REVIEW

Genetics of Alzheimer disease in the pre- and post-GWAS era

Nilüfer Ertekin-Taner*

Abstract

Since the 1990s, the genetics of Alzheimer disease (AD) has been an active area of research. The identification of deterministic mutations in the *APP*, *PSEN1*, and *PSEN2* genes responsible for early-onset autosomal dominant familial forms of AD led to a better understanding of the pathophysiology of this disease. In the past decade, the plethora of candidate genes and regions emerging from genetic linkage and smaller-scale association studies yielded intriguing 'hits' that have often proven difficult to replicate consistently. In the last two years, 11 published genome-wide association studies (GWASs) in AD confirmed the universally accepted role of *APOE* as a genetic risk factor for late-onset AD as well as generating additional candidate genes that require confirmation. It is unclear whether GWASs, though a promising novel approach in the genetics of complex diseases, can help explain most of the underlying genetic risk for AD. This review provides a brief summary of the genetic studies in AD preceding the GWAS era, with the main focus on the findings from recent GWASs. Potential approaches that could provide further insight into the genetics of AD in the post-GWAS era are also discussed.

Genetics of Alzheimer disease in the pre-GWAS era

Early-onset familial Alzheimer disease

Alzheimer disease (AD), the most common dementia, currently affects an estimated 35 million patients worldwide [1] and is characterized by extracellular accumulation of the amyloid β (Aβ) peptide in senile plaques and intracellular accumulation of the abnormally hyper-phosphorylated tau forming neurofibrillary tangles in the brain [2,3]. There is evidence from familial aggregation, transmission patterns, and twin studies that AD has a

*Correspondence: taner.nilufer@mayo.edu
Mayo Clinic Florida, Departments of Neurology and Neuroscience, 4500 San Pablo Road, Birdsall 210, Jacksonville, FL 32224 USA

substantial genetic component that has an estimated heritability of 58% to 79% [4], and the lifetime risk of AD in first-degree relatives of patients may be twice that of the general population [5]. Families with autosomal dominant transmission of AD were described in the literature in the 1980s [6]. In the early 1990s, segregation analysis studies suggested the presence of Mendelian, autosomal, dominant risk factors underlying the risk of early-onset AD, whereas a more complex model possibly involving polygenes and environmental factors emerged for late-onset AD (LOAD) [7,8]. The identification of homology in the Aβ peptide isolated from brains of patients with AD and trisomy 21 (Down syndrome) and localization of the amyloid precursor protein (*APP*) to chromosome 21 [9,10], where linkage to disease risk was mapped in early-onset familial AD (EOFAD) families [11,12], led to the discovery of first autosomal dominant missense mutations in *APP* segregating with disease risk [13]. This was followed by identifications of autosomal dominant EOFAD mutations in the presenilin 1 (*PSEN1*) [14] and *PSEN2* [15,16] genes, on chromosomes 14 and 1, respectively. The summary of EOFAD mutations is maintained at the Alzheimer Disease and Frontotemporal Dementia Mutation Database [17,18]. Accordingly, there are currently 32 mutations in *APP*, 177 in *PSEN1*, and 14 in *PSEN2*, identified in 86, 392, and 23 families, respectively. Collectively, the EOFAD mutations in these three genes account for less than 1% of all AD.

Aβ, the major peptide constituent of senile plaques, is cleaved from APP first by β-secretase, then by the γ-secretase complex, of which presenilin is a required component. Despite the rarity of *APP* and *PSEN* mutations, their functional evaluation in transfected cells [19,20], transgenic animals [21,22], and human plasma [23] identified elevations in Aβ levels, increased Aβ42/Aβ40 ratio, or fibrillogenesis, which constituted the cornerstone of the amyloid cascade hypothesis [24]. Accordingly, increases in the toxic forms of Aβ lead to a cascade of events – including inflammation, synaptic loss, ionic imbalance, and abnormal phosphorylation of proteins (including tau) – culminating in cell death and underlying clinical dementia. There exist alternative hypotheses suggesting tau [25] or dominant negative loss

of presenilin function [26] as additional pathophysiologic mechanisms underlying AD.

Late-onset Alzheimer disease

The identification of a chromosome 19 risk locus in LOAD families [27] was followed by the discovery of a higher frequency of the *APOE* ε4 allele in LOAD patients compared with controls [28,29]. *APOE* is a component of senile plaques [30], binds Aβ [31], can influence neuritic plaque formation in transgenic mouse models of AD in an isoform-specific fashion [32], and is thought to contribute to both Aβ clearance and deposition in the brain [33]. *In vitro* and *in vivo* studies also suggest a role for *APOE* in isoform-specific synaptogenesis and cognition, neurotoxicity, tau hyperphosphorylation, neuro-inflammation, and brain metabolism, although these non-Aβ-related mechanisms require further investigation [34]. Unlike the EOFAD mutations that are fully penetrant, *APOE* ε4 is a genetic risk factor that is neither necessary nor sufficient for the development of AD (see [35] for review of *APOE* in LOAD). The odds that *APOE* ε3/ε4 genotype carriers have AD is estimated to be two to four times greater than that of *APOE* ε3/ε3 carriers, according to population-based association studies in subjects of European origin. The odds ratio (OR) increases to approximately 6 to 30 in the *APOE* ε4/ε4 genotype carriers. Although there is evidence of a risk effect of *APOE* ε4 in non-Europeans, the estimated effect sizes are smaller with less consistent results in African-American and Hispanic subjects, which may suggest different underlying genetic or environmental factors or both for these ethnic groups. The effect of *APOE* ε4 appears to be age-dependent, with the strongest effect observed before age 70. The use of *APOE* as a diagnostic [36] or predictive factor in clinical practice is not warranted. Population attributable risk (PAR) of AD due to a genetic factor describes the difference in rate of AD in the population between those who are carriers of the genetic risk versus those who are not. Another way to consider PAR is the amount of decrease that would be observed in the incidence of AD if the genetic factor could be eliminated. It is important to note that the PAR due to a combination of risk factors is usually smaller than the sum of PARs from multiple risk factors because a person with disease may have more than one risk factor (that is, the cause of disease in any given case may be attributable to more than one factor). The estimated LOAD PAR for *APOE* is 20% to 70% [4,35]. This relatively wide range of PAR estimates for *APOE* likely reflects the difference in study types (population-based versus clinic-based, longitudinal versus cross-sectional), age range of the population, and other characteristics such as family history. PAR estimates may have an upward bias in family-based studies and younger populations given that

APOE has a stronger effect in younger ages (from younger than age 70 to eighties) and longitudinal studies may provide more accurate estimates of PAR of *APOE* where estimates of 20% were obtained [35,37,38]. Statistical approaches suggesting the presence of AD loci in addition to *APOE* [39] led to intensive research efforts in the genetics of LOAD.

Linkage and smaller-scale association studies

Generation of the genetic and physical map of the human genome [40] largely based on polymorphic tandem repeat regions led to the first-generation genome-wide scans in AD as well as many other diseases. Ten independent whole-genome linkage and four association studies in AD were completed using microsatellite markers (reviewed in [35]) between 1997 and 2006. These studies typically used approximately 200 to 400 microsatellite markers, covering the genome at every 5 to 16 centimorgans (1 cM is approximately equal to 1 million base pairs). The whole-genome linkage studies were conducted in AD families or sibships collectively composed of about 100 to 2,000 subjects. Whole-genome association studies were conducted on AD case-control series composed of a small number of subjects (n is approximately equal to 10 to 210, including an approximately equal number of AD patients and controls). These studies were hypothesis-free (or hypothesis-generating) in that they provided a survey of the whole genome without requiring any prior knowledge or theory about the genes/genomic regions for their role in AD. All of the studies that were done in subjects of European origin identified the *APOE* region on chromosome 19 as a risk region in AD. Importantly, however, they also detected other genetic regions, at times with a signal stronger than that of *APOE*. Some of these regions were detected in multiple independent studies, strongly suggesting the presence of non-*APOE* genetic factors underlying the risk of AD. Loci on chromosomes 6, 9, 10, and 12 [29,35] yielded multiple, independent, strong signals that led to subsequent fine-mapping efforts, which have proven difficult due to the typically large regions identified in these studies covering tens of millions of base pairs. Nonetheless, these findings generated positional and functional candidate genes that were assessed in association studies.

The identification of single-nucleotide polymorphisms (SNPs) in the human genome [41] and development of high-throughput SNP genotyping technologies galvanized such studies to understand the underlying genetics of common, complex diseases. Since the 1990s, more than 1,000 association studies have been carried out in nearly 600 AD candidate genes, regularly updated in the AlzGene website [42], which is a database of genetic linkage and association studies in AD [29]. Some of the AD candidate gene studies led to the identification of

genes with plausible functions in AD pathophysiology and evidence of association with AD risk in relatively large series. Sortilin-related receptor (*SORL1*) [43] and calcium homeostasis modulator 1 (*CALHM1*) [44] are two such genes. Variations in both genes were assessed in more than 3,000 to 4,000 subjects from multiple independent family-based or case control series in these studies. Furthermore, considerable functional analyses implicated SORL1 in APP sorting [43] and CALHM1 in calcium signaling that also influences Aβ levels [44]. Despite these initial results and some supportive follow-up studies from independent groups, there have also been reports of negative or weaker association for these genes summarized in AlzGene. Inconsistent replication of original association findings has been the rule rather than the exception in AD, except for *APOE*, and even for candidate genes with convincing functional data and thorough genetic assessment with positive association results in the initial study.

Potential problems in candidate gene association studies of AD and proposed solutions are discussed elsewhere [29,45,46] and will be briefly mentioned here. The potential causes for inconsistent replication of association studies in AD (and other common, complex diseases) include initial false-positive results, false-negative follow-up results, and heterogeneity in phenotype, genotype, or environment. Multiple testing, population stratification, genotyping errors, and initial small sample size are potential reasons for false-positive findings in the original study. Underpowered studies that are too small to detect a modest effect size can lead to false-negative follow-up studies. Meta-analyses of association studies in AD [29] and other common diseases reveal modest estimated effect sizes with ORs of less than 2.0 (or greater than 0.5 for protective alleles) [47]. It is estimated that thousands to tens of thousands of subjects are required to have sufficient power to detect such effect sizes, a prerequisite that has typically not been fulfilled for many association studies in AD until recently.

Genetic heterogeneity in the form of both different disease genes and different alleles (in the same genes) underlying disease risk could account for failure of replication in association studies of common, complex diseases. The situation becomes even more complicated when potential gene-gene and gene-environment interactions are considered. Finally, given that association studies rely on the existing linkage disequilibrium (LD) between the genotyped markers and the actual functional disease susceptibility variant (typically not genotyped), heterogeneity in the extent of this LD between different study populations could also account for inconsistent results between different studies. Study of isolated populations or disease groups or both with distinct or extreme phenotypes could be used in an attempt to decrease genic, allelic, phenotypic, and environmental heterogeneity but may come at the expense of limited generalizability of study results. Use of multiple, informative, putative functional markers; testing for and correcting population substructure; and use of biologically relevant, quantitative phenotypes (endophenotypes) in addition to the disease phenotype are additional approaches to address heterogeneity and potentially to increase power in association studies. Testing multiple independent series internally prior to publication, supplementing genetic data with functional assays, and choice of candidate genes and variants with increased *a priori* probability of association based on biology and position of the gene can increase the probability of success in association studies of common diseases such as AD.

Genetics of Alzheimer disease in the GWAS era

The International HapMap Project, launched in October 2002, led to the generation of a database of the common variations (defined as minor allele frequency of greater than 0.05) and the underlying LD structure in the human genome [48,49] that provided the foundation for the genome-wide association studies (GWASs) that use high-throughput genotyping platforms composed of common SNP markers that tag a subset of the known common SNPs in the human genome. For example, the Affymetrix GeneChip 500K platform (Affymetrix, Santa Clara, CA, USA) has 68% coverage of the phase II HapMap SNPs at r^2 of at least 0.8 in the subjects of European ancestry (CEU), whereas the Illumina Hap300 platform (Illumina, Inc., San Diego, CA, USA) has 77% coverage [49]. Furthermore, coverage is less in non-European populations, up to 1% of the common variations are untaggable by other SNPs using LD, and there likely exist millions of 'rare' SNPs not covered by these platforms. Despite their limitations, GWASs allowed a higher-resolution screen of the human genome for common diseases and traits. In the last two years, 11 GWASs in LOAD have been published. In this section, brief summaries of each of these studies are provided, followed by collective conclusions.

Study-specific characteristics and results of the late-onset Alzheimer disease GWASs

The study designs and results of the 11 LOAD GWASs are depicted in detail in Tables 1 and 2, respectively [50-61]. The first published LOAD GWAS used a select set of 17,343 SNPs from 11,211 genes that were chosen based on their likelihood of being functional polymorphisms [50]. These SNPs were weighted heavily toward missense mutations, although there were also variants in transcription factor-binding sites, introns, intergenic regions as well as other putative functional SNPs. Authors used a multi-stage design, in which the

Table 1. Late-onset Alzheimer disease genome-wide association study: summary of methodologies.

Reference	Ethnicity/ Source	Samples	Study design	Genotyping platform	SNPs[a]	Discovery series		Replication series		Follow-up criteria
						AD patients	Controls	AD patients	Controls	
Grupe et al. [50], February 2007	UK/US	Case control	1 discovery (pooled DNA) and 5 replication series (1 pooled)	Gene-based putative functional polymorphisms	17,343 (in 11,211 genes)	380	396	1,428	1,666	$P <0.075$ or $P <0.1$ and biology (stage 1) $P <0.15$ and same risk allele (stage 2) $P <0.05$ in UK1 and WU and P meta <0.005 (stage 3)
Coon et al. [51], April 2007	US/The Netherlands	Case control	Single-stage study	Affymetrix 500K	502,627	664	422	-	-	Single-stage
Reiman et al. [52], June 2007[b]	US/The 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. [53], January 2008	Canada/UK	Case control	1 discovery and 1 replication series	Affymetrix 500K	469,438	753	736	418	249	Top 120 SNPs
Abraham et al. [54], September 2008	UK	Case control	Single-stage study (first pooled, then individual genotyping)	Illumina HumanHap300 and Illumina Sentrix HumanHap240S	561,494	1,082	1,239	-	1,400	Passed three criteria and could be genotyped (stage 1) $P \leq 0.05$ (stage 2)
Bertram et al. [55], November 2008	US	Family-based	1 discovery and 3 replication series	Affymetrix 500K	404,604	941	404	1,767	838	Significance after weighted-Bonferroni correction
Beecham et al. [56], January 2009	US	Case control	1 discovery and 1 replication series	Illumina HumanHap550	532,000	492	496	238	220	Significance after FDR-BUM criteria
Feulner et al. [57], January 2009	Germany	Case control	Single-stage study	Illumina HumanHap550	555,000	491	479	-	-	Single-stage
Poduslo et al. [58], January 2009[b]	US	Family-based and case control	1 discovery and 2 replication series	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. [59], February 2009[b]	US	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. [60], September 2009[b]	US/Europe	Case control	13 discovery and 5 replication series	Illumina 610-quad, Illumina 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. [61], September 2009[b]	Europe	Case control	1 discovery and 4 replication series (15 centers)	Illumina Human 610-Quad BeadChip	537,029	2,032	5,328	3,978	3,297	$P <10^{-5}$

The study designs of the 11 independent late-onset Alzheimer disease genome-wide association studies are depicted. The Coon et al. [51] and Reiman et al. [52] studies are overlapping. Abraham et al. [54], Grupe et al. [50], and Harold et al. [60] also have overlapping samples. Carrasquillo et al. [59] contributed data to the Harold et al. [60] study. [a]Number of single-nucleotide polymorphisms (SNPs) in the initial genotyping stage. [b]Studies that yield non-*APOE* associations that are significant at the genome-wide level after Bonferroni corrections. AD, Alzheimer disease; CEPH, Centre d'Etude du Polymorphisme Humain; FDR-BUM, false discovery rate–beta uniform mixture.

Table 2. Late-onset Alzheimer disease genome-wide association study: summary of results.

Reference	Non-ApoE hits			ApoE-related hits	
	Gene symbol	P value[a]	Odds ratio[a]	P value[a]	Odds ratio[a]
Grupe *et al.* [50]	*GALP, TNK1, chr14q32.13, PCK1, LMNA, PGBD1, LOC651924, chr7p15.2, THEM5, MYH13, CTSS, UBD, BCR, AGC1, TRAK2, EBF3*	0.001 to 5.0×10^{-5}	1.07 to 1.2	7.6×10^{-5} to 1.0×10^{-8}	1.19 to 2.73
Coon *et al.* [51]				1.1×10^{-39}	4.01
Reiman *et al.* [52]	*GAB2*	9.7×10^{-11}	4.06	-	-
Li *et al.* [53]	*GOLPH2, chr9p24.3, chr15q21.2*	9.8×10^{-3} to $4.5 \times 10^{-6,b}$	0.46 to 3.23[b]	2.3×10^{-44}	-
Abraham *et al.* [54]	*LRAT*	3.4×10^{-6} to 6.1×10^{-7}	1.2 to 1.3	4.8×10^{-6} to 4.0×10^{-14}	-
Bertram *et al.* [55]	*chr14q31.2, chr19q13.41*	6.0×10^{-6} to 2.0×10^{-6}	1.1 to 1.4[c]	5.70×10^{-14}	-
Beecham *et al.* [56]	*12q13*	3.40×10^{-7}	-	-	-
Feulner *et al.* [57]	*MAPT, SORL1, CHRNB2, CH25H, GAB2, PGBD1, PCK1, LMNA*	0.05 to 6.8×10^{-3}	-	$<1.0 \times 10^{-6}$ to $<1.0 \times 10^{-40}$	-
Poduslo *et al.* [58]	*TRPC4AP*	3.85×10^{-10} to $5.63 \times 10^{-11,c}$ 0.03^d	1.56[d]	-	-
Carrasquillo *et al.* [59]	*PCDH11X*	3.8×10^{-8} (0.08 to 5.4×10^{-13})[e]	1.29 (1.17 to 1.75)[e]	5.9×10^{-6} to 3.7×10^{-120}	0.55 to 3.29
Harold *et al.* [60]	*CLU* *PICALM*	8.5×10^{-10} (CLU) 1.3×10^{-9} (PICALM)	0.86 (CLU) 0.86 (PICALM)	3.4×10^{-8} to 1.8×10^{-157}	0.63 to 2.5
Lambert *et al.* [61]	*CLU* *CR1*	7.5×10^{-9} (CLU) 3.7×10^{-9} (CR1)	0.86 (CLU) 1.21 (CR1)	5.06×10^{-7} to $<2 \times 10^{-16}$	-

The results of the 11 independent late-onset Alzheimer disease genome-wide association studies are depicted. The results from the original manuscripts shown in the table are (a) from all groups combined, (b) shown separately in each series, (c) from the discovery series, (d) from the follow-up case control series, and (e) variable based on different analytical models.

SNPs were genotyped in pooled DNA from 380 LOAD and 396 control subjects in the first stage. One thousand five hundred forty-four SNPs with AD association at *P* values of less than 0.075 or less than 0.1 and in biological candidate genes were genotyped in a second pool of 376 LOAD and 344 control subjects. Of these SNPs, 119 showed association at a *P* value of less than 0.15 with the same risk allele associating in the second pool. These 119 SNPs were genotyped individually in a total of four series, including the two initial pooled series. Eighteen of these SNPs that had a *P* value of less than 0.05 in the two pooled series and meta-analysis *P* value of less than 0.005 in the four series combined were also genotyped in a fifth series. Altogether, nearly 4,000 subjects were genotyped in this study, although the multi-stage design with two pooled DNA series in the initial stages considerably reduced the cost of this study. Nonetheless, some associations may have been missed due to the reduced sensitivity of the pooling approach and the limited number of SNPs assessed. Despite these shortcomings, *APOE*-related SNPs (SNPs in *APOE* or in LD with SNPs in *APOE*) were identified with study-wide significance in this study. Fifteen additional non-*APOE* SNPs were identified with nominal significance after meta-analyses of all five series, with modest ORs of 1.07 to 1.2 in all series combined. Of these SNPs, one was a missense mutation (rs3745833) in the galanin-like peptide

precursor (*GALP*) gene on chromosome 19 which showed suggestive study-wide significance. The galanin gene has been implicated in neuronal survival, regeneration, and neuroprotection and inhibition of learning and memory through reductions in glutamate release and reduction of long-term potentiation (reviewed in [62]). Galanin expression has been shown to be upregulated in AD brains, although it is unclear whether this is a cause or effect. In this study, the authors highlighted additional genes among the final 15 – such as phosphoenolpyruvate carboxykinase 1 (*PCK1*), trafficking protein, kinesin binding 2 (*TRAK2*), and tyrosine kinase, non-receptor 1 (*TNK1*) – with potential biological relevance in AD.

The second LOAD GWAS used the Affymetrix 500K platform to genotype more than 500,000 SNPs in a histo-pathologically confirmed series of 664 LOAD patients and 422 controls and identified the *APOE* locus as the only region with SNPs that reached genome-wide significance after Bonferroni correction for multiple testing [51]. The same group reassessed their data by analyzing *APOE* ε4-positive and -negative subjects separately and dividing their histopathologic series into discovery and replication series [52]. They focused on 312,316 SNPs after quality control exclusions and also included a clinical replication cohort in their genotyping and analysis. They identified 10 SNPs in the GRB-associated binding protein 2 (*GAB2*) gene on chromosome 11q14.1

with nominally significant associations ($P = 4.56 \times 10^{-7}$ to 8.92×10^{-5}) in the *APOE* ε4 carrier subset of the neuropathological discovery cohort composed of 299 LOAD patients and 61 controls. When these 10 SNPs were assessed in the *APOE* ε4-positive subset of the neuropathological replication cohort (113 LOAD patients and 27 controls) as well as the clinical replication cohort (115 LOAD patients and 29 controls), 6 of them showed nominally significant associations. Moreover, when all three *APOE* ε4-positive groups were jointly analyzed, 5 SNPs reached genome-wide significance after Bonferroni corrections, with 1 SNP (rs2373115) achieving $P = 9.66 \times 10^{-11}$ and OR = 4.06. The *GAB2* SNPs did not associate significantly in the *APOE* ε4 non-carriers, thereby leading to reductions in the strength of association when all subjects from all series were analyzed collectively as 861 LOAD patients and 550 controls (rs2373115 SNP association had $P = 5.56 \times 10^{-4}$ and OR = 1.66). *GAB2* encodes a scaffolding protein, GRB-associated binding protein 2, which is involved in cell signaling pathways, especially in the immune system [63]. Its potential role in AD pathophysiology remains to be elucidated; however, preliminary functional studies accompanying this LOAD GWAS revealed differential expression of *GAB2* in AD versus control brains, co-localization of GAB2 with dystrophic neurites, and variation of GAB2 expression influencing tau phosphorylation [52].

The third LOAD GWAS [53] analyzed more than 400,000 SNPs from the Affymetrix 500K platform in 753 AD patients and 736 controls from the discovery series and followed up the 120 top SNPs in the 418 AD patients and 249 controls from the replication series. They identified *APOE*-related SNPs with genome-wide significance in the combined series. Of the top 120 associations, 3 non-*APOE* SNPs passed their quality control thresholds and had nominal logistic regression *P* value of less than 0.05 in their replication series. Two of these SNPs, rs7019241 and rs10868366, reside in the Golgi phosphoprotein 2 gene (*GOLPH2*) on chromosome 9, which is in a suggestive linkage region [35] and encodes a protein functionally involved in Golgi transmembrane trafficking, a subcellular localization where APP cleavage by the γ-secretase complex occurs [64]. The other SNP, rs9886784, resides within a copy number deletion polymorphism on chromosome 9. A fourth SNP, rs10519262, which resides in an intergenic region on chromosome 15, was identified with the nominal *P* cutoff of less than 0.05 using Cox proportional hazards regression analysis for age-at-onset phenotype.

Using a pooled DNA approach on 1,082 LOAD patients and 1,239 controls, the fourth GWAS tested more than 500,000 SNPs from two Illumina platforms [54]. These SNPs were submitted to three different analyses to test for allelic frequency differences in this pooled population, followed by individual genotyping in the same series. One hundred fourteen of the 237 chosen SNPs could be genotyped individually. In addition to the five *APOE*-related SNPs with a high level of significance ($P = 8.97 \times 10^{-5}$ to 5.89×10^{-9}), 74 SNPs had nominal *P* values of not more than 0.05 and were genotyped in an additional control group of 1,400 subjects. Although this is not truly an independent test, the authors used this approach to detect five non-*APOE* SNPs that became even more significant, of which rs727153 had the best *P* value ($P = 3.4 \times 10^{-6}$). This SNP resides in a haplotype block on chromosome 4, including the lecithin retinol acyltransferase gene (*LRAT*), which is implicated in the retinoid pathway [65]. Components of this pathway have been suggested to play a role in AD pathogenesis and proposed as potential drug targets. Genotyping additional SNPs in this study led to improved association signal (rs201825, $P = 6.12 \times 10^{-7}$, OR = 1.3) in the overall group, although none reached genome-wide significance after Bonferroni correction.

The first LOAD GWAS to use family-based analysis tested more than 400,000 SNPs in 941 AD patients versus 404 controls from 410 families of European descent [55]. The follow-up was done in three series composed of 1,767 AD patients versus 838 controls from 875 families, also of European descent. The authors used a family-based association approach that assesses disease status and age of onset jointly. To obtain the corrected levels of significance, they used not the traditional Bonferroni method but a weighted-Bonferroni correction approach [66] that first screens all of the markers and estimates the conditional power of each marker, followed by the family-based association tests that are corrected using the weights determined from the screening analysis. According to this paradigm, they detected four non-*APOE* SNPs with *P* values of 2.0×10^{-3} to 4.0×10^{-6}, which would not have been significant after the traditional Bonferroni correction. These four markers had nominal significance of $P = 0.3$ to 2.0×10^{-5} in the replication series. The two SNPs with nominal significance in both stages, rs11159647 and rs3826656, achieved *P* values of 2.0×10^{-6} and 6.0×10^{-6}, respectively, for AD association in the combined series, which fell short of genome-wide significance after Bonferroni correction. rs11159647 appeared to have an effect on age of onset as well. These SNPs had a nominal *P* value of less than 0.05 in another LOAD GWAS [52,53]. rs11159647 on 14q31.2 and rs3826656 on chromosome 19q13.33 reside in regions without any mapped RefSeq genes, although the latter is in a predicted gene, which overlaps with the CD33 antigen gene. This group also evaluated the top genes from the AlzGene database at the time and determined some variants at $P = 0.03$ and 0.002. Notably, *GAB2* [52] showed one of the strongest nominal association signals.

The sixth LOAD GWAS genotyped more than 500,000 SNPs from the Illumina HumanHap550 platform in 492 LOAD cases and 496 controls in addition to imputing genotypes using a previously published LOAD GWAS [52]. The authors used a modified false discovery rate approach (false discovery rate-beta uniform mixture, or FDR-BUM) [67] to declare significance at the genome-wide level. Besides *APOE*-related SNPs, some of which would be significant after Bonferroni correction, rs11610206 on 12q13, which does not reside in a RefSeq gene, was identified at genome-wide significance using the more relaxed FDR-BUM criteria (nominal $P = 1.93 \times 10^{-6}$). When genotyped in the follow-up series of 238 LOAD and 220 control subjects, this SNP was nominally significant at $P = 0.0496$ and the significance in the joint series improved to $P = 3.452 \times 10^{-7}$, falling short of significance after Bonferroni correction. Additionally, any SNP that had a *P* value of 0.0001 or less in this study or another LOAD GWAS [52], was imputed in the study that did not genotype that SNP. They identified eight SNPs with a nominal *P* value of less than 0.05 in the other study, in which the joint *P* value reached 1.51×10^{-6} for the most significant SNP in *ZNF224* on chromosome 19q13. They also evaluated the AlzGene candidates in these two LOAD GWASs by assessing the genotyped or imputed SNPs in these genes for association with AD. Twenty-five SNPs from nine genes (*ADAM12, CSF1, GBP2, KCNMA1, NOS2A, SORCS2, SORCS3, SORL1,* and *WWC1*) had nominal *P* values of 0.003 to 0.05 in the individual LOAD GWAS with *P* values of 0.0001 to 0.01 in the joint analysis. In addition to providing suggestive evidence for a number of loci in the genome, this study drew attention to the challenges of comparing multiple GWASs via imputation, in which lack of sufficient marker information could lead to false-negative results. In these two studies, this was the case with the *APOE* SNPs that were highly significant in each study but could not be imputed in the other study, leading to missing the *APOE* signal in the joint analyses.

A single-stage LOAD GWAS from Germany analyzed 491 AD patients versus 479 younger controls and focused on the SNPs for the top 10 genes from AlzGene at the time of their study and *SORL1*. Thus, although 555,000 SNPs were genotyped, this study, due to its analytical approach, should be considered a follow-up to other association studies rather than a hypothesis-generating GWAS. Both single SNP and haplotype analyses were performed. In addition to *APOE*-related SNPs, some of which achieved genome-wide significance after Bonferroni correction, SNPs in 8 of the 11 genes tested (*MAPT, SORL1, CHRNB2, CH25H, GAB2, PGBD1, PCK1,* and *LMNA*) revealed nominal significance of *P* values of 0.05 to 6.8×10^{-3} in this study. None of the non-*APOE* SNPs achieved genome-wide significance in their study, but the

authors provided follow-up analysis results that can be used in meta-analyses of all available data on these genes.

The second family-based LOAD GWAS assessed the smallest sample size to date, focusing on two extended LOAD families with 9 affected and 10 unaffected family members versus 60 unrelated controls from the CEPH (Centre d'Etude du Polymorphisme Humain) collection. In stage I, they genotyped more than 400,000 SNPs and identified association with 6 SNPs in the *TRPC4AP* (transient receptor potential cation channel, subfamily C, member 4 associated protein) gene at genome-wide significance ($P = 5.63 \times 10^{-11}$ to 3.85×10^{-10}). Because this initial analysis compared ADs from families with unrelated controls outside of the families, it cannot be considered a family-based analysis. The authors subsequently genotyped 10 SNPs in *TRPC4AP* and identified a common haplotype with increased frequency in the AD patients from the families compared with the control spouses. The same haplotype was also associated with LOAD in an unrelated case control series of 199 LOAD versus 85 control subjects. The authors suggest a functional role for this gene in AD through its interaction with proteins in the inflammatory cascade and its role in calcium homeostasis. Despite the genome-wide significant results in this study, the findings need further replication given the small sample sizes and analytic approach in the first stage, which used unrelated controls versus AD patients from the families. Indeed, this is the only GWAS that failed to identify *APOE*; this could be due to the fact that these families have risk factors that are distinct from the general population, but this could also suggest lack of sufficient power.

The third largest LOAD GWAS to date assessed 844 LOAD subjects versus 1,255 controls in the first stage, which assessed more than 300,000 SNPs on an Illumina platform. The authors followed the top 25 SNPs from this stage for association in the follow-up series of 1,547 LOAD subjects versus 1,209 controls. *APOE*-related SNPs were the only ones that were significant at the genome-wide level in the first stage. Upon combined analysis of all samples, rs5984894 in the protocadherin 11X (*PCDH11X*) gene on chromosome Xq21.31 was the only one that achieved genome-wide significance ($P = 3.8 \times 10^{-8}$, OR = 1.29). Multi-variable logistic regression analysis using male sex as a covariate and comparing male hemizygotes, female heterozygotes, and female homozygotes with the female non-carriers yielded global a *P* value of 3.9×10^{-12} for the global association. Male hemizygotes showed suggestive association ($P = 0.07$), whereas female heterozygotes and homozygotes had nominally significant associations ($P = 0.01$ and 2.0×10^{-7}, respectively), with OR estimates of 1.18, 1.26, and 1.75, respectively, for these groups. *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 expressed predominantly in the brain, suggesting their potential role in brain morphogenesis [68]. Although 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 two most recent LOAD GWASs are by far the largest ones published to date in this field [60,61]. One of the two GWASs in LOAD case control series in which non-*APOE* SNPs reached genome-wide significance after Bonferroni correction in the first stage was performed in 3,941 LOAD and 7,848 control subjects from 13 different centers in Europe and the US, where up to more than 500,000 SNPs were analyzed. In addition to the *APOE*-related SNPs that revealed genome-wide significance ($P = 4.9 \times 10^{-37}$ to 1.8×10^{-157}), rs11136000 in clusterin (*CLU* or *ApoJ*) on chromosome 8 and rs3851179 in the phosphatidylinositol-binding clathrin assembly protein (*PICALM*) gene on chromosome 11 yielded genome-wide significance with $P = 1.4 \times 10^{-9}$ and 1.9×10^{-8}, respectively. When genotyped in the follow-up series of 2,023 LOAD versus 2,340 control subjects from five centers in Europe, these SNPs had nominal significance at a P value of less than 0.05. Significance for rs11136000 in *CLU* improved to $P = 8.5 \times 10^{-10}$ (OR = 0.86) and that for *PICALM* improved to $P = 1.3 \times 10^{-9}$ (OR = 0.90) in the combined series. In addition, the authors identified more SNPs than would be expected by chance with P values of less than 1.0×10^{-5}, including an SNP in the complement receptor 1 gene (rs1408077) that was identified at genome-wide significance in the other largest LOAD GWAS to date [61] ($P = 8.3 \times 10^{-6}$). The authors tested more than 100 SNPs in their series that were identified in other LOAD GWASs and identified nominal significance at a P value of less than 0.05 for a number of these, including *PCDH11X* and *SORL1*.

The other large LOAD GWAS analyzed more than 500,000 SNPs in 2,032 LOAD versus 5,328 control subjects from France [61]. Like the other large LOAD GWASs, this study identified rs11136000 in *CLU* with genome-wide significance in the first stage ($P = 9.0 \times 10^{-8}$) in their analysis, correcting for population stratification in addition to *APOE*-related markers ($P = 5.06 \times 10^{-7}$ to less than 2×10^{-16}). Those SNPs with a P value of less than 1.0×10^{-5} were genotyped or imputed in the follow-up series composed of 3,978 LOAD versus 3,297 controls from 15 centers in four countries. *CLU* SNP rs11136000 had nominal significance in their second stage and enhanced significance of $P = 7.5 \times 10^{-9}$ (OR = 0.86) in their combined series. Another SNP, rs6656401, in complement component receptor 1 (*CR1*) also achieved genome-wide significance in the combined series with $P = 3.7 \times 10^{-9}$ (OR = 1.21). The authors also identified nominal significance for SNPs in *PICALM* ($P = 1.0 \times 10^{-2}$ to 1.0×10^{-3}) and *PCDH11X* ($0.01 < P < 0.05$) in their series, thereby providing additional evidence for these genes identified in other LOAD GWASs [59,60].

CLU encodes clusterin or ApoJ, which along with *APOE* is one of the most abundant apolipoproteins in the human brain. *In vivo* studies suggest that clusterin, like *APOE*, is involved in Aβ clearance from the brain (reviewed in [33]). There are also studies that revealed a role for clusterin in Aβ fibrillogenesis and neurotoxicity. These results raise the possibility that, in Aβ pathophysiology, clusterin may have a dual role similar to that of *APOE*. *PICALM* encodes a protein involved in clathrin-mediated endocytosis, a suggested pathway for trafficking of APP that could also influence Aβ formation [69]. There is also evidence that *PICALM* influences endocytosis of the synaptic vesicle protein VAMP2 from the plasma membrane [70], suggesting a role in synaptic function. CR1 is a receptor for the complement component C3b, which has been suggested to be involved in the peripheral clearance of Aβ [71]. Thus, all three candidate genes that emerged from the two largest LOAD GWASs to date have putative functions in the Aβ cascade and synaptic machinery (for *PICALM*). Additional non-Aβ-mediated pathophysiologic mechanisms for the proteins encoded by these genes may exist and require further investigation.

It is important to note that a number of the 'hits' identified from the LOAD GWASs reside in prior linkage regions (for example, *GALP* on chromosome 19, *LMNA* on chromosome 1, *GOLPH2* on chromosome 9, and *CLU* on chromosome 8) and that others (such as *MAPT* and *SORL1*) have previously been implicated in candidate gene association studies. Multiple lines of evidence in support of the same gene provide additional support for their role in the risk of AD; however, in general, these results need to be interpreted with caution and special attention to any sample overlaps, effect size, and direction of effect in the different studies. The AlzGene database facilitates the visualization and cross-checking of findings from different linkage and association studies [29].

Cumulative results from late-onset Alzheimer disease GWAS

GWASs in LOAD led to the generation of a multitude of novel candidate genes for follow-up. Table 3 depicts the list of genes and variants with the strongest level of association in these studies. The most consistent pattern that emerges from this table is that the effect size and strength of association for *APOE*-related variants are greater than the best non-*APOE* associations. There are multiple explanations for this finding: It is possible that

Table 3. Late-onset Alzheimer disease genome-wide association study: estimated population attributable risks of the top variants.

Reference	Non-ApoE hits (paper)					Non-ApoE hits (AlzGene meta-analysis)			ApoE-related hits				
	Best gene(s)	Best SNP(s)	MAF	OR	PAR	MAF	OR	PAR	Best gene(s)	Best SNP(s)	MAF	OR	PAR
Grupe et al. [50]	GALP	rs3745833	0.35	1.2	6.5%	0.35	1.13	4.4%	TOMM40	rs157581	0.15	2.73	20.6%
Coon et al. [51]	-	-	-	-	-	-	-	-	APOC1	rs4420638	0.18	4.01	35.1%
Reiman et al. [52]	GAB2	rs2373115	0.81	1.66	34.8%	0.82	1.22	15.3%	-	-	-	-	-
Li et al. [53]	GOLPH2	rs7019241	0.13	0.69	28.1%	-	-	-	APOC1	rs4420638	0.18	-	-
Abraham et al. [54]	LRAT	rs201825	0.45	1.3	11.9%	-	-	-	-	-	-	-	-
Bertram et al. [55]	chr14q31.2	rs11159647	0.48	1.4	16.0%	-	-	-	APOC1	rs4420638	0.18	-	-
Beecham et al. [56]	12q13	rs11610206	-	-	-	-	-	-	-	-	-	-	-
Feulner et al. [57]	LMNA	rs915179	0.35	-	-	-	-	-	TOMM40	rs157580 and rs2075650	0.4-0.14	-	-
Poduslo et al. [58]	TRPC4AP	rs6087664	-	-	-	-	-	-	-	-	-	-	-
Carrasquillo et al. [59]	PCDH11X	rs2573905	0.46	1.29	11.8%	-	-	-	TOMM40	rs2075650	0.13	3.29	22.9%
Harold et al. [60]	CLU / PICALM	rs11136000 / rs3851179	0.40 / 0.37	0.86 / 0.86	8.9% / 9.3%	0.6 / 0.36	0.85 (CLU) / 0.88 (PICALM)	9.6% / 8.0%	TOMM40	rs2075650	0.15	2.53	18.7%
Lambert et al. [61]	CLU / CR1	rs11136000 / rs6656401	0.38 / 0.19	0.86 / 1.21	9.2% / 3.8%	0.6 / 0.18	0.85 (CLU) / 1.19 (CR1)	9.6% / 3.3%	-	-	-	-	-

The minor allele frequencies (MAFs), odds ratios (ORs), and population attributable risks (PARs) of the best single-nucleotide polymorphisms (SNPs) in the top gene(s) from the 11 late-onset Alzheimer disease genome-wide association studies are shown, where available from the original publications. MAFs of the rs157581 and rs4420638 SNPs were not shown in the original publications and were obtained from the dbSNP database for subjects of European origin. The formula used to calculate PAR was F(OR − 1)/(F(OR − 1) + 1), where F = frequency of the risk allele and OR = odds ratio of risk allele. Where available, the MAFs, ORs, and PARs emerging from meta-analyses in the AlzGene website are shown for the non-*APOE* results.

all LOAD risk variants are common with modest effect sizes, as suggested by Table 3. However, the findings from the GWASs are, for the most part, not based on the true susceptibility variants but are reflective of their tagging markers, which may harbor greater heterogeneity than the former with respect to alleles and extent of LD. Thus, it remains a possibility that the actual functional susceptibility variants may have bigger effect sizes. Another possibility is the existence of rare variants with bigger effect sizes underlying risk of LOAD, particularly given the insufficient coverage of the rare variants in the human genome with the current GWAS platforms.

Table 3 provides estimates of the PARs of the strongest *APOE* and non-*APOE* variants from the existing GWASs, where necessary information is available. PARs for the *APOE*-related variants vary between 19% and 35%. The PAR estimates for the three genes identified from the two largest LOAD GWASs to date are approximately 9% for *CLU* and *PICALM* and 4% for *CR1*. The effect sizes from the original studies tend to be overestimates. A number of the GWAS hits were assessed in independent studies summarized in AlzGene, in which estimates from meta-analysis may be closer to the true effect sizes of genetic variants in the population. As seen in Table 3, PAR estimates of *GALP*, *GAB2*, *PICALM*, and *CR1* decreased upon analyses of additional series, although that of *CLU* increased slightly. Replication studies with a thousand to tens of thousands of subjects are required to validate these findings and obtain more accurate estimates of effect sizes. The current combined LOAD PARs for *APOE*, *CLU*, *PICALM*, and *CR1* are approximately 56% at most. Given that genetic factors underlie up to 80% of the risk for LOAD [4], the remaining genetic risk for AD could be due to variants in other genes such as *PCDH11X* or *GAB2*, rare or structural variants, presence of different genetic factors in non-European populations, and/or gene-gene and gene-environment interactions.

Conclusions and future directions in the post-GWAS era

Discoveries in the sequence and structure of the human genome, technical advances in high-throughput genotyping, and development of novel analytical approaches led to an explosion in the number and extent of genetic studies in common, complex diseases such as LOAD. Numerous candidate genes that emerged from these studies require further validation not only via genetic replication but also through functional assessment by *in vitro* and *in vivo* approaches. Experience from other common, complex diseases suggests that, despite their success in uncovering common genetic factors for complex diseases, GWASs may fall short of explaining all of the underlying genetics [72] because of their limitations discussed earlier. Thus, alternative approaches are necessary to identify the remaining genetic susceptibility factors. To identify the actual functional variations responsible for the current association and linkage signals as well as to detect rare and structural variations in candidate regions, next-generation sequencing is emerging as a promising approach. Additionally, novel analytical approaches that can mine available GWAS data may uncover structural variations (such as insertions, deletions, translocations, and inversions) that are associated with disease. Indeed, novel approaches that can re-assess GWAS results may provide an efficient method to generate novel information. Combining biologically relevant quantitative phenotype (endophenotype) and GWAS data may be a powerful approach that can potentially lead to the identification of genetic factors that influence disease risk and plausible biological pathways. It will also be important to jointly assess the results of GWASs and other large-scale association and linkage studies to capitalize on the cumulative knowledge that can be gained from such meta-analyses [73]. Recent advances in the genetics of LOAD provide promise in furthering our understanding of its pathophysiology. With the new discoveries in genomics, genetic studies in the post-GWAS era hold potential for uncovering the complex susceptibility factors for this disease which may translate into therapeutic potential for this devastating epidemic.

References
1. Ferri CP, Sousa R, Albanese E, Ribeiro WS, Honyashiki M: *World Alzheimer Report 2009 - Executive Summary*. Edited by Prince M, Jackson J. London: Alzheimer's Disease International; 2009:1-22.
2. Alzheimer A: **A new disease of the cortex (Ger).** *Allg Z Psychiatr* 1907, **64**:146-148.
3. Perrin RJ, Fagan AM, Holtzman DM: **Multimodal techniques for diagnosis and prognosis of Alzheimer's disease.** *Nature* 2009, **461**:916-922.
4. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL: **Role of genes and environments for explaining Alzheimer disease.** *Arch Gen Psychiatry* 2006, **63**:168-174.
5. Lautenschlager NT, Cupples LA, Rao VS, Auerbach SA, Becker R, Burke J, Chui H, Duara R, Foley EJ, Glatt SL, Green RC, Jones R, Karlinsky H, Kukull WA, Kurz A, Larson EB, Martelli K, Sadovnick AD, Volicer L, Waring SC, Growdon JH,

Farrer LA: **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:**641-650.

6. Nee LE, Polinsky RJ, Eldridge R, Weingartner H, Smallberg S, Ebert M: **A family with histologically confirmed Alzheimer's disease.** *Arch Neurol* 1983, **40:**203-208.

7. 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:**1026-1033.

8. 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.

9. 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:**877-880.

10. Tanzi RE, Gusella JF, Watkins PC, Bruns GA, St George-Hyslop P, Van Keuren ML, Patterson D, Pagan S, Kurnit DM, Neve RL: **Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.** *Science* 1987, **235:**880-884.

11. St George-Hyslop PH, Tanzi RE, Polinsky RJ, Haines JL, Nee L, Watkins PC, Myers RH, Feldman RG, Pollen D, Drachman D, Growdon J, Bruni A, Foncin J-F, Salmon D, Frommelt P, Amaducci L,Sorbi S, Piacentini S, Stewart GD, Hobbs WJ, Conneally PM, Gusella JF: **The genetic defect causing familial Alzheimer's disease maps on chromosome 21.** *Science* 1987, **235:**885-890.

12. Goate AM, Owen MJ, James LA, Mullan MJ, Rossor MN, Haynes AR, Farrall M, Lai LY, Roques P, Williamson R, Hardy JA: **Predisposing locus for Alzheimer's disease on chromosome 21.** *Lancet* 1989, **1:**352-355.

13. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, Mant Rooke K, Roques P, Talbot C, Newton P, R, Pericak-Vance M, Roses A, Willioamson R, Rossor M, Owen M, Hardy J: **Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.** *Nature* 1991, **349:**704-706.

14. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, Tsuda T, Mar L, Foncin J-F, Bruni AC, Montesi MP, Sorbi S, Rainero I, Pinessi L, Nee L, Chumakov I, Pollen D, Brookes A, Sanseau P, Polinsky RJ, Wasco W, Da Silva HAR, Haines JL, Pericak-Vance MA, Tanzi RE, Roses AD, Fraser PE, Rommens JM, St George-Hyslop PH: **Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.** *Nature* 1995, **375:**754-760.

15. Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KAB, Weber JL, Bird TD, Schellenberg GD: **A familial Alzheimer's disease locus on chromosome l.** *Science* 1995, **269:**970-973.

16. Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, Chi H, Lin C, Holman K, Tsuda T, Mar L, Sorbi S, Nacmias B, Piacentini S, Amaducci L, Chumakov I, Cohen D, Lannfelt L, Fraser PE, Rommens JM, St George-Hyslop PH: **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:**775-778.

17. **Alzheimer Disease & Frontotemporal Dementia Mutation Database** [http://www.molgen.ua.ac.be/ADMutations].

18. Theuns J, Marjaux E, Vandenbulcke M, Van Laere K, Kumar-Singh S, Bormans G, Brouwers N, Van den Broeck M, Vennekens K, Corsmit E, Cruts M, De Strooper B, Van Broeckhoven C, Vandenberghe R: **Alzheimer dementia caused by a novel mutation located in the APP C-terminal intracytosolic fragment.** *Hum Mutat* 2006, **27:**888-896.

19. Citron M, Oltersdorf T, Haass C, McConlogue L, Hung AY, Seubert P, Vigo-Pelfrey C, Lieberburg I, Selkoe DJ: **Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production.** *Nature* 1992, **360:**672-674.

20. Borchelt DR, Thinakaran G, Eckman CB, Lee MK, Davenport F, Ratovitsky T, Prada CM, Kim G, Seekins S, Yager D, Slunt HH, Wang R, Seeger M, Levey AI, Gandy SE, Copeland NG, Jenkins NA, Price DL, Younkin SG, Sisodia SS: **Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio *in vitro* and *in vivo*.** *Neuron* 1996, **17:**1005-1013.

21. Hsiao K, Chapman P, Nilsen S, Eckman C, Harigaya Y, Younkin S, Yang F, Cole G: **Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice.** *Science* 1996, **274:**99-102.

22. Duff K, Eckman C, Zehr C, Yu X, Prada CM, Perez-tur J, Hutton M, Buee L, Harigaya Y, Yager D, Morgan D, Gordon MN, Holcomb L, Refolo L, Zenk B, Hardy J, Younkin S: **Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin1.** *Nature* 1996, **383:**710-713.

23. Scheuner D, Eckman C, Jensen M, Song X, Citron M, Suzuki N, Bird TD, Hardy J, Hutton M, Kukull W, Larson E, Levy-Lahad E, Viitanen M, Peskind E, Poorkaj P, Schellenberg G, Tanzi R, Wasco W, Lannfelt L, Selkoe D, Younkin S: **Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease.** *Nat Med* 1996, **2:**864-870.

24. Hardy J, Selkoe DJ: **The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.** *Science* 2002, **297:**353-356.

25. Small SA, Duff K: **Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis.** *Neuron* 2008, **60:**534-542.

26. 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:**403-409.

27. Pericak-Vance MA, Bebout JL, Gaskell PC, Yamaoka LH, Hung WY, Alberts MJ, Walker AP, Bartlett RJ, Haynes CA, Welsh KA, Heyman EA, Clark CM, Roses AD: **Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage.** *Am J Hum Genet* 1991, **48:**1034-1050.

28. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA: **Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.** *Science* 1993, **261:**921-923.

29. 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.

30. 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:**163-166.

31. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, Roses AD: **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 U S A* 1993, **90:**1977-1981.

32. Holtzman DM, Bales KR, Tenkova T, Fagan AM, Parsadanian M, Sartorius LJ, Mackey B, Olney J, McKeel D, Wozniak D, Paul SM: **Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease.** *Proc Natl Acad Sci U S A* 2000, **97:**2892-2897.

33. Holtzman DM: *In vivo* **effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology.** *J Mol Neurosci* 2004, **23:**247-254.

34. Kim J, Basak JM, Holtzman DM: **The role of apolipoprotein E in Alzheimer's disease.** *Neuron* 2009, **63:**287-303.

35. Ertekin-Taner N: **Genetics of Alzheimer's disease: a centennial review.** *Neurol Clin* 2007, **25:**611-667.

36. Knopman DS, DeKosky ST, Cummings JL, Chui H, Corey-Bloom J, Relkin N, Small GW, Miller B, Stevens JC: **Practice parameter: diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology.** *Neurology* 2001, **56:**1143-1153.

37. Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, van Duijn CM: **Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study.** *Arch Neurol* 1998, **55:**964-968.

38. Breitner JC, Wyse BW, Anthony JC, Welsh-Bohmer KA, Steffens DC, Norton MC, Tschanz JT, Plassman BL, Meyer MR, Skoog I, Khachaturian A: **APOE-epsilon4 count predicts age when prevalence of AD increases, then declines: the Cache County Study.** *Neurology* 1999, **53:**321-331.

39. Warwick Daw E, Payami H, Nemens EJ, Nochlin D, Bird TD, Schellenberg GD, Wijsman EM: **The number of trait loci in late-onset Alzheimer disease.** *Am J Hum Genet* 2000, **66:**196-204.

40. **A comprehensive genetic linkage map of the human genome. NIH/CEPH Collaborative Mapping Group.** *Science* 1992, **258:**148-162.

41. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, *et al*.: **A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.** *Nature* 2001, **409:**928-933.

42. **AlzGene homepage** [http://www.alzgene.org].

43. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, Katayama T, Baldwin CT, Cheng R, Hasegawa H, Chen F, Shibata N, Lunetta KL, Pardossi-Piquard R, Bohm C, Wakutani Y, Cupples LA, Cuenco KT, Green RC, Pinessi L, Rainero I,

Sorbi S, Bruni A, Duara R, Friedland RP, Inzelberg R, Hampe W, Bujo H, Song YQ, Andersen OM, *et al.*: **The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease.** *Nat Genet* 2007, **39**:168-177.

44. Dreses-Werringloer U, Lambert JC, Vingtdeux V, Zhao H, Vais H, Siebert A, Jain A, Koppel J, Rovelet-Lecrux A, Hannequin D, Pasquier F, Galimberti D, Scarpini E, Mann D, Lendon C, Campion D, Amouyel P, Davies P, Foskett JK, Campagne F, Marambaud P: **A polymorphism in CALHM1 influences Ca2+ homeostasis, Abeta levels, and Alzheimer's disease risk.** *Cell* 2008, **133**:1149-1161.

45. 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* 2003, **33**:177-182.

46. Newton-Cheh C, Hirschhorn JN: **Genetic association studies of complex traits: design and analysis issues.** *Mutat Res* 2005, **573**:54-69.

47. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG: **Replication validity of genetic association studies.** *Nat Genet* 2001, **29**:306-309.

48. International HapMap Consortium: **A haplotype map of the human genome.** *Nature* 2005, **437**:1299-1320.

49. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, *et al.*: **A second generation human haplotype map of over 3.1 million SNPs.** *Nature* 2007, **449**:851-861.

50. Grupe A, Abraham R, Li Y, Rowland C, Hollingworth P, Morgan A, Jehu L, Segurado R, Stone D, Schadt E, Karnoub M, Nowotny P, Tacey K, Catanese J, Sninsky J, Brayne C, Rubinsztein D, Gill M, Lawlor B, Lovestone S, Holmans P, O'Donovan M, Morris JC, Thal L, Goate A, Owen M, Williams J: **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**:865-873.

51. Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Lince DH, Zismann VL, Beach TG, Leung D, Bryden L, Halperin RF, Marlowe L, Kaleem M, Walker DG, Ravid R, Heward CB, Rogers J, Papassotiropoulos A, Reiman EM, Hardy J, Stephan DA: **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**:613-618.

52. Reiman EM, Webster JA, Myers AJ, Hardy J, Dunckley T, Zismann VL, Joshipura KD, Pearson JV, Hu-Lince D, Huentelman MJ, Craig DW, Coon KD, Liang WS, Herbert RH, Beach T, Rohrer KC, Zhao AS, Leung D, Bryden L, Marlowe L, Kaleem M, Mastroeni D, Grover A, Heward CB, Ravid R, Rogers J, Hutton ML, Melquist S, Petersen RC, Alexander GE, *et al.*: **GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers.** *Neuron* 2007, **54**:713-720.

53. Li H, Wetten S, Li L, St Jean PL, Upmanyu R, Surh L, Hosford D, Barnes MR, Briley JD, Borrie M, Coletta N, Delisle R, Dhalla D, Ehm MG, Feldman HH, Fornazzari L, Gauthier S, Goodgame N, Guzman D, Hammond S, Hollingworth P, Hsiung GY, Johnson J, Kelly DD, Keren R, Kertesz A, King KS, Lovestone S, Loy-English I, Matthews PM, *et al.*: **Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease.** *Arch Neurol* 2008, **65**:45-53.

54. Abraham R, Moskvina V, Sims R, Hollingworth P, Morgan A, Georgieva L, Dowzell K, Cichon S, Hillmer AM, O'Donovan MC, Williams J, Owen MJ, Kirov G: **A genome-wide association study for late-onset Alzheimer's disease using DNA pooling.** *BMC Med Genomics* 2008, **1**:44.

55. Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M, Hogan MF, Schjeide BM, Hooli B, Divito J, Ionita I, Jiang H, Laird N, Moscarillo T, Ohlsen KL, Elliott K, Wang X, Hu-Lince D, Ryder M, Murphy A, Wagner SL, Blacker D, Becker KD, Tanzi RE: **Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE.** *Am J Hum Genet* 2008, **83**:623-632.

56. Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, Pericak-Vance MA: **Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease.** *Am J Hum Genet* 2009, **84**:35-43.

57. Feulner TM, Laws SM, Friedrich P, Wagenpfeil S, Wurst SH, Riehle C, Kuhn KA, Krawczak M, Schreiber S, Nikolaus S, Förstl H, Kurz A, Riemenschneider M: **Examination of the current top candidate genes for AD in a genome-wide association study.** *Mol Psychiatry* 2009, Jan 6. [Epub ahead of print].

58. 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.

59. Carrasquillo MM, Zou F, Pankratz VS, Wilcox SL, Ma L, Walker LP, Younkin SG, Younkin CS, Younkin LH, Bisceglio GD, Ertekin-Taner N, Crook JE, Dickson DW, Petersen RC, Graff-Radford NR, Younkin SG: **Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease.** *Nat Genet* 2009, **41**:192-198.

60. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, *et al.*: **Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease.** *Nat Genet* 2009, **41**:1088-1093.

61. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, Letenneur L, Bettens K, Berr C, Pasquier F, Fiévet N, Barberger-Gateau P, Engelborghs S, De Deyn P, Mateo I, Franck A, Helisalmi S, Porcellini E, Hanon O; European Alzheimer's Disease Initiative Investigators, de Pancorbo MM, Lendon C, Dufouil C, Jaillard C, Leveillard T, Alvarez V, *et al.*: **Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease.** *Nat Genet* 2009, **41**:1094-1099.

62. Lang R, Gundlach AL, Kofler B: **The galanin peptide family: receptor pharmacology, pleiotropic biological actions, and implications in health and disease.** *Pharmacol Ther* 2007, **115**:177-207.

63. 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.

64. Baulac S, LaVoie MJ, Kimberly WT, Strahle J, Wolfe MS, Selkoe DJ, Xia W: **Functional gamma-secretase complex assembly in Golgi/trans-Golgi network: interactions among presenilin, nicastrin, Aph1, Pen-2, and gamma-secretase substrates.** *Neurobiol Dis* 2003, **14**:194-204.

65. Goodman AB: **Retinoid receptors, transporters, and metabolizers as therapeutic targets in late onset Alzheimer disease.** *J Cell Physiol* 2006, **209**:598-603.

66. Ionita-Laza I, McQueen MB, Laird NM, Lange C: **Genomewide weighted hypothesis testing in family-based association studies, with an application to a 100K scan.** *Am J Hum Genet* 2007, **81**:607-614.

67. Pounds S, Morris SW: **Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values.** *Bioinformatics* 2003, **19**:1236-1242.

68. 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**:906-914.

69. Wu F, Yao PJ: **Clathrin-mediated endocytosis and Alzheimer's disease: an update.** *Ageing Res Rev* 2009, **8**:147-149.

70. Harel A, Wu F, Mattson MP, Morris CM, Yao PJ: **Evidence for CALM in directing VAMP2 trafficking.** *Traffic* 2008, **9**:417-429.

71. Rogers J, Li R, Mastroeni D, Grover A, Leonard B, Ahern G, Cao P, Kolody H, Vedders L, Kolb WP, Sabbagh M: **Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes.** *Neurobiol Aging* 2006, **27**:1733-1739.

72. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461**:747-753.

73. Khoury MJ, Bertram L, Boffetta P, Butterworth AS, Chanock SJ, Dolan SM, Fortier I, Garcia-Closas M, Gwinn M, Higgins JP, Janssens AC, Ostell J, Owen RP, Pagon RA, Rebbeck TR, Rothman N, Bernstein JL, Burton PR, Campbell H, Chockalingam A, Furberg H, Little J, O'Brien TR, Seminara D, Vineis P, Winn DM, Yu W, Ioannidis JP: **Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases.** *Am J Epidemiol* 2009, **170**:269-279.