The Genetics of Dementia

Janice L. Farlow, BS, BA¹ and Tatiana Foroud, PhD¹
¹Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana

Abstract

Over the past decade, there has been a dramatic evolution of genetic methodologies that can be used to identify genes contributing to disease. Initially, the focus was primarily on classical linkage analysis; more recently, genomewide association studies, and high-throughput whole genome and whole exome sequencing have provided efficient approaches to detect common and rare variation contributing to disease risk. Application of these methodologies to dementias has led to the nomination of dozens of causative and susceptibility genes, solidifying the recognition that genetic factors are important contributors to the disease processes. In this review, the authors focus on current knowledge of the genetics of Alzheimer's disease and frontotemporal lobar degeneration. A working understanding of the genes relevant to common dementias will become increasingly critical, as options for genetic testing and eventually gene-specific therapeutics are developed.

Keywords

genetic risk factors; dementia; Alzheimer's disease; frontotemporal lobar degeneration

Overview of Modern Genetic Technologies

The past few decades have seen exponential advances in the ability to identify genes contributing to disease. Beginning with the identification of microsatellite markers and then single nucleotide polymorphisms (SNPs) in the early 1990s, researchers were able to interrogate the genome to detect genes causing Mendelian disorders. Linkage analysis was the most common use of these markers and sought to identify chromosomal regions shared by affected members within a family. It was hypothesized that a gene that caused the disease lay within these shared chromosomal regions. The result of linkage analysis was a rapid identification of the genetic cause for many Mendelian disorders. This in turn has provided key insights into disease pathogenesis and in some cases has led to novel therapeutics.

Over the past decade, advancing genetic technologies have allowed millions of SNPs to be efficiently interrogated using genotyping arrays. These arrays include primarily common variation and have allowed scientists to explore new hypotheses about the role of genetics in
human disease. For example, studies can now be designed to identify genetic factors that increase or decrease the risk of disease. The most common approach that has been used is a genomewide association study (GWAS), which compares a sample of cases with disease to controls who do not have the disease. Because many genetic risk factors have modest individual effect on disease risk, large samples of cases and controls are needed to obtain robust, reproducible associations. More recently, meta-analysis has been widely used to combine results from multiple studies. This has resulted in studies using data from thousands of cases and controls to identify SNPs that are associated with disease risk. A GWAS is based on the premise of linkage disequilibrium, meaning that many SNPs are tested and used as proxies for the entire genome. Thus, analyses can determine that a SNP is associated with disease risk, but cannot confirm that this particular SNP functionally contributes to disease risk.

In the past few years, technology has further advanced to allow sequencing of DNA to be cost effective on a large scale. Rather than limiting sequencing to a gene or a small region of a chromosome, it is now possible to sequence the entire genome (whole genome sequencing) or only the gene-coding regions, the gene exons (whole exome sequencing). The ability to perform such extensive DNA sequencing has allowed researchers to examine the role of rare variation in disease. For example, whole exome sequencing within a family or across many families can be used to identify genes in which rare variants are found at greater frequency among the affected individuals as compared with control subjects who do not have the disease. Whole exome sequencing is often focused on the identification of rare, functional variants, which may alter the structure or function of the resulting protein. The types of variants that can be identified with this method are not limited to single nucleotide variants, but can also include small insertions and deletions.

Whole exome and whole genome sequencing presents a challenge to researchers due to the large number of sequence variants that are identified (~ 40,000 per exome and 3 million per genome). Large numbers of subjects are required to identify candidate genes in complex diseases like Alzheimer's disease (AD), and the cost is often prohibitive. New bioinformation technologies are constantly in development for analysis of sequencing experiments, but further effort will be required to validate, compare, and combine these tools. Almost all of these genetic tools and technologies have been used to study dementia. This review will describe what has been learned about the genetics of specific types of dementia, Alzheimer's disease, and frontotemporal lobar degeneration (FTLD).

The Genetics of Alzheimer's Disease

Alzheimer's disease is the most common neurodegenerative disorder among the elderly, as well as the most common cause of dementia overall. The key neuropathologic findings in AD are amyloid plaques composed of β-amyloid (Aβ) and neurofibrillary tangles consisting of the tau protein. Alzheimer's disease is likely caused by a combination of genetic and environmental factors. Alzheimer's disease is typically categorized into early onset (EOAD; onset < 60 years) and late-onset (onset ≥60 years). Early-onset Alzheimer's disease makes up less than 5% of cases, but because a subset of families has a clear pattern of Mendelian inheritance, mutations have been identified in several causative genes. The identification of
these genes has significantly advanced the understanding of the biological underpinnings of all AD cases.

**Early-Onset Alzheimer’s Disease**

Researchers observed that among families with EOAD, a significant number displayed autosomal dominant inheritance. This allowed linkage analysis to be successfully employed to identify three causative genes (APP, PSEN1, and PSEN2).

*APP* was the first EOAD gene to be identified.\(^1,2\) It is located on chromosome 21 and encodes the \(\beta\)-amyloid precursor protein. The \(\beta\)-amyloid precursor protein is a neuronal protein that is sequentially cleaved by two specific proteases (\(\beta\)-secretase and \(\gamma\)-secretase) to release A\(\beta\). Even before *APP* was identified, biochemical analysis had identified cerebral A\(\beta\) deposits in the brains of patients with AD. In addition, it had been found that patients with Down’s syndrome were more likely to develop AD; therefore, it was predicted that a gene on chromosome 21 would function to produce A\(\beta\) and would cause AD when mutated.\(^3\)

Mutations in *APP* are found in less than a quarter of EOAD patients. A variety of different mutations have been found in *APP* including point mutations, gene duplications, and gene deletions. All *APP* mutations reported to be pathogenic are located at secretase sites or inside the sequence for the resulting A\(\beta\) peptide, in or near exons 16 and 17 (Fig. 1).\(^4\)

The majority of EOAD mutations are found in the presenilin 1 (*PSEN1*) gene, located on chromosome 14 and discovered a few years after *APP*.\(^5\) There have been 185 different mutations found in this gene alone, most clustering in the transmembrane portion of the protein. Sequence homology with *PSEN1* prompted the relatively rapid identification of the presenilin 2 (*PSEN2*) gene on chromosome 1.\(^6,7\) Mutations in this gene are quite rare and were initially due to a founder mutation among individuals of Volga-German ancestry. Mutations in *PSEN2* have now been found in individuals of other ancestry as well.\(^8-10\) Of those that have been reported as pathogenic mutations, the distribution within resultant protein domains is similar to that of *PSEN1*. Together, presenilin 1 (*PSEN1*) and presenilin 2 (*PSEN2*) serve as critical catalytic components of the \(\gamma\)-secretase.\(^11\)

Over 200 different mutations have been identified in these three EOAD genes, and most are fully penetrant (Table 1; consult the Alzheimer Disease and Frontotemporal Dementia Mutation database, [http://www.molegen.ua.ac.be/ADMutations](http://www.molegen.ua.ac.be/ADMutations), for updated information).\(^12\) Disease-causing mutations in these three genes all lead to an increase in extracellular A\(\beta\) or a relative overabundance of the A\(\beta_{42}\) species, which is thought to increase A\(\beta\) aggregation and amyloid fibril formation.\(^13\) The discovery of *APP*, *PSEN1*, and *PSEN2* contributed to the amyloid hypothesis, which posits that amyloid plays a critical role in AD pathology.

**Late-Onset Alzheimer Disease**

Although most LOAD cases do not have a clear heritable etiology, genetic factors still play an important role in disease risk. The lifetime risk of developing AD is ~ 10%, and is twice that for individuals having an affected first-degree relative with AD.\(^14\) Studies have estimated that the genetic contribution to LOAD susceptibility (i.e., heritability) is as high as
80%. Of note, no single gene has been identified that has a causative effect on disease risk in LOAD, which would be analogous to the role of APP, PSEN1, or PSEN2 in EOAD.

The apolipoprotein E gene (APOE) was the first gene discovered as a risk factor for LOAD, and it remains the identified locus that confers the greatest susceptibility to the disease. APOE is a well-known gene involved in lipid metabolism, and it is thought to be important in Aβ clearance, as well. There are three isoforms of APOE (ε2, ε3, ε4) that differ by a single amino acid substitution. It has been shown that this minor difference leads to different levels of clearance of Aβ, with the ε4 allele of APOE (APOE4) significantly decreasing the speed of clearance at the blood–brain barrier.

The ε4 allele of APOE (APOE4) increases the risk of AD with reported odds ratios ranging from 4 to 15. The effect of APOE4 is consistently present across populations and ethnicities. The APOE4 effect is additive, with those having two APOE4 alleles having a 15-fold greater risk of AD as compared with the 4-times increased risk for those carrying a single copy of the risk allele. Meanwhile, the ε2 allele (APOE2) decreases the risk of AD by approximately a quarter, and is considered a protective allele. The ε3 allele of APOE (APOE3), the most common of the 3 APOE isoforms, has not been shown to have a detrimental or protective effect. Importantly, the APOE2 and APOE4 alleles modify the risk of AD, but neither is causative (i.e., sufficient to cause or protect against AD alone).

Other than APOE, almost 50 additional genetic risk factors have now been linked to AD. Varying levels of support exist for each gene. Nine genes have been labeled as established risk factors for AD, based on genome-wide significance (p value ≤ 5 × 10^-8) in meta-analyses and replication in multiple studies. Much work remains to characterize exact variants in or near these susceptibility genes; however, the genes have been linked to the metabolism of Aβ and synaptic function (APOE, BIN1, CD33, CLU, CR1, PICALM), the immune system and inflammation (ABCA7, CD33, CLU, CR1, MS4A), and lipid metabolism (ABCA7, APOE, BIN1, CLU, PICALM). These susceptibility genes, with the exception of APOE, all confer a relatively small risk for AD, with odds ratios of 0.85 to 1.2 (Table 2). The proportion of LOAD cases that could be prevented if these nine (GWAS-identified) genetic risk factors were eliminated, referred to as the cumulative population attributable risk, is estimated to be as high as 35%. Studies are actively underway to identify additional disease-related associations in LOAD.

**Genetics of Frontotemporal Lobar Degeneration**

Frontotemporal lobar degeneration (FTLD) is a term that describes the neuropathology underlying several different clinical syndromes, all characterized by progressive damage to the frontal and/or temporal lobes of the brain. As a result, behavior and language are progressively impaired in individuals affected with FTLD. Frontotemporal lobar degeneration typically presents at an earlier age than AD (~ 45–65 years of age) and represents the second most common cause of dementia in those under 65 years of age. Almost half of individuals with FTLD have other affected family members, and an autosomal dominant mode of transmission can be identified in ~ 10% of patients.
Six unrelated genes with over 150 reported mutations have been implicated in FTLD (Table 3; Alzheimer Disease and Frontotemporal Dementia Mutation database). Mutations in three genes (C9orf72, GRN, MAPT) are the most common, with rarer causative mutations found in the remaining three genes (CHMP2B, TARDBP, VCP). Individuals with mutations in C9orf72, GRN, or MAPT often present with the most common frontotemporal dementia clinical syndrome, the behavioral variant (bvFTD). Together, mutations in these three genes account for at least 17% of familial FTLD.27 Although there are strong correlations between the mutated gene and the patient's neuropathology, there is typically not a clear association between the implicated gene and clinical phenotype.27

Microtubule-associated protein tau (MAPT), located on chromosome 17, is involved in microtubule assembly and maintenance, and has been implicated in FTLD, AD, and other neurologic diseases.28–30 Over 40 mutations have been reported in the gene since its discovery in 1998, mostly in exons 9–13 (Fig. 2). Mutations result in a change in the ratio of the three amino acid (3R) and four amino acid (4R) isoforms and are believed to increase the tendency of tau to form neurotoxic aggregates. Mutations in MAPT are found in ~ 9 to 21% of FTD patients.31

Some FTLD families demonstrated genetic linkage to chromosome 17, but surprisingly no mutations in MAPT could be found. It was later discovered that many of these families had mutations in progranulin (GRN), located only 6 Mb from MAPT.32,33 Over 60 disease-producing mutations have been reported in GRN. Most result in a premature stop codon and the resulting mRNA is degraded through nonsense-mediated decay resulting in haploinsufficiency.34 Of note, there is substantial clinical variability in patients carrying a GRN mutation. Mutations have been reported among individuals diagnosed clinically with language variants and bvFTD, AD, corticobasal degeneration, and mild cognitive impairment. The age of onset has varied widely, even within families.35 GRN mutations are found in ~ 4 to 23% of all FTLD patients.31

Recently, expansions of a GGGGCC hexanucleotide repeat in the noncoding portion (first intron) of C9orf72 on chromosome 9 have been associated with both FTLD and amyotrophic lateral sclerosis (ALS).36–39 The repeat number, typically less than 25 units in normal individuals, is expanded to greater than 60 units in FTLD. The expanded number of repeats results in the loss of one alternatively spliced transcript. However, the function of this lost transcript is not yet known. Estimations for the percentage of familial ALS and familial FTLD cases in North American populations attributable to the C9orf72 mutation range from 23.5 to 36.2% and 6.3 to 19.1%, respectively.39 The C9orf72 mutation appears to be more common among northern European populations.35

Mutations in the other three genes, chromatin modifying 2B (CHMP2B), valosin-containing protein 1 (VCP-1) and TARDNA-binding protein 43 encoding gene (TARBDP), are rare causes of FTD, found in only a few families each. The pheno-type in each is quite different and may also be associated with non-FTD phenotypes. For example, some VCP-1 mutations are associated with Paget's disease of the bone.35

Semin Neurol. Author manuscript; available in PMC 2014 November 14.
There has only been one systematic FTLD GWAS, which reported \textit{TMEM106B} on chromosome 7 as a potential susceptibility gene for FTLD cases with TAR-DNA-binding protein inclusions.\textsuperscript{40} The identified SNPs were associated with increased \textit{TMEM106B} expression, but the mechanism of upregulation and the development of FTLD has not been established.

**Genetic Counseling and Testing**

Definitive diagnoses for AD and FTD can only be made at autopsy, but genetics may be used to increase diagnostic and prognostic accuracy. With the current limitations regarding disease-modifying therapies for neurodegenerative diseases and the complex ethical considerations involved in genetic testing, clinicians must carefully consider the appropriateness of genetic analyses for patients. Because both positive and negative findings have significant implications for the patient and family, thorough genetic counseling is advised before and after any genetic testing is performed. For genetic tests that are not commercially available, the tests may be available in the academic research setting. As more is learned about the genetics of AD and FTLD, commercial availability of genetic testing, and even direct to consumer testing, will likely increase. As treatments are developed, testing options will likely expand further. Clinicians will need to be familiar with these options in an effort to counsel patients.

At present, routine DNA testing in AD is advisable only for EOAD, if the presenting affected individual is a young or middle-aged adult (onset < 60 years). Testing for mutations in \textit{PSEN 1}, \textit{PSEN 2}, and \textit{APP} is currently commercially available. \textit{APOE} genotyping is also available, and there are ongoing studies assessing the advantages and disadvantages of \textit{APOE} genotyping. Most groups do not currently recommend the test for predictive risk assessment.\textsuperscript{41}

With the heterogeneity of FTLD, the recommendations for genetic testing in symptomatic patients and their families are more complex. Considerations include the type of FTLD, the extent of family history, and the availability and findings of autopsy tissue from an affected family member. Genetic testing is available for \textit{MAPT}, \textit{GRN}, \textit{VCP}, and \textit{C9orf72} in the United States. A full algorithm for when genetic testing should be recommended has been previously reported.\textsuperscript{42}

**Conclusion and Outlook**

Although much has been discovered about the genetics of dementia, the genetic contribution to the vast majority of sporadic and some Mendelian cases of dementia has not yet been elucidated. Modern genetic technologies have helped identify causative and risk loci, but much work remains to characterize the biological mechanisms of these genes and the disease-producing mutations. As additional genetic loci are recognized and characterized, correlations between genotypes and phenotypes will become better understood. Phenotypes will not only include presenting symptoms, but also biomarkers such as structural and functional imaging and composition of the cerebrospinal fluid. Methodologies for screening, diagnosis, and prognosis will improve as additional genes and mutations are identified and correlated with increasingly specific phenotypes. Most importantly, advancing knowledge of
underlying genetic factors in neurodegenerative dementias will inform the pursuit of preventative and disease-modifying therapies.

Acknowledgments

This publication was made possible by Grant Number TL1 TR000162 (A. Shekhar, PI) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award, as well as the National Cell Repository for Alzheimer's Disease, which receives government support under the cooperative agreement grant U24 AG21886.

References

23. Bertram L, Tanzi RE. Thirty years of Alzheimer's disease genetics: the implications of systematic
22266883]
[PubMed: 21460841]
27. Sieben A, Van Langenhove T, Engelborghs S, et al. The genetics and neuropathology of
22890575]
29. Poorkaj P, Bird TD, Wijsman E, et al. Tau is a candidate gene for chromosome 17 frontotemporal
30. Spillantini MG, Murrell JR, Goedert M, Farlow MR, Klug A, Ghetti B. Mutation in the tau gene in
familial multiple system tauopathy with presenile dementia. Proc Natl Acad Sci U S A. 1998;
31. Cohn-Hokke PE, Elting MW, Pijnenburg YA, van Swieten JC. Genetics of dementia: update and
16862116]
[PubMed: 16862115]
34. Gass J, Cannon A, Mackenzie IR, et al. Mutations in progranulin are a major cause of ubiquitin-
16950801]
[PubMed: 22536193]
repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron.
Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amytrophic
[PubMed: 22154785]


Fig. 1.
Location of mutations within the *APP* protein product as generated from the Alzheimer Disease and Frontotemporal Dementia Mutation database.\(^\text{12}\) Depicted is the chain of amino acids produced from the *APP* gene. Green indicates sites of reported nonpathogenic mutations, and red indicates sites of reported pathogenic mutations. Sites of cleavage by secretases are shown in blue.
Fig. 2.
Location of mutations within the MAPT protein product as generated from the Alzheimer Disease and Frontotemporal Dementia Mutation database. Depicted is the chain of amino acids produced from the MAPT gene. Green indicates sites of reported nonpathogenic mutations, and red indicates sites of reported pathogenic mutations. Exon boundaries are marked by Ex*, and microtubule-binding domains are shown in blue.
Table 1

Mendelian genes causing familial early-onset Alzheimer's disease (EOAD)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Chromosome location</th>
<th>Mutation frequency in EOAD patients(^a)</th>
<th>Types of mutations reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>APP</td>
<td>B-amyloid precursor protein</td>
<td>21q21.2</td>
<td>6–26%</td>
<td>Missense, duplications, deletions</td>
</tr>
<tr>
<td>PSEN1</td>
<td>Presenilin 1</td>
<td>14q24.3</td>
<td>28–61%</td>
<td>Missense, deletions, insertions</td>
</tr>
<tr>
<td>PSEN2</td>
<td>Presenilin 2</td>
<td>1q31–42</td>
<td>Rare</td>
<td>Missense</td>
</tr>
</tbody>
</table>

\(^a\)Mutation frequencies reported by Cohn-Hokke et al.\(^{31}\)
Table 2

Established risk loci for late-onset Alzheimer’s disease

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Chromosome location</th>
<th>Effect size / Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA7</td>
<td>ATP-binding cassette, subfamily A, member 7</td>
<td>19p13.3</td>
<td>1.23 (1.18–1.28)</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
<td>19q13.32</td>
<td>ε3ε4 genotype 3.2 (2.8–3.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ε4ε4 genotype 14.9 (10.8–20.6)</td>
</tr>
<tr>
<td>BIN1</td>
<td>Bridging integrator 1</td>
<td>2q14.3</td>
<td>1.17 (1.13–1.20)</td>
</tr>
<tr>
<td>CD2AP</td>
<td>CD2-associated protein</td>
<td>6p12.3</td>
<td>1.12 (1.08–1.16)</td>
</tr>
<tr>
<td>CD33</td>
<td>CD33 molecule</td>
<td>19q13.41</td>
<td>0.85 (0.86–0.92)</td>
</tr>
<tr>
<td>CLU</td>
<td>Clusterin</td>
<td>8p21.1</td>
<td>0.89 (0.86–0.91)</td>
</tr>
<tr>
<td>CR1</td>
<td>Complement component receptor 1</td>
<td>1q32.2</td>
<td>1.19 (1.09–1.30)</td>
</tr>
<tr>
<td>MS4A4E, MS4A6A</td>
<td>Membrane-spanning 4-domains, subfamily A,</td>
<td>11q12.2</td>
<td>0.90 (0.88–0.93)</td>
</tr>
<tr>
<td></td>
<td>member 4E and 6A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PICALM</td>
<td>Phosphatidylinositol binding clathrin assembly protein</td>
<td>11q14.2</td>
<td>0.88 (0.86–0.91)</td>
</tr>
</tbody>
</table>

*Effect sizes calculated from AlzGene meta-analysis.*

---

Semin Neurol. Author manuscript; available in PMC 2014 November 14.
### Table 3

Mendelian genes causing frontotemporal lobar degeneration (FTLD)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Location</th>
<th>Mutation frequency in FTLD patients&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Types of mutations reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPT</td>
<td>Microtubule-associated protein tau</td>
<td>17q21.1</td>
<td>9–21%</td>
<td>Missense, deletions</td>
</tr>
<tr>
<td>CRN</td>
<td>Programin</td>
<td>17q21.31</td>
<td>4–23%</td>
<td>Missense, deletions, insertions</td>
</tr>
<tr>
<td>C9orf72</td>
<td>Chromosome 9 open reading frame 72</td>
<td>9p21</td>
<td>18–30%</td>
<td>Hexanucleotide repeat expansion</td>
</tr>
<tr>
<td>TARDBP</td>
<td>TAR DNA binding protein</td>
<td>1p36.22</td>
<td>Rare</td>
<td>Missense</td>
</tr>
<tr>
<td>VCP</td>
<td>Valosin-containing protein</td>
<td>9p13.3</td>
<td>Rare</td>
<td>Missense</td>
</tr>
<tr>
<td>CHMP2B</td>
<td>Charged multivesicular body protein 2B</td>
<td>3p11.2</td>
<td>Rare</td>
<td>Missense</td>
</tr>
</tbody>
</table>

<sup>a</sup> Mutation frequencies calculated by Cohn-Hokke PE et al.\textsuperscript{31}

*Semin Neurol.* Author manuscript; available in PMC 2014 November 14.