In This Issue: Entering the Era of High-Density Genome-Wide Association Studies

In a previous issue of *The Journal of Clinical Psychiatry,* my colleagues and I reviewed progress in the genetic, transcriptomic, and proteomic study of Alzheimer's disease (AD).[1] In this issue, we report initial findings from the first genome-wide survey of more than 500,000 single nucleotide polymorphisms (SNPs) in late-onset AD cases and controls. This study provides empirical support for the unparalleled contribution of the apolipoprotein E (*APOE*) gene to the risk of late-onset AD, it demonstrates the promise of increasingly high-density genome-wide association studies in the discovery of previously elusive susceptibility genes for AD and other common phenotypes, and it underscores some of the methodological challenges that remain to be addressed in this important endeavor.

Besides older age, the *APOE* ε4 allele is by far the best established risk factor for late-onset AD. Located on chromosome 19, the *APOE* gene has 3 common variants, ε2, ε3, and ε4, which give rise to 6 possible genotypes: ε2/ε2, ε2/ε3, ε3/ε3 (the most common genotype), ε2/ε4, ε3/ε4, and ε4/ε4. Each additional copy of the ε4 allele in a person's *APOE* genotype is associated with a higher risk of late-onset AD and a slightly younger median age at dementia onset.[2,3] (Conversely, the ε2 allele is associated with a decreased risk of AD and an older median age at dementia onset. [3])

Twin studies suggest that there are several other susceptibility genes that, along with *APOE* variants, could account for as much as 80% of the risk for late-onset AD.[4] Discovery of the remaining AD susceptibility genes could help improve our understanding of the molecular basis of AD, inform the discovery of new treatments, distinguish AD subtypes in clinically meaningful ways (including the response to treatment), and identify the persons most likely to benefit from future prevention therapies. Unfortunately, the other late-onset AD susceptibility genes have been elusive. As recently noted, there have been almost a thousand reports implicating or refuting hundreds of putative susceptibility genes for late-onset AD.[5] The problem is not unique to AD: geneticists have struggled to find susceptibility genes for most common polygenic phenotypes, including all of the common psychiatric disorders.

One limitation has been the paucity of genetic markers available to read the entire genome, for these markers have not been close together enough to make inferences about the SNPs that reside between them. It has been proposed that at least 300,000 SNPs across the genome would be needed to perform genome-wide association studies in unrelated cases and controls.[6] New technologies now permit the simultaneous assessment of more than 300,000 SNPs; they will very soon permit the assessment of more than a million SNPs, and the numbers continue to grow. Another limitation may be inaccuracies in the classification of cases and controls. At autopsy, about 10% of clinically characterized AD patients may not meet neuropathologic criteria for AD and about a third of elderly persons without clinically significant cognitive impairment may have neuropathologic features of the disorder at autopsy. In this report, we surveyed about 500,000 SNPs across the genome in more than one thousand clinically characterized and neuropathologically confirmed AD cases and controls. Since the *APOE* variants are not represented on the genotyping platform used in this survey (but

were assessed independently), we sought to determine if there were any surveyed SNPs sufficiently close to the *APOE* variants to detect an association with AD.

In this study, AD was highly associated with a SNP right next-door to (i.e., 14 kilobase pairs away from) the *APOE* ε4 allele. The association was extraordinarily significant (p value = 1×10^{-39} for this SNP and 1×10^{-44} for the neighboring *APOE* ε4 allele itself), and it was much more significant than that with every other surveyed SNP in the human genome. Still, the association with *APOE* would have been missed were it not for a single SNP in the genotyping platform, indicating the need for significantly higher-density genotyping platforms to avoid missing relevant variants between any 2 consecutive markers.

This study has several important implications. First, it provides empirical support for the notion that the *APOE* ε4 allele is unparalleled in its contribution to AD risk and for the need to develop more sophisticated approaches to identify and confirm the remaining AD susceptibility genes. Second, it demonstrates the value of studying clinically characterized and neuropathologically confirmed AD cases and controls, it clarifies the odds ratios for each *APOE* genotype in a Caucasian population, and it shows an even higher association with *APOE* ε4 gene dose (i.e., the number of ε4 alleles in a person's *APOE* genotype) in this population than those studies that did not require neuropathologic criteria for the classification of each subject group. Third, it welcomes us to the era of high-density genome-association studies in unrelated AD cases and controls while, as previously noted, indicating the need for even higher-density genotyping platforms, which continue to be developed at a rapid pace. Finally, this study demonstrates the need to address several other challenges in the genetic study of AD.

To identify the remaining susceptibility genes for late-onset AD, and those for other common and genetically complex phenotypes, researchers will need to perform ultra-high-density genome-wide association studies in an even larger number of cases and controls and to confirm new findings in independent samples. (It was difficult to find a very large group of expired elderly controls confirmed to be clinically and neuropathologically unaffected by AD for this purpose. New fibrillar amyloid imaging techniques could be used in living persons to identify a larger number of AD patients and elderly control subjects with and without AD neuropathology for genetic studies.) As we have previously suggested,[7,8] more sophisticated ways are also needed to analyze the wealth of data generated in genome-wide association studies of AD. For instance, there is a need to better account for differences in genetic background, account for the contribution of different SNPs within the same gene, and investigate the combined contribution—not just the individual contribution—of different SNPs.

As researchers enter the era of ultra-high-density whole-genome association studies, the use of compound genetic analyses in sufficiently large samples of well-characterized cases and controls could provide the best way to solve the genetic mysteries of AD, common psychiatric disorders, and a range of other important medical problems.

REFERENCES

1. Papassotiropoulos A, Fountoulakis M, Dunckley T, et al. Genetics, transcriptomics, and proteomics of Alzheimer's disease. J Clin Psychiatry 2006;67:652–670
2. Corder EH, Saunders AM, Risch NJ, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat Genet 1994;7:180–184
3. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA 1997;278:1349–1356
4. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 2006;63:168–174
5. Bertram L, McQueen MB, Mullin K, et al. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 2007;39:17–23
6. Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 1999;22:139–144
7. Papassotiropoulos A, Wollmer MA, Tsolaki M, et al. A cluster of cholesterol-related genes confers susceptibility for Alzheimer's disease. J Clin Psychiatry 2005;66:940–947
8. Reiman EM. In this issue: genetics, cholesterol, and the risk of Alzheimer's disease. J Clin Psychiatry 2005;66:938–939

Eric M. Reiman, M.D.
Deputy Editor