Genetics of dementia in a Finnish cohort

Pétra Pasanen1,2, Liisa Myllykangas3, Minna Pöyhönen4, Anna Kiviharju8, Maija Siitonen1, John Hardy5, José Bras5,6,7, Anders Paetau3, Pentti J. Tienari8,9, Rita Guerreiro5,6,7, and Auli Verkkoniemi-Ahola8

Running title: Genetics of dementia in Finnish families

Author affiliations:
1 Department of Medical Biochemistry and Genetics, Institute of Biomedicine, University of Turku, Turku, Finland
2 Tyks Microbiology and Genetics, Department of Medical Genetics, Turku University Hospital, Turku, Finland
3 Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
4 Department of Clinical Genetics, Helsinki University Central Hospital and Department of Medical Genetics, University of Helsinki, Helsinki
5 Department of Molecular Neuroscience, Institute of Neurology, University College London, London, UK
6 UK Dementia Research Institute at UCL
7 Department of Medical Sciences and Institute of Biomedicine – iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal
8 Clinical Neurosciences, Neurology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
9 Research Programs Unit, Molecular Neurology, University of Helsinki, Helsinki, Finland

Corresponding author: Pétra Pasanen
Address: Kiinamyllynkatu 10 FIN-20520 Turku
Phone: +358 2 333 7233
e-mail: petra.pasanen@utu.fi

Conflict of interest

The authors declare that they have no conflict of interest.
Abstract

Alzheimer’s disease (AD) and frontotemporal dementia (FTD) are the two most common neurodegenerative dementias. Variants in APP, PSEN1 and PSEN2 are typically linked to early-onset AD, and several genetic risk loci are associated with late-onset AD. Inherited FTD can be caused by hexanucleotide expansions in C9orf72, or variants in GRN, MAPT or CHMP2B. Several other genes have also been linked to FTD or FTD with motor neuron disease.

Here we describe a cohort of 60 Finnish families with possible inherited dementia. Our aim was to clarify the genetic background of dementia in this cohort by analysing both known dementia-associated genes and searching for rare or novel segregating variants with exome sequencing. C9orf72 repeat expansions were detected in 12 (20%) of the 60 families, including, in addition to FTD, a family with neuropathologically verified AD. Twelve families (10 with AD and 2 with FTD) with representative samples from affected and unaffected subjects and without C9orf72 expansions were selected for whole exome sequencing. Exome sequencing did not reveal any variants that could be regarded unequivocally causative of the disease in the families, but revealed potentially damaging variants in UNC13C and MARCH4.

Key words: Alzheimer’s disease, frontotemporal dementia, C9orf72, MARCH4, UNC13C, CLU
Introduction

Alzheimer’s disease (AD) and frontotemporal dementia (FTD) are the two most common neurodegenerative dementias. AD is characterized by progressive loss of memory, typically presenting with deficits in anterograde episodic memory. Other cognitive functions, such as language, executive functions and visuospatial functions, deteriorate as the disease progresses. Most AD patients first develop symptoms after 65 years of age (late-onset AD, LOAD), while less than 10% of patients present with early-onset AD (EOAD). Autosomal dominant inheritance and rare cases of autosomal recessive inheritance are seen in the EOAD group, due to variants in the amyloid precursor protein (APP), presenilin 1 (PSEN1) and presenilin 2 (PSEN2) genes. Familial cases are also seen in LOAD with an estimated heritability of 58 to 79%. More than 20 disease-associated loci have been detected in genome-wide association studies (GWAS) and meta-analyses but, apart from the APOE ε4 and TREM2 p.(R47H) risk alleles, most of these only have a modest effect. Although rare variants in APP, PSEN1 and PSEN2 have been detected in LOAD patients by targeted resequencing, variants affecting function are rare in LOAD and its pathobiology reflects the interplay of predisposing genetic variants and environmental factors. Sequencing studies have also shown that rare variants can be found in dementia-associated loci identified through GWAS.

In contrast to AD, FTD is more commonly observed in patients younger than 65 years. FTD may present with changes in personality and behaviour (behavioural-variant FTD) or language difficulties (non-fluent variant primary progressive aphasia and semantic-variant progressive aphasia). Up to 40% of patients have a positive family history with autosomal dominant inheritance in 10%. The most commonly mutated genes are C9orf72, GRN, MAPT, TARDBP, FUS, VCP, CHMP2B, UBQLN2, TBK1, SQSTM1 and CCNF, have been detected in patients with either FTD and motor neuron disease (FTD-MND) or pure FTD. C9orf72 expansions are prevalent
in Finnish patients with FTD or ALS, accounting for 48% and 46% of familial FTD and ALS, respectively, and 19% and 21% of sporadic FTD and ALS\textsuperscript{13}.

Based on the known functions of disease-associated genes, several pathways involved in the pathogenesis of AD and FTD have been identified. In AD, these include the amyloid β pathway, the immune system (\textit{CLU, CR1, ABCA7, CD33, EPHA1, the MS4A gene cluster}), synaptic activity (\textit{PICALM, CD33, CD2AP, EPHA1, BIN1}) and lipid metabolism (\textit{CLU, ABCA7}) (reviewed in \textsuperscript{28}). In FTD, the disease-implicated pathways include RNA processing and transcription regulation (\textit{C9orf72, TARDBP, FUS}), microtubule function (\textit{MAPT}), immune response (\textit{GRN}), lysosome-mediated and ubiquitin-mediated protein degradation and autophagy (\textit{GRN, VCP, CHMP2B}) (reviewed in \textsuperscript{11}).

Here we describe a cohort of 60 Finnish families with possible inherited dementia. Our aim was to clarify the genetic background of dementia in this cohort by analysing both known dementia-associated genes and searching for rare or novel segregating variants using exome sequencing. We show that \textit{C9orf72} hexanucleotide repeat expansions are common in this cohort but variants affecting function of the other most common AD and FTD genes are not accountable for the disease in these families. We also present rare variants that segregate with AD and FTD in small families. Although no definite conclusion can be achieved regarding the causal involvement of these rare variants, these should be taken into account in future studies trying to identify the genetic cause of familial dementias.

\textbf{Subjects and methods}

\textbf{Study cohort}

The study cohort is comprised of affected and unaffected members from 60 Finnish families with possible inherited dementia. The families were recruited from neurology clinics in the
Helsinki and Uusimaa hospital district (Southern Finland) and via an advertisement in a national newspaper in the late 1990s. The recruitment method proved particularly successful, resulting in 60 families suited for the study. A total of 364 blood-derived DNA samples (107 from affected patients and 257 from unaffected family members) were available from the families. A prerequisite for participation was a positive family history with two or more living first-degree family members affected by dementia.

The clinical diagnosis was AD in most families (n = 38), FTD in 10 families, dementia with Lewy bodies (DLB) in one family and unspecified dementia in 11 families (Supplementary table 1). The diagnoses were based on clinical findings and brain CT imaging studies.

Extensive neuropsychological studies had been performed for some of the patients. Liquor biomarkers were not available at the time of patient recruitment and sample collection.

Neuropathological data was available from patients belonging to seven of the 60 families.

The ages at onset are listed in Supplementary table 2.

**Ethical aspects**

Informed consent was obtained from all participants. The study was approved by the Ethics committee of Neurology department at HUCH (4.6.1997 and 11.1.2012) and the HUCH Ethics Committee of Medicine (Dnro104/13/03/01/14). Approval for using patient tissue specimens was given by Valvira (Dnro 2855/06.01.03.01/2012). Approval for using medical records and autopsy reports of the patients living outside the HUS district was obtained from National Institute for Health and Welfare (Dnro THL/701/5.05.00/2013).

**Methods**

EDTA blood samples were drawn after obtaining informed consent from the participants. Both affected patients and unaffected family members were recruited for the study. DNA
was extracted using standard protocols. Overview of the study scheme is shown in Figure 1. *APOE* genotypes were determined by PCR and *Cfo*I digestion as described in Myllykangas *et al.* 29. Screening of *C9orf72* expansions was done by repeat primed PCR as described by Renton *et al.* 13.

Expansion was defined by two criteria that had to be fulfilled: 1. Characteristic saw-tooth pattern in repeat-primed PCR extending over 30 G\(_4\)C\(_2\) repeats on capillary electrophoresis of the PCR products, and 2. Lack of large allele (>30 repeats) amplicon in standard PCR across the repeat region. Standard PCR across the repeat region was performed using LongAmp Taq Reaction Buffer (New England Biolabs) with the following PCR primers:

\[
5'-GGA\ GGG\ AAA\ CAA\ CCG\ CAG\ CCT\ GTA\ G-3'\text{ and } 5' -ATG\ CCG\ CCT\ CCT\ CAC\ TCA\ CCC\ ACT\ CG-3',
\]

1.8 M of Betaine. The PCR products were run on 2 % agarose gels.

We evaluated the pedigrees for the availability of samples from both affected and unaffected individuals as well as the availability of neuropathological data. Twelve families without *C9orf72* expansions (2 with FTD and 10 with AD) were selected for further genetic studies. The FTD families were screened for variants in *GRN* and the AD families for variants in exons 16 and 17 of *APP* and the coding regions of *PSEN1* and *PSEN2*. Exons and flanking splice site regions were amplified by PCR and the purified PCR products were sequenced in both directions using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, CA, USA). All primer sequences and PCR conditions are available upon request.

Large structural and copy number variants were excluded by using HumanOmniExpress Bead Chip (Illumina, San Diego, CA, USA). Loci known to have copy number variants that associate with dementia (such as *APP* and *SNCA*) were checked visually. In addition, the data was analysed with CNVPartition in Genome Studio (Illumina, San Diego, CA, USA) to
detect large (> 50 kb) CNVs. The identified CNVs were checked against the Database of Genomic Variants (DGV) \(^{30}\). Whole exome sequencing (WES) of selected individuals was done at University College London (UCL, London, UK). Exome enrichment was performed using TruSeq Exome Capture kit (Illumina, San Diego, CA, USA). Sequencing was performed on a HiSeq 2000 (Illumina, San Diego, CA, USA). Reads were aligned to GRCh37/hg19 using BWA, variants called according to GATK best practice guidelines and annotated with ANNOVAR \(^ {31}\). *In silico* pathogenicity predictions of nonsynonymous variants were done with SIFT \(^ {32}\), Polyphen2 \(^ {33}\), MutationTaster \(^ {34}\), MutationAssessor \(^ {35}\), and CADD \(^ {36}\). Variants were filtered against population databases (1000Genomes, ESP and ExAC) and prioritised based on variant type (missense, nonsense, splice site, frameshift, non-frameshift) and predicted pathogenicity. We concentrated on variants found in genes implicated in GWAS or genes that are highly expressed in the brain. We also assessed the known functions of the genes of interest. Selected variants (shared by the affected members in each family but not present in the unaffected family member, if appropriate sample was available) from WES were confirmed with Sanger sequencing and their segregation tested in a family setting. These variants and the associated phenotypes have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) with accession numbers SCV000576395, SCV000576396, SCV000576397, SCV000576398, and SCV000576399. We also checked the frequencies of these selected variants in SISu, a database of sequence variants in Finns (Sequencing Initiative Suomi project (SISu), Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland (URL: http://sisuproject.fi), SISu v4.1, accessed in 09/2016).
Results

APOE genotyping

APOE genotyping was performed for 364 samples from 60 families. There were 34 individuals homozygous for the risk APOE genotype ε4 (20/107 affected, 18.7%; 14/257 unaffected, 5.4%). A total of 166 individuals were heterozygotes APOE ε3/ε4 (56/107 affected, 52.3%; 110/257 unaffected, 42.8%). The most common allele, APOE ε3, is not associated with an increased risk for AD, and it was detected in 28 of 107 (26.1%) affected patients and in 123 of 257 (47.8%) unaffected individuals. Five unaffected individuals (5/257, 1.99%) were APOE ε2/ε4 heterozygotes and one unaffected (1/257, 0.3%) was ε2/ε3. No genotype was obtained for seven samples (3 affected and 4 unaffected). The genotypes in each family are shown in Supplementary table 2.

49 C9orf72

All 60 families were included in the C9orf72 hexanucleotide expansion screening. Expansions were detected in 12 of the 60 families (20%). The distribution of expansions in affected and unaffected individuals in each family is shown in Supplementary table 2. Clinical diagnosis was FTD in seven families, AD or variant AD in four families and degenerative dementia (ALS, DLB or AD resembling syndromes) in one family (Table 1). The proportions of C9orf72 expansions in each diagnosis group in shown in Figure 2. As FTD or FTD/ALS are the typical clinical phenotypes associated with C9orf72 expansions, we only describe the five families with more atypical presentations in detail. No additional information was available from family Fam-62.
Table 1. Families with $C9orf72$ expansions. Two families (Fam-18 and Fam-73) marked with an asterisk may have been initially misdiagnosed as AD. In Fam-25 (marked with two asterisks) the neuropathological diagnoses are based on haematoxylin & eosin and ancillary stainings. No immunohistochemical stainings for FTLD were available at the time of the neuropathological examination.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Clinical diagnosis</th>
<th>Neuropathological diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fam-18</td>
<td>AD*+FTD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-22</td>
<td>FTD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-25</td>
<td>FTD</td>
<td>FTLD + AD-type lesions**</td>
</tr>
<tr>
<td>Fam-27</td>
<td>dementia (AD)</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-31</td>
<td>degenerative dementia (ALS, DLB or AD resembling syndromes)</td>
<td>AD</td>
</tr>
<tr>
<td>Fam-33</td>
<td>FTD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-39</td>
<td>FTD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-50</td>
<td>AD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-62</td>
<td>AD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-71</td>
<td>FTD</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-73</td>
<td>variant AD*</td>
<td>not available</td>
</tr>
<tr>
<td>Fam-77</td>
<td>FTD/ALS</td>
<td>not available</td>
</tr>
</tbody>
</table>

The index patient of family Fam-18 developed symptoms at 65 years. Clinical presentation was compatible with AD, but CT and MRI studies showed frontal atrophy. Clinical data from other family members was not available.

Patient records of two affected patients from family Fam-27 were available for review. Both patients had clinical AD with late onset.

Neuropathological data of two affected patients from family Fam-31 were available. The index patient had late-onset AD and the diagnosis was verified neuropathologically post-mortem in 1998 using the methods available at that time. Re-analysis of the archived formalin-fixed, paraffin-embedded sample revealed Braak stage V tau-pathology and
CERAD stage B beta-amyloid load. TDP-43 staining was negative. p62 positive inclusions were observed in the granular cerebellar cells. No DNA sample was available for study. Formalin-fixed brain tissue from temporal lobe of another patient of the family was available for immunohistochemistry. TDP-43 staining was negative, but moderate to severe tau-pathology suggestive of AD was observed. This patient was shown to harbor a C9orf72 expansion and carried the APOE ε4/4 genotype.

Family Fam-50 included several affected individuals with onset of disease after 70 years of age. DNA sample was available from one of them. One patient had visual hallucinations as the first symptom and subsequently developed loss of concentration and memory deficit. Neuropsychological examination was consistent with large-scale impairment and visual defect.

The index patient of family Fam-73 was diagnosed with variant AD. However, brain SPECT was suspective of FTD. The index patient’s sister had been diagnosed with ALS. Thus, the actual diagnostic spectrum of this family is consistent with FTD and ALS.

Further studies on 12 families without C9orf72 expansions

Exclusion of GRN, APP, PSEN1, PSEN2

Twelve families without C9orf72 expansions and representative samples from both affected and unaffected individuals were selected for further studies (Figure 1.). The clinical diagnoses in these families were FTD (two families: Fam-13 and Fam-59) and AD (10 families: Fam-15, Fam-29, Fam-32, Fam-35, Fam-38, Fam-49, Fam-52, Fam-55, Fam-56, and Fam-57). Sanger sequencing did not reveal any causal variants in GRN (FTD), APP exons 16 and 17, or coding regions of PSEN1 and PSEN2 (AD families).

Exclusion of large structural and copy number variants by SNP microarray
None of the 12 index patients had a duplication of APP. Neither did we identify any deletions or duplications involving other known dementia-associated genes, such as SNCA.

**Whole exome sequencing**

Whole exome sequencing (WES) data was generated for at least two affected patients from each of the 12 families. The oldest unaffected family members from whom a DNA sample was available (7 families) were exome sequenced as controls. We concentrated on rare variants identified in WES shared by the affected patients in each family but not seen in the analysed healthy family members, when available (list of rare variants in each family can be found in a Supplementary table 3). Confirmation and segregation analyses were done with Sanger sequencing. A large number of shared rare variants were identified in each family, but we concentrated on variants in GWAS hit genes or in genes with known functions possibly relevant for neurodegeneration. The validated variants are listed in Table 2 and presented in detail below.

Table 2. Rare variants identified in exome sequencing and validated in a family setting.

<table>
<thead>
<tr>
<th>Family</th>
<th>CLU</th>
<th>PCDH11X</th>
<th>UNC13C</th>
<th>MARCH4</th>
<th>MARCH4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fam-56</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
</tr>
<tr>
<td>Fam-15</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
</tr>
<tr>
<td>Fam-49</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
</tr>
<tr>
<td>Fam-59</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
</tr>
<tr>
<td>Fam-13</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
<td>AD</td>
</tr>
</tbody>
</table>

Minor allele frequencies (MAFs) of each variant in 1000Genomes, ESP, SiSU and ExAC are shown. Different prediction programs (SIFT, PolyPhen, MutationTaster, MutationAssessor, CADD) were used to estimate the deleteriousness of the variants.
<table>
<thead>
<tr>
<th>SIFT</th>
<th>tolerated</th>
<th>Damaging</th>
<th>-</th>
<th>tolerated</th>
<th>damaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>PolyPhen</td>
<td>Benign</td>
<td>possibly damaging</td>
<td>-</td>
<td>benign</td>
<td>damaging</td>
</tr>
<tr>
<td>MutationTaster</td>
<td>Neutral</td>
<td>damaging</td>
<td>-</td>
<td>damaging</td>
<td>damaging</td>
</tr>
<tr>
<td>MutationAssessor</td>
<td>medium effect</td>
<td>neutral</td>
<td>-</td>
<td>neutral</td>
<td>neutral</td>
</tr>
<tr>
<td>CADD phred score</td>
<td>8.521</td>
<td>14.33</td>
<td>-</td>
<td>2.416</td>
<td>19.62</td>
</tr>
<tr>
<td>1000G</td>
<td>0.005</td>
<td>absent</td>
<td>Absent</td>
<td>0.000599042</td>
<td>absent</td>
</tr>
<tr>
<td>ESP</td>
<td>0.0027</td>
<td>absent</td>
<td>Absent</td>
<td>7.7e-05</td>
<td>absent</td>
</tr>
<tr>
<td>SISu</td>
<td>0.000698324</td>
<td>0.00017454</td>
<td>Singleton</td>
<td>0.00124611</td>
<td>singleton</td>
</tr>
<tr>
<td>ExAC (Finnish)</td>
<td>0.00121</td>
<td>singleton</td>
<td>Absent</td>
<td>0.001285</td>
<td>absent</td>
</tr>
<tr>
<td>ExAC (total)</td>
<td>0.001673</td>
<td>singleton</td>
<td>Absent</td>
<td>0.0002082</td>
<td>singleton</td>
</tr>
<tr>
<td>Significance</td>
<td>likely benign</td>
<td>possibly benign</td>
<td>possibly deleterious</td>
<td>likely benign</td>
<td>possibly deleterious</td>
</tr>
</tbody>
</table>

**WES findings in AD families**

**CLU**

A heterozygous *CLU* c.608C>T, p.(Thr203Ile) variant (rs41276297) was identified in two affected patients of family Fam-56. The variant was not detected in the four unaffected family members from whom a sample was available (Figure 3a). This variant is a previously known, rare variant with a frequency of 0.00121 in Finnish samples in ExAC. Polyphen and SIFT predicted no deleterious effect. This variant has also been detected in British AD samples (reported as p.T255I) with a frequency of 0.003 as well as in unaffected controls (frequency 0.006)\(^{37}\). One of the affected individuals also carried one *APOE* ϵ4 allele, while the other was homozygous for ϵ3.

**PCDH11X**

Affected members of family Fam-15 carried a heterozygous c.2279A>T, p.(Asp760Val) variant in *PCDH11X*. This variant (rs781770086) is present as a singleton in ExAC and in SISu. Sanger sequencing confirmed the variant in the two affected patients (II:4 and III:4).

However, segregation analysis showed that the variant was also present in two currently
unaffected individuals (III:2 and III:5) and in one individual (III: 6) with unclear status. The remaining two unaffected family members (II:7 and III:7) did not carry the variant (Figure 3b). Only one of the affected individuals carried APOE ε4.

**UNC13C**

In family Fam-49, a heterozygous 3-bp deletion in UNC13C, c.1324_1326del, p.(443del) was detected in two affected patients. The variant was not seen in two unaffected family members (Figure 3c). This in-frame deletion variant (rs746069739) is present as a singleton in ExAC and in SISu (low-quality). In addition to the UNC13C variant, both affected individuals also carried one APOE ε4 allele.

**WES findings in FTD families**

**MARCH4**

Two affected patients from the FTD family Fam-13 carried a heterozygous c.631A>G, p.(Lys211Glu) variant, (rs756981946) in MARCH4. This variant was absent from the unaffected family members (Figure 3d). The APOE genotypes of the two affected individuals of Fam-13 were ε3/4 and ε3/3. Two affected members of the second FTD family, Fam-59, carried a heterozygous c.39C>G variant, p.(Trp13Cys) (rs145386484) in MARCH4. Segregation analysis showed that this variant was also present in 7 currently unaffected family members (ages 45 – 73 years) and absent in other 9 unaffected family members (Figure 3e). All studied individuals in Fam-59 were homozygous for APOE ε3.

MARCH4 p.(Lys211Glu) variant is present in ExAC as a singleton in a non-Finnish European sample and in SISu as a singleton (identifier: rs756981946). In silico predictions gave the following results: Polyphen2 predicted the variant to be probably damaging (score 0.995), SIFT tolerated (score 0.29), MutationTaster damaging (score 1.000), MutationAssessor
medium effect (score 2.22), CADD Phred-like scaled C-score was 19.62. These data
demonstrate that the variant is extremely rare and suggest that it might alter the normal
function of MARCH4.

The p.(Trp13Cys) variant is more common as it is reported in ExAC with a frequency of
0.001285 in Finnish samples. It has also been detected in other populations: European
(3/56914), South Asian (8/13234), African (3/7972) and Latino (1/9762). SIFT predicted this
variant to be tolerated (score 0.29), Polyphen2 benign (score 0.00), MutationTaster damaging
(score 0.981), MutationAssessor neutral (score -0.55), and CADD Phred-like scaled C-score
was 2.416. These predictions along with the fact that it was present in individuals over 70
years of age suggest that p.(Trp13Cys) might be a rare neutral variant.

Discussion

In contrast to early-onset AD, late-onset Alzheimer’s disease is rarely caused by segregating
variants in families. The strongest identified risk factor is the APOE ε4 allele. In few cases,
variants in APP, PSEN1 and PSEN2 have been reported in LOAD families⁹. GWAS studies
have identified approximately 20 loci associated with predisposition to AD but finding
variants that actually have a biological effect has proven difficult. In FTD, variants in
C9orf72, MAPT and GRN account for up to 60% of familial cases while variants in other
genes are rare ¹¹.

In addition to the ALS/FTD entity, C9orf72 expansions have been linked to several other
clinical manifestations including AD, Parkinson’s disease and Huntington’s disease
phenocopies (reviewed in ³⁸). We detected C9orf72 expansions in 7/60 (11.6%) families with
either FTD or FTD/ALS but also in 3/60 (5%) families with clinical AD. In one family, Fam-
31, neuropathological examinations disclosed moderate to severe AD tau-pathology, and no
TDP-43-positive inclusions were seen. However, p62-positive inclusions were present in the
cerebellum, consistent with the C9orf72 expansion. Several earlier reports have described C9orf72 expansions in either clinically diagnosed or neuropathologically confirmed AD. It is possible that the AD pathology is at least partly attributable to APOE as the one affected individual with C9orf72 expansion and AD-type neuropathology was homozygous for the APOE ε4 allele (Supplementary table 2). Previous work has shown that C9orf72 expansions are seen in ~30% of Finnish FTD patients and in 48.1% of familial FTD. Our results confirm this finding and suggest that C9orf72 expansions may manifest as clinical AD and some patients may also show concomitant AD pathology at the neuropathological examination. Previous studies on C9orf72 expansions in AD patients have suggested that the clinical or neuropathological classification as AD may have been incorrect, and this appeared to be the case in some of our families with clinical diagnosis of AD.

No variants in APP exons 16 and 17 or the coding regions of PSEN1 and PSEN2 were observed in the 10 AD families selected from our cohort. Whole-genome genotyping also showed no clearly causative CNVs. Both results are in agreement with previous studies. Only a few PSEN1 variants have been reported in Finnish AD families: two families carry the ‘Cotton-wool’ variant, Δ9Finn (c.869_955del) p.(Met146Val) has been reported in a Swedish family of Finnish descent and p.(Pro264Leu) in one family. Screening of APP, PSEN1 and PSEN2 in a cohort of 140 EOAD patients revealed no variants that might affect function. In addition, duplication of APP was not detected in a cohort of 64 Finnish EOAD patients.

GRN sequencing and exome sequencing did not reveal any pathogenic variants in the two FTD families without C9orf72 expansions. In agreement with our results, previous work suggests that GRN variants are rare among Finnish FTD patients.
Exome sequencing revealed rare, potentially relevant variants in five families. Two variants were in genes previously linked to AD (CLU and PCDH11X) while three variants were in genes (UNC13C and MARCH4) that have not been directly linked to dementia but could be important in maintaining normal neuronal function.

In 2010, a large GWAS study indicated that PCDH11X was linked to LOAD in a combined American Caucasian cohort. However, subsequent studies in different populations failed to confirm the findings of Carrasquillo et al. Recently, Jiao et al. reported a SNP in PCDH11X to confer a risk to LOAD. Thus, the possible role of PCDH11X in AD susceptibility is still somewhat unclear. Our results show that the rare p.(Asp760Val) variant is present in all affected individuals of family Fam-15 but also in two asymptomatic individuals and in one subject with unclear status.

The role of CLU as an AD risk gene has been established in independent datasets. We noted co-segregation of a rare CLU variant and dementia in an AD family (Fam-56). Even though rare non-synonymous and small insertion/deletion variants have been reported to increase AD risk, the p.(Thr203Ile) variant is predictably not deleterious, but at present we cannot exclude its possible role in AD risk.

Two AD patients from family Fam-49 shared a 3bp in-frame deletion in UNC13C. The UNC13C gene is highly expressed in brain. Experimental evidence from cat and mouse models have suggested that its mammalian homologue, Munc13-3, has a role in controlling critical-period neuronal plasticity in visual cortex. Gene expression studies in human AD and control brain samples showed increased UNC13C expression in hippocampal CA3 compared to CA1 in Alzheimer patients. This implicates that UNC13C might have a neuroprotective role in the brain. The rare variant found in family 49 removes one amino
acid residue but does not disturb the reading frame. Both affected patients were also heterozygous for the \textit{APOE} ε4 allele, a likely risk factor in this family.

A rare segregating missense variant in \textit{MARCH4} was identified in the FTD family Fam-13. \textit{MARCH4} is a member of membrane-associated RING-CH family of ubiquitin E3 ligases. These ligases function in the last step of ubiquitination by recruiting the ubiquitin carrying E2 enzyme and transferring ubiquitin from E2 to the target protein\textsuperscript{63}. \textit{MARCH4} is predominantly expressed in the adult human brain\textsuperscript{64}. The ubiquitin-related protein degradation pathway has been implicated in many neurodegenerative diseases, including FTD. Recent work by Williams and coworkers described variants in a component of the ubiquitin E3 ligase complex, \textit{CCNF}, in a large ALS/FTD family and a few singleton patients\textsuperscript{27}. Although the \textit{MARCH4} variant segregated with FTD in our small family, we cannot exclude the possibility that we merely identified a rare neutral variant in a gene with function that could fit in the model of FTD pathogenesis.

\textit{C9orf72} repeat expansions are common among Finnish FTD patients and our results indicate that expansions may also be seen in patients with clinical and neuropathological diagnoses of AD. Our results suggest that unknown genetic factors are likely to be responsible for a proportion of familial dementia in the Finnish cohort, but definitely causal or risk variants in novel genes are yet to be identified. Exome sequencing is an efficient way to search for rare coding variants, but thus far only few segregating risk variants (e.g. \textit{TREM2} p.(Arg47His)\textsuperscript{65} and \textit{TTC3} p.(Ser1038Cys)\textsuperscript{66}) have been described in LOAD families. Our results corroborate the view that even in large LOAD families with multiple affected individuals the disease is likely caused by combination of multiple genetic and environmental risk factors.

The \textit{APOE} ε4 risk allele can be assumed to account for multiple affected individuals in several of the AD families in our study.
We detected rare segregating coding variants in *UNC13C* in an AD family and in *MARCH4* in an FTD family. However, replication in larger familial and case-control datasets and functional assays would be needed to prove their causality. The limitation or our study is the relatively small number of patients. Thus, we could only aim to find highly penetrant pathogenic variants. In addition, exome sequencing does not enable the identification of non-coding variants that might affect splicing or gene expression.

While our exome sequencing approach failed to identify any clearly causal variants in the 12 families, we believe that the rare variants found in our cohort will be of interest for other dementia researchers. Thus, we presented all the variants and genes of potential interest in the hope this may be useful for future studies and can facilitate analyses in other families and datasets.

**Acknowledgements**

Professors M. Haltia and R. Sulkava are acknowledged for their valuable contribution in the collection of the families. This study was financially supported by the Päivikki and Sakari Sohberg Foundation (LM, PP), the Academy of Finland (LM), the Pirkko and Veikko Mäkelä foundation (PP), Sigrid Juselius Foundation (PJT) and Helsinki University Hospital (PJT, LM). This work was supported in part by the Alzheimer’s Society and Alzheimer’s Research UK. Jose Bras and Rita Guerreiro are supported by fellowships from the Alzheimer’s Society. To the memory of study nurse and friend Raija Ahlfors.

**References**


Legends to the figures

Figure 1. Schematic presentation of the study describing the workflow of genetic examinations. WES = whole exome sequencing.

Figure 2. Proportions of C9orf72 expansions in clinical frontotemporal dementia (FTD), Alzheimer’s disease (AD) and unspecified dementia in a cohort of 60 families. DLB = dementia with Lewy bodies.

Figure 3. Pedigrees of the families with rare variants verified by Sanger sequencing. DNA samples were available from individuals marked with an asterisk. APOE genotypes are also marked in the pedigree.

a. Family Fam-56 with the CLU p.(Thr203Ile) variant. Heterozygous variant (-/+), homozygous wild-type allele (-/-).

b. Family Fam-15 with the PCDH11X p.(Asp760Val) variant. Heterozygous variant (-/+), homozygous wild-type allele (-/-), hemizygous variant (+), hemizygous wild-type allele (-).

c. Family Fam-49 with the UNC13C p.(443del) variant. Heterozygous variant (-/+), homozygous wild-type allele (-/-).

d. Family Fam-13 with the MARCH4 p.(Lys211Glu) variant. Heterozygous variant (-/+), homozygous wild-type allele (-/-).

e. Family Fam-59 with the MARCH4 p.(Trp13Cys) variant. Heterozygous variant (-/+), homozygous wild-type allele (-/-).