051. [PubMed: 14564669]

Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, Bis JC, Smith AV, Carassquillo MM, Lambert JC, Harold D, Schrijvers EMC, Ramirez-Lorca R, Debette S, Longstreth WT, Janssens A, Pankratz VS, Dartigues JF, Hollingworth P, Aspelund T, Hernandez I, Beiser A, Kuller LH, Koudstaal PJ, Dickson DW, Tzourio C, Abraham R, Antunez C, Du YC, Rotter JI, Aulchenko YS, Harris TB, Petersen RC, Berr C, Owen MJ, Lopez-Arrieta J, Varadarajan BN, Becker JT, Rivadeneira F, Nalls MA, Graff-Radford NR, Campion D, Auerbach S, Rice K, Hofman A, Jonsson PV, Schmidt H, Lathrop M, Mosley TH, Au R, Psaty BM, Uitterlinden AG, Farrer LA, Lumley T, Ruiz A, Williams J, Amouyel P, Younkin SG, Wolf PA, Launer LJ, Lopez OL, van Duijn CM, Breteler MMB. Genome-wide Analysis of Genetic Loci

Associated With Alzheimer Disease. Journal of the American Medical Association. 2010; 303(18): 1832–1840. [PubMed: 20460622]

Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, Tsuda T, Mar L, Foncin JF, Bruni AC, Montesi MP, Sorbi S, Rainero I, Pinessi L, Nee L, Chumakov I, Pollen D, Brookes A, Sanseau P, Polinsky RJ, Wasco W, Dasilva HAR, Haines JL, Pericak-Vance MA, Tanzi RE, Roses AD, Fraser PE, Rommens JM, St George-Hyslop PH. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995; 375:754–760. [PubMed: 7596406]

Sillen A, Andrade J, Lilius L, Forsell C, Axelman K, Odeberg J, Winblad B, Graff C. Expanded high-resolution genetic study of 109 Swedish families with Alzheimer's disease. European Journal of Human Genetics. 2008; 16(2):202–208. [PubMed: 17957224]

Slooter AJC, Cruts M, Kalmijn S, Hofman A, Breteler MMB, Van Broeckhoven C, van Duijn CM. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: The Rotterdam study. Archives of Neurology. 1998; 55(7):964–968. [PubMed: 9678314]

Wijsman EM, Daw EW, Yu CE, Payami H, Steinbart EJ, Nochlin D, Conlon EM, Bird TD, Schellenberg GD. Evidence for a novel late-onset Alzheimer's disease locus on chromosome 19p13.2. American Journal of Human Genetics. 2004; 75:398–409. [PubMed: 15248153]

Wijsman EM, Daw EW, Yu X, Steinbart EJ, Nochlin D, Bird TD, Schellenberg G. APOE and other loci affect age-at-onset in Alzheimer's disease families with PS2 mutation. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2005; 132B:14–20.

Wijsman EM, Pankratz ND, Choi Y, Rothstein JH, Faber KM, Cheng R, Lee JH, Bird TD, Bennett DA, Diaz-Arrastia R, Goate AM, Farlow MR, Ghetti B, Sweet RA, Foroud TM, Mayeux R. Genome wide association of familial late onset Alzheimer's disease replicates *BIN1* and *CLU*, and nominates *CUGBP2* in interaction with *APOE*. PLoS Genetics. 2011; 7(2):e1001308. [PubMed: 21379329]

Wijsman EM, Rothstein JH, Igo RP, Brunzell JD, Motulsky AG, Jarvik GP. Linkage and association analyses identify a candidate region for apoB level on chromosome 4q32.3 in FCHL families. Human Genetics. 2010; 127(6):705–719. [PubMed: 20383777]

Wijsman EM, Yu D. Joint Oligogenic Segregation and Linkage Analysis using Bayesian Markov chain Monte Carlo Methods. Molecular Biotechnology. 2004; 28:205–226. [PubMed: 15542922]
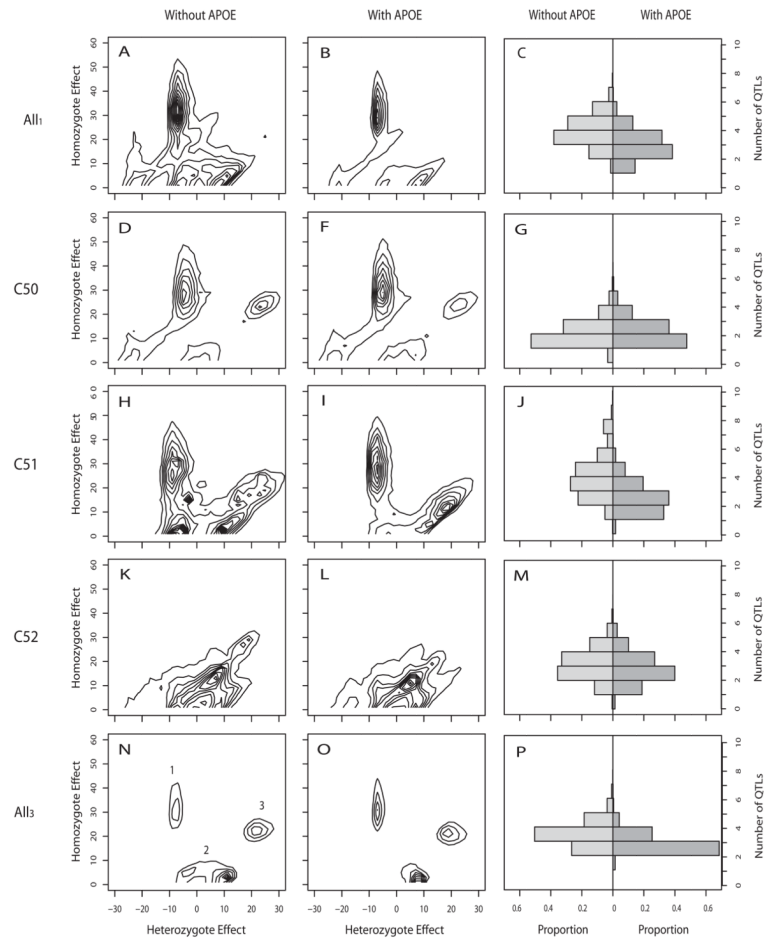
Figure 1.
QTL models from MCMC oligogenic segregation analysis. Columns 1 and 2 show probability density plots for QTL models obtained without (column 1) and with (column 2) adjustment for *APOE*. Column 3 depicts the posterior distribution of the number of QTLs without and with adjustment for *APOE*. Rows 1 (panels A–C) and 5 (panels N–P) show results from analysis of all centers jointly, with row 5 showing results that allowed for different QTL allele frequencies among centers. Rows 2–4 present results of analyses on individual centers, with Center 50 (panels D–G), Center 51 (panels H–J), and Center 52 (panels K–M).
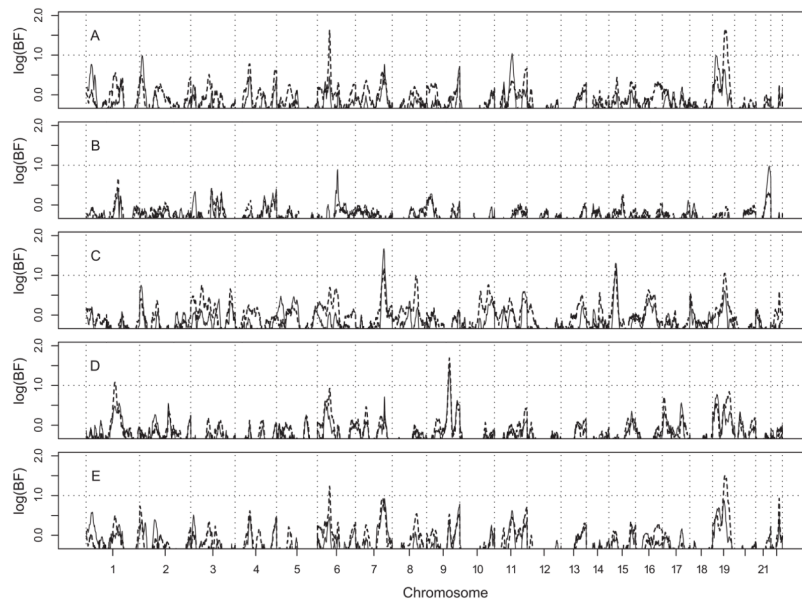
Figure 2.
Genome scan for age-at-onset of AD. For each panel, dashed line: log(BF) obtained without adjustment for *APOE*; Solid line: log(BF) obtained with adjustment for *APOE* as a major gene covariate. Panel A: All_1; B: Center 50; C: Center 51; D: Center 52; and E: All_3. Vertical dotted lines depict boundaries between chromosomes; horizontal dotted line is at logBF=1. See Supplementary Figure 2 for a color version.
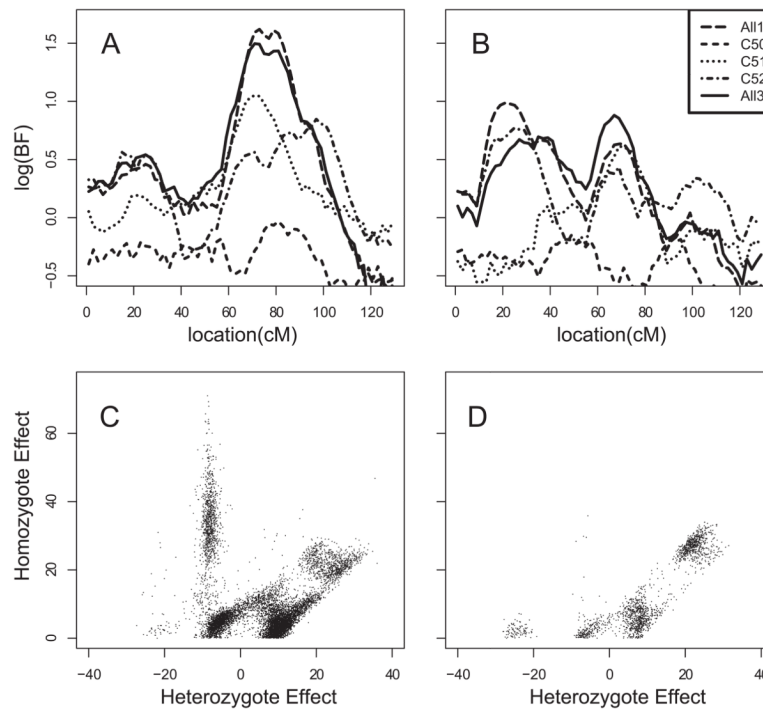
Figure 3.
Multipoint linkage analysis of chromosome 19. A: Linkage analysis without adjustment for *APOE* for all five analysis configurations; B: Linkage analysis with adjustment for *APOE* for all five analysis configurations; C: Scatterplot showing realized QTL model parameters across MCMC iterations in analysis of All$_3$ without adjustment for *APOE* for the interval from 55 to 95cM; D: Realized QTL model parameters across MCMC iterations in analysis of All$_3$ with adjustment for *APOE* for the interval from 55 to 95cM.
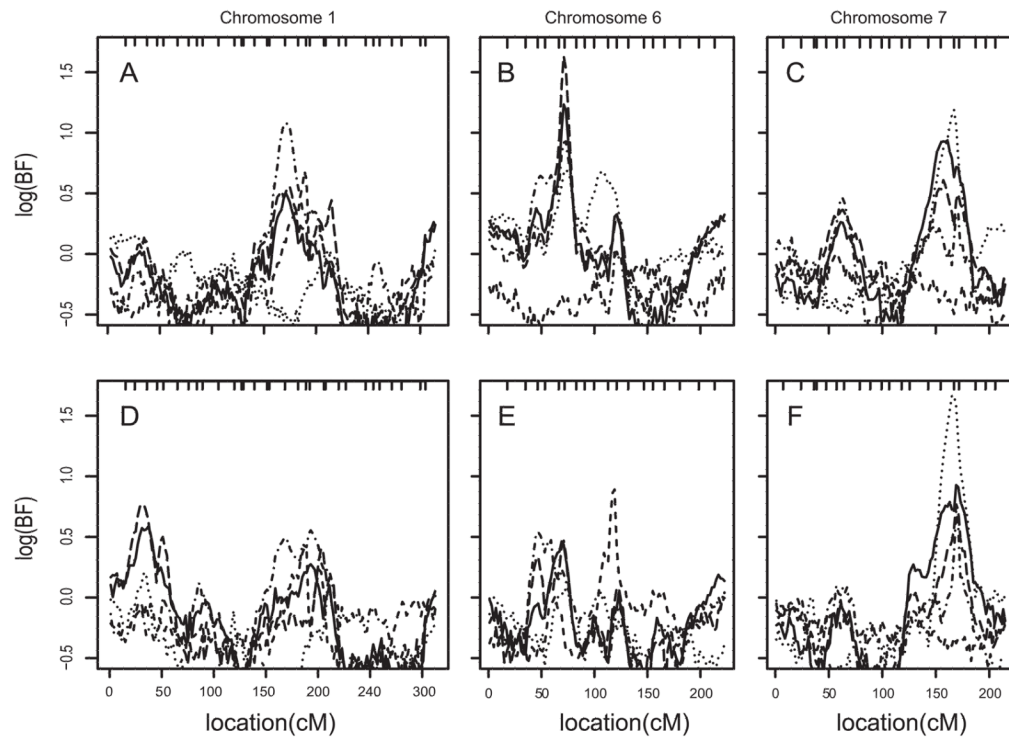
Figure 4.
MCMC multipoint oligogenic linkage analyses for chromosome 1 (panels A, D), 6 (panels B, E) and 7 (panels C, F), with upper row (panels A–C) carried out without adjustment for *APOE*, and lower row (panels D–F) carried out with adjustment for *APOE*. Analysis configurations are represented by long dashed line: All$_1$; short dashed line: Center 50; dot line: Center 51; dot dashed line: Center 52; solid line: All$_3$. See Supplementary Figure 3 for a color version.

Figure 5.
Score represents LOD score from non-parametric linkage analysis with 303 all Caucasian ASPs (solid line) or BF from MCMC multipoint linkage analysis with the same 303 ASPs (dashed line) and under analysis of All_3 (dotted line).
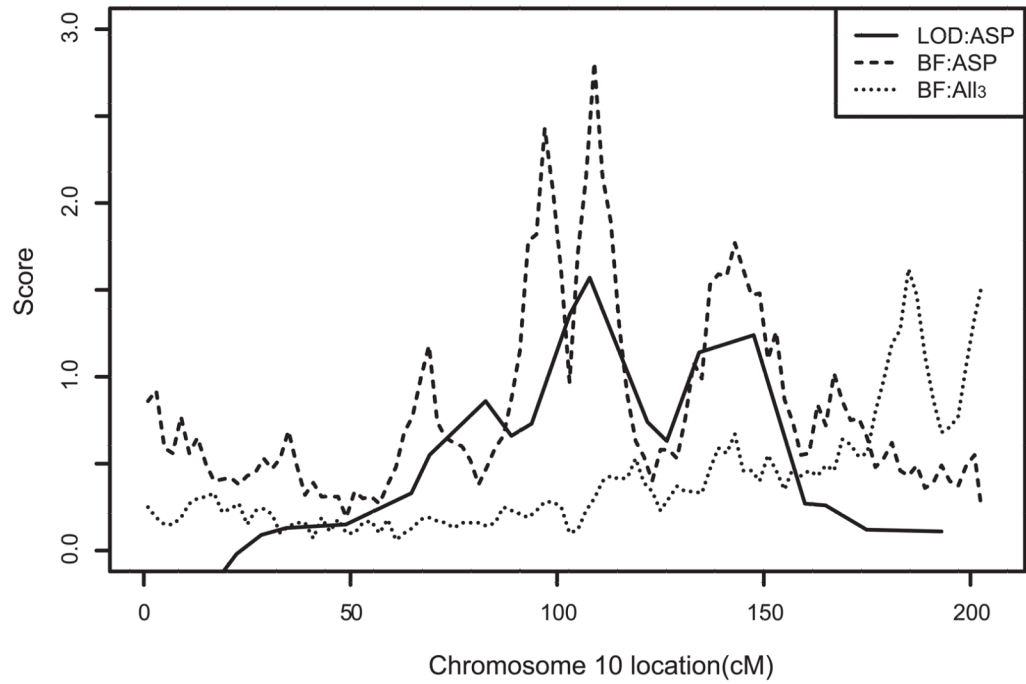
Table 1

Sample characteristics

Center	Peds	Diagnosis				Pedigree size			Age				Genotyped		APOE allele frequency	
		Obs	Un	Aff	Miss	Mean	Range	All	Un	Aff	Scan	APOE	ε3	ε4
C50	134	549	224	325	519	6.79	4–52	70.3	68.9 [29–99]	71.3 [34–88]	521	516	0.52 (0.02)	0.44 (0.02)
C51	124	545	278	267	164	5.32	4–13	71.4	70.2 [29–105]	72.6 [41–93]	375	376	0.63 (0.02)	0.33 (0.02)
C52	139	482	144	338	313	5.53	4–13	71.5	71.2 [46–91]	71.7 [42–97]	477	466	0.55 (0.02)	0.41 (0.02)
All	397	1576	646	930	996	5.89	4–52	71.1	70.0 [29–105]	71.8 [34–97]	1373	1358	0.58 (0.01)	0.39 (0.01)

Peds: Pedigrees; Obs: Subjects with observed diagnosis information; Un: Unaffected subjects; Aff: Affected subjects; Miss: Subjects with missing diagnosis in the analysis.

Note: Ages for unaffected subjects are the latest age for which they were observed

Means and ranges are reported for the age of unaffected and affected samples.

Means and standard deviations are reported for APOE allele frequencies.

Table 2

Summary of QTL models

Center	Model	Subpop	Without *APOE* covariate			With *APOE* covariate		
			p_A	$\mu_{AB}\text{-}\mu_{AA}$	$\mu_{BB}\text{-}\mu_{AA}$	p_A	$\mu_{AB}\text{-}\mu_{AA}$	$\mu_{BB}\text{-}\mu_{AA}$
All$_1$	1	-	0.76 (0.06)	−6.91 (1.40)	34.86 (7.87)	0.76 (0.08)	−6.37 (1.59)	33.28 (7.49)
	2	-	0.39 (0.24)	10.91 (3.44)	5.47 (3.19)	0.31 (0.20)	11.47 (4.11)	6.52 (4.40)
	3	-	NA	NA	NA	NA	NA	NA
C50	1	-	0.65 (0.25)	−3.29 (3.01)	31.44 (8.55)	0.69 (0.18)	−4.85 (1.92)	32.49 (8.12)
	2	-	0.48 (0.32)	2.78 (5.10)	4.49 (3.02)	0.50 (0.29)	1.87 (4.54)	3.46 (2.26)
	3	-	0.19 (0.12)	23.55 (3.54)	23.68 (3.24)	0.18 (0.14)	23.68 (3.58)	24.34 (3.33)
C51	1	-	0.73 (0.10)	−8.25 (2.81)	29.63 (8.03)	0.71 (0.11)	−6.83 (2.10)	30.62 (8.41)
	2	-	0.43 (0.21)	10.42 (4.48)	4.71 (2.74)	NA	NA	NA
	3	-	0.21 (0.13)	20.24 (5.46)	18.01 (5.37)	0.26 (0.13)	17.26 (4.75)	12.63 (5.35)
C52	1	-	NA	NA	NA	NA	NA	NA
	2	-	0.46 (0.28)	10.08 (3.06)	5.14 (2.90)	0.52 (0.26)	3.23 (3.42)	9.88 (2.74)
	3	-	0.15 (0.17)	16.85 (3.21)	27.04 (3.42)	NA	NA	NA
All$_3$	1	C50	0.70 (0.05)	−7.53 (1.34)	34.37 (7.87)	0.71 (0.04)	−6.68 (1.16)	33.22 (7.21)
		C51	0.69 (0.05)			0.68 (0.04)		
		C52	0.91 (0.04)			0.90 (0.05)		
All$_3$	2	C50	0.25 (0.29)	9.93 (2.00)	3.13 (1.64)	0.41 (0.24)	7.89 (1.27)	2.48 (1.74)
		C51	0.46 (0.14)			0.43 (0.14)		
		C52	0.76 (0.32)			0.71 (0.34)		
All$_3$	3	C50	0.20 (0.04)	22.43 (2.94)	22.65 (2.40)	0.19 (0.04)	20.21 (3.02)	21.11 (2.68)
		C51	0.18 (0.05)			0.17 (0.05)		
		C52	0.17 (0.05)			0.21 (0.05)		

Center: analysis unit used for analysis as defined in text; Model: model number; Subpop: Subpop: center that is assumed to have a unique QTL allele frequency but with the same genotype differences as the other subpopulations in the All$_3$ center analysis; p_A: A allele frequency for QTL model; $\mu_{AB}\text{-}\mu_{AA}$ and $\mu_{BB}\text{-}\mu_{AA}$: difference in genotype means for AA and AB genotypes and for AA and BB genotypes, respectively; NA: not available in model space; dashes: not relevant

Table 3

Log of Maximum of Bayes Factor (location) without (NIMH) and with (NIMH+EO) joint analysis with early-onset pedigree.

Center	APOE cov.	Chromosome 1		Chromosome 7	
		NIMH	NIMH + EO	NIMH	NIMH + EO
C50	−	0.68 (187cM)	0.92 (189cM)		
	+	0.41 (187cM)	0.46 (182cM)		
C51	−			1.19 (167cM)	1.43 (167cM)
	+			1.67 (167cM)	1.91 (167cM)
C52	−	1.08 (169cM)	1.10 (170cM)	0.33 (171cM)	0.53 (171cM)
	+	0.55 (193cM)	1.21 (195cM)	0.71 (169cM)	0.31 (171cM)
All₃	−	0.50 (169cM)	0.30 (176cM)	0.94 (161cM)	1.18 (162cM)
	+	0.27 (193cM)	0.47 (186cM)	0.93 (169cM)	0.93 (167cM)

EO: Early-onset; Center: analysis unit used for analysis, as defined in text; cov.: covariate; cM: centimorgan location on the chromosomal meiotic map.