Genetics of Alzheimer's Disease: Lessons Learned in Two Decades

Alzheimer Hastalığının Genetiği: Son 20 Yılda Öğrenilen Dersler

Nilüfer Ertekin Taner

Mayo Klinik Florida, Nöroloji ve Nörobilim Bölümleri, Jacksonville, Florida, Amerika Birleşik Devletleri

Turk Norol Derg 2010;16:1-11

ÖZET

Alzheimer hastalığı (AH) en sık rastlanan demans türüdür. 2010 yılında tüm demansların dünyada 35 milyondan fazla kişiyi etkilemesi beklenmektedir. Etkili tedaviler olmaksızın, bu salgının 2050 yılında tüm dünyada 115 milyondan fazla hasta sayısına ulaşacağı hesaplanmaktadır. Genetik çalışmalar hastalığın patofizyolojisini anlamaya yarayarak, olası tedavi, semptom öncesi tanı ve önlemlere yol açabilir. 1990 yılından bu yana AH'ın altta yatan genetik ögesi hakkında önemli oranda kanıt birikmiştir. Erken başlangıçlı ailesel AH'a yol açan otozomal dominant mutasyonlar taşıyan üç gen, AH'ın %1'inden daha azını açıklamaktadır. Geç başlangıçlı AH'daki genel kabul gören tek risk faktörü olan apolipoprotein ε4, bu hastalığın genetik riskinin yalnızca bir kısmını açıklar. Genetik bağlantı ve ilişki çalışmalarında birçok aday gen bölgesi bulunmasına rağmen, bu sonuçlar bağımsız çalışmalarda çoğunlukla tekrarlanamamıştır. Bunun nedeni, en azından kısmen, genetik heterojenlik, düşük etkili genetik faktörler ve yetersiz güçte olan çalışmalardır. Yüz binlerce tekli nükleotid polimorfizmi ile binlerce kişinin incelendiği genom çapında ilişki çalışmaları, AH gibi karmaşık genetiğe sahip hastalıkların altında yatan yaygın risk varyasyonlarının bulunması için olası güçlü bir yaklaşım olarak görülmektedir. Günümüzde geç başlangıçlı AH'da 11 tane genom çapında ilişki çalışması tamamlanmış ve izlenmesi gereken yeni aday genetik bölge ve genlerin bulunmasına yol açmıştır. Bu çalışmalar AH'da yeni, orta çapta etkili, olası genetik faktörlerle ilgili kanıtlar sağlamalarına rağmen, bu hastalığın hesaplanan riskinin tümünü açıklamaya yetmemektedir. Bunun olası nedenleri, şu ana kadar bulunmuş genetik faktörlere ek olarak, AH'ın altında yatan ve daha geleneksel genetik bağlantı ve ilişki analizi yöntemleriyle bulunamayabilen, düşük etkili, nadir ve/veya yapısal (strüktürel) DNA polimorfizmleri olabilir. Yeni kuşak sıralama (sekanslama), sayısal endofenotip, kopya sayısı varyasyonları ve meta-analiz gibi alternatif yaklaşımların bu ek genetik risk faktörlerinin bulunması için gerekli olabileceği düşünülmektedir. Bu derlemede AH'ın genetiği hakkında günümüzde bilinenlerin bir özeti sunulacak ve son 20 yılda öğrenilen dersler ışığında gelecekteki genetik çalışmalar için yaklaşım modelleri tartışılacaktır.

Anahtar Kelimeler: Alzheimer hastalığı, genetik, bağlantı (genetik), ilişki.

ABSTRACT

Genetics of Alzheimer's Disease: Lessons Learned in Two Decades

Nilüfer Ertekin Taner

Departments of Neurology and Neuroscience, Mayo Clinic Florida, Jacksonville, Florida, USA

Alzheimer's disease (AD) is the most common type of dementia. It is estimated that more than 35 million people worldwide will suffer from dementia in 2010. Without effective therapies, this epidemic is expected to affect more than 115 million patients worldwide by 2050. Genetic studies can help us understand the disease pathophysiology, thereby providing potential therapeutic, presymptomatic predictive and preventative avenues. Since 1990, there has been evidence for a substantial genetic component underlying the risk for AD. Three genes with autosomal dominant mutations lead to early-onset familial AD, which explains less than 1% of all AD. Apolipoprotein ε4, the only widely accepted genetic risk factor for late-onset AD, accounts for only a portion of this risk. Genetic linkage and association studies have identified multiple candidate gene regions, although many resulting candidate genes suffer from lack of replication, at least partially due to underpowered studies in the setting of genetic heterogeneity and small-to-moderate effect size. Genome-wide association studies that assess hundreds of thousands of single-nucleotide polymorphisms (SNPs) in thousands of subjects have been viewed as a potentially powerful approach in uncovering common risk variations for genetically complex diseases such as AD. To date, 11 independent genome-wide association studies have been completed in late-onset AD (LOAD) that led to candidate regions and genes for follow-up. These studies provide evidence for novel, plausible genetic risk factors for AD, but still fail to account for all of the estimated risk. Additional genetic risk factors of even smaller effect size, rare variants and/or structural DNA polymorphisms may exist, which may escape detection by conventional methods. Alternative approaches such as next-generation sequencing, use of quantitative endophenotypes, copy number variation analyses, and meta-analyses may be required. This review summarizes the current knowledge on the genetics of AD and suggests a framework for future genetic studies utilizing the lessons learned over the past two decades.

Key Words: Alzheimer disease, genetics, linkage (genetics), association.

Alzheimer's disease (AD) is the most common form of dementia that is characterized by memory decline but also impairment in other cognitive areas, including language, executive function and visuospatial abilities. The definitive diagnosis of AD is done by pathology, which is characterized by senile plaques composed predominantly of extracellular accumulation of the amyloid β (Aβ) peptide and neurofibrillary tangles formed by intracellular accumulation of the abnormally hyperphosphorylated microtubule associated protein, tau (1,2). More than 100 years after its description, AD is an epidemic with major medical, social and economical impact (1). The number of patients with dementia is expected to exceed 35 million in 2010 and 115 million in 2050 unless effective therapies are identified (3). The estimated yearly cost of dementia in 2005 was US $ 315.4 billion, with 77% of the cost attributable to the high-income countries (4). Whereas formal, institutional care accounts for much of the cost in high income countries, informal and mostly in-home care is the underlying major cost in middle- or low-income countries.

It is estimated that even a modest therapy that would delay the onset of this disease by only six months could lead to 380,000 fewer people with AD and an annual savings of US $ 18 billion per year just in the United States (U.S.), 50 years after initiation of such a therapy (5). It is clear that development of effective therapies for a disease requires a thorough understanding of its pathophysiology, risk and protective factors. Table 1 summarizes the

Table 1. Risk and protective factors for Alzheimer's disease

Risk factors
Age
Female sex
African-American, Hispanic ethnicity
Vascular risk factors (hypertension, high cholesterol, diabetes, smoking)
Head trauma
Poor socioeconomic status
Low education
Environmental/occupational exposures
Protective factors
Exercise
Higher education
Alcohol in moderation
Diet (fish oil, Mediterranean diet)
Mental/social activity
Medications (cholesterol-lowering agents, non-steroidal anti-inflammatory medications)*
Vitamins (Vitamins B6, B12 and folic acid: lower homocysteine; Vitamin E: antioxidant)*
Estrogen*

* No evidence from combined clinical trials to support the use of these medications, vitamins or hormones in Alzheimer's disease.

risk and protective factors for AD proposed through epidemiological, laboratory and clinical trial studies (6,7). As seen from this table, many of these factors are unmodifiable (age, sex, ethnicity) or require lifelong life habit and/or socioeconomic modifications (education, diet, exercise, socioeconomic status) that may be difficult to achieve. Recognizing the underlying genetic component of a disease and identification of genetic risk and protective factors constitute an important additional approach to elucidating the disease mechanism. This knowledge could aid in the development of novel therapeutic approaches by identifying druggable targets. Furthermore, genetic risk and protective factors could potentially be used as biomarkers to determine at-risk populations in which to commence drug therapy in the pre-symptomatic stage, much like blood cholesterol levels are used to decide who should be on lipid-lowering agents before any signs or symptoms of cardiovascular/cerebrovascular disease emerge. Current ongoing therapeutic trials in AD are beyond the scope of this review; however, it should be mentioned that many experimental therapies under investigation stem from the knowledge gained by genetic studies coupled with functional laboratory approaches (8). Thus, genetic studies play a central role in the pathophysiology-prediction/prevention-cure paradigm for AD. In this review, we discuss the current knowledge on the genetics of AD and potential future approaches. Given the vast number of publications in the area of AD genetics, this review will focus on the key studies under each of the headlines, and will refer to data-summarizing websites and other review studies, where applicable.

Evidence for a Genetic Component in AD

The initial evidence for an underlying genetic risk for AD comes from three lines of studies, namely, familial aggregation, transmission pattern and twin studies. The large, Multi-Institutional Research in Alzheimer Genetic Epidemiology (MIRAGE) project estimated the risk of AD among 12.971 first-degree relatives of 1694 probable or definite AD patients, using survival analysis, and found the cumulative lifetime risk to be approximately twice that of the general population (~39% by age 96 years) (9). The risk was increased both for early-onset AD (EOAD) and late-onset AD (LOAD) relatives. Although the cumulative risk of dementia in African-American first-degree relatives of AD patients was found to be higher than that of the white population in the MIRAGE study, given the elevated risk in the spouses of the African-American population, the familial aggregation risk was similar in both ethnic groups (10).

Segregation analysis method can help distinguish transmissible environmental factor(s) from genetic factors and identify the underlying mode of inheritance for a disease. Segregation analyses studies in AD pedigrees revealed a Mendelian autosomal dominant transmission pattern for EOAD and a more complex model in LOAD, suggestive of multiple genes and possibly environmental effect (11,12). Twin studies, especially utilizing the Scandinavian twin registries, have provided heritability estimates for this disease. In the largest twin study to date, 392 twin pairs from the Swedish Twin Registry, where one or both members had AD, were assessed. The age-adjusted heritability of AD was found be 58-79% based on the analytical model used (13).

Genetics of Early-Onset Familial AD (EOFAD)

Three genes with autosomal dominant mutations that lead to EOFAD were identified, namely amyloid precursor protein (*APP*) on chromosome 21, presenilin 1 (*PSEN1*) on chromosome 14 and presenilin 2 (*PSEN2*) on chromosome 1. The details of all EOFAD mutations can be found in the Alzheimer Disease and Frontotemporal Dementia Mutation Database (http://www.molgen.ua.ac.be/ADMutations).

***APP* mutations:** The first evidence suggesting that chromosome 21 harbored an AD genetic risk region came from reports that patients with trisomy 21 (Down syndrome) invariably developed AD-like brain histopathology if they lived past age 40 (14). Aβ peptide isolated from brains of patients with AD and Down syndrome were found to be homologous (15). Identification of linkage to chromosome 21 in EOFAD families and mapping of APP to the same locus were followed by the identification of the first missense mutation in APP leading to EOFAD (16-20). In total, 32 EOFAD mutations in APP were reported in 86 families according to the Alzheimer Disease and Frontotemporal Dementia Mutation Database (http://www.molgen.ua.ac.be/ADMutations), including recently identified APP duplications (21). Functional evaluation of these mutations demonstrated their role in elevating Aβ levels (Aβ42 elevated with or without Aβ40 elevations) or increasing its fibrillogenesis (22). Age of onset in EOFAD patients with APP mutations ranges between 35-67 years, with some mutations leading to amyloid angiopathy and intracerebral hemorrhages. All EOFAD APP mutations are fully penetrant and constitute ~0.1% of all AD (23).

***PSEN* mutations:** Genome searches in EOFAD families led to the identification of genetic risk loci on chromosomes 14 and 1 (24,25). In 1995, EOFAD mutations in *PSEN1* on chromosome 14 and *PSEN2* on chromosome 1 were identified (26-28). With 177 mutations in 392 families, *PSEN1* mutations are the most common cause of EOFAD. Conversely, only 14 *PSEN2* mutations have been identified in 23 families. Like the EOFAD *APP* mutations, all *PSEN* mutations are fully penetrant. The mean age of onset of EOFAD caused by *PSEN1* mutations is ~45 (range: 24-60 years) and that of *PSEN2* is ~52 years (40-85 years).

All *PSEN* mutations lead to elevations in Aβ42 and/or decreased Aβ40, thereby increasing the Aβ42/Aβ40 ratio in favor of the more pathogenic form of Aβ (29,30). Collectively, EOFAD *PSEN* mutations constitute ~0.61% of all AD.

EOFAD mutations, amyloid cascade hypothesis and beyond: The major peptide constituent of senile plaques, Aβ, is cleaved from APP by two enzymatic processes, first by β-secretase (cleaves at the extracellular N-terminal domain) and then by the γ-secretase complex (cleaves at the intracellular C-terminal domain) (Figure 1). The γ-secretase enzymatic complex is composed of four different proteins, of which presenilin, which is a nine-transmembrane protein, is a required component. Identification of EOFAD mutations in the substrate and enzyme for Aβ, which invariably influence either its production or fibrillogenic potential, led to the amyloid cascade hypothesis, which suggests that increases in the toxic forms of Aβ lead to a cascade of events including inflammation, synaptic loss, ionic imbalance, and abnormal phosphorylation (possibly leading to neurofibrillary tangles), culminating in cell death, which is the likely pathologic substrate of clinical dementia (31). There are still many unknowns in this hypothesis, including exactly which form of Aβ (plaques, oligomers) constitutes the toxic species and the inter-relationship between Aβ and tau. There are also alternative hypotheses that suggest that a dominant-negative loss of function in presenilins, which have a multitude of functions besides γ-secretase cleavage of *APP*, may underlie the toxicity of both *APP* and *PSEN* mutations (29,30). Nonetheless, a large body of evidence implicating Aβ as a central player in the neurodegeneration in AD has led to multiple therapeutic trials focused on various steps of the amyloid cascade (8,31).

Genetics of late-onset AD (LOAD): The most common form of AD is LOAD, which is an active area of genetic research. Although Apolipoprotein E ε4 (ApoE ε4) is the only widely accepted genetic risk factor for LOAD, many promising genes emerging from linkage, candidate gene and genome-wide association studies (GWAS) are under investigation.

ApoE ε4: In 1991, linkage studies in AD families led to the identification of a risk locus on chromosome 19, with an especially strong effect in those AD families of late-onset (32). Identification of ApoE in senile plaques in AD brains, the discovery of its binding to Aβ with high avidity in vitro, and increased frequency of the ApoE ε4 allele in LOAD patients compared to controls established ApoE ε4 as a genetic risk factor for LOAD (33-36). The major studies focused on the impact of ApoE for population risk for AD and its potential role as a diagnostic and premorbid marker in AD are discussed in detail elsewhere and will be briefly mentioned here (23). Population-based genetic association studies in Caucasians using ApoE

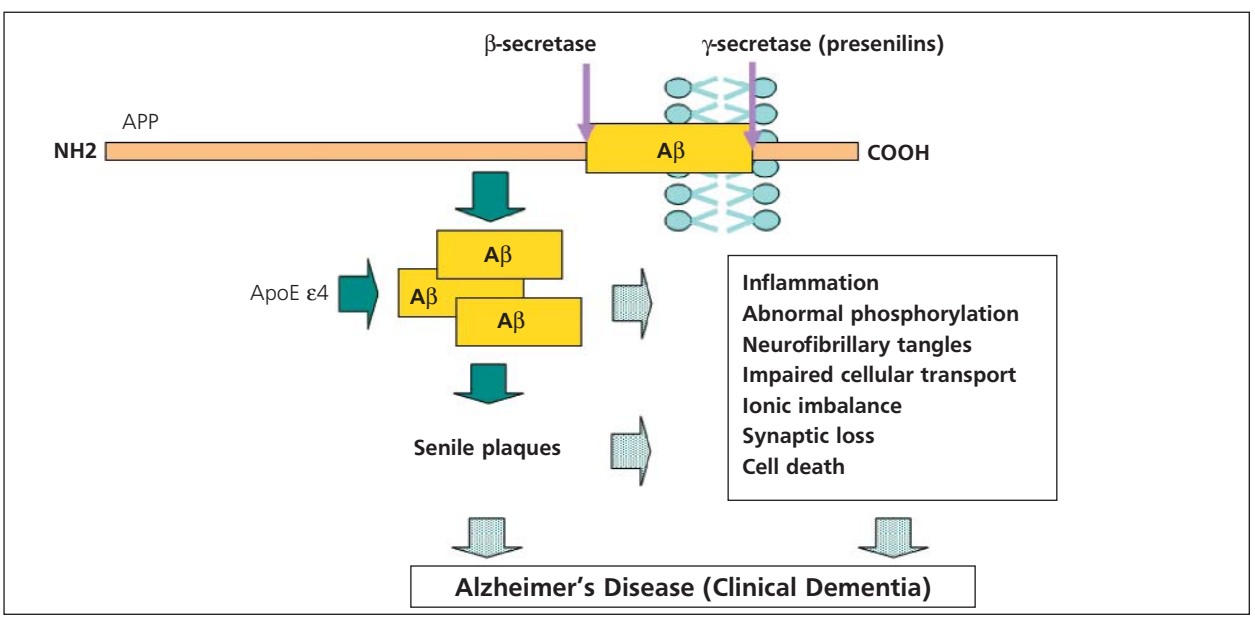

Figure 1. Pathophysiology of Alzheimer's disease - a simplified view: The known EOFAD (APP and presenilins) and LOAD (ApoE) gene products and their roles in the pathophysiology of AD are depicted. APP is the substrate from which Aβ is cleaved via β and γ-secretases. Presenilins are an integral component of the γ-secretase complex. ApoE ε4 leads to increased accumulation of Aβ. Toxic forms of Aβ (likely oligomers) lead to a number of detrimental processes shown in the box. The light arrows depict associations for which there is evidence but the details of which are still being worked out. There are likely numerous relationships between the events shown in the box, which is an over-simplified depiction.

ε3/ε3 genotype as the reference group revealed ~2-4 times increased risk [odds ratio (OR)] of AD in the ApoE ε3/ε4 genotype carriers and ~6-30 times increased risk in the ApoE ε4/ε4 genotype carriers. Although there was evidence of increased risk in other ethnic groups, including African-Americans and Hispanic populations, findings were less consistent among the different studies with smaller estimated effect sizes, suggesting different genetic and/or environmental risk factors at play for these non-Caucasian populations. The effect of ApoE ε4 appears to be age-dependent, with the strongest effect observed before age 70.

Unlike the deterministic, Mendelian EOFAD mutations, ApoE ε4 is a genetic risk modifier in LOAD, and hence is neither required nor sufficient for its development. ApoE genotyping does not contribute substantially to the diagnosis of AD, and its use in clinical practice is not recommended. There is evidence that ApoE genotypes can be useful predictive markers in longitudinally assessing development of AD from mild cognitive impairment (pre-dementia state). Currently, ApoE genotyping in mild cognitive impairment is pursued only for research purposes, and its use in clinical practice is not substantiated.

Whole genome linkage and association studies with microsatellite markers: The risk for AD or dementia attributable to ApoE ε4 is estimated to be 20-70%, strongly suggesting the presence of other factors accountable for the substantial genetic component of this disease (13,23). Between 1997 and 2006, 10 independent whole genome linkage (families or sib-pairs) and four association (case-controls) studies were carried out using microsatellite markers, reviewed in detail previously (23). Microsatellite markers are regions of repeating base-pair units in the genome with polymorphic (variable) repeat lengths. In these studies, typically ~200-400 microsatellite markers covering the whole genome at every 5-16 million base pairs (centiMorgans = cM) were genotyped in 10s-100s of AD families or sibships (n~100s-2000), followed by statistical analyses to identify areas of the genome that are shared between affected family members more often than would be expected by chance (linkage analyses). The four association studies were done on a mere 10-210 subjects (approximately equal numbers of cases and controls), comparing allelic frequencies between cases and controls. In general, the signals for the putative genetic loci for LOAD covered broad regions spanning several-tens of millions of bases and thus pose a challenge for downstream fine-mapping analyses.

All but two studies (one in inbred Arabs, another in Caribbean-Hispanics), implicated the ApoE locus on chromosome 19 as a genetic risk region for LOAD. Multiple other genetic loci emerged from these studies, some of which yielded signals stronger than that of the chromosome 19, ApoE locus. Some genetic loci were identified by multiple groups in independent studies. These results strongly suggest the presence of genetic risk factors, besides ApoE, in LOAD. The linkage and association studies using microsatellite markers revealed multiple, independent, strong signals on chromosomes 6, 9, 10 and 12, summarized previously and also on the AlzGene website (www.alzgene.org) (23,36). These results led to analyses of a multitude of candidate genes in these regions.

Candidate gene studies: Starting in the 1990s, there has been a plethora of literature focused on the assessment of candidate AD genes for their association with AD risk and/or its endophenotypes (such as age at onset, Aβ or tau levels, etc.). A summary of these studies and meta-analyses of the association studies, where possible, are provided in a regularly updated online database (AlzGene; www.alzgene.org) (36). According to the November 27, 2009 freeze of this database, 1236 studies have been published on 2335 polymorphisms in 598 AD candidate genes. Attempting a summary of these findings is beyond the scope of this review. Instead, a discussion of the potential problems in candidate gene studies and possible solutions is provided in this section.

In candidate AD gene studies, one or more genes are selected to be studied either because they are suitable biological candidates (based on in vitro, in vivo studies or their theoretical role in the pathophysiology of AD) or due to their physical location (based on whole genome linkage or association studies with microsatellite markers, and more recently GWAS), or both. Association studies in AD candidate genes have suffered from lack of consistent replication. Multiple reasons could account for this discouraging outcome, including a) Initial false-positive result, b) Small sample sizes that are underpowered to identify genetic factors of modest effect sizes (false-negative) and c) Genetic and environmental heterogeneity in the different study populations (36,37).

False-positive results could arise from multiple testing, population stratification, initial small sample size, genotyping errors, and not correcting for outliers in quantitative trait analyses. False-negative follow-up studies could be due to small sample sizes. Meta-analyses of available AD association studies revealed an estimated OR of < 2.0 for the putative AD risk variants summarized in the AlzGene database (36). It is estimated that thousands to 10.000s of samples are required to achieve sufficient power to detect associations for such modest effects. Until recently, most AD candidate gene association studies were conducted on much smaller sample sizes, which could be one explanation for the lack of replication. Typically, initial studies tend to overestimate the effect size of a genetic variant, a phenomenon known as the "winner's curse" (38). Given this, the true sample sizes required to capture the

effect of a genetic variant may be larger than what is estimated from the original study. Heterogeneity between the different study populations is another potential cause of false-negative findings. Association studies commonly test genetic variants that are markers (rather than the actual functional disease polymorphism) by taking advantage of existing linkage disequilibrium (LD) in the genome. The extent and strength of LD may be different in different populations. Furthermore, different risk genes, different risk alleles in the same gene, and different environmental and gene x environmental influences underlying disease risk in different populations also account for the heterogeneity between study populations.

Approaches to overcome these potential problems in association studies include: a) Careful selection of candidate genes and variants with increased a priori probability of association based on biology and position of the gene, b) corrections for multiple testing, c) Increasing study sizes guided by power calculations, d) Internal replication of findings in multiple series prior to publication, e) Testing and correction for population substructure, f) Attempt to decrease heterogeneity by using strict clinical/pathological disease criteria, subgroup analyses and Use of functional endophenotypes (quantitative biological phenotype), g) Use of multiple, informative and putative functional genetic markers, and h) Supplementing the genetic data by functional/biological analyses (36-38). Despite the lack of consistent replications, there are AD candidate genes with promising genetic association results and compelling underlying biological relevance (36).

Genome-wide association studies (GWAS): The most recent approach in uncovering the genetic underpinnings of AD, especially of late-onset, is GWAS. These studies are similar to microsatellite-based whole genome association studies in that they, too, are hypothesis-generating rather than being hypothesis-based. Unlike candidate gene association studies that focus on a handful of genes to test a hypothesis, GWAS are surveys of the whole genome for association signals. Single-nucleotide polymorphisms (SNPs) that capture information about the variation in the whole genome through the existing LD form the basis of GWAS. In the last two years, 10 independent GWAS have been conducted in LOAD (39-49). The study designs and results of these studies are summarized in Table 2. An eleventh study from Germany, not shown in Table 2, assessed 970 subjects and reported only their results on 11 LOAD candidate genes, none of which reached genome-wide significance (50).

There are some features that are common to these studies. First, except for the initial study, which was based on gene-centric, putative functional polymorphisms, they utilized ~300.000-600.000 SNPs on arrays (39). All but two studies utilized a case-control design (43,45). Two of the studies from Table 2 were performed on overlapping populations (40,49). Except for one study that utilized a small sample size from two extended LOAD pedigrees, the first seven studies had discovery series of ~700-2000 subjects (39-46). The same studies assessed ~400-3000 additional subjects to follow-up their initial findings. The last two studies merit special attention because of their increased power due to their study-wide combined sample sizes of > 14.000-16.000 subjects, which is more than three times the size of the next largest study (46-48). All studies were carried out in Caucasian populations from North America or Europe.

Two of the studies carried out their initial genotyping in pooled DNA, followed by individual genotyping of "interesting results" (39,42). This approach, while cost-effective, may have led to decreased sensitivity and false-negatives. The threshold for deeming results as "worthy of follow-up" differed between the studies. Two studies followed up an arbitrary number of their top hits in their replication series, whereas others utilized a significance threshold varying from "significant at genome- wide level" (i.e. corrected for the hundreds of thousands of SNPs genotyped) to a more relaxed arbitrary p-value cut-off (41,46). This variability stems from the fact that the extent of follow-up is based on cost, and that additional candidate gene leads could emerge if more SNPs can be assessed in follow-up series.

In all but the smallest study, ApoE-related SNPs reached genome-wide significance with p values ranging from 10^{-8} to 1.8×10^{-157} and ORs from ~2-4 (45). Harold et al.'s study is the only LOAD GWAS in which non-ApoE SNPs in two genes, *CLU* and *PICALM*, reached genome-wide significance in the first stage of the study (47). This is the largest GWAS to date, with > 11.000 subjects in the first stage. Importantly, the second largest LOAD GWAS also identified *CLU* at genome-wide significance in the combined first and second stage samples, in addition to CR1 (48). *CLU* encodes clusterin or ApoJ, one of the most abundant apolipoproteins in the human brain along with ApoE. In vivo studies suggest that clusterin, like ApoE, is involved in Aβ clearance from the brain (51). Although *PICALM* and *CR1* achieved the required significance thresholds in only one of the two largest LOAD GWAS, there was still evidence of association in the other study, providing additional support for these two genes. *PICALM* encodes a protein involved in clathrin-mediated endocytosis, a suggested pathway for trafficking of APP that could also influence Aβ formation (52). CR1 is a receptor for the complement component C3b, which has been suggested to be involved in the peripheral clearance of Aβ (53). Thus, all three candidate genes that emerged from the two largest LOAD GWAS to date have putative functions in the Aβ cascade.

Table 2a. Genome-wide association studies in AD - study designs

Reference	Ethnicity/Source	Samples	Study design	Genotyping platform	SNPs (a)	Discovery series ADs	Discovery series Controls	Replication series ADs	Replication series Controls	Follow-up criteria
Grupe et al.	UK/USA	Case-control	1 discovery (pooled DNA) and 5 replication series (1 pooled)	Gene-based putative functional polymorphisms	17.343	380	396	1428	1666	p values 0.15-0.005
Coon et al.	USA/Netherlands	Case-control	Single stage study	Affymetrix 500K	502.627	664	422	-	-	Overlaps with Reiman et al.
Reiman et al.	**USA/Netherlands**	**Case-control**	**1 discovery and 2 replication series**	**Affymetrix 500K**	**312.316**	**446**	**290**	**415**	**260**	**All SNPs genotyped in both stages**
Li et al.	Canada/UK	Case-control	1 discovery and 1 replication series	Affymetrix 500K	469.438	753	736	418	249	120 top SNPs
Abraham et al.	UK	Case-control	Single stage study (first pooled, then individual genotyping)	Illumina HumanHap300 and Illumina Sentrix HumanHap240S	561.494	1.082	1.239	-	1400	p≤ 0.05
Betram et al.	USA	Family-based	1 discovery and 3 replication series	Affymetrix 500K	484.522	941	404	1.767	838	Significance after weighted-Bonferroni correction
Beecham et al.	USA	Case-control	1 discovery and 1 replication series	Illumina HumanHap 550	532.000	492	498	238	220	Significance after FDR-BUM criteria
Podusio et al.	**USA**	**Family-based and case-control**	**1 discovery and 2 replication series (family-based and case-controls)**	**Affymetrix 500K**	**469.218**	**19 (family members)**	**60 (CEPH)**	**140 (family members) 199 (unrelated)**	**85 (unrelated)**	**Genome-wide significance after Bonferroni correction**
Carrasquillo et al.	**USA**	**Case-control**	**3 discovery and 4 replication series**	**Illumina HumanHap300**	**313.504**	**844**	**1.255**	**1.547**	**1.209**	**25 top SNPs**
Harold et al.	**USA/Europe**	**Case-control**	**13 discovery and 5 replication series**	**Illumina Human 610-Quad, HumanHap550 or HumanHap300**	**529.205 (up to)**	**3.941**	**7.848**	**2.023**	**2.340**	**Genome-wide significance after Bonferroni correction + 12 other CLU/PICALM SNPs**
Lambert et al.	**Europe**	**Case-control**	**1 discovery and 4 replication series (from 15 centers)**	**Illumina Human 610-Quad**	**537.029**	**2.032**	**5.328**	**3.978**	**3.297**	**< E-4**

The third largest GWAS identified an X-linked candidate gene *PCDH11X* encoding protocadherin 11, X-linked, at genome-wide significance in the combined Stage 1 and 2 series. *PCDH11X* is the first X-chromosomal candidate AD gene identified in a LOAD GWAS. Protocadherins belong to the superfamily of cadherins that are involved in cell adhesion, cell signaling and neural development. Protocadherins are predominantly expressed in the brain, suggesting their potential role in brain morphogenesis (54). Carrasquillo et al. found the largest effect of *PCDH11X* variants in female homozygotes, followed by female heterozygotes and then male hemizygotes (46). Though the functional role of this gene in AD needs to be established and the genetic effect confirmed through additional studies, it is an intriguing hypothesis that this X-chromosomal gene could explain the increased risk of AD in women.

The only other LOAD GWAS that yielded a signal significant at the genome-wide level after the conservative Bonferroni correction is the *GAB2* region identified in ApoE- ε4 positive subjects (40). GAB2 encodes a scaffolding protein, Grb2-associated binding protein 2, which is involved in cell signaling pathways, especially in the immune system (55). Its potential role in AD pathophysiology remains to be elucidated; however, preliminary functional studies revealed differential expression of GAB2 in AD vs. control brains, co-localization of GAB2 with dystrophic neurites and variation of GAB2 expression influencing tau phosphorylation (40).

The list of promising findings from all LOAD GWAS published to date is shown in Table 2b. One pattern that emerges from this table is that the strength of associations for ApoE-related SNPs and their effect sizes are bigger than those of non-ApoE-related SNPs. This may sug-

Table 2b. Genome-wide association studies in AD-results

		Non-ApoE hits		ApoE-related hits	
Reference	**Gene Symbol**	**p (b)**	**OR (b)**	**p (b)**	**OR (b)**
Grupe et al., February 2007	GALP, TNK1, chr14q32.13, PCK1, LMNA, PGBD1, LOC651924, chr7p15.2, THEM5, MYH13, CTSS, UBD, BCR, AGC1, TRAK2, EBF3	0.001 to 5.0E-5	1.07-1.2	7.6E-5 to 1.0E-8	1.19-2.73
Coon et al., April 2007	-	-	-	1.1E-39	4.01
Reiman et al., June 2007	**GAB2**	**9.7E-11**	**4.06**	**-**	**-**
Li et al., January 2008	GOLPH2; chr9p24.3; chr15q21.2	9.8E-3 to 4.5E-6 (c)	0.46-3.23 (c)	2.3E-44	-
Abraham et al., September 2008	LRAT	3.4E-6 to 6.1 E-7	1.2-1.3	4.8E-6 to 4.0E-14	-
Betram et al., November 2008	chr14q31.2; chr19q13.41	6.0E-6 to 2.0E-6	1.1-1.4 (d)	5.70E-14	-
Beecham et al., January 2009	12q13	3.40E-07	-	-	-
Poduslo, January 2009	**TRPC4AP**	**3.85E-10 to 5.63E-11 (d) 0.03 (f)**	**1.56 (f)**	**-**	**-**
Carrasquillo et al., February 2009	**PCDH11X**	**3.8E-8 (0.08 to 5.4E-13) (e)**	**1.29 (1.17-1.75) (e)**	**5.9E-6 to 3.7E-120**	**0.55-3.29**
Harold et al., September 2009	**CLU PICALM**	**8.5E-10 (CLU) 1.3E-9 (PICALM)**	**0.86 (CLU) 0.86 (PICALM)**	**3.4E-8 to 1.8E-157**	**0.63-2.5**
Lambert et al., September 2009	**CLU CR1**	**7.5E-9 (CLU) 3.7E-9 (CR1)**	**0.86 (CLU) 1.21 (CR1)**	**5.06E-7 to < 2E-16**	**-**

Table 2 Genome-wide association studies in AD, a. Study Designs, b. Results. The study designs and results of the 10 independent LOAD GWAS studies are depicted (Coon et al. and Reiman et al. studies are overlapping). The studies that yield non-ApoE associations that are significant at the genome-wide level after Bonferroni corrections are in bold. (a) Number of SNPs in the initial genotyping stage. (b) Results from all groups combined. (c) Results shown separately in each series. (d) Results from discovery series. (e) Results vary based on different analytical models. (f) Results in follow-up case-control series.

gest that the non-ApoE genetic factors underlying LOAD may be common variants with smaller effect sizes than *ApoE*. For this reason, it will likely require thousands to tens of thousands of subjects to have sufficient power to validate these results.

Based on the two large LOAD GWAS, the population attributable risks of *CLU, PICALM* and *CR1* are ~ 9%, 9% and 4%. Given that the population attributable risk for ApoE-related findings from the LOAD GWAS is ~19-35%, the novel findings from the two recent, most powerful LOAD GWAS and ApoE can account for at most 57% of the population attributable risk of AD. Furthermore, because the effect sizes of these novel genes are likely overestimates in these initial studies, the true risk explained by them is most likely smaller than these estimates. The remaining genetic risk for AD could be due to variants in other genes such as *PCDH11X* or *GAB2*. Some of the other possibilities that explain the missing genetic component underlying AD include presence of rare variants with larger effect sizes, structural variations, presence of different genetic factors in non-Caucasian populations, and gene-gene and gene-environment interactions.

Future studies: With the identification of numerous AD candidate genes, especially through the LOAD GWAS, these findings need further validation via additional genetic association studies as well as functional assessment by in vitro and in vivo approaches. The list of replication studies on these candidate genes are continually updated in AlzGene (36). These genetic studies need to be interpreted with caution, ensuring that they are empowered to detect the modest effect sizes suggested by the original studies. Use of biologically relevant quantitative phenotypes (endophenotypes) may be an important, potentially powerful alternative approach in genetic studies of common and complex diseases. Since the use of plasma Aβ levels as the first endophenotype in LOAD genetics, additional endophenotypes, such as cognitive measures and neuroimaging phenotypes, have led to the identification of candidate genes and regions in AD (56-58). The combination of existing endophenotype information on subjects who already have genotypes at the genome-wide level will allow further utilization of GWAS data. It is becoming increasingly clear that variations, other than SNPs, may underlie our complex traits and diseases (59). Copy number variations (CNVs), such as insertions, deletions, translocations and inversions, could theoretically account for some of the underlying risk for AD. Potential associations with CNVs need to be assessed both via mining available GWAS data utilizing specialized analytic tools as well as use of novel genotyping platforms specifically geared towards capturing CNVs. As more and more GWAS and other large scale association and linkage studies are becoming available, it will be important to jointly assess the results of these studies to capitalize on the cumulative knowledge that can be gained from such meta-analyses (60). It is important to recognize that the variations where AD association or linkage is demonstrated are likely markers and not the actual functional polymorphisms. This may account for the lack of replication across studies. Thus, the true functional polymorphisms need to be searched using the next-generation sequencing approach that can provide resolution of a genetic sequence down to a single base pair. The last two decades witnessed discoveries in the genetics of AD that shed light on the pathophysiology of this disease and yielded intriguing leads for follow-up. With the advent of novel technological and analytic approaches, use of increasing sample sizes and combined analyses of existing data, genetics of AD is expected to contribute to our understanding of this disease, which may translate into therapeutic and preventative advances for this 21st century epidemic.

REFERENCES

1. *Alzheimer A. A new disease of the cortex (Ger). Allg Z Psychiatr 1907;64:146-8.*
2. *Perrin RJ, Fagan AM, Holtzman DM. Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. Nature 2009;461:916-22.*
3. *Ferri CP, Sousa R, Albanese E, Ribeiro WS, Honyashiki M. World Alzheimer Report 2009 Executive Summary. In: Prince M, Jackson J (eds). Alzheimer's Disease International, 2009:1-22.*
4. *Wimo A, Winblad B, Jonsson L. An estimate of the total worldwide societal costs of dementia in 2005. Alzheimers Dement 2007;3:81-91.*
5. *Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. Am J Public Health 1998;88:1337-42.*
6. *Mayeux, R. Epidemiology of neurodegeneration. Annu Rev Neurosci 2003;26:81-104.*
7. *Whalley LJ, Dick FD, McNeill G. A life-course approach to the aetiology of late-onset dementias. Lancet Neurol 2006;5:87-96.*
8. *Rafii MS, Aisen PS. Recent developments in Alzheimer's disease therapeutics. BMC Med 2009;7:7.*
9. *Lautenschlager NT, Cupples LA, Rao VS, Auerbach SA, Becker R, Burke J, et al. Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: What is in store for the oldest old? Neurology 1996;46:641-50.*
10. *Green RC, Cupples LA, Go R, Benke KS, Edeki T, Griffith PA, et al. Risk of dementia among white and African American relatives of patients with Alzheimer disease. JAMA 2002;287: 329-36.*
11. *Farrer LA, Myers RH, Connor L, Cupples LA, Growdon JH. Segregation analysis reveals evidence of a major gene for Alzheimer disease. Am J Hum Genet 1991;48:1026-33.*
12. *Rao VS, van Duijn CM, Connor-Lacke L, Cupples LA, Growdon JH, Farrer LA. Multiple etiologies for Alzheimer disease are revealed by segregation analysis. Am J Hum Genet 1994;55:991-1000.*

13. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 2006;63:168-74.
14. Pary RJ, Rajendran G, Stonecipher A. Overview of Alzheimer's disease in down syndrome. In: Prasher VP (ed). Down Syndrome and Alzheimer's Disease Biological Correlates. Oxen, UK: Radcliffe Publishing, 2006:2-13.
15. Glenner GG, Wong CW. Alzheimer's disease and Down's syndrome: Sharing of a unique cerebrovascular amyloid fibril protein. Biochem Biophys Res Commun 1984;122:1131-5.
16. St George-Hyslop PH, Tanzi RE, Polinsky RJ, Haines JL, Nee L, Watkins PC, et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science 1987;235: 885-90.
17. Goate AM, Haynes AR, Owen MJ, Farrall M, James LA, Lai LY, et al. Predisposing locus for Alzheimer's disease on chromosome 21. Lancet 1989;1:352-5.
18. Goldgaber D, Lerman MI, McBride OW, Saffiotti U, Gajdusek DC. Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. Science 1987;235:877-80.
19. Tanzi RE, Gusella JF, Watkins PC, Bruns GA, St George-Hyslop P, Van Keuren ML, et al. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. Science 1987;235:880-4.
20. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 1991;349:704-6.
21. Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet 2006;38:24-6.
22. Theuns J, Marjaux E, Vandenbulcke M, Van Laere K, Kumar-Singh S, Bormans G, et al. Alzheimer dementia caused by a novel mutation located in the APP C-terminal intracytosolic fragment. Hum Mutat 2006;27:888-96.
23. Ertekin-Taner N. Genetics of Alzheimer's disease: A centennial review. Neurol Clin 2007;25:611-67.
24. Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science 1992;258: 668-71.
25. Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, Weber JL, et al. A familial Alzheimer's disease locus on chromosome 1. Science 1995;269:970-3.
26. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature 1995;375: 754-60.
27. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science 1995;269:973-7.
28. Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature 1995;376:775-8.
29. Shen J, Kelleher RJ 3rd. The presenilin hypothesis of Alzheimer's disease: Evidence for a loss-of-function pathogenic mechanism. Proc Natl Acad Sci U S A 2007;104:403-9.
30. Wakabayashi T, De Strooper B. Presenilins: Members of the gamma-secretase quartets, but part-time soloists too. Physiology (Bethesda) 2008;23:194-204.
31. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. Science 2002;297:353-6.
32. Pericak-Vance MA, Bebout JL, Gaskell PC, Yamaoka LH, Hung WY, Alberts MJ. Linkage studies in familial Alzheimer disease: Evidence for chromosome 19 linkage. Am J Hum Genet 1991;48:1034-50.
33. Namba Y, Tomonaga M, Kawasaki H, Otomo E, Ikeda K. Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease. Brain Res 1991;541: 163-6.
34. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS. Apolipoprotein E: High-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc Natl Acad Sci USA 1993;90:1977-81.
35. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993;261:921-3.
36. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: The AlzGene database. Nat Genet 2007;39:17-23.
37. Newton-Cheh C, Hirschhorn JN. Genetic association studies of complex traits: Design and analysis issues. Mutat Res 2005;573:54-69.
38. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 2007;33:177-82.
39. Grupe A, Abraham R, Li Y, Rowland C, Hollingworth P, Morgan A, et al. Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. Hum Mol Genet 2007;16:865-73.
40. Reiman EM, Webster JA, Myers AJ, Hardy J, Dunckley T, Zismann VL, et al. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron 2007;54:713-20.
41. Li H, Wetten S, Li L, St Jean PL, Upmanyu R, Surh L, et al. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. Arch Neurol 2008;65: 45-53.
42. Abraham R, Moskvina V, Sims R, Hollingworth P, Morgan A, Georgieva L, et al. A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. BMC Med Genomics 2008;1:44.
43. Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M, Hogan MF, et al. Genome-wide Association Analysis Reveals Putative Alzheimer's Disease Susceptibility Loci in Addition to APOE. Am J Hum Genet 2008.
44. Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, et al. Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. Am J Hum Genet 2009;84:35-43.
45. Poduslo SE, Huang R, Huang J, Smith S. Genome screen of late-onset Alzheimer's extended pedigrees identifies TRPC4AP by haplotype analysis. Am J Med Genet B Neuropsychiatr Genet 2009;150B:50-55.

46. *Carrasquillo MM, Zou F, Pankratz VS, Wilcox SL, Ma L, Walker LP, et al. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. Nat Genet 2009;41:192-198.*

47. *Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 2009.*

48. *Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 2009.*

49. *Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Lince DH, et al. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. J Clin Psychiatry 2007;68:613-8.*

50. *Feulner TM, Laws SM, Friedrich P, Wagenpfeil S, Wurst SH, Riehle C, et al. Examination of the current top candidate genes for AD in a genome-wide association study. Mol Psychiatry 2009.*

51. *Holtzman DM. In vivo effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology. J Mol Neurosci 2004;23:247-54.*

52. *Wu F, Yao PJ. Clathrin-mediated endocytosis and Alzheimer's disease: An update. Ageing Res Rev 2009;8:147-9.*

53. *Rogers J, Li R, Mastroeni D, Grover A, Leonard B, Ahern G, et al. Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes. Neurobiol Aging 2006;27:1733-9.*

54. *Blanco P, Sargent CA, Boucher CA, Mitchell M, Affara NA. Conservation of PCDHX in mammals; expression of human X/Y genes predominantly in brain. Mamm Genome 2000;11:906-14.*

55. *Sarmay G, Angyal A, Kertesz A, Maus M, Medgyesi D. The multiple function of Grb2 associated binder (Gab) adaptor/scaffolding protein in immune cell signaling. Immunol Lett 2006;104:76-82.*

56. *Ertekin-Taner N, Graff-Radford N, Younkin LH, Eckman C, Baker M, Adamson J, et al. Linkage of plasma Aß42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. Science 2000;290:2303-4.*

57. *Papassotiropoulos A, Stephan DA, Huentelman MJ, Hoerndli FJ, Craig DW, Pearson JV, et al. Common Kibra alleles are associated with human memory performance. Science 2006;314: 475-8.*

58. *Potkin SG, Guffanti G, Lakatos A, Turner JA, Kruggel F, Fallon JH, et al. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. PLoS One 2009;4:6501.*

59. *Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.*

60. *Khoury MJ, Bertram L, Boffetta P, Butterworth AS, Chanock SJ, Dolan SM, et al. Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. Am J Epidemiol 2009;170:269-79.*

Yazışma Adresi/Address for Correspondence

Yrd. Doç. Dr. Nilüfer Ertekin Taner

Departments of Neurology and Neuroscience,
Mayo Clinic Florida,
Jacksonville, Florida/USA

E-posta: taner.nilufer@mayo.edu

geliş tarihi/received 03/12/2009
kabul ediliş tarihi/accepted for publication 15/12/2009