Pathogenic mutations in *APP*, *PSEN1*, *PSEN2*, *GRN* and *MAPT* in families with late onset Alzheimer disease.

Cruchaga C[1], Haller G[1], Chakraverty S[1], Mayo K[1], Vallania FL[2], Mitra RD[2], Faber K[3], Williamson J[4], Bird T[5], Diaz-Arrastia R[6], Foroud T[3], Boeve B[7], Graff-Radford N[8], Mayeux R[4], and Goate A[1,2*] for the NIA-LOAD/NCRAD Family Study Consortium

1 Department of Psychiatry & Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University, St. Louis, MO, USA
2 Department of Genetics, Washington University, St. Louis, MO, USA
3 Department of Medical and Molecular Genetics, Indiana University, Indianapolis
4 Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University College of Physicians and Surgeons, New York, New York
5 VA Medical Center and Departments of Neurology and Medicine, University of Washington, Seattle, WA, USA
6 Department of Neurology, University of Texas Southwestern Medical Center, Dallas, Texas
7 Department of Neurology Mayo Clinic, Rochester, Minnesota
8 Department of Neurology, Mayo Clinic, Jacksonville, Florida

[*]To whom correspondence should be addressed at: Department of Psychiatry, Washington University School of Medicine, 660 South Euclid Avenue B8134, St. Louis, MO 63110. E-mail: goatea@psychiatry.wustl.edu, tel. 314-362-8691, fax. 314-747-2983

ABSTRACT

Background: Mutations in the genes encoding amyloid beta precursor protein (*APP*), presenilin 1 and 2 (*PSEN1, PSEN2*) are causes of autosomal dominant early onset Alzheimer disease, while mutations in the *microtubule associated protein tau (MAPT)* and *progranulin (GRN)* genes lead to familial forms of frontotemporal dementia (FTD). Late onset Alzheimer disease (LOAD) is not generally associated with mutations in these genes; however, families densely affected with LOAD were not the focus in earlier screening studies.

Methods: From a series of 1,431 families multiply affected by LOAD, we selected one affected individual from each of 439 families with a reported history of dementia in four or more individuals and used next-generation sequencing to identify pathogenic mutations in *APP*, *PSEN1*, *PSEN2*, *MAPT* and *GRN*.

Results: We identified eleven sequence variants, two each in *APP* and *PSEN1*, three in *MAPT* and four in *GRN*, in a total of 22 families (5% of the total families) that are pathogenic or likely pathogenic, based on the segregation data. Six of these variants were novel and five were previously reported.

Conclusions: At least 5% of families multiply affected by LOAD carry pathogenic mutations in *APP*, *PSEN1*, *PSEN2*, *MAPT* or *GRN*. Mutations in these genes are typically considered to be causes of early onset autosomal dominant Alzheimer disease and frontotemporal dementia. These results raise practical implications for the management of patients with familial late onset dementias, regarding appropriate diagnoses, counseling and treatment.

Mutations in *amyloid-beta precursor protein* (*APP)*, *presenilin 1* and 2 (*PSEN1, PSEN2*), cause autosomal dominant Alzheimer disease (AD) with typically onset typically prior to age 60. Families harboring these mutations usually have an extensive multigenerational history of disease. Mutations in *progranulin* (*GRN*) are an established cause of familial frontotemporal dementia (FTD), a form of dementia associated primarily with changes in behavior and language, but have also been observed in a series of clinically diagnosed AD.[1] Similarly, mutations in *microtubule associated protein tau* (*MAPT*) are a rare cause of familial FTD but mutation carriers can present with clinical symptoms indistinguishable from AD.[2-6] A common haplotype in *MAPT* has been associated with risk for several tauopathies, including AD.[7-11]

The only well established genetic risk factor for late onset Alzheimer disease (LOAD) is the ε4 allele of *Apolipoprotein E* (*APOE4*).[12] Recent genome-wide association studies have provided evidence for several novel risk genes of small effect size[13-15]. These studies, however, focused on the effect of common variants on risk for disease without addressing the role of rare variants. We have previously reported a *PSEN1* mutation in a family with a multigenerational history of LOAD, suggesting that causative mutations may be more common than previously anticipated.[16] We hypothesized that mutations in genes contributing to Mendelian forms of early onset AD and FTD might also be a cause of dementia in families multiply affected by late onset disease. To test this hypothesis, we identified 439 (30.7%) families with a history of dementia in four or more family members from a series of 1,431 families participating in the National Institute of Aging Late Onset Alzheimer Disease Study (NIA-LOAD Family Study). Using the next-generation pooled sequencing methodology,[17, 18] we screened these densely affected LOAD families for mutations in *APP, PSEN1, PSEN2, MAPT and GRN*.

METHODS

Patients and study Design

The NIA-LOAD Family Study recruited families with multiple members with LOAD from throughout the United States. We selected a subset of families for sequencing that met the following criteria; at least 4 family members reported with AD. We then narrowed this subset further by requiring at least 2 affected family members reporting an age of onset of 65 years or older. If there were not 2 affected family members with an age of onset over 65 then we required an average age of onset in the family of 60 years or older (n=439 families; Table 1).

We selected for sequencing the youngest affected family member with the most definitive diagnosis, choosing individuals with autopsy confirmed disease over those with probable or possible disease (NINCDS-ADRDA). Thus, in some families the age at onset for the sequenced sample was lower than 60 years, but the mean for the entire family was always equal to or higher than 60 years (Table 2). Among the individuals chosen for the sequencing, 26.9% met criteria for definite LOAD (autopsy confirmed), 71.4% probable and 1.5% possible LOAD (NINCDS-ADRDA).[19]

DNA sequencing and genotyping

We used the next-generation, pooled-DNA method described by Druley et al.,[17] to identify sequence variants in *APP, PSEN1, PSEN2, MAPT* and *GRN*. A more detailed description of this method including the quality control steps (Supplementary table 1) is provided in the supplementary materials. Sequence variants identified in the pooled DNA experiment were confirmed and the specific sample carrying the variant identified by direct genotyping using standards procedures. We genotyped the confirmed variants in all available family members to determine whether the sequence variant segregated with disease. We focused only on rare

(minor allele frequency <5%) non-synonymous, nonsense and splice-site variants. Common variants and synonymous variants were not followed up. A summary of the synonymous variants identified in this study is in supplementary table 2.

Bioinformatics

The AD and FTD mutation database (http://www.molgen.ua.ac.be/ADMutations/) was used to identify sequence variants found in previous studies of early-onset familial dementia and to determine whether or not they were considered to be pathogenic. The pilot data from 1000 genomes project (http://www.1000genomes.org/page.php) were used to compare the frequency of novel and rare (minor allele frequency less than 5%) non-synonymous, nonsense and splice site variants in each gene.

Statistical Analysis

Two-tailed t-tests and the Mann-Whitney U test were used to compare the age at onset and the number of affected individuals between the different groups. The effect of variants on AAO was analyzed, using a Cox proportional hazards model (proc PHREG, SAS), including gender in the model and family as an aggregation term.[20-25] Fisher's exact test was used to compare the frequency of sequence variants identified in the sequenced LOAD samples to those found in the 1000 genomes project.

RESULTS

We identified a total of 33 nonsense, missense and splice-site sequence variants in 60/439 (13.7%) individuals, including four known pathogenic variants in eight individuals (2.0 %), and six novel likely-pathogenic variants in twelve individuals (2.7%). An additional variant in *GRN*, reported previously as non-pathogenic, segregated perfectly with disease status and may be also pathogenic. Pathogenic or likely-pathogenic variants were found in all genes except *PSEN2*.

Pathogenic or likely pathogenic mutations.

We identified eleven sequence variants, two each in *APP* and *PSEN1*, three in *MAPT* and four in *GRN*, in a total of 22 families (5% of the total families) that are pathogenic or likely to be pathogenic, based on segregation data. Six of these variants are novel and five were previously reported (Table 4). The sequence variants showing perfect segregation were the previously reported pathogenic G206A mutation in *PSEN1*,[26, 27] the previously reported pathogenic *GRN* R110X[28, 29] and the *GRN* G515A variant. In each family, all of the genotyped affected individuals carry the sequence variants, but none of the unaffected individuals.

The *PSEN1* A79V mutation was found in four of the 439 sequenced samples (Table 3 and 4), including the most densely affected family (28 affected individuals, Table 1A). The sequenced individual from this family had a definitive diagnosis of LOAD and an age at onset of 76 years, which was similar to the mean age at onset in the entire family. Although A79V is a known pathogenic mutation, seen in families with early onset AD[27, 30], the mutation did not show perfect segregation with disease in this family (Figure 1A). One of the four affected individuals (age at onset 77 years) did not carry the mutation, representing a phenocopy (alternative causes of LOAD). In addition, three unaffected samples carry the mutation but with a mean age at last assessment more than ten years below the mean onset of disease in the family these likely represent presymptomatic individuals.

Two known pathogenic mutations in *GRN* were identified in this clinical series of LOAD families. Both mutations, R110X and R493X, create a premature termination codon resulting in nonsense-mediated mRNA decay and have been observed in FTD kindreds [29, 31]. The R493X

mutation, identified in a sample that came from a family with four affected individuals with a clinical diagnosis of LOAD, did not segregate with disease status (Table 4, Figure 1B). It was present in a single demented individual (age at onset =70), but was absent in another three demented individuals (age at onset 78.3 ± 1.15). The fact that they have a much higher age at onset and that they don't carry the mutation suggests that they are phenocopies. Four unaffected individuals are also mutation carriers. However, they are younger than the onset of the affected mutation carrier. As in the A79V family these individuals could be presymptomatic or could indicate reduced penetrance.

For the novel variants *APP* E559K and N660Y, *MAPT* V224G (exon 4A) G201S (exon 9) and A229T (exon 9), some of the variant carriers were asymptomatic at the time of the last assessment (Table 4), but in all cases, the age at last assessment of the unaffected carriers is lower than the oldest age at onset for the affected carriers. The E599K (*APP*) variant was found in two sequenced samples, one of which had a definite diagnosis of LOAD. Among the other family members, this variant was found in three affected samples (earliest age at onset=79 years). The two unaffected carriers were younger (mean age 70 ± 1 years) than the earliest age at onset of the affected samples. The *MAPT*-A229T variant was found in four families, two of them with five affected individuals and the other two families with four affected individuals. Globally, all the affected genotyped samples (six) carried the mutation. Three additional unaffected samples were also carriers, but they were younger than the mean onset for the affected samples (Table 4, figure 1C). This suggests that the unaffected carriers may be presymptomatic or these variants.

In the families with *APP* E599K and *GRN* c.1414-1G>T variants we also found a single affected individual per mutation who did not carry the sequence variant. For the *GRN* c.1414-1G>T, family the age at onset of the affected individual is higher than the mean onset for the affected individuals carrying the sequence variant, suggesting that this individual may be a phenocopy. The *GRN* splice-site variant c.1414-1G>T affects the last nucleotide of intron 10 within a core splice-site. This variant was found in two families, one of which had 20 affected individuals. Bioinformatic analyses indicate that this variant alters an acceptor splice-site (NetGene2 Prediction score for the wild type variant 1.00 vs. 0.00 for the mutant) causing exon skipping. A recent screen for *GRN* pathogenic mutations identified a similar variant (c.1414-2A>G) affecting the same core splice-site. *Ex vivo* splicing assays confirmed that the mutation c.1414-2A>G affects splicing of the exon.[28] The segregation data, our bioinformatic analyses and the bibliographic data strongly suggest that the c.1414-1G>T variant is likely to be pathogenic.

None of the sequenced individuals with pathogenic or likely pathogenic mutations in *GRN* or *MAPT* had neuropathologic confirmation of diagnosis.

Non-pathogenic or likely non-pathogenic variants.

We also identified a total of ten (five novel) sequence variants (15 families) which clearly did not segregate with disease suggesting that these are not pathogenic. Previously, some of these variants (*PSEN2* R71W and M174V, and *MAPT* R5H) have been classified as pathogenic in the AD and FTD mutation database; however the genetic and functional data has not been supportive.

In one family with eight affected and five genotyped individuals only one, of five affected individuals was a mutation carrier (*PSEN2 M174V*), but this individual had the earliest age at onset in that family (table 4). Two of them were younger than the carrier (ages 49 and 54 years), but the third mutation carrier was 79 years. Take together these results suggest that either the *PSEN2*-M174V variant has reduced penetrance, and the four other affected individuals

represent phenocopies, or alternatively, the mutation is not pathogenic.

The *PSEN2*-R71W variant was found in six different families, with between four and nine affected individuals, and it did not segregate with disease (Table 4). However the affected carriers presented a significantly lower age at onset than the non-carriers affected (70.2 vs 76.7 respectively; p=0.0005). The *PSEN2*-R62H variant has been previously identified in several familial AD kindreds[30], and also found to be relatively frequent among African populations[32]. We identified this variant in six of 439 sequenced samples, but it did not segregate with disease within families (table 4). Similarly to the *PSEN2*-R71W, the affected carriers for the *PSEN2*-R62H variant did have an earlier age at onset than the affected non-carriers (71 vs 75 years, respectively, p=0.0019; Table 4). These results suggest that these variants may be not pathogenic, but could be disease modifiers.

We also identified a family with a mutation in *MAPT*-R5H, a variant that has been classified as pathogenic by the AD and FTD mutation databases. This mutation was found in a Japanese family with eight affected individuals, but surprisingly it did not segregate with disease status. Only a single affected individual was a mutation carrier suggesting that this variant in *MAPT* may not be pathogenic.

Additional novel variants *APP*-G322A, *GRN*-P85A and T268M and *MAPT*-R168C (exon 4A) and A152T (exon 7) also did not segregate with disease and are likely to be non-pathogenic.

Variants with unknown pathogenicity

For the novel variants *APP* (G191E and V340M), *PSEN1* (P7L) *GRN* D135V, M207T and V519M and *MAPT* S427F (exon 6), segregation within families was inconclusive, possibly because of the small number of family members sampled. However, none of these variants were found in the 1000 genomes project. In fact, no missense variants have been found in *APP* in 185 individuals sequenced as part of the 1000 genomes pilot projects. The number of rare variants in these genes in this dataset was much higher than observed in the randomly selected individuals of the 1000 genomes project from three different populations (CEU, YRI and CHBJPT; $p=5.41x10^{-7}$; OR= 3.35; 95%CI=2.04-5.45) most likely reflecting the increased recurrence risk among densely affected families. It is possible therefore that many of these new variants could be pathogenic or be disease modifiers.

An additional of 67 common non-synonymous and synonymous variants were identified in these genes (supplementary table 2).

Comparison of the individuals and families with and without sequence variants

The age at onset of the sequenced individuals who carried variant alleles was significantly lower than that for sequenced individuals with normal gene sequence (67.8 years vs 70.5 years p=0.004), but there was no difference in the mean age at onset in families with compared to those without sequenced variants. The major distinguishing feature differentiating the families with sequence variants from those with normal gene sequence was the reported number of affected individuals per family. We found that the families with sequence variants had significantly more affected individuals than families with without sequence variants (average = 6.88 vs 5.62; p=0.0008; Supplementary Figures 1-3, Supplementary table 3).

DISCUSSION

This is the first study to systematically screen for pathogenic mutations and rare variants in *APP*, *PSEN1*, *PSEN2*, *MAPT* and *GRN* in a large series of well-characterized families densely affected by LOAD. Pathogenic mutations in these genes have previously been linked to familial early onset forms of dementia. We identified two previously known pathogenic mutations in

PSEN1,[26, 27, 30, 33] and two clearly pathogenic mutations in *GRN* (R110X, R493X).[6, 28, 31, 34-37] In some families the pathogenic mutations (*PSEN1*-A79V, and *GRN*-R493X, Figure 1) did not completely segregate with disease status illustrating how phenocopies, potential pre-symptomatic individuals and reduced penetrance complicate the interpretation of sequence variants in LOAD. While we can identify the phenocopies in families with known mutations, it is more challenging in families with novel variants, because phenocopies might imply that a variant does not segregate with disease. Putative pathogenic variants in genes that cause LOAD could have a less severe effect on protein function due to other genetic or environmental modifiers that are not present when these same variants result in early-onset disease. This would lead to the occurrence of reduced penetrance in LOAD families. Indeed, several cases of reduced penetrance have been reported in families with *PSEN1* and *GRN* mutations.[5, 34, 38-40] Even with these problems, our segregation data suggest that the novel *APP* N660Y, E599K, the *GRN* c.1414-1G>T and *MAPT* V224G, G201S and A229T variants are likely to be pathogenic variants.

The absence of consistent segregation of mutations with disease in this study suggests that the variants *PSEN2* variants R62H, R71W and M174V and *MAPT* R5H are not causative mutations.[30, 32, 41] Although, it is possible that they could represent risk factors or disease modifiers.

These results have very important clinical implications. We have demonstrated that pathogenic mutations in *APP*, *PSEN1*, *PSEN2*, *MAPT* and *GRN* can be present in individuals with any age at onset. Among the 439 sequenced samples, that represent families densely affected by disease (four or more affected individuals), 22 (5%) carried a pathogenic or likely-pathogenic mutation. A total of 60 individuals (13.7%) had a missense, splice-site or nonsense mutation in these genes (Table 3 and 4). The scarcity of available DNA samples from other members of families with LOAD made it difficult to determine whether or not eight novel variants identified were pathogenic. However the fact that we found a significant increase of rare variants in these genes compared with the 1000 genomes project ($p=5.41x10^{-7}$; OR= 3.35; 95%CI=2.04-5.45) and because it is known that these genes are involved in neurodegeneration we anticipate that some of these variants may be pathogenic. To confirm this hypothesis more genetic and functional analyses will be necessaries.

The recurrence risk in family members peaks with an age-at-onset of 85 years in the proband.[42, 43] Several families with more than 10 affected individuals have no mutations in the genes studied here. Thus it is very likely that we have underestimated the number of causative mutations in LOAD and that there are novel AD causative genes yet to be discovered. Worldwide there are a total of 865 early onset dementia families carrying pathogenic mutations in these genes reported in the AD and FTD mutation databases. Although the proportion of early onset patients with causal mutations is higher when compared to patients with familial LOAD, we estimate that numerically there may be many more patients with familial LOAD carrying causal pathogenic mutations. Approximately 10-20% of all patients with LOAD have a family history of dementia. Of those, approximately one-third have a sibling with disease and another one-third of these have a history of four or more affected family members. Thus, 0.054-0.30% of all patients with familial LOAD in the United States may carry a pathogenic mutation in these genes, which represents 2,862-15,900 cases/families. This estimation suggests that there are at least 3-18 times more LOAD families carrying pathogenic mutations in these genes than worldwide early onset families.

Another important finding in this study is the observation of clinically diagnosed AD families with previously reported and likely novel pathogenic mutations in *GRN* and *MAPT*. Mutations in

these genes are typically associated with frontotemporal lobar degeneration[29, 31] The result here indicates that families carrying mutations in *GRN* or *MAPT* can present with disease that is clinically indistinguishable from probable AD even in specialist memory disorder clinics, which is analogous to the observations that "AD mutations" can present with frontotemporal lobar degeneration[44-46]. Clinical syndromic diagnoses obviously reflect the presence/absence of particular features, which in turn reflects the underlying topography of degeneration. The fact that certain *GRN* and *MAPT* mutations cause an AD-like phenotype suggests that certain mutations may have some degree of topographic specificity (ie, mesial temporal lobe degeneration greater than inferolateral and frontal degeneration), or that other factors influence the topography of degeneration, or both.

None of the individuals sequenced in this series and found to have *GRN* and *MAPT* variants were autopsy confirmed as LOAD. Therefore, pathologic confirmation will be necessary to establish the ultimate diagnoses. In this series the frequency of pathogenic mutations and rare sequence variants in *GRN* and *MAPT* is similar to the frequency of variants in the AD-related genes (*APP, PSEN1* and *2*). Identification of families and individuals carrying mutations in genes associated with frontotemporal lobar degeneration will be important for clinical management of these patients, particularly as therapies are developed that target the specific pathophysiologic processes of these disorders.

The results here highlight the importance of genetic evaluation and the importance of counseling in clinical practice. This work provides some general guidelines to identify individuals and families that should be prioritized for mutation screening. Although we did not find clear differences between the families carrying missense, splice-site or nonsense variants compared with the families without novel variants, the number of reported affected individuals and the presence of at least one affected individual with an early age at onset were strong indicators of the presence of a pathogenic variant. It also indicates that *MAPT* and *GRN*, genes typically associated with frontotemporal lobar degeneration should also be included when no autopsy is available. Finally, this work shows that the age at onset of disease is highly variable even in individuals with the same mutation and that reduced penetrance and phenocopies are likely to be common illustrating the importance of genetic counseling in families with late onset AD.

Figure legends

Figure 1: Pedigrees for some of the sequenced families illustrating the presence of phenocopies and low penetrance mutations or the presence of presymptomatic cases. A) Pedigree for a family with the PSEN1 A79V mutation. B) Pedigree for a family with the GRN R493X mutation. C) Pedigree for a family with the MAPT A229 mutation. AO indicates the subject or family report of age of onset of symptoms. AE the age of last evaluation. MCI: Mild cognitive impairment or questionable dementia by family report. + symbol indicates that the subject is positive for the indicated mutation. - symbol indicates that the subject is negative for the indicated mutation. A number inside of a diamond indicates the number of subjects have the same status. Fully shaded circles or squares indicate confirmed AD by autopsy. Three/fourth shaded symbol indicates probable AD diagnosed using NINCDS-ADRDA criteria. One/fourth shaded symbol indicates that the family reports this individual has AD.

ACKNOWLEDGEMENTS

This study was supported by federal grants U24 AG026395 (The NIA-LOAD Family Study) and R37AG15473 (RM). Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors, including the Alzheimer's Disease Centers who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

This work was supported by grants from NIH (AG16208, P01AG03991, P50AG05681, P01AG026276), the Barnes-Jewish Hospital Foundation, the Department of Veterans Affairs, and an anonymous foundation~~,~~ the Barnes-Jewish Hospital Foundation, the Department of Veterans Affairs research funds. CC has a fellowship from "Fundacion Alfonso Martin Escudero".

REFERENCES

1. Cortini F, Fenoglio C, Guidi I, et al. Novel exon 1 progranulin gene variant in Alzheimer's disease. Eur J Neurol 2008;15:1111-7.
2. Ludolph AC, Kassubek J, Landwehrmeyer BG, et al. Tauopathies with parkinsonism: clinical spectrum, neuropathologic basis, biological markers, and treatment options. Eur J Neurol 2009;16:297-309.
3. Reed LA, Grabowski TJ, Schmidt ML, et al. Autosomal dominant dementia with widespread neurofibrillary tangles. Ann Neurol 1997;42:564-72.
4. Kelley BJ, Haidar W, Boeve BF, et al. Alzheimer disease-like phenotype associated with the c.154delA mutation in progranulin. Arch Neurol 2010;67:171-7.
5. Foster NL, Heidebrink JL, Clark CM, et al. FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. Brain 2007;130:2616-35.
6. Rademakers R, Baker M, Gass J, et al. Phenotypic variability associated with progranulin haploinsufficiency in patients with the common 1477C-->T (Arg493X) mutation: an international initiative. Lancet Neurol 2007;6:857-68.
7. Myers AJ, Kaleem M, Marlowe L, et al. The H1c haplotype at the MAPT locus is associated with Alzheimer's disease. Hum Mol Genet 2005;14:2399-404.
8. Kauwe JS, Cruchaga C, Mayo K, et al. Variation in MAPT is associated with cerebrospinal fluid tau levels in the presence of amyloid-beta deposition. Proc Natl Acad Sci U S A 2008;105:8050-4.
9. Skipper L, Wilkes K, Toft M, et al. Linkage disequilibrium and association of MAPT H1 in Parkinson disease. Am J Hum Genet 2004;75:669-77.
10. Zabetian CP, Hutter CM, Factor SA, et al. Association analysis of MAPT H1 haplotype and subhaplotypes in Parkinson's disease. Ann Neurol 2007;62:137-44.
11. Pastor P, Ezquerra M, Tolosa E, et al. Further extension of the H1 haplotype associated with progressive supranuclear palsy. Mov Disord 2002;17:550-6.
12. Bertram L, McQueen M, Mullin K, Blacker D, Tanzi R. The AlzGene Database. Alzheimer Research Forum. Available at: http://wwwalzgeneorg Accessed 12/05/2005.
13. Lambert JC, Heath S, Even G, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 2009;41:1094-9.
14. Seshadri S, Fitzpatrick AL, Ikram MA, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. Jama 2010;303:1832-40.
15. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 2009;41:1088-93.
16. Kauwe JS, Jacquart S, Chakraverty S, et al. Extreme cerebrospinal fluid amyloid beta levels identify family with late-onset Alzheimer's disease presenilin 1 mutation. Ann Neurol 2007;61:446-53.
17. Druley TE, Vallania FL, Wegner DJ, et al. Quantification of rare allelic variants from pooled genomic DNA. Nat Methods 2009;6:263-5.
18. Vallania F. SPLINTER. Genome Research 2010 2010.
19. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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 1984;34:939-44.
20. Zhao H. Family-based association studies. Stat Methods Med Res 2000;9:563-87.
21. Blacker D, Wilcox MA, Laird NM, et al. Alpha-2 macroglobulin is genetically associated with Alzheimer disease. Nat Genet 1998;19:357-60.
22. Siegmund KD, Langholz B, Kraft P, Thomas DC. Testing linkage disequilibrium in sibships. Am J Hum Genet 2000;67:244-8.
23. Kraft P, Thomas DC. Case-sibling gene-association studies for diseases with variable age at onset. Stat Med 2004;23:3697-712.
24. Romas SN, Mayeux R, Tang MX, et al. No association between a presenilin 1 polymorphism and Alzheimer disease. Arch Neurol-Chicago 2000;57:699-702.
25. Cruchaga C, Kauwe JS, Mayo K, et al. SNPs Associated with Cerebrospinal Fluid Phospho-Tau

Levels Influence Rate of Decline in Alzheimer's Disease. PLoS Genet 2010;6.
26. Athan ES, Williamson J, Ciappa A, et al. A founder mutation in presenilin 1 causing early-onset Alzheimer disease in unrelated Caribbean Hispanic families. Jama 2001;286:2257-63.
27. Rogaeva EA, Fafel KC, Song YQ, et al. Screening for PS1 mutations in a referral-based series of AD cases: 21 novel mutations. Neurology 2001;57:621-5.
28. Yu CE, Bird TD, Bekris LM, et al. The spectrum of mutations in progranulin: a collaborative study screening 545 cases of neurodegeneration. Arch Neurol 2010;67:161-70.
29. Van Deerlin VM, Wood EM, Moore P, et al. Clinical, genetic, and pathologic characteristics of patients with frontotemporal dementia and progranulin mutations. Arch Neurol 2007;64:1148-53.
30. Cruts M, van Duijn CM, Backhovens H, et al. Estimation of the genetic contribution of presenilin-1 and -2 mutations in a population-based study of presenile Alzheimer disease. Hum Mol Genet 1998;7:43-51.
31. Huey ED, Grafman J, Wassermann EM, et al. Characteristics of frontotemporal dementia patients with a Progranulin mutation. Ann Neurol 2006;60:374-80.
32. Guerreiro RJ, Baquero M, Blesa R, et al. Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. Neurobiol Aging 2010;31:725-31.
33. Brickell KL, Leverenz JB, Steinbart EJ, et al. Clinicopathological concordance and discordance in three monozygotic twin pairs with familial Alzheimer's disease. J Neurol Neurosurg Psychiatry 2007;78:1050-5.
34. Gass J, Cannon A, Mackenzie IR, et al. Mutations in progranulin are a major cause of ubiquitin-positive frontotemporal lobar degeneration. Hum Mol Genet 2006;15:2988-3001.
35. Pickering-Brown SM, Baker M, Gass J, et al. Mutations in progranulin explain atypical phenotypes with variants in MAPT. Brain 2006;129:3124-6.
36. Mesulam M, Johnson N, Krefft TA, et al. Progranulin mutations in primary progressive aphasia: the PPA1 and PPA3 families. Arch Neurol 2007;64:43-7.
37. Spina S, Murrell JR, Huey ED, et al. Clinicopathologic features of frontotemporal dementia with Progranulin sequence variation. Neurology 2007.
38. Rossor MN, Fox NC, Beck J, Campbell TC, Collinge J. Incomplete penetrance of familial Alzheimer's disease in a pedigree with a novel presenilin-1 gene mutation. Lancet 1996;347:1560.
39. Llado A, Fortea J, Ojea T, et al. A novel PSEN1 mutation (K239N) associated with Alzheimer's disease with wide range age of onset and slow progression. Eur J Neurol 2010;17:994-6.
40. Le Ber I, van der Zee J, Hannequin D, et al. Progranulin null mutations in both sporadic and familial frontotemporal dementia. Hum Mutat 2007;28:846-55.
41. Hayashi S, Toyoshima Y, Hasegawa M, et al. Late-onset frontotemporal dementia with a novel exon 1 (Arg5His) tau gene mutation. Ann Neurol 2002;51:525-30.
42. Silverman JM, Smith CJ, Marin DB, Mohs RC, Propper CB. Familial patterns of risk in very late-onset Alzheimer disease. Arch Gen Psychiatry 2003;60:190-7.
43. Silverman JM, Ciresi G, Smith CJ, Marin DB, Schnaider-Beeri M. Variability of familial risk of Alzheimer disease across the late life span. Arch Gen Psychiatry 2005;62:565-73.
44. Mendez MF, McMurtray A. Frontotemporal dementia-like phenotypes associated with presenilin-1 mutations. Am J Alzheimers Dis Other Demen 2006;21:281-6.
45. Bernardi L, Tomaino C, Anfossi M, et al. Novel PSEN1 and PGRN mutations in early-onset familial frontotemporal dementia. Neurobiol Aging 2009;30:1825-33.
46. Marcon G, Di Fede G, Giaccone G, et al. A novel Italian presenilin 2 gene mutation with prevalent behavioral phenotype. J Alzheimers Dis 2009;16:509-11.

TABLES

Table 1: Distribution of the number of families by the number of affected individuals in each family.

# affected	# families	%	# sequence variants	%
28	1	0.23	1	1.67
20	2	0.46	2	3.33
19	1	0.23	0	0.00
17	1	0.23	2	3.33
16	1	0.23	0	0.00
15	1	0.23	0	0.00
14	1	0.23	0	0.00
13	5	1.14	1	1.67
12	3	0.68	0	0.00
11	4	0.91	0	0.00
10	13	2.96	2	3.33
9	14	3.19	2	3.33
8	27	6.15	5	8.33
7	32	7.29	6	10.00
6	54	12.30	5	8.33
5	107	24.37	9	15.00
4	172	39.18	25	41.67
TOTAL	439		60	

The number of families with different numbers of affected individuals is shown. More than 50% of the families have 4 or 5 affected individuals. The number of sequence variants found in each group is also shown.

Table 2: Demographic data for the cohort

AAO (years)	Sequenced sample	69.92 ± 8.37 (range 30-92)
	Families (mean age at onset)	72.80 ± 5.62 (range 60-89)
Ethnicity	European ancestry	76.5%
	White- Hispanics	14.6%
	Black- Hispanics	2.7%
	African-American	2.7%
	others	3.6%
Diagnosis	Definite AD	26.94%
	Probable AD	71.46%
	Possible AD	1.59%

The mean, the standard deviation and the range for the age at onset (AAO) for the sequenced samples and the families are shown. For the families the AAO represents the mean AAO of the affected individuals of each family

Table 3: Number of sequence variants found in the 439 sequenced samples.

	Previously reported						New**							
	Pathogenic*		Path. Unknown*		Not Path.*		Likely Pathogenic*		Path. Unknown*		Not Path.*		total	
	#mut	# fam	#mut	# fam	#mut	# fam	#mut	# fam	#mut	# fam	#mut	# fam	#mut	# fam
APP	0	0	0	0	1	1	2	3	2	2	1	1	6	7
PSEN1	2	7	0	0	0	0	0	0	1	1	0	0	3	8
PSEN2	0	0	3	13	0	0	0	0	0	0	0	0	3	13
MAPT	0	0	1	1	0	0	3	7	1	1	2	6	7	15
GRN	3	3	1	1	4	4	1	2	3	4	2	3	14	17
Total	5	10	5	15	5	5	6	12	7	8	5	10	33	60
		2.3%		3.4%		1.1%		2.7%		1.8%		2.3%		13.7%

- *Based on our segregation analyses

Only the rare (MAF<0.05) non-synonymous, splice-site and nonsense sequence variants are shown.

Table 4: List of all the rare non-synonymous, splice site and nonsense sequence variants identified.

Gene	Change	# of fam.	Status	Affected		Unaffected	
				Carriers	Non-carriers	Carriers	Non-carriers
Pathogenic or likely pathogenic							
APP	E599K	2	Novel	4 82.2 ± 3.2	1 83	2 70 ± 1	6 66.83 ± 5.3
	N660Y	1	Novel	3 61.6 ± 3	0 -	1 62	1 69
PSEN1	A79V	4	Previously reported Pathogenic	10 68.9 ± 8.5	1 77	3 60 ± 4.3	13 70.62 ± 12.95
	G206A	3	Previously reported Pathogenic	6 60.8 ± 6.7	0 -	0 -	1 40
MAPT	V224G	2	Novel	6 71.6 ±11.3	0 -	1 62	2 76 ± 9.9.
	G201S	1	Novel	2 74 ± 0	0 -	2 61 ±5.7	2 49 ± 4.2
	A229T	4	Novel	6 73.6 ±7.2	0 -	3 70.3 ± 1.15	4 70.5 ± 9.26
GRN	R110X	1	Previously reported Pathogenic	3 66.3 ± 2.1	0 -	0 -	5 72.2 ± 5.26
	c.1414-1G>T	2	Novel	8 65.8 ± 11.7	1 74	0 -	1 70
	R493X	1	Previously reported Pathogenic	1 70	3 78.3 ± 1.15	4 56.3 ± 4.04	3 56.3 ± 11.06
	G515A	1	Previously reported Not pathogenic	3 80± 7.5	0 -	0 -	1 66
Non-pathogenic or likely non-pathogenic							
APP	G322A	1	Novel	4 68.2 ±18	7 75.3 ± 10	1 56	2 65 ± 17
	A673T	1	Previously reported Not pathogenic	1 67	2 84.2 ± 6.3	0 -	1 72
PSEN2	R62H	6	Previously reported Unclear	12 71.1 ± 7.7	8 75.0 ± 8.4	8 59.5 ± 16.6	35 61.7 ± 13.47
	R71W	6	Previously reported Pathogenic	6 70 ± 5.07	8 76.7 ± 7.3	3 67 ± 14.52	10 71.45 ± 11.85
	M174V	1	Previously reported Pathogenic	1 60	4 67.5 ± 1.2	3 60.7 ± 16	17 50.5 ± 14

MAPT	R5H	1	Previously reported Pathogenic	1 85	1 80	4 77 ± 7.4	1 72
	R168C	1	Novel	1 63	4 75.5 ± 6.8	3 49.3 ± 17	28 58.5 ±11.8
	A152T	5	Novel	8 72.9 ± 10	8 72.0 ± 7.44	2 52.5 ± 13.4	16 62.4 ± 14.7
GRN	P85A	2	Novel	3 72 ± 11	3 75 ± 11	0	1 59
	V141I	1	Previously reported Not pathogenic	2 81.5 ±2	1 79	1 55	2 69 ±19.8
	T268M	1	Novel	1 64	2 69 ± 1.41	5 64.8 ± 14.4	7 62.1 ± 13
	A324T	1	Previously reported Not pathogenic	2 67.0 ±5	1 69	2 60 ± 2.8	0 -
	D376N	1	Previously reported No pathogenic	1 63	0 -	0 -	0 -
	R433Q	1	Previously reported Not pathogenic	1 65	0 -	0 -	0 -
Unknown							
APP	G191E	1	Novel	1 84	0 -	0 -	0 -
	V340M	1	Novel	2 66.0 ± 5	1 66	1 67	1 80
PSEN1	P7L	1	Novel	2 72.5 ± 3.5	1 79	0 -	5 57 ± 14.3
MAPT	S427F	1	Novel	2 76.5 ± 12	1 77	1 72	0 -
GRN	D135V	1	Novel	1 57	0 -	0 -	0 -
	M207T	1	Novel	1 66	1 70	0 -	1 76
	V514M	1	Novel	1 72	0 -	0 -	0 -
	V519M	2	Novel	2 74.5 ± 7	0 -	1 73	0 -

List of the non-synonymous, splice and nonsense variants identified in the 439 sequences samples. The identified variants were genotyped in all the available family samples. The number of affected carriers, non-carries and the un-affected carriers, non-carries, as well as the mean age at onset and the standard deviation for the affected and the age at the last assessment for the unaffected individuals are shown.

The variants were classified as pathogenic, non-pathogenic or unknown based on our segregation analyses and previous reports