

Am. J. Hum. Genet. 78:1084, 2006

Genomewide Association, Parkinson Disease, and *PARK10*

To the Editor:

Genomewide linkage analysis of rare familial forms of parkinsonism has identified mutations in seven genes, revealing a clinicopathologically and genetically heterogeneous syndrome.¹ Less progress has been made in the more typical late-onset form of Parkinson disease (PD [MIM 168600]), although the recently identified *LRRK2* (MIM 609007) G2019S substitution is estimated to account for ~1% of sporadic PD cases.² Common polymorphisms of familial genes may also influence susceptibility to idiopathic PD.^{3,4} Of the 198,345 SNPs successfully genotyped in the recent genomewide association (GWA) study, 26 had notably different allele frequencies between patients and controls in both tiers ($P < .01$).⁵ Fifteen of these SNPs had opposite directions of effect (disease risk or protection) in tiers 1 and 2. The remaining 11 SNPs were proposed as markers for new genes/chromosomal loci that influence susceptibility to PD. In addition, two SNPs in tier 2 (*rs682705* and *rs7520966*) were highlighted in the *PARK10* locus (MIM 606852), which nominated the gene *LOC200008* in disease susceptibility.

The *PARK10* locus on chromosome 1p32 was originally identified in a genomewide linkage analysis of 117 patients from 51 Icelandic families (maximum $Z_{lr} = 4.8$ at *D1S231*, with a LOD-1, 7.6-cM support interval from *D1S2874* to *D1S475*).⁶ Iceland has a well-characterized genealogy that is powerful for family-based linkage studies. The ancestral founders of Iceland have Scandinavian patrilineal inheritance with a minor Celtic matrilineal component.⁷ Assuming that the *PARK10* mutation predates the Icelandic settlement, we reasoned that the 1p32 susceptibility gene might be more readily found in patients with PD originating from Scandinavian or Celtic populations. In parallel to the study of Maraganore et al.,⁵ we have been mapping the *PARK10* locus. Genotypes from 28 SNPs (including *rs682705* and *rs7520966*) within a 132-kb region of chromosome 1p32 located around the *LOC200008* gene have been analyzed in two well-characterized case-control series

from Norway and Ireland. In addition, we attempted to replicate findings for the two *PARK10* SNPs in a U.S. series collected at the Mayo Clinic in Jacksonville, FL. We then employed all three case-control series to investigate the genotype/allele frequencies of the main 11 SNPs nominated to influence PD susceptibility.⁵ Power was comparable to the original study (>80% at $\alpha = 0.05$ for odds ratios [ORs] >2.0 and for disease-allele frequencies >0.035), and genotyping call rates were >95% for all markers (table 1).

In total, Norwegian samples included 676 subjects (cases and controls) with a mean age (\pm SD) of 70 ± 11 years, Irish samples included 372 subjects with a mean age (\pm SD) of 61 ± 13 years, and the U.S. samples included 522 subjects with a mean age (\pm SD) of 71 ± 10 years. All patients were examined and were observed longitudinally by a movement-disorders neurologist (J.O.A., J.M.G., D.G., T.L., Z.K.W., and R.J.U.), and they were given a diagnosis of PD in accordance with published criteria.⁸ Each patient was individually matched, on the basis of age (± 4 years) and ethnicity, to an unrelated control without evidence of neurological disease. The ethical review boards at each institution involved approved the study, and all participants provided informed consent.

SNP genotyping was performed using TaqMan chemistry on an ABI7900 genetic analyzer; in cases where genotype data was available for only one subject of a matched pair, the other subject was retained in the analysis. For the controls in each population, χ^2 tests of Hardy-Weinberg equilibrium (HWE) were implemented using Haploview.⁹ Optimal SNP coverage for association analysis of the *LOC200008* gene was determined empirically by the construction of linkage-disequilibrium (LD) maps in Norwegian and Irish samples, onto which haplotype blocks were assigned (fig. 1).^{10,11} ORs for disease association, with corresponding 95% CIs, were subsequently calculated using logistic-regression models adjusted for age and sex. Overall ORs combining data from

Table 1

Genotype/Allele Frequencies of the Main 11 SNPs Nominated to Influence PD Susceptibility

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

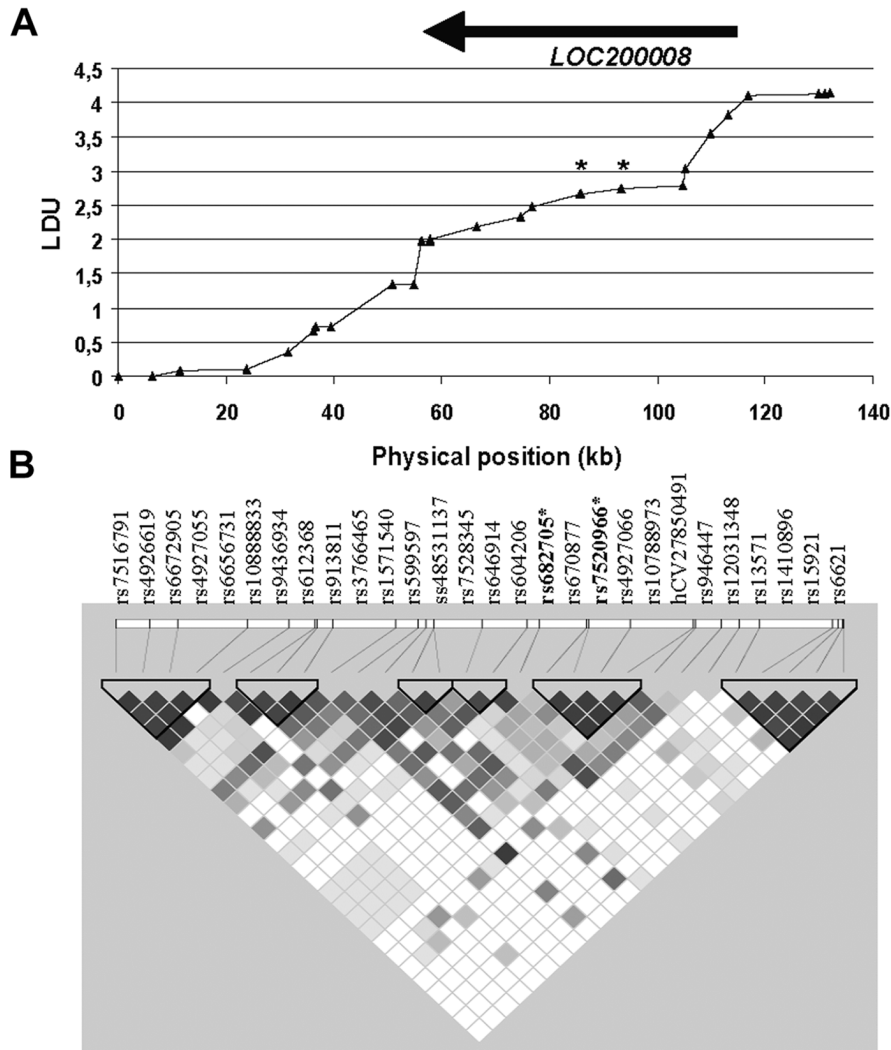


Figure 1 Metric LD map and haplotype block structure of the investigated region. *A*, LD map providing information about LD patterns in the investigated candidate region, through locations expressed in LDUs. LDUs have an inverse relationship with LD, with regions of extensive recombination having many LDUs. The physical position of the gene in the region *LOC200008* is marked with an arrow. All 28 SNPs genotyped are reported, although the symbols (\blacktriangle) may be obscured for SNPs that lie in close physical proximity and high LD. SNPs *rs682705* and *rs7520966* are denoted by an asterisk (*). *B*, LD structure of the candidate region. *Black* and *dark gray cells*, strong LD; *gray cells*, intermediate; and *light gray* and *white cells*, evidence for historical recombination. The haplotype block structure of the region is defined according to Gabriel et al.¹⁰ An asterisk denotes SNPs *rs682705* and *rs7520966*. The LD map and haplotype structure were constructed using genotypes from the Norwegian sample. Similar results were obtained for the Irish population.

all three sites were additionally adjusted for site. Previous studies have nominated the *PARK10* locus as an age-at-onset modifier in PD¹²; thus, we also assessed the influence of 1p32 SNPs variability on this disease trait, using linear-regression models adjusted for sex.

There was no evidence of association with PD for any of the 28 genotyped 1p32 SNPs in our study (all SNP $P > .05$ after applying Bonferroni correction in both population samples). Haplotype frequencies between patients and controls were not significantly different for

the haplotype blocks identified; nor was the age at onset in patients associated with any single marker or haplotype (all corrected $P > .05$). Of note, the ancestral recombination and haplotype blocks apparent within Norwegian and Irish samples were comparable for this interval at this marker resolution. The average number of SNPs per LD unit (LDU) was 6.8 (mean LDU between markers 0.15, range 0–0.63), indicating that the number of SNPs genotyped within and flanking *LOC200008* should be sufficient for examination of the region.¹¹ In

addition, the two *PARK10* SNPs showed no significant association within the U.S. series ($P > .05$). None of the other 11 SNPs nominated by the GWA study had different allele frequencies or genotype distributions between affected subjects and matched controls (all SNP $P > .05$ in all populations independently or as a combined sample set) (table 2). There was no evidence of departure from HWE in controls ($P > .01$ in all population controls).

Our study indicates that genetic variability within the *LOC200008* gene is unlikely to explain the *PARK10* susceptibility locus for PD. Sadly, the lack of disease association and replication in an independent U.S. series of comparable power suggests that the original findings may be spurious. Failure to nominate *LOC200008* as the *PARK10* gene in our population samples provides empirical support for statistical caveats concerning GWA studies. Implicit in multiple testing is false discovery, even in well-designed studies, and there are several potential sources of bias.¹³ Of note, neither *PARK10* SNP *rs682705* nor *rs7520966* fulfilled the main criterion for being genotyped in tier 2 ($P < .01$ in tier 1 overall analysis), but each was included with a less stringent association criterion ($P < .05$ in tier 1 overall analysis) because of its physical position within a *PARK* locus. Interestingly, the combined P value for *rs682705* ($P = 9.07 \times 10^{-6}$) is the second-lowest P value of the overall study, even though it did not fulfill the inclusion criteria. Individual-level data from the GWA study is not yet available, but, in our study, these two SNPs also appear to be in LD (pairwise $r^2 > 0.9$), as suggested by Maraganore et al.⁵; in addition, the minor-allele frequencies (MAFs) of the two SNPs are comparable across studies and populations. The former suggests less-than-optimal haplotype tagging in the initial study, whereas the latter argues against technical errors in genotyping, but neither provides sufficient explanation for the positive findings observed elsewhere.⁵

We found no evidence of direct association between the 11 SNPs nominated in the GWA study and disease in the three independent populations or in a combined sample group ($n = 1,570$) (table 2). However, for these loci, we did not employ a gene-based approach (nor did we fine-map each region as with *PARK10*), as advocated elsewhere¹⁴; we await the results of further replication studies. Of note, in the study by Maraganore et al.,⁵ the *rs7702187* SNP within *SEMA5A* (MIM 609297) had the lowest combined P value ($P = 7.62 \times 10^{-6}$); however, a total of 53 SNPs were examined in this gene in tier 1. Only *rs7702187* was significant before correction ($P = .001$), which supports the possibly spurious nature of this and the other associations. The MAFs observed in our three populations and in that of the GWA study are comparable, which argues against population bias/heterogeneity (table 2).

The number of SNPs highlighted in each tier of the original study is consistent with chance—that is, 1% of SNPs use a significance level of $P < .01$. None of the P values obtained by Maraganore et al.⁵ meets a Bonferroni correction for multiple testing, although this standard may be too conservative in GWA, since it fails to account for LD and incorrectly assumes that chromosomal markers are independent. A consensus on the most appropriate correction for multiple testing has yet to be reached. Now that genomewide data sets have been generated, there exists the possibility to use these to develop appropriate statistical methods to identify true positive results.¹⁵

In the interim, we recommend that enthusiasm for positive findings should be tempered by the strength of the evidence, the population-attributable risk, and the differences in SNP allele/genotype frequencies between cases and controls. If allele frequencies are significantly different, genomic controls might be used to assess population substructure. It is important that future studies

Table 2

11 SNPs Nominated in GWA Study as Genetic Susceptibility Loci for PD

dbSNP ACCESSION NUMBER	CHROMOSOME	POSITION (bp)	THIS STUDY					MARAGANORE ET AL. ⁵					
			Control MAF			Estimated OR (95% CI) ^a	Combined P ($n = 1,570$)	Control MAF		P		Estimated OR (95% CI) ^b	Combined P ($n = 1,550$)
			Ireland	Norway	United States			Tier 1	Tier 2	Tier 1	Tier 2		
<i>rs7702187</i>	5p15.2	9385281	.17	.18	.18	.88 (.74–1.06)	.18	.18	.20	.001	.002	1.74 (1.36–2.24)	7.62×10^{-6}
<i>rs10200894</i>	2q36	228642637	.13	.11	.09	.96 (.77–1.21)	.74	.12	.13	.009	.001	1.84 (1.38–2.45)	1.70×10^{-5}
<i>rs2313982</i>	4q31.1	139145665	.05	.11	.09	.93 (.73–1.18)	.54	.07	.06	.006	.002	2.01 (1.44–2.79)	1.79×10^{-5}
<i>rs17329669</i>	7p14	36625169	.13	.12	.14	1.01 (.82–1.24)	.92	.13	.11	.008	.001	1.71 (1.33–2.21)	2.30×10^{-5}
<i>rs7723605</i>	5p15.3	5407615	.13	.14	.13	.91 (.75–1.12)	.38	.12	.09	.010	.002	1.78 (1.35–2.35)	3.30×10^{-5}
<i>ss46548856</i>	10q21	58986929	.09	.08	.11	.93 (.73–1.19)	.58	.09	.11	.003	.002	1.88 (1.38–2.57)	3.65×10^{-5}
<i>rs16851009</i>	2q24	166456214	.11	.11	.12	.95 (.76–1.18)	.63	.09	.08	.002	.009	1.84 (1.36–2.49)	4.17×10^{-5}
<i>rs2245218</i>	1p36.2	13885132	.19	.17	.15	.95 (.79–1.14)	.57	.11	.13	.002	.002	1.67 (1.29–2.14)	4.61×10^{-5}
<i>rs7878232</i>	Xq28	150516943	.25	.23	.25	1.10 (.97–1.25)	.15	.29	.26	.003	.010	1.38 (1.17–1.62)	6.87×10^{-5}
<i>rs1509269</i>	4q31.1	139111329	.08	.13	.13	.94 (.76–1.17)	.58	.10	.09	.005	.008	1.71 (1.30–2.26)	9.21×10^{-5}
<i>rs11737074</i>	4q27	125438978	.21	.20	.21	1.05 (.89–1.25)	.55	.19	.19	.007	.005	1.50 (1.21–1.86)	1.55×10^{-4}

NOTE.—In this study, MAFs are not significantly different between the populations. No P values are corrected for multiple testing. SNPs are ordered by combined P value, per Maraganore et al.⁵

^a The direction of effect of the estimated OR observed in this study for each SNP is shown (i.e., >1 risk and <1 protective).

^b Estimated ORs in the study by Maraganore et al.⁵ do not indicate the direction of effect relative to the MAF.

employ multiple independent sample series, each with sufficient power to verify significant genetic associations, before publication.¹⁶ However, lack of evidence for an association is not the same as evidence against one; thus, lack of replication should also be interpreted with caution.

Over the few next years, the number of GWA studies will increase, and it is important to learn from the experiences gained by the few studies performed to date. Although our negative findings suggest that the conclusions drawn from the study by Maraganore et al.⁵ might be based on spurious associations, further analysis of individual-level raw data is now necessary. The recent identification of a complement factor H polymorphism in age-related macular degeneration in a GWA study and the identical findings by two other groups using other study designs demonstrates that this approach can be used successfully.^{17–19} It may be that, because of the heterogeneous nature of PD, associations with a gestalt phenotype are masked by background variation in SNP informativeness, population strata, and insufficient power. It is, therefore, crucial that future associations are validated and that analysis is performed to resolve the underlying cause of association in the sample population. GWA studies may still provide direction for the genetic analysis of heterogeneous complex traits, but, in the short term, they may exacerbate the problem of replication failure in association studies.

Acknowledgments

The authors thank the patients and families for their participation in this study. Mayo Clinic Jacksonville is an M. K. Udall Parkinson's Disease Research Center of Excellence (National Institute of Neurological Disorders and Stroke grant P01 NS40256), and the authors thank all collaborators at the Udall Center. This study was also supported by the National Institutes of Health grant R01 NS36960, the Research Council of Norway grant 153487/V50, and Reberg's Legacy. The Ireland research consortium was supported by a Programme for Research in Third-Level Institutions neurosciences award and by the Research and Development Office of the Health and Personal Social Services. O.A.R. and M.T. are partly funded by National Parkinson Foundation and Parkinson's Disease Foundation fellowships, respectively. We thank Minnie Schreiber for technical assistance.

MATTHEW J. FARRER,¹ KRISTOFFER HAUGARVOLL,^{1,2}
OWEN A. ROSS,¹ JEREMY T. STONE,¹
NICOLE M. MILKOVIC,¹ STEPHANIE A. COBB,¹
ANDREW J. WHITTLE,¹ SARAH J. LINCOLN,¹
MARY M. HULIHAN,¹ MICHAEL G. HECKMAN,¹
LINDA R. WHITE,² JAN O. AASLY,³
J. MARK GIBSON,⁴ DAVID GOSAL,^{5,6}
TIMOTHY LYNCH,^{5,6} ZBIGNIEW K. WSZOLEK,¹
RYAN J. UITTI,¹ AND MATHIAS TOFT^{1,2}

¹Departments of Neuroscience and Neurology, Mayo Clinic College of Medicine, Jacksonville, FL;

²Department of Neuroscience, Norwegian University of Science and Technology, and ³Department of Neurology, St. Olav's Hospital, Trondheim, Norway;

⁴Department of Neurology, Royal Victoria Hospital, Belfast, Ireland; ⁵Department of Neurology, Mater Misericordiae Hospital, and ⁶The Conway Institute, University College, Dublin, Ireland

Web Resources

The URLs for data presented herein are as follows:

dbSNP, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=snp>

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for PD, *LRRK2*, *PARK10*, and *SEMA5A*)

References

1. Ross OA, Farrer MJ (2005) Pathophysiology, pleiotropy and paradigm shifts: genetic lessons from Parkinson's disease. *Biochem Soc Trans* 33:586–590
2. Kachergus J, Mata IF, Hulihan M, Taylor JP, Lincoln S, Aasly J, Gibson JM, Ross OA, Lynch T, Wiley J, Payami H, Nutt J, Maraganore DM, Czyzewski K, Styczynska M, Wszolek ZK, Farrer MJ, Toft M (2005) Identification of a novel *LRRK2* mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. *Am J Hum Genet* 76:672–680
3. Skipper L, Wilkes K, Toft M, Baker M, Lincoln S, Hulihan M, Ross OA, Hutton M, Aasly J, Farrer M (2004) Linkage disequilibrium and association of *MAPT* H1 in Parkinson disease. *Am J Hum Genet* 75:669–677
4. Pals P, Lincoln S, Manning J, Heckman M, Skipper L, Hulihan M, Van den Broeck M, De Pooter T, Cras P, Crook J, Van Broeckhoven C, Farrer MJ (2004) α -Synuclein promoter confers susceptibility to Parkinson's disease. *Ann Neurol* 56:591–595
5. 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
6. Hicks AA, Petursson H, Jonsson T, Stefansson H, Johannsdottir HS, Sainz J, Frigge ML, Kong A, Gulcher JR, Stefansson K, Sveinbjornsdottir S (2002) A susceptibility gene for late-onset idiopathic Parkinson's disease. *Ann Neurol* 52:549–555
7. Helgason A, Sigurethardottir S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, Stefansson K (2000) Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* 67:697–717

8. Gelb DJ, Oliver E, Gilman S (1999) Diagnostic criteria for Parkinson disease. *Arch Neurol* 56:33–39
9. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265
10. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–2229
11. Tapper WJ, Maniatis N, Morton NE, Collins A (2003) A metric linkage disequilibrium map of a human chromosome. *Ann Hum Genet* 67:487–494
12. Li YJ, Scott WK, Hedges DJ, Zhang F, Gaskell PC, Nance MA, Watts RL, et al (2002) Age at onset in two common neurodegenerative diseases is genetically controlled. *Am J Hum Genet* 70: 985–993
13. Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6:109–118
14. Neale BM, Sham PC (2004) The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 75:353–362
15. Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6:95–108
16. Hattersley AT, McCarthy MI (2005) What makes a good genetic association study? *Lancet* 366:1315–1323
17. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh, J (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–389
18. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Noureddine M, Gilbert JR, Schetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA (2005) Complement factor H variant increases the risk of age-related macular degeneration. *Science* 308:419–421
19. Edwards AO 3rd, Ritter R, Abel KJ, Manning A, Panhuysen C, Farrer LA (2005) Complement factor H polymorphism and age-related macular degeneration. *Science* 308:421–424

Address for correspondence and reprints: Dr. Owen A. Ross, Department of Neuroscience, Mayo Clinic College of Medicine, Birdsall Building, 4500 San Pablo Road, Jacksonville, FL 32224. E-mail: ross.owen@mayo.edu
 © 2006 by The American Society of Human Genetics. All rights reserved.
 0002-9297/2006/7806-0021\$15.00