

## ARTICLE

# Expanded high-resolution genetic study of 109 Swedish families with Alzheimer's disease

Anna Sillén<sup>1</sup>, Jorge Andrade<sup>2</sup>, Lena Lilius<sup>1</sup>, Charlotte Forsell<sup>1</sup>, Karin Axelman<sup>1</sup>,  
Jacob Odeberg<sup>2</sup>, Bengt Winblad<sup>1</sup> and Caroline Graff<sup>\*,1</sup>

<sup>1</sup>Department of Neurobiology, Care Sciences and Society, Karolinska Institutet Dainippon Sumitomo Pharma Alzheimer Center (KASPAC), Karolinska Institutet, Huddinge, Sweden; <sup>2</sup>Department of Biotechnology, AlbaNova University Center, Royal Institute of Technology (KTH), Stockholm, Sweden

Alzheimer's disease (AD) is a neurodegenerative disease that affects approximately 20 million persons all over the world. There are both sporadic and familial forms of AD. We have previously reported a genome-wide linkage analysis on 71 Swedish AD families using 365 genotyped microsatellite markers. In this study, we increased the number of individuals included in the original 71 analysed families besides adding 38 new families. These 109 families were genotyped for 1100 novel microsatellite markers. The present study reports on the linkage data generated from the non-overlapping genotypes from the first genome scan and the genotypes of the present scan, which results in a total of 1289 successfully genotyped markers at an average density of 2.85 cM on 468 individuals from 109 AD families. Non-parametric linkage analysis yielded a significant multipoint LOD score in chromosome 19q13, the region harbouring the major susceptibility gene *APOE*, both for the whole set of families (LOD = 5.0) and the *APOE*  $\epsilon$ 4-positive subgroup made up of 63 families (LOD = 5.3). Other suggestive linkage peaks that were observed in the original genome scan of 71 Swedish AD families were not detected in this extended analysis, and the previously reported linkage signals in chromosomes 9, 10 and 12 were not replicated.

*European Journal of Human Genetics* (2008) 16, 202–208; doi:10.1038/sj.ejhg.5201946; published online 24 October 2007

**Keywords:** linkage analysis; familial dementia; *APOE*; Alzheimer's disease; genome scan; lod score

## Introduction

The most common cause of dementia in the elderly is Alzheimer's disease (AD), characterized by a progressive decline of cognitive functions, particularly memory. The neuropathological hallmarks of the disease are numerous cortical amyloid plaques formed by aggregation of amyloid  $\beta$ -peptide ( $A\beta$ ), neurofibrillary tangles containing hyper-

phosphorylated tau protein and atrophy. There are different hypotheses underlying the pathology of AD and the amyloid cascade hypothesis is the most prevalent.<sup>1–3</sup> Oligomerization and accumulation of the  $A\beta$ 1–40 and  $A\beta$ 1–42 peptides, as well as failure of normal degradation of the peptides, are the key events in this amyloid cascade hypothesis. Sporadic AD accounts for the majority of cases and familial AD (FAD, defined as at least two affected first-degree relatives) constitutes a smaller fraction of all AD cases. Less than 1% of all AD cases are autosomal dominant, early-onset (before 65 years of age) monogenic forms caused by mutations in one of the three known genes: amyloid precursor protein (*APP*, OMIM: no. 104300<sup>4,5</sup>), presenilin 1 (*PSEN1*, OMIM: no. 607822<sup>6</sup>) and presenilin 2 (*PSEN2*, OMIM: no. 606889<sup>7,8</sup>). The functions

\*Correspondence: Dr C Graff, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet Dainippon Sumitomo Pharma Alzheimer Center (KASPAC), Karolinska Institutet, Novum level 5, Huddinge SE-14157, Sweden.

Tel: +46 8 585 836 19; Fax: +46 8 585 836 10;

E-mail: Caroline.Graff@ki.se

Received 14 May 2007; revised 26 September 2007; accepted 27 September 2007; published online 24 October 2007

of these three genes fit well with the amyloid cascade hypothesis concerning the AD aetiology, as the encoded proteins are all involved in APP processing and/or are reported to increase A $\beta$  production.<sup>9–12</sup> In contrast, the  $\epsilon$ 4 allele of the *APOE* gene is the only consistently confirmed genetic risk-factor for the sporadic forms of the disease, which predominantly acts by reducing the age at onset of disease in a dose-dependent manner.<sup>13</sup>

There are today a handful of published articles involving linkage analysis on FAD, and even though different research groups have confirmed linkage to regions in chromosomes 6,<sup>14–18</sup> 9,<sup>14,16–20</sup> 10q<sup>14,16–18,21</sup> and 12p,<sup>14,15,17</sup> no new genes with causal mutations or with increased effects on AD risk have been identified. Other genomic regions that have been shown to be linked to or associated with AD in previous studies are 1q23–31, 2p12–q11, 4p16, 4q35, 5p13–15 and 19q13.<sup>14,16–19,21–25</sup> For recent reviews and meta-analyses see <http://www.alzforum.org/res/com/gen/alzgene/linkage.asp>, Bertram *et al*,<sup>26</sup> and Kamboh.<sup>27</sup>

Recently, we published the first Swedish genome-wide scan on 71 AD families (GS1) and found a significant linkage peak corresponding to the *APOE* region in chromosome 19q13.<sup>28</sup> In the present study (GS2), we have added additional affected and unaffected individuals to the original 71 families as well as 38 novel families and increased the density of genotyped markers (from an average intermarker distance of 8.97 cM to one marker per 2.85 cM) by genotyping 1102 novel microsatellites. Thus, we report the results from a genome-wide linkage analysis based on the data generated by genotyping 1289 markers on 468 individuals from 109 Swedish AD families.

## Materials and methods

### Samples

The families were selected from our research registry of neurodegenerative dementias at Department of NVS, Karolinska Institutet, Huddinge, Sweden. The families were recruited either through referrals from primary caregivers, memory clinics or by self-referrals from all of Sweden. The inclusion criteria for the study were a positive family history for dementia (at least two affected first-degree relatives) and that DNA was available on at least two affected relatives in each family. The family history of dementia, the age at onset (the age when first symptoms appeared) and the disease course was based on medical records, autopsy reports, if available, and genealogy as well as through interviews of relatives. DNA was available on 292 affected (188 included in GS1) and 176 unaffected individuals from 109 families (Table 1). The average number of unaffected individuals genotyped in each family was 1 (range 0–13). The sample set included 284 women (64.4% affected) and 184 men (59.2% affected), with an average age at onset (AAO) of  $68.7 \pm 7.5$  ( $\pm 1SD$ ) years.

**Table 1** Pedigree size distribution of the AD families included in the study

No. of families	No. of affected individuals genotyped in the families
2	7
4	6
7	5
7	4
25	3
64	2
109	

At least one unaffected family member was genotyped in each family except for 24 families where DNA was not available.

Thirty-five of the families had an early-onset (family mean AAO  $\leq 65$  years of age), and 87 families had a history of affected individuals in two or more generations, consistent with a dominant inheritance pattern.

All families contained at least one affected family member with a clinical diagnosis of AD according to NINCDS-ADRDA criteria<sup>29</sup> and/or neuropathological diagnosis of AD. In the majority of families, all cases had a clinical diagnosis of AD. In 24 families, there was at least one neuropathological diagnosis of AD according to CERAD criteria.<sup>30,31</sup> However, there were several families with more than one clinical dementia diagnosis both in a single individual and/or in different cases from the same family. Thus, several cases had a diagnosis of AD in combination with atypical AD signs such as predominant frontal symptoms, Lewy body signs, Parkinsonism and vascular components. A group of 18 families had a clinical diagnosis of 'mixed AD'. The mixed AD families contained families where there were cases with a clinical diagnosis of AD as well as cases with a vascular dementia diagnosis (VaD, including stroke with dementia and multi-infarction dementia) or cases with a combination of both diagnoses. Furthermore, there were families with two different neuropathological diagnoses in two different siblings. Table 2 shows the distribution of the sub-phenotypes among the families based on the diagnoses presented in the affected individuals in the family: definitive AD (when at least one family member had autopsy-confirmed AD), clinical AD (all cases had clinical AD), AD and vascular dementia; when both AD and VaD were present in the family and/or a single case had mixed dementia diagnosis, AD and frontotemporal dementia, AD and Lewy body dementia and AD and Parkinson's disease.

Of the 109 families, 63 were *APOE*  $\epsilon$ 4-positive (all affected carried at least one  $\epsilon$ 4 allele) and 46 were  $\epsilon$ 4 negative (at least one affected did not carry any  $\epsilon$ 4 alleles). Families with known mutations in the *APP*, *PSEN1* and *PSEN2* genes were excluded. Two of the included markers (D21S1914 and D21S1442) are located within 1.6 Mb of the

**Table 2** Subgrouping of the 109 AD families according to phenotypes present in the family

Subgroup	Phenotype	No. of families
Autopsy-confirmed AD	At least one family member with autopsy-confirmed AD	24
Clinical AD	All cases having clinical AD	65
AD and VaD	Both AD and vascular dementia present in the family	18
AD and FTD	Both AD and frontotemporal dementia present in the family	4
AD and LBD	Both AD and Lewy body dementia present in the family	1
AD and PD	Both AD and Parkinson's disease present in the family	1

Note that the sum is greater than 109 because some of the autopsy-confirmed families are also represented in other subgroups.

*APP* gene, and any duplication resulting in three alleles would be detected in the quality-check procedure. However, duplicated segments smaller than 1.6 Mb and/or duplications resulting in two alleles would not be detected by this method, and thus quantitative PCR for detection of *APP* duplications is ongoing. Mutation screening in the granulin gene (*GRN*, MIM no. 138945) has not been performed. Genealogical studies have been made on the majority of the 109 families and shared ancestries have been identified in 12 families (ie 12 of the families have a genetic history in common with one other family). Three families, presently living in Sweden, originated from Finland in earlier generations, two from Norway and two from Germany, married into Swedish families.

Blood samples were collected after informed consent by participating individuals or next of kin, and the study was approved by the local Ethics Committee at the Karolinska Institutet, Huddinge, Sweden. DNA was extracted according to standard protocols. Whole genome amplification was performed on a total of 78 DNA samples either by us using the GenomiPhi DNA Amplification kit (GE Healthcare BioSciences AB, Uppsala, Sweden) or by deCODE genetics Inc., Iceland, because of small starting amounts of genetic material. The markers were genotyped at the genotyping service; deCODE genetics Inc., Iceland, using their 1000 marker panel. DNA samples from 468 individuals were genotyped for 1102 markers. Combining the genotypes from GS1 (187 non-redundant markers genotyped in 188 affected individuals) and the present genome-wide scan (GS2), there was a total of 1289 genotyped markers leading to an average intermarker distance of 2.85 cM. Intermarker distances and their order were obtained from deCODE and by combining the deCODE map<sup>32</sup> with publicly available genetic maps from Marshfield and Généthron.<sup>33,34</sup> Marker allele frequencies were estimated from the families, which tends to give conservative results.<sup>35</sup> The graphical presentation of the multipoint (mpt) results were converted to

$Z_{lr}$  scores, which reflect the sign of  $d$ hat unlike Allegro LOD scores.

### Statistical analyses

Linkage analysis was performed using the information of 1289 markers on the whole set of 109 families and in the *APOE*-stratified groups (63 *APOE*  $\epsilon$ 4-positive families and 46 *APOE*  $\epsilon$ 4-negative families) using the Allegro version 1.2,<sup>36</sup> applying both mpt and singlepoint (spt) analyses. All unaffected siblings were coded as having unknown disease status in the calculations. In the subset of 46  $\epsilon$ 4-negative families, the affection status for affected individuals carrying an  $\epsilon$ 4 allele was set as 'unknown'. Thereby, these individuals do not contribute to the LOD score, but they add information about phase. Non-parametric allele sharing LOD scores,  $Z_{lr}$  scores and NPL scores were obtained. The non-parametric model was used, as the pedigrees show mixed patterns of disease inheritance and the true underlying inheritance model is unknown. We used the exponential model due to its higher robustness when handling pedigrees of different sizes, and scoring function  $S_{pairs}$ , as suggested by McPeck,<sup>37</sup> when there is no clear disease inheritance model. Taking the family size differences into consideration (ranging from 2 to 23 bits), the family weighting option 'power: 0.5' was used as suggested in the Allegro manual. Parametric analysis allowed for heterogeneity assuming 5, 65 and 80% penetrances for homozygous wild type, heterozygotes and homozygosity for the disease allele, respectively, and a disease allele frequency of 5% was also performed using the parametric option in Allegro. Three of the largest families had to be cut in size (younger genotyped persons with unknown disease status were removed) to fit the size limitations of 25 bits in the Allegro program and also to save computer time. Linkage analyses calculations of chromosome 19 were performed both with and without the *APOE* genotypes.

Simulation analysis under the null hypothesis of no linkage across the whole genome was performed 1000 times with simulated genotypes using the same marker map, allele frequencies and pedigree structure and assuming 7% of missing genotypic data. We developed a grid-aware computer implementation of the Allegro program, by which many thousands of calculations can be executed in parallel and thereby save months of computer time.<sup>38</sup> To estimate the empirical genome-wide significance (GWS) level, the three highest obtained LOD scores for the total ( $N=109$  families), and the *APOE*  $\epsilon$ 4-stratified groups ( $N=63$  and  $N=46$ , respectively) were used as threshold levels, which were compared to the number of times these LOD scores were estimated in the simulated data. A GWS probability in which  $P$ -value is less than 0.05, in other words an occurrence of a LOD score equal to or higher than a given threshold once per 20 genome scans, was used as the definition of significance, in agreement with Lander and Kruglyak.<sup>39</sup>

Power calculations on the 109 families were performed using the AllegroSim option assuming heterogeneity with  $\alpha=30\%$  (ie 30% of families linked to other loci), 7% missing genotypes and the authentic pedigree structures.

## Results

We report the results of a follow-up genome scan in Swedish AD families. In this study, we genotyped 1102 microsatellite markers in 486 individuals from 109 AD families with a success rate of 96%, and possible genotyping errors were minimized by checking for Mendelian inconsistencies before starting the linkage analysis. These genotypes were combined with the non-redundant genotypes from our previous study,<sup>28</sup> resulting in information from a total of 1289 genotyped markers. All linkage data presented in table format were acquired without the *APOE* genotypes. Supplementary Figure 1 illustrates the  $Z_{lr}$  score curves for all analysed chromosomes. Table 3 presents a summary of the results with the highest obtained non-parametric mpt/spt LOD scores for the full set of families ( $N=109$ ), the *APOE*  $\epsilon 4$ -positive ( $N=63$ ) and  $\epsilon 4$ -negative ( $N=46$ ) stratified families. Additional results with all mpt LOD scores  $\geq 1$  are listed in Supplementary Table 1 and all obtained spt LOD scores  $\geq 1.5$  are listed in Supplementary Table 2. The entire set of 109 families generated its highest linkage peak in 19q13.33, with a significant mpt LOD score of 5.05 and a significant spt LOD score of 3.86. The second highest spt LOD score for the 109 families was located in 4q25 and did not reach significance: spt LOD = 2.31 (GWS  $P=0.31$ ), corresponding mpt LOD = 0.16 ( $P=1$ ).

The highest and most significant mpt LOD score in the *APOE*  $\epsilon 4$ -positive families was 5.31 ( $P=0.0011$ ) located at marker D19S903 approximately 0.35 cM from the *APOE* locus. The only other LOD score that reached GWS in the  $\epsilon 4$ -positive families was obtained in 6p24 spt LOD = 3.21 ( $P=0.044$ ). However, its matching mpt score was low and insignificant (LOD = 0.23) as well as the spt LOD scores for the flanking markers. The third highest obtained spt LOD score in this group was 2.31 ( $P=0.28$ ) at the same peak marker D4S2989 in 4q25 as in the whole set of families.

The corresponding mpt LOD score (0.07) in 4q25 was insignificant ( $P=1$ ) as well as the spt LOD scores for the flanking markers.

None of the obtained linkage peaks reached significance in the subset of 46 *APOE*  $\epsilon 4$ -negative families. The LOD scores in chromosome 19 were not significantly different when *APOE* genotypes were included in the analysis, except for the 63 *APOE*  $\epsilon 4$ -positive families, where the maximum spt LOD score increased from 5.3 to 9.3 at the *APOE* marker.

Parametric analysis allowing for heterogeneity did not identify any other loci besides *APOE*, and the number of families linked to *APOE* was 76% (data not shown).

## Discussion

By increasing the number of participating families in the present study from 71 to 109 families, we hoped to increase the genetic information enough to obtain stronger linkage signals in the suggestive linkage regions observed in the original genome scan.<sup>28</sup> A little surprisingly, the only significant linkage peak obtained by mpt analysis in the full set of 109 Swedish AD families (mpt LOD = 5.05,  $P=0.015$ ) was still a reflection of the known *APOE* gene in chromosome 19q13. Sixty-three of the 109 families (58%), in which all affected carried at least one  $\epsilon 4$  allele, generated a significant mpt LOD of 5.31 at a distance of 0.35 cM centromeric to the *APOE* gene even in the absence of the *APOE* genotypes in the analysis. The data suggest that our analysed family material is under a very strong influence of the  $\epsilon 4$  allele, as supported by both non-parametric and parametric linkage analysis. The AAO was 68.7 years for all families, which is also the age at which the *APOE* gene has been reported to exert its strongest effect.<sup>40</sup> Besides the peak in 19q13.33, the only other LOD score that reached the level of significance was in 6p24 for the subset of 63 *APOE*  $\epsilon 4$ -positive families, with a spt LOD score of 3.21 ( $P=0.044$ ) at marker D6S1279. However, the weak mpt LOD score and insignificant spt LOD scores for the flanking markers (Supplementary Table 2) imply that this is most likely a nonsignificant spurious effect. Furthermore, the

**Table 3** A summary of the single highest obtained multipoint (mpt) and the single highest obtained singlepoint (spt) non-parametric LOD scores in the complete family material, the *APOE*  $\epsilon 4$ -positive and *APOE*  $\epsilon 4$ -negative families, respectively, are listed with their corresponding genome-wide significant values ( $P$ -values)

Chromosome	Stratum (No. of families)	Peak marker	Max LOD (P-value)	Max $Z_{lr}$	Analysis
19	All families (109)	Close to D19S900	<b>5.05 (0.015)</b>	4.82	Mpt
4	All families (109)	D4S2989	2.31 (0.31)	3.26	Spt
19	<i>APOE</i> $\epsilon 4$ positive (63)	Close to D19S903	<b>5.31 (0.0011)</b>	4.95	Mpt
6	<i>APOE</i> $\epsilon 4$ positive (63)	D6S1279	<b>3.21 (0.044)</b>	3.84	Spt
19	<i>APOE</i> $\epsilon 4$ negative (46)	D19S246	1.06 (0.86)	2.21	Mpt
4	<i>APOE</i> $\epsilon 4$ negative (46)	D4S426	1.86 (0.62)	2.93	Spt

Significant LOD scores in bold. Note that the corresponding mpt LOD score in chromosomes 4 and 6 is low and nonsignificant, as illustrated in Supplementary Table 2.

6p24 region has not been reported earlier to be linked to AD, although findings of several smaller linkage peaks (LODs between 1.5 and 1.9) in the adjacent chromosomal regions 6p21 and 6q21 have been described.<sup>14,15,17,18</sup> In addition to the already discussed highly significant linkage to 19q13.33, the 4q25 region appeared as a suggestive peak both in the original (spt LOD = 2.35,  $P = 0.35$  in the  $\epsilon 4$ -positive subgroup) and this extended (spt LOD = 2.31,  $P = 0.28$  in the  $\epsilon 4$ -positive subgroup, and LOD = 2.31,  $P = 0.31$  in 109 families) genome-wide scan for AD susceptibility genes. Although simulation analyses and the spt LOD scores of flanking markers (Supplementary Table 2) suggest that these LOD scores are not significant, one may be cautious to rule out the possibility of a susceptibility gene in this part of the genome as at least one candidate gene, *COL25A1*, is located in this region.<sup>41,42</sup> Furthermore, an ongoing association study of the *COL25A1* gene indicates that it may contribute to the genetic risk of developing AD (C Forsell *et al*, manuscript in preparation).

It is also noteworthy that we did not find linkage to the reported linkage peaks in sib-pairs to chromosomes 9, 10 or 12.<sup>14–21</sup> It is unlikely that the reason for this inconsistency is a reflection of the relative contribution of *APOE* in the different study populations, as the study by Blacker *et al*<sup>18</sup> also had strong evidence of linkage to the *APOE* gene (LOD 7.7), similar to our study. However, the stage 2 study by Myers *et al*<sup>17</sup> only reported a weak-linkage signal in the *APOE* region (LOD below 1.6). Furthermore, neither chromosome 17 (granulin, *GRN*) nor chromosome 21 (*APP*) generated any suggestive linkage peaks in our linkage analysis, suggesting that these two genes have no major causative effect in our AD population.

The analysed families were not diagnosed solely with AD, but with a combination of dementias. It is possible that this mixture made the family set too heterogeneous genetically. However, it has also been hypothesized that the same genes may contribute to neurodegeneration in general, but that the different combination of susceptibility genes will result in different phenotypes.<sup>43–45</sup> Furthermore, there are examples of varied clinical phenotypes in family members carrying the same *APP* duplication,<sup>46</sup> *GRN* mutation<sup>47</sup> and in families segregating frontotemporal dementia with motor neuron disease,<sup>48</sup> and we clearly have families with variable presenting symptoms as well as neuropathological changes among siblings. Another explanation for the lack of significant linkage might be the complex segregation pattern in some of our AD families. Dementia was present in both paternal and maternal lines for 11 of the families, adding further to the complexity of the involved pathogenic genes. Looking back on the published genetic studies of familial AD, the lack of success in identifying additional genes harbouring causative mutations gives a sense of the large problems and difficulties in determining the aetiology of the disease. The

answer may lie in improving the clinical classification of the different sub-phenotypes of dementia, using only autopsy-confirmed AD cases in the linkage analysis as Gordon *et al*<sup>23</sup> did in their suggested 'gold-standard' method or by mapping quantitative traits such as measurable biomarkers, for example, protein levels in cerebrospinal fluid or variables obtained by brain imaging using MRI.<sup>49</sup> Finally, a whole genome association study design may prove to be a more fruitful approach for identifying additional susceptibility genes in AD. The first reported high-density genome-wide SNP association study of 1086 definitive AD cases showed the strongest association to an SNP located 14 kb from the *APOE*  $\epsilon 4$  variant,<sup>50</sup> but no other significant loci were repeated. The major obstacle regardless of approach is probably that the number of cases required may be 10- to 100-fold that which has been used in single studies so far.<sup>51</sup> Follow-up studies are ongoing in the subset of 24 families with autopsy-confirmed AD and in the 18 families with 'mixed AD' as well as in families with shared ancestries.

#### Acknowledgements

We acknowledge the contribution of the neuropathological examinations by Associate Professors Inger Nennesmo and Nenad Bogdanovic. We also express sincere gratitude to the participating families. This study was supported by Dainippon Sumitomo Pharma Co., Ltd.

#### Electronic-database information

deCODE Genetic Map:

<http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?ORG=hum&MAPS=ideoqr,decode>

Généthon Genetic Map:

<http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?taxid=9606&chr=7&maps=ideoqr,thon>

Marshfield Genetic Map:

<http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?ORG=hum&MAPS=ideoqr,marsh>

#### References

- 1 Selkoe DJ: Cell biology of the amyloid beta-protein precursor and the mechanism of Alzheimer's disease. *Annu Rev Cell Biol* 1994; **10**: 373–403.
- 2 Haass C, Selkoe DJ: Cellular processing of beta-amyloid precursor protein and the genesis of amyloid beta-peptide. *Cell* 1993; **75**: 1039–1042.
- 3 Hardy JA, Higgins GA: Alzheimer's disease: the amyloid cascade hypothesis. *Science* 1992; **256**: 184–185.
- 4 Goate A, Chartier-Harlin M-C, Mullan M *et al*: Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991; **349**: 704–706.
- 5 Rovelet-Lecrux A, Hannequin D, Raux G *et al*: APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* 2006; **38**: 24–26.
- 6 Sherrington R, Rogaev EI, Liang Y *et al*: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995; **375**: 754–760.

- 7 Rogaev EI, Sherrington R, Rogaeva EA *et al*: 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* 1995; **376**: 775–778.
- 8 Levy-Lahad E, Wijsman EM, Nemens E *et al*: A familial Alzheimer's disease locus on chromosome 1. *Science* 1995; **269**: 970–973.
- 9 Haass C, Hung AY, Selkoe DJ, Teplow DB: Mutations associated with a locus for familial Alzheimer's disease result in alternative processing of amyloid beta-protein precursor. *J Biol Chem* 1994; **269**: 17741–17748.
- 10 Citron M, Westaway D, Xia W *et al*: Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nat Med* 1997; **3**: 67–72.
- 11 Citron M, Vigo-Pelfrey C, Teplow DB *et al*: Excessive production of amyloid beta-protein by peripheral cells of symptomatic and presymptomatic patients carrying the Swedish familial Alzheimer's disease mutation. *Proc Natl Acad Sci USA* 1994; **91**: 11993–11997.
- 12 Thinakaran G, Borchelt DR, Lee MK *et al*: Endoproteolysis of presenilin 1 and accumulation of processed derivatives *in vivo*. *Neuron* 1996; **17**: 181–190.
- 13 Blacker D, Haines JL, Rodes L *et al*: ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. *Neurology* 1997; **48**: 139–147.
- 14 Curtis D, North BV, Sham PC: A novel method of two-locus linkage analysis applied to a genome scan for late onset Alzheimer's disease. *Ann Hum Genet* 2001; **65**: 473–481.
- 15 Pericak-Vance MA, Bass MP, Yamaoka LH *et al*: Complete genomic screen in late-onset familial Alzheimer disease. Evidence for a new locus on chromosome 12. *JAMA* 1997; **278**: 1237–1241.
- 16 Olson JM, Goddard KA, Dudek DM: A second locus for very-late-onset Alzheimer disease: a genome scan reveals linkage to 20p and epistasis between 20p and the amyloid precursor protein region. *Am J Hum Genet* 2002; **71**: 154–161.
- 17 Myers A, Wavrant De-Vrieze F, Holmans P *et al*: Full genome screen for Alzheimer disease: stage II analysis. *Am J Med Genet* 2002; **114**: 235–244.
- 18 Blacker D, Bertram L, Saunders AJ *et al*: Results of a high-resolution genome screen of 437 Alzheimer's disease families. *Hum Mol Genet* 2003; **12**: 23–32.
- 19 Pericak-Vance MA, Grubber J, Bailey LR *et al*: Identification of novel genes in late-onset Alzheimer's disease. *Exp Gerontol* 2000; **35**: 1343–1352.
- 20 Scott WK, Hauser ER, Schmechel DE *et al*: Ordered-subsets linkage analysis detects novel Alzheimer disease loci on chromosomes 2q34 and 15q22. *Am J Hum Genet* 2003; **73**: 1041–1051.
- 21 Li YJ, Scott WK, Hedges DJ *et al*: Age at onset in two common neurodegenerative diseases is genetically controlled. *Am J Hum Genet* 2002; **70**: 985–993.
- 22 Hiltunen M, Mannermaa A, Thompson D *et al*: Genome-wide linkage disequilibrium mapping of late-onset Alzheimer's disease in Finland. *Neurology* 2001; **57**: 1663–1668.
- 23 Gordon D, Haynes C, Finch SJ, Brown AM: Increase in linkage information by stratification of pedigree data into gold-standard and standard diagnoses: application to the NIMH Alzheimer Disease Genetics Initiative Dataset. *Hum Hered* 2006; **61**: 97–103.
- 24 Lee JH, Mayeux R, Mayo D *et al*: Fine mapping of 10q and 18q for familial Alzheimer's disease in Caribbean Hispanics. *Mol Psychiatry* 2004; **9**: 1042–1051.
- 25 Farrer LA, Bowirrat A, Friedland RP, Waraska K, Korczyn AD, Baldwin CT: Identification of multiple loci for Alzheimer disease in a consanguineous Israeli–Arab community. *Hum Mol Genet* 2003; **12**: 415–422.
- 26 Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE: Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 2007; **39**: 17–23.
- 27 Kamboh MI: Molecular genetics of late-onset Alzheimer's disease. *Ann Hum Genet* 2004; **68**: 381–404.
- 28 Sillén A, Forsell C, Lilius L *et al*: Genome scan on Swedish Alzheimer's disease families. *Mol Psychiatry* 2006; **11**: 182–186.
- 29 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ARDRA work group under the auspices of Department of Health and Human Services task force on Alzheimer's disease. *Neurology* 1984; **34**: 939–944.
- 30 Tariot PN, Mack JL, Patterson MB *et al*: The behavior rating scale for dementia of the consortium to establish a registry for Alzheimer's disease. The behavioral pathology committee of the consortium to establish a registry for Alzheimer's disease. *Am J Psychiatry* 1995; **152**: 1349–1357.
- 31 Mirra SS, Heyman A, McKeel D *et al*: The consortium to establish a registry for Alzheimer's disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* 1991; **41**: 479–486.
- 32 Kong A, Gudbjartsson D, Sainz J *et al*: A high-resolution recombination map of the human genome. *Nat Genet* 2002; **31**: 241–247.
- 33 Broman KW, Murray JC, Sheffield VC, White RL, Weber JL: Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998; **63**: 861–869.
- 34 Dib C, Faure S, Fizames C *et al*: A comprehensive genetic map of the human genome based on 5264 microsatellites. *Nature* 1996; **380**: 152–154.
- 35 Ott J: Strategies for characterizing highly polymorphic markers in human gene mapping. *Am J Hum Genet* 1992; **51**: 283–290.
- 36 Gudbjartsson DF, Jonasson K, Frigge ML, Kong A: Allegro, a new computer program for multipoint linkage analysis. *Nat Genet* 2000; **25**: 12–13.
- 37 McPeck MS: Optimal allele-sharing statistics for genetic mapping using affected relatives. *Genet Epidemiol* 1999; **16**: 225–249.
- 38 Andrade J, Andersen M, Sillén A, Graff C, Odeberg J: The use of grid computing to drive data-intensive genetic research. *Eur J Hum Genet* 2007; **15**: 694–702.
- 39 Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
- 40 Farrer LA, Cupples LA, Haines JL *et al*: 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* 1997; **278**: 1349–1356.
- 41 Söderberg L, Zhukareva V, Bogdanovic N *et al*: Molecular identification of AMY, an Alzheimer disease amyloid-associated protein. *J Neuropathol Exp Neurol* 2003; **62**: 1108–1117.
- 42 Hashimoto T, Wakabayashi T, Watanabe A *et al*: CLAC: a novel Alzheimer amyloid plaque component derived from a transmembrane precursor, CLAC-P/collagen type XXV. *EMBO J* 2002; **21**: 1524–1534.
- 43 Dermaut B, Kumar-Singh S, Engelborghs S *et al*: A novel presenilin 1 mutation associated with Pick's disease but not beta-amyloid plaques. *Ann Neurol* 2004; **55**: 617–626.
- 44 Grabowski TJ, Cho HS, Vonsattel JP, Rebeck GW, Greenberg SM: Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy. *Ann Neurol* 2001; **49**: 697–705.
- 45 Raux G, Gantier R, Thomas-Anterion C *et al*: Dementia with prominent frontotemporal features associated with L113P presenilin 1 mutation. *Neurology* 2000; **55**: 1577–1578.
- 46 Cabrejo L, Guyant-Marchal L, Laquerriere A *et al*: Phenotype associated with APP duplication in five families. *Brain* 2006; **129**: 2966–2976.
- 47 Benussi L, Binetti G, Sina E *et al*: A novel deletion in progranulin gene is associated with FTDP-17 and CBS. *Neurobiol Aging* 2006 Dec 5; [Epub ahead of print].
- 48 Talbot K, Ansorge O: Recent advances in the genetics of amyotrophic lateral sclerosis and frontotemporal dementia:

- common pathways in neurodegenerative disease. *Hum Mol Genet* 2006; **15**: R182–R187.
- 49 Ramani A, Jensen JH, Helpert JA: Quantitative MR imaging in Alzheimer disease. *Radiology* 2006; **241**: 26–44.
- 50 Coon KD, Myers AJ, Craig DW *et al*: A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry* 2007; **68**: 613–618.
- 51 Teng J, Risch N: The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases. II. Individual genotyping. *Genome Res* 1999; **9**: 234–241.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)