Published in final edited form as:
Lancet Neurol. 2009 November ; 8(11): 977–979. doi:10.1016/S1474-4422(09)70277-7.

Beyond Proof of Principle: New Genes for Alzheimer Disease Through Collaboration

Margaret A. Pericak-Vance[1,*] and **Jonathan L. Haines**[2]

[1] John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami FL 33138

[2] Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville TN 37232

Alzheimer disease (AD) is the leading cause of dementia in the elderly with over 5 million individuals affected with AD in the U.S., a number projected to quadruple by 2050 as the population ages [1]. AD has a complex and largely undescribed etiology with strong genetic determinants. Until now only four unequivocal genes carrying risk for AD have been identified. Three of these, the amyloid precursor protein [APP] [2] and the presenilin 1 and 2 [PS1 and PS2] genes [3–5], were identified using the classical Mendelian positional cloning paradigm most prominently applied in the 1990's. This success was facilitated by highly-penetrant autosomal dominant inheritance in early-onset AD families. While these three genes explain the majority of early-onset familial AD and their identification represents a tremendous accomplishment, collectively they account for less than 2% of all AD cases.

The genetic architecture underlying the far more common late-onset AD (LOAD; age at onset ≥ 60 years of age) [6] is much more complex. The sibling recurrence risk (λs)for AD is surprisingly consistent across studies [7–9] with a range of about 4–5. The confluence of biology [10, 11] and genetic mapping [12] facilitated the identification of the association between the apolipoprotein E (*APOE*) gene (the *APOE-4* allele increases risk; the *APOE-2* allele decreases risk) in both familial late-onset and sporadic AD patients [13–15]. *APOE* is the single most significant genetic risk factor identified for LOAD, the fourth of the identified AD genes.

The finding of an association of *APOE* with AD initially ignited the field, but with the exception of variations like the CFH Y402H polymorphism in age related macular degeneration [16–18], few such strong effects in complex diseases have been seen since identified. Since 1993, attempts to identify additional LOAD loci have taken multiple approaches using the best available technologies including genome-wide linkage studies (GWLS) and tests of association for individual candidate genes.

Multiple GWLS for LOAD were published between 1997 and 2006 [19–29]. While some chromosomal regions have been studied extensively (most notably on chromosomes 9, 10, and 12), no consistently replicated LOAD gene has yet been identified using this method. Several reasons account for the limited results including the generally small datasets, the inability of the then available molecular genotyping technologies to capture all the segregation information in the families, [30] and the sensitivity of linkage studies to

[*]Corresponding author: Margaret A Pericak-Vance, Ph.D., Associate Dean for Human Genomic Programs, Dr. John T. Macdonald Foundation Professor of Human Genomics, Director, John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, 1501 NW 10th Avenue, BRB-314 (M860), Miami, Florida 33136, 305 243 2308 (office), 305 243 2396 (fax), mpericak@med.miami.edu.

underlying locus heterogeneity when using datasets consisting of a large number of small families.

Association studies for specific candidate genes selected due to their known (or more often, hypothesized) biological function relevant to AD have been performed for over 350 polymorphisms (www.alzgene.org). For several reasons these studies also have been largely unsuccessful. First, our knowledge of gene function is still very limited and it has been difficult to make direct observations of altered gene function or expression in AD tissues. Second, the sample sizes and single-stage study designs have generally been too small for the moderate effect sizes and the substantial locus heterogeneity that we now know underlie LOAD. Third, the level of genomic detail in single nucleotide polymorphisms (SNP) and in copy number variations (CNV) content that could be interrogated was low. These issues conspired to make replication of any true effect difficult and generation of false positive results rampant. Thus, while some of these reported associations are likely to be important, it is not surprising that the overall evidence for any one of these loci is mixed with results from the majority of studies refuting any association.

The (nearly) complete characterization of the consensus human sequence has greatly increased our ability to identify and describe the genomic structure of genes and variation in those genes between different individuals and species. Of even more import for disease gene studies is the vast pool of characterized common differences among people provided by HapMap data [31–33] that allows a genome wide association study (GWAS) design to be implemented by genotyping 100,000–1,000,000 SNPs with high fidelity and low cost per genotype. GWAS have already been successful in over 150 different phenotypes with over 400 different new polymorphisms associated with disease (http://www.genome.gov/26525384). Clearly such an approach is successful in finding at least some of the underlying genetic variation responsible for disease risk. It is also becoming clear, however, that the effect sizes of almost all of these variations are quite small (odds ratios ranging from 1.1–1.4, with most in the 1.15–1.3 range) and rarely explain more than a tiny fraction of the overall genetic effect in any common disease [34].

There are now 10 published GWAS in AD and most use unrelated cases and controls while a few have used family datasets. Following the pattern of most diseases, the initial five GWAS studies used available sample sets and had somewhat limited power [35–39]. A sixth report [40] used previously reported data [35] stratified by *APOE* genotype, a seventh report used a “gene based” screen [41], and an eighth report used a DNA pooling scheme [42] rather than genotyping individual samples. All these studies confirmed the strong effect of *APOE*, and while several SNPs achieved genome-wide significance, it is clear that the remaining genetic risk loci in AD have population-level effects much smaller than *APOE*. Two general conclusions can be made from the existing GWAS studies. First, there is no other genetic effect as pervasive and strong as *APOE* and second, the total genetic effect explained by the additional identified SNPs is still very small (1–2%) and a large proportion of the genetic effect remains unexplained.

The two most recent publications by Harold et al (2009) and Lambert et al. (2009) [43, 44] represent the first of the next generation, large sample size GWAS studies designed in part to overcome the power problem. These studies used the collaborative model that brings together datasets from multiple research groups. Harold et al. included over 16,000 total individuals with over 5,900 cases and 10,000 controls in their two stage analysis. With such a commanding data set, they identified SNPs in two genes with genome wide significance, the *CLU gene* (clusterin, which is also referred to as apolipoprotein J (*APOJ*)) and the *PICALM* gene (phosphatidylinositol binding clathrin assembly protein). Both of these SNPs were identified in the initial dataset (Stage 1) and were replicated in a second independent

dataset (stage 2) with p-values of 8.5×10^{-10}, odds ratio=0.86 and 1.3×10^{-9}, odds ratio =0.86, respectively. In the same issue, Lambert et al. (2009) performed a similarly powered independent study of 6,000 cases and over 8,600 controls. They also employed a two stage design and also found genome wide significance in SNPs in two genes. Significantly, the most significant SNP identified in the Lambert et al (2009) study was in *CLU*. Their second gene meeting genome wide significance was *CR1* (complement component (3b/4b) receptor 1). As with Harold et al. (2009) the odds ratio for *CLU* was 0.86 with a p-value of 7.5×10^{-9} and for *CR1* the p-value was 3.7×10^{-9} with an odds ratio = 1.21. *CLU*, the consensus candidate identified between the two studies is an excellent functional candidate. Like its predecessor, *APOE, CLU* is expressed in cerebrospinal fluid, found in amyloid plaques and can bind beta amyloid (Aβ). The two genes share such functionality that together with the strong statistical support a compelling story emerges in support of *CLU* as a new AD risk locus, albeit with an effect size much smaller than *APOE*. The remaining two genes of interest, *PICALM* and *CR1* also receive cross support from the two studies but do not emerge as candidates as strong as *CLU*. In addition both studies unequivocally conclude that additional AD genes remain to be found.

Collectively, these data represent a significant advance in the search for the genetic underpinnings of AD and confirm that GWAS is a powerful and exciting tool for geneticists as they continue to describe the genetic architecture of AD. However, a word of caution is still needed. None of these genes were identified in the earlier GWAS studies as important players, and a reanalysis of these earlier data to determine their level of support for these new genes is needed. Additional large datasets must be examined to test the reproducibility and generalizabiily of these effects. Even in studies generating results with p-values of this magnitude, future studies may not consistently replicate these effects, and the effect sizes may be even smaller than the initial reports. As demonstrated by Harold et al., and Lambert et al., developing the necessary samples sizes almost always requires collaboration among multiple investigators. New consortia such as the Alzheimer Disease Genetics Consortium (ADGC) funded by the National Institute on Aging are essential and will contribute significantly in determining the true role for these genes and identifying the remaining genetic effects in AD.

References

1. 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. [PubMed: 9736873]
2. 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. [PubMed: 1671712]
3. 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. [PubMed: 7638622]
4. 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 Aug 31; 376(654):775–8. [PubMed: 7651536]
5. 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 Jun 29; 375(653):754–60. [PubMed: 7596406]
6. Pericak-Vance MA, Haines JL. Genetic susceptibility to Alzheimer disease. Trends Genet. 1995 Dec; 11(12):504–8. [PubMed: 8533168]
7. Breitner JC, Silverman JM, Mohs RC, Davis KL. Familial aggregation in Alzheimer's disease: comparison of risk among relatives of early-and late-onset cases, and among male and female relatives in successive generations. Neurology. 1988 Feb; 38(2):207–12. [PubMed: 3340281]

8. Sadovnick AD, Irwin ME, Baird PA, Beattie BL. Genetic studies on an Alzheimer clinic population. Genet Epidemiol. 1989; 6(5):633–43. [PubMed: 2591733]
9. Hirst C, Yee IM, Sadovnick AD. Familial risks for Alzheimer disease from a population-based series. Genet Epidemiol. 1994; 11(4):365–74. [PubMed: 7813898]
10. Namba Y, Tamonaga M, Kawasaki H, Otomo E, Ikeda K. Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and cru plaque amyloid in Creutzfeldt-Jakob disease. Brain Res. 1991; 541:163–6. [PubMed: 2029618]
11. Wisniewski T, Frangione B. Apolipoprotein E: A pathological chaperone protein in patients with cerebral and systemic amyloid. Neurosci Let. 1992; 135:235–8. [PubMed: 1625800]
12. Pericak-Vance MA, Bebout JL, Gaskell PC Jr, Yamaoka LH, Hung WY, Alberts MJ, et al. Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. Am J Hum Genet. 1991 Jun.48:1034–50. [PubMed: 2035524]
13. 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 Aug 13; 261(5123):921–3. [PubMed: 8346443]
14. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology. 1993 Aug.43:1467–72. [PubMed: 8350998]
15. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. 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 Mar 01.90:1977–81. [PubMed: 8446617]
16. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, et al. Complement factor H variant increases the risk of age-related macular degeneration. Science. 2005 Apr 15; 308(572): 419–21. [PubMed: 15761120]
17. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005 Apr 15; 308(5720):385–9. [PubMed: 15761122]
18. Edwards AO, Ritter R 3rd, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. Science. 2005 Apr 15; 308(5720):421–4. [PubMed: 15761121]
19. Pericak-Vance MA, Bass MP, Yamaoka LH, Gaskell PC, Scott WK, Terwedow HA, et al. Complete genomic screen in late-onset familial Alzheimer disease. Evidence for a new locus on chromosome 12. JAMA. 1997 Oct 15; 278(1):1237–41. [PubMed: 9333264]
20. Kehoe P, Vrieze Wavrant-De, Crook R, Wu WS, Holmans P, Fenton I, et al. A full genome scan for late onset Alzheimer's disease. Hum Mol Genet. 1999; 8:237–45. [PubMed: 9931331]
21. Pericak-Vance MA, Grubber J, Bailey LR, Hedges D, West S, Kemmerer B, et al. Identification of novel genes in late-onset Alzheimer disease. Exp Gerontol. 2000; 35:1343–52. [PubMed: 11113612]
22. Hiltunen M, Mannermaa A, Thompson D, Easton D, Pirskanen M, Helisalmi S, et al. Genome-wide linkage disequilibrium mapping of late-onset Alzheimer's disease in Finland. Neurology. 2001 Nov 13.57:1663–8. [PubMed: 11706108]
23. Mayeux R, Lee JH, Romas SN, Mayo D, Santana V, Williamson J, et al. Chromosome-12 mapping of late-onset Alzheimer disease among Caribbean Hispanics. Am J Hum Genet. 2002 Jan.70:237–43. [PubMed: 11715112]
24. Myers A, Wavrant De-Vrieze F, Holmans P, Hamshere M, Crook R, Compton D, et al. Full genome screen for Alzheimer disease: stage II analysis. Am J Med Genet. 2002 Mar 08.114:235–44. [PubMed: 11857588]
25. Blacker D, Bertram L, Saunders AJ, Moscarillo TJ, Albert MS, Wiener H, et al. Results of a high-resolution genome screen of 437 Alzheimer's Disease families. Hum Mol Genet. 2003 Jan 01.12:23–32. [PubMed: 12490529]
26. Farrer LA, Bowirrat A, Friedland RP, Waraska K, Korczyn AD, Baldwin CT. Identification of multiple loci for Alzheimer disease in a consanguineous Israeli-Arab community. Hum Mol Genet. 2003 Feb 15.12:415–22. [PubMed: 12566388]

27. Mayeux R. Dissecting the relative influences of genes and the environment in Alzheimer's disease. Ann Neurol. 2004 Feb.55:156–8. [PubMed: 14755716]
28. Ashley-Koch AE, Shao Y, Rimmler JB, Gaskell PC, Welsh-Bohmer KA, Jackson CE, et al. An autosomal genomic screen for dementia in an extended Amish family. Neurosci Lett. 2005 May 13.379:199–204. [PubMed: 15843063]
29. Hahs DW, McCauley J, Crunk A, McFarland L, Gaskell P, Jiang L, et al. A genome-wide linkage analysis of dementia in the Amish. American Journal of Medical Genetics (Neuropsychiatric Genetics). 2006 Jan 02.
30. Sawcer SJ, Maranian M, Singlehurst S, Yeo T, Compston A, Daly MJ, et al. Enhancing linkage analysis of complex disorders: an evaluation of high-density genotyping. Hum Mol Genet. 2004 Sep 01; 13(1):1943–9. [PubMed: 15238506]
31. International HapMap Consortium. The International HapMap Project. Nature. 2003 Dec 18; 426(6968):789–96. [PubMed: 14685227]
32. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005 Oct 27; 437(7063):1299–320. [PubMed: 16255080]
33. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007 Oct 18; 449(7164):851–61. [PubMed: 17943122]
34. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. J Clin Invest. 2008; 118(5):1590–605. [PubMed: 18451988]
35. 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 Apr; 68(4):613–8. [PubMed: 17474819]
36. 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 Nov; 83(5):623–32. [PubMed: 18976728]
37. 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 Jan; 65(1):45–53. [PubMed: 17998437]
38. 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 Jan; 84(1):35–43. [PubMed: 19118814]
39. 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 Feb; 41(2):192–8. [PubMed: 19136949]
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 Jun 7; 54(5):713–20. [PubMed: 17553421]
41. 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 Apr 15; 16(8):865–73. [PubMed: 17317784]
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 Sep 29.1:44. [PubMed: 18823527]
43. 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 Sep 6.
44. 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 Sep 6.