Molecular Genetics of Alzheimer's Disease

Peter H. St George-Hyslop

Application of genetic paradigms to Alzheimer's disease (AD) has led to confirmation that genetic factors play a role in this disease. Additionally, researchers now understand that AD is genetically heterogeneous and that some genetic isoforms appear to have similar or related biochemical consequences. Genetic epidemiologic studies indicate that first-degree relatives of AD probands have an age-dependent risk for AD ≅38% by age 90 years (range 10% to 50%). This incidence strongly suggests that transmission may be more complicated than a simple autosomal dominant trait. Nevertheless, a small proportion of AD cases with unequivocal autosomal dominant transmission have been identified. Studies of these autosomal dominant familial AD (FAD) pedigrees have thus far identified four distinct FAD genes. The β-amyloid precursor protein (βAPP) gene (on chromosome 21), the presenilin 1 (PS1) gene (on chromosome 14), and the presenilin 2 (PS2) gene (on chromosome 1) gene are all associated with early-onset AD. Missense mutations in these genes cause abnormal βAPP processing with resultant overproduction of $A\beta_{42}$ peptides. In addition, the ϵ4 allele of apolipoprotein E (APOE) is associated with a increased risk for late-onset AD. Although attempts to develop symptomatic treatments based on neurotransmitter replacement continue, some laboratories are attempting to design treatments that will modulate production or disposition of Aβ peptides. Biol Psychiatry 2000;47: 183–199 © *2000 Society of Biological Psychiatry*

Key Words: Presenilins, βAPP, secretases, Aβ peptides

Overview

Investigations into Alzheimer's disease (AD) can be used to demonstrate how molecular genetics has been applied to a complex disease with multiple genetic etiologies. Preliminary data on the biology of the genetic etiologies seem to suggest that the genes identified to date act through a common biochemical pathway, which has become the target of attempts to develop therapies and to develop animal models that will further elucidate the biochemistry of this disease. Moreover, the molecular genetic tools that have been discovered can be applied to the clinical setting. There are some very clear differences between AD, on the one hand, and schizophrenia and bipolar illness, on the other. In contrast to the latter two diseases, in AD there are severe pedigrees with unambiguous simple autosomal dominant transmission. Furthermore, AD has a very clear neuropathology, which includes the presence of intraneuronal neurofibrillary tangles (Figure 1), hyperphosphorylated forms of a microtubule-associated protein known as *tau*. More importantly, there is the presence in the extracellular space of amyloid plaques (Figure 2) composed of a number of proteins, principal among which is a 40–42 amino acid peptide known as the amyloid β peptide (Aβ). The Aβ itself is a proteolytic derivative of a longer protein known as the β amyloid precursor protein (βAPP).

The presence of neurofibrillary tangles and amyloid plaques as hallmarks makes diagnosis somewhat easier than in some other disorders. These pathological attributes are also advantageous in that they provide biological clues that can be used to supplement the molecular genetics. In other words, they suggest likely candidate genes. The clinical symptoms of AD arise from neuronal dysfunction and probably also neuronal death, resulting from perturbation in intermediary biochemical processes due, in turn, to environmental or genetic causes or to a combination of both.

It appears from a number of genetic epidemiology studies that about 30–50% of the population risk for AD can be attributed to genetic factors (Breitner et al 1988; Farrer et al 1991; Heyman et al 1984; Katzman et al 1994; Lautenschlager et al 1996; Rocca et al 1986; Silverman et al 1994). Furthermore, the risk of developing AD increases with age in first-degree relatives of patients with AD (Rocca et al 1986). The most comprehensive recent study suggests an age-dependent risk curve asymptotic to a final risk of 38% by age 85 years (Lautenschlager et al 1996). The latter study, as well as several other earlier epidemiologic studies, makes it difficult to assign a pure Mendelian mode of transmission in the majority of AD cases. Instead, these studies imply that the majority of cases of familially aggregated AD probably reflect a complex mode of transmission, such as one or more common independent, but incompletely penetrant, single autosomal gene defects; or a mode of transmission in which genetic and environmental factors interact (Farrer et al 1991). Nevertheless, in a small proportion of AD cases (~10%), the disease appears to be transmitted as a pure

From the Centre for Research in Neurodegenerative Disease, Department of Medicine, University of Toronto and the Department of Medicine (Neurology), The Toronto Hospital, Toronto, Canada.

Address reprint requests to P.H. St George-Hyslop, MD, PhD, FRCP(C), Centre for Research in Neurodegenerative Disease, Department of Medicine, University of Toronto, Tanz Neuroscience Building, 6 Queen's Park Crescent, Toronto, Ontario, Canada M5S 1A8.

Received September 20, 1999; revised December 1, 1999; accepted December 1, 1999.

0006-3223/00/$20.00
PII S0006-3223(99)00301-7

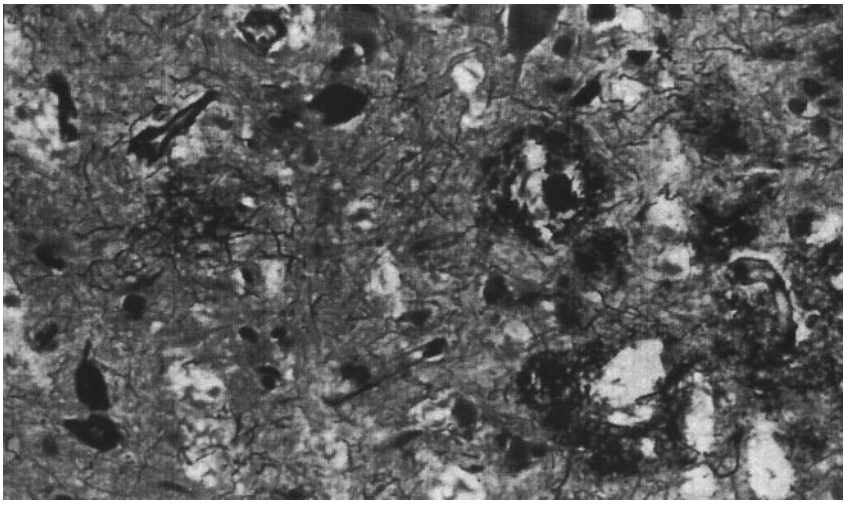

Figure 1. Intraneuronal fibrillary tangles characteristic of Alzheimer's disease contain hyperphosphorylated forms of the microtubule-associated protein known as *tau*.

autosomal dominant Mendelian trait with age-dependent but high penetrance.

Analysis of such pedigrees has led to the discovery of four different genetic loci (St George-Hyslop et al 1990; Figure 3) associated with inherited susceptibility to AD that are reviewed here. Of the four genetic loci that have been identified as being the genetic determinants of AD, three of them are associated predominately with early-onset forms of autosomal dominant familial AD. Apolipoprotein E (APOE) and particularly the $\epsilon 4$ allele of APOE on chromosome 19 have been associated with later onset forms of AD. The four genes themselves do not account for all of the genetic factors; other genes remain to be identified.

The Amyloid Precursor Protein

The βAPP gene, the first AD susceptibility gene to be identified, encodes a transmembrane protein that is glycosylated and in its longest isoform contains 770 amino acids (Goldgaber et al 1987; Kang et al 1987; Robakis 1988; Tanzi et al 1987). The key portion of this protein spans the end of the extracellular domain and the beginning of the transmembrane domain. The βAPP undergoes a series of endoproteolytic cleavages (reviewed in Selkoe 1994) (Figure 4) during its processing. There are three proteolytic cleavage events that surround or are within the AβP domain. One of these cleavages results from the putative membrane-associated α-secretase, which cleaves βAPP_{695} within the Aβ peptide domain between residues Lys_{687} and Leu_{688} (residues 16 and 17 of Aβ) and liberates the extracellular N-terminus of βAPP (which was previously identified as protease nexin II B). This pathway is nonamyloidogenic, because the cleavage precludes the formation of AβP.

The other cleavage pathway involves the recently cloned β-secretase (Vassar et al 1999) and the putative γ-secretases that give rise to a series of peptides that

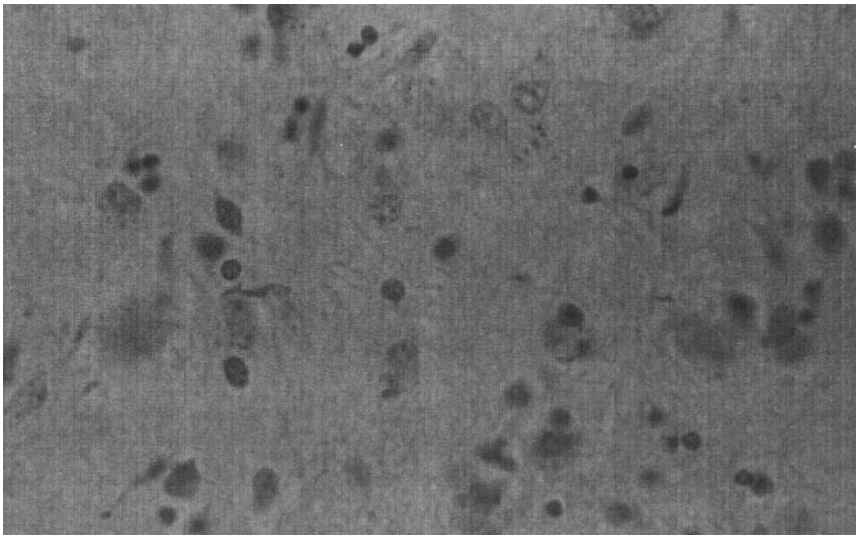

Figure 2. Extracellular amyloid plaques characteristic of Alzheimer's disease are composed principally of amyloid beta protein.

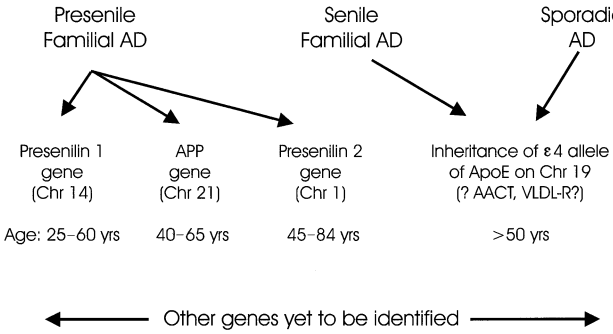

Figure 3. Age of onset distributions for the four known Alzheimer's disease genes.

contain the 40–42 amino acid AβP (β-secretase cleaves between Met_{671} and Asp_{672}; γ-secretase cleaves after Ile_{712}, Thr_{714}, or Val_{715} to generate $Aβ_{40}$, $Aβ_{42}$ or $Aβ_{43}$, respectively). AβP ending at residue 40 is the predominant isoform produced during normal metabolism of βAPP (Haass et al 1992; Jarrett and Lansbury 1993; Lorenzo and Yankner 1994; Pike et al 1993; Shoji et al 1992; Yankner et al 1990). Current evidence suggests that $Aβ_{40}$ is predominantly produced in endosomal–lysosomal systems (Golde et al 1992; Haass et al 1992a). Aβ peptides ending at residue 42 or 43 ($Aβ_{42/43}$ or long-tailed Aβ), on the other hand, are thought to be more fibrillogenic and more neurotoxic (Lorenzo et al 1994). These longer isoforms may be generated at intracellular sites such as the endoplasmic reticulum and cis-Golgi which, in neurons at least, are distinct from the sites of $Aβ_{40}$ synthesis (Cook et al 1997; Hartmann et al 1997; Wild-Bode et al 1997), although lipid-rich raft domains have also been suggested (Lee et al 1998). These secretases, and especially the specific γ-secretases, giving rise to the more fibrillogenic and potentially neurotoxic long-tailed $Aβ_{1–42}$, appear to play a central role in the pathogenesis of at least the genetic forms of AD associated with presenilin (PS) and βAPP mutations (see below and reviews) (Selkoe 1994; Yankner 1996).

Mutational studies of the βAPP gene first identified a Glu693Gln missense mutation of the βAPP gene (codon numbering of the $βAPP_{770}$ isoform) in affected and at-risk

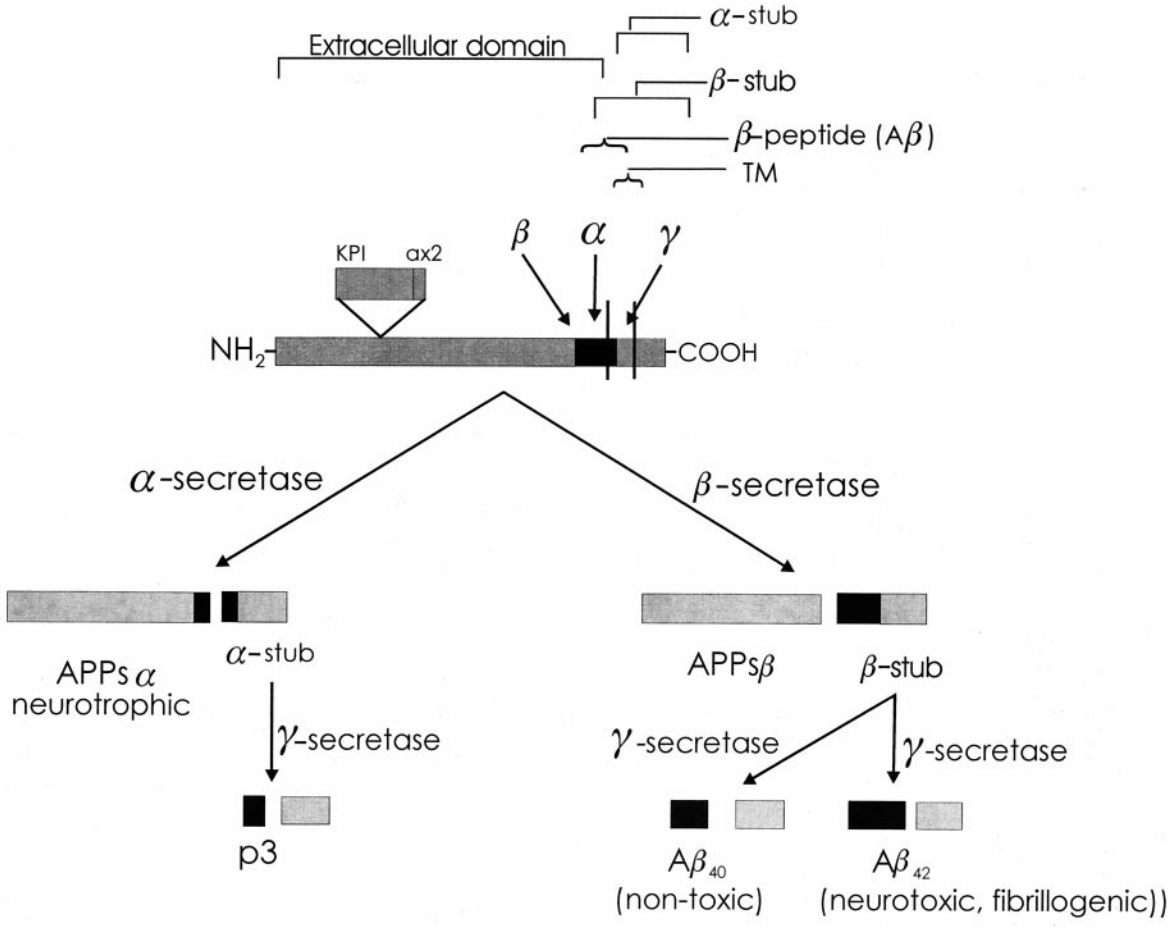

Figure 4. The beta amyloid precursor protein undergoes a series of endoproteolytic cleavages central to Alzheimer's disease pathogenesis.

Table 1. Missense Mutations in the βAPP Gene

Codon	Mutation	Phenotype
665	Gln→Asp	Late onset AD—no segregation
670/671	Lys–Met→Asn–Leu	FAD; increased Aβ production
673	Ala→Thr	No disease phenotype
692	Ala→Gly	FAD + cerebral hemorrhage; increased Aβ
693	Glu→Gly	Late onset AD, no segregation
	Glu→Gln	HCHWA-D
713	Ala→Val	Schizophrenia—no segregation
	Ala→Thr	AD—no segregation
716	Ile→Val	FAD
717	Val→Ile	FAD; increased long Aβ isoforms
	Val→Phe	FAD
	Val→Gly	FAD

FAD, familial Alzheimer's disease; HCHWA-D, amyloidosis of the Dutch type.
Data from St George-Hyslop 1997.

members of families with hereditary cerebral hemorrhage with amyloidosis of the Dutch type (HCHWA-D; Levy et al 1990). Subsequently, several different missense mutations were also found in exons 16 and 17 of the βAPP gene in families with early-onset AD (Table 1). Although some of these missense mutations are probably not pathogenic, the missense mutations at codon 670/671 (Swedish mutation; Mullan et al 1992), at codon 692 (Flemish mutation; Hendriks et al 1992), at codon 716 (Eckman et al 1997), and at codon 717 (Chartier-Harlin et al 1991; Goate et al 1991; Karlinsky et al 1992; Murrell et al 1991; Naruse et al 1991) are thought to be pathogenic. The mutations at codon 670/671 and at codon 692 are rare, having been seen only in single families. Mutations at codon 717 have been seen in approximately 20 unrelated pedigrees from different ethnic origins. The majority of codon 717 mutations have been seen in Anglo-Saxon, Italian, and Japanese subjects.

All of the clearly pathogenic mutations cluster close to the β-secretase site after Met_{671} (Lys670Asn/Met671Leu), α-secretase site after Lys_{687} (Ala692Gly and Glu693Gln), or γ-secretase site after Thr_{714} (Ile716Val, and Val717Ile). This led to the hypothesis that these mutations might influence the processing of βAPP (Mullan et al 1992). Indeed, mutations at codons 716 and 717 lead to a selective increase in the production of Aβ peptides ending at residue 42/43 (Cai et al 1993; Citron et al 1992; Eckman et al 1997; Haass et al 1994, 1995; Susuki et al 1994). The Lyc670Asn/Met671Leu mutation, on the other hand, appears to augment the production of both $Aβ_{40}$ and $Aβ_{42/43}$ (Citron et al 1994), whereas the Ala692Gly mutation has a more complicated effect on βAPP processing causing impaired α-secretase cleavage, increased heterogeneity of secreted Aβ species, and increased hydrophobicity of the Aβ (Haass et al 1994). The Ala692Gly mutation also has clinical features in some cases similar to those of HCHWA-D, and in other cases more similar to AD but with somewhat subtle differences in the size of the amyloid cores. The Glu693Gln mutation causes an increased propensity for Aβ to form fibrils (Wisniewski et al 1991).

The relative or absolute overproduction of AβP, and in particular of $Aβ_{42}$, as an effect that leads to neurodegeneration, is an attractive hypothesis. This hypothesis is supported by the observation that mutations in PS1 and PS2 (see below) also affect βAPP metabolism and lead to overproduction to $Aβ_{42}$. Moreover, $Aβ_{42}$ peptides have an enhanced propensity to fibril formation (Jarrett et al 1993; Lorenzo et al 1994; Pike et al 1993) and this conformational change is necessary to change the inert (or even marginally neurotrophic) soluble Aβ into toxic Aβ (Lorenzo et al 1994; Pike et al 1993). Multiple molecular mechanisms have been postulated to explain the neurotoxic effects of Aβ. These include induction of apoptosis by either direct effects on cell membranes or by indirect effects, such as potentiation of neurotoxic effects of excitatory amino acids, oxidative stress, and increases in intracellular calcium and free radicals (Arispe et al 1993; Mattson et al 1992, 1995). At least one experimental result, however, has suggested that βAPP mutations could also cause apoptosis by a mechanism that is independent of AβP production (Okamoto et al 1995; Yamatsuji et al 1996). Furthermore, it is also worth emphasizing that overproduction of AβP (as opposed to increased deposition and/or decreased clearance) is not a universal feature of all cases of sporadic AD (Scheuner et al 1996).

Apolipoprotein E

Genetic linkage studies in pedigrees with predominantly late-onset, familially aggregated AD provided suggestive evidence ($z = +2.5$ at $\theta = 0$) for the existence of a

second AD susceptibility locus near the markers *BCL3* and *ATP1A3* that maps to chromosome 19q12–q13 (Pericak-Vance et al 1991). Subsequent candidate gene analysis revealed that the gene likely responsible for the genetic locus on chromosome 19 was the APOE gene (Strittmatter et al 1993a). The APOE gene in humans contains three common coding sequence polymorphisms. The most common variant, ϵ3, reflects the presence of a cysteine at codon 112 and arginine at codon 158 and is present in approximately 75% of Caucasians. A second coding sequence variant, ϵ4, reflects substitution of arginine for cysteine at codon 112, and is present in approximately 15% of Caucasians. The third variant, ϵ2, contains cysteine at codons 112 and 158, and is present in approximately 10% of Caucasians. Analysis of these polymorphisms in normal control populations and in patients with AD has consistently shown that 1) there is an increased frequency of the ϵ4 allele in patients with AD (ϵ4 allele frequency in AD is approximately 40%) (Saunders et al 1993); and 2) that there is a smaller, reduced frequency of the ϵ2 allele (to about 2% in AD) (Corder et al 1994). More significantly, there is a dose-dependent relationship between the number of copies of ϵ4 and the age of onset of AD, such that homozygous ϵ4/ϵ4 subjects have an earlier onset (mean age $<$70 years) than heterozygous ϵ4 subjects (mean age of onset for ϵ2/ϵ3 is $>$90 years) (Corder et al 1994). Subjects with an ϵ2 allele, on the other hand, have a later onset (Corder et al 1993). The association between ϵ4 and AD has been robustly confirmed in numerous studies and in several different ethnic groups (reviewed in Roses 1996). The association is weaker with advanced age of onset, and the putative protective role of the ϵ2 allele is less clear at younger ages of onset (where ϵ2 may even be associated with a more aggressive course) (Rebeck et al 1994; van Duijn et al 1994). Currently, the major exceptions to the association of APOE ϵ4 with AD arise from studies in African Americans and Hispanics that have generated conflicting results (Hendrie et al 1995; Maestre et al 1995; Tang et al 1998). It remains unclear whether these conflicting results reflect the effects of statistical confounders or whether there is a true lack of association between AD and APOE ϵ4 in subsets of these populations.

Although the association between APOE ϵ4 and AD is robust, it is not entirely specific. Observations in patients with head injury (Mayeux et al 1995; Roses and Saunders 1995), spontaneous intracerebral hemorrhage (Alberts et al 1995), and those undergoing elective cardiac bypass surgery (Newman et al 1995), all suggest a poorer outcome for patients with the ϵ4 allele. There is also evidence for synergistic effects of a history of head injury and APOE ϵ4 on risk for AD (Mayeux et al 1995), such that patients with APOE ϵ4 and a head injury have a 10-fold increase in risk for AD compared to a 2-fold increase with APOE ϵ4 alone, and no increase for head injury alone. There is also a confirmed association between the ϵ4 allele and the Lewy body variant of AD, which has a subtly different clinical phenotype from classical AD (e.g., more frequent hallucinations, sensitivity to neuroleptics) (Olichney et al 1996).

The mechanism by which the ϵ4 allele is associated with an earlier onset of AD and by which the ϵ2 allele is associated with a later onset is unclear. The genetic data (the association of AD exclusively with the ϵ4 allele, a protective effect for the ϵ2 allele, and dose-dependent relationship between the ϵ4 copy number and age of onset of AD) provide a strong argument that ϵ4/ϵ2 polymorphisms are actual biological effectors in the APOE gene. At the current, however, time the possibility of mutations/polymorphisms in nearby sequences in linkage disequilibrium with ϵ4/ϵ2 cannot be entirely excluded. Nevertheless, a large body of biochemical evidence has been accumulated to support various hypotheses on how these APOE coding sequence polymorphisms might promote/protect against AD. The most obvious hypothesis is that APOE ϵ4/ϵ2 polymorphisms might influence the production, distribution, or clearance of the AβP. This hypothesis is supported by observations that the genotype at APOE accounts for some of the variation in age of onset in subjects carrying the βAPP Val717Ile mutation (but not the βAPP_{692} mutation) suggesting a direct biochemical interaction between APOE and βAPP or its metabolic products (Nacmias et al 1995; Sorbi et al 1995; St George-Hyslop et al 1994; Van Broeckhoven et al 1994). Second, subjects with one or more APOE ϵ4 alleles have a higher AβP plaque burden than do subjects with no ϵ4 alleles (Schmechel et al 1993). In vitro studies suggest that delipidated APOE ϵ4 binds Aβ more avidly than APOE ϵ3 (Strittmatter et al 1993a, 1993b). Third, there is evidence that both APOE and Aβ may be cleared through the lipoprotein-related (LRP) receptor and that APOE ϵ4 and the Aβ peptide may compete for clearance through the LRP receptor (Kounnas et al 1995). Finally, transgenic mice that have an intact endogenous APOE gene and that over-express human βAPP with the Val717phe mutant under the control of platelet-derived growth factor (PDGF) β-subunit promoter develop profuse deposits of extracellular Aβ by 9 months of age. In contrast, when the same transgene is expressed in an $APOE^{-/-}$ background, there is a dramatic reduction in extracellular Aβ deposition, thus supporting the hypothetical role for APOE in sequestering extracellular Aβ (Bales et al 1997).

There is some biochemical evidence to suggest a relationship between APOE and neurofibrillary tangles and synaptic density. In vitro APOE isoform-specific binding experiments with tau and the microtubule-associated pro-

tein (MAP) 2 suggests that ϵ3 binds to both *tau* and MAP2 better than does the ϵ4 isoform (Huang et al 1994, 1995; Strittmatter et al 1994). This has led to suggestions that ϵ2 and ϵ3 may protect and sequester MAPs better than ϵ4, thereby reducing the ability of *tau* to bind to microtubules, become hyperphosphorylated, and form paired helical filaments.

Finally, it has been suggested that APOE may be involved in synaptic plasticity during regeneration and repair, and that the ϵ4 allele is less efficient in this role. Thus, APOE knockout mice also show an age-dependent decrease in synaptic density and spontaneous Aβ peptide aggregation within astrocytic processes (Masliah et al 1995). Several types of neural tissue culture cells demonstrate decreased neurite outgrowth in the presence of APOE ϵ4 in the media rather than APOE ϵ3 (Nathan et al 1994, 1995). Perhaps then, the APOE isoforms might differentially affect synapse formation in response to injury, learning, and aging.

In addition to the ϵ2/ϵ3/ϵ4 coding sequence polymorphisms, several polymorphisms have been discovered in the 5′-promoter of the APOE gene. In some studies (Bullido et al 1998; Lambert et al 1998), these polymorphisms are suspected of increasing risk for AD independently of APOE ϵ4 and to cause this increased risk for AD by altering the transcriptional activity of APOE (although direct proof of the latter in vivo remains to be obtained). Several other independent studies have been unable to replicate these findings, although they do confirm that the B491 A/T polymorphism at least, is in linkage disequilibrium with the APOE ϵ2/ϵ3/ϵ4 polymorphism (Town et al 1998; Song et al 1998).

Presenilin 1

Genetic linkage studies mapped a locus associated with a very aggressive early-onset AD to a series of polymorphic markers located on chromosome 14q24.3 (D14S43, D14S71, D14S77, and D14S53) (Schellenberg et al 1992; St George-Hyslop et al 1992; Van Broeckhoven et al 1992). The actual disease gene (PS1) was subsequently isolated using a positional cloning strategy (Sherrington et al 1995) and a homologue (PS2) was then mapped to chromosome 1.

The chromosome 14 AD3 subtype gene, PS1, is highly conserved in evolution, being present in *Caenorhabditis elegans* (Levitan and Greenwald 1995) and *Drosophila melanogaster* (Boulianne et al 1997) and encodes a polytopic integral membrane protein with between 6 and 10 possible transmembrane domains (see below and Figure 5). The PS1 gene is transcribed at low levels in many different cell types, both within the central nervous system (CNS) and also in non-neurological tissues (Sherrington et al 1995). In the CNS, PS1 transcripts can be detected by in situ hybridization in the neocortex (especially in cortical neurons in layers II and IV), neurons of the CA1-CA3 fields of the hippocampus, granule cell neurons of the dentate gyrus, subiculum, cerebellar Purkinje and granule cells and deep nuclei, as well as lesser amounts in the olfactory bulb, the striatum, some brainstem nuclei, and the thalamus (Lee et al 1996).

Immunoblotting and immunohistochemical studies suggest that the PS1 protein is approximately 50 kDa in size and is predominantly located within intracellular membranes in the endoplasmic reticulum, perinuclear envelope, the Golgi apparatus and some as yet uncharacterized intracytoplasmic vesicles (De Strooper et al 1997; Walter et al 1996). Studies of the topology of PS1 suggest that the N-terminus, the TM6-TM7 loop, and the C-terminus are located in the cytoplasm (DeStrooper et al 1997; Doan et al 1996; Lehmann et al 1997; Li and Greenwald 1996). Studies of the PS1 protein in brain tissue, as well as in many other peripheral tissues, reveal that only very small amounts of the PS1 holoprotein exist within the cell at any given time (Podlisny et al 1997; Thinakaran et al 1996). Instead, the holoprotein is actively catabolized, by at least two different proteolytic mechanisms. One of these mechanisms appears to involve the proteasome (Fraser et al 1998). Another proteolytic mechanism involves a series of heterogeneous endoproteolytic cleavages near residue 290 within the TM6-TM7 loop domain (Podlisny et al 1997; Thinakaran et al 1996). This endoproteolytic cleavage generates a series of N- and C-terminal heterogeneous fragments of approximately 35 kDa and 18–20 kDa, respectively. Remarkably, the stoichiometry of the N- and C-terminal fragments is tightly maintained on a 1:1 ratio, and the absolute abundance of the N- and C-terminal fragments is also tightly regulated, such that artificial over-expression of PS1 results in only a modest increase in N- and C-terminal fragments (Thinakaran et al 1996). This has led to the suggestion that the process of endoproteolytic cleavage involves a tightly regulated saturable process (Thinakaran et al 1998). A third proteolytic mechanism acting upon the presenilin holoproteins involves members of the caspase 3 family of proteases. Activation of apoptosis by a variety of means that culminates in activation of caspase 3 results in cleavage of PS1 near residue aspartate 345 and the equivalent aspartate residue in PS2 (Asp 329) (Brockhaus et al 1998; Grunberg et al 1998; Kim et al 1997). It is currently unclear whether caspase-mediated endoproteolytic cleavage of the presenilins is actively involved in the regulation of apoptotic signal pathways (preliminary data suggest an anti-apoptotic effect for the C-terminal derivative of PS1 cleaved by caspase (Vito et al 1997), or whether the presenilins simply represent innocent bystanders that are cleaved by the caspase

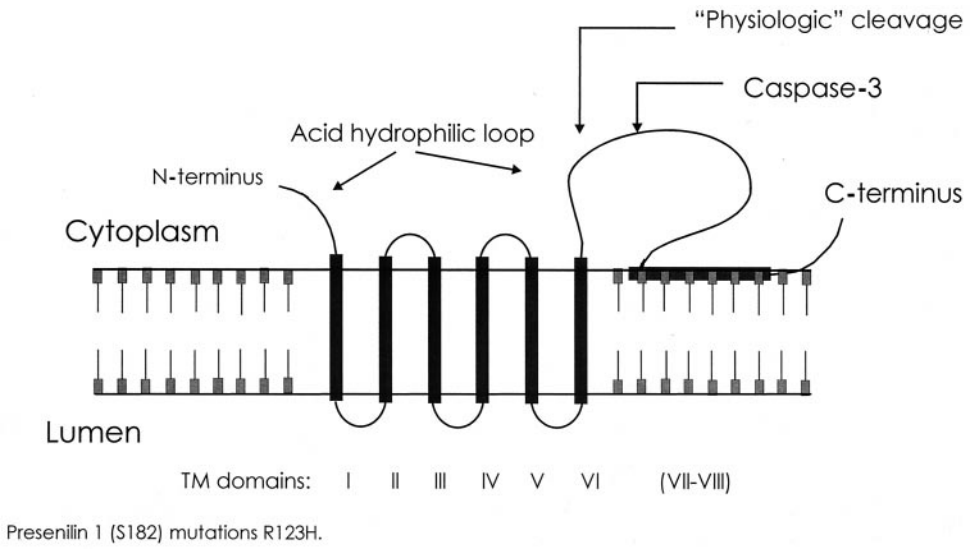

Figure 5. A putative structure of the presenilin 1 gene product: a polytopic integral membrane protein with 6–10 possible transmembrane domains.

enzyme once apoptosis is activated. Certainly, caspase-mediated cleavage of the presenilins is not required for the effects of the presenilins on amyloidogenesis and *Notch* signaling (Baumeister et al 1997; Brockhaus et al 1998).

Both the holoprotein and its endoproteolytic fragments exist as components of independent high molecular weight, multimeric protein complexes. Thus, the holoprotein appears to be a component of an ~180 kDa complex, which is predominantly resident within the rough endoplasmic reticulum (rER) (Capell et al 1998; Thinakaran et al 1998; Yu et al 1998). Both the N-terminal fragment (NTF) and the C-terminal fragment (CTF) associate with each other as heterodimeric components of a larger (~250 kDa) multimeric protein complex that is resident in the endoplasmic reticulum, Golgi apparatus, and some additional intracellular membranous domains whose identity has not been entirely clarified (Capell et al 1998; Lee et al 1998; Thinakaran et al 1998; Yu et al 1998). It has been suggested that the incorporation of the holoprotein and its subsequent endoproteolytic cleavage and incorporation into a larger complex represents a rate-limiting step in the processing pathway (Thinakaran et al 1997). Once incorporated into these high molecular weight complexes, however, the endoproteolytic fragments remain together with a stable 1:1 stochiometry and with very long half-lives (Capell et al 1998; Thinakaran et al 1998; Yu et al 1998). Holoprotein monomers that fail to get incorporated into these complexes are rapidly degraded with a half-life of less than one hour via a proteasome-dependent mechanism (Fraser et al 1998). Incorporation of the presenilin proteins and their derivatives into these complexes is necessary for the biological activity of the presenilins (Tomita et al 1998; Figure 6).

The identity of the other components of the presenilin complexes is currently under investigation; however, a combination of yeast-two-hybrid, co-immunoprecipitation, immunohistochemistry, and biochemical fractionation studies reveals that in peripheral tissues and in brain, the presenilins associate with β-catenin, a member of the armadillo protein superfamily (Yu et al 1998; Zhou et al 1997). In brain, the presenilins also associate with a novel armadillo protein termed Neuronal Plakophillin-related Armadillo Protein (NPRAP) or δ-catenin (Yu et al 1998; Zhou et al 1997). The functional significance of the presenilin–armadillo interactions is not entirely clear because the armadillo proteins have diverse functions, ranging from a structural role in stabilization of intercellular junctions (including synapses), intracellular transduction of receptor-mediated signals (e.g., Wnt and certain growth factors), to participation in apoptotic cell death pathways. There is evidence from some laboratories that presenilins may also directly interact with a number of other proteins such as βAPP (Weidemann et al 1997; Xia et al 1997) and Filamin Binding Protein (Zhang et al 1998); however, not all laboratories confirm these latter results (Thinakaran et al 1998).

The normative function of PS1 has not yet been fully defined. Functional analogies have been made to the weakly homologous SPE4 protein of *C. elegans,* which is involved in maintenance of a Golgi-derived membranous organelle thought important in partitioning of protein and cell membrane products in the maturing spermatocyte of *C. elegans* (L'Hernault et al 1992). This has led to speculation that the PS1 protein might subserve a similar role in protein and membrane trafficking (Sherrington et al 1995). This hypothesis is supported by two different lines

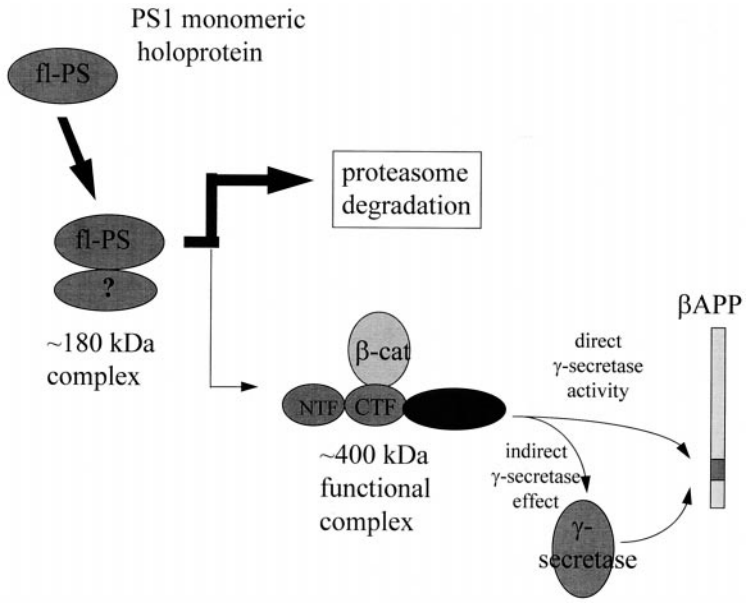

Figure 6. Biological processing of presenilins.

of direct experimental evidence. First, ablation of functional PS1 expression by homozygous targeted disruption of the murine PS1 gene causes aberrant processing of βAPP with the failure of γ-secretase cleavage of the C-terminal stubs of βAPP derived from α-secretase or β-secretase cleavage (De Strooper et al 1998). Failure of γ-secretase cleavage results in the accumulation of uncleaved α-secretase or β-secretase stubs in a variety of intracellular loci including the endoplasmic reticulum, Golgi bodies, and lysosomes (De Strooper et al 1998). Second, as described earlier, the presenilins form multimeric protein complexes with the armadillo protein β-catenin (Yu et al 1998; Zhou et al 1997), and missense mutations in PS1 and PS2 cause mistrafficking of this presenilin ligand (Nishimura et al 1999).

Other putative roles for PS1 have included a role in the regulation of intercellular signal transduction during development, in apoptosis, and possibly in intracellular calcium ion homeostasis. The former suggestion arose because null mutations in a second presenilin orthologue in *C. elegans (Sel12)* exert a suppressor effect on abnormalities in vulva progenitor cell fate decisions induced by activated *Notch* mutants (Levitan and Greenwald 1995). *Notch* is involved in intercellular signaling during development. *Sel12* protein shows stronger amino acid sequence identity to the human presenilin proteins than does SPE4. A role for mammalian presenilins in *Notch*-mediated signal transduction is further supported by the fact that homozygous targeted knock-out of the murine PS1 protein (using homologous recombination) causes embryonic lethality around day E13 and is associated with 1) severe developmental defects in somite formation and axial skeleton formation; 2) the occurrence of cerebral hemorrhage; and 3) reduced *Notch* and *delta* transcription in selected cell types (Shen et al 1997; Wong et al 1997). Similar phenotypes have been observed in mice with targeted knock-outs of the murine *Notch1* and DLL1 genes, supporting the hypothesis that PS1 has either a direct or an indirect role in intercellular signal transduction (Conlon et al 1995).

To date, more than 40 different mutations have been discovered in the PS1 gene (Table 2). The majority of these mutations are missense mutations giving rise to the substitution of a single amino acid. These mutations are predominantly located in highly conserved transmembrane domains; at/near putative membrane interfaces; or in the N-terminal hydrophobic or C-terminal hydrophobic residues of the putative TM6-TM7 loop domain. Two splicing defect mutations have been identified. One involves a point mutation in the splice acceptor site at the 5′-end of exon 10 (in some exon numbering systems, exon 10 is labeled exon 9) (Kwok et al 1997; Perez-Tur et al 1995; Sato et al 1998). Because exon 9 and exon 11 are in-frame, this mutation allows exon 9 to be fused in-frame with exon 11, thereby removing a series of charged residues at the apex of the hydrophillic acidically charged TM6-TM7 loop domain. The second splice site mutation arises from deletion of a G nucleotide from the splice donor site at the 3′-end end of exon 5 (E. Rogaeva et al, unpublished data, 1999; Tysoe et al 1998). This results in use of an alternate GT splice donor site four base pairs further downstream in intron 5, which causes the in-frame insertion of an extra tyrosine residue within the luminal TM1-TM2 loop domain (E. Rogaeva et al, unpublished data, 1999).

The wide scattering of missense mutations has led to speculation that the effect of most of the familial Alzhei-

Table 2. Missense Mutations in the Presenilin Genes

		Presenilin 1 (S182)	
Codon	Location	Mutation	Phenotype
82	TM1	Val6Leu	FAD, onset 55 years
96	TM1	Val6Phe	FAD
115	TM16TM2 loop	Tyr6His	FAD, onset 37 years
117	TM16TM2 loop	Pro6Leu	FAD onset 28 years
120	TM16TM2 loop	Glu6Asp	FAD, onset 48 years
139	TM2	Met6Thr	FAD, onset 49 years
139	TM2	Met6Val	FAD, onset 40 years
143	TM2	Ile6Thr	FAD, onset 35 years
146	TM2	Met6Leu	FAD, onset 45 years
146	TM2	Met6Val	FAD, onset 38 years
146	TM2	Met6Ile	FAD, onset 40 years
163	TM3 interface	His6Arg	FAD, onset 50 years
163	TM3 interface	His6Tyr	FAD, onset 47 years
171	TM3	Leu6Pro	FAD, onset 40 years
209	TM4 interface	Gly6Val	FAD
213	TM4 interface	Ile6Thr	FAD
231	TM5	Ala6Thr	FAD, onset 52 years
233	TM5	Met6Thr	FAD, onset 35 years
235	TM5	Leu6Pro	FAD, onset 32 years
246	TM6	Ala6Glu	FAD, onset 55 years
260	TM6	Ala6Val	FAD, onset 40 years
263	TM6–TM7 loop	Cys6Arg	FAD, onset 47 years
264	TM6–TM7 loop	Pro6Leu	FAD, onset 45 years
267	TM6–TM7 loop	Pro6Ser	FAD, onset 35 years
280	TM6–TM7 loop	Glu6Ala	FAD, onset 47 years
280	TM6–TM7 loop	Glu6Gly	FAD, onset 42 years
285	TM6–TM7 loop	Ala6Val	FAD, onset 50 years
286	TM6–TM7 loop	Leu6Val	FAD, onset 50 years
del291-319	TM6–TM7 loop short loop		FAD,
384	TM6–TM7 loop	Gly6Ala	FAD, onset 35 years
392	TM6–TM7 loop	Leu6Val	FAD, onset 25–40 years
410	TM7	Cys6Tyr	FAD, onset 48 years
141	TM2	Asn6Ile	FAD, onset 50–65 years
239	TM5	Met6Val	FAD, onset variable 45–84 yrs

FAD, familial Alzheimer's disease.
Data from St George-Hyslop 1997.

mer's disease- (FAD-) related mutations is a gain of function effect (Van Broeckhoven 1995). This is partially borne out by two observations in PS1 gene knock-out animals ($PS1^{-/-}$). First, these animals have a phenotype of early perinatal mortality without evidence of AD (Shen et al 1997; Wong et al 1997). This loss of function phenotype in $PS1^{-/-}$ animals can be completely rescued by both wild-type and mutant PS1 transgenes (Davis et al 1998; Qian et al 1998). Second, $PS1^{-/-}$ mice have a defect in βAPP processing manifested by the failure of γ-secretase cleavage and the accumulation of the C-terminal stubs of βAPP following α- and β-secretase cleavage (α- and β-stubs) (De Strooper et al 1998). This defect in βAPP processing is completely reversed by both wild-type and mutant PS1 transgenes. A gain of function is imparted by the mutant transgenes because, in addition, it also induces an increase in $A\beta_{42}$ which, as described below, is a consistent biochemical effect of PS1 mutations (Davis et al 1998; Qian et al 1998). Studies using human PS1 cDNAs in complementation assays of mutant *Sel12* in *C. elegans,* however, suggest that the wild-type human PS1, but not mutant human PS1 cDNAs, are able to complement the loss-of-function sel12 mutants (Baumeister et al 1997; Levitan et al 1996). The latter result argues that the human PS1 mutants may not be fully functional (but do not fully preclude a gain of function effect as well).

Regardless of whether PS1 (and PS2) mutations cause a gain of function or a loss of function, it seems likely that one effect is to alter the processing of βAPP by preferentially favoring the production of potentially toxic long-tailed Aβ peptides ending at residue 42 or 43 (Borchelt et al 1996; Citron et al 1997; Duff et al 1996; Martins et al 1995; Scheuner et al 1996). Thus, fibroblasts from heterozygous carriers of PS1 mutations, various cell lines

transfected with βAPP and PS1 cDNAs, as well as the brain from transgenic mice mutant PS1 transgenes all contain or secrete increased quantities of long AβP isoforms with only a variable but minor increase in short-tailed Aβ peptides (Borchelt et al 1996; Citron et al 1997; Duff et al 1996; Martins et al 1995; Scheuner et al 1996). Direct measurements of AβP isoforms in the postmortem brain tissue of patients dying with PS1-linked FAD also show marked increases in the amount of long-tailed Aβ isoforms compared to control brain tissue and to brain tissue from subjects with sporadic AD (Tamaoka et al 1998).

Presenilin 2

During the cloning of the PS1 gene on chromosome 14, a very similar sequence was identified in the public nucleotide sequence databases (Levy-Lahad et al 1995; Rogaeva et al 1995). Further analysis revealed that this similar nucleotide sequence was derived from a gene on chromosome 1q42.1. The gene encodes a polypeptide whose open reading frame contained 448 amino acids and has substantial amino acid sequence identity with that of the PS1 protein (overall identity approximately 60%); however, in contrast to PS1, which is expressed more or less homogeneously in the brain and in peripheral tissues, PS2 is maximally expressed in cardiac muscle, skeletal muscle, and pancreas. Not surprisingly, the predicted topology of PS2 is similar to that of PS1, and it also forms similar but independent multimeric protein complexes that contain β-catenin (Capell et al 1998; Yu et al 1998).

Mutational analyses uncovered two different missense mutations in the PS2 gene in families segregating early-onset forms of AD. In the first mutation (Asn141Ile), which was detected in a proportion of families of Volga German ancestry, in which the FAD locus had been independently mapped by genetic linkage studies to chromosome 1 (Levy-Lahad et al 1995; Rogaeva et al 1995). The second mutation (Met239Val) was discovered in an Italian pedigree (Rogaeva et al 1995). In contrast to the frequency of PS1 mutations, however, screening of large data sets reveals that PS2 mutations are likely to be rare (Sherrington et al 1996).

Another profound difference between the PS2 mutations and those in the βAPP and PS1 genes is that the phenotype associated with PS2 mutations is much more variable (Bird et al 1996; Sherrington et al 1996). The range of age of onset in heterozygous carriers of PS2 mutations is between 40 and 85 years of age, and there is at least one instance of apparent nonpenetrance in an asymptomatic octogenarian transmitting the disease to affected offspring (Sherrington et al 1996).

Given the strong similarities in structure and in amino acid sequence of the PS1 and PS2 proteins, it would seem likely that PS1 and PS2 have similar or overlapping activities. This hypothesis is supported by the fact that PS2 mutations, like PS1 mutations, increase the secretion of long-tailed Aβ peptides (Citron et al 1997; Scheuner et al 1996); however, in contrast to $PS1^{-/-}$ mice, $PS2^{-/-}$ mice show no major developmental or βAPP processing defects (Herreman et al 1999).

The Chromosome 12 Locus

The observation that PS1, PS2, βAPP, and APOE mutations/polymorphisms accounted for approximately one half of the genetic variance of AD prompted a number of groups to undertake genome surveys to identify other susceptibility loci. In two independent data sets of late-onset AD pedigrees ($n = 16$; $n = 38$), a novel AD locus was identified in the pericentriometic region of chromosome 12 between the markers D12S1042 and D12S390 ($z = 3.5$) (Pericak-Vance et al 1997). Follow-up studies in a dataset of 53 independent pedigrees confirmed the presence of an AD susceptibility gene within a larger (approximately 60 cM) region of chromosome 12 in a subset of pedigrees (Rogaeva et al 1998). The subset of pedigrees that do not show linkage to chromosome 12 presumably reflect the presence of one or more genetic susceptibility factors elsewhere in the genome.

The identity of the chromosome 12 locus at the moment is unclear. Several candidate genes including α-2-macroglobulin (A2M), low-density lipoprotein receptor related protein (LRP1), and ARF2, etc., map within this interval. Some of these genes (e.g., ARF3, Wnt1, plakophilin 2, ITR2) have been excluded by the failure to find nucleotide sequence changes in their open reading frames that are either enriched in, or unique to, patients with the chromosome 12 form of AD. Some biochemical studies have suggested a role for A2M protein in AD (through its ability to bind Aβ and through competition with both Aβ and APOE for clearance through the LRP1 receptor) and have led to speculation that A2M might be the site of mutations associated with AD (see Blacker et al 1998 for a summary.). Considerable support for this concept (which would link Aβ, APOE, A2M, and LRP1 in one biochemical cascade) was provided when a preliminary family-based association study detected an association between AD and an defects similar to those seen with mutations in PS1 and βAPP. Other genes may result in incompletely penetrant autosomal dominant traits like that associated with PS2; however, it is likely that a significant proportion of the remaining genes will have effects similar to that of APOE in which the ultimate phenotype is influenced by the presence or absence of other genetic and environmental risk factors.

Attempts to identify novel AD susceptibility genes have followed two strategies. One strategy, a continuation of the conventional positional cloning strategy, has attempted to show co-segregation of marker alleles with the disease phenotype (identity by descent) in pedigrees multiply affected by AD. These studies have assessed co-segregation using both conventional parametric lod score methods as well as newer nonparametric methods (Kruglyak et al 1996; Pericak-Vance 1997; Rogaeva et al 1998). Because these methods work best for pedigrees multiply affected with AD, however, increasing emphasis has been placed upon the use of simple case-control studies (such as those so effectively used to discover the association between AD and APOE). More recently the case-control method for discovering allelic associations has been supplemented by novel statistical methods (family-based association methods) that allow the examination of allele sharing between affected siblings compared to unaffected siblings, such as the sibship disequilibrium transmission test (SDT) and the sib-transmission disequilibrium test (S-TDT) (Blacker et al 1998; Spielman and Ewens 1998). These family-based association methods examine the parental alleles transmitted to unaffected siblings as a source of control chromosome information that, theoretically, is better matched for ethnicity and genetic background.

The conventional case-control allelic association tests have yielded positive results on a significant number of genes, many of which are plausible biochemical candidate genes; however, most of these studies have not received the same robust replication as the association between AD and the ϵ4 allele of APOE. As a result, it has been difficult to discern whether the reported associations are true but perhaps limited to particular subsets of AD or whether they represent statistically significant but biologically incorrect results. The possibility of biologically false-positive results in allelic association studies is a well-recognized problem in human genetics that can arise when the test and control populations are not drawn from identical genetic backgrounds (i.e., due to population stratification). Recently, in an attempt to rectify the high false-positive rate for simple allelic association studies, some studies have begun to use the family-based association methods even though they have not yet been fully validated (see Blacker et al 1998).

A partial list of candidate genes provisionally identified as putative AD susceptibility loci include homozygosity for the "AA" allele of an intronic polymorphism α1-chymotrypsin (Kamboh et al 1995), "A5 repeat" allele of an intronic insertion-deletion polymorphism in the very low-density lipoprotein receptor (Okuizumi et al 1995), neutral coding sequence and intronic polymorphisms in low-density lipoprotein receptor related protein (Kamboh et al 1998; Kang et al 1997), homozygosity for the common "A1" allele of an intronic polymorphisms in PS1 (Wragg et al 1996), PS2, K-variant of butyrylcholinesterase (Lehmann et al 1997), homozygosity for the Val/Val variant of the Val443Ile polymorphism in bleomycin hydroxylase (Montoya et al 1998), etc. Most of these candidate genes, however, have not received the same widespread confirmation as did APOE ϵ4 when tested in independent but comparable data sets (Brindle et al 1998; Clatworthy et al 1997; Crawford et al 1998; Fallin et al 1997; Haines et al 1996; Hollenbach et al 1998; Lendon et al 1997; Scott et al 1996; Song et al 1998; Town et al 1998; Wavrant-DeVrieze et al 1997; Woodward et al 1998).

Animal Models

Analysis of the biochemical effects of the known FAP genes suggests that they have a common biochemical effect related to βAPP processing and extracellular disposition of AβP. These findings form a basis for rational attempts to design treatments for AD: to inhibit γ-secretase, to inhibit the production of Aβ, or to prevent the aggregation of the Aβ into fibrils. The FAD genes identified to date, particularly βAPP and PS1 and PS2, can be used to make murine models that will help further explore the pathogenesis and test various treatments. Models with mutant transgenes and βAPP have been produced to date, and these animals generally develop very profuse deposits of Aβ. The animal model is not yet a full model of AD. There is some suggestion that these mice have dystrophic neurons in the Aβ plaques and they may have subtle behavioral changes, and electrophysiological changes but no neurodegeneration. Transgenes expressing mutant PS1 and PS2 have so far just produced electrophysiological abnormalities but no overt pathology. If mutant PSI transgenic mice are crossed with mutant βAPP mice, accelerated pathology with diffuse AβP deposits results (see above). In addition, crossing βAPP mutant mice with $APOE^{-/-}$ mice reveals that APOE probably has an effect on βAPP processing (see above).

Genetic Testing in AD

These genetic substrates can be used for clinical diagnostic testing in AD. The βAPP, PS1 and PS2 genes can be employed in both presymptomatic and symptomatic genetic testing (Figure 7) under very careful conditions using a paradigm for counseling similar to that for Huntington's disease. These genes have an early age of onset; they are virtually fully penetrant, and the age of onset at which the disease develops in a carrier of a presenilin mutation or a βAPP mutation is consistent in a given family (generally within 5–10 years of the mean age of onset for that