Letter to the Editor

Am. J. Hum. Genet. 78:1081, 2006

Considerations for Genomewide Association Studies in Parkinson Disease

To the Editor:

Although the magnitude of a genetic component of Parkinson disease (PD [MIM 168600]) remains to be determined, the disease has already shown remarkable genetic heterogeneity, with at least five monogenic forms identified, the most common of which is *LRRK2* (MIM 609007).[1] In this issue of *The American Journal of Human Genetics,* four investigative teams[2–5] report that they have sought to replicate the findings from a genomewide association (GWA) study of PD affection by Maraganore et al.[6] Taken together, these four studies appear to provide substantial evidence that none of the SNPs originally featured as potential PD loci are convincingly replicated and that all may be false positives. Furthermore, that the *LRRK2* gene was not identified may be considered a false-negative result. This conclusion is both disappointing and discouraging. The original study invested heavily in this venture, with 443 sibling pairs ($n = 886$) discordant for PD typed in tier 1 for 198,345 SNPs (172,420,019 genotype calls) and a tier 2 follow-up typing the strongest 1,892 SNPs in 332 matched case-control unrelated pairs (1,176,772 genotypes). Because this report is among the first GWA studies and because the effort appears to have failed to produce the desired objective, it is worth examining the implications for GWA studies in general and, specifically, the significance of this study for PD.

First, let's examine the original report. Tier 1 of the original study is founded upon sibling pairs discordant for PD recruited from the Mayo Clinic in Rochester, MN. The sample is composed of individuals substantially of northern and central European descent. Discordant sibling pairs were selected to limit false-positive results due to population stratification bias.[7] Population differences between case and control samples are recognized as the primary source of false-positive associations, and, clearly, every effort to minimize these effects is to be encouraged. However, in PD there is substantial evidence for reduced penetrance,[8] and the disease etiology is most likely a complex interaction of genetic and environmental factors.[9] Thus, the selection of randomly ascertained PD cases (often termed "sporadic") may include a substantial proportion of cases with little or no genetic basis for disease, and, even among familial cases, many unaffected siblings may carry PD risk alleles but remain unaffected for lack of critical environmental exposure, for essential modifying genes, or for follow-up to an advanced age. Case identification in tier 1 should focus on the selection of those most likely to carry the inherited form of the disease, whereas controls should be likely non–gene carriers drawn from the same population. Concerns for population stratification might best be addressed in tier 2 by the genotyping of families of tier 1 cases and by family-based association studies. SNPs showing association in these first phases can be typed in a second unrelated case-control sample as a tier 3, with case enrichment for familial disease when possible.

Fundamentally, scientific discovery relies first and foremost upon the independent replication of results. Investigators seeking to replicate the findings of association studies need to consider whether their sample provides an appropriate forum for the investigation. Because the overwhelming majority of SNPs in GWA studies will not be functionally related to the disease, one cannot reasonably expect that linkage-disequilibrium patterns will generalize across diverse ethnic groups. Thus, one may expect that there may not be replication for samples recruited from a restricted geographic region (e.g., Taiwan[2]). Whereas most of these replication samples are composed of Europeans (e.g., from Finland,[2] Norway and Ireland,[3] and the United Kingdom[4]), a few reveal minor-allele frequencies that vary from the original sample and that may deserve further study. Enrichment for familial PD would also be important, since none of these replication studies is described as familial PD.

Genomewide linkage studies have generally not been successful in finding genes responsible for common complex diseases, and whether GWA studies will prove to be more successful remains to be determined. There is at least one important positive precedent of the Maraganore et al.[6] study. Notably, all of their single-SNP association results (minor-allele frequencies and *P* values) are available in two online text files (available from http://www.journals.uchicago.edu/AJHG/journal/issues/

v77n5/42619/tableS2new.txt and http://www.journals.uchicago.edu/AJHG/journal/issues/v77n5/42619/tableS3new.txt) in the online-only version of the original article.[6] These results can be readily downloaded and searched for evidence of association with other interesting PD candidate genes. Maraganore and colleagues, with the Michael J. Fox Foundation, have the opportunity to establish a precedent for making the entire GWA study available online, since one may reasonably expect that true PD risk alleles may be found among the SNPs with lesser levels of statistical significance. The jury is still out on whether this GWA study holds important insights for PD.

DR. RICHARD H. MYERS
Boston University School of Medicine
Boston

Web Resources

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/entrez/Omim/ (for PD and *LRRK2*)

References

1. Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, Kachergus J, Hulihan H, Uitti RJ, Calne DB, et al (2004) Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron 44:601–607
2. Clarimon J, Scholz S, Fung H-C, Hardy J, Eerola J, Hellström O, Chen C-M, Wu Y-R, Tienari PJ, Singleton A (2006) Conflicting results regarding the semaphorin gene (*SEMA5A*) and the risk for Parkinson disease. Am J Hum Genet 78:1082–1084 (in this issue)
3. Farrer MJ, Haugarvoll K, Ross OA, Stone JT, Milkovic NM, Cobb SA, Whittle AJ, Lincoln SJ, Hulihan MM, Heckman MG, White LR, Aasly JO, Gibson JM, Gosal D, Lynch T, Wszolek ZK, Uitti RJ, Toft M (2006) Genomewide association, Parkinson disease, and *PARK10*. Am J Hum Genet 78:1084–1088 (in this issue)
4. Goris A, Williams-Gray CH, Foltynie T, Compston DAS, Barker RA, Sawcer SJ (2006) No evidence for association with Parkinson disease for 13 single-nucleotide polymorphisms identified by whole-genome association screening. Am J Hum Genet 78:1088–1090 (in this issue)
5. Li Y, Rowland C, Schrodi S, Laird W, Tacey K, Ross D, Leong D, Catanese J, Sninsky J, Grupe A (2006) A case-control association study of the 12 single-nucleotide polymorphisms implicated in Parkinson disease by a recent genome scan. Am J Hum Genet 78:1090–1092 (in this issue)
6. Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA, Pant PVK, Frazer KA, Cox DR, Ballinger DG (2005) High-resolution whole-genome association study of Parkinson disease. Am J Hum Genet 77:685–693
7. Maraganore DM (2005) Blood is thicker than water: the strengths of family-based case-control studies. Neurology 64:408–409
8. Maher NE, Currie LJ, Lazzarini AM, Wilk JB, Taylor CA, Saint-Hilaire MH, Feldman RG, Golbe LI, Wooten GF, Myers RH (2002) A segregation analysis of Parkinson disease revealing evidence for a major causative gene. Am J Med Genet 109:191–197
9. Dekker MC, Bonifati V, van Duijn CM (2003) Parkinson's disease: piecing together a genetic jigsaw. Brain 126:1722–1733

Address for correspondence and reprints: Dr. Richard H. Myers, Department of Neurology, E-304, Boston University School of Medicine, 715 Albany Street, Boston, MA 02118. E-mail: rmyers@bu.edu

0002-9297/2006/7806-0019$15.00