

# Genetic risk variants in African Americans with multiple sclerosis

Noriko Isobe, MD, PhD  
Pierre-Antoine Gourraud,  
PhD, MPH  
Hanne F. Harbo, MD,  
PhD  
Stacy J. Caillier, BS  
Adam Santaniello, BS  
Pouya Khankhanian, BA  
Martin Maiers, MS  
Stephen Spellman, MBS  
Nezih Cereb, MD  
SooYoung Yang, PhD  
Marcelo J. Pando, PhD  
Laura Piccio, MD, PhD  
Anne H. Cross, MD  
Philip L. De Jager, MD,  
PhD  
Bruce A.C. Cree, MD,  
PhD, MCR  
Stephen L. Hauser, MD  
Jorge R. Oksenberg, PhD

Correspondence to  
Dr. Oksenberg;  
Jorge.Oksenberg@ucsf.edu

## ABSTRACT

**Objectives:** To assess the association of established multiple sclerosis (MS) risk variants in 3,254 African Americans (1,162 cases and 2,092 controls).

**Methods:** Human leukocyte antigen (HLA)-DRB1, HLA-DQB1, and HLA-A alleles were typed by molecular techniques. Single nucleotide polymorphism (SNP) genotyping was conducted for 76 MS-associated SNPs and 52 ancestry informative marker SNPs selected throughout the genome. Self-declared ancestry was refined by principal component analysis of the ancestry informative marker SNPs. An ancestry-adjusted multivariate model was applied to assess genetic associations.

**Results:** The following major histocompatibility complex risk alleles were replicated: HLA-DRB1\*15:01 (odds ratio [OR] = 2.02 [95% confidence interval: 1.54-2.63],  $p = 2.50e-07$ ), HLA-DRB1\*03:01 (OR = 1.58 [1.29-1.94],  $p = 1.11e-05$ ), as well as HLA-DRB1\*04:05 (OR = 2.35 [1.26-4.37],  $p = 0.007$ ) and the African-specific risk allele of HLA-DRB1\*15:03 (OR = 1.26 [1.05-1.51],  $p = 0.012$ ). The protective association of HLA-A\*02:01 was confirmed (OR = 0.72 [0.55-0.93],  $p = 0.013$ ). None of the HLA-DQB1 alleles were associated with MS. Using a significance threshold of  $p < 0.01$ , outside the major histocompatibility complex region, 8 MS SNPs were also found to be associated with MS in African Americans.

**Conclusion:** MS genetic risk in African Americans only partially overlaps with that of Europeans and could explain the difference of MS prevalence between populations. *Neurology*® 2013;81:219-227

## GLOSSARY

**AIM** = ancestry informative marker; **CI** = confidence interval; **GWAS** = genome-wide association studies; **HLA** = human leukocyte antigen; **LD** = linkage disequilibrium; **MAF** = minor allele frequencies; **MHC** = major histocompatibility complex; **MS** = multiple sclerosis; **OR** = odds ratio; **PC** = principal component; **SNP** = single nucleotide polymorphism.

Genome-wide association studies (GWAS) have contributed considerably to the understanding of multiple sclerosis (MS) susceptibility through the identification of genetic variants influencing risk and quantitation of their effects. To date, the results of 9 GWAS<sup>1-9</sup> and 2 meta-analyses<sup>10,11</sup> have been reported, generating a roster of nearly 60 candidate genes in support of a polygenic model of pathogenesis<sup>12,13</sup> driven primarily by relatively common alleles. Most noteworthy is a recent multicenter GWAS<sup>8</sup> that replicated nearly all of the previously GWAS-suggested associations together with the identification of 29 novel susceptibility loci. Immunologically relevant genes, including epistatic signals across the major histocompatibility complex (MHC) region, dominate the genomic signature of this disease. Invariably, all reported MS GWAS focused on high-prevalence datasets of European descent.<sup>14</sup> The transferability of these results to other ancestral groups remains to be addressed.

Supplemental data at  
[www.neurology.org](http://www.neurology.org)

From the Department of Neurology (N.I., P.-A.G., S.J.C., A.S., P.K., B.A.C.C., S.L.H., J.R.O.), School of Medicine, University of California, San Francisco; Department of Neurology (H.F.H.), Oslo University Hospital and University of Oslo, Norway; National Marrow Donor Program (M.M.), Minneapolis; Center for International Blood and Marrow Transplant Research (S.S.), Minneapolis, MN; Histogenetics LLC (N.C., S.Y.Y.), Ossining, NY; Department of Pathology (M.J.P.), Stanford University School of Medicine, Stanford, CA; Department of Neurology (L.P., A.H.C.), Washington University, St. Louis, MO; Program in Translational NeuroPsychiatric Genomics (P.L.D.), Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston; Harvard Medical School (P.L.D.), Boston; and Broad Institute of Harvard University and Massachusetts (P.L.D.), Institute of Technology, Cambridge, MA.

The views expressed in this article do not reflect the official policy or position of the NIH, the Department of the Navy, the Department of Defense, or any other agency of the US government.

Go to [Neurology.org](http://Neurology.org) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

African Americans have a lower disease risk compared with Europeans and white Americans, but appear to carry a greater risk of ambulatory disability.<sup>15,16</sup> The differences in genome linkage disequilibrium (LD) patterns and clinical phenotype between African Americans and white Americans were successfully leveraged to probe and refine the genetics of MS using conventional association, admixture, and genotype-phenotype studies.<sup>15–21</sup> Here, we extend these studies and test the updated MS genetic map in a well-characterized African American MS dataset. After controlling for admixture, we confirm the multiallelic human leukocyte antigen (HLA)-DRB1 association, but only 8 of 63 tested non-MHC MS loci replicated in this dataset.

**METHODS Subjects.** DNA samples from 1,162 African American patients with MS and 2,092 controls, and 577 white American cases and 461 controls were obtained through a recruitment effort initiated in 2000 using stringent inclusion and exclusion criteria as previously described.<sup>21</sup> The African American control dataset included 1,178 de-identified samples from the National Marrow Donor Program. The African American dataset overlaps approximately 48% with our previous report.<sup>21</sup>

**Standard protocol approvals, registrations, and patient consents.** The University of California at San Francisco Institutional Review Board and the National Marrow Donor Program Institutional Review Board approved this study. For all non-identified individuals participating in this study, written informed consent was obtained.

**Single nucleotide polymorphism genotyping.** Fifty-two ancestry informative marker (AIM) single nucleotide polymorphisms (SNPs) and 76 MS SNPs were genotyped using the TaqMan OpenArray genotyping technology (Life Technologies, Carlsbad, CA). AIM SNPs were selected from previously published studies<sup>22–27</sup> to cover all chromosomes. MS SNPs were selected from the latest GWAS<sup>8</sup> and meta-analysis.<sup>11</sup> Samples were loaded into customized TaqMan OpenArray genotyping plates with the OpenArray Autoloader and amplified in a Dual Flat Block GeneAmp PCR System 9700, as recommended by the manufacturer. The OpenArray Genotyping Analysis Software was used for assigning genotypes.

**Quality control.** SNPs had to meet several criteria to be included in the analysis: less than 10% missing genotypes, Hardy-Weinberg proportion test  $p > 0.001$  in controls, and no significant SNP call rate differences between patients and controls in each population ( $p > 0.001$ ), whereas samples needed to have missing genotype calls of less than 10%. This standard quality control removed 10 SNPs in total (7.8%): 2 AIM SNPs because of low SNP call rates, 7 AIM SNPs because of violation of Hardy-Weinberg proportion, and one MS SNP, rs2150702, located in *MLANA*, because of a significant difference in call rates between cases and controls in African Americans. The average sample call rate for 75 MS SNPs was  $99.3\% \pm 1.4\%$  in African Americans and  $99.4\% \pm 1.2\%$  in white Americans, whereas the call rate for 43 AIM SNPs

was  $99.0\% \pm 2.4\%$  in both populations. Altogether, 118 SNPs, 1,144 African American MS cases, 1,986 African American controls, 569 white American MS cases, and 460 white American controls were included in the association analysis.

**HLA allele typing.** Four-digit HLA-DRB1 allele data were available for 454 African American patients with MS and 1,124 controls. For an additional 487 patients and 470 controls, only HLA-DRB1\*15:01, \*15:03, \*03:01, and \*03:02 genotypes were available. Four-digit HLA-DQB1 allele data were available for 455 patients and 520 controls. Additionally, to explore the association of putative protective allele HLA-A\*02:01,<sup>8,28</sup> we generated HLA-A genotypes for 456 patients and 1,199 controls. In this dataset, 70.0% of HLA-DRB1 typing, 91.1% of HLA-DQB1 typing, and 95.3% of HLA-A\*02:01 typing were obtained by the sequence-based typing technique, and the remaining data were typed by sequence-specific oligonucleotide hybridization or sequence-specific priming.

**Molecular assessment of genetic ancestry.** The 43 AIM SNPs (including 5 in the extended MHC region) exceeding established quality-control thresholds were set in a principal component (PC) analysis to identify the degree of European ancestry admixture in each individual. Estimates were checked against our previously reported percentages of European ancestry<sup>19,20,29</sup> using Pearson correlation test. To assess population stratification, the values of the major PCs between cases and controls in both of the populations were compared using Wilcoxon rank sum test.

**Genetic association analysis.** An additive model of multivariate logistic regression was applied to assess the association of the genetic variants with MS. As covariates, the ancestral PC values and HLA class II alleles were used when relevant. For each MS variant, the Cochran Heterogeneity Q Test was performed to test effect size differences between African Americans and Europeans. Finally, to evaluate the statistical power embedded in the African American dataset, 1,000 bootstrap simulations were performed in the white reference dataset. For each simulation, the number of MS SNPs validated by association analysis was returned and distribution of the numbers from those simulations was compared with the actual number of validated MS SNPs in the African American dataset. All of the analyses were conducted using R software (version 2.13.0; R Foundation for Statistical Computing, Vienna, Austria). Odds ratios (ORs) were used to reflect effect sizes of MS risk. Confidence intervals (CIs) are given at 95%, and all of the  $p$  values are provided uncorrected.

**RESULTS Clinical and demographic features of the enrolled participants.** In this study, African American patients with MS had a higher female to male ratio and younger age at onset compared with white Americans (table 1). The proportion of primary progressive or progressive relapsing MS was larger and the MS severity score was significantly higher in African American patients than in those of European origin, which can be partially explained by the lower proportion of progressive MS in this white American group compared with that of white Americans in the literature.<sup>8</sup> Interestingly, family history of MS as reported by the proband was 17.3%, which is within the range of that reported for European datasets.<sup>30</sup>

PC analysis was used to assess ancestry and admixture, and to control for the pervasive effects of

**Table 1** Clinical and demographic features of the enrolled participants

	African Americans (n = 3,130)	White Americans (n = 1,029)
MS, n (M:F)	1,144 (1:3.7)	569 (1:2.2)
Controls, n (M:F)	1,986 (1:3.4)	460 (1:2.1)
Age at onset, y, mean ± SD (n)	32.59 ± 10.22 (1,060)	33.58 ± 9.42 (557)
Disease course		
CIS + RR/SP (%)	682/257 (91.0)	491/48 (96.6)
PP + PR (%)	93 (9.0)	19 (3.4)
MSSS, mean ± SD (n)	5.90 ± 2.39 (939)	4.27 ± 2.06 (557)
% European ancestry, <sup>a</sup> mean ± SD (n)	20.36 ± 10.90 (2,355) <sup>b</sup>	

Abbreviations: AIM = ancestry informative marker; CIS = clinically isolated syndrome; MS = multiple sclerosis; MSSS = multiple sclerosis severity score; PC = principal component; PP = primary progressive; PR = progressive relapsing; RR = relapsing-remitting; SNP = single nucleotide polymorphism; SP = secondary progressive.

<sup>a</sup>The % European ancestry was imputed using the previously reported values<sup>18–20,29</sup> and the values of the PC1 of the PC analysis with 43 AIM SNPs. In further analyses, we applied the original values of PC1 and PC2.

<sup>b</sup>In African American controls, the mean value of % European ancestry was 20.33% ± 11.41%.

population stratification in follow-up analyses. The first component of the PC analysis explained 50.77% of the interindividual European ancestry variance as informed by the 43 AIM SNPs (figure e-1A on the *Neurology*<sup>®</sup> Web site at [www.neurology.org](http://www.neurology.org)). PC2 explained 2.30% of the variances, and the biplot showed that the main contributors for PC2 were SNPs located in the extended MHC region (figure e-1B). PC1 was correlated with the previously reported percentage of European ancestry in African Americans with MS<sup>18</sup> ( $R = 0.84$ ,  $p < 2.2e-16$ ), and the mean percentage of European ancestry in African Americans was 20.4%. From the scree plot and biplot, both PC1 and PC2 were used to account for population stratification in the association analysis: PC1 as signals of genome-wide and PC2 as major MHC signals for ancestry. The values of PC1 differentiated the African American and white American groups ( $p < 2.2e-16$ ), but were not different according to the affectation status in either population (figure e-1C). The variance of PC1 was smaller in cases than in controls within African Americans ( $p = 7.22e-06$ ), suggesting higher heterogeneity in the African American control dataset. PC2 values were not statistically different based on the affectation status in African Americans, but a trend was observed in white Americans ( $p = 0.039$ ) (figure e-1D).

**Association analysis for HLA alleles.** HLA-DRB1\*15:01 and HLA-DRB1\*03:01 were significantly associated, as expected, with disease risk in African Americans (table 2). Adjustment of the association by the PC1 and (MHC-driven) PC2 retained the statistical significance. When fitting these 2 alleles in the multivariate

model, HLA-DRB1\*15:03 frequencies were elevated in African Americans with MS, as previously shown.<sup>31</sup> HLA-DRB1\*04:05 was confirmed as a risk allele, whereas 2 alleles, HLA-DRB1\*11:01 and HLA-DRB1\*04:01, were found to be associated with reduced MS risk. No association of HLA-DQB1 alleles was detected for African American MS (table 3). Additionally, association analysis for HLA-A\*02:01 showed a protective effect for carriers, independent of the top 2 HLA-DRB1 risk alleles of HLA-DRB1\*15:01 and HLA-DRB1\*03:01 (OR = 0.72 [95% CI 0.55–0.93],  $p = 0.013$ ).

In the association analysis for SNPs in the MHC region, rs3129889, tagging HLA-DRB1, was the most significantly associated with MS in both African Americans and white Americans, but with a different effect size between the 2 populations, i.e., a larger effect in whites (Q test  $p = 3.91e-07$ ). No association was found in African Americans for the HLA-B SNP, rs2523393, tagging HLA-B\*44 in Europeans (table e-1).<sup>10</sup>

**Replication study of the MS-associated SNPs in non-MHC regions.** Among the 73 MS SNPs in non-MHC regions that passed quality control, 15 SNPs were replicated in African Americans with  $p$  values  $< 0.05$  (tables 4 and e-1). Using the  $p < 0.01$  threshold, 8 MS-validated risk markers—rs7200786 (*CLEC16A*), rs1335532 (*CD58*), rs13333054 (*IRF8*), rs180515 (*RPS6KB1*), rs4613763 (*PTGER4*), rs1800693 (*TNFRSF1A*), rs7238078 (*MALTI1*), and rs669607 (no gene)—were replicated in the African American dataset. Among those validated SNPs in non-MHC regions, no significant differences were detected for each effect size between African Americans and populations from European descent as reported in the literature (Q test  $p > 0.1$ ). However, even though rs2300603 (*BATF*) was replicated ( $p = 0.017$ ), the direction of the effect size was reversed in African Americans (OR = 0.81) compared with that reported in Europeans (Q test  $p = 0.002$ ).

In the white American dataset, 4 non-MHC SNPs were replicated with  $p$  values  $< 0.01$ . To investigate the role of sample size in our analysis, bootstrapping of the white American dataset was conducted up to the sample size of the African Americans. The mean number of the statistically associated non-MHC SNPs in each simulation in white Americans was higher (mean 23.8) than the number of the statistically associated SNPs in African Americans ( $n = 8$ ), suggesting that the observed disparity in the number of validated SNPs between African Americans and white Americans is not entirely attributable to the size of the datasets.

**Application of the frequently associated SNPs to the 1000 Genomes Project population.** To address the hypothesis that the worldwide distribution of MS correlates with

**Table 2** Association of HLA-DRB1 alleles for MS in African Americans<sup>a</sup>

HLA-DRB1	MS, <sup>b</sup> n (%)	Controls, <sup>b</sup> n (%)	Multivariate model								
			Monovariate model			Ancestry-adjusted <sup>c</sup>			Ancestry- and MHC signal-adjusted <sup>d</sup>		
			OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
15:01	124 (6.6)	112 (3.5)	1.93	1.48–2.51	9.70e-07	2.06	1.57–2.69	1.44e-07	NA	NA	NA
03:01	195 (10.4)	224 (7.0)	1.54	1.26–1.89	3.01e-05	1.56	1.27–1.91	2.20e-05	NA	NA	NA
04:05	19 (2.1)	23 (1.0)	2.09	1.13–3.88	0.019	2.05	1.10–3.81	0.023	2.35	1.26–4.37	0.007
15:03	252 (13.4)	369 (11.6)	1.20	1.00–1.43	0.049	1.18	0.99–1.41	0.069	1.26	1.05–1.51	0.012
11:01	52 (5.7)	203 (9.0)	0.61	0.44–0.84	0.002	0.61	0.44–0.84	0.002	0.68	0.49–0.93	0.017
04:01	9 (1.0)	52 (2.3)	0.43	0.21–0.87	0.019	0.43	0.21–0.89	0.022	0.48	0.23–0.97	0.041
16:02	18 (2.0)	29 (1.3)	1.56	0.86–2.84	0.146	1.54	0.84–2.80	0.159	1.81	0.99–3.30	0.054
14:01	11 (1.2)	52 (2.3)	0.53	0.28–1.01	0.054	0.53	0.28–1.01	0.053	0.55	0.29–1.05	0.070
09:01	19 (2.1)	81 (3.6)	0.57	0.35–0.95	0.031	0.58	0.35–0.96	0.033	0.63	0.38–1.05	0.079
08:04	52 (5.7)	104 (4.6)	1.25	0.89–1.76	0.198	1.25	0.88–1.76	0.208	1.34	0.95–1.90	0.094
11:02	23 (2.5)	91 (4.1)	0.61	0.38–0.98	0.039	0.61	0.38–0.97	0.038	0.67	0.42–1.08	0.100
07:01	97 (10.7)	220 (9.8)	1.10	0.86–1.43	0.444	1.10	0.86–1.43	0.445	1.23	0.95–1.59	0.124
13:03	22 (2.4)	80 (3.6)	0.67	0.41–1.08	0.102	0.67	0.41–1.08	0.100	0.71	0.44–1.16	0.170
01:02	27 (3.0)	91 (4.1)	0.73	0.47–1.13	0.154	0.72	0.47–1.12	0.148	0.79	0.51–1.23	0.304
03:02	121 (6.4)	207 (6.5)	0.99	0.78–1.25	0.929	0.97	0.77–1.23	0.827	1.07	0.85–1.35	0.575
13:01	41 (4.5)	117 (5.2)	0.86	0.60–1.24	0.425	0.87	0.60–1.24	0.434	0.96	0.66–1.38	0.806
10:01	15 (1.7)	43 (1.9)	0.86	0.47–1.56	0.618	0.85	0.47–1.55	0.605	0.94	0.52–1.72	0.845
13:02	57 (6.3)	153 (6.8)	0.92	0.67–1.26	0.590	0.92	0.67–1.26	0.592	0.97	0.71–1.34	0.869
12:01	34 (3.7)	93 (4.1)	0.90	0.60–1.35	0.610	0.89	0.59–1.33	0.570	0.97	0.65–1.45	0.876
01:01	20 (2.2)	55 (2.5)	0.90	0.55–1.50	0.692	0.92	0.55–1.52	0.739	1.03	0.62–1.72	0.899

Abbreviations: CI = confidence interval; HLA = human leukocyte antigen; MHC = major histocompatibility complex; MS = multiple sclerosis; NA = not applicable; OR = odds ratio; PC = principal component.

<sup>a</sup>All of the alleles in which frequencies in controls are  $\geq 1.0\%$  are shown.

<sup>b</sup>For HLA-DRB1\*15:01, \*15:03, \*03:01, and \*03:02, 941 patients and 1,594 controls were enrolled, whereas for the other HLA-DRB1 alleles, data from 454 patients and 1,124 controls were included.

<sup>c</sup>Adjusted by the values of PC1 (OR = 1.01 [95% CI 0.96–1.06],  $p = 0.769$ ) and PC2 (OR = 0.94 [95% CI 0.85–1.05],  $p = 0.261$ ).

<sup>d</sup>Adjusted by the values of PC1 (OR = 0.97 [95% CI 0.91–1.03],  $p = 0.343$ ) and the doses of HLA-DRB1\*15:01 (OR = 2.02 [95% CI 1.54–2.63],  $p = 2.50e-07$ ) and HLA-DRB1\*03:01 (OR = 1.58 [95% CI 1.29–1.94],  $p = 1.11e-05$ ).

the frequency of the risk variants in each ancestral group, the cumulative genetic risk scores accounting for the minimal set of the 8 non-MHC SNPs with  $p$  values  $< 0.01$  in African Americans were computed for samples included in the 1000 Genomes Project dataset.<sup>32</sup> The scores were significantly different according to the populations (figure 1). The African group had the lowest genetic risk scores and there was an increasing gradient from Africans to African Americans, Asians, Hispanics, and Europeans. This distribution mimics the worldwide disease prevalence. Additionally, the genetic burden scores of our MS case-control datasets fit within the observed ancestral gradient, and patients were more loaded than controls in both African Americans and white Americans ( $p = 4.11e-13$  and  $p = 0.031$ , respectively).

**DISCUSSION** This is the largest genetic study of an MS dataset of non-European ancestry. After adjusting

for admixture, we replicated established HLA-DRB1 allelic associations as well as 8 of the 73 validated non-MHC MS SNPs discovered in European populations. African American patients with MS carry the 2 most prominent susceptibility HLA-DRB1 alleles found in Europeans, HLA-DRB1\*15:01 and HLA-DRB1\*03:01, as well as the African allele HLA-DRB1\*15:03.<sup>8,20,31</sup> The role of HLA-DRB1\*04:05 observed in this study was also reported in other populations, including Japanese and Sardinian.<sup>33,34</sup> We also confirmed the protective association of HLA-A\*02:01 in African Americans, independent of the HLA-DRB1 risk alleles. Also consistent with previous reports,<sup>35</sup> a lower frequency for HLA-DRB1\*11 was observed. Although the protective association of HLA-DRB1\*09:01 lost statistical significance after adjustment, the direction of the effect matched previous results in Asians.<sup>33,36</sup> No association of HLA-DQB1 alleles was detected for African American



**Table 3 Association of HLA-DQB1 alleles for MS in African Americans<sup>a</sup>**

HLA-DQB1	MS, <sup>b</sup> n (%)	Controls, <sup>b</sup> n (%)	Multivariate model								
			Monovariate model			Ancestry-adjusted <sup>c</sup>			Ancestry- and MHC signal-adjusted <sup>d</sup>		
			OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
02 <sup>e</sup>	221 (24.3)	225 (21.6)	1.17	0.94–1.45	0.157	1.16	0.94–1.44	0.173	0.99	0.76–1.30	0.936
03:01	128 (14.1)	176 (16.9)	0.79	0.61–1.02	0.074	0.79	0.61–1.02	0.073	0.86	0.66–1.13	0.292
03:02	40 (4.4)	36 (3.5)	1.26	0.81–1.97	0.305	1.28	0.82–2.00	0.279	1.29	0.82–2.04	0.278
03:03	8 (0.9)	11 (1.1)	0.83	0.33–2.08	0.688	0.85	0.34–2.14	0.727	1.05	0.39–2.86	0.916
04:02	72 (7.9)	82 (7.9)	1.00	0.72–1.40	0.982	0.99	0.71–1.39	0.969	1.16	0.81–1.66	0.430
05:01	125 (13.7)	154 (14.8)	0.92	0.71–1.18	0.502	0.91	0.71–1.18	0.481	0.94	0.72–1.22	0.628
05:02	24 (2.6)	23 (2.2)	1.20	0.67–2.16	0.536	1.20	0.67–2.16	0.543	1.10	0.61–1.98	0.761
05:03	9 (1.0)	14 (1.4)	0.75	0.33–1.69	0.484	0.75	0.33–1.70	0.494	0.79	0.34–1.83	0.579
06:02	207 (22.8)	234 (22.5)	1.01	0.82–1.25	0.896	1.02	0.82–1.26	0.869	0.97	0.76–1.25	0.829
06:03	21 (2.3)	27 (2.6)	0.88	0.49–1.59	0.678	0.91	0.51–1.64	0.760	1.17	0.61–2.24	0.644
06:04	17 (1.9)	24 (2.3)	0.81	0.44–1.51	0.507	0.83	0.44–1.54	0.551	0.91	0.47–1.75	0.772
06:09	33 (3.6)	30 (2.9)	1.26	0.77–2.06	0.364	1.25	0.76–2.05	0.385	1.27	0.76–2.12	0.361

Abbreviations: CI = confidence interval; HLA = human leukocyte antigen; MHC = major histocompatibility complex; MS = multiple sclerosis; OR = odds ratio; PC = principal component.

<sup>a</sup> All of the alleles in which frequencies in controls are  $\geq 1.0\%$  are shown.

<sup>b</sup> For HLA-DQB1 alleles, data from 455 patients and 520 controls were included.

<sup>c</sup> Adjusted by the values of PC1 (OR = 1.01 [95% CI 0.96–1.06],  $p = 0.769$ ) and PC2 (OR = 0.94 [95% CI 0.85–1.05],  $p = 0.261$ ).

<sup>d</sup> Adjusted by the values of PC1 (OR = 0.97 [95% CI 0.91–1.03],  $p = 0.343$ ) and the doses of HLA-DRB1\*15:01 (OR = 2.02 [95% CI 1.54–2.63],  $p = 2.50 \times 10^{-7}$ ) and HLA-DRB1\*03:01 (OR = 1.58 [95% CI 1.29–1.94],  $p = 1.11 \times 10^{-5}$ ).

<sup>e</sup> All the alleles with HLA-DQB1\*02 were assembled together.

patients with MS, consistent with our previous study highlighting the primary role of HLA-DRB1 vs HLA-DQB1.<sup>31</sup> Interestingly, all of these MS-associated alleles are found in various HLA haplotypes in different populations,<sup>13,37</sup> consistent with a primary role for HLA-DRB1 and HLA-A alleles rather than other HLA genes. When the effect size of HLA-DRB1\*15 was compared in our African American and white American datasets, it was different (Q test  $p = 1.65 \times 10^{-6}$ ). Whether this difference reflects a true biological effect of the MHC in these 2 populations remains to be addressed.

Outside the MHC, the number of statistically associated MS SNPs in African Americans was far lower than the mean number of MS-associated SNPs in the bootstrapped white American dataset. There are 3 possible explanations for this partial replication. First, differences in minor allele frequencies (MAF), especially for those SNPs with lower MAF in African Americans, could have affected the power to detect the associations. Indeed, among the MS SNPs with a MAF less than 15% in African Americans, only the SNP rs3129889 in the HLA-DRB1 region was replicated. Second, the smaller size of African haplotype blocks<sup>32,38</sup> may have caused some of the MS SNPs originally identified in Europeans to fail tagging putative causative variants in African Americans. For example, our previous study in African Americans showed

associations for 2 SNPs in *EVI5* (rs10735781 and rs6680578).<sup>21</sup> Although the SNP genotyped in this study (rs11810217) locates between rs10735781 and rs6680578 and appears to be within the same haplotype block in Europeans (measures of LD,  $R^2$ , and  $d'$  are strong among the 3 SNPs), the  $R^2$  between rs11810217 and the other variants, rs10735781 and rs6680578, is extremely low in African Americans. Similarly, we can attribute the lack of replication for *CD6* and *TYK2*, despite suggestive association signals in our previous report,<sup>21</sup> to SNP selection and LD haplotype differences between African Americans and Europeans. The observed opposite direction of the association of the SNP rs2300603 in *BATF* can be interpreted as a possible difference in the LD pattern between this SNP and the putative causative variants in African Americans, considering the observed smaller haplotype blocks at the locus in African Americans than in Europeans.<sup>39</sup> It is important to note that for the 57 non-MHC SNPs reported by Sawcer et al.,<sup>8</sup> the number of possible genes corresponding to one SNP signal ranges between 1 and 33. Finally, genetic heterogeneity across populations is an additional conceivable explanation for absence of replication, at least for some of the loci (table e-2). It is of interest that the individual accumulation of genetic risk alleles across populations for this minimal set of common risk alleles matches the disease prevalence in each ancestral group,

**Table 4** Replication study in African Americans for MS-associated SNPs in Europeans<sup>a</sup>

rsID	Chr	GOI	African Americans				White Americans				Literature <sup>d</sup> OR (95% CI)	Meta-analysis	
			No. (%) of allele <sup>b</sup> in		Ancestry-adjusted <sup>c</sup>		No. (%) of allele <sup>b</sup> in		Ancestry-adjusted <sup>c</sup>			HET <sup>e</sup> p	Global OR <sup>f</sup> (95% CI)
			MS	Controls	OR (95% CI)	p	MS	Controls	OR (95% CI)	p			
rs3129889	6	HLA-DRB1	175 (7.7)	193 (4.9)	1.65 (1.33–2.05)	5.30e-06	282 (25.1)	86 (9.4)	3.33 (2.54–4.36)	3.27e-18	2.97 (2.77–3.18)	3.91e-07	2.23 (1.26–3.97)
rs7200786	16	CLEC16A	711 (31.1)	1,045 (26.3)	1.26 (1.12–1.41)	6.43e-05	537 (47.4)	439 (47.8)	0.98 (0.83–1.17)	0.863	1.15 (1.11–1.20)	0.141	1.16 (1.12–1.20)
rs1335532	1	CD58	1,257 (55.1)	1,987 (50.5)	1.21 (1.09–1.34)	4.41e-04	1,019 (89.9)	789 (86.5)	1.40 (1.07–1.85)	0.016	1.18 (1.12–1.24)	0.669	1.19 (1.13–1.24)
rs13333054	16	IRF8	556 (24.4)	829 (20.9)	1.22 (1.08–1.37)	0.002	267 (23.5)	192 (20.9)	1.16 (0.94–1.43)	0.169	1.12 (1.08–1.17)	0.182	1.13 (1.09–1.17)
rs180515	17	RPS6KB1	661 (29.4)	995 (25.8)	1.20 (1.07–1.35)	0.002	372 (33.3)	301 (32.7)	1.03 (0.85–1.24)	0.764	1.05 (0.99–1.12)	0.047	1.08 (1.02–1.14)
rs4613763	5	PTGER4	496 (21.8)	739 (18.6)	1.22 (1.07–1.38)	0.003	136 (12.0)	114 (12.4)	0.96 (0.74–1.25)	0.778	1.21 (1.15–1.28)	0.907	1.21 (1.15–1.27)
rs1800693	12	TNFRSF1A	904 (40.9)	1,406 (37.1)	1.17 (1.05–1.30)	0.004	500 (44.5)	385 (42.6)	1.08 (0.91–1.29)	0.388	1.12 (1.08–1.16)	0.447	1.12 (1.09–1.16)
rs7238078	18	MALT1	1,746 (77.8)	2,931 (74.7)	1.19 (1.05–1.34)	0.007	862 (76.4)	681 (75.2)	1.08 (0.87–1.33)	0.487	1.14 (1.06–1.23)	0.556	1.15 (1.08–1.23)
rs669607	3	No gene	429 (18.8)	640 (16.1)	1.21 (1.05–1.38)	0.007	562 (49.6)	425 (46.3)	1.14 (0.96–1.36)	0.140	1.15 (1.08–1.23)	0.510	1.16 (1.09–1.23)
rs2248359	20	CYP24A1	894 (39.3)	1,427 (36.1)	1.14 (1.03–1.27)	0.013	660 (58.1)	549 (59.9)	0.93 (0.78–1.11)	0.409	1.11 (1.04–1.19)	0.675	1.12 (1.06–1.18)
rs2300603	14	BATF	2,039 (89.2)	3,606 (91.1)	0.81 (0.68–0.96)	0.017	856 (75.5)	696 (76.0)	0.97 (0.79–1.19)	0.785	1.08 (1.01–1.16)	0.002	0.95 (0.71–1.25)
rs11129295	3	EOMES	826 (36.1)	1,318 (33.2)	1.14 (1.02–1.27)	0.020	495 (43.5)	344 (37.5)	1.27 (1.07–1.51)	0.007	1.09 (1.02–1.16)	0.489	1.10 (1.04–1.17)
rs4648356	1	MMEL1	1,270 (55.9)	2,073 (53.0)	1.13 (1.02–1.25)	0.024	797 (70.0)	609 (66.6)	1.17 (0.97–1.41)	0.100	1.16 (1.12–1.21)	0.637	1.16 (1.12–1.20)
rs12048904	1	VCAM1	1,690 (74.4)	2,829 (71.8)	1.14 (1.02–1.28)	0.027	467 (41.6)	366 (40.1)	1.06 (0.89–1.27)	0.511	1.08 (1.01–1.15)	0.418	1.09 (1.03–1.16)
rs13192841	6	OLIG3	441 (19.3)	679 (17.1)	1.16 (1.01–1.32)	0.032	339 (29.8)	291 (31.7)	0.92 (0.76–1.11)	0.363	1.10 (1.06–1.15)	0.457	1.10 (1.06–1.15)
rs10201872	2	SP140	239 (10.5)	354 (8.9)	1.20 (1.01–1.44)	0.040	207 (18.3)	165 (17.9)	1.03 (0.82–1.29)	0.821	1.15 (1.06–1.24)	0.667	1.16 (1.08–1.24)

Abbreviations: AIM = ancestry informative marker; Chr = chromosome; CI = confidence interval; GOI = gene of interest; HET = heterogeneity; MS = multiple sclerosis; OR = odds ratio; PC = principal component; SNP = single nucleotide polymorphism.

<sup>a</sup>Results of all MS SNPs are available in table e-1.

<sup>b</sup>The allele frequency of the one reported to be a risk allele in Europeans.

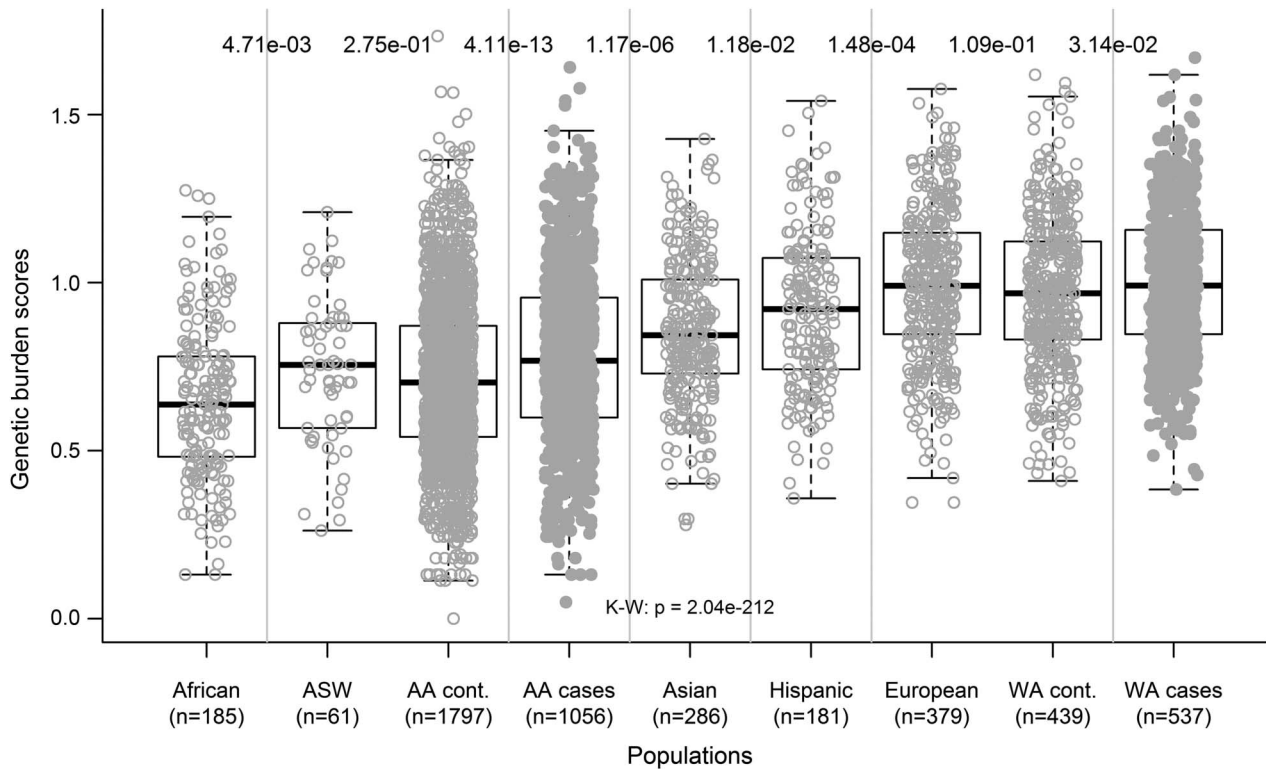
<sup>c</sup>Adjusted by PC1 values. For those individuals without PC1 values because of missing genotype for at least one AIM SNP, the mean PC1 values of the individual's population were applied.

<sup>d</sup>References 8 and 11.

<sup>e</sup>Cochrane Heterogeneity Q Test.

<sup>f</sup>For SNPs in which the p values of the Cochrane Heterogeneity Q Test were <0.01, a random-effect model was applied to calculate the global ORs and 95% CIs; otherwise, a fixed-effect model was applied.

**Figure 1** MS genetic burden scores in the 1000 Genomes Project population



Based on the 8 non-MHC MS SNPs for which the association with MS were replicated in African Americans ( $p < 0.01$ ), aggregation of those SNPs was computed for individuals in the 1000 Genomes Project dataset, as previously described.<sup>40</sup> Luhya in Webuye, Kenya, and Yoruba in Ibadan, Nigeria, are combined as “African”; Han Chinese in Beijing, China, Han Chinese South, and Japanese in Tokyo, Japan, are combined as “Asian”; Colombian in Medellin, Colombia, Mexican Ancestry in Los Angeles, CA, and Puerto Rican in Puerto Rico are combined as “Hispanic”; and Toscani in Italy, Iberian populations in Spain, Finnish from Finland, British from England and Scotland, and Northern and Western Europeans from Utah are combined as “European.” Additionally, genetic burden scores of our case-control dataset in AA and WA are included. A worldwide gradient of genetic burden scores was seen, matching MS prevalence in the world. Patients were more loaded than controls in both AA and WA. The  $p$  values at the top of the figure indicate the statistical significance of comparisons in groups next to each other (Wilcoxon rank sum test). The  $p$  value at the middle, bottom was calculated by Kruskal-Wallis test. Uncorrected  $p$  values are shown. AA = African Americans; ASW = African ancestry from southwest United States; K-W = Kruskal-Wallis test; MHC = major histocompatibility complex; MS = multiple sclerosis; SNP = single nucleotide polymorphism; WA = white Americans.

suggesting that the global frequency of the disease is influenced, at least in part, by the distribution of risk alleles and that these validated SNPs tag causative variants across different populations.

Building on the success of the GWAS approach, meta-analyses and dense genotyping in large datasets will most likely result in the rapid identification of additional risk alleles. Studies comparing the genomes of different ancestral groups may prove to be highly informative by leveraging the differences in genome LD patterns and clinical phenotypes.

#### AUTHOR CONTRIBUTIONS

Dr. Isobe participated in study design, performed experiments and statistical analysis, and drafted and edited the manuscript. Dr. Gourraud participated in study design, supervised and performed the statistical analyses, and critically revised the manuscript. Dr. Harbo conducted experiments and critically revised the manuscript. Ms. Caillier was responsible for sample preparation and management. Mr. Santaniello was responsible for data management. Mr. Khankhanian supported the statistical analyses. Mr. Maiers, Mr. Spellman, Dr. Cereb, and Dr. Yang contributed samples and data and critically revised the manuscript. Dr. Pando contributed data and critically revised the manuscript.

Dr. Piccio, Dr. Cross, and Dr. De Jager contributed samples and critically revised the manuscript. Dr. Cree and Dr. Hauser contributed to the phenotypic characterization of the study participants and critically revised the manuscript. Dr. Oksenberg designed and supervised the study and edited the manuscript.

#### ACKNOWLEDGMENT

The authors thank the MS patients and healthy controls who participated in this study. The authors acknowledge the contributions of H. Mousavi and R. Guerrero (UCSF) for sample processing and management, R. O’Shea and J. More (UCSF) for the management of the clinical data, and L. Maddireddy (UCSF) for support in writing R commands.

#### STUDY FUNDING

This study was supported by NIH grants RO1NS076492 and RO1NS046297. Recruitment of study participants and sample acquisition were supported by NMSS RG2899-D11 and RC2 GM093080. The National Marrow Donor Program is supported in part by funding from the Office of Naval Research N00014-11-1-0339 and the Health Resources and Services Administration (HRSA/DHHS contract HHS234200637020C).

#### DISCLOSURE

N. Isobe received fellowship from the Uehara Memorial Foundation and has received funding from a postdoctoral fellowship for research abroad from the Japan Society for the Promotion of Science. P.-A. Gourraud

is a recipient of the Nancy Davis Foundation young investigator award. H. Harbo is supported by the Research Council of Norway (grants 189639/196776). S. Caillier, A. Santaniello, P. Khankhanian, M. Maier, and S. Spellman report no disclosures. N. Cereb is a CEO and cofounder of Histogenetics LLC, which performed HLA typing for this manuscript. S. Yang is a chairman and founder of Histogenetics LLC, which performed HLA typing for this manuscript. M. Pando reports no disclosures. L. Piccio is a Harry Weaver Neuroscience Scholar of the National MS Society (JF2144A2/1) and received support from the Fondazione Italiana Sclerosi Multipla (FISM) (2009/R/33). A. Cross reports no disclosures. P. De Jager is a Harry Weaver Neuroscience Scholar of the National MS Society (JF2138A1). B. Cree, S. Hauser, and J. Oksenberg report no disclosures. Go to [Neurology.org](http://Neurology.org) for full disclosures.

Received November 30, 2012. Accepted in final form April 4, 2013.

## REFERENCES

- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–678.
- Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007;357:851–862.
- Comabella M, Craig DW, Camina-Tato M, et al. Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. *PLoS One* 2008;3:e3490.
- Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009;18:767–778.
- Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009;41:824–828.
- Jakkula E, Leppä V, Sulonen AM, et al. Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. *Am J Hum Genet* 2010;86:285–291.
- Sanna S, Pitzalis M, Zoledziewska M, et al. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet* 2010;42:495–497.
- Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011;476:214–219.
- Matesanz F, Gonzalez-Perez A, Lucas M, et al. Genome-wide association study of multiple sclerosis confirms a novel locus at 5p13.1. *PLoS One* 2012;7:e36140.
- De Jager PL, Jia X, Wang J, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009;41:776–782.
- Patsopoulos NA, Esposito F, Reischl J, et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* 2011;70:897–912.
- Bush WS, Sawcer SJ, De Jager PL, et al. Evidence for polygenic susceptibility to multiple sclerosis: the shape of things to come. *Am J Hum Genet* 2010;86:621–625.
- Gourraud PA, Harbo HF, Hauser SL, Baranzini SE. The genetics of multiple sclerosis: an up-to-date review. *Immunol Rev* 2012;248:87–103.
- Pugliatti M, Sotgiu S, Rosati G. The worldwide prevalence of multiple sclerosis. *Clin Neurol Neurosurg* 2002;104:182–191.
- Cree BA, Khan O, Bourdette D, et al. Clinical characteristics of African Americans vs Caucasian Americans with multiple sclerosis. *Neurology* 2004;63:2039–2045.
- Cree BA, Reich DE, Khan O, et al. Modification of multiple sclerosis phenotypes by African ancestry at HLA. *Arch Neurol* 2009;66:226–233.
- Barcellos LF, Begovich AB, Reynolds RL, et al. Linkage and association with the NOS2A locus on chromosome 17q11 in multiple sclerosis. *Ann Neurol* 2004;55:793–800.
- Reich D, Patterson N, De Jager PL, et al. A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nat Genet* 2005;37:1113–1118.
- Caillier SJ, Briggs F, Cree BA, et al. Uncoupling the roles of HLA-DRB1 and HLA-DRB5 genes in multiple sclerosis. *J Immunol* 2008;181:5473–5480.
- McElroy JP, Cree BA, Caillier SJ, et al. Refining the association of MHC with multiple sclerosis in African Americans. *Hum Mol Genet* 2010;19:3080–3088.
- Johnson BA, Wang J, Taylor EM, et al. Multiple sclerosis susceptibility alleles in African Americans. *Genes Immun* 2010;11:343–350.
- Smith MW, Patterson N, Lautenberger JA, et al. A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004;74:1001–1013.
- Seldin MF, Price AL. Application of ancestry informative markers to association studies in European Americans. *PLoS Genet* 2008;4:e5.
- Ruiz-Narvaez EA, Rosenberg L, Wise LA, Reich D, Palmer JR. Validation of a small set of ancestral informative markers for control of population admixture in African Americans. *Am J Epidemiol* 2011;173:587–592.
- Lao O, Vallone PM, Coble MD, et al. Evaluating self-declared ancestry of U.S. Americans with autosomal, Y-chromosomal and mitochondrial DNA. *Hum Mutat* 2010;31:E1875–E1893.
- Kidd JR, Friedlaender FR, Speed WC, Pakstis AJ, De La Vega FM, Kidd KK. Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples. *Investig Genet* 2011;2:1.
- Chiang CW, Gajdos ZK, Korn JM, et al. Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. *PLoS Genet* 2010;6:e1000866.
- Brynedal B, Duvefelt K, Jonassdottir G, et al. HLA-A confers an HLA-DRB1 independent influence on the risk of multiple sclerosis. *PLoS One* 2007;2:e664.
- Patterson N, Hattangadi N, Lane B, et al. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet* 2004;74:979–1000.
- Kahana E. Epidemiologic studies of multiple sclerosis: a review. *Biomed Pharmacother* 2000;54:100–102.
- Oksenberg JR, Barcellos LF, Cree BA, et al. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. *Am J Hum Genet* 2004;74:160–167.
- Abecasis GR, Altshuler D, Auton A, et al; 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061–1073.
- Matsuoka T, Matsushita T, Osoegawa M, et al. Association of the HLA-DRB1 alleles with characteristic MRI features of Asian multiple sclerosis. *Mult Scler* 2008;14:1181–1190.
- Cocco E, Sardu C, Pieroni E, et al. HLA-DRB1-DQB1 haplotypes confer susceptibility and resistance to multiple sclerosis in Sardinia. *PLoS One* 2012;7:e33972.



35. Ramagopalan SV, Morris AP, Dymont DA, et al. The inheritance of resistance alleles in multiple sclerosis. *PLoS Genet* 2007;3:1607–1613.
36. Qiu W, James I, Carroll WM, Mastaglia FL, Kermode AG. HLA-DR allele polymorphism and multiple sclerosis in Chinese populations: a meta-analysis. *Mult Scler* 2011;17:382–388.
37. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. *Hum Immunol* 2007;68:779–788.
38. Tishkoff SA, Kidd KK. Implications of biogeography of human populations for ‘race’ and medicine. *Nat Genet* 2004;36:S21–S27.
39. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265.
40. Gourraud PA, McElroy JP, Caillier SJ, et al. Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. *Ann Neurol* 2011;69:65–74.

## ***Neurology*<sup>®</sup> WriteClick: Join the Debate!**

The editors encourage comments about recent articles through WriteClick:

Go to [www.neurology.org](http://www.neurology.org) and click on the “WriteClick” tab at the top of the page. Responses will be posted within 72 hours of submission.

Before using WriteClick, remember the following:

- WriteClick is restricted to comments about studies published in *Neurology* within the last eight weeks
- Read previously posted comments; redundant comments will not be posted
- Your submission must be 200 words or less and have a maximum of five references; reference one must be the article on which you are commenting
- You can include a maximum of five authors (including yourself)

## **Spring AAN Webinars: Help for Your Practice, CME for Your Career**

The American Academy of Neurology offers cost-effective Practice Management Webinars that can be attended live or through convenient recordings posted online five days after the event. AAN members can save 25% on all regular webinars! Plus, physicians can earn 1.5 valuable CME credits for each webinar, and administrators receive a certificate of completion. Mark your calendar for upcoming programs and register today for these and other 2013 webinars at [www.aan.com/view/pmw13](http://www.aan.com/view/pmw13):

- |            |   |
|------------|---|
| Online Now | <b>Correct Coding for Chemodenervation</b>  |
| Online Now | <b>Working with NPs and PAs to Maximize Office Productivity</b>                         |
| Online Now | <b>Remaining Relevant in the Changing Health Care Payment and Care Delivery Systems</b> |
| Online Now | <b>Coding Accurately for Epilepsy</b>   |
| Online Now | <b>E/M: Minimize Mistakes, Maximize Reimbursement</b>                                   |