We have genotyped 292 affected sibling pairs (ASPs) with Alzheimer’s disease (AD) according to NINCDS–ADRDA diagnostic criteria and with onset ages of ≥65 years using 237 microsatellite markers separated by an average distance of 16.3 cM. Data were analysed by SPLINK and MAPMAKER/SIBS on the whole sample of 292 ASPs and subsets of 162 ASPs where both members possessed an apolipoprotein E (APOE) ε4 allele and 63 pairs where neither possessed an ε4 allele. Sixteen peaks with a multipoint lod score (MLS) >1 either in the whole sample, the ε4-positive or -negative subgroups were observed on chromosomes 1 (two peaks), 2, 5, 6, 9 (two peaks), 10 (two peaks), 12, 13, 14, 19, 21 and X (two peaks). Simulation studies revealed that these findings exceeded those expected by chance, although many are likely to be false positives. The highest lod scores on chromosomes 1 (MLS 2.67), 9 (MLS 2.38), 10 (MLS 2.27) and 19 (MLS 1.79) fulfill Lander and Kruglyak’s definition of ‘suggestive’in that they would be expected to occur by chance once or less per genome scan. Several other peaks were only marginally less significant than this, in particular those on chromosomes 14 (MLS 2.16), 5 (MLS 2.00), 12, close to α2-macroglobulin (MLS 1.91), and 21, close to amyloid precursor protein (MLS 1.77). This is the largest genome scan to date in AD and shows for the first time that this is a genetically complex disorder involving several, perhaps many, genes in addition to APOE. Moreover, our data will be of interest to those hoping to identify positional candidate genes using information emerging from neurobiological studies of AD.

INTRODUCTION

Alzheimer’s disease (AD) is a progressive neurodegenerative disorder that occurs predominantly in later life. It is the commonest cause of dementia and represents the fourth largest cause of death in the developed world (1).

To date, four genes have been implicated in the aetiology of AD. Mutations in three of these, coding for the amyloid precursor protein (APP) (2), presenilin 1 (PS-1) (3,4) and presenilin 2 (5) account for most cases of autosomal dominant familial AD (FAD) (6). However, FAD accounts for <1% of all cases of AD. The inheritance of common forms of the disorder appears considerably more complex and probably reflects the co-action or interaction of several or many genes together with environmental factors. One gene that is clearly implicated in this form of the disorder is that encoding apolipoprotein E (APOE). The ε4 allele of APOE, although neither necessary nor sufficient to cause AD, is associated with increased risk of both early and late onset disease (7). The effect of ε4 appears additive such that heterozygotes and homozygotes are, respectively, approximately three and eight times more likely to be affected than controls (7). However, variation at the APOE locus accounts for at most ∼50% of the genetic variation in liability (8) to develop the disorder and there must be other genetic variants that account for the remaining risk.

A number of strategies are available for mapping genetically complex traits (9). Traditional lod score analysis in multiplex pedigrees is best suited to forms of a disorder showing obvious Mendelian inheritance. This approach is clearly applicable to FAD, where successes have already been evident. However, the
lod score method is at its weakest when the mode of transmission is complex and the genetic parameters cannot be accurately specified. It is generally agreed that the best way to proceed under such circumstances is by a combination of allele sharing linkage methods in pairs of affected siblings or other relative pairs and association studies (9, 10). A number of candidate gene association studies have been performed in AD since the identification of the APOE locus. Some positive findings have been claimed (11–30) but unfortunately none of these has been consistently confirmed. These inconsistencies are likely to be due to a number of factors that bedevil genetic association studies, including heterogeneity, issues of statistical power, multiple testing and population stratification (31). Moreover, at the present time, association studies can only be based on testing of genes whose candidature is suggested by an existing understanding of the pathophysiology. In contrast, a systematic genome screen using allele sharing linkage methods offers the potential to identify novel pathogenic pathways and mechanisms. Consequently we have carried out a two-stage genome scan in 600 affected sibling pairs (ASPs) with AD. In the first phase we have genotyped 292 ASPs with a 20 cM grid of markers and report the results here.

RESULTS

A multipoint ASP analysis was performed on the entire dataset of 292 ASPs using MAPMAKER/SIBS (32). The sample was also stratified for analysis on the basis of whether both (162 ASPs) or neither (63 ASPs) members of an ASP possessed at least one APOE ε4 allele in order to maximize power to detect loci acting epistatically or heterogeneously with respect to APOE. APOE genotypes were not included in lod score calculations. This is a computationally simple, model-free approach to two-locus analysis which does not require the specification of unknown parameters such as gene frequencies, penetrances and interactions and which allows the multilocus approach implemented in MAPMAKER/SIBS to be used. The resulting multipoint lod scores (MLS) are shown in Figure 1, together with the regions where exclusion analysis based on a λs of 1.4 gave lod scores of ≥2 or less. A genetic effect of size λs = 1.4 was chosen since this is approximately equal to that given by APOE (see below). This could be excluded from ~28% of the genome, whereas a λs of 2 could be excluded from ~80% of the genome (data not shown).

The multipoint results are also summarized in Table 1, which shows data from the 16 peaks with an MLS ≥1 in either the whole sample, the ε4-positive (ε4+ve) or ε4-negative (ε4–ve) subgroups. These regions were observed on chromosomes 1 (two peaks), 2, 5, 6, 9 (two peaks), 10 (two peaks), 12, 13, 14, 19, 21 and X (two peaks). Only peaks on chromosomes 1, 5, 9, 10 and 19 gave a MLS >1 in the whole sample. The remaining peaks were observed in either the ε4–ve (chromosomes 1, 10, 12, 21 and X) or the ε4+ve (chromosomes 2, 6, 13 and 14) sib pairs only. Pointwise and genome-wide significance levels calculated by simulation for all three samples analysed are shown in Table 2. It can be seen that the number of observed regions (i.e. MLSs) exceeding a given lod score is greater than would be expected by chance. Furthermore, the highest lod scores on chromosomes 1 (MLS 2.67), 9 (MLS 2.38), 10 (MLS 2.27) and 19 (MLS 1.79) fulfill Lander and Kruglyak’s definition of ‘suggestive’ in that they would be expected to occur by chance once or less per genome scan (32). Several other peaks were only marginally less significant than this, in particular those on chromosomes 14 (MLS 2.16), 5 (MLS 2.00), 12 (MLS 1.91) and 21 (MLS 1.77). We did not observe any ‘significant’ linkages (i.e. genome-wide occurrence probability <0.05) since our simulation results indicate that a lod score of ~3.6 would be required for this (Table 2).

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Whole sample</th>
<th>IBD Both ε4+ve</th>
<th>IBD Both ε4–ve</th>
<th>IBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>0.02</td>
<td>0.00</td>
<td>1.35</td>
<td>0.67</td>
</tr>
<tr>
<td>1B</td>
<td>1.33</td>
<td>0.56</td>
<td>2.67</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>0.49</td>
<td>1.01</td>
<td>0.57</td>
<td>0.36</td>
</tr>
<tr>
<td>5</td>
<td>1.07</td>
<td>0.55</td>
<td>2.00</td>
<td>0.61</td>
</tr>
<tr>
<td>6</td>
<td>0.48</td>
<td>1.40</td>
<td>0.60</td>
<td>0.85</td>
</tr>
<tr>
<td>9A</td>
<td>0.00</td>
<td>0.00</td>
<td>1.12</td>
<td>0.64</td>
</tr>
<tr>
<td>9B</td>
<td>1.15</td>
<td>0.56</td>
<td>2.38</td>
<td>0.62</td>
</tr>
<tr>
<td>10A</td>
<td>2.27</td>
<td>0.60</td>
<td>2.25</td>
<td>0.61</td>
</tr>
<tr>
<td>10B</td>
<td>0.2</td>
<td>0.00</td>
<td>1.17</td>
<td>0.63</td>
</tr>
<tr>
<td>12</td>
<td>0.89</td>
<td>0.20</td>
<td>1.91</td>
<td>0.63</td>
</tr>
<tr>
<td>13</td>
<td>0.20</td>
<td>1.20</td>
<td>0.58</td>
<td>0.00</td>
</tr>
<tr>
<td>14</td>
<td>0.30</td>
<td>2.16</td>
<td>0.60</td>
<td>0.00</td>
</tr>
<tr>
<td>19</td>
<td>1.79</td>
<td>0.57</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>21</td>
<td>0.25</td>
<td>0.50</td>
<td>1.77</td>
<td>0.65</td>
</tr>
<tr>
<td>XA</td>
<td>0.10</td>
<td>0.05</td>
<td>1.45</td>
<td>0.72</td>
</tr>
<tr>
<td>XB</td>
<td>0.05</td>
<td>0.20</td>
<td>1.93</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Data from the 16 peaks with an MLS ≥1 in either the whole sample, the ε4+ve or ε4–ve subgroups are shown.

DISCUSSION

In a genome screen of 292 sib pairs with late onset AD, we observed 16 loci (i.e. MLSs) with lod scores ≥1, which exceeded the number expected by chance. The regions of interest occur on chromosomes 1 (two peaks), 2, 5, 6, 9 (two peaks), 10 (two
The Duke group, also using non-parametric linkage detected APP polymorphisms that predispose to late onset AD. This study: APOE, APP and AD. Only three of these are located within the peaks identified in genes' for which evidence exists implicating their involvement in the region. It is well known that mutations in APP can cause in typical late onset AD and this might indicate genetic variability in APP expression (40). In addition, recent genetic analysis of a case of Down syndrome (DS) due to non-disjunction has firmly implicated triplication of APP in the pathogenesis of AD in DS (41). Together with our data, these findings suggest that a full genetic analysis of the APP gene should now be a priority. A2M is a serum pan-protease inhibitor which has been implicated in AD because of its ability to mediate the clearance and degradation of βA (42,43) and because, like ApoE, it is a ligand of low density lipoprotein receptor-related peptide (44). Blacker et al. (29) have recently demonstrated an association between a pentanucleotide deletion in the 5’ splice site of exon 18 of A2M and AD using a subset of the NIMH sample typed in the present study. This association appeared independent of APOE genotype and this would appear to correlate well with our observation that linkage to the A2M region was seen only in ε4-ve families. Our data on chromosome 12 have been reported separately (35), but we are now able to place them in the context of a full genome scan. The maximum MLS of 1.91 in this region is less than the threshold calculated by simulation of 2.25 for suggestive linkage according to Lander and Kruglyak (32), but was the second highest score obtained in the ε4-ve families. None of the other candidate genes shown in Figure 1 are located within the lod score peaks in our study. However, the power of linkage methods to detect genes of small effect is limited (45). Indeed only PS-1 and α1-antichymotrypsin fall within the 2-2.5 lod score threshold for sufficient effect size equivalent to or greater than APOE (λc = 1.4).

This study is the first stage of a two-stage genome scan for late onset AD in 600 sibling pairs. The 16 areas identified will form the basis of analysis in stage 2. This will comprise further analysis of these areas using the original and new markers spaced at ~5 cM intervals within each region in a total of 600 sibling pairs with late onset Alzheimer’s disease. We have calculated that this two-stage study will have a power of >0.80 to detect a locus of effect size λc ≥ 1.5. Regions showing evidence of linkage at stage 2 will be tested further using positional cloning and candidate gene approaches.

This is the largest genome scan to date in Alzheimer’s disease and shows for the first time that this is a genetically complex disorder involving several, perhaps many, genes in addition to APOE. We have provided evidence consistent with a role of A2M as a susceptibility locus and preliminary data suggesting that variation in the regulatory regions of A2M may also influence predisposition to AD. The other areas identified in our study, in particular those on chromosomes 1, 9 and 10, will be of considerable interest to those trying to identify positional candidate genes using information emerging from neurobiological studies of AD.

**MATERIALS AND METHODS**

**Families**

The families used were selected from those collected by the NIMH–AD Genetic Consortium (46). From within this family series 230 families were selected based on the following criteria: at least two affected siblings with probable or definite AD according to the NINCDS–ADRDA diagnostic criteria (47) with onset ages of ≥65 years sampled and available for genotyping (diagnosis of definite AD requires neuropathological data which are usually obtained post-mortem). Within these families the probable (n = 417) or definitely affected (n = 79) siblings were genotyped. In families where there were more than two affected siblings (definite or probable) available, all of them were used. In families where there were just two affected siblings but unaffected individuals were available, the oldest of the unaffected individuals was also sampled so that the genotype data from this unaffected

peaks), 12, 13, 14, 19, 21 and X (two peaks). The chromosome 19 result appears attributable to the APOE locus, with marker D19S571 (15 cM distal of APOE) showing a lod score of 1.79. This complies with Lander and Kruglyak’s (32) definition of ‘suggestive’ linkage and indicates that our study would have detected a susceptibility gene for AD in this region had association with APOE ε4 not already been reported. In addition we obtained a single point lod score of 2.02 (IBD 0.64) with marker D19S412 (2 cM distal of the APOE locus).

Recently Pericak-Vance et al. (33,34) reported the results of a two-stage genome screen in late onset FAD in 54 families with multiple affected members. This first stage was carried out in 16 of the largest families, in which DNA was available from 52 affected individuals. Fifteen regions of interest were followed up in a further 38 families, in which DNA was available from 89 affected individuals. Interestingly, no evidence for linkage around the APOE locus was obtained despite the enrichment of markers in this region. On follow-up, four regions were identified showing possible or suggestive linkage on chromosomes 4, 6, 12 and 20, with the strongest evidence being found in the pericentromeric region of chromosome 12 (peak MLS 3.5). We found little evidence to support these findings on chromosome 12 in our dataset (35), although we did find some evidence for linkage on 12p (see below). Unfortunately, Pericak-Vance et al. provide no information on the specific locations of the other areas of interest on chromosomes 4, 6 and 20 (33,34). We did not observe evidence of linkage on chromosomes 4 or 20 but did find a lod score of 1.4 in ε4+ve ASPs on chromosome 6. However, while we cannot assess whether our chromosome 6 region of interest corresponds to that of Pericak-Vance et al. (34), it is of interest that it contains much of the HLA region and that associations between AD and HLA-A2 have been reported (20). From initial findings of a separate genome survey for AD susceptibility genes, Zubenko et al. (36) also reported a region of interest on the X chromosome, near DXS1047. No support for linkage to this region was obtained in our study. Indeed, we could exclude a locus of λc ≥ 1.4 from this region.

Figure 1 also shows the positions of a number of ‘candidate genes’ for which evidence exists implicating their involvement in AD. Only three of these are located within the peaks identified in this study: APOE, APP and α2-macroglobulin (A2M). APOE is the only unequivocally established genetic risk factor for late onset AD and it is therefore reassuring that we were able to obtain evidence for linkage at this locus without enriching for markers in the region. It is well known that mutations in APP can cause autosomal dominant early onset familial AD (2). Direct analyses of the coding sequence (37) and the promoter (37,38) have not detected APP polymorphisms that predispose to late onset AD. However, the Duke group, also using non-parametric linkage methods, have provided evidence that a locus predisposing to late onset AD might reside in this region of chromosome 21 (39). Our findings suggest that further attention should now be paid to the possibility of a susceptibility locus either within the regulatory regions of APP or, less parsimoniously, in a gene close by. It is therefore of interest that high levels of βA are sometimes found in typical late onset AD and this might indicate genetic variability in APP expression (40). In addition, recent genetic analysis of a case of Down syndrome (DS) due to non-disjunction has firmly implicated triplication of APP in the pathogenesis of AD in DS (41). Together with our data, these findings suggest that a full genetic analysis of the APP gene should now be a priority. A2M
Figure 1. Multipoint lod scores.
Figure 1. Continued
Figure 1. Continued
Figure 1. Continued
individual could be used to check for genotyping errors. These 230 families yielded a total of 292 ASPs, in only 17 of which were both definitively affected.

**Genotyping**

Lymphoblastoid cell lines were generated from peripheral blood leukocytes and DNA extracted using routine methods. Semi-automated fluorescent genotyping was undertaken using the ABI Genescan/Genotyper system by comparison of the fragment sizes with an internal standard. All genotypes were scored blind as to phenotype. Genotype data from all three participating centres was sent to a centralized database in Cardiff (48). The database (MEGABASE) was used to check the binning of alleles, convert allele sizes to whole numbers and to check for non-Mendelian inheritance where possible. MEGABASE stored all relevant genotypic/phenotypic data and produced all necessary files for statistical analysis.

**Markers**

A total of 237 microsatellite markers, obtained from Genethon, CHLC and GDB (106 tri- and 110 tetranucleotides) were typed in an average of 253.8 (159–288) sib pairs. The average distance between adjacent markers was 16.3 cM (1–40 cM) and the average heterozygosity was 0.75 (0.39–0.89).

**Statistical analysis**

The program SPLINK (49) was used to compute single point MLSs under the ‘possible triangle’ restrictions (50) and to calculate allele frequency estimates for each marker. These were used in the multipoint analyses, which were carried out using MAPMAKER/SIBS (32) on the whole sample of 292 sib pairs, the subset of 162 sib pairs where both members possessed an ε4 allele and the subset of 63 pairs where neither possessed an ε4 allele. A multipoint exclusion map was also obtained for the whole sample using MAPMAKER/SIBS. For the purposes of this analysis, the disease susceptibility model was parameterized in terms of λε, the relative risk to siblings of a case (51). A λε of 1.4 was used, since this is approximately equal to that given by APOE, assuming a population gene frequency of 0.15, a relative risk of 4 for ε4 heterozygotes and a relative risk of 10 for ε4 homozygotes, similar to observations in a number of Caucasian populations (7). Higher values of λε were also tested, although the results are not shown here. Genome-wide significance levels were obtained by simulation in the following way: 500 replicate samples of each chromosome were simulated under the null hypothesis of no linkage, using the observed allele frequencies and ensuring that the individuals typed at each locus were the same as in the original dataset. Stratification into identity by descent; PS-1, presenilin 1.

**ACKNOWLEDGEMENTS**

The samples used in this study were selected from those collected by the NIMH Alzheimer’s Genetics Initiative and were banked at the Coriell Cell Repository. Full sample IDs are available from the authors. Genotyping and data analysis were supported by a grant from the MRC (UK), the Alzheimer’s Association, the Metropolitan Life Foundation, the Mayo Foundation and the Mayo/USF Program Project Grant. A.G. is the recipient of an NIH career development award (AG00634).

**ABBREVIATIONS**

A2M, α2-macroglobulin; AD, Alzheimer’s disease; ASPs, affected sibling pairs; APP, amyloid precursor protein; APOE, apolipoprotein E; DS, Down syndrome; ε4+, ε4-positive; ε4–, ε4-negative; FAD, familial autosomal dominant AD; IBD, identity by descent; PS-1, presenilin 1.

**REFERENCES**


16. Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T., Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T., Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T., Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T., Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T., Matsushita, S., Arai, H., Muramatsu, T., Makimoto, K., Nakagawa, T.


