Alzheimer’s Disease Genetics: From the Bench to the Clinic

Celeste M. Karch, PhD,¹ Carlos Cruchaga, PhD,¹ and Alison Goate, DPhil²

¹Department of Psychiatry and Hope Center for Neurological Disorders
Washington University School of Medicine
St. Louis, Missouri

²Department of Neuroscience
Icahn School of Medicine at Mount Sinai
New York, New York
Introduction
Alzheimer's disease (AD) is defined clinically by a gradual decline in memory and other cognitive functions, and neuropathologically by gross atrophy of the brain and the accumulation of extracellular amyloid plaques and intracellular neurofibrillary tangles. Genetic, biochemical, and neuropathological data suggest that aggregation of \( \beta \)-amyloid (A\( \beta \)) is central to initiating AD pathogenesis (Hardy and Selkoe, 2002). A\( \beta \) is a proteolytic fragment of the amyloid precursor protein (APP), generated as a result of sequential cleavage by \( \beta \)- and \( \gamma \)-secretases. Deposition of extracellular amyloid plaques is followed by the accumulation of neurofibrillary tangles in neuronal cell bodies and associated processes. Neurofibrillary tangles are composed of hyperphosphorylated tau aggregates. The presence of neurofibrillary tangles in AD brains is strongly correlated with neuronal dysfunction and disease progression (Holtzman et al., 2011). The amyloid cascade hypothesis posits that changes in APP and/or A\( \beta \) homeostasis lead to the aggregation of A\( \beta \) and deposition in plaques and that these events are sufficient to initiate the cascade of pathological and clinical changes associated with AD, including the aggregation of tau protein in neurofibrillary tangles (Hardy, 1997).

Linkage Studies
Dominantly inherited mutations in \( \beta \)-amyloid precursor protein (APP), presenilin 1 (PSEN1), and presenilin 2 (PSEN2) cause early-onset AD. These genes, as well as APOE, were identified through genetic linkage studies in families.

APP
APP encodes a ubiquitously expressed type I transmembrane protein. The majority of APP is proteolyzed by \( \alpha \)- and \( \gamma \)-secretases, leading to cleavage of APP within the A\( \beta \) domain. The result is nonpathogenic fragments: sAPP\( \alpha \) and \( \alpha \)-C-terminal fragment (CTF). Alternatively, APP can be cleaved through sequential proteolytic cleavage by \( \beta \)- and \( \gamma \)-secretases to generate A\( \beta \) peptides: sAPP\( \beta \), and \( \beta \)-CTF. Cell surface APP is internalized, allowing A\( \beta \) to be generated in the endocytic pathway and secreted into the extracellular space. Dominant mutations in APP account for approximately 14% of early-onset autosomal dominant cases of AD, and more than 30 mutations have been described (Alzheimer Disease & Frontotemporal Dementia Mutation Database, n.d.). Two recessive APP mutations, A673V and E693\( \Delta \), also reportedly cause early-onset AD. The majority of mutations in APP cluster in the region that is adjacent to or within the A\( \beta \) domain.

Mutations in APP have revealed many important aspects of the molecular mechanisms underlying AD pathogenesis. The Swedish mutation (KM670/671NL) increases plasma A\( \beta \) levels by twofold to threefold by altering \( \beta \)-secretase cleavage efficiency. Duplications of APP and the surrounding sequence are also associated with early-onset AD. Families carrying these duplications exhibit classic AD neuropathology and cerebral amyloid angiopathy. Additionally, individuals with Down syndrome, which results from trisomy of chromosome 21, develop AD neuropathology. Individuals with partial trisomy of chromosome 21, which does not include the APP gene, fail to develop AD neuropathology. Thus, excess A\( \beta \) production is sufficient to cause AD. Several APP mutations cluster at or after the C-terminal portion of the A\( \beta \) domain. These mutations alter \( \gamma \)-secretase function, leading to a shift in APP processing that increases the highly amyloidogenic A\( \beta \)42 fragment while reducing the A\( \beta \)40 fragment. The result is an altered A\( \beta \)42/A\( \beta \)40 ratio without a change in total A\( \beta \) levels (Bergmans and De Strooper, 2010). Because A\( \beta \)42 is more prone to aggregate than A\( \beta \)40, these findings suggest that A\( \beta \) aggregation is a critical component of AD pathogenesis. In contrast, mutations such as the Arctic mutation (E693G) and the Dutch mutation (E693Q) that occur within the A\( \beta \) domain likely increase the aggregation rate of the mutant peptide. Individuals carrying these mutations develop hereditary cerebral hemorrhage with amyloidosis, which is characterized by predominant vascular A\( \beta \) deposition with diffuse plaques in the parenchymal tissue. These mutations provide further evidence that A\( \beta \) aggregation is a critical process in AD pathogenesis. Genetic changes that lead to altered APP processing and A\( \beta \) accumulation may produce variable neurological and neurovascular phenotypes.

PSEN1 and PSEN2
PSEN1 and PSEN2 are critical components of the \( \gamma \)-secretase complex, which cleaves APP into A\( \beta \) fragments, and localize in the endoplasmic reticulum and Golgi apparatus. As many as 185 dominant, pathogenic mutations have been identified in PSEN1, accounting for approximately 80% of autosomal dominant AD (ADAD) cases; 13 pathogenic mutations have been identified in PSEN2, accounting for approximately 5% of ADAD cases (Alzheimer Disease & Frontotemporal Dementia Mutation Database, n.d.).
Dementia Mutation Database, n.d.). PSEN1 and PSEN2 mutations are distributed throughout the protein, with some clustering occurring in the transmembrane domains (Guerreiro et al., 2010). Dominantly inherited mutations in PSEN1, as well as APP, have also been identified in late-onset AD cases with a strong family history of disease. These families may carry additional genetic variants that delay age at onset of the normally fully penetrant disease mutation. Mutations in PSEN1 and PSEN2 alter the proteolytic site preference of γ-secretase. Whereas wild-type γ-secretase cleaves APP to generate more Aβ40 than Aβ42, mutations in the γ-secretase components lead to enhanced production of the Aβ42 fragment. In turn, because Aβ42 is highly aggregation-prone, the increase in the Aβ42 fragment is likely sufficient to cause AD.

Apolipoprotein E

Apolipoprotein E (APOE) encodes a pleiotropic glycoprotein that is highly expressed in liver, brain, and macrophages, where it plays a role in mobilization and redistribution of cholesterol. APOE has also been implicated in neuronal growth and repair, nerve regeneration, immune response, and activation of lipolytic enzymes (Mahley and Rall, 2000). APOE occurs as three isoforms that differ at two amino acid residues (112 and 158): APOEε2, APOEε3, and APOEε4. APOEε3 is the most common APOE isoform, occurring in approximately 72% of the population. Family-based methods originally identified a genetic linkage between AD and the region of chromosome 19 that contains the APOE gene. APOEε4 increases risk in familial and sporadic early-onset and late-onset AD, increasing risk threefold for heterozygous carriers and increasing risk 8-fold to 10-fold for homozygous carriers (Farrer et al., 1997). APOEε4 also has a dose-dependent effect on age at onset. Interestingly, APOEε2 decreases the risk for late-onset AD and delays age at onset (Corder et al., 1994). APOE binds to Aβ, influencing the clearance of soluble Aβ and Aβ aggregation (Castellano et al., 2011). APOE4 binds to Aβ more rapidly than does APOE3, resulting in accelerated fibril formation. APOE also regulates Aβ metabolism indirectly by interacting with low-density lipoprotein receptor-related protein 1 receptors. In vitro, APOE influences the amount and structure of intraparenchymal Aβ deposits in an isoform-dependent manner. Thus, the major risk gene associated with AD likely influences Aβ metabolism as a mechanism of pathogenicity.

Genome-Wide Association Studies

Only 50% of individuals with AD carry an APOEε4 allele, and only approximately 2% carry a pathogenic mutation, suggesting that other genetic factors must contribute to risk for the disease. The first genome-wide association study (GWAS) for AD to use thousands of AD cases and elderly nondemented controls successfully generated replicable associations for several new genetic risk factors, including clusterin (CLU), phosphatidylinositol-binding clathrin assembly protein (PICALM), complement receptor 1 (CR1), and bridging integrator protein 1 (BIN1) (Harold et al., 2009; Lambert et al., 2009; Seshadri et al., 2010). Two subsequent studies that each included more than 8000 cases and a similar number of controls identified four additional loci with genome-wide significant evidence for association (<5 × 10⁻⁴) (Hollingworth et al., 2011; Naj et al., 2011) (Fig. 1).

Recently, a meta-analysis of GWAS data from 74,046 individuals from four large consortia confirmed these associations and reported 11 new susceptibility loci for AD (Lambert et al., 2013) (Fig. 1). For most of these loci, the specific functional variants and genes remain to be identified. These studies employed samples of individuals of European descent, but two additional GWASs examined African American and Asian (Korean and Japanese) case-control series. In addition to APOE, the most significant finding in the African-American study of 5896 cases and controls was of single nucleotide polymorphisms (SNPs) in ABCA7, which reached genome-wide significance and exhibited a somewhat larger effect than in European samples (Reitz et al., 2013). The study in Japanese and Korean subjects used a multistage strategy to identify a novel locus, SORL1, which was replicated in a large European American cohort (Lambert et al., 2013; Miyashita et al., 2013). Together, these studies demonstrate that some of these loci show similar effects across populations (e.g., BIN1), while others appear to have a bigger impact in some populations (e.g., ABCA7). As larger datasets become available for other populations, a more complete picture of the population-specific loci and those that are shared across populations will more fully resolve.

The common SNPs identified in these GWASs alter risk by 10–15%, suggesting that the effect of these risk alleles is much smaller than that of APOEε4, unless they are tagging rare alleles of larger effect. For the most part, the functional alleles responsible for each of these associations remain to be determined.
For some of the GWAS loci, several genes are found within the associated region, any one of which may contain the functional risk variant. However, many of these AD risk loci have putative functions in the immune system (CLU, CR1, ABCA7, CD33, and EPHA1); four are involved in processes at the cell membrane, including endocytosis (PICALM, BIN1, CD33, and CD2AP); and three are involved in lipid biology (APOE, CLU, and ABCA7) (Fig. 1). Although some of the new genes appear to be involved in Aβ metabolism (e.g., CLU and PICALM), the fact that others likely influence inflammation, endocytosis, and lipid biology suggests that targeting specific components of these pathways may lead to novel directions for drug discovery and treatment.

Sequencing Studies
The most recent GWAS for late-onset AD combined several large datasets and identified more than 20 loci associated with AD risk. However, each of these new loci account for only a small percentage of the genetic variance that contributes to AD susceptibility. Thus, a large proportion of the genetic heritability for AD has not been explained by GWASs. It has been hypothesized that low-frequency (minor allele frequency [MAF] 1–5%) and rare variants (MAF < 1%) might explain the missing genetic heritability. As was the case with linkage studies and GWASs, initial sequencing efforts focused on candidate genes in small populations until the arrival of next-generation sequencing technology, which allowed for low-cost, high-throughput sequencing and ushered in the era of whole-exome and whole-genome sequencing. Even with these new advances, the first studies using this technology also focused on the pathogenic AD genes. Subsequent genome-wide studies have led to the identification of novel variants and genes.

One powerful approach to identifying AD risk variants with large effect sizes has been to study families with a strong history of late-onset AD. Studies sponsored
by the National Institutes of Mental Health and the National Institute on Aging have invested in the collection and characterization of large family datasets with at least two late-onset AD cases. Large sequencing studies in these cohorts have shown that mutations in APP, PSEN1, and PSEN2 also cause AD in 1.5–2% of these cases (Cruchaga et al., 2012). In addition to identifying known and likely pathogenic mutations in APP and PSEN1, these studies have revealed variants that may modify AD risk. For example, the PSEN2 variants R62H and R71W were found to be associated with a lower age at onset, between 5 and 7 years earlier than in noncarriers (Benitez et al., 2013). PSEN1 E318G (rs17125721) was previously classified as nonpathogenic because it does not segregate according to disease status in some families, but analyses in large case-control series have demonstrated that in APOEε4 carriers, PSEN1-E318G is associated with a 10-fold increased risk of developing AD and an earlier age at onset of AD (Benitez et al., 2013). Together, these results suggest that some variants in PSEN1 and PSEN2 are risk factors for AD rather than fully penetrant, causative mutations.

**Protective APP Variants**

It is well known that APOEε4 is associated with increased AD risk and that APOEε2 decreases AD risk, demonstrating that different variants within the same gene can have opposing effects on disease risk. However, most genetic studies of AD have focused on identifying variants that increase AD risk. Jonsson et al. reported that rs63750847 (APP A673T) was associated with reduced risk of AD (odds ratio = 0.23) in the Icelandic population (Jonsson et al., 2012). Interestingly, this mutation occurs at the same amino acid residue as a recessive mutation reported to cause early-onset AD (APP A673V). This mutation is near the proteolytic cleavage site of BACE1 (position 2 in the Aβ peptide) and results in impaired BACE1 cleavage of APP in the A673T carriers and reduction of Aβ40 and Aβ42 in vitro (Jonsson et al., 2012). This study provides a proof of principle for the hypothesis that reducing the β-cleavage of APP may protect against AD.

**New AD Risk Genes**

**TREM2**

Using whole-exome sequencing and whole-genome sequencing strategies, two groups simultaneously reported a low-frequency variant (TREM2 R47H; rs75932628) in triggering receptor expressed on myeloid cells 2 protein (TREM2) that was associated with a twofold to threefold increase in AD risk (Guerreiro et al., 2013; Jonsson et al., 2013). Guerreiro et al. described several rare variants in TREM2 that were observed more frequently in cases than in controls, suggesting that there may be multiple rare variants in TREM2 that increase risk for AD. This theory is supported by data from several subsequent studies in European and African American populations. Some studies also suggest that TREM2 R47H could be associated with Parkinson’s disease, frontotemporal dementia, and amyotrophic lateral sclerosis, but this association remains controversial. In previous studies, rare homozygous loss-of-function mutations in TREM2 were associated with an autosomal recessive form of early-onset dementia: polycystic lipomembranous osteodysplasia with sclerosis, leukoencephalopathy (Paloneva et al., 2003). Subsequent studies found TREM2 mutations in three patients with frontotemporal-like dementia without any bone-associated symptoms (Guerreiro et al., 2013). These observations highlight a recurring theme in neurodegenerative diseases: Heterozygous and homozygous mutations in the same gene can lead to clinically distinct disorders.

TREM2 is a type 1 transmembrane receptor protein expressed on myeloid cells, including microglia, monocyte-derived dendritic cells, osteoclasts, and bone marrow–derived macrophages. In the brain, TREM2 is expressed primarily on microglia and has been shown to control two signaling pathways: regulation of phagocytosis and suppression of inflammation reactivity. TREM2 deficiency and haploinsufficiency in a mouse model of β-amyloid deposition augment Aβ accumulation owing to a dysfunctional response of microglia, which fail to cluster around Aβ plaques and become apoptotic (Wang et al., 2015). This study also demonstrated that TREM2 senses a broad array of anionic and zwitterionic lipids known to associate with fibrillar Aβ in lipid membranes and to be exposed on the surface of damaged neurons. Remarkably, the R47H risk factor for AD impairs TREM2 detection of lipid ligands. Thus, TREM2 detects damage-associated lipid patterns associated with neurodegeneration, sustaining the microglial response to Aβ accumulation. The identification of coding variants in TREM2 that increase the risk for AD supports the role of the immune response and inflammation in AD pathogenesis. Interestingly, coding variants in other genes in the TREM gene family may also modify AD disease risk, as is the case for a common protective variant in TREML2 (Benitez et al., 2014).
PLD3, ABCA7, UNC5C, AKAP9, and SORL1

Applying a family-based design to identify AD risk genes has revealed several rare, missense variants in phospholipase D3 (PLD3), ATP-binding cassette subfamily 1, member 7 (ABCA7), Netrin receptor gene (UNC5C), A kinase anchor protein 9 (AKAP9), and sortilin-related receptor, L (DLR class) A repeats (SORL1) in individual studies. ABCA7 and SORL1 had previously been implicated by GWAS, which suggests that both common and rare variants of these genes modulate risk. Several loss-of-function alleles have been reported in ABCA7 that are found at higher frequency in AD cases compared with controls, implicating ABCA7 function in protection against developing AD. For PLD3, UNC5C, and AKAP9, additional studies in large cohorts will be necessary to confirm whether or not these genes influence the risk for AD.

Disease Mechanisms Implicated by Studies of Genetic Risk Factors for AD

Early-onset AD mutations in APP, PSEN1, and PSEN2 lead to altered production or ratios of Aβ isoforms in the brain, while APOE influences Aβ clearance and aggregation; both observations support the hypothesis that Aβ levels are critical for disease pathogenesis (Hardy, 1997). GWASs have now identified polymorphisms in or near more than 20 genes that are associated with AD risk. The identification of common variants that have small effects on AD risk has created a broader picture of the processes and pathways involved in AD risk, including lipid metabolism, the inflammatory response, and endocytosis (Fig. 1). Whole-genome and exome-sequencing studies have also identified risk alleles in TREM2, ABCA7, UNC5C, SORL1, and PLD3. The identification of rare variants in the population that have moderate-to-large effects on AD risk will be most valuable for identifying pathways that are central to AD pathogenesis.

Future Directions for Genetic Studies

Substantial progress has been made during the past five years toward understanding the genetic architecture of AD because of the technological advances in genotyping and sequencing that have made it feasible to genotype or sequence thousands of individuals. GWASs have now identified more than 20 loci that influence risk for AD. It is clear from these studies that, with the exception of APOE, common risk alleles for AD have modest effects on risk individually but point to a number of important disease pathways. In other disorders in which even larger samples have been genotyped, more risk loci have been identified, suggesting that there is some value to continuing to increase the number of samples with GWAS data. Whereas very large sample sizes have been examined for people of European descent, this is not true for other populations. Large GWASs in these populations may identify genes not detected in Europeans. In addition, other AD-related phenotypes, including rate of progression of AD, rate of cognitive decline, and age at onset of AD, have yet to be examined in the largest datasets.

An alternative approach to increasing sample size is to use gene-based or network-based analyses to identify the genes that, while not of genome-wide significance, are overrepresented among SNPs, which have low p values in GWASs. This approach has identified additional significant genes that further implicate immune regulation, energy metabolism, and protein degradation in AD risk (Escott-Price et al., 2014). Another approach has compared GWAS data across related disorders to identify common underlying genes that demonstrate a genetic link between cardiovascular disease and AD risk. A remaining challenge for the field is to follow up the GWAS loci to identify the specific functional alleles and mechanisms underlying the GWAS signals. We anticipate that similar success will be observed as large-scale sequencing projects begin to yield results. Early studies in small discovery datasets have already identified several promising genes that modify AD risk. So far, these studies have focused largely on whole-exome sequencing and have identified coding variations that raise or lower the risk for AD by as much as twofold to threefold. This has the advantage of making functional follow-up of the genes identified through sequencing substantially easier that those identified in GWAS.

Several groups are performing whole-exome or whole-genome sequencing in unrelated AD cases and controls, and in families with multiple members affected by AD. It is anticipated that whole-exome or whole-genome sequencing will be available on more than 20,000 well-characterized samples before the end of 2015. Both the GWAS data and the sequencing data will be made available through public repositories such as dbGaP (the database of
Genotypes and Phenotypes, available at http://www.ncbi.nlm.nih.gov/gap). Combining these data with other large datasets, such as transcriptomics from human brain tissue, immune cells, or induced pluripotent stem cell-derived neurons and glia, will provide further insight into the disease pathways, and regulatory nodes within these pathways, that may provide druggable targets for future therapies. An early application of this integrative network-based approach used brain tissue from more than 1600 AD cases and nondemented individuals to identify an immune-specific and microglia-specific module that is dominated by genes involved in pathogen phagocytosis (Forabosco et al., 2013; Zhang et al., 2013). This module contains TREM2 (identified by whole-exome sequencing) but shows that TYROBP (aka DAP12, which binds to TREM2) is a key regulator of the network and is upregulated in AD brains. Mouse microglia overexpressing intact or truncated TYROBP reveal expression changes that significantly overlap with the human brain TYROBP network. This causal network structure was able to predict responses to gene perturbations, and thus presents a useful framework for testing models of disease mechanisms underlying AD.

Acknowledgments
C.M.K. is supported by a Career Development Award from the National Institutes of Health (NIH) (Grant K01 AG046374); A.M.G. received support from the Ronald M. Loeb Center for Alzheimer's Disease and the NIH (Grant AG035083); and C.C. received support from the NIH (Grant AG044546). This paper was adapted from Karch CM, Cruchaga C, Goate AM (2014) Alzheimer's disease genetics: from the bench to the clinic. Neuron 83:11–26.

References
Alzheimer Disease & Frontotemporal Dementia Mutation Database (n.d.). Available at http://www.molgen.vib-ua.be/ADMutations


