Next Generation Sequencing analysis in early onset dementia patients

Cristian Bonvicini\textsuperscript{a*,b}, Catia Scassellati\textsuperscript{a*}, Luisa Benussi\textsuperscript{a}, Emilio Di Maria\textsuperscript{b} Carlo Maj\textsuperscript{c,d}, Miriam Ciani\textsuperscript{a,e}, Silvia Fostinelli\textsuperscript{i}, Anna Mega\textsuperscript{l}, Martina Bocchetta\textsuperscript{f,g}, Gaetana Lanzi\textsuperscript{h}, Edoardo Giacopuzzi\textsuperscript{c,e}, Sergio Ferraboli\textsuperscript{e}, Michela Pievani\textsuperscript{f}, Virginia Fedi\textsuperscript{i}, Carlo Alberto Defanti\textsuperscript{l}, Silvia Giliani\textsuperscript{h}, Alzheimer’s Disease Neuroimaging Initiative*, Giovanni Battista Frisoni\textsuperscript{f,m}, Roberta Ghidoni\textsuperscript{a}, Massimo Gennarelli\textsuperscript{c,e}

\textsuperscript{a}Molecular Markers Laboratory, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
\textsuperscript{b}Department of Health Sciences, University of Genova and Division of Medical Genetics, Galliera Hospital, Genova, Italy
\textsuperscript{c}Genetics Unit, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
\textsuperscript{d}Institute for Genomic Statistics and Bioinformatics, Bonn, Germany
\textsuperscript{e}Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy
\textsuperscript{f}Laboratory Alzheimer's Neuroimaging and Epidemiology, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
\textsuperscript{g}Dementia Research Centre, Department of Neurodegenerative Disease, Institute of Neurology, University College London, London, United Kingdom
\textsuperscript{h}A. Nocivelli” Institute for Molecular Medicine Spedali Civili and University of Brescia, Brescia Italy
\textsuperscript{i}Fondazione Case Serena, Pontoglio, Brescia, Italy
\textsuperscript{m}Fondazione Europea Ricerca Biomedica, Centro di Eccellenza Alzheimer, Ospedale Briolini Gazzaniga, Bergamo, Italy.
\textsuperscript{m}Laboratory of Neuroimaging of Aging (LANVIE), University Hospitals and University of Geneva, Geneva, Switzerland; Department of Internal Medicine, University Hospitals and University of Geneva, Geneva, Switzerland.
\textsuperscript{*}Authors have contributed equally to this work.
Running title: NGS and early onset dementia
Abstract

Background: Early onset dementias (EOD) are rare neurodegenerative dementias that present before 65 years. Genetic factors have a substantially higher pathogenetic contribution in EOD patients than in late onset dementia.

Objective: To identify known and/or novel rare variants in major candidate genes associated to EOD by high-throughput sequencing. Common-risk variants of apolipoprotein E (APOE) and prion protein (PRNP) genes were also assessed.

Methods: We studied 22 EOD patients recruited in Memory Clinics, in the context of studies investigating genetic forms of dementia. Two methodological approaches were applied for the target-Next Generation Sequencing (NGS) analysis of these patients. In addition, we performed progranulin plasma dosage, C9Orf72 hexanucleotide repeat expansion analysis and APOE genotyping.

Results: We detected three rare known pathogenic mutations in the GRN and PSEN2 genes and eleven unknown-impact mutations in the GRN, VCP, MAPT, FUS, TREM2, NOTCH3 genes. Six patients were carriers of only common risk variants (APOE and PRNP), and one did not show any risk mutation/variant. Overall, 69% (n=9) of our EAOD patients, compared with 34% (n=13) of sporadic late-onset Alzheimer Disease (AD) patients and 27% (n=73) of non-affected controls (ADNI, whole genome data), were carriers of at least two rare/common risk variants in the analyzed candidate genes panel, excluding the full penetrant mutations.

Conclusion: Our results show that EOD patients are characterized by polygenic mutations/risk alleles and that the genetic load without full penetrant mutations stays on higher than that observed in late onset AD forms. Thus, we recommend the screening of all causative genes in larger EOD cohorts.

This study suggests that EOD patients without full penetrant mutations are characterized by higher probability to carry polygenic risk alleles that patients with LOAD forms. This finding is in line with recently reported evidence, thus suggesting that the genetic risk factors identified in LOAD might modulate the risk also in EOAD.
Keywords

Alzheimer’s disease; Frontotemporal dementia; Lewy body dementia; early onset dementia; Next Generation Sequencing; rare mutations; common variants
INTRODUCTION

The term early onset dementias (EOD) refers to a group of progressive neurodegenerative diseases (e.g. Alzheimer’s disease (AD), frontotemporal dementia (FTD) or dementia with Lewy Body (LBD), affecting individuals aged between 45 and 65 years and it represents roughly 5% of dementia cases [1]. The symptoms of EOD are similar to those of late-onset AD and FTD. However, EOD is thought to be more severe and typically causes a rapid decline in health [2,3].

Both AD and FTD are pathologically heterogeneous disorders, characterized by a complex genetic architecture that is not yet completely understood. The heritability rates of the different dementia subtypes range from 40 to 80% with EOD showing a higher genetic component than late-onset dementia (for review [4]).

AD is clinically characterized by memory impairment and pathologically by the presence of amyloid β peptide (the precursor of which is encoded by the APP gene) plaques and intraneuronal tangles of hyperphosphorylated forms of tau (a microtubule-associated protein encoded by the MAPT gene). The risk AD spectrum is composed of Mendelian genetic traits, genetic population risk factors (susceptibility genes), and nongenetic risk factors such as low cognitive reserve and head trauma [5,6]. The apolipoprotein E gene (APOE) ε4 allele is a known population risk factor [7] that has been found to increase the risk of early onset Alzheimer’s disease (EOAD) [8]. Since its discovery, over 550 susceptibility genes have been suggested to increase the risk of AD [9], though the impact of most of these genes seems to be much lower than that of APOE [10,11]. In particular, the common variants with small individual effects jointly modify the risk and age at onset of AD and dementia, showing a stronger effect in carriers homozygous for APOE ε4 [12].

Three genes have been identified to carry causative mutations for familial EOAD: amyloid precursor protein (APP), presenilin 1 (PSEN1), and presenilin 2 (PSEN2) (for review [11]). The estimated mutation frequencies of these three genes are 1% for APP, 6% for PSEN1 and 1% for PSEN2. Together, they explain a genetic background of only 5-10% of EOAD patients, leaving a large group of autosomal dominant pedigrees genetically unexplained (for review [13]). This finding suggests that additional causal genes remain to be identified.
FTD is characterized by personality changes, language impairment, and deficits of executive functions associated with frontal and temporal lobe degeneration. At least nine autosomal dominant genetic traits have been associated with this pathology: mutations in MAPT, in the progranulin gene (GRN), and in the hexanucleotide repeat expansion C9orf72 genes are the most common, with the highest prevalence of GRN mutations found in populations of northern Italy [14-17]. GRN null mutations cause protein haploinsufficiency, leading to a significant decrease in the circulating progranulin levels in plasma, serum and cerebrospinal fluid (CSF) of mutation carriers [18-20]. Mutations in valosin-containing protein (VCP), TAR DNA-binding protein 43 (TARDBP), charged multivesicular body protein 2B (CHMP2B), fused in sarcoma (FUS), dynactin (DCTN1) and triggering receptor expressed on myeloid cell (TREM2) are rarer causes of this pathology [4,21]. Mutations in VCP [22], TARDBP [23], and TREM2 [24] have been observed in Italian families with a history of FTD.

Interestingly, mutations in some of these genes, such as MAPT, GRN and C9orf72 have also been detected at low frequencies in AD patients, supporting the notion that a genetic heterogeneity exists for these diseases and that both diseases could form an AD-FTD disease continuum (for review [13]). An AD-like phenotype has also been described with the presence of a nonsense mutation in the prion protein gene (PRNP p.Q160*), which is responsible for inherited neurodegenerative spongiform encephalopathies [25]. In addition, the common coding polymorphism, methionine (M) to valine (V) at position 129 (M129V) in PRNP has been associated with EOAD, where the risk is higher for the VV genotype and is increased in patients with a positive family history [26].

The recent development of extremely powerful, massively, parallel DNA sequencing technologies allows for the systematic screening of individual genomes for DNA sequence variations at base-pair resolution, enabling researchers to address the missing hereditability question and, thus, to uncover novel and/or potentially pathogenic rare variants in candidate genes. As previously documented [27-29], targeted re-sequencing of a clinically significant gene panel may represent a powerful and cost-time-effective technique compared to the previously used sequential Sanger sequencing.
On this basis, we aimed to identify known and/or novel rare variants in candidate genes using next
generation sequencing (NGS) and to evaluate the contribution of the common-risk variants in the APOE and
PRNP genes in a selected cohort of Italian EOD patients.

Recently, Cruchaga C et al. [30], confirmed that the genetic factors identified in LOAD modulate the risk
also in EOAD cohorts, where the burden of these risk variants is associated with familial clustering and
earlier onset of AD. In the present study, we estimated the genetic load in EOAD and LOAD, by identifying
known and novel, both rare and common risk variants in candidate genes. We applied next generation
sequencing (NGS) analysis in a selected retrospective cohort of Italian EOD patients and compared the
frequencies of variants found with those estimated in samples from the ADNI database.

MATERIALS AND METHODS

Participants

A retrospective sample of patients were recruited in the context of studies investigating genetic forms of
dementia at IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia Italy, Fondazione Case
Serena, Pontoglio, Brescia, Italy and Fondazione Europea Ricerca Biomedica, Centro di Eccellenza
Alzheimer, Ospedale Briolini Gazzaniga, Bergamo, Italy. Specifically, twenty-two patients fulfilled the
following inclusion criteria for the present study: i) phenotype of AD, FTD, or LBD and ii) early disease
onset (<65 years old), or iii) family history suggestive of an autosomal dominant genetic form of dementia
(i.e., high or medium risk of identifying a mutation according to Loy and Woods criteria [31,32], as
described below). Family history was collected through interviews with a first-degree relative or the spouse
of the proband. The clinical and medical history of each family member was collected, and all of the
available documentation for affected members was acquired. The probability of identifying a genetic
mutation for AD or FTD was estimated considering the family medical history, the number of first and
second-degree affected family members and the age of symptom onset, according to the criteria developed
by Loy and colleagues [31]. According to Loy et al.’s criteria for AD, we defined a probability of
identifying a genetic mutation of ≥86% as a high risk, a probability of 68-85% as a medium risk, a
probability of 15-67% as a low risk, and a probability <15% as apparently sporadic/unknown significance.
Considering the same criteria for FTD patients, we considered an $\geq 88\%$ probability of identifying a genetic mutation as a high risk, a 31-41% probability as a medium risk, and a probability <13% as a low risk. FTD pedigrees were also scored according to Wood’s pedigrees classifications criteria [32,33].

All participants were of Italian ancestry. Demographic features and clinical data (age at onset, MMSE) are reported in Table 1. As the sample was retrospectively pooled for the analysis, a standard protocol for biomarker characterization was not applied. When available, information on disease biomarkers status was included. Sixteen patients underwent one of the following examinations as part of their diagnostic exam: magnetic resonance imaging (MRI), positron emission tomography (FDG-PET) or Single Photon Emission Computed Tomography (SPECT), and/or lumbar puncture. MRI and PET/SPECT scans were visually evaluated to determine medial-temporal atrophy and hypometabolism, respectively. CSF samples were processed with local procedures to determine the level of Ab, tau and p-tau. Positive diagnosis was determined based on established cut-offs.

Blood samples were collected from all patients. DNA and plasma were obtained according to standard procedures. Patients provided written informed consent. This study was approved by the local ethics committee (CEIOC, 62/2013).

**NGS panel analysis screening**

Genomic DNA was extracted from whole-blood samples with a commercially available kit according to standard procedures (GENTRA Minneapolis, MN, USA).

Due to logistics issues, some samples were analysed through the use of the Ion Torrent PGM (Thermo Fisher Scientific, Waltham, MA USA) sequencer as NGS platform, by using a candidates genes panel, already described in Beck et al [27]. Briefly, for library construction, 5 ng of genomic DNA were amplified using the Ion Ampliseq Dementia Research gene panel (Ampliseq™, Thermo Fisher Scientific, Waltham, MA USA), and the Ion Ampliseq™ Library kit 2.0, according to manufacturer’s instructions. The generated amplicon library includes PRNP, PSEN1, PSEN2, APP (Amyloid Beta A4 Precursor Protein), GRN, MAPT, TREM2, CHMP2B, CSF1R (Colony Stimulating Factor 1 Receptor), FUS, ITM2B (Integral Membrane Protein 2B), NOTCH3 (Notch 3), SERPINI1 (Serpine Peptidase Inhibitor, Clade I (Neuroserpin), Member 1), TARDBP, TYROBP (TYRO Protein Tyrosine Kinase Binding Protein), VCP, SQSTM1 (Sequestosome 1).
Amplicons were ligated to Ion Torrent Barcodes/adapters P1 using DNA ligase. A first step of Agencourt AMPure XP bead (Beckman Coulter Inc., Brea CA, USA) purification was followed by nick-translation of adapter-ligated products and PCR-amplification. A second purification step using AMPure beads was performed and the concentration and size of the libraries were determined using an Agilent BioAnalyzer DNA High-sensitivity LabChip (Agilent Technologies, Santa Clara, CA USA). After dilution to 100 pM, libraries were clonally amplified on Ion sphere™ particles (ISP) by emulsion PCR with the Ion PGM™ template OT2 200 kit on the IonOne Touch 2 instrument according to the manufacturer’s instructions (Thermo Fisher Scientific, Waltham, MA USA). ISP were enriched using the Ion One Touch ES module, loaded on an Ion 314 chip kit V2 and sequenced with an Ion Torrent PGM System (Thermo Fisher Scientific, Waltham, MA USA).

The remaining samples were analysed with the Illumina MiSeq platform and the TruSight One Sequencing Panel (Illumina, Inc., San Diego, CA, USA). This panel includes 125,395 probes targeting a 12-Mb region spanning 4,813 genes, among which the genes investigated by Ion Torrent PGM (PRNP, PSEN1, PSEN2, APP, GRN, MAPT, TREM2, CHMP2B, CSF1R, FUS, ITM2B, NOTCH3, SERPINI1, TARDBP, TYROBP, VCP, SQSTM1). The data regarding the other all genes were not used. The obtained sequence reads were aligned to the hg19 human reference sequence using the Burrow–Wheeler Aligner (BWA version 0.7.12). Duplicated reads were removed with Picard tools. Local realignment, recalibration, and variant calling were conducted with the Genome Analysis Tool Kit (GATK version 3.30). In order to have comparable results between the two sequencing approaches, we extracted from TruSight One Variant Call Format file (i.e., VCF file), the variants located in the regions sequenced by Ion Torrent PGM panel using BEDTools [34].

**APOE genotyping**

Genetic variation at the APOE locus was determined by using the SNaPshot technique [35]. Briefly, assays for the APOE polymorphisms were performed using PCR reactions, which were subsequently combined to perform a single SNaPshot reaction. The amplification assay was designed with the following forward and reverse primers: APOE F: 5’ CCAAGGAGCTGCAGGCGGCGCA 3’ and APOE R: 5’ GCCCCGGCCTGTTACACTGCCA 3’. A product of PCR-amplification was used as a template in the SNaPshot Multiplex assay. The following specific primers were used: SNAP APOE112: 5’
ACTGCACCAGGCGGCCGC 3’ and SNAP APOE158: 5’ATGCCGATGACCTGCAGAAG 3’. Finally, the samples were analysed, and allele peaks were determined using the ABI 3130xl genetic analyzer and the GeneMapper 4.0 program (Applied Biosystems, Foster City, CA, USA).

*C9Orf72 hexanucleotide repeat expansion*

PCR sizing of the GGGGCC hexanucleotide repeat was performed using previously published primers [36] on the ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA, USA). The PCR reaction was carried out in a mixture containing 5% dimethylsulfoxide and 7-deaza-2-deoxy GTP in substitution for dGTP. Allele identification and scoring were performed using GeneMapper v4.0 software (Applied Biosystems).

*GRN plasma level measurement*

Plasma progranulin levels were measured in duplicate using an ELISA kit (Human Progranulin ELISA Kit, AdipoGen Inc., Seoul, Korea).

*Statistical and Bioinformatics analyses*

To classify a variant as rare, its frequency should be lower than 1% in at least one of the three reference databases (1000 Genomes Project [http://www.internationalgenome.org/], Exome Sequencing Project [http://evs.gs.washington.edu/EVS/] and Exome Aggregation Consortium [http://exac.broadinstitute.org]) [37].

In order to predict the functional consequences of non-synonymous variations, we exploited eight different bioinformatics tools, namely: SIFT, PolyPhen-2, FATHMM, phyloP, MutationTaster, LRT, and CADD and GERP++ [38-45]. A variant is classified as damaging if for at least three tools the mutation is predicted to be deleterious.

Finally, to evaluate the mutation rate of the candidate genes selected in the NGS panel, we considered the gene damage index (GDI, a genome-wide, gene-level metric of the mutational damage that has accumulated in the general population), according to Itan et al. [46].

*ADNI whole genome data:*

As a genetic replication cohort, we considered whole genome data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). From the whole genome data, we extracted
the variants within the regions included in our sequencing panel and we applied the same variants annotation and classification performed for our sample.

The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of Mild Cognitive Impairment (MCI) and early AD. The aim of ADNI project is to collect, validate and utilize heterogeneous clinical and biological data (including MRI and PET images, genetics, cognitive tests, CSF and blood biomarkers) to study the progression of AD. For up-to-date information see www.adni-info.org.

RESULTS

Target screening: Plasma progranulin and C9ORF72 analyses

As a first step, progranulin plasma levels were assayed to screen for GRN null mutations (Table 1). One FTD patient was found to have progranulin plasma levels lower than the optimized cut-off value for null mutations detection of 61.55 ng/ml [19,47]. For 5 samples, it was no possible to detect the progranulin levels, due to the lack of plasma samples from these patients.

None of the EOD patients carried the pathogenic hexanucleotide repeat expansion of C9ORF72. All patients were found to have less than 12 repeats [48].

NGS screening: identification of known and unknown rare variants

Through the target re-sequencing of the 17 candidate genes panel, we identified fourteen rare variants in 68% of the selected EOD cases (15 patients) (Table 1). These variants were defined as “pathogenic” when previously described in the literature, and they were classified as “damaging” by bioinformatic tools or as “unknown impact” when no data were available in literature and no deleterious effect was predicted by bioinformatics tools.

Pathogenic variants

Among the 14 identified rare variants, three were classified as pathogenic and damaging and have been described in Italian pedigrees unrelated to the patients analysed in the present study. Two variants were located in the PSEN2 gene (p.M239V=rs28936379 code case: 26_1, [49]; p.M239I=rs63749884 code cases:
30_1, L031, [50]); and one variant was a p.L271fs null mutation in the GRN gene (code cases: 29_1; L029) (Table 1).

The PSEN2-M239V (Table 1, code: 26_1) was identified in an AD subject belonging to a pedigree with four affected family members. The patient presented with behavioural disturbances at the age of 50 years and likely inherited the disease from an affected parent who also showed behavioural disorders at a similar age. The affected parent’s twin sibling was also affected by a neurodegenerative disease, and a second-degree relative developed dementia before the age of 44 years.

The PSEN2-M239I variant (Table 1, code: 30_1) was identified in an AD patient with disease onset at an age of 54 years. The proband belongs to a pedigree with four affected individuals in two generations. The patient likely inherited the disease from an affected parent who was diagnosed with AD before the age of 54 years. The two siblings of this individual presented with behavioural problems that progressed to dementia before the age of 65 years.

Another patient carrying the same M239I mutation (Table 1, code: L031) was diagnosed as having AD at 66 years of age (onset of symptoms at 64 years). Her father and brother were affected with AD with onset at 66 and 62 years of age, respectively.

The GRN L271fs mutation was identified in a pedigree with seven affected subjects in two generations (Table 1, index case code: 29_1). The presence of this mutation was confirmed in other two siblings with FTD, characterized by age of onset of 60 and 61 years, respectively.

Same L271fs mutation was identified also in another pedigree, where the index case (Table 1, code: L029) showed behavioural and speech disturbances at 62 years of age, diagnosed as FTD at 63 years.

Finally, an heterozygous R93C mutation of the VCP gene was detected in a patient (Table 1, code: 36_1) affected by Paget’s disease of bone diagnosed at 44 of age and no signs of dementia at the last examination (47 years). Family history showed multiple individuals with FTD and Paget’s disease with autosomal dominant inheritance.

Variants of unknown impact

Two variants in the TREM2 gene (R62H=rs143332484, D87N=rs142232675, [51]) (Table 1) were observed in three of the 22 patients (14%), including a patient with LBD.
Moreover, in GRN gene, we found variants in the 5'UTR (rs76783532) and two rare missense variants (V77I=rs148531161; R19W=rs63750723; Table 1), which have been reported as “pathogenic nature unclear” and “not pathogenic”, respectively, in online database (www.molgen.en.ua.ac.be/FTDmutations). Accordingly, these missense mutations did not influence the progranulin level in plasma (Table 1).

Other variants reported in online databases were those at 3’UTR of the FUS (rs80301724) and a missense mutation (V1183M=rs10408676) in NOTCH3 gene resulting “damaging” according to bioinformatics tools. Moreover, we detected rare mutations which have not been reported in the literature and databases. In particular, we found a mutation localized in the 5’UTR region and a missense mutation R93C (NM_007126: exon3: c.C277T:p.R93C) in the VCP gene, a deletion (NM_001203251:c.*55delA) in MAPT gene at 3’UTR region, a splicing variant (NM_000435: exon24: c.3838-1G>T) in NOTCH3 gene.

Screening of known common variants

All EOD patients were genotyped for the rs429358 and rs7412 polymorphisms in the APOE gene and the rs1799990 polymorphism in the PRNP gene. APOE ε4 and PRNP 129Val are known to be risk alleles. The results indicated that the frequencies of the APOE ε4 carriers and the PRNP 129Val carriers were 59% for both. Twenty (91%) of the 22 subjects carried at least one of the two risk variants; of these subjects, 27% (six out of 22) carried both the APOE ε4 and the PRNP 129Val risk alleles. Six patients were carriers of at least one risk allele (APOE ε4 or PRNP 129Val) (Table 1).

One patient did not show any rare or common risk variants (Table 1).

Comparison to whole genome data

We compared and calculated the rare and common risk variants frequencies of the 17 candidate genes in our panel, with the data from ADNI whole genome sequencing database (ADNI, Alzheimer Disease Neuroimaging Initiative). In ADNI database, sporadic late-onset (38 patients), early onset (7 patients) AD cases and 272 controls subjects were available. In our sample, the results indicated that, excluding the full
penetrant mutations present in five patients, 9 patients (69%) showed ≥ 2 rare/common risk variants, compared with 13 (34%) and 73 (27%) observed in ADNI late-onset AD cases (LOAD) and controls, respectively (Chi-squared test: $\chi^2 = 15.8$, df=4, p=0.003; Table 2). If we add to our sample, the seven EOAD patients coming from ADNI database (n tot=20), the percentage of genetic load did not change (65%, n=13; Chi-squared test: $\chi^2 = 17.9$, df=4, p=0.001; Table 2).

DISCUSSION

This study suggests that EOD patients without full penetrant mutations are characterized by higher probability to carry polygenic risk alleles that patients with LOAD forms. This finding is in line with recently reported evidence [30], thus suggesting that the genetic risk factors identified in LOAD might modulate the risk also in EOAD.

Moreover, this study reports a high-throughput targeted re-sequencing of candidate dementia disease genes (PRNP, PSEN1, PSEN2, APP, GRN, MAPT, TREM2, CHMP2B, CSF1R, FUS, ITM2B, NOTCH3, SERPINI1, TARDBP, TYROBP, VCP and SQSTM1) in a selected cohort of EOD patients. We confirmed the role of GRN and PSEN2 genes in EOD, with the involvement of specific rare mutations already known. We also detected additional rare variants of unknown impact, located in the 5'/3' UTRs regulatory gene region of the GRN, VCP, MAPT and FUS genes, missense mutations in TREM2, GRN, NOTCH3 and VCP, genes and a splicing variant in NOTCH3 gene. According to the gene damage index (GDI) [46], mutations in these genes showed a value of “medium”, suggesting that these genes are not frequently mutated in healthy populations. This finding further implies that mutations in these genes could be disease-causing.

Rare mutations of known significance

Twenty three percent of the cases carried one pathogenic mutation for dementia. In line with previous studies [14], mutations in the PSEN2 gene, a rare cause of dementia worldwide, were frequent in our cohort (14%). Looking at the geographical distribution of PSEN2 mutations described to date, it is noteworthy that 80% of these mutations were uncovered in two southern European countries, Italy and Spain. Thus, we can speculate that the non-homogeneous distribution of pathogenic mutations might be a result of genetic drift.
The GRN p.L271fs mutation is one of the most common GRN mutations worldwide. An analysis of this mutation in northern Italy showed that almost all families can be traced to a single founder. The origin of the mutation was dated to the Middle Ages at the turn of the first millennium, which explains the high frequency of this mutation in this geographic area [52].

**Rare mutations of unknown significance**

The role for novel variants of unknown significance in both common and rare dementia-associated genes was not exhaustively elucidated. Recently, novel, likely pathogenic variants were described in Italian patients with dementia [53].

We found two AD patients carrying the R62H mutation in the TREM2 gene, which has an unknown impact. A recent review on the correlation between TREM2 and AD [54], showed a meta-analytic association of this mutation with the late onset form of the disease. Our results also showed its involvement in the early onset form of AD. Since the two patients were also homozygous for the APOE ε4 allele, this finding suggests interactions between TREM2 and APOE, as already demonstrated in vitro [55,56]. TREM2 is a lipid sensor that interacts with several AD risk factors involved in lipid metabolism, including APOE, which could decrease the threshold of disease occurrence [57].

In the specific case of two AD patients carrying the GRN p.V77I and R19W mutations, there is evidence of AD pathology in imaging and from biofluid biomarkers (Table 1). Since these missense mutations do not affect the progranulin levels, a pathogenic role of these mutations seems unlikely. However, we cannot exclude that they might have a pathogenic role other than “loss of function”, as no functional studies have been performed. Their presence in AD patients might rather indicate that this gene could be implicated also in the pathophysiological mechanisms leading to AD dementia.

Our analyses showed the presence of additional rare variants located in the 5'/3' UTR regulatory gene region of the GRN, VCP, MAPT and FUS genes. A recent study reported that 3'UTR SNPs, such as rs80301724 in the FUS gene, are present in microRNA binding sites and could impact the post-transcriptional regulation, resulting in overexpression of the protein [58]. Also missense mutations in TREM2, GRN, NOTCH3 and VCP genes and a splicing variant in NOTCH3 gene were detected but, except for some information from bioinformatic tools, their specific functional impact was not assessed. The involvement of NOTCH3 gene in
dementia patients is interesting, both with a missense and with a splicing variant. This gene encodes a single-pass trans-membrane protein of 2321 amino acids, predominantly expressed in vascular smooth muscle cells in adults. It is well documented that NOTCH3 mutations play a critical role in the pathogenetic mechanism of vascular smooth muscle cell degeneration linked to CADASIL, one of the most common hereditary forms of stroke [59]. A recent hypothesis of AD [60] suggested that in CADASIL triggering events in the pathogenic cascade are not amyloid deposits but damaged blood vessels caused by inflammatory reactions that lead to ischemia, amyloid accumulation, axonal degeneration, synaptic loss, and eventually irreversible neuronal cell death. Inflammation and blood vessel damage are well recognized complications of AD, but what causes them and why the cerebral microvasculature is affected is still under debate [60]. Mutations in NOTCH3 gene are known to provoke inflammatory reactions and damage the brain in a wide variety of diseases [59], thus it is possible that one or more mutations in this gene may damage the microvasculature of the brain eventually leading that leads to dementia. The V1183M mutation was classified as a polymorphism in an Italian population [61], though the A allele frequency observed was 0.006.

Inclusion body myopathy with Paget’s disease of bone and IBMPFD is a recently identified autosomal dominant disorder due to mutations in the VCP gene affecting muscle, bone and brain. Interestingly, in our cohort we found the R93C (47832C>T) mutation in the VCP gene already described in patients with IBMPFD [62-65].

Common variants

In this study, we investigated the most established common risk variant for AD, the APOE ε4 haplotype. The functional role of this polymorphism in AD pathogenesis is unclear. However, there is now strong evidence that APOE ε4 could affect amyloid deposition [66]. Consistent with this evidence, in our cohort, ε4 haplotype carriers showed abnormal CSF biomarker levels of amyloid protein (Table 1). Consistent with this evidence, in our cohort all APOE4 carriers with available CSF were amyloid positive (Table 1), except for the case code 19_1 of which we discuss separately (see below). The frequency of the ε4/ε4 genotype (n=4,
17%, exact confidence interval 3-32%) was higher than that commonly observed in the Caucasian population (1000 genomes), which is reported to be 2%.

Moreover, we investigated the non-synonymous polymorphism p.Met129Val in the PRNP gene. Although there are no data on the functional effect of this polymorphism, we observed that the frequency of the risk variant allele G/Val was higher (59%) than the frequency reported in the general European population (frequency G/Val=33% reported by the Exome Aggregation Consortium). A recent meta-analysis showed that the p.Met129Val allele was associated with decreased disease risk in late-onset AD, but not in EOAD [67].

Six patients from the present cohort carry only common risk variants. They could be sporadic cases with onset at the extreme end of expected age range. However, the hypothesis that EOD is caused by mutations in genes not included in the NGS panel cannot be ruled out. In this regard, whole genome sequencing could foster the investigation of additional genetic factors underlying apparently sporadic EOD. Nonetheless, this task was beyond the scope of the present work.

Additional observations

Interestingly, fifteen (68%) of the 22 patients carried at least one rare variant (TREM2, GRN, PSEN2, MAPT, VCP, NOTCH3 or FUS). Among these, fourteen subjects carried also a common variant (APOE and/or PRNP). This result supports the hypothesis that EOD results from the interconnected mechanisms leading to neurodegeneration, where multiple genes can be implicated in one or more systems. Indeed, recent biochemical approaches [55,56] have shown interactions among these genes, such as between TREM2 and APOE in vitro. These results strongly implicate a potential additive/synergic effect in EOD forms linked to the variable inter- and intra-familiar expressivity. To indirectly assess this effect, we found through the ADNI database that, excluding the full penetrant mutations, 71% of our sample showed ≥ 2 rare/common risk variants, as compared to 34% and 27% in sporadic late-onset AD patients and controls, respectively. This indicates that the EOD is more often associated with rare variants or risk alleles, and this could be useful in the genotype-phenotype correlations. Moreover, <10% of subjects, in our cohort, and in late onset AD patients, compared to 25% in a control group, were not carriers of any of the examined variants, which strengthens the idea of using an NGS whole/exome genome approach in a larger sample.
One AD patient with a very early age of onset (41 years) showed neither rare nor common-risk variants. Even an exome clinic investigation detected no rare or common risk variants. The family history was negative, as no other first-degree relatives were affected (neither the parents nor three siblings, two of whom were older than the patient, presented with the disease). Misdiagnosis is unlikely since this patient was positive for all AD biomarkers (abnormal CSF amyloid and tau levels, hypometabolism on FDG-PET, and medial temporal atrophy on MRI). Although we considered the possibility that the patient may show an extreme early-onset presentation of sporadic AD, this finding suggests that additional genes could be implicated in EOD, which strengthens the evidence that the panel of candidate genes needs to be expanded in the future.

For the first time, the D87N mutation in the TREM2 gene was detected in a LBD patient with early onset. LBD is the second most common form of dementia after AD, with a prevalence rate of 4% in the general population [68]. The core symptoms of LBD include sleep disturbances, hallucinations, and cognitive deficits, accompanied within the first year by Parkinsonian motor symptoms. A recent twin study did not show a strong support for a genetic contribution to LBD. However, other studies have demonstrated that LBD aggregates in families and may have an autosomal inheritance pattern (for review [4]). To date, a few genetic markers have been identified. For instance, duplication and SNPs within α, β, γ–synuclein genes have been associated with increased risk of LBD [69,70]. Moreover, mutations in the glucocerebrosidase (GBA) gene are more common in LBD, in addition to mutations in the MAPT or leucine rich repeat kinase 2 (LRRK2) genes (for review [4]). Only one genome-wide linkage study has been performed among patients with familial LBD. A locus on chromosome 2q35-q36 was identified, though none of the genes in this region could explain the relation with LBD [71]. Although further confirmation is needed, the presence of a TREM2 mutation in an LBD patient adds a new actor to its genetic architecture. Mutations in TREM2, a microglial receptor, can lead to aberrant innate immune cell signalling, contributing to the initiation and propagation of several neurodegenerative phenotypes [72-83], including LBD. Moreover, this LBD patient was a carrier of the GG (Val/Val) PRNP risk genotype. This finding is in agreement with a previous study [84] that described a patient carrying the M232R mutation in the PRNP gene who developed dementia and
died six years after onset. An autopsy revealed the patient had dementia with Lewy bodies, not Creutzfeldt-Jakob disease.

Conclusions

This study confirms the role of GRN and PSEN2 mutations in EOD, in the Italian population and provides evidence for roles of novel rare mutations located in the 5'/3' UTRs regulatory gene region of the GRN, VCP, MAPT and FUS genes, missense mutations in TREM2, GRN, NOTCH3 and VCP, genes and a splicing variant in NOTCH3 gene, with a “medium” GDI value. As previously observed, mutations in the PSEN2 gene, a rare cause of dementia worldwide, are frequent in Italian patients. We also confirmed that mutations in GRN gene were present in both FTD and AD phenotypes. Moreover, six patients were carriers of only common risk variants (APOE and PRNP), and one patient did not show any mutation/variant. Overall, 69% (n=9) of our EAOD patients, compared with 34% (n=13) of sporadic LOAD patients and 27% (n=73) of non-affected controls, were carriers of at least two rare/common risk variants in the analysed candidates’ genes panel.

Our results show that EOD patients are characterized by polygenic mutations/risk alleles and that the genetic load without full penetrant mutations stays on higher than that observed in late-onset AD forms. Thus, as recently supported by Bartoletti Stella et al. we recommend the screening of all causative/candidate genes in larger EOD cohorts.

Though our findings are consistent with results obtained from large cohorts [12], independent replications in larger samples are warranted. To further validate the role of polygenic risk variants in EOD, a systematic screening of rare and common variants in dementia-associated genes should be implemented in prospective cohorts with full clinical and biomarker characterization.

ACKNOWLEDGMENTS

This work was supported by the Italian Ministry of Health, Ricerca Corrente, RF-2010-2319722. We also acknowledge the University of Genova for partial support (Fondi ricerca Ateneo 2015 to EDM) and Dr. Irene Maria Briata for technical assistance.
ADNI data collection and sharing was funded by the Alzheimer's Disease Neuroimaging Initiative (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

*ADNI data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

All authors declare that they have no conflict of interest.
REFERENCES


Table 1 Demographic, clinical features and presence of rare/common risk variants in candidate genes in early onset dementia

<table>
<thead>
<tr>
<th>code</th>
<th>gender</th>
<th>age at onset</th>
<th>Clinical diagnosis</th>
<th>Genetic risk</th>
<th>MMSE</th>
<th>Ab</th>
<th>tau</th>
<th>p-tau</th>
<th>atrophy</th>
<th>H-M</th>
<th>PRCI N (mg/ml)</th>
<th>Gene</th>
<th>eDNA Change</th>
<th>Protein Change</th>
<th>Function</th>
<th>dSNP</th>
<th>Pedogenic</th>
<th>SIFT</th>
<th>PP2</th>
<th>Pratt</th>
<th>FATHMM</th>
<th>CADD</th>
<th>GERP++</th>
<th>phylop</th>
<th>APOE</th>
<th>PRNP (M23N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4_1</td>
<td>F</td>
<td>60</td>
<td>AD</td>
<td>Apparently sporadic</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>130</td>
<td>TREM2</td>
<td>G185A</td>
<td>R62H</td>
<td>Missense</td>
<td>rs143332484</td>
<td>unknown</td>
<td>T</td>
<td>B</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>8.315</td>
<td>-0.214</td>
<td>-0.173</td>
<td>e4/e4</td>
<td>Met/Met</td>
</tr>
<tr>
<td>5_1</td>
<td>M</td>
<td>57</td>
<td>AD</td>
<td>Unknown significance</td>
<td>10/30</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>104</td>
<td>GRN</td>
<td>G229A</td>
<td>V77I</td>
<td>Missense</td>
<td>rs148531161</td>
<td>unknown</td>
<td>T</td>
<td>B</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>10.3</td>
<td>-3.85</td>
<td>-0.864</td>
<td>e3/e4</td>
<td>Met/Met</td>
</tr>
<tr>
<td>15_1</td>
<td>M</td>
<td>45</td>
<td>AD</td>
<td>Unknown significance</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Val</td>
</tr>
<tr>
<td>16_1</td>
<td>M</td>
<td>51</td>
<td>AD</td>
<td>Unknown significance</td>
<td>4/30</td>
<td>n.a</td>
<td>+</td>
<td>n.a</td>
<td>-</td>
<td>97</td>
<td>n.a</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>111</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19_1</td>
<td>F</td>
<td>41</td>
<td>AD</td>
<td>Apparently sporadic</td>
<td>n.a.</td>
<td>+</td>
<td>+</td>
<td>n.a</td>
<td>-</td>
<td>115</td>
<td>n.a</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20_1</td>
<td>F</td>
<td>57</td>
<td>AD</td>
<td>Unknown significance</td>
<td>25/30</td>
<td>+</td>
<td>+</td>
<td>n.a</td>
<td>-</td>
<td>97</td>
<td>n.a</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21_1</td>
<td>F</td>
<td>48</td>
<td>AD</td>
<td>Unknown significance</td>
<td>23/30</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>104</td>
<td>FUS*</td>
<td>*41G&gt;A</td>
<td>3'UTR</td>
<td>n.a.</td>
<td>rs80301724</td>
<td>n.a.</td>
<td>e4/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25_1</td>
<td>M</td>
<td>52</td>
<td>AD</td>
<td>Medium to high</td>
<td>26/30</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>104</td>
<td>PSEN2</td>
<td>A715G</td>
<td>M239V</td>
<td>Missense</td>
<td>rs28936379</td>
<td>yes</td>
<td>D</td>
<td>P</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>19.5</td>
<td>4.92</td>
<td>9.213</td>
<td>e3/e3</td>
<td>Met/Val</td>
</tr>
<tr>
<td>26_1</td>
<td>F</td>
<td>50</td>
<td>AD</td>
<td>Medium</td>
<td>14/30</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>+</td>
<td>86</td>
<td>PSEN2</td>
<td>G717A</td>
<td>M239I</td>
<td>Missense</td>
<td>rs63749884</td>
<td>yes</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>27.8</td>
<td>4.92</td>
<td>9.734</td>
<td>e3/e3</td>
<td>Met/Val</td>
</tr>
<tr>
<td>30_1</td>
<td>F</td>
<td>54</td>
<td>AD</td>
<td>Medium</td>
<td>12/30</td>
<td>+</td>
<td>n.a</td>
<td>n.a</td>
<td>-</td>
<td>85</td>
<td>MAPT</td>
<td>*55delA</td>
<td>3'UTR</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>e4/e4</td>
<td>Val/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31_2</td>
<td>F</td>
<td>60</td>
<td>AD</td>
<td>Medium</td>
<td>22/30</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>+</td>
<td>102</td>
<td>TREM2</td>
<td>G185A</td>
<td>R62H</td>
<td>Missense</td>
<td>rs143332484</td>
<td>unknown</td>
<td>T</td>
<td>B</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>8.315</td>
<td>-0.214</td>
<td>-0.173</td>
<td>e4/e4</td>
<td>Val/Val</td>
</tr>
<tr>
<td>11_1</td>
<td>F</td>
<td>51</td>
<td>FTD</td>
<td>High</td>
<td>23/30</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
<td>n.a</td>
<td>108</td>
<td>GRN</td>
<td>-3895G&gt;T</td>
<td>5'UTR</td>
<td>n.a.</td>
<td>rs76783532</td>
<td>n.a.</td>
<td>e3/e3</td>
<td>Met/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29_1</td>
<td>M</td>
<td>56</td>
<td>FTD</td>
<td>High</td>
<td>39</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>111_11414del</td>
<td>L271fs</td>
<td>Deletion</td>
<td>n.a.</td>
<td>yes</td>
<td>n.a.</td>
<td>e3/e3</td>
<td>Met/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14_1</td>
<td>M</td>
<td>57</td>
<td>LBD</td>
<td>Unknown significance</td>
<td>5/30</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>e3/e3</td>
<td>Val/Val</td>
</tr>
<tr>
<td>34_1</td>
<td>F</td>
<td>49</td>
<td>AD</td>
<td>Medium to high</td>
<td>15/30</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>149</td>
<td>GRN</td>
<td>C55T</td>
<td>R19W</td>
<td>Missense</td>
<td>rs63750723</td>
<td>unknown</td>
<td>T</td>
<td>B</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>8.7</td>
<td>0.13</td>
<td>0.17</td>
<td>e3/e4</td>
<td>Met/Val</td>
</tr>
<tr>
<td>35_12</td>
<td>F</td>
<td>54</td>
<td>AD</td>
<td>Medium</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>111_11414del</td>
<td>L271fs</td>
<td>Deletion</td>
<td>n.a.</td>
<td>yes</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36_1</td>
<td>M</td>
<td>57</td>
<td>FTD-BMPF</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>145</td>
<td>VCP</td>
<td>C277T</td>
<td>R93C</td>
<td>Missense</td>
<td>unknown</td>
<td>P</td>
<td>D</td>
<td>M</td>
<td>D</td>
<td>3.8</td>
<td>19.39</td>
<td>2.70</td>
<td>e3/e4</td>
<td>Met/Val</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37_2</td>
<td>F</td>
<td>62</td>
<td>AD</td>
<td>Low</td>
<td>n.a.</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>138</td>
<td>NOTCH3</td>
<td>G3547A</td>
<td>V1183M</td>
<td>Missense</td>
<td>rs10408676</td>
<td>D</td>
<td>D</td>
<td>M</td>
<td>D</td>
<td>3.9</td>
<td>3.09</td>
<td>2.32</td>
<td>e3/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1413p2</td>
<td>M</td>
<td>62</td>
<td>Familial Dementia</td>
<td>Unknown significance</td>
<td>5/30</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>98</td>
<td>n.a</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>+</td>
<td>n.a.</td>
<td>e3/e4</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>Gender</td>
<td>Age</td>
<td>Status</td>
<td>Diagnosis</td>
<td>Age at Onset</td>
<td>Early Onset</td>
<td>Huntington’s Disease</td>
<td>Sporadic</td>
<td>Mutation Type</td>
<td>Genetic Risk</td>
<td>Clinical Details</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>----------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L028</td>
<td>F</td>
<td>63</td>
<td>AD</td>
<td>Low</td>
<td>161 NOTCH3</td>
<td>3838-1G&gt;T</td>
<td>+ n.a. - +</td>
<td>e4/e4</td>
<td>Met/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L029</td>
<td>F</td>
<td>62</td>
<td>FTD</td>
<td>High</td>
<td>811_814del</td>
<td>L271fs</td>
<td>Deletion n.a. yes</td>
<td>c3/e3</td>
<td>Met/Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L031</td>
<td>F</td>
<td>64</td>
<td>AD</td>
<td>Medium</td>
<td>4/27 PSEN2</td>
<td>G717A</td>
<td>M239I rs63749884</td>
<td>D D D A D</td>
<td>27.8 4.92 9.734</td>
<td>e3/e3 Met/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Keys: APOE, Apolipoprotein E (E4 risk allele); PRNP, Prion protein (G/val risk allele); PGRN/GRN, Progranulin; PSEN2, Presenilin 2; TREM2, Triggering receptor expressed on myeloid cell; FUS, Fused in sarcoma; MAPT, Microtubule-associated protein TAU; VCP, Valosin-containing protein; A, known damaging; B, benign; D, damaging; N, neutral; P, possibly damaging; T, tolerant; na, not available; CADD, CADD phred-scaled (a score of 20 means 1% percentile highest scores of whole genome); dbSNP, single nucleotide polymorphism database (rs number); GERP, genomic evolutionary rate profiling score; GERP++, evolution score; MT, mutation taster; Phylop, phyloP100way vertebrate; PP2, PolyPhen2; SIFT, Sorting Intolerant from Tolerant; FATHMM, Functional Analysis through Hidden Markov Models (v2.3); LRT, Likelihood Ratio Test; AD, Alzheimer Disease; FTD, Fronto Temporal Dementia; LBD, Levy Body Dementia; FTD-IBMPFD, Inclusion Body Myopathy-Paget’s disease of bone; “Genetic risk” denotes the estimated probability of identifying a genetic mutation base on Loy and Wood criteria for AD and FTD; “Unknown Significance” denotes lack of information about diagnosis or clinical details; “Apparently sporadic” indicates no other affected case in the family; H-M, Hypo metabolism.
Table 2. Comparison of frequencies of rare (excluding the full penetrant) and common risk variants in our sample (This study) versus controls, late onset and early onset AD patients (LOAD and EOAD respectively) obtained from ADNI database

<table>
<thead>
<tr>
<th>Variants</th>
<th>n=0</th>
<th>n=1</th>
<th>n≥2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>Only rare</td>
<td>Only common</td>
</tr>
<tr>
<td>Groups</td>
<td>Sample (N)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>ADNI</td>
<td>Controls (272)</td>
<td>67 (25)</td>
<td>9 (7)</td>
</tr>
<tr>
<td></td>
<td>LOAD (38)</td>
<td>3 (8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>EOAD (7)</td>
<td>1 (14)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>This study</td>
<td>AD (13)</td>
<td>1 (8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>ADNI + This study</td>
<td>EOAD + AD (20)</td>
<td>2 (10)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

$\chi^2$-squared test: Controls vs LOAD-ADNI vs AD-our study, $\chi^2=15.8$, df=4, p=0.003

$\chi^2$-squared test: Controls vs LOAD-ADNI vs EOAD-ADNI+AD-our study, $\chi^2=17.9$, df=4, p=0.001