Identification of multiple loci for Alzheimer disease in a consanguineous Israeli–Arab community

Lindsay A. Farrer1–5,* Abdalla Bowirrat1, Robert P. Friedland7, Kristin Waraska6, Amos D. Korczyn8 and Clinton T. Baldwin1,6

1Department of Medicine (Genetics Program), 2Department of Neurology, 3Department of Genetics and Genomics, 4Department of Epidemiology and 5Department of Biostatistics, and 6Center for Human Genetics, Boston University School of Medicine and Public Health, Boston, MA 02118, USA, 7Department of Neurology, Case Western Reserve University, Cleveland, OH 44106, USA and 8Department of Neurology, Tel Aviv University, Tel Aviv, Israel

Received October 30, 2002; Revised and Accepted December 16, 2002

We have observed an unusually high prevalence of dementia of the Alzheimer type (DAT) in Wadi Ara, an inbred Arab community in northern Israel comprising ~850 persons over the age of 60 years. Family studies revealed that more than one-third of the DAT cases are members of one hamula (tribal group) within Wadi Ara. To map chromosomal loci contributing to DAT susceptibility, we conducted a 10 cM scan in a series of five cases and five controls selected from this hamula. Markers from 18 chromosomal regions showed significant allelic association with DAT ($P < 0.05$). Locations on chromosomes 2, 9 and 10 remained significant after testing additional affected and non-demented individuals. Significant associations were also observed for markers on chromosome 12 which overlap with a locus implicated in previous genome scans. Analysis of allele frequency distributions for 12 markers spanning 20 cM on chromosome 9 narrowed the possible location of an DAT susceptibility gene to a 13 cM interval between D9S157 and D9S259 (most significant result: $P = 2.3 \times 10^{-7}$). Analysis of 14 markers spanning 24 cM on chromosome 12 narrowed the possible location to a 14 cM interval distal to the LRP1 locus (most significant result: $P = 1.3 \times 10^{-6}$). Evidence for linkage on chromosome 9 stemmed primarily from excess homozygosity of marker alleles in cases compared with controls, suggesting that the gene at this location behaves in either a recessive or additive fashion. The unique characteristics of this community together with the emergent human genome data should allow for the rapid identification of DAT genes in these candidate regions.

INTRODUCTION

Alzheimer’s disease (AD) is a progressive, neurodegenerative disease characterized by memory loss, language deterioration, impaired visuo-spatial skills and poor judgment, and neuropathologically by extracellular amyloid plaque deposition intra neuronal neurofibrillary tangles and neuronal loss. AD usually begins after age 65 however, its onset may occur as early as age 30. The early-onset forms of the disease appear to have a strong genetic component, and mutations in several genes have been identified including the amyloid precursor gene (APP; 104760), presenilin-1 (PSEN1; 104311), and presenilin-2 (PSEN2; 600759).

The basis of the more common, late-onset form of AD is less well understood. The genetic risk factor with the highest attributable risk for AD is the $\varepsilon4$ allele of the APOE gene on chromosome 19q (1). Among Caucasians, the odds of AD for $\varepsilon4$ homozygotes and for $\varepsilon3/\varepsilon4$ heterozygotes are 14.9 and 3.2, respectively, greater than that associated with $\varepsilon3$ homozygotes (2). Many other genes have been studied because of the possible role of their gene products in the disease, or because of their proximity to known AD loci. The evidence implicating other susceptibility genes is much less compelling. Associations have been reported for no fewer than 40 genes (3,4), yet none has been clearly established as an AD risk factor. Linkage studies have implicated regions on chromosomes 9 (5), 10 (6–8), and 12 (9–11). Candidate genes in these regions include insulin-degrading enzyme (IDE) on chromosome 10 and $\varepsilon2$-macroglobulin (A2M) and low-density lipoprotein receptor-related gene (LRP1) on chromosome 12 (8,12,13), although none has been proven to be an AD gene (14–17).

*To whom correspondence should be addressed at: Genetics Program L-320, Boston University School of Medicine, 715 Albany Street, Boston, MA 02118, USA. Tel: +1 617 6385393; Fax: +617 6384275; Email: farrer@bu.edu
In this report, we examined the genetic basis of dementia of the Alzheimer type (DAT) in a unique community. In a population based study, we screened all residents ages 60 years and older (n = 821) of Wadi Ara, an Arab community near Tel-Aviv in northern Israel and observed a prevalence higher than that found in Israel, China, Europe or the USA, even after adjustment for age, education and gender (18,19). This unusually high prevalence is apparently not due to an increased frequency of the APOE e4 allele, which is actually reduced in this community (0.02 for non-demented elders) as compared with other Caucasians (20). We hypothesized that the increased prevalence of DAT in Wadi Ara is associated with a limited number of unique susceptibility alleles whose frequency was enriched by inbreeding. The current residents of Wadi Ara are members of ~14 hamulas (family groups). Until recently, there has been minimal immigration or emigration from the community. We designed and carried out an efficient 10 cM genome scan by exploiting the unique population structure. Results of this study confirm the existence of DAT loci on chromosomes 9, 10 and 12 and provide evidence for a new DAT locus on chromosome 2.

RESULTS

Genome scan

The high prevalence of DAT in Wadi Ara may, of course, be explained by environmental or genetic factors (or both). The genetic hypothesis was bolstered by examination of 187 pedigrees showing a smaller number of surnames in those families containing DAT cases than in those without DAT cases. In fact, more than one-third of the 168 prevalent DAT cases were from one of the 14 hamulas. Despite the large concentration of DAT cases in this hamula, few individual pedigrees contained more than one living affected member. Therefore, traditional family-based linkage study approaches were not feasible. However, given the high degree of inbreeding in Wadi Ara evidenced by frequent first-cousin and second-cousin marriages, and the limited number of founders, we reasoned that prevalent DAT susceptibility alleles may be identical by descent and, therefore, frequency distributions of alleles or genotypes for markers linked to a DAT locus would differ between DAT cases and controls. A 10 cM genome scan was conducted by genotyping 375 markers from the ABI genome scan panel in DNAs from five DAT cases (ages 68–85 years) and five non-demented controls (ages 63–85 years) from different families in the hamula with the highest disease prevalence. Markers at 18 linkage regions on 13 different chromosomes showed a significant difference in the allele frequency distribution between DAT cases and controls (Table 1). In many instances, the difference was not attributable to a specific allele. There was no evidence for either excess marker homozygosity in cases compared with controls or association at the genotype level.

Follow-up studies

Markers whose allele frequency distribution differed significantly (P < 0.05) between the cases and controls were genotyped in an enlarged sample of 100 DAT cases and 110 controls selected from the entire cohort. Results which remained significant in the enlarged sample were pursued by analysis of additional markers spanning the positive signals. Significant allelic association remained for locations on chromosomes 2, 9 and 10 only (Fig. 1). We also genotyped several markers in a region on chromosome 12 implicated in other linkage studies (5,9,10,11), but not in our genome scan, and evidence of allelic association was also found at this location (Fig. 1).

The region showing association on chromosome 2 is a 1.3 cM interval between D2S305 and D2S2201. The most significant allelic distribution difference was observed with D2S305 (global P = 0.01) at position 40.7 (see Fig. 1). At the genotype level a more remarkable difference was found with D2S310 located 1 cM away. For this marker, 10% of DAT cases but 0% of controls were carriers of the 135 bp allele (P = 0.005).

On chromosome 9, significant allelic association was observed with four markers located in the interval between 32 and 45 cM (Fig. 1). The associations with three of these markers (D9S925, D9S9162 and D9S259) were due to over-representation of a specific allele. Further examination revealed significant excess homozygosity of these marker alleles and a marker allele at D9S171 in DAT cases compared with non-demented subjects (Table 2). The observation of alleles with sizes 244 bp or greater in 16.5% of DAT cases but 0% of controls was carriers of the 135 bp allele (P = 0.005).

Taking into account the results from the genome scan and previous linkage studies, we evaluated intensively a 90 cM interval on chromosome 10 (Fig. 1). Although significant allelic associations were observed with markers located in separate linkage groups [e.g. D10S1426 @ 59 cM (P = 0.007), D10S1225 @ 80.8 cM (P < 0.02), D10S13765 @ 108.8 cM (P = 0.004)], the most conservative conclusion is that an DAT susceptibility gene is located in the interval between 105 cM and...
and 115.3 cM which contains four contiguous markers (D10S1686, D10S1765, D10S1753 and D10S583) showing evidence for allelic association (Fig. 1) and genotypic association involving carriers of the associated allele (data not shown). Concomitant allelic and genotypic association was observed for only two other chromosome 10 markers (D10S1426 and D10S208).

Our genotyping efforts on chromosome 12 focused on the region near and distal to LRP1, which was implicated in

![Figure 1](http://hmg.oxfordjournals.org/)

Table 2. Evidence for an additive or recessive AD locus on chromosome 9

<table>
<thead>
<tr>
<th>Locus</th>
<th>Map position</th>
<th>Genotypes</th>
<th>AD (%)</th>
<th>Control (%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D9S156</td>
<td>29.5</td>
<td>131/131 and 145/145</td>
<td>31</td>
<td>9</td>
<td>n.s.</td>
</tr>
<tr>
<td>D9S925</td>
<td>32</td>
<td>181/181</td>
<td>15</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>D9S162</td>
<td>33.3</td>
<td>173/173</td>
<td>32</td>
<td>15</td>
<td>0.02</td>
</tr>
<tr>
<td>D9S171</td>
<td>42</td>
<td>127/127</td>
<td>36</td>
<td>20</td>
<td>0.01</td>
</tr>
<tr>
<td>D9S259</td>
<td>45.1</td>
<td>126/126</td>
<td>17</td>
<td>2</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Figure 1. Association between AD and markers on chromosomes 2, 9, 10 and 12. Allelic association test results only are plotted. Significance levels are expressed on a log scale.
linkage and association studies of outbred populations (10,13,22). Whereas no association was evident among markers within 5 cM of LRP1, six markers between 5 and 20 cM distal to LRP1 were associated with DAT in this population (Fig. 1). Allele 216 at D12S1686 was present in 44% of DAT cases but in only 27% of controls ($P=0.006$). Three alleles (256/258/260) at D12S1722 accounted for 14% of DAT chromosomes but were absent in controls ($P=1.3\times10^{-6}$). Allele 216 at D12S326 was significantly over-represented ($P=0.0008$) in chromosomes of DAT cases (40%) compared with controls (26%).

**DISCUSSION**

We designed an unconventional 10 cM genome scan to identify loci for DAT in a highly inbred Israeli–Arab community having the highest known prevalence of DAT in the world (18). The history of this region suggests that the residents of this community were descendants of a small number of founder individuals. Since then, it has remained a relatively closed society with a high degree of consanguinity. Because only 10 distantly related individuals (five DAT cases and five nondemented controls) from one extended kindred were
genotyped, this approach is both efficient and economical relying, of course, on the assumption that the genes contributing to DAT in Wadi Ara are identical by descent and prevalent among affected members. The genome scan with 375 microsatellite markers revealed evidence of association (P < 0.05) with 18 locations, however, after testing additional markers in a larger number of people, only association to chromosomes 2, 9 and 10 remained.

Arguably, our strategy for the genome scan has a high false-negative rate because it failed to detect the DAT locus on chromosome 12. Motivated by linkage findings from previous studies of outbred populations, we genotyped many closely spaced markers in the region adjacent and distal to LRPL1 and observed significant association with DAT. Careful inspection of the results from the genome scan revealed that two of the markers flanking the linked region (D12S83 at 76.5 cM and D12S351 at 96.7 cM) failed to show association in the genome scan or the follow up (Fig. 1). Unfortunately, the genotyping assay for the one marker in this interval (D12S326 at 81.6 cM)—which yielded a very significant result in the follow-up analysis—failed in the genome scan, thus accounting for the false-negative result for this locus. The large number of alleles per locus (Table 3) may have further increased the false-negative rate. These observations illustrate a potentially serious limitation of a 10 cM genome scan using few subjects. Nonetheless, this strategy successfully detected the loci on chromosomes 2, 9 and 10.

Results of this study attest to the utility of genetic isolates, including those from this region of the Middle East (23–25), for mapping genes responsible for disease susceptibility in outbred Caucasian populations. Our strategy, evaluating linkage disequilibrium between microsatellite marker alleles and DAT, differs from traditional linkage approaches in genetic isolates in several respects. First, we focused our efforts on regions identified by a genome scan including only 10 individuals. Second, the DAT patients in the current cohort, all purportedly descendants from 14 founder matings, are members of distinct pedigrees. Gene mapping designs which rely on linkage or family-based association approaches (23–25) are thus not suitable for our situation in which relationships between individuals in the sample are distant or unclear. Third, although the efficacy of linkage disequilibrium mapping in isolated founder populations for recessive traits was demonstrated by Hästbacka and colleagues (26), the efficacy of this approach has not been proven for complex traits. Notably, the allele-frequency-dependent homozygosity mapping method, which enabled the localization of a gene for recessive deafness in Bali using only 13 affected individuals (27), does not require evaluation of DNA from parents and is thus perhaps well suited to the Wadi Ara population, however, it is effectively limited to recessive traits.

All four chromosomal regions implicated in this study have been reported previously suggesting that they may be authentic loci rather than chance findings. The marker on chromosome 2 showing the most significant association with DAT (D2S305) is located about 6–10 cM from two markers showing linkage disequilibrium in the case–control sample from Finland (28). The chromosome 9 location overlaps with peaks observed in two other studies with overlapping data sets (5,29) and the chromosome 10 location with several other studies (7,8,29,30). While we detected evidence of association to chromosome 12 in this population, the precise region is 20–35 cM distal to the linkage peaks observed by others (9–11,22,31). Thus, it is unclear if the region we identified on chromosome 12 is the same as that reported in other studies.

Because multiple markers were tested on chromosomes 2, 9, 10 and 12, these findings were reconsidered after transforming the P-values for the most significant result at each location to LOD scores, Z. According to guidelines for interpreting linkage results (32), the locations on chromosomes 9 and 12 are highly significant (Z = 5.5 and 4.8, respectively) and the chromosome 10 locus is certainly suggestive for linkage (Z = 2.1). The most significant result on chromosome 2 (Z = 1.5) falls just below the cutoff of 1.7 proposed for ‘suggestive linkage’; however, our finding would still be considered significant because it replicates a prior linkage finding (28).

The identification of multiple loci in a genetic isolate was unexpected. One or more of the loci might represent a false-positive result, although as discussed above all four loci correspond to approximate locations identified in studies of other populations. A more likely explanation is the unfortunate aggregation of multiple susceptibility loci in a single community. Such heterogeneity is consistent with the remarkably high prevalence of DAT in Wadi Ara (18). Alternatively, the high prevalence of DAT can be explained by an oligogenic mode of transmission requiring coincident inheritance of susceptibility alleles at multiple loci. Discrimination among these hypotheses will require simultaneous examination of marker haplotypes or, preferably, the actual susceptibility alleles from each locus in groups of affected and non-demented individuals.

While Wadi Ara has one of the highest prevalence rates of DAT in the world (18), the very low frequency of the APOE ε4 allele in both non-demented individuals (2.4%) and DAT cases (3.6%) indicates that the genetic basis of the disorder in this population is due entirely to other loci (20). A low impact of APOE on AD risk has been reported in Yoruba living in Nigeria, but the baseline frequency of ε4 is more than 20% and the incidence of AD is low (33,34). In contrast, Pericak-Vance and colleagues observed only the APOE 3/3 genotype among AD cases in the Amish, an inbred religious sect (35). However, the Amish have a lower prevalence of dementia than Wadi Ara and even most outbred Caucasian populations of European origin. Non-genetic factors including diet, low levels of education and physical activity may also be partly responsible for the high prevalence of the disorder in Wadi Ara (36).

We were surprised by the high degree of marker polymorphism suggesting that there might be a greater degree of admixture with other populations than predicted (see Table 3). Alternatively, the large number of alleles may be due to the relative instability of microsatellite markers from generation to generation. Interestingly, substantial marker polymorphism was also observed in a study of AD in 98 members from a genetic isolate in Finland (28). In the Finnish study, global tests for allelic association were significant at the 0.05 level for 21 out of 366 markers tested genome-wide, but two-thirds of these regions were not confirmed by follow-up analysis with additional linked markers. The very high proportion of marker heterozygosity, while perhaps enabling the detection of
individual marker association, hindered meaningful haplotype analysis because without parental data to establish linkage phase the power for estimating haplotypes in this sample is poor, even with the assistance of maximum likelihood approaches (37). A more efficacious strategy for assessing linkage disequilibrium over small intervals defined by haplotypes and, ultimately, for identifying candidate genes, is the evaluation of single-nucleotide polymorphisms (SNPs).

The results of our study implicate several potential AD genes. APOB and ADAM17 are located in the candidate region on chromosome 2. APOB is the main apolipoprotein of chylomicrons and low-density lipoproteins (LDL) and is important for lipid and cholesterol metabolism. ADAM17 is a disintegrin-metalloprotease that is responsible for S2 cleavage of NOTCH1. This cleavage makes NOTCH1 susceptible to cleavage by a gamma-secretase-like protease, leading to activation of the notch pathway (38). This is analogous to the action of presenilins that also act through the notch pathway.

The peak on chromosome 10 between D10S1686 and D10S583 includes several genes previously associated with AD including cholesterol 25-hydroxylase (CH25H) which encodes a membrane protein that synthesizes oxysterol regulators of lipid metabolism (39), insulin-degrading enzyme (IDE), a neutral metalloproteinase and has been shown to degrade beta-amyloid (40), an important protein found in the plaques seen in AD, and urinary plasminogen activator (PLAU), a protease that converts plasminogen to plasmin and is elevated in AD patients (41).

The observation of excess chromosome 9 marker homozygosity among DAT patients is consistent with recessive inheritance. All of the known genes for familial early-onset AD (i.e. APP, PS1 and PS2) are expressed in an autosomal dominant fashion, and the effect of APOE e4—the one bona fide gene for late-onset AD—on AD risk is additive (2,42). Although most genetic modeling studies based on data from outbred families have not found evidence for recessive inheritance (43–45), a recessive model was not rejected in an analysis of families lacking the APOE e4 allele (46). Consanguineous populations provide a unique opportunity to discern recessively behaving genes for diseases manifesting in late life.

A potential caveat of this study is the lack of neuropathological evidence for AD in any patient from this community. Unfortunately, brain autopsy could not be performed because it violates religious customs. It is therefore possible that subjects with other dementing illnesses, especially vascular dementia, may have been misclassified as DAT. However, this concern is lessened by corroborating evidence from brain MRI scans performed on a portion of DAT patients (see Methods). Diagnostic validity is also supported by the finding of linkage to three regions previously implicated as harboring AD loci.

The identification of chromosomal segments showing association or linkage is only the first step toward discovery of genetic defects that increase susceptibility to AD. Together with the large amount of gene information available, it is now feasible to systematically identify SNPs in each of the potential candidate genes and evaluate their association with AD. Given the SNP map is considerably more dense than the microsatellite marker map, it is likely the candidate regions can be reduced significantly by haplotype analysis using small blocks of SNPs.

### MATERIALS AND METHODS

#### Subjects

Subjects were ascertained from a prevalence study of dementia conducted in Wadi Ara, a geographically defined area in northern Israel comprising three Arab villages (18). Among the 853 residents ages 60 years and older on prevalence day (October 1, 1995), 821 were available and consented to participate in the study. These subjects were interviewed and examined by an Arabic-speaking physician (A.B.) who had been previously trained in a memory clinic. Each subject underwent a battery of standard cognitive tests modified to fit the cultural and linguistic characteristics of this community (18). The diagnosis of DAT was determined using DSM-IV criteria (47). A diagnosis of AD is conventionally reserved for patients diagnosed using NINCDS/ADRDA criteria (48). Persons with mild cognitive impairment and demented subjects whose medical history and laboratory and cognitive test results suggested the presence of other illnesses such as vascular dementia, Parkinson disease, normal-pressure hydrocephalus or pseudodementia (depression), were excluded. Evidence of ischemic stroke or white matter disease was present in only 2/15 DAT patients (ages ranging from 71 to 91 years) who had a brain MRI scan, suggesting that relatively few subjects were misclassified. Peripheral blood samples were obtained between 1997 and 2000 from the 650 survivors for DNA and biochemical analysis.
DNA analysis

DNA was extracted from whole blood according to procedures described previously (18). Fluorescently labeled primers for the microsatellite markers used in this study were purchased from Applied Biosystem (ABI) and used in a 10 μl PCR reaction with 5 ng of genomic DNA. Products of different size and fluorescent label were combined and applied to either an ABI 377 Gel-based DNA sequencer or an ABI 3700 capillary DNA sequencer. The genotyper program (ABI) was used to determine allele sizes. Laboratory technicians were blinded to subject identifying information associated with any sample ID number. Each plate contained multiple control samples to ensure quality genotyping and allele determination.

Statistical analysis

Allele and genotype associations were assessed using SAS (version 8.02) software. Exact tests were used to evaluate associations in the genome scan and global \( \chi^2 \) tests were used in follow-up studies in the enlarged sample. Because the number of alleles and genotypes for these microsatellite markers far exceeded that expected for genetic isolate (as many as 19 alleles and 60 genotypes), the frequency distributions for all global tests were inspected visually to suggest one or two single-allele and collapsed-genotype group tests. All reported significant \( P \)-values (\( P \leq 0.05 \)) are based on these latter tests, unless indicated otherwise, and are not adjusted for multiple testing. To correct for testing multiple markers on the same chromosome in the follow-up studies, \( P \)-values were transformed to lod scores using the equivalence

\[
\text{LOD} = \left( \frac{\text{probit}(1 - P\text{-value})}{\sqrt{2 \times \log_{10}}} \right)^2
\]

which is adapted from allele-sharing models developed by Kong and Cox (49). These LOD scores were then interpreted according to the criteria recommended by Lander and Kruglyak (32). Marker positions were obtained from the online Genethon Linkage Map available at www.ncbi.nlm.nih.gov/mapview/map_search.cgi

ACKNOWLEDGEMENTS

We thank Dr Yoav Chapman and Dr Miriam Birnbaum for performing DNA extraction and APOE genotyping on many of the samples and Corey Adams for initiating the genotyping effort in this study. This work was supported in part by grants from the National Institutes of Health (U01-AG17173) and the Institute on the Study of Aging.

REFERENCES


