American Journal of Medical Genetics (Neuropsychiatric Genetics) 114:235–244 (2002)

Rapid Publication

Full Genome Screen for Alzheimer Disease: Stage II Analysis

Amanda Myers,[1] Fabienne Wavrant De-Vrieze,[2] Peter Holmans,[3] Marian Hamshere,[4] Richard Crook,[2] Danielle Compton,[2] Helen Marshall,[1] David Meyer,[1] Shantia Shears,[1] Jeremy Booth,[1] Dzanan Ramic,[1] Heather Knowles,[1] John C. Morris,[5] Nigel Williams,[4] Nadine Norton,[4] Richard Abraham,[4] Pat Kehoe,[4] Hywel Williams,[4] Varuni Rudrasingham,[4] Francis Rice,[4] Peter Giles,[4] Nigel Tunstall,[5] Lesley Jones,[4] Simon Lovestone,[6] Julie Williams,[4] Michael J. Owen,[4] John Hardy,[2] and Alison Goate[1,7]*

[1]*Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri*
[2]*Laboratory for Neurogenetics, Birdsall Building, Mayo Clinic Jacksonville, Jacksonville, Florida*
[3]*MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Cambridge, U.K.*
[4]*Department of Psychological Medicine, University of Wales College of Medicine, Cardiff, U.K.*
[5]*Department of Neurology, Washington University School of Medicine, St. Louis, Missouri*
[6]*Institute of Psychiatry, De Crespigny Park, London, U.K.*
[7]*Department of Genetics, Washington University School of Medicine, St. Louis, Missouri*

We performed a two-stage genome screen to search for novel risk factors for late-onset Alzheimer disease (AD). The first stage involved genotyping 292 affected sibling pairs using 237 markers spaced at approximately 20 cM intervals throughout the genome. In the second stage, we genotyped 451 affected sibling pairs (ASPs) with an additional 91 markers, in the 16 regions where the multipoint LOD score was greater than 1 in stage I. Ten regions maintained LOD scores in excess of 1 in stage II, on chromosomes 1 (peak B), 5, 6, 9 (peaks A and B), 10, 12, 19, 21, and X. Our strongest evidence for linkage was on chromosome 10, where we obtained a peak multipoint LOD score (MLS) of 3.9. The linked region on chromosome 10 spans approximately 44 cM from D10S1426 (59 cM) to D10S2327 (103 cM). To narrow this region, we tested for linkage disequilibrium with several of the stage II microsatellite markers. Of the seven markers we tested in family-based and case control samples, the only nominally positive association we found was with the 167 bp allele of marker D10S1217 (chi-square = 7.11, *P* = 0.045, df = 1). © 2002 Wiley-Liss, Inc.

KEY WORDS: genome screen; genome scan; late-onset Alzheimer disease

INTRODUCTION

While the genetics of early-onset autosomal dominant Alzheimer disease (AD) is fairly well understood [Goate et al., 1991; Levy-Lahad et al., 1995; Sherrington et al., 1995], our understanding of the more common late-onset disorder remains much less complete. Segregation and risk analysis suggest a large familial component to the disease; the estimated cumulative risk to first-degree relatives of AD-affected probands approaches 50% by age 90 compared to a disease risk of 10–15% in the general population [Breitner, 1990]. To date, only the apolipoprotein ϵ4 allele has been linked to increased risk for late-onset AD. While it is clear that *APOE* is a major risk factor for AD, epidemiological studies estimate that 42–68% of late-onset AD cases do not have an *APOE* ϵ4 allele, indicating that additional genetic and environmental factors are involved in this form of the disease [Henderson et al., 1995; Martins et al., 1995; Kukull et al., 1996].

To search for novel late-onset AD risk factors, we used a two-stage affected sibling pair (ASP) approach that maintains power while focusing on the regions most likely to contain susceptibility genes [Holmans and Craddock, 1997]. Nonparametric allele-sharing methods were used to test for linkage. Unlike parametric methods, which specify a mode of transmission,

Amanda Myers and Fabienne Wavrant De-Vrieze contributed equally to this work.

*Correspondence to: Alison Goate, Department of Psychiatry, B8134 Washington University School of Medicine, 660 S. Euclid, St. Louis, MO 63110. E-mail: goate@icarus.wustl.edu

Received 5 October 2001; Accepted 9 October 2001

© 2002 Wiley-Liss, Inc.
DOI 10.1002/ajmg.10183

allele frequencies, and penetrances of the susceptibility locus, nonparametric analyses are for the most part model-free and are thereby more suited to the analysis of diseases with a complex mode of transmission like late-onset AD.

In the first stage of our screen [Kehoe et al., 1999], we genotyped 237 markers in 292 ASPs, yielding a mean intermarker interval of 16.3 cM. We tested for linkage in the whole sample (292 sibling pairs) and in two subseries stratified by the presence or absence of an *APOE* ϵ4 allele. *APOE* ϵ4-positive sibling pairs were siblingships in which each sibling possessed at least one ϵ4 allele (162 ASPs) and *APOE* ϵ4-negative sibling pairs were pairs in which neither sibling had any ϵ4 alleles (63 ASPs). We found 16 multipoint LOD scores (MLSs) ≥ 1 on 12 chromosomes. Simulation studies indicated that this number of peaks exceeded that expected by chance (for a LOD = 1, P value = 0.025, expected number per genome scan = 12.5). Four peaks were suggestive of linkage, as defined by Lander and Kruglyak [1995], in that they would be expected to occur by chance once or less per genome scan. These peaks were located on chromosomes 1 (D1S1675 at 149 cM), 9 (D9S176 at 105 cM,), 10 (D10S1211 at 82 cM,), and 19 (D19S571 at 84 cM, 24 cM from *APOE*).

The second stage of our screen involved typing 91 additional markers within our 16 stage I peaks in an extended series of 451 ASPs. In addition to follow-up of our stage I linkage, we attempted to narrow the linkage peak further on chromosome 10 by looking for linkage disequilibrium (LD). In this study, we present these data with a comparison of the findings with those of other groups who have used similar strategies, together with a practical assessment of the strengths and weaknesses of this and related approaches.

MATERIALS AND METHODS

Samples

Affected sibling pair series. One hundred seventy-four additional affected sibling pairs (ASPs) were obtained using the same selection criteria as in our stage I study; each sibling used for analysis had to have an age of onset greater than or equal to 65 years and a diagnosis of definite or probable AD according to NINCDS-ADRDA diagnostic criteria. Diagnoses of previously collected samples were updated and any individuals who no longer had either definite or probable AD were excluded. Additionally, in both stages only Caucasian families were selected to reduce potential genetic heterogeneity and allelic frequency differences caused by ethnic origin. If there were more than two affected siblings within a family, all siblings were sampled and genotyped. Stage I ASPs were selected from families ascertained by the NIMH-AD Genetics Initiative (see appendix A). Stage II sibling pairs were obtained from the NIMH, Indiana Alzheimer Disease Center National Cell Repository (NIA), and by the U.K. members of our collaborative group. The 94 ASPs from the U.K. were ascertained through contacts with clinical services. Participants provided informed consent according to procedures approved by local and national ethics committees. All siblings had parents born in the U.K. Each was interviewed using standard measures that have been validated against postmortem diagnosis and show a positive predictive value of over 90% for detecting AD pathology [Holmes et al., 1999].

Discordant sibling pair series. A series of discordant sibling pairs (DSPs) was used to test for linkage disequilibrium. All unaffected individuals in the DSP set were selected on the basis of the following criterion: to avoid possible censoring, only unaffected individuals that were shown to be cognitively normal at an age greater than the oldest age of onset for their family were selected. Only the NIMH sample had unaffected individuals who fit this criterion, therefore only NIMH families were used for the DSP series. We obtained unaffected siblings from the same families that were used to find our stage I linkage and collected additional affected and unaffected sibling pairs. Our total DSP sample size is 308 individuals from 101 families with a total of 132 unaffected siblings and 176 affected, 160 of those DSPs were from families used in stage I.

Case control series. We have used a group of 92 Caucasian cases (mean age of onset, 75 ± 6.2) and 94 age-, gender-, and ethnicity-matched controls (mean age at last assessment, 78 ± 7.8) to look for LD. These samples were obtained from the Memory and Aging Project (MAP) from the Washington University Alzheimer's Disease Research Center (ADRC). The subjects are participants in the Washington University ADRC patient registry. All participants are enrolled in longitudinal studies of the natural course of AD and healthy aging in the Washington University MAP and have been described previously [Kwon et al., 2000].

Genotyping

Samples were genotyped using the methods described in our previous study [Kehoe et al., 1999]. Ninety-one additional markers were typed to reduce the average within peak intermarker interval to approximately 5 cM. These markers were selected from the CHLC (http://lpg.nci.nih.gov/CHLC/index.html) and CEPH (http://www.cephb.fr/cgi-bin/wdb/ceph/systeme/forme) databases and marker order was determined from the Marshfield (http://www.marshmed.org/genetics/mapmarkers/maps/indexmap.html) and Southampton maps (ftp://cedar.genetics.soton.ac.uk/public_html/gmap.html). Intermarker intervals were determined from the Marshfield maps.

Statistical Analysis

Linkage studies. Stage II data and our edited stage I sample were analyzed using the same methods. The program SPLINK [Holmans and Clayton, 1995] was used for single marker assessment with Holmans and Clayton's possible triangle restrictions and also to estimate marker allele frequencies for use in the multipoint analyses. Multipoint affected sibling pair linkage analyses were carried out using MAPMAKER/SIBS [Kruglyak and Lander, 1995]. Pairs taken from multiplex siblingships were not down-weighted in

either the two-point or the multipoint analyses. Three analyses were performed: on the whole sample (451 pairs in 349 siblingships), on pairs where both siblings possessed at least one *APOE* ϵ4 allele (280 pairs in 217 siblingships), and on pairs where neither sibling possessed an ϵ4 allele (76 pairs in 66 siblingships).

Error detection for linkage study. Three procedures were used to test for genotyping errors. First, visible inheritance errors were determined for each marker. In the absence of genotyped parents, many genotyping errors will be undetectable, making the visible error rate a severe underestimate of the true error rate. Therefore, error rates were estimated using a second method. Replicate sample sets were simulated using the actual people genotyped and the allele frequencies at each locus, randomly introducing errors at a chosen fixed rate. The average expected number of visible errors by simulation at the given error rate was counted for each locus, and this process was repeated with different fixed error rates. The final stimulated error rate was the error rate in which the expected number of simulated visible errors most closely matched the actual number of visible errors. Markers within the peak regions from stage I giving an estimated error rate $\geq 5\%$ were replaced. Finally, data were checked using the program SIBMED [Douglas et al., 2000] to eliminate genotypes that did not give visible inconsistencies but were nevertheless unlikely given the allele frequencies and marker maps.

Genomewide significance for linkage study. Simulation procedures were also used to determine genomewide significance. In regions that were followed up in stage II, replicates of the whole second-stage sample were simulated, in the absence of disease loci, using the same marker maps, allele frequencies, and family structures used in the analyses, and replicate stage I samples were extracted. For regions that were not followed up, a second-stage marker grid was approximated by filling in the gaps larger than 5 cM from the stage I screen with markers consisting of four equifrequent alleles. The three analyses (whole sample, ϵ4-positive, and ϵ4-negative) were performed on each replicate stage II sample and its corresponding stage I sample. This procedure was repeated for 5,000 replicates and expected numbers of stage II peaks reaching a given height in any of the three analyses, given that at least one analysis gave a stage I LOD ≥ 1 was calculated.

APOE covariate analysis. Interactions between the putative susceptibility gene on chromosome 10 and *APOE* were investigated using a method suggested by Rice [1997]. The probabilities of an affected pair sharing a particular parental allele identical-by-descent (IBD) were modeled as a logistic regression on a categorical variable with three levels, one for each of the three types of pairwise *APOE* genotype designations (++, +−, −−, with + being defined as possessing at least one ϵ4 allele). A maximum MLS was calculated allowing the IBD probabilities in the three groups to differ. The IBD probabilities were not restricted to $\geq 50\%$, since differences in the sharing between groups were being tested, rather than the overall IBD. Chromosomewide significance of the maximum MLS was assessed by randomly permuting the pairwise *APOE* designations among the affected pairs and repeating the calculations.

Association studies. The program CLUMP [Sham and Curtis, 1995] was used to analyze the case control series for association using seven microsatellite markers from chromosome 10. CLUMP accounts for dependency between tests of individual alleles and low allele frequencies by using Monte-Carlo methods to assess significance of test statistics. TRANSMIT [Clayton, 1999] was used to analyze association in the DSP sample. Microsatellites were assessed for association by testing all alleles simultaneously and also by testing each allele against all other alleles for that marker. *P* values for the DSP analysis were obtained by bootstrap simulation to allow for siblingships with multiple affected individuals.

Power analysis for linkage (ASP) sample. Replicate chromosomes were simulated for the 451 sibling pairs in stage II assuming a 5 cM grid of markers with four equifrequent alleles (heterozygosity = 0.75). The stage I sample was obtained by extracting the genotypes at every fourth marker (20 cM grid) in the first 270 sibling pairs. MLS analyses were performed on the stage I and II samples and the power was defined as the proportion of replicates in which the LOD score at the true location of the disease locus exceeded preset criteria in both stages. The analyses were repeated for disease loci in a number of positions relative to the marker grid [Holmans and Craddock, 1997] and the power averaged.

RESULTS

For the second stage of our screen, we reexamined all diagnoses for the ASPs within our stage I screen. Any individuals who no longer had either definite or probable AD were excluded from this study. Fifteen sibling pairs had to be dropped due either to changes in diagnoses (i.e., they were originally diagnosed as probable AD and this changed to possible AD) or to problems with age of onset (onset below 65 years). Additionally, some of the chromosomal maps changed considerably from the time we performed our first scan. This resulted in several of the markers on chromosome 12 mapping to the same chromosomal location. Markers from this chromosome that had more inheritance errors or for which the map order was less certain were dropped. The combined effect of these changes led to different stage I results, which are shown in the dashed lines on Figure 1. As a general rule, peak locations did not change (except in cases where the peaks disappeared, as on chromosome 2), but some multipoint LOD scores did. It is notable that the results from chromosome 2, 6, 9 (peak A), 12, 14, and X dropped below the criterion for follow-up (MLS ≥ 1 in stage I; Fig. 1, dashed lines) with the new diagnoses and map changes. In contrast, the MLS on chromosome 10 increased from 2.3 to 3.5 with these changes.

In stage II, we collected an additional 174 sibling pairs, 80 from the NIA series and 94 sibling pairs from a

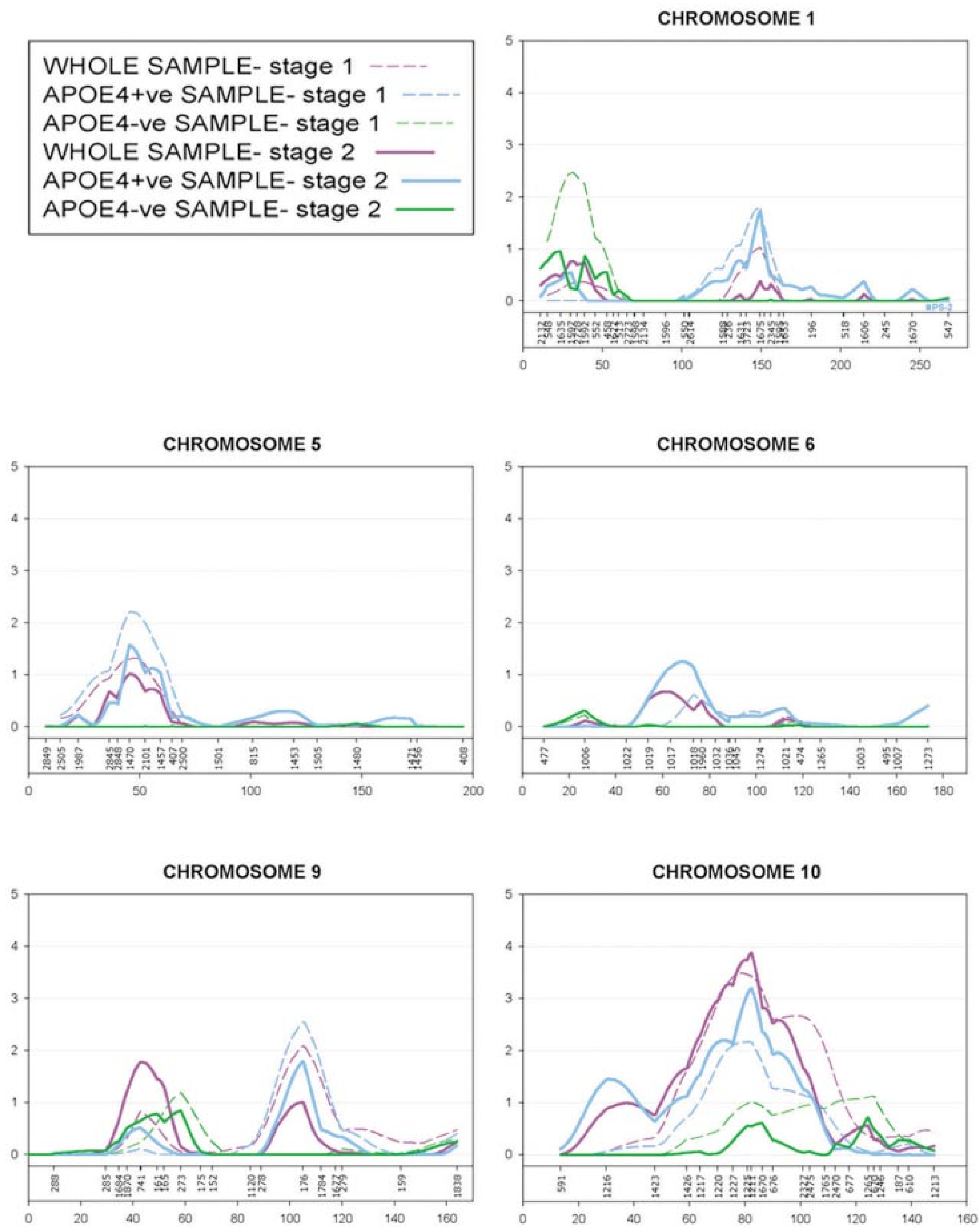

Fig. 1. Multipoint analysis using MAPMAKER/SIBS [Kruglyak and Lander, 1995] for all chromosomes that had maximum MLSs ≥ 1 in our stage I screen [Kehoe et al., 1999]. Results from chromosome 2 were completely negative with both the reanalysis of stage I and the stage II analysis and are therefore not shown. For all graphs, the X-axis represents the cM distance from the pter of each chromosome and the Y-axis is the MLS. Presenilin 1 (PS-1, chromosome 14), presenilin 2 (PS-2, chromosome 1), amyloid precursor protein (APP, chromosome 21), α-2 macrogobulin (A2M), and low-density lipoprotein-related protein 1 (LRP 1, chromosome 12) were mapped relative to our linkage results using the NCBI map viewer locations (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/).

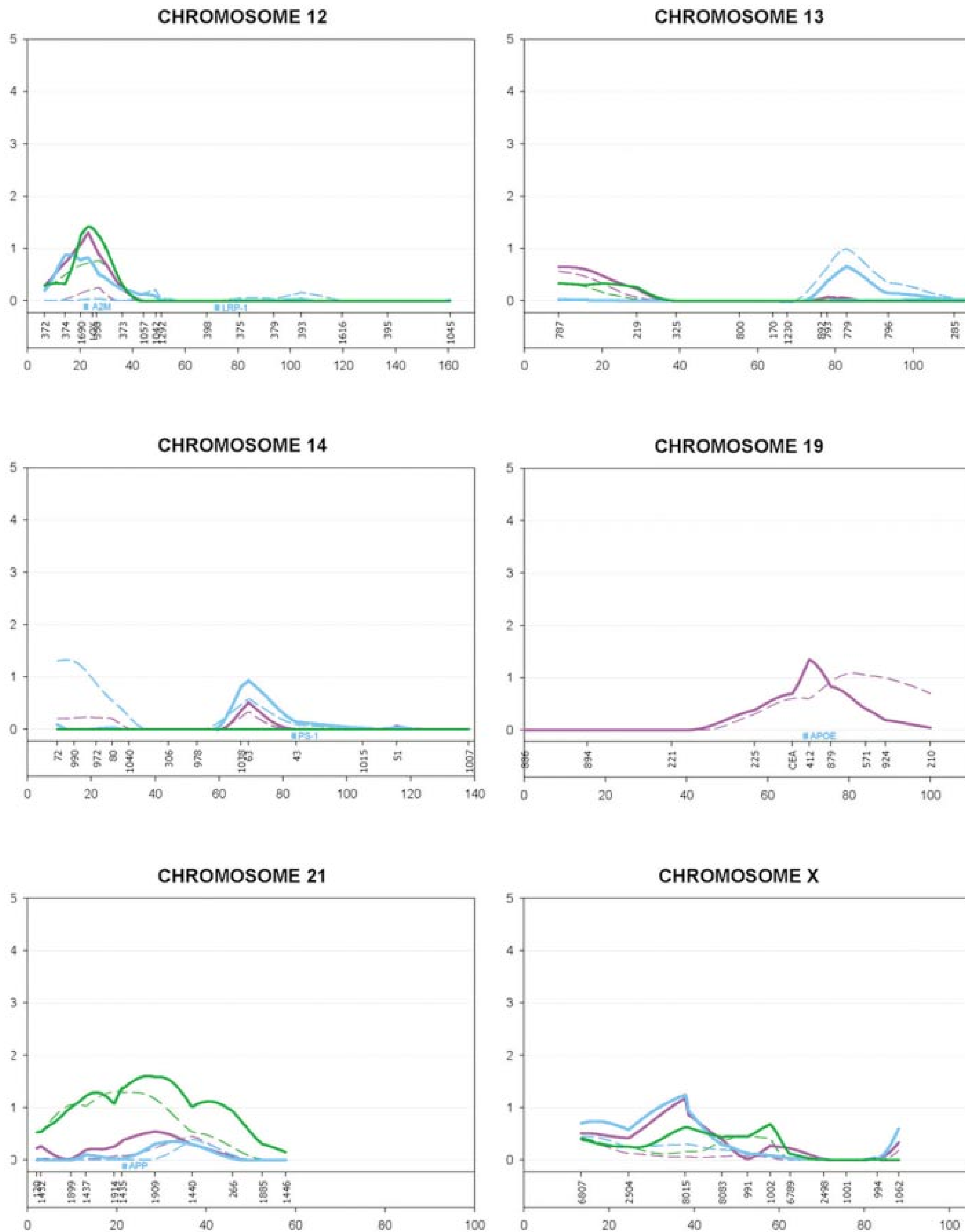

Fig. 1. (*Continued*)

series collected in England and Wales to make a total sample of 451 sibling pairs. These additional sibling pairs were collected using the same criteria as in stage I. The sample was stratified by *APOE* ϵ4 genotype as in stage I. Sample characteristics for both stages and all subgroups analyzed are listed in Table I.

Ninety-one new markers (mean heterozygosity = 0.76) were genotyped in this second stage. Seventy-five of the 91 new markers were located within the peaks from stage I, and 16 filled gaps greater than 20 cM outside the peaks. We also genotyped the 60 stage I markers that were within the peaks on the additional

TABLE I. Summary Statistics of the Affected Sibling Pair Sample Used to Assess Linkage in Stage I and II

Sample	Mean age of onset (SD)	Siblingships with ≥ 2 affecteds	Number of ASPs (% female)	Number of *APOE4*$^{+}$ ASPs[a]	Number of *APOE4*$^{-}$ ASPs[b]
NIMH(stage I)[c]	74.9 (5.8)	212	277 (74%)	160	56
NIA (stage II)	73.4 (5.3)	64	80 (65%)	66	5
U.K. (stage II)	76.1 (6.61)	73	94 (79%)	54	15
Stage II totals[d]		349	451	280	76

[a]Sibling ships in which both siblings possessed at least one *APOE* ϵ4 allele.
[b]Sibling ships in which neither sibling possessed an *APOE* ϵ4 allele.
[c]Numbers listed for both sample sets are totals after individuals who did not have a diagnosis of probable or definite AD were eliminated.
[d]Stage II totals are the sum statistics for all three samples (NIMH, NIA, and U.K.).

174 sibling pairs added in stage II. We reduced our average intermarker distance within the peaks to 5.92 cM and the intermarker distance outside the peaks to 12.8 cM. Markers were assessed for visible Mendelian errors. Additionally, since we were using a sibling pair series and no parental DNA was typed, we also estimated error rates for each new marker by simulation. Five markers (CEA on chromosome 19, D19S571, D19S412, D21S1009, and DXS1002) gave estimated error rates above 5% and therefore were regenotyped. One marker was accidentally genotyped by two sites and 21 genotypes did not match, yielding a genuine between-site genotyping error rate of less than 1% for this marker.

As in our first study [Kehoe et al., 1999], multipoint ASP analysis was performed on all follow-up chromosomes using MAPMAKER/SIBS [Kruglyak and Lander, 1995]. The reanalysis of our stage I data was performed using the individuals from our first study that continued to fit our stringent diagnostic criteria and the markers from our original stage I analysis. The stage II analysis was performed on the combined set of both the stage I sibling pairs and the additional sibling pairs collected from the NIA and the U.K. Complete chromosomes were analyzed in stage II, not just the regions underneath the peaks from stage I. The results from the reanalysis of our stage I data and our stage II analysis of the whole sample and *APOE* ϵ4-positive and -negative subgroups are shown in Figure 1 summarized in Table II. Table III lists the results of two-point analysis using the program SPLINK [Holmans and Clayton, 1995] for all markers that gave two-point LOD scores ≥ 1.

The peak MLS scores on chromosomes 1 (peak B), 6, 9 (peak A), 10, 12, 19, 21, and X all either increased or remained the same as the peaks from our reanalysis of the stage I genotypes. In contrast, the peak multipoint scores on chromosomes 1 (peak A), 2, 5, 9 (peak B), 13, and 14 all decreased relative to the reanalyzed stage I data; however, the chromosome 5 and 9B peaks remained above an MLS of 1 in stage II. In general, peak location drifted only slightly (i.e., chromosome 6 peak stage I = 73 cM and stage II = 63 cM) and these shifts reflect insignificant changes; simulation studies have shown that linkage can drift by as much as 20–30 cM from study to study [Roberts et al., 1999].

The peak on chromosome 19 mapped to within 2 cM of *APOE*, the only known risk locus for late-onset AD. To determine whether the linkage on chromosome 19 could be explained solely by *APOE* genotype, we compared the number of *APOE* ϵ4 alleles in individuals who shared alleles IBD at D19S412 (IBD ≥ 1.5), the marker closest to our linkage peak, to the number of ϵ4 alleles in individuals who did not share at D19S412 (IBD ≤ 0.5). There was a significant difference in the number of ϵ4 alleles in the sharers versus the non-sharers (one-sided $P = 0.001$; Table IV).

Amyloid precursor protein (APP) maps within the APOE4 negative peak on chromosome 21 in both stage I and II. Our previous analysis of sibling pairs from the NIMH series (stage I) found that siblingpairs lacking E4 alleles were more likely to share alleles at markers within APP, indicating that APP might also be involved in late-onset disease [Wavrant De-Vrieze et al., 1999]. Additionally, covariate analysis of the same series

TABLE II. Maximum MLSs ≥ 1 From Stage II Linkage Screen

Chromosome	cM	Nearest marker(s)	Sample		
			Whole	*APOE4*$^{+}$	*APOE4*$^{-}$
1	149.2	D1S1675		1.7	
5	45.34	D5S1470	1.0	1.6	
6	68.21	D6S1017–D6S1018		1.3	
9A	43.38	D9S741	1.8		
9B	105	D9S176	1.0	1.8	
10	82.21	D10S1211	3.9	3.2	
12	23.07	LOX-1	1.3		1.4
19	70.14	D19S412	1.3		
21	27.05	D21S1909			1.6
X	37.87	DXS8015	1.2	1.2	

TABLE III. Two-Point LOD Scores ≥ 1*

Chromosome	Locus	cM	Hetero-zygosity of marker	Number of families analyzed	Number of sibling pairs analyzed	Whole sample		*APOE4*$^+$ sample		*APOE4*$^-$ sample	
						IBD	LOD	IBD	LOD	IBD	LOD
1A[a]	D1S1592	38.51	0.62	305	387	0.57	1.7	0.55	0.5	0.69	2.0
	D1S552	45.33	0.72	297	372	0.52	0.2	0.5	0.0	0.59	0.4
	D1S458	52.7	0.73	302	373	0.53	0.3	0.5	0.0	0.65	1.1
1B[a]	D1S1675	149.2	0.7	296	370	0.55	1.1	0.58	1.9	0.5	0.0
	D1S534	151.88	0.86	310	393	0.53	0.6	0.55	1.3	0.55	0.3
2[b]	D2S125	260.65	0.83	203	262	0.55	1.0	0.53	0.1	0.54	0.1
5[a]	D5S1470	45.34	0.85	319	401	0.56	2.1	0.6	2.8	0.5	0.0
	D5S2101	52.55	0.87	275	355	0.53	0.6	0.53	0.4	0.53	0.1
	D5S1457	59.3	0.75	322	408	0.55	1.2	0.58	2.0	0.5	0.0
6[a]	D6S1018	73.13	0.47	310	400	0.53	0.3	0.6	1.5	0.5	0.0
	D6S1960	76.62	0.73	294	375	0.53	0.5	0.56	1.1	0.51	0.0
6[b]	D6S1021	112.2	0.71	172	214	0.58	1.2	0.63	1.9	0.55	0.1
9A[a]	D9S1870	37.58	0.8	329	419	0.54	0.9	0.51	0.1	0.62	1.1
	D9S741	42.73	0.79	331	425	0.56	1.7	0.54	0.6	0.56	0.2
	D9S169	49.2	0.84	248	311	0.54	0.7	0.53	0.3	0.58	0.4
	D9S161	51.81	0.79	253	320	0.55	1.0	0.52	0.1	0.59	0.5
9B[a]	D9S176	105.02	0.85	329	422	0.56	1.5	0.59	2.3	0.5	0.0
10[a]	D10S1216	30	0.77	205	258	0.55	0.6	0.58	1.1	0.5	0.0
10[a]	D10S1217	63.83	0.52	291	368	0.56	0.9	0.57	0.9	0.5	0.0
	D10S1220	70.23	0.61	285	367	0.57	1.3	0.59	1.6	0.5	0.0
	D10S1227	75.57	0.73	301	372	0.55	0.7	0.52	0.1	0.58	0.3
	D10S1225	80.77	0.76	282	354	0.55	1.0	0.56	1.1	0.53	0.1
	D10S1211	82.21	0.66	310	394	0.62	4.1	0.63	3.6	0.63	0.7
	D10S1670	86.2	0.76	321	411	0.56	1.3	0.58	1.7	0.56	0.2
	D10S676	90.01	0.79	307	382	0.55	1.1	0.55	0.7	0.55	0.1
	D10S2327	100.92	0.68	329	409	0.55	1.0	0.57	1.3	0.5	0.0
	D10S2475	103.43	0.77	326	417	0.56	1.4	0.57	1.3	0.55	0.2
19[a]	D19S412	70.14	0.83	298	383	0.56	1.6				
21[b]	D21S1432	2.99	0.72	305	384	0.57	1.5	0.58	1.1	0.59	0.4
21[a]	D21S1435	21.13	0.78	307	390	0.54	0.7	0.51	0.1	0.67	2.0
21[a]	D21S1909	28.48	0.82	318	405	0.53	0.5	0.53	0.3	0.63	1.4
XA[a]	DXS8015	37.87	0.79	195	246		1.3		1.6		
XB[b]	DXS1002	57.91	0.74	321	410		1.2				1.2

*Adjacent markers with LODs < 1 are shown for peaks. *APOE4*$^+$ denotes siblingships where each sibling has at least 1 ϵ4 allele. *APOE4*$^-$ denotes siblingship where neither sibling has any ϵ4 allele.
[a]Two-point location matches multipoint peak location.
[b]Two-point location does not match multipoint peak location.

concluded that the APP locus may predispose to AD in the very eldery [Olson et al., 2001].

Our best results in both stages I and II were on chromosome 10, where we initially obtained an MLS of 2.3 [Kehoe et al., 1999]. This MLS increased to 3.5 when we reevaluated the stage I diagnoses and used updated maps. In stage II, the MLS increased further to 3.9. All of these peaks occurred at D10S1211, around 82.2 cM on our maps. Furthermore, there was a 44 cM region in which eight markers yielded two-point LOD scores approaching 1 in the whole sample analysis (Table III). All eight markers had elevated IBD allele sharing. Covariate analysis of these results and *APOE* genotype were not significant ($P = 0.75$), further indicating that this locus affects risk regardless of *APOE* ϵ4 genotype.

TABLE IV. Analysis of Linkage on Chromosome 19 and *APOE* Genotype*

Number of alleles shared IBD at D19S412	Number of *APOE* alleles		
	ϵ2	ϵ3	ϵ4
2-sharers[a]{FN3}	14	281	261
0-sharers[b]{FN4}	17	218	123

*Odds ratio for the presence of ϵ4 alleles in 2-vs. 0-sharers = 1.69 (95% CI = 1.20–2.21), *P* (one-sided) = 0.0001.
[a]2-sharers correspond to sibling pairs in which the proportion of estimated alleles shared was ≥ 1.5.
[b]0-sharers correspond to sibling pairs in which the proportion of estimated alleles shared was ≤ 0.5.

TABLE V. Simulated Genomewide *P*-values

Stage II LOD	Expected # of peaks	Genomewide *P*
1	8.3	1
1.5	3.6	0.93
2	1.5	0.69
2.2[a]	1	0.56
2.5	0.6	0.38
3	0.24	0.19
3.6[b]	0.06	0.05
4	0.04	0.03
Chromosome 19 (single analysis)		
1.3	1.5	0.7

[a]LOD score that would meet Lander and Kruglyak's [1995] definition of a suggestive peak.
[b]*P*-value corresponding to a significant peak.

TABLE VI. Results From Association Study Testing Seven Microsatellite Markers Underneath Linkage Peak on Chromosome 10 in a Case Control Sample and a Discordant Sibling Pair Sample*

			Case control results		DSP results[a]	
Locus	cM	Number of alleles	Chi-square	Simulated *P*-value	Chi square (df)	bootstrap *P*-value
D10S1217	63.83	9	4.03	0.892	8.38 (5)	0.1[b]
D10S1220	70.23	7	9.23	0.111	1.98 (5)	0.75
D10S1227	75.57	9	8.97	0.331	10.04 (7)	0.17
D10S1225	80.77	9	9.52	0.29	7.72 (5)	0.31
D10S1211	82.21	6	0.945	0.967	3.74 (6)	0.81
D10S1670	86.2	15	12.29	0.639	16.84 (10)	0.13
D10S676	90.01	9	10.51	0.215	7.52 (8)	0.69

*Note that all alleles with frequencies $\leq 1\%$ were pooled in the DSP analysis.
[a]Within discordant sibling pair (DSP) sample APOE chi-square for APOE $\epsilon 4$ allele = 28.45 (2 df), $P < 0.0002$.
[b]For D10S1217 allele, 167 had an individual chi-square of 7.11 (bootstrap $P = 0.045$).

Pointwise P values and genomewide significance levels were determined by simulation and the results are shown in Table V. These calculations indicate that we would expect to obtain an MLS of 2.1 once per genome scan in the absence of linkage and thus any MLS ≥ 2.1 would satisfy the definition of suggestive linkage by Lander and Kruglyak [1995]. None of our stage II results except the peak on chromosome 10 were above an MLS of 2.1. Yet the results from chromosome 10 reached a genomewide significance of $P \leq 0.05$ (by simulation, an MLS of 3.6 corresponds to a $P = 0.05$) and thus fulfilled the definition of significant linkage by Lander and Kruglyak [1995].

We have begun to follow up our significant linkage on chromosome 10 by looking for linkage disequilibrium within this region. We typed five markers (D10S1220–D10S1670) within the -1 LOD interval and the two markers (D10S1217 and D10S676) flanking the -1 LOD interval on two sample sets. The first was a series of discordant sibling pairs from the NIMH series. The second was a series of case control DNA from theMemory and Aging Project from the Washington University Alzheimer's Disease Research Center. We typed 314 DSPs from 101 families and 92 cases and 94 controls. Using the programs TRANSMIT (DSP analysis [Clayton, 1999]) and CLUMP (case control analysis [Sham and Curtis, 1995]), we did not find significant association in either sample, with six of the markers we typed within a 26 cM region around our peak. These results are shown in Table VI. The 167 bp allele of marker D10S1217 gave a modest association in the DSP sample (chi-square = 7.11, $P = 0.045$, df = 1, allowing for multiple alleles tested). Given that seven marker loci were analyzed, this result is likely to be a type I error, particularly since this same allele showed no association in our case control sample (chi-square = 0.24, df = 1). We are now creating a denser (0.5 cM) LD map of microsatellite markers spanning the -1 LOD interval surrounding our peak (D10S1220–D10S1670).

DISCUSSION

In our stage I full genome screen for novel loci involved in risk for late-onset AD, we found 16 regions with a maximum MLS ≥ 1. We have tested these results in stage II by adding additional markers in the regions underneath our peaks and by genotyping additional samples. The stage II maximum MLSs on chromosomes 1 (peak B), 6, 9 (peak A), 10, 12, 19, 21, and X all increased or remained the same as our reanalysis of the stage I genotypes. All of our other peaks decreased or disappeared, with the caveat that the chromosome 9 B-peak and the peak on chromosome 5 still have maximum MLSs ≥ 1. Our most consistent linkage evidence occurs on chromosome 10, with a peak MLS of 3.9 at marker D10S1211. However, we were unable to obtain any significant associations using the microsatellite markers underneath our chromosome 10 linkage peak in two independent samples.

For this study, we employed a two-stage nonparametric linkage approach. This method has several advantages. It required far fewer genotypes than a one-stage approach. If we had screened the entire genome at 5 cM intervals, we would have had to collect 700,000 genotypes instead of the 225,000 we completed, yet we still obtained linkage information at 5 cM intervals in the second stage of our screen. We have also used allele-sharing nonparametric methods to analyze our data, which make far fewer assumptions about modes of inheritance, allele frequencies, and disease penetrances and thus are more suited to our analysis, since those statistics are currently unknown for late-onset AD. Additionally, a greater proportion of the population is sampled with this approach than would be the case if we used only large pedigrees with many generations of affected individuals. Therefore, using this approach allows for the sample to be more representative of the general AD population and also enables a larger sample to be collected. Lastly, sibling pair studies are advantageous in that the same sample can be used to test for both linkage (in the absence of association) and association, provided that unaffected siblings are collected as we have done in our DSP sample.

One limitation of our current study is the potential for type II error. The original power calculations for this two-stage study estimated that for an $\alpha < 0.05$, 600 sibling pairs were needed to give an 80% likelihood of detecting a locus effect of $\lambda_s = 1.5$, the estimated effect

size of the APOE locus. Because of our strict diagnostic criteria, our total sample size for our stage II analysis was 451 sibling pairs. Simulations using a sample size similar to our stage II analysis (270 ASPs stage I, 451 ASPs total) give a 68% likelihood of finding both a stage I MLS = 1 and a stage II MLS = 1. This is a particular problem for our *APOE* ϵ4-negative sample, which contained only siblingships where neither sibling possessed an ϵ4 allele. There were only 56 *APOE* ϵ4-negative sibling pairs in stage I and 77 sibling pairs in stage II, yielding only a 16% likelihood of finding both a stage I MLS = 1 and a stage II MLS = 1. Thus, this sample is likely to detect only relatively large effects. In this context, it is reassuring to note that we did detect some evidence for linkage in the region containing *APOE*.

Two other linkage studies have reported results from genomewide screens for late-onset AD risk loci. Pericak-Vance et al. [1997, 1998] performed a two-stage screen, first genotyping markers at 10 cM intervals in 16 families, then examining regions with a LOD score ≥ 1 or a $P < 0.05$ in an additional 38 families. Recently, this group has rescreened the entire genome in 466 families [Pericak-Vance et al., 2000]. These studies employed both parametric and nonparametric linkage methods. This group also stratified their sample by *APOE* genotype, but used different criteria than in our study. Another study by Zubenko et al. [1998] looked for association between AD and markers spaced 10 cM apart in a series of autopsy-confirmed cases and controls. While this group screened 391 markers throughout the entire genome, the likely extent of LD around these markers ($\sim$200 Kb) means that much of the genome was not covered by their analysis.

Similar to our results, Pericak-Vance and colleagues have found linkage (either two-point parametric or multipoint nonparametric) on chromosomes 5, 6, 9, 10, 12, and 19. In their first study [Pericak-Vance et al., 1997, 1998], they obtained one two-point LOD score on chromosome 6 near our linkage, yet in their second screen [Pericak-Vance et al., 2000] their new chromosome 6 results are located 45 cM and 122 cM away from our peak. Additionally, their chromosome 12 results are not in the same location as ours; the first study [Pericak-Vance et al., 1997] reported a peak at D12S1042 and this peak has moved even further away from our linkage in a follow-up study of their results [Scott et al., 2000]. In contrast, their results on chromosomes 5 (their second peak), 9 (our peak A, their first peak), 10, and 19 appear to overlap with those from our current study in that they are located within the region spanned by our multipoint linkage peaks. In their second study [Pericak-Vance et al., 2000], their most significant linkage was located at marker D9S741, the same marker with which we obtained a peak MLS of 1.7 (their MLS = 4.31, autopsy confirmed subset). It should be noted that there is significant overlap between our sample and that of Pericak-Vance et al. [2000], perhaps as much as 80% of our sample is shared with theirs; therefore, our results are not an independent confirmation of their findings.

In their genome association study, Zubenko et al. [1998] found six markers with alleles giving significant associations, including *APOE*. Besides *APOE*, none of these markers are located underneath linkage peaks from our screen. The only significant marker in their study that maps near our results is D10S1423, which maps at the edge of our chromosome 10 linkage peak, around 35 cM from our maximum MLS. In our sample, however, we do not find any association with this marker (chi-square = 2.92, df = 7, bootstrap $P = 0.86$) despite showing robust linkage to chromosome 10.

Two other studies published recently assessed only chromosome 10 for risk loci for late-onset AD. Both Ertekin-Taner et al. [2000] and Bertram et al. [2000] were looking at specific regions of chromosome 10; however, each found different results. Bertram et al. [2000] genotyped six markers surrounding the *IDE* gene, since *IDE* is thought to be a good candidate gene for late-onset AD due to its putative role in the degradation of β-amyloid [Vekrellis et al., 2000]. They also genotyped marker D10S1225, a marker within our linkage peak. They found a maximum Z-score for the likelihood ratio of 2.1 at D10S1710 in the subset of their sample, which had ages of onset ≥ 65. This marker is located 9 cM away from *IDE* and 42 cM away from our peak; however, our original stage I data and our stage I reanalysis gave a minor peak within this region (MLS = 1.1, APOE ϵ4-negative sample, 2.2 cM away from the peak reported by Bertram et al. [2000]). For marker D10S1225, Bertram et al. [2000] obtained a Z-score of 1.6, only slightly lower than their best result.

Ertekin-Taner et al. [2000] assessed a quantitative trait related to AD. It has been shown that the processing of APP is a critical step in the pathogenesis of AD. Most mutations in the known genes involved in autosomal dominant inherited forms of AD (*APP*, *PS1*, and *PS2*) cause an increase in Aβ42 levels in patients [Scheuner et al., 1996] (see Ancolio et al. [1999] for exception), transgenic mice [Hsiao et al., 1996; Citron et al., 1997], and cell lines [Citron et al., 1992, 1997] by affecting the processing of APP. Ertekin-Taner et al. [2000] collected families with increased plasma Aβ42 levels and screened them for quantitative trait loci (QTL) controlling Aβ levels. They hypothesized that risk loci for high plasma Aβ would also be risk loci for AD and thus began their screen by focussing on the regions we had found in our first study [Kehoe et al., 1999]. They genotyped markers on chromosomes 1, 5, 9, 10, and 19 in cognitively normal individuals from families with measured plasma Aβ42 levels. Using the program SOLAR [Almasy and Blangero, 1998], the only region with a maximum MLS ≥ 1.00 was on chromosome 10, where they obtained a maximum MLS of 3.9 in a subset of five families in which the proband had plasma Aβ42 levels in the 90th percentile of the normal range. This peak maps 1.5 cM away from our linkage peak, giving us cause to believe that these two peaks are representative of the same locus. If this were true, the putative chromosome 10 gene would be causing late-onset AD by affecting Aβ42 levels.

In conclusion, we have replicated some of our stage I results along with the results on chromosomes 5, 9, 10, and 19 from other groups. Our only stage II peak meeting the criteria of Lander and Kruglyak [1995] was on chromosome 10, which was significant with a simulated genomewide $P \leq 0.05$. We are currently following these chromosome 10 results by using finer LD mapping and by testing for association with candidate genes mapping within the area.

ACKNOWLEDGMENTS

We thank the patients and their families who participated in these studies. A.G. and J.H. are supported by NIH grants AG16208; A.G. and J.C.M. by NIH grant AG5681. The M.J.O. and the J.W. groups were supported by the Medical Research Council (U.K.) and the Alzheimer's Research Trust.

APPENDIX A

Many data and biomaterials were collected in three projects that participated in the National Institute of Mental Health (NIMH) Alzheimer's Disease Genetics Initiative. From 1991 to 1998, the principal investigators and coinvestigators were as follows: Massachusetts General Hospital, Boston, MA, U01 MH46281, Marilyn S. Albert and Deborah Blacker; Johns Hopkins University, Baltimore, MD, U01 MH46290, Susan Bassett, Gary A. Chase, and Marshal F. Folstein; University of Alabama, Birmingham, AL, U01 MH46373, Rodney C.P. Go and Lindy E. Harrell.

REFERENCES

Almasy L, Blangero J. 1998. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198–1211.

Ancolio K, et al. 1999. Unusual phenotypic alteration of beta amyloid precursor protein (betaAPP) maturation by a new Val-715 → Met betaAPP-770 mutation responsible for probable early-onset Alzheimer's disease. Proc Natl Acad Sci USA 96:4119–4124.

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

Breitner JC. 1990. Life table methods and assessment of familial risk in Alzheimer's disease. Arch Gen Psychiatry 47:395–396.

Citron M, Ottersdor FT, Haass C, McConlogue L, Hung AY, Seubert P, Vigo-Pelfrey C, Lieberburg I, Selkoe DJ. 1992. Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. Nature 360:672–674.

Citron M, Westaway D, Xia W, Carlson G, Diehl T, Levesque G, Johnson-Woa K, Lee M, Seubert P, Davis A, Kholodenko D, Motter R, Sherrington R, Perry B, Yao H, Strome R, Lieberburg I, Rommens J, Kim S, Schenk D, Fraser P, St. George Hyslop P, Selkoe DJ. 1997. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. Nat Med 3:67–72.

Clayton D. 1999. A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. Am J Hum Genet 65:1170–1177.

Douglas JA, Boehnke M, Lange K. 2000. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. Am J Hum Genet 66:1287–1297.

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

Goate A, et al. 1991. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 349:704–706.

Henderson AS, et al. 1995. Apolipoprotein E allele epsilon 4, dementia, and cognitive decline in a population sample. Lancet 346:1387–1390.

Holmans P, 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.

Holmans P, Craddock N. 1997. Efficient strategies for genome scanning using maximum-likelihood affected-sib-pair analysis. Am J Hum Genet 60:657–666.

Holmes C, et al. 1999. Validity of current clinical criteria for Alzheimer's disease, vascular dementia and dementia with Lewy bodies. Br J Psychiatry 174:45–50.

Hsiao K, et al. 1996. Correlative memory deficits, abeta elevation, and amyloid plaques in transgenic mice. Science 274:99–102.

Kehoe P, et al. 1999. A full genome scan for late onset Alzheimer's disease. Hum Mol Genet 8:237–245.

Kruglyak L, Lander ES. 1995. Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454.

Kukull WA, et al. 1996. Apolipoprotein E in Alzheimer's disease risk and case detection: a case control study. J Clin Epidemiol 49:1143–1148.

Kwon J, Nowotny P, Shah PK, Chakraverty S, Norton J, Morris JC, Goate AM. 2000. tau polymorphisms are not associatied with Alzheimer's disease. Neurosci Lett 284:77–80.

Lander E, Kruglyak L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247.

Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, Weber JL, Bird TD, Schellgberg GD. 1995. A familial Alzheimer's disease locus on chromosome 1. Science 269:970–973.

Martins RN, et al. 1995. ApoE genotypes in Australia: roles in early and late onset Alzheimer's disease and Down's syndrome. Neuroreport 6:1513–1516.

Olson J, Goddard K, Dudek D. 2001. The amyloid precursor protein locus and very-late-onset Alzheimer disease. Am J Hum Genet 69:895–899.

Pericak-Vance MA, et al. 1997. Complete genomic screen in late-onset familial Alzheimer disease: evidence for a new locus on chromosome 12. JAMA 278:1237–1241.

Pericak-Vance MA, et al. 1998. Complete genomic screen in late-onset familial Alzheimer's disease. Neurobiol Aging 19 (suppl 1):S39–S42.

Pericak-Vance MA, et al. 2000. Identification of novel genes in late-onset Alzheimer's disease. Exp Gerontol 35:1343–1352.

Rice JP. 1997. The role of meta-analysis in linkage studies of complex traits. Am J Med Genet 74:112–114.

Roberts SB, Maclean CJ, Neale MC, Eaves LJ, Kendler KS. 1999. Replication of linkage studies of complex traits: an examination of variation in location estimates. Am J Hum Genet 65:876–884.

Scheuner D, et al. 1996. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. Nat Med 2:864–870.

Scott WK, et al. 2000. Fine mapping of the chromosome 12 late-onset Alzheimer disease locus: potential genetic and phenotypic heterogeneity. Am J Hum Genet 66:922–932.

Sham PC, Curtis D. 1995. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. Ann Hum Genet 59:97–105.

Sherrington R, et al. 1995. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature 375:754–760.

Vekrellis K, et al. 2000. Neurons regulate extracellular levels of amyloid beta-protein via proteolysis by insulin-degrading enzyme. J Neurosci 20:1657–1665.

Wavrant-De Vrieze F, Crook R, Holmans P, Kehoe P, Owen MJ, Williams J, Roehl K, Laliiri DK, Shears S, Booth J, Wu W, Goate A, Chartier-Harlin MC, Hardy J, Perez-Tur J. 1999. Genetics variability at the APP locus may contribute to late-onset Alzheimer's disease. Neurosci Lett 269:67–70.

Zubenko GS, Hughes HB, Stiffler JS, Hurtt MR, Kaplan BB. 1998. A genome survey for novel Alzheimer disease risk loci: results at 10-cM resolution. Genomics 50:121–128.