

The Association Between Genetic Variants in *SORL1* and Alzheimer Disease in an Urban, Multiethnic, Community-Based Cohort

Joseph H. Lee, PhD; Rong Cheng, PhD; Nicole Schupf, PhD; Jennifer Manly, PhD; Rafael Lantigua, MD; Yaakov Stern, PhD; Ekaterina Rogaeva, PhD; Yosuke Wakutani, PhD; Lindsay Farrer, PhD; Peter St. George-Hyslop, MD; Richard Mayeux, MD, MSc

Objective: To investigate the association between Alzheimer disease (AD) and variant alleles in *SORL1* using a series of single nucleotide polymorphisms (SNPs) in an urban, multiethnic, community-based population.

Design: We used a nested case-control analysis in a population-based, prospective study of aging and dementia in Medicare recipients, 65 years and older.

Setting: Northern Manhattan, NY.

Participants: There were 296 patients with probable AD and 428 healthy, elderly controls. The participants were African American (34%), Caribbean Hispanic (51%), or non-Hispanic white (15%).

Main Outcome Measures: We genotyped all 29 SNPs in *SORL1* that were examined in the earlier report. We assessed allelic association with AD using standard case-control methods, which included apolipoprotein E genotype as a covariate.

Results: Several individual SNPs and SNP haplotypes were significantly associated with AD in this prospectively col-

lected community-based cohort, confirming the previously reported positive association of *SORL1* with AD. Single nucleotide polymorphism 12, near the 5' region, was associated with AD in African American and Hispanic individuals. Two SNPs in the 3' region were also associated with AD in African American (SNP 26) and non-Hispanic white (SNP 20) individuals. A single haplotype in the 3' region was associated with AD in Hispanic individuals. However, several different haplotypes were associated with AD in African American and white individuals, including the TTC haplotypes at SNPs 23 through 25 ($P = .035$), which was significantly associated with AD in the North European white individuals in our previous report.

Conclusions: This study confirms the association between genetic variants in *SORL1* and AD. While the associations observed in these data sets overlap with those previously reported, the finding of novel SNP and haplotype associations suggests that there may be extensive allelic heterogeneity in *SORL1*. Broad regions of the *SORL1* gene will therefore need to be scrutinized for functional pathogenic variants.

Arch Neurol. 2007;64:501-506

IN A PREVIOUS STUDY, WE FOUND that genetic variants within the *SORL1* gene were associated with an increased risk of Alzheimer disease (AD).¹ To ensure the validity of that study, we investigated the association using a variety of specialized data sets, including multiplex family samples and clinic-based case-control cohorts from genetically and culturally distinct populations. This strategy allowed the discovery and replication of individual single nucleotide polymorphism (SNP) and haplotypic associations in 2 distinct regions of *SORL1* in multiple independent cohorts. To explore the observation further, we have reinvestigated the association between *SORL1* and AD in an independent collection of patients with

probable AD and healthy, elderly controls from a longitudinal prospective study of aging and dementia in multiethnic communities in northern Manhattan, NY. The investigation of a randomly sampled, prospectively studied, community-based cohort has advantages and disadvantages in

*For editorial comment
see page 479*

replicating genetic association studies. The diagnoses for both affected and normal status are highly secure and are relatively unaffected by referral bias inherent in clinic-based series (eg, enrichment of patients from centers with known interests in genetics). However, sampling from commu-

Author Affiliations are listed at the end of this article.

Table 1. Characteristics of Genotyped Cases vs Controls in 3 Ethnic Groups Residing in Northern Manhattan, NY*

	Caribbean Hispanic†		African American†		White Non-Hispanic European†	
	Case	Control	Case	Control	Case	Control
Sample size (N = 724)	178	194	88	158	30	76
Age at onset, y, mean (SD)‡	80.9 (6.8)	79.1 (5.1)	83.3 (7.3)	81.1 (6.1)	84.4 (8.0)	82.7 (7.2)
Education, y, mean (SD)	4.8 (4.0)	7.6 (4.5)	8.5 (3.7)	11.7 (3.7)	10.0 (3.7)	13.0 (3.0)
Women, %	73.0	73.7	78.4	73.4	80.0	75.0
APOE ε4, %	21.1	11.3	18.8	17.1	13.3	11.8

Abbreviation: APOE, apolipoprotein E.

*Demographic characteristics of the cases and controls from the prospective, longitudinal, multiethnic, community-based study of aging and dementia in northern Manhattan, who met diagnostic criteria for either probable Alzheimer disease or for normal aging on at least 2 occasions separated by 18 months.

†Ethnicity based on self-report.

‡For prevalent cases, we used the age at which the first complaint was reported, if available; otherwise, age at first evaluation was used to conservatively approximate age at onset.

nity-based series ablates the ability to collate patients and controls according to their true genetic backgrounds. Nevertheless, we reasoned that, on balance, the investigation of samples from this study would serve as a credible source of independent replication and might provide valuable initial insights from which to formulate future questions about population and relative risks from *SORL1* variants and the degree of allelic heterogeneity in the *SORL1* AD locus.

METHODS

SUBJECTS AND SETTING

The patients with AD and elderly controls without dementia assessed herein were participants in a prospective study of aging and dementia in Medicare recipients, 65 years and older, residing in northern Manhattan (Washington Heights, Hamilton Heights, and Inwood). This epidemiological study originally consisted of a stratified random sample of 50% of all persons older than 65 years that was obtained from the Health Care Finance Administration. Individuals were sent a letter from the Health Care Finance Administration explaining that they had been selected to participate in a study of aging by investigators at Columbia University. The sampling procedures have been described in detail elsewhere.^{2,3} Each participant underwent an in-person interview of general health and functional ability at the time of entry into the study followed by a standardized assessment, including medical history, physical and neurological examination, and a neuropsychological battery especially developed for this community.⁴⁻⁶ Ethnic group was classified by participant's self-report using the format of the 1990 US Census.⁷ Participants were asked if they considered themselves to be white, black, or other and then asked if they were Hispanic. Participants were recruited at 2 points (1992-1994 and 1999-2002). They have been followed up at approximately 18-month intervals with similar assessments at each interval. The institutional review boards of Columbia University Medical Center and the New York Psychiatric Institute approved recruitment, informed consent, and study procedures.

In the present study, to maximize diagnostic accuracy, we included all patients with probable AD who had a Clinical Dementia Rating Scale score⁸ of 1 or higher and who had been followed up, using the criteria described later, on at least 2 occasions (**Table 1**). Similarly, the healthy, elderly controls included subjects who were also followed up on at least 2 occasions sepa-

rated by approximately 18 months and who had had no evidence of AD or mild cognitive impairment on either assessment (Table 1). Because our prior study revealed evidence for allelic heterogeneity, with different SNPs and haplotypes showing association with AD in data sets with different ancestries, we collated cases and controls in the present study into 3 nested subsets: Caribbean Hispanic (178 cases, 194 controls); African American (88 cases, 158 controls); and white non-Hispanic European (30 case, 76 controls). While these cohorts are small, statistical power estimates, assuming the parameters from our initial study (eg, SNP 8, allele frequency of 0.39 in cases; odds ratio of 1.5), reveal that the current study had 98% power to detect significant allelic association between genetic variants in *SORL1* and AD at an α of .05 for Caribbean Hispanic individuals; 90% power for African American individuals; and 53% power for white individuals, based on the model by Gordon and colleagues.⁹ For rarer SNPs (eg, SNP 23, allele frequency of 0.125; odds ratio of 2), the current study had 84% power for Caribbean Hispanic individuals, 66% for African American individuals, and 32% for white individuals.

CLINICAL ASSESSMENT AND NEUROLOGICAL DIAGNOSIS

All participants received structured neurological and functional assessments by physicians and underwent a standardized neuropsychological battery that included measures of memory, orientation, language, abstract reasoning, and visuospatial ability.^{5,6} The diagnosis of dementia was established at a consensus conference that included neurologists, neuropsychologists, and psychiatrists and based on all available information gathered from the initial and follow-up assessments and medical records. The diagnosis was based on the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for probable AD.^{10,11} The diagnosis of dementia required evidence of cognitive decline, including memory impairment on the neuropsychological test battery as well as evidence of impairment in social or occupational function (Clinical Dementia Rating >1.0).⁸

GENOTYPING

Genotyping was performed using the GenomeLab SNPstream System and primer sets as previously described.¹ One hundred DNA samples were genotyped twice for every SNP marker (concordance rate >99%). Apolipoprotein E (APOE) was genotyped as

(.029 ≤ *P* ≤ .016). Single nucleotide polymorphism 12, which is located 12.2 kilobases (kb) from the SNP 8 through 10 cluster associated with AD in multiple data sets in our initial article,¹ was significantly associated with AD in African American and Caribbean Hispanic individuals. The T allele at SNP 12 was significantly associated with AD in Caribbean Hispanic individuals (*P* = .029). In non-Hispanic white individuals, the T allele was associated with AD but was not significant (*P* = .20). The T allele was also associated with AD in both the Mayo Clinic autopsy cohort (*P* = .003) and in the overall Mayo Clinic case-control cohort (*P* = .046) in our initial report of an association between *SORL1* and AD (Table 4¹). In African American individuals, however, the C allele at SNP 12 was associated with AD (*P* = .016). Single nucleotide polymorphism 20 was significantly associated with AD in the non-Hispanic white individuals (*P* = .025, G allele). Single nucleotide polymorphism 20 is closely flanked by both SNP 19 (−93 base pairs [bp]), which was associated with AD in several data sets in our report, and by SNP 21 (+5966 bp), which was associated with AD in the North European case-control data set.¹ Single nucleotide polymorphism 26 was associated with AD in the African American cohort (*P* = .019, G allele) and is located +6015 bp from the SNP 23 through 25 cluster that had previously shown a haplotypic association with AD in the MIRAGE African American sibships and in 3 different white data sets.¹

SORL1 HAPLOTYPE ASSOCIATION

Haplotype analyses using a sliding window size of 3 contiguous SNPs demonstrated several haplotypic associations in all 3 ethnic group data sets (Table 3). Although these haplotypes were distributed across the *SORL1* gene and varied in frequency (0.01–0.037), several of them clustered. The common TTC haplotype at SNPs 23, 24, and 25 was associated with AD in non-Hispanic white individuals (*P* = .035). This same haplotype had previously been shown to be associated with AD in the North European familial AD (FAD) data set and the Israeli Arab, North European, and Mayo Clinic–Jacksonville case-control data sets in our initial report.¹ In this same region of the *SORL1* gene, the frequent CCA haplotype at SNPs 24, 25, and 26 was robustly associated with AD in African American individuals (haplotype *P* = .0006, empirical *P* = .0005, and global *P* = .01), while the common CTG haplotype at these same SNPs had a borderline protective effect (*P* = .06). Intriguingly, these same haplotypes were also previously shown to be associated with AD in the MIRAGE African American cohort in the previous study.¹ However, in the MIRAGE African American cohort, the CCA haplotype was protective while the CTG haplotype was deleterious.¹ We interpret these opposite allelic associations to mean that 1 or more risk alleles nearby are in linkage disequilibrium with these haplotypes in this region in African American individuals.

In the central region of *SORL1*, several frequent, overlapping haplotypes at SNPs 16 through 22 were also associated with AD. Thus, the haplotypes ATA at SNPs 16 through 18, TAG at SNPs 17 through 19, AGG at SNPs

18 through 20, GGC at SNPs 19 through 21, and GCC at SNPs 20 through 22 were significantly associated with AD (<.001 ≤ haplotype *P* ≤ .032). The ATA haplotype at SNPs at 16 through 18 was also associated with risk for AD in the North European FAD data set in the initial report (haplotype frequency, 0.218; *z* = 2.794; haplotype *P* = .005; global *P* = .057; Supplementary Table 5¹).

At the 5′ end of the gene, several low-frequency haplotypes at SNPs 1 through 6 and SNPs 8 through 13 were also associated with AD in all 3 cohorts. The CGC haplotype at SNPs 8 through 10, which was previously associated with AD in the Caribbean Hispanic FAD, Israeli Arab, and North European case-control data sets in the initial report,¹ was not replicated in the sporadic Caribbean Hispanic case-control samples herein. Nevertheless, at SNPs 8 through 10, both the CAT haplotype in white non-Hispanic individuals (haplotype *P* = .014) and the CGT haplotype in African American individuals (haplotype *P* = .016) were associated with AD.

COMMENT

Our results independently confirm the previous conclusion that multiple genetic variants in *SORL1* are associated with AD. We have directly replicated the previously reported association between AD and both the TTC haplotype at SNPs 23, 24, and 25 and the ATA haplotype at SNPs 16 through 18 among white individuals in the present and the previous report.¹ We have also shown that SNPs 26 through 28 display haplotypic association with AD in African American individuals in both the present data set and in the previously studied MIRAGE African American cohort,¹ although the same haplotypes had opposite effects on risk for AD in African American individuals from these 2 data sets. A similar situation exists at SNPs 8 through 10 where different haplotypes were associated with AD in the present study and the previous study (CAT in white non-Hispanic individuals and CGT in African American individuals herein and CGC in data sets from our previous study¹). The failure to detect an association at SNPs 8 through 10 in the sporadic Caribbean Hispanic AD cases from this population-based cohort in Washington Heights does not contradict the findings from the earlier study of the Caribbean Hispanic FAD pedigrees. Indeed, weak allelic associations at SNP 12 (+11.7 kb from SNPs 8–10) and weak haplotypic associations at SNPs 4 through 6 (−7037 bp from SNPs 8–10) were observed in the 5′ region of the *SORL1* gene in this data set. There are several possible explanations for the disparity in location of the association. First, there may be further allelic heterogeneity within the Caribbean Hispanic population. Second, it is conceivable that genetic factors such as *SORL1* play a relatively smaller role in community-based sporadic AD in Caribbean Hispanic individuals and are observable only with larger sample sizes or when investigated in sibships multiply affected with AD.

Taken together, the current results suggest 2 conclusions. First, the association between AD and variants in the *SORL1* gene is likely to be correct. Second, the discovery of significant association in multiple regions of the gene

and the discovery of different AD-associated haplotypes in different data sets support the notion that there may be a high degree of allelic heterogeneity, with disease-associated variants occurring on multiple different haplotypic backgrounds. This situation differs markedly from the circumstances observed with *APOE* $\epsilon 4$.¹⁶ Two practical considerations arise from these conclusions. Further replication studies will be needed to assess cohorts with as few founders as possible. Second attempts to identify the pathogenic variants in *SORL1* will likely have to investigate larger regions of the *SORL1* gene than simply just between SNPs 8 through 10 and 22 through 25.

Accepted for Publication: January 24, 2007.

Author Affiliations: Taub Institute on Alzheimer's Disease and the Aging Brain (Drs Lee, Cheng, Schupf, Manly, Lantigua, Stern, and Mayeux), Gertrude H. Sergievsky Center (Drs Lee, Cheng, Schupf, Manly, Stern, and Mayeux); Departments of Neurology (Drs Stern and Mayeux), Psychiatry (Drs Stern and Mayeux), and Medicine (Dr Lantigua), College of Physicians and Surgeons, and Department of Epidemiology, School of Public Health (Drs Lee and Mayeux), Columbia University, New York, NY; Centre for Research in Neurodegenerative Diseases, Department of Medicine, University of Toronto, and Toronto Western Hospital Research Institute, Toronto, Ontario (Drs Rogaeva, Wakutani, and St. George-Hyslop); and Departments of Medicine (Genetics Program), Neurology, Genetics and Genomics, Epidemiology, and Biostatistics, Boston University Schools of Medicine and Public Health, Boston, Mass (Dr Farrer). **Correspondence:** Richard Mayeux, MD, MSc, Gertrude H. Sergievsky Center, Columbia University College of Physicians and Surgeons, 630 W 168th St, New York, NY 10032 (rpm2@columbia.edu).

Author Contributions: *Study concept and design:* Lee, Farrer, St. George-Hyslop, and Mayeux. *Acquisition of data:* Schupf, Manly, Lantigua, Stern, Rogaeva, Wakutani, St. George-Hyslop, and Mayeux. *Analysis and interpretation of data:* Lee, Cheng, Schupf, Manly, Stern, Rogaeva, Wakutani, Farrer, St. George-Hyslop, and Mayeux. *Drafting of the manuscript:* Lee, Schupf, St. George-Hyslop, and Mayeux. *Critical revision of the manuscript for important intellectual content:* Lee, Cheng, Manly, Lantigua, Stern, Rogaeva, Wakutani, Farrer, St. George-Hyslop, and Mayeux. *Statistical analysis:* Lee, Cheng, Schupf, and Farrer. *Obtained funding:* Manly, Rogaeva, Wakutani, St. George-Hyslop, and Mayeux. *Administrative, technical, and material support:* Manly and Lantigua. *Study supervision:* Lee, Schupf, Stern, Rogaeva, Wakutani, St. George-Hyslop, and Mayeux.

Financial Disclosure: None reported.

Funding/Support: This work was supported by grants P01-AG07232 and R37-AG15473 (Dr Mayeux) and R01-AG09029 (Dr Farrer) from the National Institutes of Health; the National Institute on Aging; the Alzheimer Association; the Blanchett Hooker Rockefeller Foundation; Charles S. Robertson Gift from the Banbury Fund (Dr Mayeux); the Canadian Institutes of Health Research; the Howard Hughes Medical Institute; the Canadian Institutes of Health Research–Japan Science and Technology Trust; the Alzheimer Society of Ontario; the Canada Foundation for Innovation; the Ontario Research and Development Challenge Fund; the Ontario Mental Health Foundation; Genome Canada; and the Alzheimer Society of Canada (Dr St. George-Hyslop).

REFERENCES

1. Rogaeva E, Meng Y, Lee JH, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet.* 2007;39:168-177.
2. Gurland BJ, Wilder DE, Lantigua R, et al. Rates of dementia in three ethnorracial groups. *Int J Geriatr Psychiatry.* 1999;14:481-493.
3. Tang MX, Cross P, Andrews H, et al. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology.* 2001;56:49-56.
4. Pittman J, Andrews H, Tatemichi T, et al. Diagnosis of dementia in a heterogeneous population: a comparison of paradigm-based diagnosis and physician's diagnosis. *Arch Neurol.* 1992;49:461-467.
5. Stern Y, Andrews H, Pittman J, et al. Diagnosis of dementia in a heterogeneous population: development of a neuropsychological paradigm-based diagnosis of dementia and quantified correction for the effects of education. *Arch Neurol.* 1992;49:453-460.
6. Stricks L, Pittman J, Jacobs DM, Sano M, Stern Y. Normative data for a brief neuropsychological battery administered to English- and Spanish-speaking community-dwelling elders. *J Int Neuropsychol Soc.* 1998;4:311-318.
7. *Census of Population and Housing Summary Tape File 1, Technical Documentation.* Washington, DC: Bureau of the Census; 1991.
8. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry.* 1982;140:566-572.
9. Gordon D, Heath SC, Liu X, Ott J. A transmission/disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. *Am J Hum Genet.* 2001;69:371-380.
10. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Third Edition.* Washington, DC: American Psychiatric Association; 1980.
11. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology.* 1984;34:939-944.
12. Hixson JE, Vernier DT. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with HhaI. *J Lipid Res.* 1990;31:545-548.
13. Maestre G, Ottman R, Stern Y, et al. Apolipoprotein E and Alzheimer's disease: ethnic variation in genotypic risks. *Ann Neurol.* 1995;37:254-259.
14. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21:263-265.
15. Ott J. *Analysis of Human Genetic Linkage.* Baltimore, Md: Johns Hopkins University Press; 1999.
16. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology.* 1993;43:1467-1472.