Human Molecular Genetics, 2003, Vol. 12, No. 23 **3133–3143**
DOI: 10.1093/hmg/ddg343

Fine mapping of the α-T catenin gene to a quantitative trait locus on chromosome 10 in late-onset Alzheimer's disease pedigrees

Nilüfer Ertekin-Taner[1,†], James Ronald[1,†], Hideaki Asahara[1], Linda Younkin[1], Maria Hella[1], Shushant Jain[1], Eugene Gnida[1], Samuel Younkin[1], Daniel Fadale[1], Yasumasa Ohyagi[2], Adam Singleton[1], Leah Scanlin[1], Mariza de Andrade[3], Ronald Petersen[4], Neill Graff-Radford[5], Michael Hutton[1] and Steven Younkin[1,*]

[1]Department of Neuroscience, Mayo Clinic Jacksonville, Jacksonville, FL, USA, [2]Department of Neurology, Neurological Institute, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan, [3]Department of Health Sciences Research, Division of Biostatistics, Mayo Clinic Rochester, Rochester, MN, USA, [4]Department of Neurology, Mayo Clinic Rochester, Rochester, MN, USA and [5]Department of Neurology, Mayo Clinic Jacksonville, Jacksonville, FL, USA

Received July 29, 2003; Revised September 23, 2003; Accepted October 7, 2003

Using plasma amyloid β protein (Aβ42) levels as an intermediate, quantitative phenotype for late onset Alzheimer's disease (LOAD), we previously obtained significant linkage at ∼80 cM on chromosome 10. Linkage to the same region was obtained independently in a study of affected LOAD sib-pairs. Together, these two studies provide strong evidence for a novel LOAD locus on chromosome 10 that acts to increase Aβ42. VR22 is a large (1.7 Mb) gene located at 80 cM that encodes α-T catenin, which is a binding partner of β catenin. This makes VR22 an attractive candidate gene because β catenin interacts with presenilin 1, which has many mutations that elevate Aβ42 and cause early onset familial AD. We identified two intronic VR22 SNPs (4360 and 4783) in strong linkage disequilibrium (LD) that showed highly significant association ($P = 0.0001$ and 0.0006) with plasma Aβ42 in 10 extended LOAD families. This association clearly contributed to the linkage at ∼80 cM because the lod scores decreased when linkage analysis was performed conditional upon the VR22 association. This association replicated in another independent set of 12 LOAD families ($P = 0.04$ for 4783 and $P = 0.08$ for 4360). Bounding of the association region using multiple SNPs showed VR22 to be the only confirmed gene within the region of association. These findings indicate that VR22 has variant(s) which influence Aβ42 and contribute to the previously reported linkage for plasma Aβ42 in LOAD families.

INTRODUCTION

Late-onset Alzheimer's disease (LOAD) is a complex genetic disorder (1) with an onset defined as >60–65 years of age. Twin studies have suggested that as much as 74% of the risk for LOAD may be due to genetic factors (2,3). It is well established that ApoE plays an important role in the genetics of LOAD (4,5). A recent study has estimated, however, that four additional loci make a contribution to the variance in LOAD age at onset that have an effect size similar to or greater than that of ApoE (6). Given the heterogeneity that is likely if at least five loci contribute substantially to the genetics of LOAD, it is not surprising that the large number of single gene association studies that have been performed have so far failed to identify variants that associate reproducibly with LOAD. Other reasons for failure of replication are possible gene–gene, gene–environment interactions that can be difficult to account for and poor study design.

Several groups (7–9) have suggested that quantitative, intermediate phenotypes may be useful in unraveling the

*To whom correspondence should be addressed at: Mayo Clinic Jacksonville, 4500 San Pablo Road, Birdsall 2, Jacksonville, FL 32224, USA. Tel: +1 9049537356; Fax: +1 9049537370; Email: younkin.steven@mayo.edu
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

complex genetics of disorders such as LOAD. Clinical disease states such as LOAD are likely to result from an uncertain, complex sequence of pathophysiologic events whereas intermediate phenotypes may be more directly related to the action of a set of genes that are more easily identified. Thus the identification of genes with variants that contribute to specific intermediate phenotypes may be a tractable way to identify at least some of the genes that contribute to diseases governed by complex genetics. To optimize the utility of this approach, it is important to select intermediate phenotypes that are highly heritable and have a close biological link to the disease in question.

In our studies, we have focused on levels of the amyloid β protein (Aβ), specifically Aβ42, as a quantitative intermediate phenotype for LOAD. Aβ is a secreted protein (10–12) derived from a set of large precursor proteins collectively referred to as the amyloid β protein precursor. Secreted Aβ, which is produced by virtually all cells, is readily detected in human plasma and cerebrospinal fluid. Most secreted Aβ has forty amino acids (Aβ40) but a small fraction (5–10%) has two additional amino acids at the carboxyl terminus (Aβ42) (13–15). In the brains of all patients with AD, a large amount of Aβ aggregates and is deposited in senile plaques, much of it in the form of highly insoluble amyloid fibrils. To make a definite diagnosis of AD, large numbers of senile plaques and neurofibrillary tangles must be demonstrated in the brain at autopsy (16). In all LOAD patients, a large amount of the Aβ deposited in senile plaques ends at Aβ42. In one-third of cases, Aβ ending at Aβ42 is essentially the only form of Aβ deposited, in one-third it is the predominant form and in one-third a large amount of Aβ40 is also deposited (17). It is worth noting that synthetic Aβ40 and Aβ42 both spontaneously assemble into amyloid fibrils *in vitro*, but Aβ42 forms amyloid fibrils far more rapidly than Aβ40 (18,19). The brain has very high levels of APP mRNA; the concentration of Aβ in cerebrospinal fluid is 50 times higher than in plasma; and mixed fetal brain cultures produce far more Aβ than other types of cultured cells (12,20,21). Thus it is likely that senile plaques form specifically in the brain because Aβ levels are higher in the brain than in other organs.

The APP, presenilin 1 and presenilin 2 genes have many, fully penetrant mutations that cause early onset familial AD (EOFAD) (22–25). Apart from its early onset and causation by specific mutations, EOFAD is clinically and pathologically indistinguishable from LOAD. We and others have shown that the EOFAD mutations all increase Aβ42 and sometimes increase Aβ40 as well (15,26–30). Importantly, we observed plasma Aβ42 elevations both in EOFAD patients and in young, cognitively normal carriers of EOFAD mutations (31). We also observed Aβ42 elevations in plasma from non-demented relatives of LOAD patients between the ages of 20 and 65 (32). Importantly, these elevations were not associated with the ApoE4 allele or with variants in any of the EOFAD genes (32). Finally, Mayeux *et al.* (33) found in a longitudinal study of unrelated elderly individuals, that those who subsequently developed AD had higher plasma levels of Aβ42 than did those who remained free of dementia. Collectively these findings show that Aβ42 is closely linked to AD. To investigate the heritability of plasma Aβ, we analyzed extended LOAD families and found that plasma Aβ40 and Aβ42 levels both have substantial heritability (34).

Given these findings that the level of plasma Aβ42 was likely to be useful as an intermediate, quantitative phenotype for LOAD, we performed variance components linkage analysis on this trait in a group of extended LOAD pedigrees. This analysis showed significant linkage to a locus at ∼80 cM on chromosome 10 (35). Linkage to the same region was obtained independently in a study of affected LOAD sib-pairs (36). Together, these two studies provide strong evidence for a novel LOAD locus on chromosome 10 that acts to increase Aβ42. The sib-pair group estimated the effect of this locus to be equivalent to that of ApoE, suggesting that it was likely to be a major risk factor for LOAD. A third group found linkage in a more downstream region of chromosome 10 and focused on the insulin-degrading enzyme (IDE) as a possible candidate gene in this region (37). More recently, a fourth group using age of onset as a quantitative phenotype in AD and Parkinson's disease (PD) families found suggestive linkage to chromosome 10 at 133–135 cM (38).

An important advantage of the variance components approach employed here is that candidate genes, identified through their likely effect on Aβ42, can be evaluated systematically to determine how much various within-gene variants contribute to linkage. In this study, we analyzed VR22 for variants that contribute to the linkage signal on chromosome 10. Both its location and function make VR22 a strong candidate gene. Located at 80 cM at the peak of the previously reported linkage signal, VR22 encodes a novel α catenin, α-T catenin, which was recently cloned by Jannsens *et al.* (39). The α catenins interact closely with β catenin and play an important role in intercellular adhesion as strong intercellular adhesion depends on linkage of the cadherins to the actin cytoskeleton via the α/β catenin complex. α-T catenin is closely homologous to α-N catenin, which is neuron specific, and to α-E catenin, which is expressed in all tissues. α-T catenin is expressed at high levels in testis and heart, but it is also expressed in human and mouse brain (39). α-T catenin has overall sequence identity of 56 and 58% with α-E and α-N catenin respectively. Importantly, Jannsens *et al.* (39) have shown that that α-T catenin interacts with β catenin in a functionally effective fashion. This makes VR22 a strong functional candidate because β catenin interacts with presenilin 1 (40), which has many mutations that increase Aβ42 (26,29,31) and cause early onset familial AD (24). Zhang *et al.* (40) has shown that loss of signalling through the β catenin pathway leads to enhanced neuronal vulnerability to Aβ. In addition, β catenin and presenilin forms a complex that leads to stabilization of the former, such that destabilization of this complex leads to β catenin degradation and enhanced neuronal apoptosis. Presenilin mutations were shown to lead to destabilization of the β catenin–presenilin complex, degradation of the β catenin and enhanced Aβ production. Given that α catenins provide the link between β catenin and its binding counterparts (such as presenilins) to the actin cytoskeleton, it is possible to postulate that mutations in α catenins could lead to disruption of the stability of the β catenin–presenilin complex resulting in enhanced neuronal vulnerability of cells to Aβ and increased Aβ production.

Table 1. Association of VR22 SNPs with plasma Aβ42 levels in extended LOAD families

Sets of families analyzed				10 LOAD families	12 LOAD families	22 LOAD families
Number of subjects				204	88	292
Analysis conditional on linkage[a]				Yes	no	no
Name	Alleles[b]	Location[c]	Frequency[d]	*P*-value for association		
4825	C/A	61136997	0.62/0.38	0.06	0.98	0.10
4783	A/G	61137039	0.76/0.24	0.0006	0.04	0.0000059
4360	T/C	61137462	0.74/0.26	0.0001	0.08	0.0000021

[a]To control for linkage in the 10 LOAD families, the MIBD that yielded the maximum multipoint lod score was used.
[b]The common allele/rare allele.
[c]The locations are in base pairs and determined according to the chromosome 10 genomic map in Celera.
[d]The frequency of the common allele/rare allele.

RESULTS

Association between VR22 SNPs and plasma Aβ42

In our initial assessment of VR22, we examined three SNPs [4825 (C/A), 4783 (A/G) and 4360 (T/C)] in the intron between exons 10 and 11. These SNPs (Table 1), which spanned 465 bp, were genotyped in the 10 extended LOAD families previously analyzed for linkage (35). In order to minimize age and dementia-related environmental effects on plasma Aβ42 levels, we performed all of our analyses on the 20–65 year age group, as in our previous studies (34,35). The effect of each of the three VR22 SNPs was analyzed by encoding it as a covariate and testing its effect on the variance of plasma Aβ42 in a regression style approach implemented in SOLAR (Table 1). To control for the effect of linkage, the linkage component was included in both the null and alternative models tested. The 4360 (T/C) and 4783 (A/G) SNPs, which were in strong linkage disequilibrium, both showed strong association ($P = 0.0001$ and 0.0006, respectively). The 4825 SNP showed only marginal association ($P = 0.06$, Table 1). All three SNPs were also examined in a second set of 12 additional LOAD families (Table 1) (34). The VR22 4783 SNP showed significant association ($P = 0.04$), VR22 4360 SNP yielded marginal association ($P = 0.08$) and VR22 4825 SNP did not yield significant association ($P = 0.98$) in this additional and smaller dataset. When all 22 families were combined, the significance of association reached $P = 5.9 \times 10^{-6}$ for 4783 and 2.1×10^{-6} for 4360 SNPs (Table 1). Figure 1 depicts the box plots for the 10*logAβ42 levels in the 22 families stratified by the most significant VR22 4360 SNP genotypes. Consistent with the association results, a genotype dependent effect on plasma Aβ levels is observed, where 4360 TT genotype is associated with the highest Aβ42 levels, CT with intermediate levels and CC with lowest levels.

To determine if other SNPs in this region of the VR22 gene show stronger association, we analyzed 10 additional SNPs for a total of 13 SNPs analyzed in a region spanning 8.8 kb (Table 2, hCV3096482–hCV11295092). All of the SNPs in this 8.8 kb region were in linkage disequilibrium. Many SNPs in this region showed nominally significant association, but none showed association stronger than that observed with 4360 and 4783 (Table 1).

Bounding of the association between VR22 SNPs and plasma Aβ42

To bound the region of VR22 4360/4783 association, we evaluated a total of 49 additional SNPs spaced on either side of the 4360 and 4783 SNPs. These SNPs were selected to be in several groups, which were designed to span the entire VR22 gene at ~250 kb intervals as well as the region 3′ of VR22. The known exons in *VR22* were 43.3 kb (median) away from the closest SNP in the closest SNP group (mean = 53.8 kb, 22–149 kb). Given that the extent of LD in the human genome was estimated in different studies to be anywhere between 10 and >100 kb (41–45); and that this extent in any one genomic region is not predictable *a priori*, it is possible to miss association signals due to variants in-between our SNP groups. However, given the size of *VR22*, and our extent of coverage throughout and outside this gene, association signals we did identify with the 4360–4783 *VR22* SNPs will be well-bounded within *VR22* with this SNP genotyping design. Three to six SNPs were analyzed in each group, and linkage disequilibrium between the most significantly associating 4360 SNP and the other SNPs was analyzed in the 22 combined families (Fig. 2A). Using a covariate style regression approach, each SNP was analyzed for association with plasma Aβ42 in the 10, 12 and 22 combined LOAD families (Table 2 and Fig. 2B).

On both the 3 and the 5 sides of the strongly associating 4360 SNP, LD with flanking SNPs fell to essentially baseline levels within the large VR22 gene. On the 5 side, SNP association with Aβ42 also decreased to baseline levels within the VR22 gene (Fig. 2B). On the 3 side, the significance of SNP association decreased by more than 2 orders of magnitude (from 0.0001 for 4360 to 0.03 for the most significantly associated SNP in the nearest SNP group hCV1380042) (Fig. 2B and Table 2). Association was at baseline levels on the 3 side ~1 000 kb from the 3 margin of VR22.

This bounding approach effectively limits the region of potential association on both the 3 and 5 sides of 4360 to within VR22. Thus, our data provide strong evidence that the association of the 4360 and 4783 SNPs with plasma Aβ42 is due to variant(s) in the VR22 gene.

Given that a total of 51 SNPs were genotyped, the issue of multiple testing can be raised. Many of the SNPs are in linkage disequilibrium and therefore are not independent from one another (Fig. 2B). Therefore, a simple Bonferroni correction

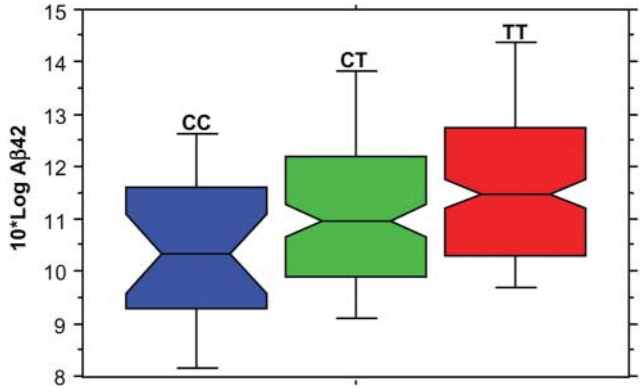

Figure 1. Dose-dependent effect of VR22 4360 SNP on plasma Aβ42 levels in extended LOAD families. Box plots depicting the median (middle part of the box), 25th percentile (lower limit of box), 75th percentile (upper limit of box), 10th percentile (lower end of line) and 90th percentile (upper end of line) for the 10*logAβ42 levels of the 20–65-year-old family members in the 22 LOAD families. The plots are depicted separately for VR22 4360 genotypes TT (red), CT (green) and CC (blue).

under the assumption of 51 independent tests is overly conservative for this study. Nonetheless, even when Bonferroni correction is applied, the 4783 and 4360 SNPs are still significant in the 10 families ($P = 0.03$ and 0.005) and the 22 combined families ($P = 0.0003$ and 0.0001). The association in the smaller set of 12 families are no longer significant after this overly conservative correction ($P = 2$ and 4 for 4783 and 4360, respectively), nonetheless their inclusion in the analysis leads to improved evidence for association.

Contribution of VR22 association to linkage at 80 cM

We next investigated the contribution of the VR22 4360/4783 association to the linkage that we previously reported. Our initial linkage findings were obtained using seven markers covering 14 cM on chromosome 10. To analyze the entirety of chromosome 10, we genotyped 26 additional markers spanning 165 cM in the 10 LOAD families. When all 10 families were combined, the best two point lod scores were obtained with 4783 (2.62) and 4360 (2.17). The peak described by the multipoint lod scores broadened and flattened considerably with the incorporation of the additional markers, but the maximum multipoint lod score continued to be at ~80 cM (Fig. 3).

The reason for the broader and flatter peak in this fine-mapping analysis may be several-fold. First, we utilized the MCMC MIBD estimation implemented in the Simwalk2 package (46,47) for the multipoint linkage results presented in this paper as opposed to the regression-based MIBD estimation algorithm implemented in SOLAR, which was used in our original paper (34, see also Subjects and Methods). Second, genotyping of additional markers leads to more information regarding recombination that could potentially lead to a decrease in the original linkage signal obtained by fewer markers. Third, utilizing a larger number of markers increases the likelihood of map or genotyping errors, despite the implementation of strict genotype-quality control measures. Indeed, others have also found that fine-mapping with more markers can lead to a lowering in their original lod scores obtained with a lesser number of markers (48,49).

To evaluate the extent to which the 4360 and 4783 SNPs [or functional SNP(s) associated with this tightly linked pair] account for the linkage observed, we repeated linkage analysis after regressing out the effect of the VR22 4360 association by including it as a covariate in the linkage model. After regressing out the effect of this SNP, the multipoint linkage signal diminished substantially in the 10 combined families (Fig. 3). This marked diminution in lod scores indicates that a substantial proportion of the linkage signal for plasma Aβ42 is accounted for by the VR22 4360 association. The residual linkage signal that persisted after regressing out the VR22 4360 association indicates that there are likely to be additional variants that contribute to linkage, though this residual signal may also be due to random variation. The diminution in the maximum multipoint lod scores, upon inclusion of the 4360 SNP as a covariate, is also evident from the comparison of the variance component model that excludes this covariate (Table 3B) to the one that includes it (Table 3C). The *P*-value for the major gene effect at the maximum multipoint locus goes from 0.004 to 0.16. The same trend is observed when the twopoint linkage for the 4360 SNP is assessed with and without including 4360 as a covariate (Table 3D and E). The *P*-value for the major gene at the 4360 locus loses significance from $P = 0.009$ to $P = 0.038$.

DISCUSSION

In this study, we analyzed the VR22 gene on chromosome 10 to determine if it has functional SNP(s) that contribute to the linkage signal for plasma Aβ42 that we previously reported in extended LOAD families. We found two VR22 SNPs, 4360 and 4783, which showed significant, robust association with plasma Aβ42 in 20–65-year-old family members of the extended LOAD families. These two SNPs are 423 bp apart, and they are in such strong LD that they give virtually identical results. The 4360 and 4783 SNPs are located in the 177 kb intron between exons 10 and 11 at a distance of ~1.4 kb from exon 11. We identified 11 additional SNPs within 6 kb of 4360/4783, eight in the intron between exons 10 and 11 and three in the intron between exons 11 and 12, but none of these SNPs gave more significant association than 4360/4783. Careful bounding showed that the large (1.7 Mb) VR22 gene is the only confirmed gene within the region of 4360/4783 association. Analyses in which 4360 was used as a covariate in both the null and linkage models reduced the multipoint lod score indicating that the VR22 4360/4783 association contributes to linkage. The large reduction observed when all 10 families were analyzed suggests that the 4360/4783 association contributes substantially to the linkage signal observed in these families. The presence of residual linkage after regressing out the effect of VR22 4360 SNP indicates, however, that the 4360 association does not entirely account for linkage. Thus our data provide substantial evidence that VR22 variant(s) contribute to the linkage signal we previously reported. Though the residual linkage signal may be due to random variation, it could also indicate that there may be additional variants that contribute to linkage. Whether these variants are in

Table 2. *P*-values for association with plasma Aβ42 in 20–65-year-old subjects from extended LOAD families

Sets of families analyzed				10 LOAD families	12 LOAD families	22 LOAD families
Number of subjects				204	88	292
Analysis conditional on linkage[a]				yes	no	no
Name	Alleles[b]	Location[c]	Frequency[d]	*P*-value for association		
Group1						
hCV1938265	C/T	59846790	0.62/0.38	0.62	0.83	0.33
hCV3132915	A/C	59849109	0.63/0.37	0.74	0.84	0.44
hCV3132902	A/G	59854715	0.77/0.23	0.67	0.96	0.93
hCV3132900	G/A	59855508	0.62/0.38	0.74	0.85	0.79
Group2						
hCV1380042	C/T	60955105	0.61/0.39	0.03	0.02	0.002
hCV1380044	T/C	60957122	0.72/0.28	0.05	0.14	0.02
hCV1380046	T/A	60960024	0.78/0.22	0.23	0.30	0.26
Group3						
hCV3096482	G/A	61132928	0.62/0.38	0.08	0.88	0.14
hCV3096481[e]	C/T	61133181	0.55/0.45	0.03	0.68	0.06
hCV3096480	A/G	61134482	0.84/0.16	0.22	0.09	0.03
4825	C/A	61136997	0.62/0.38	0.06	0.98	0.10
4783	A/G	61137039	0.76/0.24	0.0006	0.04	0.0000059
4360	T/C	61137462	0.74/0.26	0.0001	0.08	0.0000021
hCV3096478	T/C	61137817	0.65/0.35	0.03	0.94	0.06
hCV3096476	G/A	61137976	0.89/0.11	0.33	0.28	0.12
hCV3096475	A/G	61139207	0.78/0.22	0.05	0.93	0.13
hCV11295087	T/C	61139465	0.88/0.12	0.80	0.11	0.16
1332TC	T/C	61140488	0.67/0.33	0.15	0.18	0.06
hCV11295091	C/G	61141464	0.92/0.08	0.21	0.02	0.02
hCV11295092	A/C	61141737	0.69/0.31	0.18	0.22	0.07
Group 4						
hCV1848564	T/C	61212363	0.79/0.21	0.62	0.52	0.31
hCV1848569	T/C	61217316	0.92/0.08	0.14	0.04	0.03
hCV1848570[e]	C/G	61218129	0.52/0.48	0.73	0.22	0.26
hCV1848573	C/T	61220578	0.65/0.35	0.32	0.94	0.45
hCV1848575	G/A	61223700	0.92/0.08	0.10	0.04	0.02
Group 5						
hCV11848223	A/T	61449609	0.53/0.47	0.66	0.22	0.96
hCV397228	T/C	61456935	0.85/0.15	0.88	0.24	0.59
hCV11848234	G/A	61457375	0.93/0.07	0.55	0.03	0.06
hCV11848235[e]	A/T	61457837	0.91/0.09	0.36	0.0003	0.01
hCV397225[e]	G/A	61459168	0.90/0.10	0.97	0.96	0.99
Group 6						
hCV1635418	A/G	61724305	0.88/0.12	0.36	0.29	0.83
hCV1635417	T/C	61724697	0.88/0.12	0.14	0.27	0.49
hCV1635413	T/G	61730993	0.88/0.12	0.51	0.10	0.64
Group 7						
hCV1974464	T/C	61999466	0.78/0.22	0.59	0.33	0.80
hCV1974463	C/A	61999560	0.77/0.23	0.30	0.21	0.73
hCV1974461	T/C	62002888	0.78/0.22	0.16	0.16	0.53
hCV1974459	G/A	62007913	0.79/0.21	0.13	0.47	0.29
Group 8						
hCV1853789	T/C	62252091	0.89/0.11	0.70	0.05	0.12
hCV1853790	A/G	62252217	0.89/0.11	0.63	0.02	0.09
hCV1903901	G/C	62252793	0.72/0.28	0.58	0.71	0.30
hCV1853792	C/G	62254617	0.89/0.11	0.73	0.04	0.12
hCV1853793	G/A	62255127	0.90/0.10	0.53	0.05	0.09
hCV1853794	C/T	62255454	0.87/0.13	0.97	0.42	0.59
Group 9						
hCV3023768	G/A	62529141	0.90/0.10	0.25	0.47	0.56
hCV3023767	G/C	62529663	0.96/0.04	0.34	0.54	0.39
hCV3023766	C/A	62530064	0.96/0.04	0.69	0.87	0.63
hCV1431646	G/A	62531335	0.96/0.04	0.51	0.82	0.41

Table 2. (*Continued*)

Sets of families analyzed				10 LOAD families	12 LOAD families	22 LOAD families
Number of subjects				204	88	292
Analysis conditional on linkage[a]				yes	no	no
Name	Alleles[b]	Location[c]	Frequency[d]	*P*-value for association		
Group 10						
hCV1340375	G/A	62926002	0.60/0.40	0.93	0.86	0.95
hCV1340376	G/A	62926743	0.84/0.16	0.18	0.51	0.57
hCV1340395	C/T	62945396	0.61/0.39	0.60	0.90	0.68
hCV1340396	G/C	62948122	0.62/0.38	0.64	0.83	0.65

[a]To control for linkage in the 10 LOAD families, the MIBD that yielded the maximum multipoint lod score was used.
[b]The common allele/rare allele.
[c]The locations are in base pairs and determined according to the chromosome 10 genomic map in Celera.
[d]The frequency of the common allele/rare allele. These are obtained using the founder population in the extended families ($n = 127$–149 for various SNPs).
[e]Of the 51 SNPs tested, these four SNPs were in Hardy–Weinberg disequilibrium. All other SNPs were in Hardy–Weinberg equilibrium. HWE was tested using the observed and expected genotype frequencies in the founder population.

the large VR22 gene or in nearby gene(s) remains to be determined.

Since they are intronic SNPs located ~1.4 kb from the nearest exon, 4360 and 4783 are unlikely to have a functional effect, and their association with Aβ42 presumably occurs because they are in LD with one or several functional VR22 SNPs. Our sequencing of all VR22 exons in the only reported VR22 transcript (39) failed to identify any potentially functional SNP(s) that could account for the 4360/4783 association. It is possible that there are functional SNPs that affect the splicing of known or alternatively spliced exons, in strong LD with 4360/4783.

The basic premise in studying plasma Aβ42 as an intermediate phenotype is that genes with variants that influence plasma Aβ42 will be genes with variants that influence risk for LOAD. The results reported here are useful because they identify VR22 as a gene with variants that influence plasma Aβ42, thereby increasing the likelihood that VR22 is the gene (or one of the genes) in the 80 cM region of chromosome 10 with variants that influence risk for LOAD. An important, unanswered question is how well the effect of a SNP on plasma Aβ42 will correlate with its effect on the AD phenotype. It is unlikely that the change in plasma Aβ42 caused by a specific genetic variant directly increases risk for AD; increased risk for AD is probably produced by a related effect that the variant has within the brain. Thus it is conceivable, perhaps even likely, that a specific variant with a strong influence on plasma Aβ42 may sometimes have little or no influence on the AD phenotype and vice versa. It is only by thoroughly studying several genes and elucidating the effects of the entire set of functional variants on the plasma Aβ42 and AD phenotypes that this question can be answered. It may be that effects on the two phenotypes correlate well for the variants in some genes and poorly in others.

The results reported here indicate that VR22 variants significantly associate with plasma Aβ42 and account for a substantial portion of our linkage signal on chromosome 10. Additional analysis is needed to identify the entire set of functional variants that account for the Aβ42 linkage at 80 cM. Once they are identified it will be important to test their effects on Aβ production and VR22 gene function by biological experiments; as well as to determine whether these functional variants modify the risk for AD.

SUBJECTS AND METHODS

Subjects

We collected 10 extended LOAD families composed of 292 individuals, 205 of whom were between the ages of 20 and 65. Four of these families were collected via a LOAD patient who had extremely high plasma Aβ42 and/or Aβ40 levels ('extremes', top 10th percentile of the 545 AD patients analyzed). The remaining six families were ascertained via a LOAD proband, who had multiple relatives with LOAD, without prior Aβ measurements ('non-extremes'). One of the probands from the 'non-extreme' families was determined to have extremely high plasma Aβ levels after the family collection; therefore this family was grouped with the 'extreme families' subsequently. Except for this one family, the distinction between the 'extreme' and 'non-extreme' families were made *a priori*, before the family collection, based on the plasma Aβ42 levels of the proband. The detailed collection strategies and family profiles for these families are provided elsewhere (34,35). In addition to these 10 families, 12 independent extended LOAD pedigrees were collected. These additional pedigrees, all of which had a proband who was a first degree relative of a LOAD patient and who had elevated plasma Aβ levels (34), contributed 88 subjects between the ages of 20 and 65. This study was approved by the appropriate institutional review board; and appropriate informed consent was obtained from all participants.

Microsatellite genotyping

DNA was extracted from peripheral blood leukocytes using routine methods. Genotypes for the microsatellites were obtained using ABI 377 and 3100 sequencers and associated Genescan/Genotyper software packages. Single nucleotide polymorphisms (SNPs) were genotyped using the TaqMan chemistry and software for designing primers and probes implemented within the ABI PRISM 7900 HT Sequence Detection System. Pedigree structure, phenotypic and genotypic information were maintained in PEDSYS (http://www.sfbr.org/sfbr/public/software/pedsys/pedsys.html) (50), which also produced output files for the analyses performed with

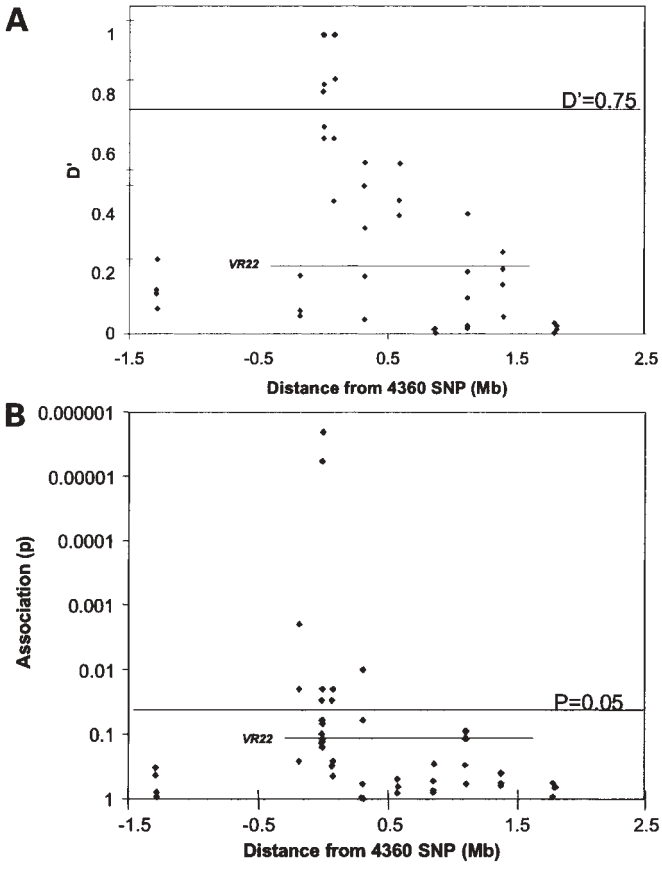

Figure 2. Boundaries of VR22 4360/4783 association. (**A**) Marker–marker linkage disequilibrium in which all SNPs are analyzed with respect to the 4360 SNP. Horizontal line shows the position of the 1.7 Mb VR22 gene. Dashed line shows D = 0.75. (**B**) Marker–phenotype (plasma Aβ42) association. Horizontal line shows the position of the 1.7 Mb VR22 gene. Dashed line shows $P = 0.05$.

SOLAR, as explained below. The order and location of 27 chromosome 10 microsatellite markers were determined from the Marshfield linkage map. Two markers, D10S1435 and D10S1211, which did not have locations on the Marshfield linkage map (51), were placed on the linkage map according to the location estimates obtained from the MAP-O-MAT program (http://compgen.rutgers.edu/mapomat/) (52). Two SNPs found in the *VR22* gene were also used as markers in linkage analysis. These SNPs, 4360TC and 4783AG, are located at 61137462 and 61137039 bp, respectively, on the chromosome 10 genomic map in the Celera database (http://www.celera.com). The two SNPs were placed on the linkage map based on their approximate centimorgan locations estimated from the physical maps.

Identification and genotyping of SNPs

The VR22 genomic and mRNA sequences were initially obtained from the National Center for Biotechnology Information (NCBI) web site using the genomic contig NT_008663 and AF091606. In order to identify SNPs, primers were designed to cover all known exons and a 7 kb intronic region between exons 10 and 11. LOAD cases, elderly control

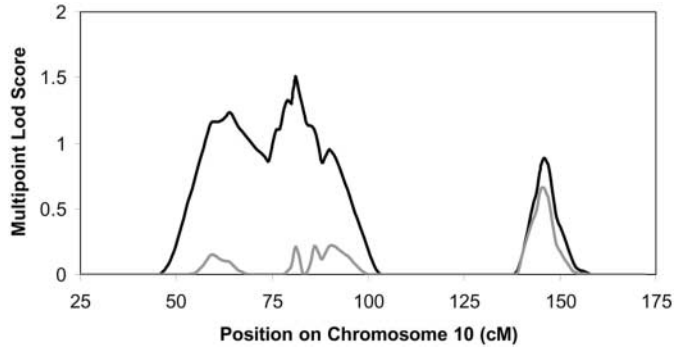

Figure 3. Effect of the VR22 SNP 4360 on the multipoint lod scores (MLS) in the LOAD families. Effect of VR22 4360 on the linkage of plasma Aβ42 in 10 LOAD families. Black line: linkage; gray line: linkage conditional upon association with the VR22 4360 SNP. The multipoint IBDs used in these analyses are generated using the 31 markers as described in the text.

subjects and family members from the extended LOAD pedigrees were then sequenced to identify polymorphisms. The initial sequencing of the VR22 exons and the 7 kb intronic region identified three SNPs in the intronic 7 kb region (4825, 4783 and 4360). Subsequently, many additional SNPS in VR22 and flanking genes were identified using the Celera database. Our aim was to cover this region using SNP groups that are located at every ~250 kb, with each group composed of 3–5 SNPs separated by 1–5 kb. A total of 51 SNPs, including VR22 4360 and 4783 were genotyped throughout VR22 and the 3 (ANXA2P3) and 5 (SIRT1) flanking genes. The SNPs within each group were found to be in strong LD with each other. The SNP locations were obtained from the Celera sequence map.

Variance components analysis

We employed the variance components methodology, implemented in the software package SOLAR (53,54) to estimate the heritability of plasma Aβ42 levels and to perform linkage and association analyses to detect quantitative trait loci (QTLs) that affect the variance of Aβ42 in our extended LOAD families. This method, described in detail elsewhere (53,54), estimates the amount of variance in a quantitative trait due to a particular genetic locus (σ^2_q), residual genetic factors (σ^2_g) and individual-specific, random environmental factors (σ^2_e), based on the phenotype covariance between arbitrary relative pairs. The covariance matrix takes the form

$$\Omega = \sigma^2_q \Pi + 2\sigma^2_g \Phi + \sigma^2_e I$$

where π is the matrix of the probabilities that each pair of individuals share genetic material identical by descent for a given chromosomal location (estimated from the genetic marker data), 2Φ is the matrix of kinship coefficients describing polygenic factors and I is the identity matrix describing sporadic environmental factors.

In addition, the effect of shared-environment (σ^2_h) on the variance of the trait can also be estimated by including a household matrix in the model. This is a matrix whose ijth element is equal to 1 if individuals i and j share a specified environment, and 0, otherwise. Our family collection is largely

Table 3. Variance components models depicting the effect of the 4360 SNP on multipoint and twopoint linkage

	Estimates	SE	*P*-values
A. No-covariate-no-linkage			
Mean	11.19	0.33	
SD	1.77	0.15	
H2R	0.36	0.17	0.008
C2	0.24	0.12	0.0004
E2	0.40	0.15	0.0002
H2Q	NA	NA	NA
βCOV	NA	NA	NA
B. No-covariate (MIBD)			
Mean	11.25	0.32	
SD	1.79	0.15	
H2R	0	NA	0.50
C2	0.22	0.12	0.001
E2	0.33	0.10	0.002
$H2Q^{mibd}$	0.46	0.13	0.004
βCOV	NA	NA	NA
C. Covariate-linkage (MIBD)			
Mean	11.30	0.32	
SD	1.69	0.15	
H2R	0.02	0.28	0.48
C2	0.26	0.13	0.0005
E2	0.46	0.18	0.0004
$H2Q^{mibd}$	0.27	0.26	0.16
βCOV	0.83	0.20	0.0001
D. No-covariate (IBD 4360)			
Mean	11.19	0.32	
SD	1.78	0.14	
H2R	0.00	NA	0.5
C2	0.20	0.12	0.00225
E2	0.21	0.11	0.00460
$H2Q^{4360}$	0.59	0.15	0.00090
βCOV	NA	NA	NA
E. Covariate-linkage (IBD 4360)			
Mean	11.26	0.32	
SD	1.69	0.15	
H2R	0	NA	0.5
C2	0.24	0.13	0.00081
E2	0.34	0.16	0.00056
$H2Q^{4360}$	0.42	0.19	0.03766
βCOV	0.83	0.21	0.00009

SE: standard error (mean and standard deviation of the quantitative phenotype, plasma Aβ42); H2R: additive genetic variance due to residual genetic factors; C2: variance due to shared family effects; E2: variance due to random environmental factors; H2Q: variance due to major gene effects (panels B and C MIBD matrix that yields the highest lod score at ∼80 cM is used, panels C and D IBD matrix for the VR22 4360 marker is used); βCOV: coefficient for the covariate (4360 covariate).

composed of extended pedigrees, where we have measured the plasma Aβ of most members from each family in a single batch of ELISAs. To rigorously account for any possible assay batch to batch variations that may affect the variance in plasma Aβ42, we included a household matrix in the models. Since almost all individuals from each family were measured in one batch, the household matrix included an element = 1 for those individuals from the same extended family and 0, otherwise. We realize that by doing so, we are probably being overly conservative. Since the shared-family (household) effect employed to account for potential batch to batch variance in the assay overlaps with the kinship effect (55), inclusion of the household matrix almost certainly overcorrects for batch to batch variance and underestimates the shared-genetic effect. For this reason, we performed the heritability and linkage analyses, both with and without the household matrix in the model. The multipoint results shown in Figures 2 and 3 include the household matrix. The analyses without the household matrix yielded slightly better lod scores and higher heritabilities in general.

The scalars σ^2_q, σ^2_g, σ^2_h and σ^2_e are estimated using maximum likelihood. For heritability estimations, we compared an environmental model where σ^2_q and σ^2_g were fixed at zero and σ^2_h and σ^2_e were estimated, to the polygenic model where σ^2_g, σ^2_h and σ^2_e were estimated. The linkage analyses were performed by comparing the polygenic model to the linkage model where σ^2_q was also estimated. The difference between the two log10 likelihoods for the models produces a lod score that is equivalent to the classical lod score of linkage analysis. We incorporated sex, age, age2, sex*age, and sex*age2 as covariates in the heritability estimations. None of these covariates were significant at $P < 0.05$ therefore, they were excluded from the multipoint linkage analysis.

In order to minimize age and dementia-related environmental effects on plasma Aβ42 levels, we performed all of our analyses on the 20–65 year age group, as in our previous studies (34,35). Since variance components methods are based on the underlying assumption that the quantitative trait being analyzed has a normal distribution, 10log(Aβ42) levels were used as the phenotypes. Extreme outliers with plasma Aβ levels ± 4 standard deviations beyond the mean were excluded from the analyses. Thus the 10log(Aβ42) phenotype analyzed had a normal distribution. Plasma Aβ measurements were done using a well established sandwich ELISA system as described previously (15). All samples were measured in duplicate. Each ELISA plate included samples from five volunteers. The results from these five standard samples were used to normalize the values obtained so as to reduce plate-to-plate and day-to-day variance.

Identity-by-descent probability estimations

The identity-by-descent (IBD) probabilities at each genotyped marker were estimated using the pairwise likelihood-based estimation method implemented in SOLAR (53). These IBD estimates were used to obtain the twopoint lod scores. Multipoint IBD probabilities were also estimated at each centimorgan between each pair of adjacent markers using the Markov chain Monte Carlo (MCMC) multipoint IBD estimation approach implemented in the Simwalk2 package (46,47). The MCMC process was run twice, once with the default length of 1000 replicates and a second time by doubling this process length. Multipoint linkage analyses using these MIBD matrices were done using the variance components methodology implemented in SOLAR. The results using the shorter and longer runs were similar. The results of the latter are presented here. Given that the computation time increases with the marker number and that our two most upstream markers yielded no evidence for linkage in any of the datasets, these two markers (D10S1435 and D10S189) were omitted from the MIBD estimations. There are two major differences between the MIBD estimations utilized in our previous report of linkage on chromosome 10 and the current study (35). The results of

the previous linkage were obtained as result of genotyping seven markers in a ~13 cM region, and the current study utilizes 31 markers spread from 28.3 to 165.3 cM on chromosome 10. Second, the MIBD estimation approaches that we used in the previous and current studies are different. The MCMC approach implemented in Simwalk2 (46,47) that we utilized in this study is different from the MIBD probability estimation algorithm implemented in SOLAR, which was the approach we utilized in our initial analysis of seven chromosome 10 markers. Sobel *et al.* (47), demonstrated significant differences between the results obtained by these two methods using the same dataset and concluded that the multipoint MCMC approach yielded more accurate MIBD probability estimates for inter-marker loci. For this reason, we utilized the MCMC approach in this study.

Tests of association

We initially tested for association between the SNPs and plasma Aβ42 in the 10 extended LOAD families using a regression based covariate approach in the variance components paradigm implemented in the computer program SOLAR (53). In order to control for the effect of linkage, we included the linkage component ($\sigma^2_q\pi$) in both the null and alternative models tested. We used the multipoint IBD matrix (π) that yielded the maximum multipoint lod score at ~80 cM in the 10 extended LOAD families. The multipoint IBD probabilities in the matrix π, were estimated in Simwalk, using the genetic marker information as described above.

To test the effect of specific SNP genotypes, we used the measured genotype approach where we included the copy number of the SNP allele as a covariate in the analysis and regressed the plasma Aβ42 phenotype on this covariate for each individual. The measured genotype approach is a classical test for differences in trait mean by genotype described theoretically previously (56,57). This approach was also utilized in previous association studies of SNPs and various quantitative traits in extended families (58,59). For individuals with an XX, XY or YY SNP genotype, we used a covariate that takes the value of 0, 1 or 2, respectively. This formulation assumes that the SNP is an additive QTL. The model that includes the covariate is tested against the model that does not include it to determine its effect on the quantitative trait of interest. Twice the difference in natural log likelihoods of these two models is distributed as a χ^2 distribution with 1 degree of freedom. In order to test for association above and beyond linkage, residual genetic and same-household effects, these three components were included in both the null and alternative models. We first tested for association between the SNPs and Aβ42 in the 10 LOAD families that we previously used to describe linkage to chromosome 10 (35). In order to control for the effect of linkage, residual genetic effects and same-household effects in this group, we estimated the effect of each SNP on the background of a model that included our 80 cM multipoint identity-by-descent matrix, as well as kinship and household matrices. We then tested each SNP for association with Aβ42 in a second, independent set of 12 LOAD families, and also analyzed the combined group of 22 families.

Linkage analysis conditional upon association

A conditional linkage analysis that accounts for association with a functional polymorphism (or one in strong LD with it) at the linkage locus will decrease or eliminate the linkage signal (58,59). For this reason, we performed linkage analysis conditional on the association with our most strongly associating SNP (4360TC). We performed this test by estimating the proportion of variance due to linkage on the background of a model that included this SNP as a covariate and comparing it to a linkage only model that excluded this covariate.

Analysis of linkage disequilibrium

We measured linkage disequilibrium (LD) between the SNPs and the haplotype groups within the 22 extended LOAD families by using the GOLD program (60). We used the expectation-maximization algorithm of Slatkin and Excoffier (61) to determine various pairwise statistical parameters for LD, such as Lewontin's standardized disequilibrium coefficient D (62). This algorithm determines LD from the founders and married-ins. LD was estimated between the VR22 4360 SNP and all other SNPs.

ACKNOWLEDGEMENTS

We are grateful to Frederick T. Pishotta for computer support and Thomas Dyer from Southwest Foundation for Biomedical Research for providing us the scripts for obtaining SOLAR-compatible mibd files from Simwalk. We acknowledge the Mayo Clinic ADRC employees for their help in the collection of samples. We are grateful to all of the individuals who participated in this study, without whom this work would not be possible. This study was supported by the NIH grants AG18023 (SGY), AG06656 (SGY), AG20903 (NET), a grant from American Federation for Aging Research Grant PD01062 (NET) and a Robert and Clarice Smith Fellowship (NET).

REFERENCES

1. Schellenberg, G.D., D'Souza, I. and Poorkaj, P. (2000) The genetics of Alzheimer's disease. *Curr. Psych. Rep.*, **2**, 158–164.
2. Bergem, A.L., Engedal, K. and Kringlen, E. (1997) The role of heredity in late-onset Alzheimer disease and vascular dementia. A twin study. *Arch. Gen. Psych.*, **54**, 264–270.
3. Gatz, M., Pedersen, N.L., Berg, S., Johansson, B., Johansson, K., Mortimer, J.A., Posner, S.F., Viitanen, M., Winblad, B. and Ahlbom, A. (1997) Heritability for Alzheimer's disease: the study of dementia in Swedish twins. *J. Gerontol. A. Biol. Sci. Med. Sci.*, **52**, M117–1125.
4. Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L. and Pericak-Vance, M.A. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921–923.
5. Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., Myers, R.H., Pericak-Vance, M.A., Risch, N. and Duijn, C.M. (1997) Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*, **278**, 1349–1356.
6. Warwick-Daw, E., Payami, H., Nemens, E.J., Nochlin, D., Bird, T.D., Schellenberg, G.D. and Wijsman, E.M. (2000) The number of trait loci in late-onset Alzheimer disease. *Am. J. Hum. Genet.*, **66**, 196–204.

7. Blangero, J., Williams, J.T. and Almasy, L. (2001) Variance component methods for detecting complex trait loci. *Adv. Genet.*, **42**, 151–181.
8. Ghosh, S. and Collins, F.S. (1996) The geneticist's approach to complex disease. *Annu. Rev. Med.*, **47**, 333–353.
9. Lernmark, A. and Ott, J. (1998) Sometimes it's hot, sometimes it's not. *Nat. Genet.*, **19**, 213–214.
10. Shoji, M., Golde, T.E., Cheung, T.T., Ghiso, J., Estus, S., Shaffer, L.M., Cai, X.-D., McKay, D., Tintner, R., Frangione, B. *et al*. (1992) Production of the Alzheimer amyloid β protein by normal proteolytic processing. *Science*, **258**, 126–129.
11. Seubert, P., Vigo-Pelfrey, C., Esch, F., Lee, M., Dovey, H., Davis, D., Sinha, S., Schlossmacher, M., Whaley, J., Swindlehurst, C. *et al*. (1992) Isolation and quantification of soluble Alzheimer's beta-peptide from biological fluids. *Nature*, **359**, 325–327.
12. Haass, C., Schlossmacher, M.G., Hung, A.Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B.L., Lieberburg, I., Koo, E.H., Schenk, D., Teplow, D.B. *et al*. (1992) Amyloid beta-peptide is produced by cultured cells during normal metabolism. *Nature*, **359**, 322–325.
13. Dovey, H.F., Suomensaari-Chrysler, S., Lieberburg, I., Sinha, S. and Keim, P.S. (1993) Cells with a familial Alzheimer's disease mutation produce authentic beta-peptide. *Neuroreport*, **4**, 1039–1042.
14. Vigo-Pelfrey, C., Lee, D., Keim, P., Lieberburg, I. and Schenk, D.B. (1993) Characterization of beta-amyloid peptide from human cerebrospinal fluid. *J. Neurochem.*, **61**, 1965–1968.
15. Suzuki, N., Cheung, T.T., Cai, X.D., Odaka, A., Otvos, L., Jr, Eckman, C., Golde, T.E. and Younkin, S.G. (1994) An increased percentage of long amyloid beta protein secreted by familial amyloid beta protein precursor (beta APP717) mutants. *Science*, **264**, 1336–1340.
16. McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D. and Stadlan, E.M. (1984) Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, **34**, 939–944.
17. Gravina, S.A., Ho, L., Eckman, C.B., Long, K.E., Otvos, L., Younkin, L.H., Suzuki, N. and Younkin, S.G. (1995) Amyloid beta protein (A beta) in Alzheimer's disease brain. Biochemical and immunocytochemical analysis with antibodies specific for forms ending at A beta40 or A beta 42(43). *J. Biol. Chem.*, **270**, 7013–7016.
18. Jarrett, J.T., Berger, E.P. and Lansbury P.T., Jr (1993) The carboxy terminus of the beta amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry*, **32**, 4693–4697.
19. Snyder, S.W., Ladror, U.S., Wade, W.S., Wang, G.T., Barrett, L.W., Matayoshi, E.D., Huffaker, H.J., Krafft, G.A. and Holzman, T.F. (1994) Amyloid-beta aggregation: selective inhibition of aggregation in mixtures of amyloid with different chain lengths. *Biophys. J.*, **67**, 1216–1228.
20. Bahmanyar, S., Higgins, G.A., Goldgaber, D., Lewis, D.A., Morrison, J.H., Wilson, M.C., Shankar, S.K. and Gajdusek, D.C. (1987) Localization of amyloid beta protein messenger RNA in brains from patients with Alzheimer's disease. *Science*, **237**, 77–80.
21. Tanzi, R.E., Gusella, J.F., Watkins, P.C., Bruns, G.A., St George-Hyslop, P., Van Keuren, M.L., Patterson, D., Pagan, S., Kurnit, D.M. and Neve, R.L. (1987) Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science*, **235**, 880–884.
22. Goate, A., Chartier-Harlin, M.C., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L. *et al*. (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, **349**, 704–706.
23. Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D.M., Oshima, J., Pettingell, W.H., Yu, C.E., Jondro, P.D., Schmidt, S.D., Wang, K. *et al*. (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science*, **269**, 973–977.
24. Rogaev, E.I., Sherrington, R., Rogaeva, E.A., Levesque, G., Ikeda, M., Liang, Y., Chi, H., Lin, C., Holman, K., Tsuda, T. *et al*. (1995) Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature*, **376**, 775–778.
25. Sherrington, R., Rogaev, E.I., Liang, Y., Rogaeva, E.A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K. *et al*. (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**, 754–760.
26. Borchelt, D.R., Thinakaran, G., Eckman, C.B., Lee, M.K., Davenport, F., Ratovitsky T., Prada, C.M., Kim, G., Seekins, S., Yager, D. *et al*. (1996) Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1–42/1–40 ratio *in vitro* and *in vivo*. *Neuron*, **17**, 1005–1013.
27. Cai, X.D., Golde, T.E. and Younkin, S.G. (1993) Release of excess amyloid beta protein from a mutant amyloid beta protein precursor. *Science*, **259**, 514–516.
28. Citron, M., Oltersdorf, T., Haass, C., McConlogue, L., Hung, A.Y., Seubert, P., Vigo-Pelfrey, C., Lieberburg, I. and Selkoe, D.J. (1992) Mutation of the beta-amyloid precursor protein in familial Alzheimer'sdisease increases beta-protein production. *Nature*, **360**, 672–674.
29. Duff, K., Eckman, C., Zehr, C., Yu, X., Prada, C.M., Perez-Tur, J., Hutton, M., Buee, L., Harigaya, Y., Yager, D. *et al*. (1996) Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin1. *Nature*, **383**, 710–713.
30. Hsiao, K., Chapman, P., Nilsen, S., Eckman, C., Harigaya, Y., Younkin, S., Yang, F. and Cole, G. (1996) Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science*, **274**, 99–102.
31. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T.D., Hardy, J., Hutton, M., Kukull, W. *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.
32. Younkin, S.G., Ertekin Taner, N., Xin, T., Baker, M., Younkin, L.H., Hutton, M. and Graff-Radford, N. (2000) Role of elevated amyloid β protein in patients with typical late onset Alzheimer's disease. *Neurobiol. Aging*, **21**, S136.
33. Mayeux, R., Tang, M.X., Jacobs, D.M., Manly, J., Bell, K., Merchant, C., Small, S.A., Stern, Y., Wisniewski, H.M. and Mehta, P.D. (1999) Plasma amyloid beta-peptide 1–42 and incipient Alzheimer's disease. *Ann. Neurol.*, **46**, 412–416.
34. Ertekin-Taner, N., Graff-Radford, N., Younkin, L., Eckman, C., Adamson, J., Schaid, D., Blangero, J., Hutton, M. and Younkin, S. (2001) Heritability of plasma Aβ levels in typical late onset Alzheimer's disease pedigrees. *Genet. Epidem.*, **21**, 19–30.
35. Ertekin-Taner, N., Graff-Radford, N., Younkin, L.H., Eckman, C., Baker, M., Adamson, J., Ronald, J., Blangero, J., Hutton, M. and Younkin, S.G. (2000) Linkage of plasma Aβ42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. *Science*, **290**, 2303–2304.
36. Myers, A., Holmans, P., Marshall, H., Kwon, J., Meyer, D., Ramic, D., Shears, S., Booth, J., DeVrieze, F.W., Crook, R. *et al*. (2000) Susceptibility locus for Alzheimer's disease on chromosome 10. *Science*, **290**, 2304–2305.
37. Bertram, L., Blacker, D., Mullin, K., Keeney, D., Jones, J., Basu, S., Yhu, S., McInnis, M.G., Go, R.C., Vekrellis, K. *et al*. (2000) Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. *Science*, **290**, 2302–2303.
38. Li, Y.J., Scott, W.K., Hedges, D.J., Zhang, F., Gaskell, P.C., Nance, M.A., Watts, R.L., Hubble, J.P., Koller, W.C., Pahwa, R. *et al*. (2002) Age at onset in two common neurodegenerative diseases is genetically controlled. *Am. J. Hum. Genet.*, **70**, 985–993.
39. Janssens, B., Goossens, S., Staes, K., Gilbert, B., van Hengel, J., Colpaert, C., Bruyneel, E., Mareel, M. and van Roy, F. (2001) alphaT-catenin: a novel tissue-specific beta-catenin-binding protein mediating strong cell-cell adhesion. *J. Cell. Sci.*, **114**, 3177–3188.
40. Zhang, Z., Hartmann, H., Do, V.M., Abramowski, D., Sturchler-Pierrat, C., Staufenbiel, M., Sommer, B., van de Wetering, M., Clevers, H., Saftig, P. *et al*. (1998) Destabilization of beta-catenin by mutations in presenilin-1 potentiates neuronal apoptosis. *Nature*, **395**, 698–702.
41. Ardlie, K.G., Kruglyak, L. and Seielstad, M. (2002) Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.*, **3**, 299–309.
42. Daly, M.J., Rioux, J.D., Schaffner, S.F., Hudson, T.J. and Lander, E.S. (2001) High-resolution haplotype structure in the human genome. *Nat. Genet.*, **29**, 229–232.
43. Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M. *et al*. (2002) The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–2229.

44. Patil, N., Berno, A.J., Hinds, D.A., Barrett, W.A., Doshi, J.M., Hacker, C.R., Kautzer, C.R., Lee, D.H., Marjoribanks, C., McDonough, D.P. *et al*. (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, **294**, 1719–1723.
45. Rioux, J.D., Daly, M.J., Silverberg, M.S., Lindblad, K., Steinhart, H., Cohen, Z., Delmonte, T., Kochner, K., Miller, K., Guschwan, S. *et al*. (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat. Genet.*, **29**, 223–228.
46. Sobel, E. and Lange, K. (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am. J. Hum. Genet.*, **58**, 1323–1337.
47. Sobel, E., Sengul, H. and Weeks, D.E. (2001) Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. *Hum. Hered.*, **52**, 121–131.
48. Haines, J.L., St George-Hyslop, P.H., Rimmler, J.B., Yamaoka, L.H., Kazantsev, A., Tanzi, R.E., Gusella, J.F., Roses, A.D. and Pericak-Vance, M.A. (1993) Inheritance of multiple loci in familial Alzheimer disease. In Corain, B., Iqbal, K., Nicolini, M., Winblad, B., Wisniewski, H. and Zatta, P. (eds), *Alzheimer's Disease: Advances in Clinical and Basic Research*. John Wiley and Sons, New York, pp. 221–226.
49. Scott, W.K., Grubber, J.M., Conneally, P.M., Small, G.W., Hulette, C.M., Rosenberg, C.K., Saunders, A.M., Roses, A.D., Haines, J.L. and Pericak-Vance, M.A. (2000) Fine mapping of the chromosome 12 late-onset Alzheimer disease locus: potential genetic and phenotypic heterogeneity. *Am. J. Hum. Genet.*, **66**, 922–932.
50. Dyke, B. and Mamelka, P. (1996) PEDSYS: a pedigree data management system, 2.0 edn. Population Genetics Laboratory Department of Genetics Southwest Foundation for Biomedical Research, San Antonio, TX.
51. Broman, K.W., Murray, J.C., Sheffield, V.C., White, R.L. and Weber, J.L. (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.*, **63**, 861–869.
52. Lander, E. and Green, P. (1987) Construction of multilocus genetic linkage maps in humans *Proc. Nat. Acad. Sci. USA*, **84**, 2363–2367.
53. Almasy, L. and Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.*, **62**, 1198–1211.
54. Blangero, J. and Almasy, L. (1996) SOLAR: sequential oligogenic linkage analysis routines. Population Genetics Laboratory, Southwest Foundation for Biomedical Research, San Antonio, TX 78228.
55. Bethony, J., Williams, J.T., Kloos, H., Blangero, J., Alves-Fraga, L., Buck, G., Michalek, A., Williams-Blangero, S., Loverde, P.T., Correa-Oliveira, R. *et al*. (2001) Exposure to Schistosoma mansoni infection in a rural area in Brazil. II: household risk factors. *Trop. Med. Int. Health*, **6**, 136–145.
56. Hopper, J.L. and Mathews, J.D. (1982) Extensions to multivariate normal models for pedigree analysis. *Ann. Hum. Genet.*, **46**, 373–383.
57. Boerwinkle, E., Chakraborty, R. and Sing, C.F. (1986) The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann. Hum. Genet.*, **50**, 181–194.
58. Soria, J.M., Almasy, L., Souto, J.C., Tirado, I., Borell, M., Mateo, J., Slifer, S., Stone, W., Blangero, J. and Fontcuberta, J. (2000) Linkage analysis demonstrates that the prothrombin G20210A mutation jointly influences plasma prothrombin levels and risk of thrombosis. *Blood*, **95**, 2780–2785.
59. Soria, J.M., Almasy, L., Souto, J.C., Bacq, D., Buil, A., Faure, A., Martinez-Marchan, E., Mateo, J., Borrell, M., Stone, W. *et al*. (2002) A quantitative-trait locus in the human factor XII gene influences both plasma factor XII levels and susceptibility to thrombotic disease. *Am. J. Hum. Genet.*, **70**, 567–574.
60. Abecasis, G.R. and Cookson, W.O. (2000) GOLD—graphical overview of linkage disequilibrium. *Bioinformatics*, **16**, 182–183.
61. Excoffier, L. and Slatkin, M. (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.*, **12**, 921–927.
62. Lewontin, R.C. and Kojima, K. (1960) The evolutionary dynamics of complex polymorphisms. *Evolution*, **14**, 450–472.