

## Report

# Linkage and Association Studies Identify a Novel Locus for Alzheimer Disease at 7q36 in a Dutch Population-Based Sample

Rosa Rademakers,<sup>1</sup> Marc Cruts,<sup>1</sup> Kristel Sleegers,<sup>1,2</sup> Bart Dermaut,<sup>1</sup> Jessie Theuns,<sup>1</sup> Yurii Aulchenko,<sup>2</sup> Stefan Weckx,<sup>1</sup> Tim De Pooter,<sup>1</sup> Marleen Van den Broeck,<sup>1</sup> Ellen Corsmit,<sup>1</sup> Peter De Rijk,<sup>1</sup> Jurgen Del-Favero,<sup>1</sup> John van Swieten,<sup>3</sup> Cornelia M. van Duijn,<sup>1,2</sup> and Christine Van Broeckhoven<sup>1</sup>

<sup>1</sup>Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology, University of Antwerp, Antwerp; and <sup>2</sup>Department of Epidemiology and Biostatistics, Erasmus Medical Center, and <sup>3</sup>Department of Neurology, University Hospital Dijkzigt, Rotterdam

We obtained conclusive linkage of Alzheimer disease (AD) with a candidate region of 19.7 cM at 7q36 in an extended multiplex family, family 1270, ascertained in a population-based study of early-onset AD in the northern Netherlands. Single-nucleotide polymorphism and haplotype association analyses of a Dutch patient-control sample further supported the linkage at 7q36. In addition, we identified a shared haplotype at 7q36 between family 1270 and three of six multiplex AD-affected families from the same geographical region, which is indicative of a founder effect and defines a priority region of 9.3 cM. Mutation analysis of coding exons of 29 candidate genes identified one linked synonymous mutation, g.38030G→C in exon 10, that affected codon 626 of the PAX transactivation domain interacting protein gene (*PAXIP1*). It remains to be determined whether *PAXIP1* has a functional role in the expression of AD in family 1270 or whether another mutation at this locus explains the observed linkage and sharing. Together, our linkage data from the informative family 1270 and the association data in the population-based early-onset AD patient-control sample strongly support the identification of a novel AD locus at 7q36 and re-emphasize the genetic heterogeneity of AD.

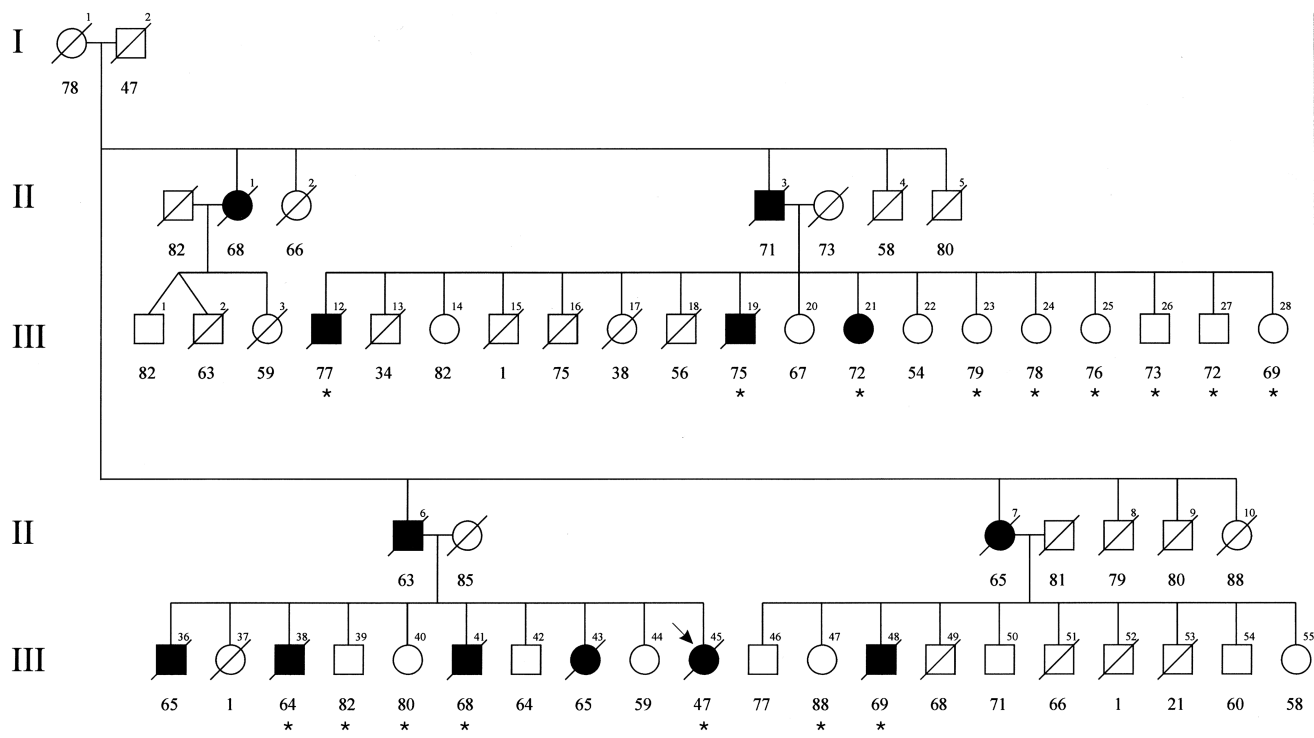
Alzheimer disease (AD [MIM 104300]) is a common neurodegenerative disorder of the CNS that is characterized by slowly progressing impairment of memory and intellectual functioning that leads to dementia and, ultimately, death. Clinical symptoms of AD most frequently arise after age 65 years (late-onset AD); however, in <1% of the patients, AD develops before age 65 years (early-onset AD) (Ott et al. 1998). A definite diagnosis of AD can be established only when pathological examination of autopsied brain shows parenchymal senile plaques composed largely of amyloid  $\beta$  and of intraneuronal neurofibrillary tangles of hyperphosphorylated tau. Thus far, positional cloning in multigeneration families with AD

has identified three genes, coding for the amyloid precursor protein (*APP* [MIM 104760]) (21q21) (Goate et al. 1991), presenilin 1 (*PSEN1* [MIM 104311]) (14q23) (Sherrington et al. 1995), and presenilin 2 (*PSEN2* [MIM 600759]) (1q42.3) (Levy-Lahad et al. 1995; Rogae et al. 1995). Currently, 172 different mutations in *APP*, *PSEN1*, and *PSEN2* have been documented in 363 families (Cruts and Van Broeckhoven 1998; Alzheimer Disease & Frontotemporal Dementia Mutation Database). In addition, the  $\epsilon 4$  allele of the apolipoprotein E gene (*APOE* [MIM 107741]) (19q13) was identified as a risk factor for both early- and late-onset AD (Saunders et al. 1993; Strittmatter et al. 1993; van Duijn et al. 1994a). For ~7%–20% of patients, AD is attributable to *APOE*  $\epsilon 4$  (Slooter et al. 1998; Warwick et al. 2000). Since the genetic cause remains unknown in a significant percentage of patients with familial AD, including large families with well-documented autosomal dominant AD (Pericak-Vance et al. 2000; Ashley-Koch et al. 2005; Jimenez-Escrig et al. 2005), it is clear that additional AD genes remain to be identified.

Received June 14, 2005; accepted for publication July 25, 2005; electronically published August 30, 2005.

Address for correspondence and reprints: Dr. Christine Van Broeckhoven, Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology (VIB8), University of Antwerp, Campus Drie Eiken, Building V, Room 0.10, Universiteitsplein 1, B-2610 Antwerpen, Belgium. E-mail: christine.vanbroeckhoven@ua.ac.be

© 2005 by The American Society of Human Genetics. All rights reserved.  
0002-9297/2005/7704-0013\$15.00



**Figure 1** Pedigree of family 1270. A blackened symbol represents a patient, and an unblacked symbol represents an unaffected individual or an at-risk individual with unknown phenotype. The Roman numbers to the left of the pedigree denote generations; Arabic numbers above the symbols denote individuals. All numbering is in accordance with that of van Duijn et al. (1994b). The Arabic numbers below the symbols denote age at onset for patients and either age at last examination or age at death for unaffected individuals and at-risk individuals with unknown phenotype. An asterisk (\*) indicates an individual who was included in the linkage analysis. The arrow identifies the proband with early-onset AD (age 47 years).

The present study was designed to identify additional loci for AD by positional cloning in multigenerational families. The source for our project was a sample of 110 patients with early-onset AD from a population-based epidemiology study aiming to ascertain all patients with early-onset AD in Rotterdam and the four northern provinces of The Netherlands (Hofman et al. 1989). From this sample, we had previously ascertained 10 multigenerational families for linkage studies (van Duijn et al. 1994b). In two families, a mutation in *PSEN1* cosegregated with the disease, whereas mutations in *APP* and *PSEN2* were absent (van Duijn et al. 1991; Cruts et al. 1998). Two other families, 1083 and 1270, were selected for further linkage studies, since simulation linkage data showed that each single family was sufficiently informative to generate significant LOD scores independently. In family 1083, genomewide linkage analysis identified significant linkage at 17q21 (multipoint LOD score 5.41 at *D17S951*;  $\theta = 0.0$ ) (Rademakers et al. 2002). Here, we report on the clinical follow-up and the genetic analy-

**Table 1**

**Clinical Characteristics of Patients of Family 1270**

PATIENT	SEX	AGE (IN YEARS) AT		APOE GENOTYPE <sup>a</sup>	DIAGNOSIS
		Disease Onset	Death		
II-1	F	68	72	...	Dementia <sup>b</sup>
II-3	M	71	77	...	MCI <sup>b</sup>
II-6	M	63	74	...	Dementia <sup>b</sup>
II-7	F	65	70	...	Dementia <sup>b</sup>
III-12	M	77	>82	$\epsilon 3\epsilon 3$	Probable AD
III-19 <sup>c</sup>	M	75	80	$\epsilon 3\epsilon 3$	MCI <sup>b</sup>
III-21 <sup>c</sup>	F	72	...	$\epsilon 4\epsilon 4$	Possible AD
III-36	M	65	70	...	Possible AD
III-38	M	64	72	$\epsilon 3\epsilon 4$	Probable AD
III-41 <sup>c</sup>	M	68	72	$\epsilon 3\epsilon 3$	Probable AD
III-43 <sup>c</sup>	F	65	71	...	Dementia <sup>b</sup>
III-45	F	47	60	$\epsilon 4\epsilon 4$	Probable AD
III-48	M	69	>75	$\epsilon 3\epsilon 4$	Probable AD

<sup>a</sup> APOE genotypes of available patients in generation III (fig. 1) were determined elsewhere (van Duijn et al. 1994a).

<sup>b</sup> Diagnosis obtained through family informant.

<sup>c</sup> Newly diagnosed patients.

---

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics*.

---

**Figure 2** Brain CT scan of patient III-21 at age 82 years. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics*.

ses of the three-generation family 1270. This family was ascertained through its 47-year-old proband with early-onset AD; a total of nine patients and seven at-risk individuals had cooperated, and blood samples had been collected for genetic linkage studies (van Duijn et al. 1994b). A clinical follow-up of family 1270 was performed by neurological examination of incident patients, interviews of first-degree relatives, and review of medical records. Diagnoses of possible or probable AD were made according to National Institute of Neurological and Communicative Diseases and Stroke–Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria (McKhann et al. 1984), and that of mild cognitive impairment (MCI) according to Honig and Mayeux (2001). For five patients who were deceased or could not be examined, diagnosis was complemented by information obtained through a family informant. Since the first report (van Duijn et al. 1994b), four family members (III-19, III-21, III-41, and III-43) became symptomatic. The updated pedigree of family 1270 is shown in figure 1; clinical data is summarized in table 1. The overall clinical picture of family 1270 was compatible with AD, with a mean ( $\pm$  SD) age at onset of  $66.8 \pm 7.4$  years ( $N = 13$ ; age range 47–77 years) and a mean duration of disease of  $6.7 \pm 2.7$  years ( $N = 10$ ; range of duration 4–13 years). In proband III-45 and most other patients, the disease initially presented with memory impairment, except for patient III-38, in whom a change of character was the initial complaint, later followed by memory loss. In all patients, the disease progressed into other areas of cognition, such as praxis and speech. Neuroimaging was available for two patients, III-48 at age 74 years and III-21 at age 82 years (table 1), which showed cortical atrophy in both patients. For the only living patient, III-21, who received the diagnosis of possible AD, a recent CT scan (at age 82 years) showed that cortical and subcortical atrophy was most notable in the temporal and frontal regions (fig. 2). No brain pathology of deceased patients was available for family 1270. Since previous studies have already described mutations in the genes encoding the microtubule-associated protein tau (*MAPT* [MIM 157140]) and the prion protein (*PRNP* [MIM 176640]) in patients who received a clinical diagnosis of AD but a pathological diagnosis, at autopsy, of frontotemporal dementia (MIM 600274) or Creutzfeldt-Jacob disease (MIM 123400) (Dermaut et al. 2000;

Rademakers et al. 2003), the identification of the pathology underlying the dementia in family 1270 will be important for the definite classification of the inherited dementia in this family.

We performed a genomewide linkage analysis of family 1270, using 400 microsatellite markers of the ABI Prism Linkage Mapping Set MD-10 Version 2 (Applied Biosystems). Microsatellite markers were analyzed on an automated ABI 377 or ABI 3700 DNA-Analyzer (Applied Biosystems). For allele identification and scoring, GENE-SCAN and GENOTYPER were used (Applied Biosystems). We calculated linkage analysis with the assumption of an autosomal dominant mode of inheritance and a disease frequency of 0.1%. A cumulative risk curve was calculated with use of mean age at onset in the family and the assumption of a maximal disease penetrance of 85% at age  $\geq 85$  years, and five disease penetrance classes were defined. Age-dependent phenocopy rates were defined on the basis of the literature (Ott et al. 1995). All patients with AD and MCI were considered affected in the linkage analysis. The phenotypes of individuals I-1 and I-2 were considered unknown. Two-point and multipoint LOD scores were calculated using MLINK and LINKMAP options of the computer package Linkage, version 5.1 (Lathrop et al. 1985), with use of equal recombination rates for males and females. Using this linkage model, we obtained a genomewide conclusive LOD score of 3.39, at  $\theta = 0.00$ , only with marker *D7S798* at 7q36 (table 2). The flanking markers, *D7S636* and *D7S2465*, also showed positive LOD scores of 2.69 and 1.70, respectively, at  $\theta = 0.00$ . A summary of all two-point LOD scores obtained in the genomewide scan is provided in figure 3. To fine map the candidate region at 7q36, we analyzed an additional 13 microsatellite markers in a 26.9-cM interval flanked by markers *D7S661* and *D7S2423*, including 9 markers from the Marshfield (Center for Medical Genetics) sex-averaged linkage map, 1 marker from the genetic map of the Cooperative Human Linkage Center, and 3 novel markers—*pRR8*, *pRR10*, and *pRR15*—identified using the Tandem Repeat Finder program (table 2) (Benson 1999). The highest two-point LOD score, 3.41, was obtained with *D7S1807* in the absence of recombinants. The maximum multipoint LOD score was 3.47 (data not shown). By use of the genotype data of 17 markers, haplotypes at 7q36 were reconstructed, with considera-

---

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics*.

---

**Figure 3** Genomewide linkage data of the 10-cM density scan of family 1270. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics*.

**Table 2**  
Two-Point LOD Scores for 7q36 Markers in Family 1270

MARKER <sup>a</sup>	DISTANCE <sup>b</sup> (cM)	LOD AT $\theta =$						
		.00	.01	.05	.10	.20	.30	.40
<i>D7S661</i>	.0	-1.31	-.93	-.40	-.15	.01	.02	.01
<i>D7S668</i>	4.4	-.68	.34	.97	1.15	1.09	.78	.32
<i>D7S505</i>	6.1	.30	1.28	1.76	1.80	1.52	1.03	.43
<i>D7S636</i>	7.2	2.69	2.65	2.46	2.22	1.68	1.08	.43
<i>pRR8</i> <sup>c</sup>	...	-1.08	-.09	.53	.73	.70	.45	.15
<i>D7S642</i>	7.2	1.26	1.29	1.32	1.26	1.01	.64	.24
<i>D7S2439</i>	8.6	2.88	2.83	2.62	2.34	1.74	1.09	.42
<i>D7S483</i>	10.1	3.32	3.27	3.04	2.74	2.08	1.34	.55
<i>D7S798</i>	13.9	3.39	3.33	3.11	2.81	2.14	1.38	.57
<i>D7S1807</i> <sup>d</sup>	...	3.41	3.36	3.14	2.84	2.16	1.40	.58
<i>D7S2546</i>	17.9	2.51	2.47	2.26	1.99	1.41	.82	.28
<i>D7S1823</i>	18.6	2.72	2.69	2.52	2.29	1.74	1.12	.44
<i>pRR10</i> <sup>c</sup>	...	1.95	1.90	1.73	1.52	1.07	.61	.20
<i>D7S2465</i>	25.1	1.70	1.67	1.56	1.41	1.06	.65	.23
<i>pRR15</i> <sup>c</sup>	...	3.11	3.06	2.85	2.57	1.95	1.26	.52
<i>D7S559</i>	26.9	1.41	1.61	1.84	1.82	1.49	.98	.38
<i>D7S2423</i>	26.9	-1.04	-.82	-.45	-.28	-.14	-.06	-.02

<sup>a</sup> Markers from the 10-cM genomewide scan are in bold italics.

<sup>b</sup> Genetic distances were deduced from the Marshfield (Center for Medical Genetics) sex-averaged linkage map.

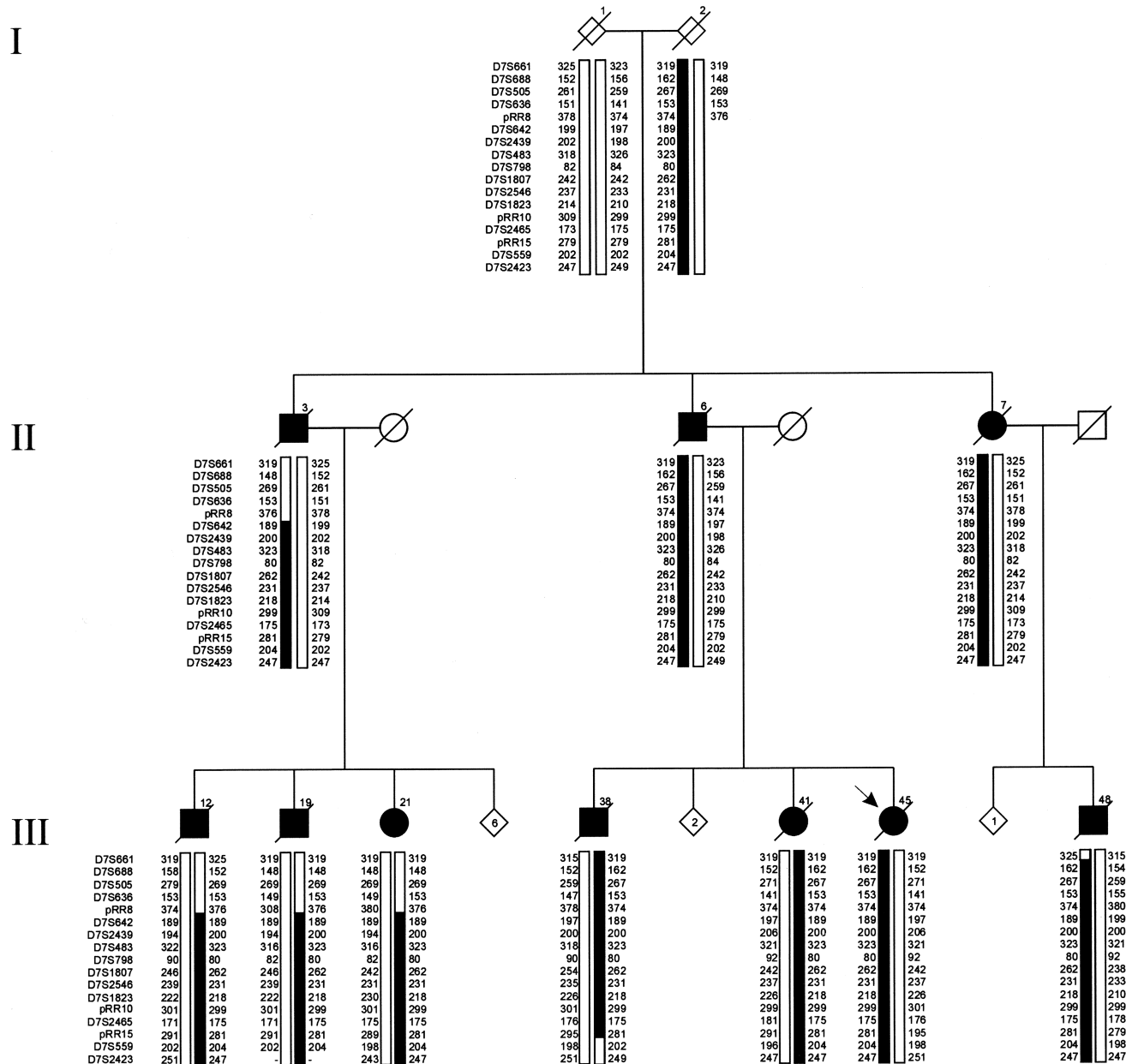
<sup>c</sup> *pRR8* is located in AC010973.6 starting at nt 93259, *pRR10* in AC078834.5 at nt 132119, and *pRR15* in AC005534.2 at nt 148849. The number of different alleles observed within family 1270 was three for *pRR10*, four for *pRR8*, and five for *pRR15*.

<sup>d</sup> Selected from the genetic map of the Cooperative Human Linkage Center.

tion of minimal intermarker recombination (fig. 4). In the linkage pedigree, a disease haplotype was present in all patients, and obligate recombinants defined a candidate region of 19.7 cM between *pRR8* (individuals III-12, III-19, and III-21) and *D7S559* (individual III-38). To estimate the contribution of the 7q36 locus in the Dutch population, we genotyped eight microsatellite markers in 102 patients with AD from the population-based early-onset AD sample, recruited from the same geographical region as family 1270 (Hofman et al. 1989), and 187 age- and sex-matched control individuals, drawn from a large Dutch population-based cohort study of chronic diseases of the elderly (Hofman et al. 1991). Markers *D7S1807* and *D7S483* were excluded from further analysis: *D7S1807* because the linked allele was not observed in any of the patients with early-onset AD, which indicated that the linked allele is extremely rare (frequency <1%) in AD-affected individuals, and *D7S483* because >20 different alleles were observed, which is indicative of genetic instability. When we calculated single-marker associations with the remaining six markers in the Dutch early-onset AD patient-control sample with use of GenePop, we identified significant allelic association with *D7S798* ( $P = .03$ ) (table 3). This association could be explained in part by an increase of the rare 80-bp allele, also linked to AD in family 1270, from 1.9% in control individuals to 4.9% in patients with early-onset AD ( $P_{\text{specific}} = .04$ ). Examination of individual genotype data

of *D7S798* revealed that 4 of the 10 patients with early-onset AD who carried the 80-bp allele at *D7S798* were probands of families with autosomal dominant AD. We estimated that, in the Dutch patient-control sample, the risk of AD for subjects carrying the 80-bp allele was 2.7 times higher (OR 2.69; 95% CI 1.01–7.17) than for noncarriers.

In the population with early-onset AD, six patients were probands of families with multiplex AD, which is consistent with autosomal dominant inheritance (van Duijn et al. 1994b); they were not included in the genomewide scan, because none had individually sufficient genetic information to detect conclusive linkage. In all six probands, the onset of AD was at age  $\leq 65$  years (van Duijn et al. 1994b). In three families (1034, 1104, and 1125), the mean age at onset of the relatives was <65 years, whereas the three other families (1027, 1094, and 1242) showed a mean onset of dementia at age >65 years (table 4). The genotypes of the families with multiplex AD were compared for allele sharing with those of family 1270 (table 4). In three families (1034, 1125, and 1242), allele sharing was observed of the rare 80-bp allele at *D7S798*, as well as the linked alleles at two neighboring markers (*D7S2439* and *D7S2546*). The frequency of the shared haplotype that was calculated on the basis of the frequencies of the linked alleles in Dutch control individuals was 0.6%. Also, all six markers were genotyped in additional patients as well as relatives of families 1034 (2



**Figure 4** Disguised linkage pedigree of family 1270 (fig. 1). Haplotypes are based on 17 informative markers at 7q36. Haplotypes for individuals from the first and second generations were inferred from genotype data of siblings and offspring. The risk haplotype was arbitrarily set for individual I-2. For confidentiality reasons, haplotypes are shown only for patients; the number of at-risk individuals included in the genotyping are indicated by numbers within diamonds.

patients), 1125 (2 patients), and 1242 (1 patient), and segregation analysis of the shared alleles was performed to define haplotypes (fig. 5). In family 1034, all three patients with AD (III-4, III-6, and III-8) inherited the 201-80-232 haplotype. Segregation analysis of family 1242 revealed that both patients (III-14 and III-20) carried the 80-232 haplotype. The lack of sufficient segregation information precluded haplotype determination in family 1125. However, allele sharing in the two ad-

ditional patients from that family identified the linked 201-80-232 alleles in one patient (III-1), whereas the other patient (III-2) shared the 201-80 alleles. Together, these studies delineated a priority region—between markers *D7S2439* and *D7S2546*, with a maximum genetic distance of 9.3 cM—that is shared by families 1270, 1034, 1242, and 1125. Two-marker sliding-window haplotype analysis in the early-onset AD patient-control sample, with use of the expectation maximization algorithm pro-

**Table 3****Allelic Frequencies and Association Analysis of Dutch Patient-Control Sample of Early-Onset AD, at 7q36**

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

vided by the Arlequin package, further revealed that the 201-80 haplotype of *D7S2439-D7S798* significantly increased from 0.5% in control individuals to 3.4% in patients with AD ( $P_{\text{specific}} = .007$ ; OR 6.52; 95% CI 1.34–31.71). The haplotype-sharing data of the four Dutch families with early-onset AD is suggestive of a common founder for AD that is linked to the 7q36 locus in the Dutch population. We documented elsewhere the founder effects within the Dutch population with early-onset AD, since all patients carrying the *PSEN1* Ala79Val mutation or the Glu318Gly polymorphism shared an extended haplotype (Cruts et al. 1998; Dermaut et al. 1999). Comparison of the onset ages in the four families with multiplex AD also identified relatively similar mean ages at onset (58.0, 61.3, 66.8, and 72.0 years), with wide intrafamilial ranges (14–30 years) that showed patients with both early- and late-onset AD in each family, which further supported a common phenotype in these families. Previous studies of families with AD have shown that the age at onset can be influenced by the *APOE* genotype, with *APOE*  $\epsilon 4$  carriers having an age at onset earlier than that of noncarriers of *APOE*  $\epsilon 4$  (Corder et al. 1993; van Duijn et al. 1994a). However, since ages at onset were 47 and 72 years for two patients homozygous for *APOE*  $\epsilon 4$  in family 1270 (table 1), it is unlikely that the broad onset range in the patients with AD in this family is explained by *APOE* genotypes.

Also, in the three families who share a disease haplotype at 7q36, all patients with available DNA were homozygous for *APOE*  $\epsilon 4$ , whereas ages at onset varied from 51 to 80 years (van Duijn et al. 1994a).

We assembled the genomic sequence of the candidate region in family 1270, between *pRR8* and *D7S559*, using publicly available finished and unfinished sequences of large-insert clones. This resulted in two contiguous sequences, of 4.03 Mb and 1.33 Mb, separated by a sequence gap of ~80 kb located between *D7S1823* and *pRR10* (table 2) (Hillier et al. 2003). In family 1270, the candidate region segregating the disease gene at 7q36 was therefore estimated at 5.44 Mb. The priority region between *D7S2439* and *D7S2546* that was shared by the four dominant families with AD corresponded to an estimated physical size of 2.94 Mb. In the course of this study, the complete sequence assembly of 7q36 became available through public genome annotation databases (National Center for Biotechnology Information [NCBI] Map Viewer; University of California–Santa Cruz [UCSC] Genome Bioinformatics). Although two gaps remained in the public genome assembly (NCBI build 34), no major inconsistencies in sequence order were observed with our independently assembled sequence contigs. Within the 5.44-Mb candidate region, we determined the location and genomic organizations of 29 genes by annotating the candidate sequence contigs on the basis of homology with genes and proteins of human and other species, ESTs, and predicted genes and exons (table 5). These included 20 validated, reviewed, and provisional RefSeq genes located in the 150.2–155.8 Mb region of NCBI build 34 of the UCSC Human Genome Browser (UCSC Genome Bioinformatics). We performed mutation analysis of 28 genes

**Table 4****Allele-Sharing Analysis of Microsatellite Markers at 7q36**

MARKER	LINKED ALLELE IN FAMILY 1270 (bp)	FREQUENCY OF LINKED ALLELE IN CONTROL INDIVIDUALS (%)	GENOTYPE (ALLELES IN BP) OF AUTOSOMAL DOMINANT FAMILY					
			1027 <sup>a</sup>	1034 <sup>b</sup>	1094 <sup>c</sup>	1104 <sup>d</sup>	1125 <sup>e</sup>	1242 <sup>f</sup>
<i>D7S642</i>	<b>189</b>	45.4	<b>189–191</b>	<b>189–191</b>	<b>189–191</b>	<b>189–197</b>	195–195	191–191
<i>D7S2439</i>	<b>201</b>	29.4	<b>201–201</b>	<b>201–201</b>	195–195	199–209	199–201	199–201
<i>D7S798</i>	<b>80</b>	1.9	82–84	<b>80–88</b>	84–84	90–92	<b>80–82</b>	<b>80–92</b>
<i>D7S2546</i>	<b>232</b>	18.6	238–238	<b>232–236</b>	195–195	199–209	<b>232–238</b>	<b>232–232</b>
<i>D7S1823</i>	<b>217</b>	12.3	<b>217–225</b>	221–225	84–84	90–92	221–229	209–237
<i>D7S2465</i>	<b>175</b>	23.5	180–184	173–176	169–173	<b>175–183</b>	173–177	173–177

NOTE.—The linked alleles of family 1270 are in bold italics. Ages at onset are taken from the publication by van Duijn et al. (1994b).

<sup>a</sup> Number of patients with dementia = 3; mean ( $\pm$  SD) age at onset = 72.3  $\pm$  17.9 years; age range 57–92 years.

<sup>b</sup> Number of patients with dementia = 4; mean age at onset = 61.3  $\pm$  6.1; age range 56–70 years.

<sup>c</sup> Number of patients with dementia = 6; mean age at onset = 70.5  $\pm$  8.0; age range 61–80 years.

<sup>d</sup> Number of patients with dementia = 9; mean age at onset = 51.9  $\pm$  9.7; age range 40–67 years.

<sup>e</sup> Number of patients with dementia = 5; mean age at onset = 58  $\pm$  8.6; age range 49–70 years.

<sup>f</sup> In family 1242, we analyzed allele sharing in patient III-14, since DNA of the proband (III-20) was no longer available (van Duijn et al. 1994b). Number of patients with dementia = 4; mean age at onset = 72  $\pm$  7.7; age range 63–80 years.

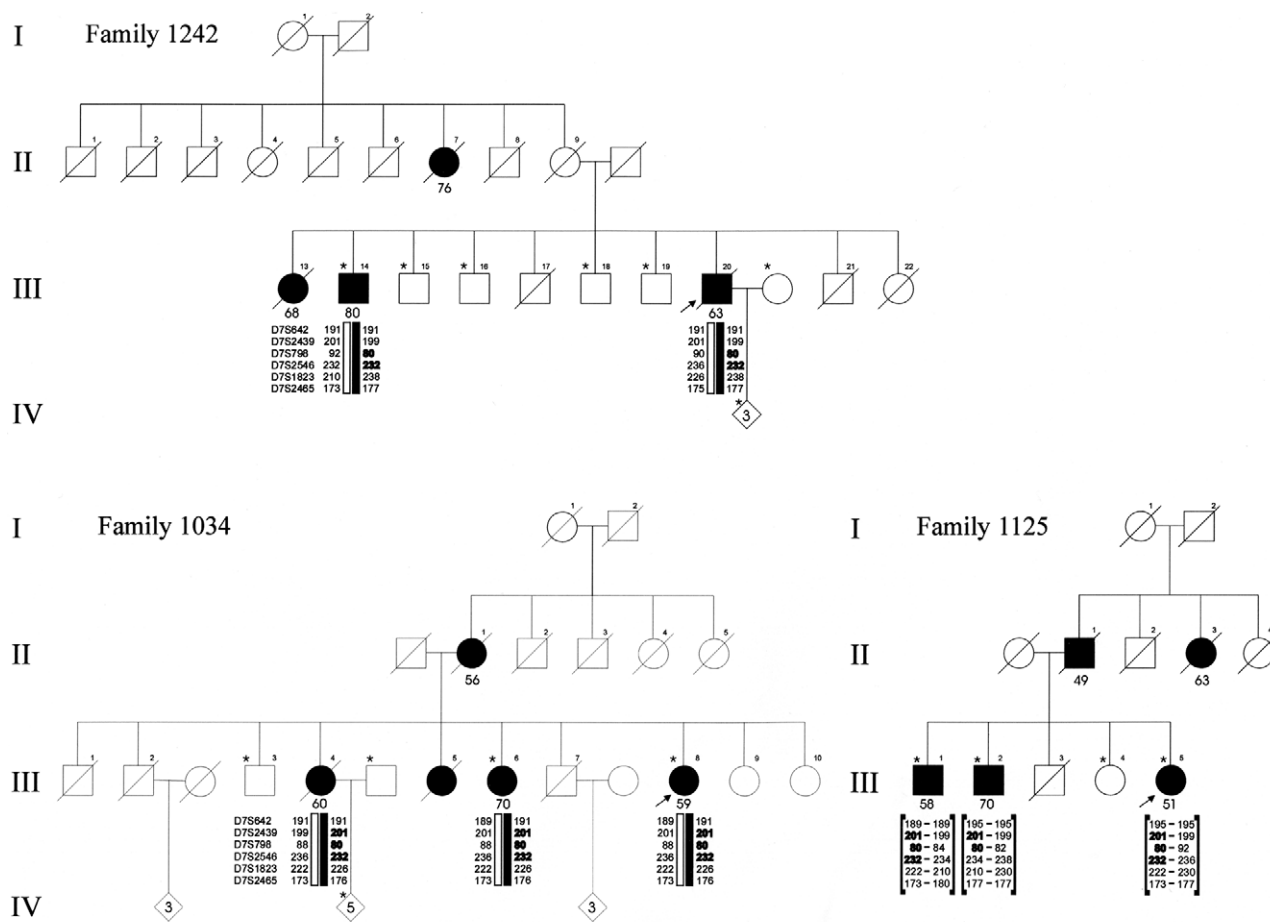
from the candidate region by direct genomic DNA (gDNA) sequencing of PCR amplicons of all coding exons in three patients (III-12, III-41, and III-48) and one control individual of family 1270. On the basis of gDNA and cDNA sequence alignment, intronic PCR primers flanking coding exons were developed using SNPbox (Weckx et al. 2005). For the gene encoding myeloid and lymphoid leukemia 3 (*MLL3*), mutation analysis was performed by cDNA sequencing in two patients (III-38 and III-41) and one control individual. DNA sequences were determined by cycle sequencing with the BigDye Cycle Sequencing kit (Applied Biosystems) and were analyzed on an ABI 3700 and ABI 3730 DNA analyzer (Applied Biosystems). In total, we identified 167 sequence variations: 24 missense

**Table 5**

**Overview of Analyzed Candidate Genes at 7q36**

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

and 21 silent mutations in 18 genes, and 122 intronic variations (data not shown). Seven single-nucleotide changes in five different genes segregated in family 1270 on the disease haplotype (table 6). However, analysis of gDNA of 40 Dutch control individuals indicated that six of the seven sequence variants were common polymorphisms, with frequencies of the linked allele varying from 10.7% to 54.0% (table 6). Only g.38030G→C in exon



**Figure 5** Linkage pedigrees of Dutch families with early-onset AD who share haplotypes at 7q36. Blackened symbols represent patients with AD; unblackened symbols represent unaffected individuals or at-risk individuals with unknown phenotype. Roman numerals to the left of the pedigree denote generations; Arabic numbers above the symbols denote individuals. All numbers are in accordance with those used in the complete pedigrees published by van Duijn et al. (1994b). For patients, the age (in years) at onset is shown below the symbol. An asterisk (\*) indicates that DNA was available for genetic analysis. The arrow identifies the proband. To protect privacy of at-risk individuals, the shared haplotypes are shown only for patients. Haplotypes for individual III-20 in family 1242 and individual III-4 in family 1034 were reconstructed from genotype data obtained from their siblings and offspring. For family 1125, only genotype data are given, since segregation data were not informative for deduction of haplotypes.

**Table 6**  
Sequence Variations Segregating on the Disease Haplotype in Family 1270

Gene <sup>a</sup> , GenBank Clone Number, and Variant	Predicted Effect on Amino Acid	Linked Allele	Frequency <sup>b</sup> in Control Individuals (%)	Restriction Enzyme <sup>c</sup>
<i>ABCF2</i> :				
AC021097.5:				
g.16531C→T	Leu→Leu	T	26.2	...
g.19567G→C	...	C	33.3	<i>HinfI</i>
<i>GALNTL5</i> :				
AC074257.5:				
g.176022A→G	Glu→Gly	G	19.8	...
<i>DPP6</i> :				
AC006019.2:				
g.57910C→T	Gly→Gly	T	18.1	...
<i>PAXIP1</i> :				
AC093726.3:				
g.38030G→C	Ala→Ala	C	.0	<i>BsbI236I</i>
<i>EN2</i> :				
AC008060.5:				
g.89332G→A	Pro→Pro	A	10.7	<i>SmaI</i>
g.89582T→C	Phe→Leu	C	54.0	<i>AluI</i>

<sup>a</sup> *ABCF2*: ATP-binding cassette subfamily F, member 2; *GALNT11*: N-acetylgalactosaminyltransferase 11; *DPP6*: dipeptidylpeptidase 6; *PAXIP1*: PAX transcription activation domain interacting protein; *EN2*: engrailed 2.

<sup>b</sup> Allele frequencies for all variations were estimated in 80 control chromosomes. For g.38030G→C in *PAXIP1*, an additional 560 control chromosomes were tested.

<sup>c</sup> Control individuals were genotyped by a PCR-RFLP assay and electrophoresis of the PCR fragments on a 2% agarose gel or, if the mutation did not affect a restriction enzyme site, by direct sequencing.

10 of *PAXIP1*, which predicted a silent Ala626 mutation, was absent in the 40 control individuals and in an additional 280 Dutch control individuals, which resulted in a C-allele frequency of <0.16% (<1/640 chromosomes). When we examined the expression profile of *PAXIP1* on a multiple-human-tissue northern blot, we identified a 4.5-kb transcript of *PAXIP1* in all tissues analyzed, except in liver, in which an alternative *PAXIP1* transcript of 1.5 kb was expressed (fig. 6). Brain expression analysis of *PAXIP1* by RT-PCR identified *PAXIP1* transcripts in seven different brain regions, including frontal and temporal regions affected in AD-affected brains (fig. 6). Although g.38030G→C is located 10 bp upstream of the intron 10 splice-donor site, a preliminary RT-PCR analysis of lymphoblast cDNA of two patients (III-38 and III-41) and one control individual of family 1270 did not suggest aberrant *PAXIP1* splicing. *PAXIP1* g.38030G→C was also not observed in the sample of patients with early-onset AD, including the three probands of the families with autosomal dominant disease who share the priority region between markers *D7S2439* and *D7S2546*. Since *PAXIP1* is outside the shared region at 0.44 Mb downstream of the associated *D7S2546*, the absence of

g.38030G→C in the three probands does not contradict a founder effect. Sequencing of PCR amplicons of all coding exons in gDNA of the probands of the six autosomal dominant families further excluded other coding genetic variants in *PAXIP1*, except for the known g.56231A→G variant (*rs3501*); this predicts a Met979Val mutation in exon 18, with a frequency of 26.9% for the G-allele, which was present in two heterozygous individuals and one homozygous autosomal dominant proband. Also, expansion of a polyglutamine repeat in *PAXIP1* exon 7 was excluded in six patients of family 1270, by Southern blot analysis of PCR-amplified exon 7 genomic fragments. Since *PAXIP1* g.38030G→C was absent in the three families with multiplex AD who potentially share a common founder with family 1270 and since we limited our mutation analysis to coding regions, we cannot exclude the possibility that a mutation in 1 of the other 28 positional candidate genes might be the underlying genetic defect in family 1270. Also, a mutation not detected by genomic sequencing of exons—for example, whole-exon deletion—remains plausible. Further, our current annotation efforts might have missed coding sequences. The assembled sequence of the candidate region still contains a gap of ~80 kb between the genes *INSIG1* and *EN2* in which additional genes might be present, although the syntenic region in mouse does not contain genes.

Together, our linkage data of the informative family 1270, the association data of the population-based patient-control sample of early-onset AD, and the haplotype-sharing data of probands of three additional families with multiplex AD strongly support the locus at 7q36 as a novel locus for AD. Mutation analysis of all 29 known genes in the candidate region identified only an exonic silent mutation in *PAXIP1*, segregating in family 1270, which was absent from 640 control chromosomes (mutant allele frequency <0.16%). However, whether *PAXIP1* has a functional role in expression of AD in family 1270 remains unclear. The novel AD locus at 7q36 provides further evidence that AD is highly genetically heterogeneous, with genes at different loci contributing to the expression of the phenotype. In addition, it can be expected that the identification of the underlying gene defect at 7q36 might largely contribute to our current understanding of the genetic etiology and pathogenesis of AD.

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics*.

**Figure 6** Expression analysis of *PAXIP1*. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics*.



## Acknowledgments

We are grateful to the family members for their cooperation in this study. We also acknowledge Gerwin Roks and Hilda Kornman, for their contributions to the genealogy and ascertainment of family 1270; Sebastiaan Engelborghs, for expert second opinion on the clinical diagnoses; and the Flanders Interuniversity Institute for Biotechnology (VIB8) Genetic Service Facility, for genetic analyses. The research was funded by the Special Research Fund of the University of Antwerp; the Fund for Scientific Research Flanders (FWO-F); Interuniversity Attraction Poles program P5/19 of the Belgian Science Policy Office; the International Alzheimer Research Foundation, Belgium; EU contract LSHM-CT-2003-503330 (APOPIS); and the Netherlands Organization for Scientific Research. R.R., M.C., and J.T. are postdoctoral fellows of the FWO-F.

## Web Resources

Accession numbers and URLs for data presented herein are as follows:

Arlequin, <http://lgb.unige.ch/arlequin/> (for estimation of haplotype frequencies)  
 Alzheimer Disease & Frontotemporal Dementia Mutation Database, <http://www.molgen.ua.ac.be/ADMutations/>  
 Center for Medical Genetics, <http://research.marshfieldclinic.org/genetics/> (for the Marshfield Medical Research Foundation and marker position and extra markers in region of interest)  
 Cooperative Human Linkage Center, <http://gai.nci.nih.gov/CHLC/> (for extra marker in region of interest)  
 GenBank, <http://www.ncbi.nlm.nih.gov/Genbank/> (for *PAXIP1* [accession numbers AC093726.3 and NM\_007349.6], *ABCF2* [accession number AC021097.5], *GALNT11* [accession number AC074257.5], *DPP6* [accession number AC006019.2], and *EN2* [accession number AC008060.5])  
 Genepop, <http://wbiomed.curtin.edu.au/genepop/> (for single SNP association analysis)  
 NCBI Map Viewer, <http://www.ncbi.nlm.nih.gov/mapview/> (for gene position as well as links to sequence and protein function information)  
 Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for AD, *APP*, *PSEN1*, *PSEN2*, *APOE*, *MAPT*, *PRNP*, frontotemporal dementia, and Creutzfeldt-Jacob disease)  
 SNPbox, <http://www.snpbox.org/>  
 UCSC Genome Bioinformatics, <http://genome.ucsc.edu/> (for gene position and sequence information)  
 VIB8 Genetic Service Facility, <http://www.vibgeneticservice.facility.be/>

## References

- Ashley-Koch AE, Shao Y, Rimmler JB, Gaskell PC, Welsh-Bohmer KA, Jackson CE, Scott WK, Haines JL, Pericak-Vance MA (2005) An autosomal genomic screen for dementia in an extended Amish family. *Neurosci Lett* 379:199–204
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580
- Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921–923
- Cruts M, Van Broeckhoven C (1998) Presenilin mutations in Alzheimer's disease. *Hum Mutat* 11:183–190
- Cruts M, van Duijn CM, Backhovens H, Van den BM, Wehnert A, Serneels S, Sherrington R, Hutton M, Hardy J, George-Hyslop PH, Hofman A, Van Broeckhoven C (1998) Estimation of the genetic contribution of presenilin-1 and -2 mutations in a population-based study of presenile Alzheimer disease. *Hum Mol Genet* 7:43–51
- Dermaut B, Croes EA, Rademakers R, Van den BM, Cruts M, Hofman A, van Duijn CM, Van Broeckhoven C (2003) PRNP Val129 homozygosity increases risk for early-onset Alzheimer's disease. *Ann Neurol* 53:409–412
- Dermaut B, Cruts M, Backhovens H, Lubke U, Van Everbroeck B, Sciot R, Dom R, Martin JJ, Van Broeckhoven C, Cras P (2000) Familial Creutzfeldt-Jakob disease in a patient carrying both a presenilin 1 missense substitution and a prion protein gene insertion. *J Neurol* 247:364–368
- Dermaut B, Cruts M, Slooter AJ, Van Gestel S, De Jonghe C, Vanderstichele H, Vanmechelen E, Breteler MM, Hofman A, van Duijn CM, Van Broeckhoven C (1999) The Glu318Gly substitution in presenilin 1 is not causally related to Alzheimer disease. *Am J Hum Genet* 64:290–292
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152–154
- Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349:704–706
- Hillier LW, Fulton RS, Fulton LA, Graves TA, Pepin KH, Wagner-McPherson C, Layman D, et al (2003) The DNA sequence of human chromosome 7. *Nature* 424:157–164
- Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA (1991) Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* 7:403–422
- Hofman A, Schulte W, Tanja TA, van Duijn CM, Haaxma R, Lameris AJ, Otten VM, Saan RJ (1989) History of dementia and Parkinson's disease in 1st-degree relatives of patients with Alzheimer's disease. *Neurology* 39:1589–1592
- Honig LS, Mayeux R (2001) Natural history of Alzheimer's disease. *Aging (Milano)* 13:171–182
- Jimenez-Escrig A, Gomez-Tortosa E, Baron M, Rabano A, Arcos-Burgos M, Guillermo PL, Yusta A, Anta P, Perez I, Hierro M, Munoz DG, Barquero S (2005) A multigenerational pedigree of late-onset Alzheimer's disease implies new genetic causes. *Brain* 128:1707–1715
- Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am J Hum Genet* 37:482–498
- Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K

- (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 269:973–977
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (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
- Ott A, Breteler MM, van Harskamp F, Claus JJ, van der Cammen TJ, Grobbee DE, Hofman A (1995) Prevalence of Alzheimer's disease and vascular dementia: association with education: the Rotterdam study. *BMJ* 310:970–973
- Ott A, Breteler MM, van Harskamp F, Stijnen T, Hofman A (1998) Incidence and risk of dementia: the Rotterdam Study. *Am J Epidemiol* 147:574–580
- Pericak-Vance MA, Grubber J, Bailey LR, Hedges D, West S, Santoro L, Kemmerer B, Hall JL, Saunders AM, Roses AD, Small GW, Scott WK, Conneally PM, Vance JM, Haines JL (2000) Identification of novel genes in late-onset Alzheimer's disease. *Exp Gerontol* 35:1343–1352
- Rademakers R, Cruts M, Dermaut B, Sleegers K, Rosso SM, Van den BM, Backhovens H, Van Swieten J, van Duijn CM, Van Broeckhoven C (2002) Tau negative frontal lobe dementia at 17q21: significant finemapping of the candidate region to a 4.8 cM interval. *Mol Psychiatry* 7:1064–1074
- Rademakers R, Dermaut B, Peeters K, Cruts M, Heutink P, Goate A, Van Broeckhoven C (2003) Tau (MAPT) mutation Arg406Trp presenting clinically with Alzheimer disease does not share a common founder in Western Europe. *Hum Mutat* 22:409–411
- Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, Chi H, Lin C, Holman K, Tsuda T (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
- Roks G, Dermaut B, Heutink P, Julliams A, Backhovens H, Van den Broeck M, Serneels S, Hofman A, Van Broeckhoven C, van Duijn CM, Cruts M (1999) Mutation screening of the tau gene in patients with early-onset Alzheimer's disease. *Neurosci Lett* 277:137–139
- Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan DR, Alberts MJ (1993) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43:1467–1472
- Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375:754–760
- Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, van Duijn CM (1998) Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol* 55:964–968
- Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, Roses AD (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci USA* 90:1977–1981
- van Duijn CM, de Knijff P, Cruts M, Wehnert A, Havekes LM, Hofman A, Van Broeckhoven C (1994a) Apolipoprotein E4 allele in a population-based study of early-onset Alzheimer's disease. *Nat Genet* 7:74–78
- van Duijn CM, Hendriks L, Cruts M, Hardy JA, Hofman A, Van Broeckhoven C (1991) Amyloid precursor protein gene mutation in early-onset Alzheimer's disease. *Lancet* 337:978
- van Duijn CM, Hendriks L, Farrer LA, Backhovens H, Cruts M, Wehnert A, Hofman A, Van Broeckhoven C (1994b) A population-based study of familial Alzheimer disease: linkage to chromosomes 14, 19, and 21. *Am J Hum Genet* 55:714–727
- Warwick DE, Payami H, Nemens EJ, Nochlin D, Bird TD, Schellenberg GD, Wijsman EM (2000) The number of trait loci in late-onset Alzheimer disease. *Am J Hum Genet* 66:196–204
- Weckx S, De Rijk P, Van Broeckhoven C, Del Favero J (2005) SNPbox, a modular software package for large-scale primer design. *Bioinformatics* 21:385–387