Human Molecular Genetics, 2007, Vol. 16, No. 22 ***2703–2712***
doi:10.1093/hmg/ddm224
Advance Access published on August 27, 2007

Genome-wide linkage analysis of 723 affected relative pairs with late-onset Alzheimer's disease

Marian L. Hamshere[1,2], Peter A. Holmans[1,2], Dimitrios Avramopoulos[3,4], Susan S. Bassett[3], Deborah Blacker[5,7], Lars Bertram[6], Howard Wiener[8], Nan Rochberg[9], Rudolph E. Tanzi[6], Amanda Myers[10], Fabienne Wavrant-De Vrièze[11], Rodney Go[8], Daniele Fallin[3], Simon Lovestone[12], John Hardy[13], Alison Goate[9], Michael O'Donovan[2], Julie Williams[1,2] and Michael J. Owen[2,*]

[1]Biostatistics and Bioinformatics Unit, and [2]Department of Psychological Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK, [3]Department of Psychiatry, and [4]McKusick Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, [5]Gerontology Research Unit, Department of Psychiatry, and [6]Genetics and Aging Research Unit, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129, USA, [7]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA, [8]Department of Epidemiology, University of Alabama at Birmingham, USA, [9]Department of Psychiatry, Washington University School of Medicine, 660 S. Euclid Avenue, St Louis, MO 63110, USA, [10]Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, USA, [11]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, 35 Convent Drive, Bethesda, MD 20892-3707, USA, [12]MRC Centre for Neurodegeneration Research, King's College London, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK and [13]Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, UCL, Queen Square, London WC1N 3BG, UK

Received August 11, 2007; Revised and Accepted August 12, 2007

Previous attempts to identify genetic loci conferring risk for late-onset Alzheimer's disease (LOAD) through linkage analysis have observed some regions of linkage in common. However, due to the sometimes-considerable overlap between the samples, some of these reports cannot be considered to be independent replications. In order to assess the strength of the evidence for linkage and to obtain the best indication of the location of susceptibility genes, we have amalgamated three large samples to give a total of 723 affected relative pairs (ARPs). Multipoint, model-free ARP linkage analysis was performed. Genome-wide significant evidence for linkage was observed on 10q21.2 (LOD = 3.3) and genome-wide suggestive evidence was observed on 9q22.33 (LOD = 2.5) and 19q13.32 (LOD = 2.0). One further region on 9p21.3 was identified with an LOD score > 1. We observe no evidence to suggest that more than one locus is responsible for the linkage to 10q21.2, although this linked region may harbour more than one susceptibility gene. Evidence of allele-sharing heterogeneity between the original collection sites was observed on chromosome 9 but not on chromosome 10 or 19. Evidence for an interaction was observed between loci on chromosomes 10 and 19. Where samples overlapped, the genotyping consistency was high, estimated to average at 97.3%. Our large-scale linkage analysis consolidates clear evidence for a susceptibility locus for LOAD on 10q21.2.

INTRODUCTION

Alzheimer's disease (AD) is a heritable (1), debilitating disorder characterized by a gradual decline in cognitive abilities. AD is estimated to account for two-thirds of individuals with dementia (2) and to affect approximately 15 million people worldwide (3). Variants of three genes (APP, PS-1 and PS-2) play a major role in the genetics of early-onset autosomal dominant AD (4–6). The great majority of AD cases are of late onset and show complex, non-Mendelian patterns of

*To whom correspondence should be addressed. Tel: +44 2920743058; Fax: +44 2920746554; Email: owenmj@cardiff.ac.uk

inheritance. Late-onset AD (LOAD) probably results from the combined effects of variations in a number of genes as well as environmental factors. Recent twin studies suggest that the heritability of AD is in the range of 60–80%, with the balance of variance attributable to unique (typically adult) environmental influences (1,7). The only genetic risk factor for LOAD to have been identified with certainty is the apolipoprotein E (*APOE*) gene. The *APOE*4 allele is highly enriched in AD cases compared with non-demented individuals, whereas the *APOE*2 allele is under-represented in cases (8). However, only 50% of AD cases carry an *APOE*4 allele (9–11), and it is clear that other genes play a role, in particular to the age at onset of illness (12). Many other positive findings have been reported from association studies of functional and positional candidate genes, but unequivocal evidence for association has yet to be obtained. A database presenting meta-analyses of current candidate gene association studies can be found on the Alzheimer Research Forum at www.alzgene.org (13).

Linkage analysis normally provides the first step in identifying regions of the genome that could harbour novel susceptibility genes. There have been a number of genome-wide linkage studies of AD using large samples of families with two or more affected relatives (14–21). The major regions identified are found on chromosomes 9p, 9q, 12p and 19q, with the most convincing evidence for linkage on 10q (14,16,22–25). These results support the possibility that a number of susceptibility genes of moderate effect size influence AD risk and which remain to be identified. However, many of the family samples used in the linkage analyses overlap, and this has made it difficult to assess the strength and consistency of the evidence in favour of linkage. We have therefore combined data from a number of the published linkage studies, taking full account of overlap. The complex two-stage design of one of the studies (14,16) resulted in the total number of affected relative pairs (ARPs) with markers genotyped for different chromosomal regions varying between 423 and 710 in the total sample of 723 ARPs, and the largest number of available ARPs genotyped was for markers on chromosome 10. Given the mounting evidence to support an AD susceptibility gene on chromosome 10, we were particularly keen to analyse this chromosome to obtain the best evidence for linkage and the most accurate estimate of gene location from the available data since previous studies have provided discrepant estimates of the location of the putative risk gene. Given that many markers were genotyped more than once in different laboratories, combining the data also allowed us to assess genotyping quality between research groups.

RESULTS

Four regions produced a maximum LOD>1 on chromosomes 9p21.3, 9q22.33, 10q21.2 and 19q13.32 (Table 1 and Fig. 1). In addition, all four regions were identified as being of interest when the genotypes from the three groups were analysed separately (see Supplementary Material, Table S1 for more details) and showed no evidence of allele heterogeneity. Chromosomes 9 and 10 showed no significant increase in maximum LOD when *APOE* was included as a covariate in the model, whereas chromosome 19 showed a potential age at onset effect.

Table 1. All maximum LOD scores >1 in the full sample and their relevant locations

Chromosome	Maximum LOD	Genetic location of maximum LOD (cM)	Nearest genotyped microsatellite(s) to maximum LOD
9p21.3	1.2	40	D9S1870 and D9S741
9q22.33	2.5	104	D9S910
10q21.2	3.3	78	D10S464
19q31.32	2.0	70	D19S412

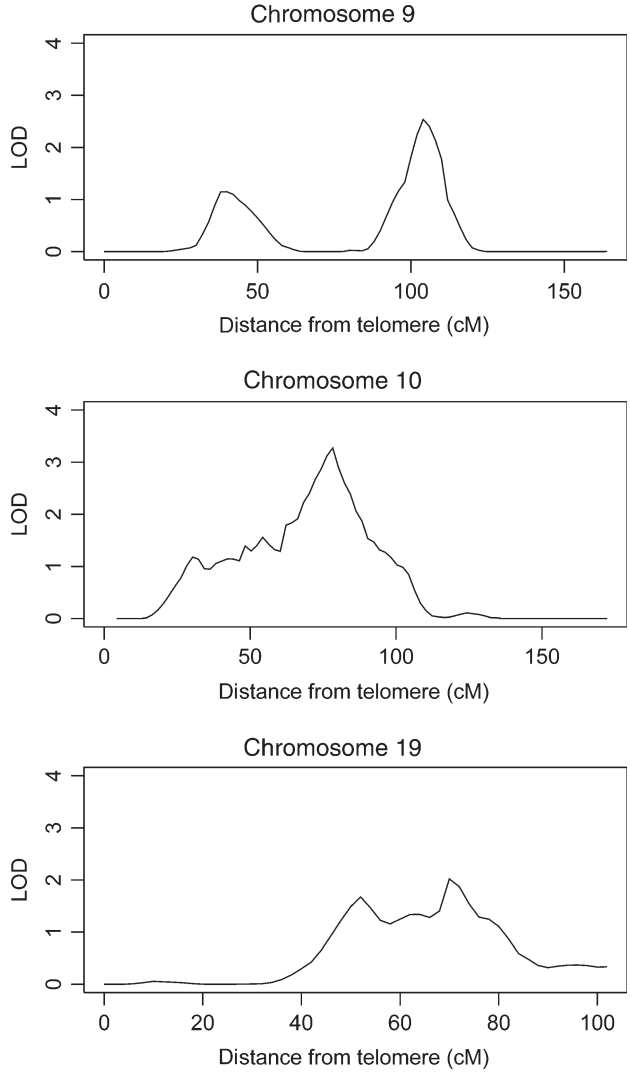

Figure 1. Plots of all chromosomes where the maximum LOD>1 in the full sample.

Chromosome 10

The highest LOD score was observed on chromosome 10 with a value of 3.3 at 78 cM (nearest marker = D10S464), with an

identity by descent (IBD) estimate of 0.58. This maximum LOD score equals the criterion required for significant genome-wide linkage (26). The LOD-1 region spans 17 cM (between 68 and 85 cM). The maximum LOD score is higher than the LOD scores obtained when analysing data from the three groups separately; note, however, that in their second stage linkage analysis, group 1 (16) had reported a maximum LOD score of 3.9 near 82 cM. The maximum LOD score for the sample studied by group 2 lies at 88 cM, which is outside the overall LOD-1 region, suggesting that inclusion of the additional data from groups 1 and 3 has provided extra information. A test of homogeneity in allele sharing between sites of collection was carried out by incorporating site as a 5-level factor in a covariate linkage analysis (see Method of Analysis). This gave a chromosome-wide *P*-value of 0.18, indicating no significant differences in allele sharing between the five sites. However, the power of such a test may be low due to the multiple degrees of freedom incurred by testing differences between all sites simultaneously. The location of the maximum LOD score when the site covariate was included was 62 cM (outside the LOD-1 region) and the estimates of IBD allele-sharing probabilities (compared with an expected value of 0.5 in the absence of linkage) were 0.51, 0.46, 0.56, 0.60 and 0.58 in the UK, NIMH (National Institute of Mental Health Genetics Initiative for Alzheimer's Disease) sites 50, 51 and 52 and the NIA (National Institute on Aging, National Cell Repository for Alzheimer's Disease) samples, respectively. Thus, there may be some heterogeneity of IBD sharing between sites despite the overall test for homogeneity being non-significant.

The analysis described in the previous paragraph does not test for homogeneity in the location of the maximum LOD score in each sample (which may indicate the presence of two or more disease-susceptibility genes in the region). For a complex trait like LOAD, testing for such homogeneity in the whole sample would be complicated by the likely presence of several pedigrees that do not segregate any disease gene in the region. Therefore, we performed a linkage analysis on the 352 pedigrees (77% of the whole sample, denoted LOD_+) which showed elevated allele sharing ($LOD > 0$) over the region 60–90 cM (covering the LOD-1 region for the linkage analysis on the whole sample, the location where there was maximum evidence for IBD heterogeneity between sites and the maximum LOD score locations from the samples studied by groups 1–3). The results are shown in Figure 2. Such an analysis is clearly invalid as a test for linkage, but may give insights into the likely location of a disease susceptibility locus. As expected, the magnitude of the peak (18.0) in the LOD_+ sample far exceeds that in the whole sample. The shape of the LOD score curve for the LOD_+ sample is similar to that from the whole sample (Fig. 2A) and shows no evidence of there being more than one disease locus in this region of chromosome 10. Although, we cannot rule out there being more than one susceptibility gene within this locus. Likewise, there seems to be no difference between collection sites in the locations of maximum IBD allele sharing among the LOD_+ pedigrees (Fig. 2B), again supporting the view that the same locus is conferring risk in all samples. Note that the LOD_+ pedigrees from the UK sample show the highest IBD, despite the IBD from the UK sample as a whole being close to 0.5 in the region. This indicates that a large proportion of pedigrees in this sample are unlikely to be segregating a disease gene in this region. The LOD-1 region is reduced to 7 cM (73–80 cM) in the LOD+ sample and this suggests that the LOD scores of the non-linked pedigrees are randomly distributed across the region.

Chromosome 9

Two distinct peaks were observed on chromosome 9. The maximum LOD of 2.5 was observed at 9q22.33 (104 cM, nearest marker = D9S910, genome-wide suggestive), and a second peak with a maximum LOD of 1.2 at 9p21.3 (40 cM, nearest markers = D9S1870 and D9S741). A test for homogeneity of allele sharing was performed between the five sites of collection. A significant difference in allele sharing was observed (chromosome-wide $P = 0.04$). Including the site of collection as a covariate maximized the LOD at 102 cM. Inspection of the allele-sharing estimates at this maximum gave UK: 0.45, NIMH 50: 0.58, NIMH 51: 0.62, NIMH 52: 0.59 and NIA: 0.50, indicating that the linkage evidence at the main peak is due to the NIMH data only. There was no statistical evidence of a difference between the three original NIMH collection sites (chromosome-wide $P = 0.78$). Given that the test statistic uses the maximum LOD scores on the chromosome, these significance levels are related to the main peak on 9q22.33. We also performed supplementary homogeneity tests on the restricted chromosome length between 24 and 64 cM, thereby targeting the secondary peak in the full sample. There was no significant difference observed between the five collection sites (chromosome-wide $P = 0.25$).

Analysing the pedigrees with reasonable evidence of linkage to the main peak ($LOD > 0$ for at least 50% of the LOD-1 region, peak location ± 6 cM, $n = 208$) produced a single peak at 102 cM. This criterion for linkage evidence was chosen to ensure definite membership to this peak rather than the secondary peak. Similarly, focusing on the secondary peak ± 6 cM region ($n = 199$) produced a single peak at 40 cM. These two peaks from the separate analyses are of similar height and in the same locations as those when analysing the full data set. Analysing the pedigrees with evidence for linkage ($LOD > 0$) for at least 50% of both the main and secondary peak regions ($n = 102$) gave LOD scores in both peak regions, approximately half the height of those seen in the separate analyses (from approximately half the number of pedigrees; Fig. 3). These results suggest that pedigrees with increased allele sharing at one of the loci have a 50% chance of having increased sharing at the other, i.e. the two loci are independent. The correlation in allele sharing between the two loci is not significantly different from what you would expect by chance, given the distance between the two loci ($P = 0.13$), suggesting that the two loci should be treated separately.

Chromosome 19

A maximum LOD score of 2.0 was observed at 19q13.32 (70 cM, nearest marker = D19S412, genome-wide suggestive). The LOD-1 region covers 44 cM (46–82 cM), including

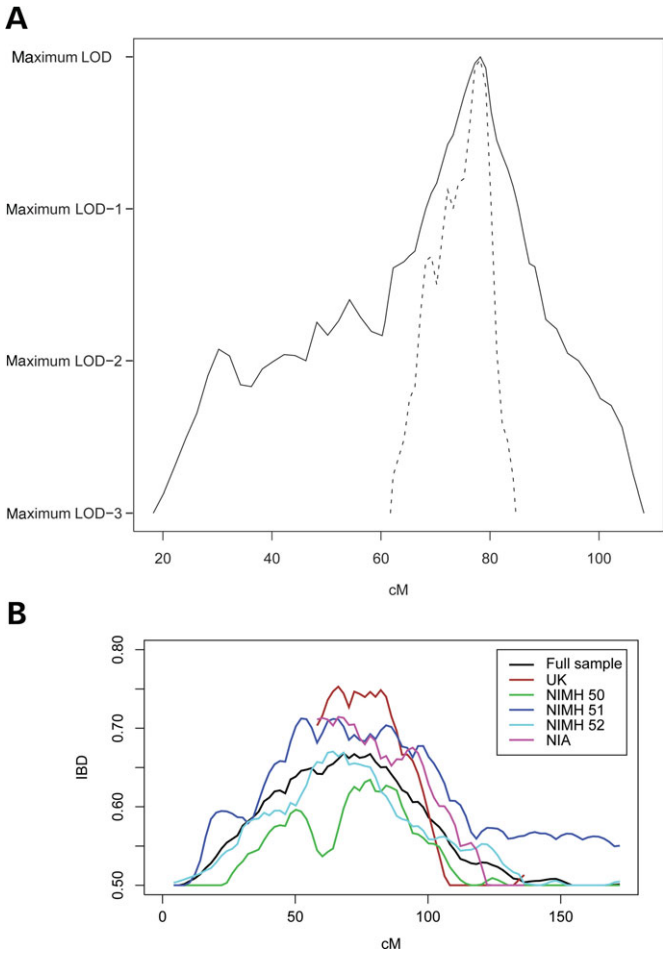

Figure 2. Chromosome 10 linkage analyses of pedigrees showing evidence for linkage (LOD_+, i.e. those pedigrees with LOD > 0) between 60 and 90 cM. (**A**) Comparison of the peak region. One can see that the peak from the LOD_+ pedigrees (dashed line) shows a better resolution than that from all the pedigrees (solid line). (**B**) IBD allele-sharing probabilities in the LOD_+ sample, split by sites of collection.

the location of the *APOE* gene. Similar evidence was observed in the two original data sets and no evidence of site of collection heterogeneity was detected. In the ARPs with data, evidence for linkage, conditional on age at onset, produced an increase in maximum LOD score of 3.8, from 2.7 to 6.3 at 70 cM, close to *APOE*. Increased allele sharing was observed in the pairs with a younger age at onset. *APOE* is known to have an effect on age at onset (27), so we corrected age at onset for *APOE*4 and repeated the analysis. The maximum age at onset covariate LOD score was reduced, but the observed effect was not eliminated (Fig. 4). This confirms that *APOE* has an effect on age at onset, but also suggests that there may be an additional locus in the region influencing age at onset.

Interaction results

ARP locus–locus interaction analyses of chromosomes 9, 10 and 19 were performed, by testing the allele-sharing probabilities of chromosome I, conditional on chromosome J. The results are presented in Figure 5. The most interesting result

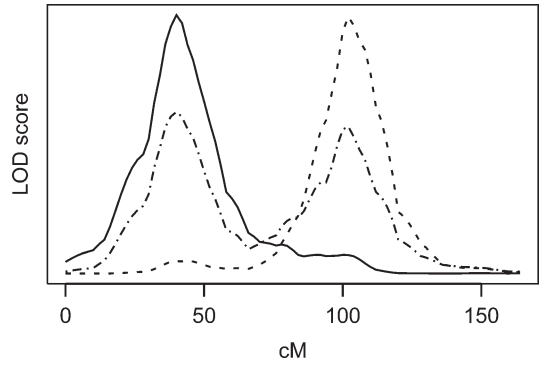

Figure 3. Chromosome 9 linkage analyses of pedigrees showing evidence for linkage (LOD>0 for at least 50% of the given region) at solid line, secondary peak ±6 cM ($n = 199$); dashed line, main peak LOD-1 region (12 cM, $n = 208$) and dashed-dotted line: both secondary peak ±6 cM and main peak LOD-1 region ($n = 102$).

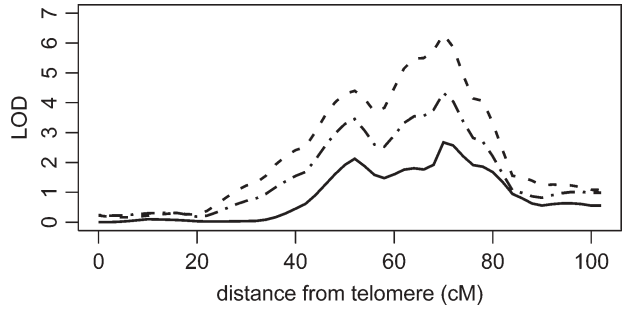

Figure 4. Chromosome 19 linkage analysis of individuals with age at onset and *APOE* data. The LOD scores represent an unconditional linkage analysis (solid line), conditional on age at onset (dashed line), conditional on age at onset after removing the effect of the presence or absence of *APOE*4 (dotted-dashed line).

was observed when analysing chromosome 10 (130 cM), conditional on chromosome 19 (38 cM). The unconditional LOD score at this point on chromosome 10 was 0.2, increasing by 1.8 to 2.0 when conditioning on chromosome 19 (point-wise $P = 0.005$). The covariate parameter is positive, indicating that at these particular locations, ARPs with increased allele sharing on chromosome 19 also have increased allele sharing on chromosome 10. A similar effect was observed when conditioning chromosome 10 on 19. The ARPs with increased allele sharing in this same region on chromosome 10 also have increased allele sharing on chromosome 9 at 66 cM, although there is no such effect between chromosomes 9 and 19.

Chromosome 12

A maximum LOD score of 0.2 at 26 cM was observed on 12p13.2 (co-localized by D12S391 and D12S358). Although we do not observe evidence to suggest that a gene for AD resides on this chromosome, some studies have observed such evidence (28,29), and so we performed further covariate analyses to see whether we could replicate the findings observed in a sample, including a large proportion of NIMH pedigrees, of Liang *et al*. (28). Liang *et al*. found linkage

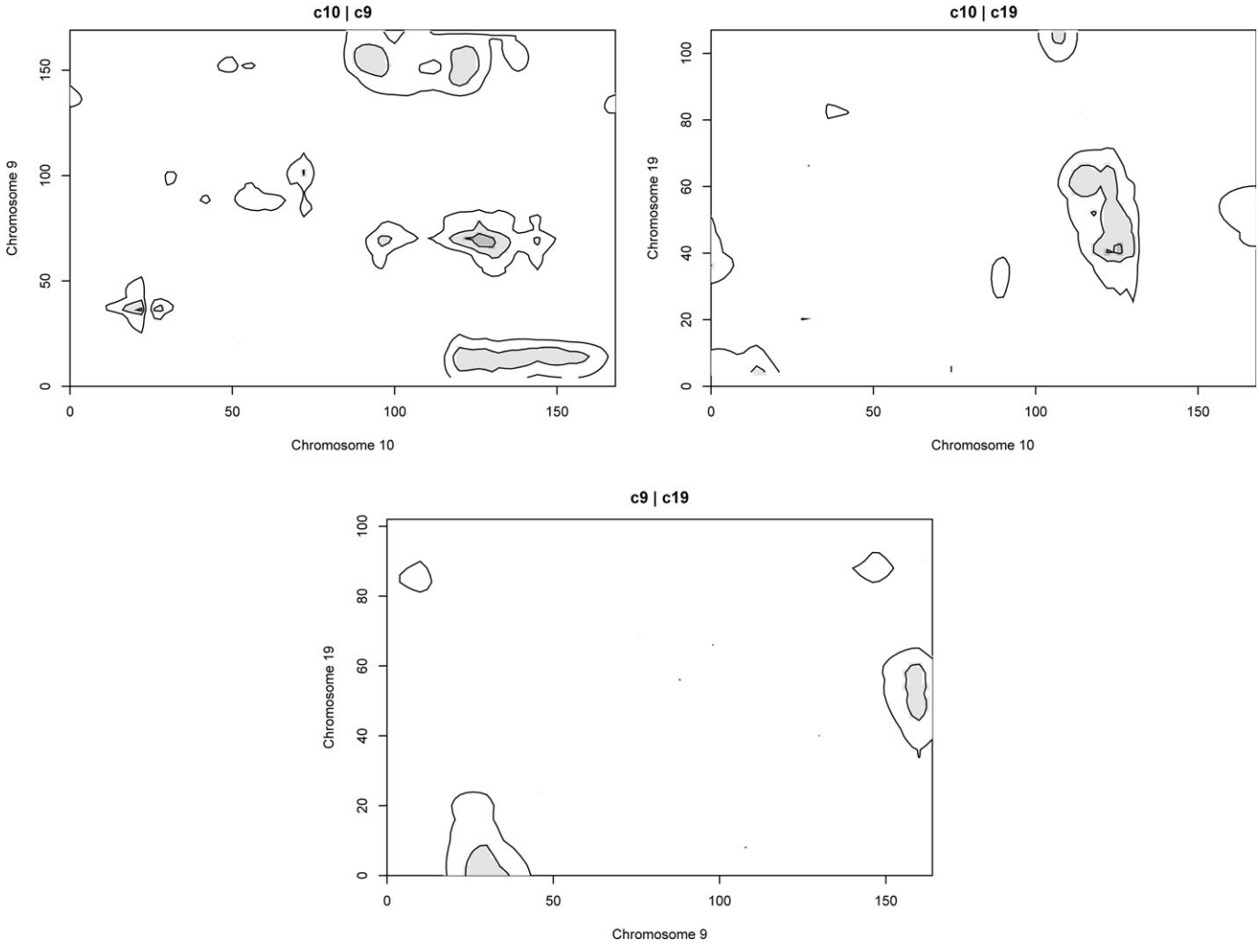

Figure 5. Interaction plots. Increases in LOD score on primary chromosome (*x*-axis) when allowing for allele sharing on a secondary chromosome (*y*-axis). Contour grades increment every increase in LOD of 0.5.

covariate effects in the 51–67 cM region when conditioning on the LOD score of an SNP (rs7070570) in the α-catenin gene (VR22) on chromosome 10 (82.5–86.2 cM, point-wise $P < 0.0001$), D9S741 (42 cM, point-wise $P = 0.20$) and the proportion of APOE ϵ4 alleles in the pair (point-wise $P = 0.04$) in an ordered subset analysis (OSA). We do not observe significant evidence of these effects in our sample, observing point-wise significance levels for the three covariates of 0.54, 0.39 and 0.58, respectively. The OSA method employed by Liang *et al*. seeks to identify a subset of the families in the sample defined by a continuous family-wise covariate that maximizes the evidence for linkage. Our method uses pair-wise measures of linkage on the chromosome being conditioned on, and so might be more sensitive in the presence of within-family linkage heterogeneity. Nevertheless, if these are true effects, we would also hope to identify them in our large sample.

Including *APOE* as a covariate maximized the LOD score at 22 cM (chromosome-wide $P = 0.53$), the IBD sharing probabilities observed at the peak were −/−: 0.58, −/+: 0.48 and +/+: 0.50 in 104, 119 and 374 ARPs, respectively. Although the *APOE* effect was not significant, elevated sharing in the −/− pairs was observed, confirming that found by group 1 (16). The IBD estimate in the *APOE* −/− pairs is of comparable magnitude with that observed at the peak on chromosome 10, although the sample size is much reduced.

Data quality

The mean (SD) genotyping consistency rate per marker was estimated to be 97.3 (2.6)%. A total of 46 012 genotypes were checked over 99 markers. In the four regions identified, the rate was estimated to be 99.4%, indicating that the genotyping in these regions is of a very high quality. Three outliers were detected with a rate < 90%, namely, D4S403 (26 cM), D9S1838 (164 cM) and D12S1045 (161 cM), which were removed from the analysis.

DISCUSSION

Multiple genome scans have found evidence for linkage to chromosomes 9, 10 and 19 (14–21). Interpretation of these data is difficult because the samples analysed overlapped in a number of instances, and so the results cannot be considered to be independent replications. We have amalgamated the raw data from three large studies of LOAD. Access to the raw data allowed us to perform a genome screen for linkage to LOAD, using the same inclusion criteria for all individuals. Further linkage analyses were also performed in regions of interest to investigate any potential heterogeneity between sites of collection, locus–locus interactions, age at onset effect and in particular with *APOE*. The main aim was to provide the best estimates for the most likely locations of genes involved with LOAD. We also aimed to assess the reliability of genotyping between the research groups. Although the data analysed here constitute the largest single sample yet assembled for a linkage study of AD on chromosome 10, there are other samples that we did not analyse. These include both samples of Caucasian origin (20,30) and samples drawn from other racial groups (21). An analysis of all available AD scans, allowing for differences in ethnic origin and other clinical phenotypes, would be worthwhile.

In our large sample, we observe evidence for linkage at 9p21.3, 9q22.33, 10q21.2 and 19q13.32. There is significant evidence to support a susceptibility locus on 10q21.2, with a strong indication that this locus operates in a number of populations, and the most likely location of the risk gene(s) is at 78 cM. Our findings also fail to support the presence of second locus on 10q. Although these findings are potentially of considerable importance in guiding further genetic studies aimed at identifying a susceptibility gene in this region, some caveats must be noted. In particular, we cannot definitively exclude the possibility that more than one locus for AD exists on chromosome 10q in spite of the fact that we did not obtain significant evidence for heterogeneity. The power of any linkage analysis aimed at identifying a second (or even third) locus in the peak region, especially without knowing a definite location of the first, is related to the distances between loci. Differentiating between one locus or two in close proximity is difficult in this setting. As the distance between the two potential loci increases, the chance of separating their respective linkage signals also increases. Larger samples of pedigrees genotyped and pooled into a single linkage analysis (as presented here) will be beneficial, as will be scans with higher genetic information content and resolution (e.g. using SNP genotypes). Although we cannot exclude the possibility that susceptibility is conferred by two genes located close together, or maybe on either side of the peak, but outside the LOD-1 region, considerations of parsimony encourage us to suggest that a single locus between 68 and 85 cM is responsible for the linkage to this region.

It has been suggested that the chromosome 10 linkage is subject to a parent of origin effect, with elevated allele sharing observed only in a subset pedigrees where AD has been inherited from the mother's family (15,31). This finding is based on a subset of the data analysed in the present study. However, since data suggesting the parent of origin of familial risk are only available for two of the NIMH family sets (sites 50 and 52), we were not able to supplement this analysis. Interestingly, the linkage obtained in this parent of origin analysis maximized at 75.6 cM, only 2 cM from our peak. Linkage to 10q21.2 (LOD = 1.7) was observed in a sample containing 457 NIMH pedigrees with age at onset >50 years (17). The evidence for linkage appeared to originate from both early- and late-onset cases. Correspondingly, in agreement with our results, Holmans *et al*. (32). found negligible effects of age at onset on the chromosome 10 linkage in the combined UK/NIMH/NIA sample of Myers *et al*. (16) when age at onset was restricted to be ≥65 years. Modest evidence for linkage was observed to 10q21.2 in an independent sample of predominantly late-onset (mean age at onset = 70.7, range = 40–89 years) AD families, using age at onset (rather than AD affection status) as the trait of interest (30). Linkage to 10q21.2 was not replicated by Sillen *et al*. (20) in a small sample of Swedish pedigrees (71 pedigrees, mean age at onset = 69 ± 6.8 years), nor by Pericak-Vance *et al*. (18) in a sample containing 286 NIMH and 118 NIA pedigrees, together with 62 pedigrees from CAP (Collaborative Alzheimer Project; Duke, UCLA, Vanderbilt, USA), (family mean age at onset >60 years, LOD < 1). As we do not have access to the entire raw data of Pericak-Vance *et al*. (18), it is impossible to say with certainty why, in an overlapping sample, the evidence for linkage to 10q21.2 differed. A combination of many factors could be the cause. The additional pedigrees used by Pericak-Vance *et al*. (18) could show no linkage, thus diluting the signal. The inclusion criteria of family mean onset of 60 years by Pericak-Vance *et al*. (18), rather than individual onset age used here, could reduce the power to detect linkage to a disease locus that affects those individuals with later onset. However, the Blacker *et al*. (17) analysis that was not constrained by onset age did show evidence for linkage to this region. Another possibility is that the genetic information content in the Pericak-Vance *et al*. (18) sample is less than that in the data analysis presented here. Lee *et al*. (21) found evidence of linkage to 10q at 138 cM (NPL = 2.02) with minimal linkage evidence to 10q21.2 (NPL < 1) in a sample of Caribbean-Hispanic pedigrees with predominantly LOAD (mean age at onset = 73.5 years). Interestingly, we observe some linkage evidence nearby (~130 cM) when interaction with chromosome 19 is taken into account. Thus, although there is good evidence for linkage to chromosome 10q, there is also likely linkage heterogeneity due to ethnic origin. Variation in age at onset and interaction with other loci may also influence linkage evidence.

The two linkages on chromosome 9, at 9p21.3 and 9q22.33, appear to reflect two susceptibility loci that act independently. Evidence exists for site of collection heterogeneity in the allele-sharing probabilities. In particular, the main peak on 9q22.33 appears to be contributed primarily by elevated allele sharing from the NIMH sample. Pericak-Vance *et al*. (18) identified a linkage peak close to the one we observed on 9p22.1 (43 versus 40 cM) in a sample that included 286 NIMH and 118 NIA pedigrees out of their total of 466 pedigrees, with a maximum LOD score of 2.97. In our analysis, linkage evidence to 9p21.3 comes mainly from the NIMH sites 50 and 51 and the NIA samples (although a formal test of

IBD differences across all five samples was non-significant). Since our sample overlaps with that of Pericak-Vance, this cannot be considered as evidence of replication. Blacker *et al*. report evidence for linkage at 55 cM (15 cM more centromeric, LOD = 1.3). Wijsman *et al*. and Sillen *et al*. found no linkage evidence to this region on chromosome 9, whereas Lee *et al*. report minimal evidence (NPL < 1).

The 9p22.1 region was the second strongest area of linkage in the Pericak-Vance study, second only to an analysis of the *APOE* locus itself. This group also obtained LOD scores in excess of 1 in the full sample at 4q32.1 (LOD = 1.30), 7q31.31 (LOD = 1.56) and 19q13.3 (LOD = 2.21), although the evidence decreased when restricting the analyses to families with at least one autopsy-confirmed AD case. We do not replicate these findings on 4q32.1 and 7q31.31. The linkage evidence on 9q22.33 was first identified in the NIMH sample by Kehoe *et al*. (14) and subsequently confirmed in overlapping late-onset samples (16,17,33). Again, in agreement with our results, Holmans *et al*. (32) found no evidence of an age-at-onset effect on linkage in AD cases with age at onset ≥65years. No evidence for linkage was reported on 9q22.33 by Pericak-Vance *et al*., Lee *et al*. or Sillen *et al*. Linkage evidence on chromosome 9 appears to show inter-sample heterogeneity.

We investigated the possibility that there are other linked variants in addition to *APOE*4 that contribute to the LOD score of 2.0 on 19q13.32, using the software LAMP (34). The test for other linked variants is not significant, although the power of this test is low for a LOD of this magnitude and when *APOE* is in such close proximity to the peak. Assuming that *APOE4* is the only linked variant in this region, the population attributable fraction associated with *APOE4* was estimated to be 0.36, with penetrances similar to an additive model.

The inclusion of *APOE* genotypes as covariates had no significant effects on the linkage analysis in any of the other linkage regions. This suggests that statistical interactions between *APOE* and genes in these regions do not have a large effect on risk of AD, although it does not preclude the possibility of biological interactions (35). We also included IBD sharing probabilities in the regions of interest as covariates in order to search for potential interaction effects on AD risk between disease susceptibility loci in these regions. A region on chromosome 10 (∼130 cM) was identified by this analysis when IBD probabilities on chromosome 19 were used as covariates. The region on chromosome 19 implicated in this interaction is at 38 cM from the pter, ∼35 cM from *APOE* itself. Although ARP analyses do not always give accurate estimates of disease-locus location (36), it seems unlikely that the interaction on 19p involves *APOE* (on 19q). This conclusion is reinforced by the lack of evidence for interaction when the *APOE* genotypes were used as covariates, despite the strong association of *APOE* with LOAD. Interestingly, the region on 19p showing interaction evidence to chromosome 10 is close to that identified by Wijsman *et al*. (30). They also concluded that linkage to this region on 19p did not reflect *APOE* effects.

We observed excellent genotype reliability between the three research groups. For diseases with a late age at onset, genome scans using mainly affected sibling pairs (ASPs) are common as parental DNAs are often unavailable. Genotyping error is difficult to detect in small pedigrees (e.g. one ASP) with no parental genotypes, as all genotypic configurations are consistent with Mendelian inheritance. The presence of random genotyping errors is likely to reduce the mean IBD allele sharing and hence the evidence for linkage. By removing genotypes that were discrepant between the three samples, we have produced a cleaner more powerful sample for analysis.

Despite the strength of the linkage finding on chromosome 10, it has not been possible to identify a susceptibility locus. Indeed, with the exception of *APOE*, no such risk variants have been identified for LOAD. There are number of possible reasons for this. First, a number of interacting susceptibility genes could reside within a linked region, of which the marginal effects are small and hence difficult to detect through traditional single marker tests of association. Second, AD shows a degree of clinical heterogeneity (37) which might relate to the underlying genetic architecture. It follows that difficulties in replicating genetic findings might result from sample differences and the field might benefit from more careful attention to phenotype definition. Third, evidence from the linkage analysis of Pericak-Vance *et al*. (18) also suggests the possibility of differences in allele-sharing estimates between clinical and neuropathological diagnoses of AD. Extended phenotypic data are not always available and sometimes inconsistently reported. Further analyses including these data will therefore only be possible following extensive data quality assessment, a subject of future work.

To conclude, our analyses show evidence for disease loci for LOAD on chromosome 10q (near 78 cM), chromosome 9q (near 104 cM), chromosome 19q (near 70 cM; likely to reflect the *APOE* gene) and, less convincingly, chromosome 9p (near 40 cM). However, there is evidence for linkage heterogeneity between samples, both directly, from our own analyses, and possibly indirectly, from the failures of other studies to replicate the linkages on chromosomes 9 and 10. This heterogeneity may also be influenced by other factors such as variation in age at onset and may explain the moderate evidence for association so far found to genes in these regions. Thus, there is a need for a large-scale analysis using all available LOAD linkage samples in order to ascertain which regions are most likely to contain AD susceptibility genes. These regions should then be subjected to detailed analysis in large association studies, using clinical covariates such as age at onset, to reduce the effects of heterogeneity.

MATERIALS AND METHODS

Sample

The data analysed were generated by three research groups. Group 1 genotyped pedigrees from the UK ($n = 82$), NIMH ($n = 212$) and NIA ($n = 72$) (14,16,25,32). Group 2 genotyped pedigrees from the NIMH ($n = 437$) (17,24). Group 3 also genotyped pedigrees from the NIMH ($n = 287$) with genotypes limited to chromosome 10 only (15,31). The three data sets have the NIMH sample in common, with 196 (64%) pedigrees and 516 (21%) individuals in this overall NIMH sample genotyped by all three groups. It should be noted that although

all three groups had access to the NIMH database of individuals, groups 2 and 3 had access to additional pedigrees that are not available on the NIMH database. In total, three collection sites contributed data to the NIMH database, and they are sites 50 (University of Alabama- Birmingham, AL, USA), 51 (Massachusetts General Hospital, MA, USA) and 52 (John Hopkins, MD, USA), all based in the USA. A summary of the number of pedigrees and ASPs from each sample is presented in Table 2. In addition to the ASPs, 57 non-sibling ARPs informative for linkage were also included in the analysis. To minimize heterogeneity, the samples were limited to individuals of Caucasian origin. The original analyses by groups 2 and 3 included pedigrees of non-Caucasian ancestry—these were excluded from the current analysis.

Table 2. The total number of ASPs (pedigrees) for each sample, split by site of origin

Sample	NIMH 50	NIMH 51	NIMH 52	UK	NIA	Total
(1)	82 (63)	69 (61)	136 (88)	109 (82)	90 (72)	486 (366)
(2)	121 (85)	82 (76)	157 (95)			360 (256)
Subtotal	130 (92)	86 (78)	171 (103)	109 (82)	90 (72)	586 (427)
(3)	154 (103)	100 (88)	209 (117)			463 (308)
Total	156 (105)	100 (88)	211 (119)	109 (82)	90 (72)	666 (466)

Diagnosis

Individuals were considered affected if they were diagnosed with definite or probable AD using the NINCDS-ADRDA (38) criteria, with an age at onset of at least 60 years. Previous presentations of the late-onset data have used a cut-off of 65 years. We relaxed this criterion to increase power, as we do not believe that this will include a significant proportion of familial autosomal dominant cases. Indeed, there was only one ARP with multiple reports of early onset AD cases (onset $<$ 65 years) in the same family. The original report by group 2 also included an analysis with onset ages of at least 50 years and an early/mixed onset analysis, and group 3 only presented analyses including cases with the full onset age range. The two groups also included some individuals with possible rather than probable or definite AD—these were also excluded from the current analysis.

Genotypic data

Group 1 first genotyped 237 microsatellite markers (average spacing = 20 cM) (14). The 16 peak regions with a maximum LOD score $>$1 were followed up with a further 91 markers (peak region average marker spacing = 10 cM) (16,25). Group 2 genotyped 381 microsatellite markers (average spacing = 9 cM). Group 3 genotyped a further 18 markers on chromosome 10 in the families ascertained at NIMH sites 50 and 52 (0–80 cM), plus an additional seven markers in the families from site 52 (0–54 cM). When the 20 markers genotyped in the NIMH families by groups 1 and 2 were taken into account, it gave a total of 45 markers on chromosome 10 (average spacing = 4 cM). Detailed information on the laboratory methods used to establish the microsatellite genotypes by each group are provided in their respective original publications (14,16,17,31). The genotypes from each group were provided in standard pedfile format, one file per chromosome. For each chromosome, one large pedfile was created by joining the original pedfiles for the given chromosome from each group, based on a standardized marker map comprising all markers. This map was derived from Marshfield (Marshfield Clinic, Marshfield genetic map, available at: www.marshfieldclinic.org/research/genetics), supplemented with information from Decode (39) and the UCSC Human Genome Browser (www.genome.ucsc.edu, May 2004 assembly, build 35) where necessary. Where the same genetic marker was genotyped by more than one group, each of the original markers remained as separate markers in the final pedfile, separated by an arbitrary distance of 0.001 cM. Individuals not genotyped at specific markers were coded as missing data, i.e. given a 0–0 genotype. This method of sample amalgamation allowed analysis of the complete data set and made no assumption that the allele numbering was consistent between groups.

Method of analysis

Given the availability of genotypic data, both within and between family relationships were confirmed statistically; see original publications for further information. Sample specific marker allele frequencies were estimated with SPLINK (40). Given that the data from all three groups comprised individuals from the NIMH sample, when the same genetic markers were genotyped by more than one group, we were able to check the reliability of the genotyping. As the genotypes from groups 2 and 3 were the same, quality control checks were performed to compare genotypes from group 1 with the combination of groups 2 and 3. Discrepant genotypes were removed from the analysis. One pedigree (53 bits) comprising three ASPs was trimmed to permit analysis within the constraints of modern computing power. Multipoint model-free ARP linkage analysis was performed to give likelihood ratio LOD scores for all chromosomes under the logistic regression model suggested by Rice (41), using IBD allele-sharing probabilities estimated at 2 cM intervals with MERLIN (42,43). Care was taken to consider marker–marker LD, most of which was imposed when the three overlapping samples were combined, i.e. the inclusion of duplicated markers, which were set to be 0.001 cM apart. Markers separated by $<$0.5 cM were grouped into clusters and estimated haplotype frequencies used in the IBD estimation to assume linkage disequilibrium within each cluster. For improved accuracy in regions of interest, the allele-sharing probabilities were estimated at 1 cM intervals. The levels of significant and suggestive evidence for linkage employed were as defined by Lander and Kruglyak (26).

Further tests were performed on chromosomes with evidence of linkage (LOD $>$ 1). Separate linkage analyses of the data from the three groups were performed. The results may differ from previous published analyses as the method of analysis and selection criteria of affected individuals differ. In addition, data quality checks removed some questionable and hence potentially influential genotypes. Homogeneity of allele sharing between sites of collection was investigated. Under the assumption that the parental alleles were inherited independently, the

probability of allele sharing IBD was modelled using logistic regression (41,44,45). Site of collection (i.e. UK, NIMH 50, 51, 52 or NIA) of each pair was included as an indicator variable. For each 2 cM position, a LOD score was obtained for the models (i) with and (ii) without the site of collection variables. Linkage peaks from standard analyses are often some distance from the disease loci (36); therefore, to allow for the multiple testing of more than one 2 cM position on the chromosome, the maximum LOD score for models (i) and (ii) were obtained over all positions analysed on the chromosome. The difference in these maximums (i–ii) was taken as the test statistic, $T_{obs} \geq 0$. A value of $T_{obs} = 0$ implies no difference in models (i) and (ii), i.e. there is no site of collection effect. A value of $T_{obs} > 0$ suggests some site of collection effect, the significance of which was assessed through simulations. The original site of collection data were permuted over all pedigrees and a value of T_{sim} for the simulated data set produced. This was repeated $n = 10\,000$ times. An estimate of the chromosome-wide significance of the test for homogeneity was assessed with $P = (r + 1)/(n + 1)$, where r is the number of times T_{sim} exceeded T_{obs}.

A similar test was performed in order to condition linkage by *APOE* genotype. Each affected individual was scored as − or +, indicating whether they carried at least one copy of the *APOE*4 allele (+) or not (−). Pairs of affected individuals were then constructed which could be either −/−, −/+ or +/+, where a −/− (+/+) pair indicates that neither (both) members of the pair carried at least one *APOE*4 allele, whereas a −/+ pair indicates that just one member had an *APOE*4 allele. The *APOE* covariate was entered into the regression model as two indicator variables. Constraints were applied to ensure that the allele sharing of the discordant pairs could not exceed that of the concordant pairs. This test used the same method as that presented in Myers *et al*. (16). As *APOE* resides on 19q, no *APOE* covariate test was performed on chromosome 19. Incorporating quantitative measures, such as age at onset and allele-sharing probabilities, into an analysis is also possible. Conditioning on allele-sharing probabilities allows us to investigate the possibility of interactions between two chromosomes. This method is considered to be less powerful than conditioning on the actual risk factor itself (e.g. the risk allele) (46), although until these risk factors are identified, conditioning on allele-sharing probabilities is considered to be a useful substitute. A test of the allele-sharing probabilities on chromosome I, conditional on J was performed. For completeness, a test for chromosome J, conditional on I was also performed, which gave similar results.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

Conflict of Interest statement. None.

FUNDING

This analysis was supported primarily by the Medical Research Council (UK) who support the Medical Research Council Co-operative group in Neuropsychiatric Genetics (M.L.H., M.J.O., J.W., M.O.D.), and a programme of AD research (J.W., M.J.O., M.O.D., P.H., S.L.). It was also supported by the Alzheimer's Research Trust (M.J.O., J.W.). The original data collection was supported by the Medical Research Council (UK) and by the National Institute on Aging, National Cell Repository for Alzheimer's Disease and National Institute of Mental Health Genetics Initiative for Alzheimer's Disease.

REFERENCES

1. Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske, A. and Pedersen, N.L. (2006) Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry*, **63**, 168–174.
2. Nussbaum, R.L. and Ellis, C.E. (2003) Alzheimer's disease and Parkinson's disease. *N. Engl. J. Med.*, **348**, 1356–1364.
3. Fratiglioni, L., De Ronchi, D. and Aguero-Torres, H. (1999) Worldwide prevalence and incidence of dementia. *Drugs Aging*, **15**, 365–375.
4. Goate, A., Chartier-Harlin, M.C., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L. *et al.* (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, **349**, 704–706.
5. Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D.M., Oshima, J., Pettingell, W.H., Yu, C.E., Jondro, P.D., Schmidt, S.D., Wang, K. *et al.* (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science*, **269**, 973–977.
6. Sherrington, R., Rogaev, E.I., Liang, Y., Rogaeva, E.A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K. *et al.* (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**, 754–760.
7. Pedersen, N.L., Gatz, M., Berg, S. and Johansson, B. (2004) How heritable is Alzheimer's disease late in life? Findings from Swedish twins. *Ann. Neurol.*, **55**, 180–185.
8. Bales, K.R., Dodart, J.C., DeMattos, R.B., Holtzman, D.M. and Paul, S.M. (2002) Apolipoprotein E, amyloid, and Alzheimer disease. *Mol. Interv.*, **2**, 363–375.
9. Henderson, A.S., Easteal, S., Jorm, A.F., Mackinnon, A.J., Korten, A.E., Christensen, H., Croft, L. and Jacomb, P.A. (1995) Apolipoprotein E allele epsilon 4, dementia, and cognitive decline in a population sample. *Lancet*, **346**, 1387–1390.
10. Martins, R.N., Clarnette, R., Fisher, C., Broe, G.A., Brooks, W.S., Montgomery, P. and Gandy, S.E. (1995) ApoE genotypes in Australia: roles in early and late onset Alzheimer's disease and Down's syndrome. *Neuroreport*, **6**, 1513–1516.
11. Kukull, W.A., Schellenberg, G.D., Bowen, J.D., McCormick, W.C., Yu, C.E., Teri, L., Thompson, J.D., O'Meara, E.S. and Larson, E.B. (1996) Apolipoprotein E in Alzheimer's disease risk and case detection: a case-control study. *J. Clin. Epidemiol.*, **49**, 1143–1148.
12. Daw, E.W., Payami, H., Nemens, E.J., Nochlin, D., Bird, T.D., Schellenberg, G.D. and Wijsman, E.M. (2000) The number of trait loci in late-onset Alzheimer disease. *Am. J. Hum. Genet.*, **66**, 196–204.
13. Bertram, L., McQueen, M.B., Mullin, K., Blacker, D. and Tanzi, R.E. (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet.*, **39**, 17–23.
14. Kehoe, P., Wavrant-De Vrieze, F., Crook, R., Wu, W.S., Holmans, P., Fenton, I., Spurlock, G., Norton, N., Williams, H., Williams, N. *et al.* (1999) A full genome scan for late onset Alzheimer's disease. *Hum. Mol. Genet.*, **8**, 237–245.
15. Bassett, S.S., Avramopoulos, D., Perry, R.T., Wiener, H., Watson, B., Jr, Go, R.C. and Fallin, M.D. (2006) Further evidence of a maternal parent-of-origin effect on chromosome 10 in late-onset Alzheimer's disease. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **141B**, 537–540.
16. Myers, A., Wavrant De-Vrieze, F., Holmans, P., Hamshere, M., Crook, R., Compton, D., Marshall, H., Meyer, D., Shears, S., Booth, J. *et al.* (2002) Full genome screen for Alzheimer disease: stage II analysis. *Am. J. Med. Genet.*, **114**, 235–244.
17. Blacker, D., Bertram, L., Saunders, A.J., Moscarillo, T.J., Albert, M.S., Wiener, H., Perry, R.T., Collins, J.S., Harrell, L.E., Go, R.C. *et al.* (2003) Results of a high-resolution genome screen of 437 Alzheimer's disease families. *Hum. Mol. Genet.*, **12**, 23–32.

18. Pericak-Vance, M.A., Grubber, J., Bailey, L.R., Hedges, D., West, S., Santoro, L., Kemmerer, B., Hall, J.L., Saunders, A.M., Roses, A.D. *et al.* (2000) Identification of novel genes in late-onset Alzheimer's disease. *Exp. Gerontol.*, **35**, 1343–1352.
19. Rademakers, R., Cruts, M., Sleegers, K., Dermaut, B., Theuns, J., Aulchenko, Y., Weckx, S., De Pooter, T., Van den Broeck, M., Corsmit, E. *et al.* (2005) Linkage and association studies identify a novel locus for Alzheimer disease at 7q36 in a Dutch population-based sample. *Am. J. Hum. Genet.*, **77**, 643–652.
20. Sillen, A., Forsell, C., Lilius, L., Axelman, K., Bjork, B.F., Onkamo, P., Kere, J., Winblad, B. and Graff, C. (2006) Genome scan on Swedish Alzheimer's disease families. *Mol. Psychiatry*, **11**, 182–186.
21. Lee, J.H., Mayeux, R., Mayo, D., Mo, J., Santana, V., Williamson, J., Flaquer, A., Ciappa, A., Rondon, H., Estevez, P. *et al.* (2004) Fine mapping of 10q and 18q for familial Alzheimer's disease in Caribbean Hispanics. *Mol. Psychiatry*, **9**, 1042–1051.
22. Bertram, L., Blacker, D., Mullin, K., Keeney, D., Jones, J., Basu, S., Yhu, S., McInnis, M.G., Go, R.C., Vekrellis, K. *et al.* (2000) Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. *Science*, **290**, 2302–2303.
23. Ertekin-Taner, N., Graff-Radford, N., Younkin, L.H., Eckman, C., Baker, M., Adamson, J., Ronald, J., Blangero, J., Hutton, M. and Younkin, S.G. (2000) Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. *Science*, **290**, 2303–2304.
24. Li, Y.J., Scott, W.K., Hedges, D.J., Zhang, F., Gaskell, P.C., Nance, M.A., Watts, R.L., Hubble, J.P., Koller, W.C., Pahwa, R. *et al.* (2002) Age at onset in two common neurodegenerative diseases is genetically controlled. *Am. J. Hum. Genet.*, **70**, 985–993.
25. Myers, A., Holmans, P., Marshall, H., Kwon, J., Meyer, D., Ramic, D., Shears, S., Booth, J., DeVrieze, F.W., Crook, R. *et al.* (2000) Susceptibility locus for Alzheimer's disease on chromosome 10. *Science*, **290**, 2304–2305.
26. Lander, E. and Kruglyak, L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.*, **11**, 241–247.
27. Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., Myers, R.H., Pericak-Vance, M.A., Risch, N. and van Duijn, C.M. (1997) Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*, **278**, 1349–1356.
28. Liang, X., Schnetz-Boutaud, N., Kenealy, S.J., Jiang, L., Bartlett, J., Lynch, B., Gaskell, P.C., Gwirtsman, H., McFarland, L., Bembe, M.L. *et al.* (2006) Covariate analysis of late-onset Alzheimer disease refines the chromosome 12 locus. *Mol. Psychiatry*, **11**, 280–285.
29. Mayeux, R., Lee, J.H., Romas, S.N., Mayo, D., Santana, V., Williamson, J., Ciappa, A., Rondon, H.Z., Estevez, P., Lantigua, R. *et al.* (2002) Chromosome-12 mapping of late-onset Alzheimer disease among Caribbean Hispanics. *Am. J. Hum. Genet.*, **70**, 237–243.
30. Wijsman, E.M., Daw, E.W., Yu, C.E., Payami, H., Steinbart, E.J., Nochlin, D., Conlon, E.M., Bird, T.D. and Schellenberg, G.D. (2004) Evidence for a novel late-onset Alzheimer disease locus on chromosome 19p13.2. *Am. J. Hum. Genet.*, **75**, 398–409.
31. Bassett, S.S., Avramopoulos, D. and Fallin, D. (2002) Evidence for parent of origin effect in late-onset Alzheimer disease. *Am. J. Med. Genet.*, **114**, 679–686.
32. Holmans, P., Hamshere, M., Hollingworth, P., Rice, F., Tunstall, N., Jones, S., Moore, P., Wavrant DeVrieze, F., Myers, A., Crook, R. *et al.* (2005) Genome screen for loci influencing age at onset and rate of decline in late onset Alzheimer's disease. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **135**, 24–32.
33. Perry, R.T., Wiener, H., Harrell, L.E., Blacker, D., Tanzi, R.E., Bertram, L., Bassett, S.S. and Go, R.C. (2007) Follow-up mapping supports the evidence for linkage in the candidate region at 9q22 in the NIMH Alzheimer's disease Genetics Initiative cohort. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **144**, 220–227.
34. Li, M., Boehnke, M. and Abecasis, G.R. (2005) Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *Am. J. Hum. Genet.*, **76**, 934–949.
35. Cordell, H.J. (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.*, **11**, 2463–2468.
36. Cordell, H.J. (2001) Sample size requirements to control for stochastic variation in magnitude and location of allele-sharing linkage statistics in affected sibling pairs. *Ann. Hum. Genet.*, **65**, 491–502.
37. Hollingworth, P., Hamshere, M.L., Moskvina, V., Dowzell, K., Moore, P.J., Foy, C., Archer, N., Lynch, A., Lovestone, S., Brayne, C. *et al.* (2006) Four components describe behavioral symptoms in 1,120 individuals with late-onset Alzheimer's disease. *J. Am. Geriatr. Soc.*, **54**, 1348–1354.
38. McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D. and Stadlan, E.M. (1984) 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*, **34**, 939–944.
39. Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G. *et al.* (2002) A high-resolution recombination map of the human genome. *Nat. Genet.*, **31**, 241–247.
40. Holmans, P. and Clayton, D. (1995) Efficiency of typing unaffected relatives in an affected-sib-pair linkage study with single-locus and multiple tightly linked markers. *Am. J. Hum. Genet.*, **57**, 1221–1232.
41. Rice, J.P. (1997) The role of meta-analysis in linkage studies of complex traits. *Am. J. Med. Genet.*, **74**, 112–114.
42. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
43. Abecasis, G.R. and Wigginton, J.E. (2005) Handling marker–marker linkage disequilibrium: pedigree analysis with clustered markers. *Am. J. Hum. Genet.*, **77**, 754–767.
44. Rice, J.P. (2001) Diagnosis as a covariate in sib-pair linkage analysis. *Am. J. Med. Genet.*, **105**, 55–56.
45. Hamshere, M.L., Macgregor, S., Moskvina, V., Nikolov, I.N. and Holmans, P.A. (2005) Covariate linkage analysis of GAW14 simulated data incorporating subclinical phenotype, sex, population, parent-of-origin, and interaction. *BMC Genet.*, **6** (Suppl 1), S45.
46. Holmans, P. (2002) Detecting gene-gene interactions using affected sib pair analysis with covariates. *Hum. Hered.*, **53**, 92–102.