Title

The genetic diversity of multiple sclerosis risk among Hispanic and African American populations living in the United States.

Journal

Multiple sclerosis (Houndmills, Basingstoke, England), 26(11)

ISSN

1352-4585

Authors

Beecham, AH
Amezcua, L
Chinea, A
et al.

Publication Date

2020-10-01

DOI

10.1177/1352458519863764

Peer reviewed

HHS Public Access

Author manuscript

Mult Scler. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

Mult Scler. 2020 October ; 26(11): 1329–1339. doi:10.1177/1352458519863764.

The genetic diversity of multiple sclerosis risk among Hispanic and African American populations living in the United States

A.H. Beecham[1,2], **L. Amezcua**[3], **A. Chinea**[4], **C.P. Manrique**[1], **C. Rubi**[4], **N. Isobe**[5], **B.T. Lund**[3], **A. Santaniello**[6], **G.W. Beecham**[1,2], **E.G. Burchard**[7], **M. Comabella**[8], **N. Patsopoulos**[9], **K. Fitzgerald**[10], **P.A. Calabresi**[10], **P. De Jager**[11], **D.V. Conti**[12], **S.R. Delgado**[13], **J.R. Oksenberg**[6], **J.L. McCauley**[1,2]

[1]John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, Florida, United States of America

[2]Dr. John T. Macdonald Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, Florida, United States of America

[3]Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America

[4]San Juan MS Center, Guaynabo, Puerto Rico, United States of America; Universidad Central del Caribe, Bayamon, Puerto Rico, United States of America

[5]Department of Neurological Therapeutics, Graduate School of Medical Sciences, Kyushu University, Fukuoka City, Japan

[6]Department of Neurology, University of California at San Francisco, San Francisco, California, United States of America

[7]Department of Medicine and Department of Bioengineering and Therapeutic Sciences, University of California at San Francisco, San Francisco, California, United States of America

[8]Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain

[9]Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America

[10]Department of Neurology and the Solomon Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

[11]Center for Translational and Computational Neuroimmunology. Department of Neurology, Columbia University Medical Center, New York, New York, United States of America

[*]Corresponding author – Jacob L. McCauley, jmccauley@med.miami.edu, Office: (305) 243-4578, Fax: (305) 243-239, Address: 1501 NW 10th Ave, Miami, FL 33136.

The Authors declare that there is no conflict of interest.

Web Resources

Genotype data for the 200 variants have been made available at www.arhms.org.

[12]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America

[13]Multiple Sclerosis Division, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida, United States of America

Abstract

Background: Substantial progress has been made towards unraveling the genetic architecture of multiple sclerosis (MS) within populations of European ancestry, but few genetic studies have focused on Hispanic and African American populations within the United States.

Objective: We sought to test the relevance of common European MS risk variants outside of the Major Histocompatibility Complex (n=200) within these populations.

Methods: Genotype data were available on 2652 Hispanics (1298 with MS, 1354 controls) and 2435 African Americans (1298 with MS, 1137 controls). We conducted single variant, pathway, and cumulative genetic risk score analyses.

Results: We found less replication than statistical power suggested, particularly among African Americans. This could be due to limited correlation between the tested and causal variants within the sample; or alternatively could indicate allelic and locus heterogeneity. Differences were observed between pathways enriched among the replicating versus all 200 variants. Although these differences should be examined in larger samples, a potential role exists for gene-environment or gene-gene interactions which alter phenotype differentially across racial and ethnic groups. Cumulative genetic risk scores were associated with MS within each study sample but showed limited diagnostic capability.

Conclusion: These findings provide a framework for fine-mapping efforts in multi-ethnic populations of MS.

Keywords

multiple sclerosis; multi-ethnic; genetics; admixture; allelic heterogeneity; locus heterogeneity; risk score; pathway analysis

Introduction

Multiple sclerosis (MS [MIM: 126200]) is a chronic neurodegenerative disease, characterized by the presence of inflammatory demyelinating lesions in the central nervous system.[1] Through genome-wide association studies and meta-analyses,[2–5] substantial progress has been made towards unraveling the genetic architecture of MS within populations of European ancestry, to date explaining ~39% of the narrow-sense heritability (19.2%; 95% CI: 18.5-19.8%).[2]

Few genetic studies of MS have been conducted in Hispanics and African Americans[6] although they represent a sizeable proportion of the United States (US) population (~16% and ~12%, respectively).[7] These few studies, particularly those in select Hispanic sub-populations, primarily focus on HLA alleles within the Major Histocompatibility Complex (MHC).[8–10] Despite the value of minority inclusion for examining population differences,

providing insight into health disparities, understanding biology, and improving care; these exclusions persist across a variety of disease phenotypes.[11] Only 19% of published genome-wide studies reported on non-European populations in 2016.[11, 12] Further, there exists a long-held belief that prevalence of MS is lower in these populations compared to populations of European ancestry. However, epidemiological evidence suggests that prevalence may be higher than previously indicated[13–15] and vary considerably by geographical region.[16, 17]

As compared to Europeans, African Americans often exhibit greater disease severity[18] and Hispanics more often present with optic neuritis, commonly at an earlier age; although geographical location and genetic admixture play a role in differences observed across Hispanic sub-populations.[19–21] The disease heterogeneity seen between and among these populations implies that their study is essential to understanding mechanisms underlying MS genetics for all people. Moreover, genetic admixture observed in Hispanics and African Americans provides unique insight into allelic and locus heterogeneity. Smaller linkage disequilibrium (LD) blocks are observed when compared to ancestral Europeans, given their African ancestral component, and the greater degree of recombination observed due to ancestral lineage.[22]

Our objective is to examine Hispanics and African Americans for relevance and diagnostic accuracy of the 200 independent non-MHC variants,[2] spanning 156 loci of 2-Megabases (Mb) each, which are associated with MS risk in European populations.

Materials and methods

Participants and Genotyping

2995 self-reported Hispanics (1558 with MS, 1437 controls) and 2630 self-reported African Americans (1427 with MS, 1203 controls) were ascertained from seven US participating institutions (Appendix) as part of the Alliance for Research in Hispanic MS (ARHMS) (www.arhms.org). A further 464 European samples from the Centre d'Esclerosi Múltiple de Catalunya (Cemcat) in Barcelona, Spain (232 with MS and 232 controls) were provided. The Institutional Review Boards at each institution approved this study, and all participants provided written informed consent prior to participation. SNP genotyping was conducted using the MS Chip, an Illumina Infinium custom genotyping array. Participants and variants were excluded using standard criteria (Appendix). In total, 2784 Hispanics (1398 with MS, 1386 controls), 2460 African Americans (1305 with MS, 1155 controls), and 406 Spaniards (198 with MS, 208 controls) remained.

Global ancestry computation

To assess global ancestry, we used ADMIXTURE[23] and reference data from native populations of the Americas within the Human Genome Diversity Project (HGDP)[24] and from Europeans and Africans within 1000 Genomes[25]. We additionally removed individuals with ≥99.9% ancestry from any one reference population, resulting in 2652 Hispanics (1298 with MS, 1354 controls) and 2435 African Americans (1298 with MS, 1137 controls) (Appendix, Table 1, S1 Table).

Association analysis: MS risk

Association between MS status and the 200 MS single nucleotide polymorphisms (SNPs) was assessed using logistic regression in the Hispanic and African American study samples separately, adjusting for global European and Native American ancestry to control for differences within and across ascertainment sites. Sensitivity analyses were conducted to further assess the effect of population stratification, by removal of individuals with ancestral extremes (Appendix). An inverse-variance meta-analysis of the Hispanic and African American study samples, under a random effects model, was performed using PLINK. In the homogeneous Spanish sample, we adjusted for the first five principal components (Appendix). SNP was modeled additively as 0, 1, or 2 copies of the risk allele. Replication was defined as marginal one-sided $p \leq 0.05$. Correlation of the observed risk allele frequency difference with Europeans and replication status (0 = no replication, 1 = replication) was examined with Pearson correlation coefficients using R v3.0.2. Quanto v1.2.4 was used to determine if observations were consistent with power (Appendix).

To determine if marginal associations within the same 2-Mb locus were independent, as they were in Europeans,[2] forward stepwise conditional logistic regression was performed. At each of the 156 loci, the most significant SNP having a one-sided $p \leq 0.05$ was identified and included as a covariate. The association analysis was then repeated for the remaining SNPs within the locus, until no more SNPs were added.

Networks and pathways

SNPs were mapped to genes using three parallel approaches (Appendix). StringDB[26] was used to construct a network of the variants which replicated in both study populations. We allowed the addition of up to 10 connecting genes (Appendix) in order to build the network. We additionally used StringDB to compute the enrichment of Gene Ontology (GO) Biological processes among the gene sets represented by all 200 variants as compared to those which replicated, on a background of all genes in the human genome.

Risk score analysis

Cumulative genetic risk scores were computed in three ways for each study sample in addition to 1000 Genomes populations (Appendix). We further computed the genetic risk scores for 1811 European individuals with MS and 477 control samples available from UCSF which were included in the original European analysis performed by the International Multiple Sclerosis Genetics Consortium (IMSGC).[2]

The first method utilized a weighted sum of previously published risk alleles from the 15 variants which replicated in both admixed study samples and demonstrated *homogeneity* of effect (heterogeneity $p > 0.05$) across populations. Weights were extracted as effect sizes from the published European study.[2] The remaining two methods incorporated all 194 independent variants (Appendix); first as a weighted sum and then as an unweighted sum of risk alleles. Scores were compared utilizing two-sample t-tests with R v3.0.2. Receiver Operating Characteristic (ROC) curves and area under the curve (AUC) were generated to assess utility of the risk score for MS prediction.

To further assess the inherent disease risk within our study samples, we used logistic regression models, adjusting for global ancestry, to compute the odds of disease for individuals in the 0-5th, 6-10th, 11-25th, 75-89th, 90-94th, and 95-100th percentile of each of the two distributions compared to individuals within the interquartile range (IQR).

Results

Distribution of ancestry

On average, our Hispanics are 74% European, 15% Native American, and 11% African; while African Americans are 20% European, 2% Native American, and 78% African. Similar distributions are seen between individuals with MS and controls (Fig 1). Site-specific differences are observed along a geographical cline for the Hispanic study sample, in part due to waves of immigration and forced relocations of native populations.[27] In contrast, site-specific differences are minimal for the African Americans (S1 Table, S1–S4 Figs).

SNP associations: MS risk

152 of 200 (76%) and 136 of 200 (68%) SNPs show directional consistency with the European risk allele in Hispanics (one-sided binomial $p = 4.37 \times 10^{-14}$) and African Americans ($p = 1.95 \times 10^{-07}$) respectively. In Hispanics with ≥80% European ancestry (462 with MS and 736 controls), directional consistency increases to 79%. African Americans show a greater risk allele frequency difference with Europeans (mean absolute frequency difference = 0.14, SD = 0.11) than Hispanics (mean absolute frequency difference = 0.04, SD = 0.03).

In total, 16 of 200 SNPs show marginal replication (one-sided $p \leq 0.05$) in both the Hispanic and African American study sample, 57 in only the Hispanic study sample, and 28 in only the African American study sample (S2 Table). Beyond adjustment for global ancestry, population stratification did not substantially affect our association results (S9 Table, Appendix). While we do not see correlation between replication status and absolute frequency difference in African Americans ($r = -0.07$, $p = 2.94 \times 10^{-01}$) or Hispanics ($r = 0.07$, $p = 3.14 \times 10^{-01}$); we do see correlation with relative frequency difference in both African Americans ($r = 0.16$, $p = 2.40 \times 10^{-02}$) and Hispanics ($r = 0.20$, $p = 4.84 \times 10^{-03}$); where more replication is seen when the European risk allele frequency is greater than the study population frequency than vice versa (S5 and S6 Figs).

We see no statistically significant evidence for heterogeneity of effect size after correction for multiple testing (Bonferroni threshold = 2.50×10^{-04}). Nominal evidence ($p \leq 0.05$, S2 Table) is observed at one variant that replicates in both study samples: intronic rs4545915 within *MALT1 paracaspase (MALT1)* at 18q21.32: published[2] European OR = 1.09, Hispanic OR = 1.29, and African American OR = 1.18. Nominal evidence for heterogeneity of effect size is additionally seen for five variants which replicate only in Hispanics (S2 Table) and for one variant which replicates only in African Americans (intergenic rs11740512 at 5p13.1: published European OR = 1.15, African American OR = 1.29). rs11740512 was mapped to *PTGER4* through regulatory networks. Five SNPs show

significance (two-sided $p \leq 0.10$), but with opposite direction of effect than Europeans. No association (one-sided $p > 0.05$) is indicated for the remaining 94 SNPs; although seven of the 94 reach marginal significance (one-sided $p \leq 0.05$) in the meta-analysis (S2 Table).

The average power across the 200 variants is 42% in Hispanics (46% in replicating, 39% in non-replicating variants), 36% in African Americans (39% in replicating, 35% in non-replicating variants), and 14% in Spaniards (15% in replicating, 14% in non-replicating variants). We observe fewer independent marginal associations than would be expected based on power in African Americans (observe 41, expect 69 with 95% CI: 57-82). In Hispanics and Spaniards, we respectively observe fewer (observe 70, expect 80 with 95% CI: 67-93) and greater (observe 32, expect 28 with 95% CI: 19-38) independent marginal associations than expected; however our observations fall within the 95% CI for expected associations (S2 Table, S4 Table, Appendix).

Independence of associations within loci

64 of 156 loci (42%) in Hispanics and 39 of 156 loci (25%) in African Americans have at least one variant which replicates with one-sided $p \leq 0.05$; with five (S5 Table) and two (S6 Table) of the loci which indicated multiple independent effects in Europeans[2] also indicating multiple statistically independent effects following conditional modeling. These results provide evidence for within population allelic heterogeneity.

There are four independent effects at the 1p22.1 locus in Europeans;[2] however, we see two of those effects (rs9887787 and rs58394161) in Hispanics and one (rs58394161) in African Americans. The primary European effect (rs11577426) shows no association in either admixed study sample, although this could be due to limited power for detection (S2 Table).

Pathway analysis

Among the genes mapped to variants (S7 Table) which replicate in both admixed study samples, pathway analysis highlights the role of ‘regulation of T cell differentiation’ (False Discovery Rate (FDR) $p = 1.69 \times 10^{-07}$) and ‘positive regulation of cytokine production’ (FDR $p = 2.39 \times 10^{-06}$) (Fig 2, S8 Table).

When considering the pathways enriched among genes mapping to all 200 variants; ‘cell activation’, ‘positive regulation of RNA metabolic process’, ‘cellular response to cytokine stimulus’, ‘response to cytokine’, ‘positive regulation of transcription’, and ‘positive regulation of RNA biosynthetic process’ all show FDR $p < 1.00 \times 10^{-07}$ (S8 Table). All but ‘positive regulation of RNA metabolic process’ are also enriched ($p < 0.05$) among genes mapping to variants which replicate in either the Hispanic or African American study sample. Several other pathways show enrichment (FDR $p < 1.00 \times 10^{-05}$) among genes mapping to all 200 variants but not the replicating variant subset (FDR $p > 0.05$); including ‘positive regulation of nucleobase-containing compound metabolic process’, ‘multi-organism process’, positive regulation of lymphocyte activation’, ‘positive regulation of cellular metabolic process’, ‘regulation of homotypic cell-cell adhesion’, ‘positive regulation of macromolecule metabolic process’, ‘regulation of response to stimulus’, and ‘leukocyte cell-cell adhesion’. Yet, we find no difference in power to detect association for variants within and outside of each of these pathways (two-sample t-test $p > 0.05$ for each pathway).

Risk score results

All three cumulative genetic risk scores show association with MS disease status in all study samples; with significance of association boosted using the 194-variant scores over the 15-variant scores in all but African Americans. For the 15-variant score, a cline is seen in the geographical distribution of the risk score with 1000G Africans (AFR) demonstrating the lowest scores, followed by Asians (ASIA), Americans (AMR), and Europeans (EUR) with the highest (Fig 3A). With the 194-variant unweighted (Fig 3B) and weighted risk scores (Fig 3C), 1000 Genomes AFR now demonstrate the highest scores, followed by populations of EUR, AMR, and ASIA with the lowest.

None of the risk scores provide perfect diagnostic MS capability (AUC = 1) (Figs 4A–4C). The AUC of the European subset[2] and Hispanics increases from the 15-variant to the 194-variant risk scores, signifying better predicting capability when using all 194 variants. However, the African American AUC decreases or remains unchanged from the 15-variant to the 194-variant risk scores, signifying no improvement in predictive capability when using all 194 variants.

Hispanics and African Americans in the 95th percentile of the 15-variant weighted risk score distribution have odds of 2.17 (95% CI: 1.48-3.19) and 2.14 (95% CI: 1.42-3.22) respectively of developing MS as compared to individuals in the IQR. The odds of developing MS for those in the 95th percentile increases using the 194-variant risk scores in Hispanics (2.69 unweighted, 2.67 weighted) but decreases in African Americans (1.61 unweighted, 1.78 weighted). These data provide further evidence for reduction in diagnostic capability in African Americans, specifically when utilizing all 194 variants (Table 2).

Discussion

MS genetic discoveries in populations of European ancestry now include 200 autosomal non-MHC variants, 32 variants within the extended MHC, and one X-chromosome variant.[2] Our primary objective was to test the relevance of the 200 MS non-MHC risk variants within US minority populations of Hispanics and African Americans. While both admixed study samples suggest an over-representation of European risk alleles among individuals with MS, we in fact see less replication than would be anticipated within African Americans. Since the 200 analyzed SNPs are unlikely to be causal and instead tag the underlying variation in European populations, the smaller LD blocks observed in African Americans could result in less correlation between the analyzed and causal variants, inherently reducing power for detection. Alternatively, limited replication could indicate that not all MS risk SNPs discovered in European populations are relevant, possibly due to locus heterogeneity or gene-environment interactions. Further, the effect sizes which are relevant for MS may be lower than those observed in Europeans for a subset of variants. This could be due to allelic heterogeneity or differing LD structure. The ‘Winner’s Curse’ could also explain our observations; by which the observed European effect sizes are inflated due to thresholding, or only reporting associations meeting a specified statistical threshold. This can result in preferential selection of effects which may be overestimated due to noise. Nonetheless, we expect this to play a minimal role due to the robustness of the multi-stage European study design[2] which resulted in narrow confidence intervals for genome-wide associations.

We see nominal evidence heterogeneity of effect size for several replicating variants, warranting further investigation in larger samples. These include an intronic SNP in *MALT1*. *MALT1* is a component of the *CARMA1-BCL10-MALT1* signalosome and encodes a caspase-like protease that influences *BCL10*-induced activation of NF-kappaB, which is essential for lymphocyte activation. Pharmacological studies have indicated that inhibition of *MALT1* protease activity attenuates experimental autoimmune encephalomyelitis in mice,[28] and has been suggested as a rational drug target for immunomodulation.[29] *MALT1* contains VDR-binding peaks and VDRE motifs for inflammatory dendritic cells,[30] suggesting that vitamin D exposure and absorption, likely to be influenced by race/ethnicity,[31] may influence expression.[32] *PTGER4*, showing heterogeneity between Europeans and African Americans, is a receptor for *prostaglandin E2 (PGE2)*. *PGE2* production may be regulated in part by vitamin D3 metabolites,[33] and vitamin D absorption may be decreased among individuals with higher skin pigmentation.[31] These results are compatible with reports that levels of 25-hydroxyvitamin D are positively correlated with European ancestry and are lower in African Americans with MS than controls.[34] These heterogeneous effects, while nominal, illustrate potential interactions between gene and environment which may occur to influence the effect of a variant on MS risk.

We note several pathways which appear to be enriched for MS in more than one population (i.e. Europeans and at least one of our admixed populations) and several which are only enriched in Europeans. For instance, the pathway involving positive regulation of RNA metabolic process was highly enriched among genes which mapped to all 200 variants, but did not show even nominal enrichment among genes which mapped to variants replicating in at least one population. This could indicate that this pathway is most relevant for the etiology of MS in populations of European ancestry. This phenomenon of a race/ethnic effect on assorted metabolic processes has been well studied in pharmacokinetics.[35] We present here evidence that further study of the influence of metabolic processes on disease risk and progression, inclusive of MS, may be warranted and provide insight which may be relevant for treatment. While these findings should be followed-up in larger samples; gene-environment or gene-gene interactions may exist which alter disease phenotype differentially across race and ethnic groups.

We find that cumulative genetic risk scores are associated with MS disease status in all study samples. The 15-variant risk score shows limited diagnostic accuracy and limited variability across populations (AUC from 0.56 to 0.60). Accuracy improves minimally for the 194-variant risk scores in Europeans (AUC = 0.70) and Hispanics (AUC = 0.67) but remains relatively unchanged for African Americans; providing additional evidence that not all variants associated with MS risk in European populations may be relevant in this US minority population. However, this could again be due to limited LD between the analyzed and causal variant in our African American population; where the true causal variant may be relevant. Both affected and unaffected African Americans have a dramatic increase in score from 15 to 194 variants, even above that of Europeans with MS; illustrating an inherent bias. Theoretical risk may increase due to increased ‘risk’ allele frequency, while no further knowledge is gained regarding disease status. The similarities observed between the 194-

unweighted and weighted risk scores indicates that while both methods present some level of bias (assuming equal or imprecise weights), the results remain robust.

While diagnostic accuracy is currently limited in all study populations, these data provide a wealth of knowledge that can be used to efficiently and accurately fine-map the published MS loci. For variants which replicate in both admixed study samples, the true causal variant is most likely tagged by the associated variant in Europeans, Hispanics, and African Americans. Likewise, for variants which replicate in only one admixed study sample, the causal variant would only be tagged by the associated variant in Europeans and the replicating population. While we do not currently have evidence to suggest association in the non-replicating population; we acknowledge that if the lack of association is due to limited LD between the analyzed and causal variant, fine-mapping in this population may be ideal. It may result in a smaller credible interval, meaning that the genetic distance in which the causal variant is contained with high probability may be smaller in this population than in another. For variants which show significant yet opposite direction of effects in an admixed study sample, we must consider causal variants that are in LD in either Europeans or the relevant minority population, given this evidence for variation in LD structure. For instance, we observe a protective effect in Hispanics of the European risk allele for rs767455 in *TNFRSF1A* (a locus was previously fine-mapped to rs1800693, with a posterior probability of 0.69 in Europeans).[5]

These findings represent the most comprehensive study of established common genetic risk alleles for MS in Hispanics and African Americans. Novel insight is provided into the relevance of previously published MS risk variants for US minority populations. Overall, these findings provide a framework for future fine-mapping efforts in multi-ethnic populations of MS. Our results are limited by a lack of genome-wide array data which would allow for incorporation of local ancestry into all analyses. The lack of fine-mapping content also limits our ability to translate association into causality, although future efforts will be focused on this work. In summary, this work highlights the importance of diversity in genomic studies to both to uncover effective therapies for all individuals burdened with disease and provide a path towards prevention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Andrea Martinez (USC), Stacy Caillier (UCSF), Dr. Bertha Fonseca and Dr. Carlos Ramirez-Calderon (Miami), Dr. Patricia De Jesus (PR), Mary Davis (VU), and Dr. Dorlan Kimbrough (JHU) for their contributions to data collection/recruitment. We would also like to thank the Fundación de Esclerosis Múltiple de Puerto Rico and Ms. Lourdes Fernández Trujillo for considerable contributions to participant recruitment across Puerto Rico. We additionally thank the individuals from the diverse MS registries across sites for their participation in the study.

We acknowledge the Center for Genome Technology within the University of Miami John P. Hussman Institute for Human Genomics for generating the MS Chip genotype data for this project and the IMSGC for design of the MS Chip array. We additionally acknowledge the Population Architecture using Genomics and Epidemiology (PAGE) Study, and in particular Genevieve Wojcik and Christopher Gignoux, for early access to MEGA array data on Native American samples from HGDP.

Funding

This work was supported by the National Institutes of Health [NCATS KL2TR000131 to L.A., NINDS 1R01NS096212 to J.L.M, RO1NS26799 to J.R.O., R01HL102487 to G.W.B.]; and the National Multiple Sclerosis Society [RG 4680A1/1 to J.L.M., RG 1611-26299 to J.R.O].

References

1. Iwanowski P and Losy J. Immunological differences between classical phenothypes of multiple sclerosis. J Neurol Sci 2015; 349: 10–14. [PubMed: 25586536]
2. International Multiple Sclerosis Genetics Consortium. The Multiple Sclerosis Genomic Map: Role of peripheral immune cells and resident microglia in susceptibility. bioRxiv 2017.
3. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011; 476: 214–219. [PubMed: 21833088]
4. Patsopoulos NA, the Bayer Pharma MS Genetics Working Group, the Steering Committees of Studies Evaluating IFNbeta-1b and a CCR1-Antagonist, ANZgene Consortium, GeneMSA, International Multiple Sclerosis Genetics Consortium and de Bakker PI. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol 2011; 70: 897–912. [PubMed: 22190364]
5. International Multiple Sclerosis Genetics Consortium (IMSGC), Beecham AH, Patsopoulos NA, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013; 45: 1353–1360. [PubMed: 24076602]
6. Isobe N, Madireddy L, Khankhanian P, et al. An ImmunoChip study of multiple sclerosis risk in African Americans. Brain 2015; 138: 1518–1530. [PubMed: 25818868]
7. U.S. Department of Commerce. Overview of Race and Hispanic Origin: 2010. U S Census Bureau - 2010 Census Briefs 2011: 1–24.
8. Alves-Leon SV, Papais-Alvarenga R, Magalhaes M, et al. Ethnicity-dependent association of HLA DRB1-DQA1-DQB1 alleles in Brazilian multiple sclerosis patients. Acta Neurol Scand 2007; 115: 306–311. [PubMed: 17489940]
9. Barquera R, Zuniga J, Hernandez-Diaz R, et al. HLA class I and class II haplotypes in admixed families from several regions of Mexico. Mol Immunol 2008; 45: 1171–1178. [PubMed: 17904223]
10. Rojas OL, Rojas-Villarraga A, Cruz-Tapias P, et al. HLA class II polymorphism in Latin American patients with multiple sclerosis. Autoimmun Rev 2010; 9: 407–413. [PubMed: 19896562]
11. Bentley AR, Callier S and Rotimi CN. Diversity and inclusion in genomic research: why the uneven progress?. J Community Genet 2017; 8: 255–266. [PubMed: 28770442]
12. Bustamante CD, Burchard EG and De la Vega FM. Genomics for the world. Nature 2011; 475: 163–165. [PubMed: 21753830]
13. Khan O, Williams MJ, Amezcua L, et al. Multiple sclerosis in US minority populations: Clinical practice insights. Neurol Clin Pract 2015; 5: 132–142. [PubMed: 26137421]
14. Langer-Gould A, Brara SM, Beaber BE, et al. Incidence of multiple sclerosis in multiple racial and ethnic groups. Neurology 2013; 80: 1734–1739. [PubMed: 23650231]
15. Wallin MT, Culpepper WJ, Coffman P, et al. The Gulf War era multiple sclerosis cohort: age and incidence rates by race, sex and service. Brain 2012; 135: 1778–1785. [PubMed: 22628389]
16. Cristiano E and Rojas JI. Multiple sclerosis epidemiology in Latin America: An updated survey. Mult Scler J Exp Transl Clin 2017; 3: 2055217317715050. [PubMed: 28638628]
17. Chinea A, Rios-Bedoya CF, Vicente I, et al. Increasing Incidence and Prevalence of Multiple Sclerosis in Puerto Rico (2013–2016). Neuroepidemiology 2017; 49: 106–112. [PubMed: 29136613]
18. Buchanan RJ, Zuniga MA, Carrillo-Zuniga G, et al. Comparisons of Latinos, African Americans, and Caucasians with multiple sclerosis. Ethn Dis 2010; 20: 451–457. [PubMed: 21305836]
19. Amezcua L, Oksenberg JR and McCauley JL. MS in self-identified Hispanic/Latino individuals living in the US. Mult Scler J Exp Transl Clin 2017; 3: 2055217317725103. [PubMed: 28979795]

20. Hadjixenofontos A, Beecham AH, Manrique CP, et al. Clinical Expression of Multiple Sclerosis in Hispanic Whites of Primarily Caribbean Ancestry. Neuroepidemiology 2015; 44: 262–268. [PubMed: 26184860]
21. Amezcua L, Beecham AH, Delgado SR, et al. Native ancestry is associated with optic neuritis and age of onset in hispanics with multiple sclerosis. Ann Clin Transl Neurol 2018; 5: 1362–1371. [PubMed: 30480030]
22. Shifman S, Kuypers J, Kokoris M, et al. Linkage disequilibrium patterns of the human genome across populations. Hum Mol Genet 2003; 12: 771–776. [PubMed: 12651872]
23. Alexander DH, Novembre J and Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009; 19: 1655–1664. [PubMed: 19648217]
24. Cavalli-Sforza LL. The Human Genome Diversity Project: past, present and future. Nat Rev Genet 2005; 6: 333–340. [PubMed: 15803201]
25. 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012; 491: 56–65. [PubMed: 23128226]
26. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447–52. [PubMed: 25352553]
27. Bryc K, Durand EY, Macpherson JM, et al. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am J Hum Genet 2015; 96: 37–53. [PubMed: 25529636]
28. Mc Guire C, Elton L, Wieghofer P, et al. Pharmacological inhibition of MALT1 protease activity protects mice in a mouse model of multiple sclerosis. J Neuroinflammation 2014; 11: 124-2094-11-124.
29. Jaworski M and Thome M. The paracaspase MALT1: biological function and potential for therapeutic inhibition. Cell Mol Life Sci 2016; 73: 459–473. [PubMed: 26507244]
30. Booth DR, Ding N, Parnell GP, et al. Cistromic and genetic evidence that the vitamin D receptor mediates susceptibility to latitude-dependent autoimmune diseases. Genes Immun 2016; 17: 213–219. [PubMed: 26986782]
31. Yao S, Zirpoli G, Bovbjerg DH, et al. Variants in the vitamin D pathway, serum levels of vitamin D, and estrogen receptor negative breast cancer among African-American women: a case-control study. Breast Cancer Res 2012; 14: R58. [PubMed: 22480149]
32. Nolan D, Castley A, Tschochner M, et al. Contributions of vitamin D response elements and HLA promoters to multiple sclerosis risk. Neurology 2012; 79: 538–546. [PubMed: 22786591]
33. Schwartz Z, Swain LD, Kelly DW, et al. Regulation of prostaglandin E2 production by vitamin D metabolites in growth zone and resting zone chondrocyte cultures is dependent on cell maturation. Bone 1992; 13: 395–401. [PubMed: 1419381]
34. Gelfand JM, Cree BA, McElroy J, et al. Vitamin D in African Americans with multiple sclerosis. Neurology 2011; 76: 1824–1830. [PubMed: 21606454]
35. Johnson JA. Predictability of the effects of race or ethnicity on pharmacokinetics of drugs. Int J Clin Pharmacol Ther 2000; 38: 53–60. [PubMed: 10706191]

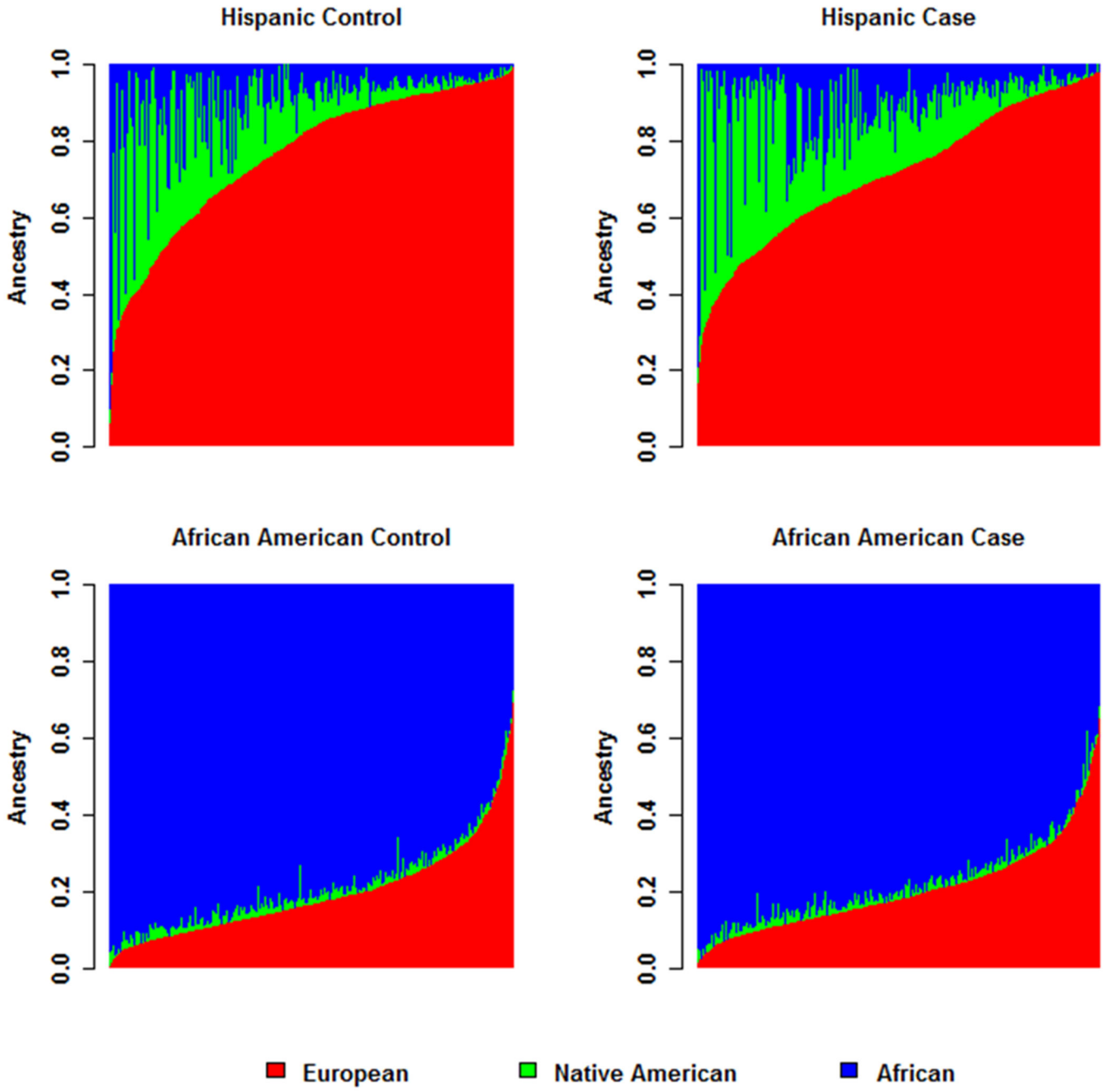

Fig 1. Global ancestry proportions for Hispanic and African American case-control samples.

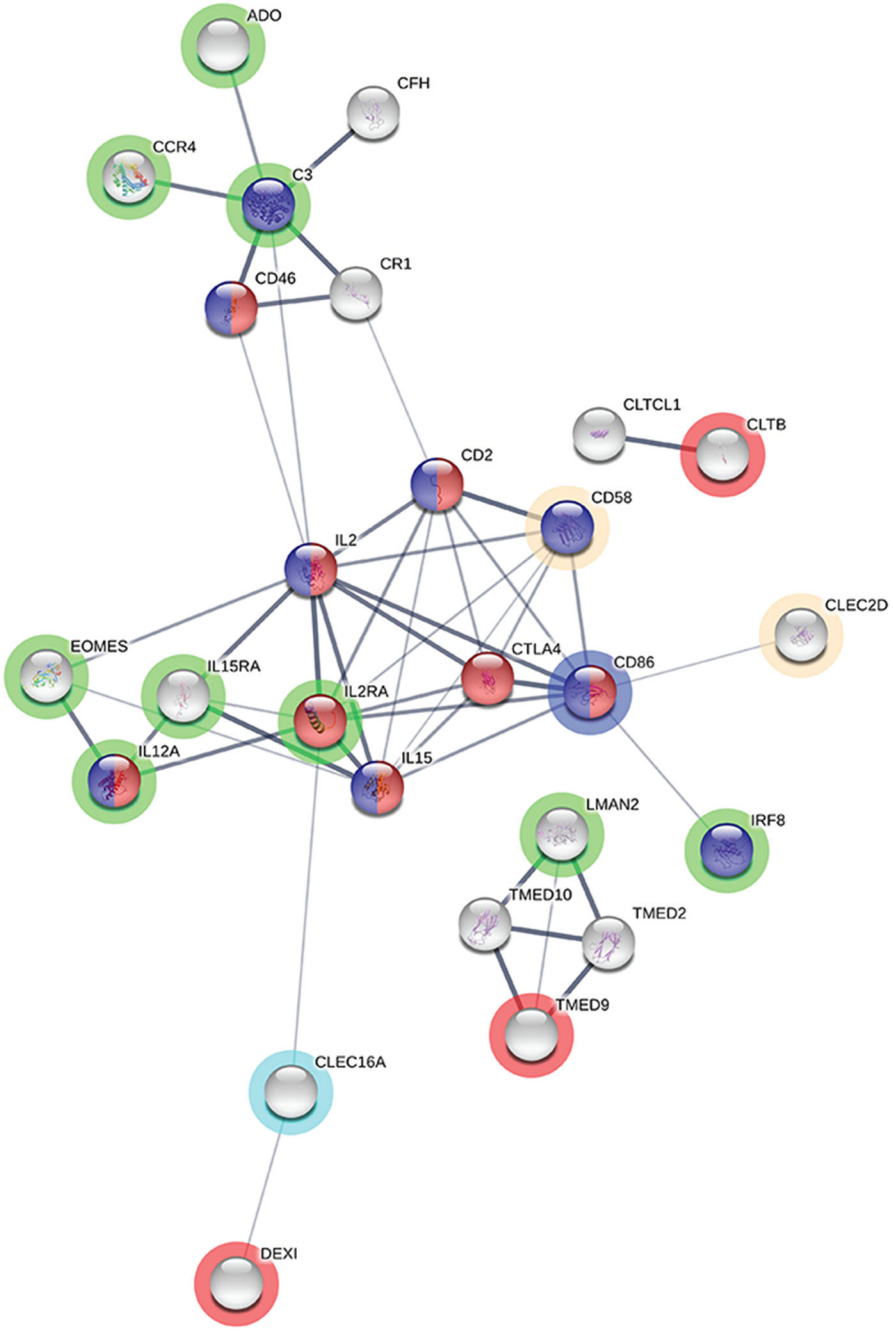

Fig 2. Gene-network for genes mapped to 16 variants which replicated in both Hispanics and African Americans.

Blue fill indicates genes which are part of the Positive Regulation of Cytokine Production pathway, and red fill indicates genes which are part of the Regulation of T Cell Differentiation pathway as defined by GO Biological processes. Border colors denote the mode of gene-mapping: green (regulatory), blue (exonic), aqua (regulatory+exonic), red (eQTL PBMC), tan (regulatory+eQTL PBMC). No border color indicates the gene was a link added by StringDB.

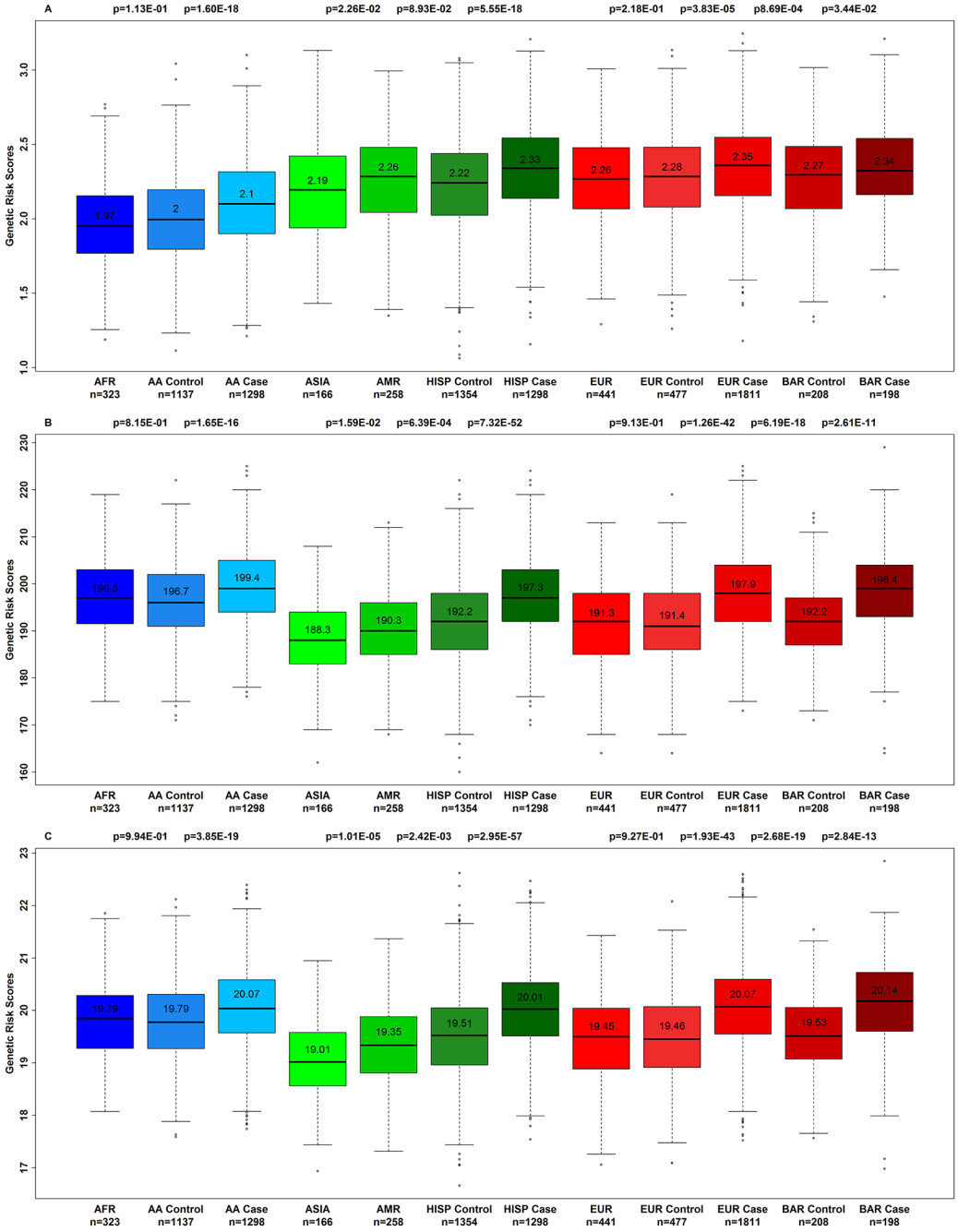

Fig 3. Distribution of the cumulative genetic risk score across populations.

(A) 15-variant weighted risk score, (B) 194-variant unweighted risk score, and (C) 194-variant weighted risk score. P-values indicate the results of a two-sample t-test between adjacent populations. AFR = 1000G Africans (ACB, ASW, LWK, and YRI), AA Control = African American controls, AA Case = African American MS cases, ASIA = 1000G Asians (CHB, CHD, and JPT), AMR = 1000G Americans (CLM, MXL, PEL, PUR), HISP Control = Hispanic controls, HISP Case = Hispanic MS cases, EUR = 1000G Europeans (CEU, FIN,

GBR, IBS, and TSI), EUR Control = UCSF European controls, EUR Case = UCSF European MS cases, BAR Control = Spanish controls, and BAR Case = Spanish MS cases.

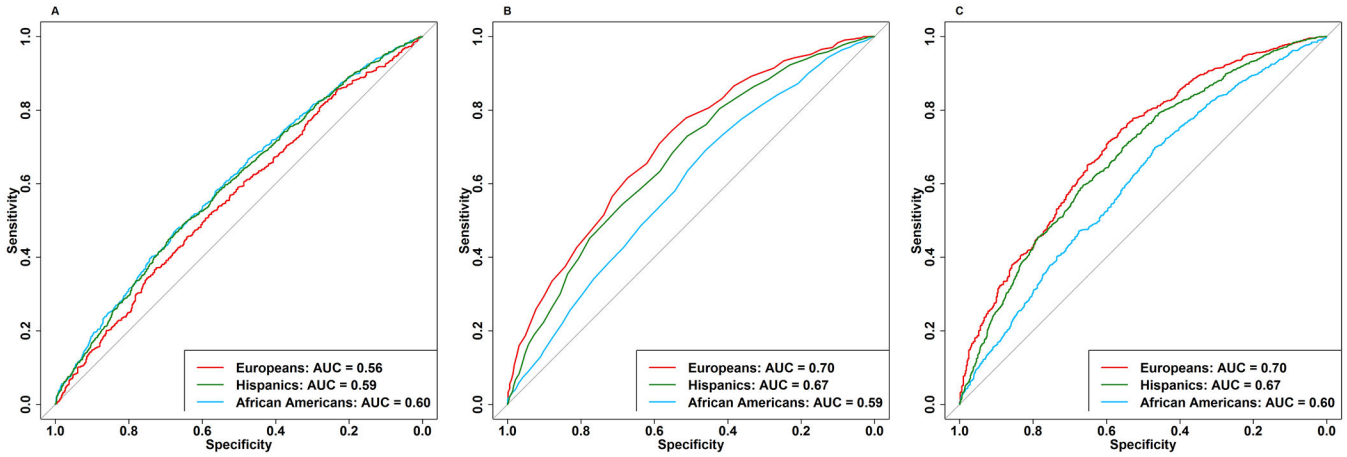

Fig 4. Receiver Operating Characteristic curves for the cumulative genetic risk score.

(A) 15-variant weighted risk score and (B) 194-variant unweighted risk score, and (C) 194-variant weighted risk score. Computed for each population with: Europeans = 1811 UCSF MS Cases and 477 UCSF Controls, Hispanics = 1298 MS Cases and 1354 Controls, and African Americans = 1298 MS Cases and 1137 Controls.

Table 1.

Sample distribution.

Site	Hispanic		African American	
	MS Case	Control	MS Case	Control
	N (%)		N (%)	
University of Miami (UM)	539 (42)	1131 (84)[a]	49 (4)	378 (33)
University of California, San Francisco (UCSF)	156 (12)	149 (11)[a]	1001 (77)	600 (53)
University of Southern California (USC)	195 (15)	0 (0)	0 (0)	0 (0)
Caribbean Neurological Center, Puerto Rico (PR)	316 (24)	6 (0)	0 (0)	0 (0)
Brigham and Women's Hospital (BWH) / Johns Hopkins (JHU)	92 (7)	68 (5)	169 (13)	5 (0)
Vanderbilt University (VU)	0 (0)	0 (0)	79 (6)	154 (14)
Total	1298	1354	1298	1137
Global Admixture	%		%	
European	71	76	20	19
African	11	11	78	79
Native American	18	13	2	2

[a]41 and 89 controls from UM and UCSF respectively are Puerto Rican in heritage, genotyped to provide an ancestral balance to the Puerto Rican cases ascertained from the Caribbean Neurological Center. All MS cases from UM were recruited through the MS Registry, while the majority (95%) of UM Hispanic controls were recruited from the Miami Cardiovascular Registry (MCR). African American controls from UM were recruited as controls from collections initiated for study of Alzheimer's disease.

Table 2.

MS association by risk score percentile.

Percentile	Hispanics		African Americans		Europeans[a]	
	OR (L95-U95)	P	OR (L95-U95)	P	OR (L95-U95)	P
15 Replicating Variants - Weighted						
0-5	0.43 (0.29-0.64)	2.72×10^{-05}	0.43 (0.29-0.64)	2.89×10^{-05}	0.68 (0.44-1.05)	8.13×10^{-02}
6-10	0.61 (0.42-0.89)	1.01×10^{-02}	0.62 (0.42-0.90)	1.25×10^{-02}	0.70 (0.45-1.09)	1.14×10^{-01}
11-25	0.74 (0.59-0.93)	1.10×10^{-02}	0.76 (0.60-0.97)	2.43×10^{-02}	0.73 (0.55-0.96)	2.57×10^{-02}
75-89	1.27 (1.01-1.60)	3.90×10^{-02}	1.45 (1.14-1.84)	2.65×10^{-03}	1.11 (0.82-1.52)	5.00×10^{-01}
90-94	1.38 (0.96-1.98)	8.52×10^{-02}	1.39 (0.95-2.04)	9.17×10^{-02}	1.25 (0.75-2.09)	3.92×10^{-01}
95-100	2.17 (1.48-3.19)	7.26×10^{-05}	2.14 (1.42-3.22)	2.80×10^{-04}	1.44 (0.85-2.46)	1.79×10^{-01}
194 Independent Variants - Unweighted						
0-5	0.25 (0.16-0.37)	8.27×10^{-11}	0.37 (0.25-0.53)	9.09×10^{-08}	0.19 (0.13-0.28)	$<2.00 \times 10^{-16}$
6-10	0.35 (0.23-0.53)	6.84×10^{-07}	0.58 (0.40-0.85)	4.64×10^{-03}	0.24 (0.16-0.37)	2.54×10^{-11}
11-25	0.50 (0.39-0.63)	2.64×10^{-09}	0.63 (0.50-0.79)	4.79×10^{-05}	0.47 (0.36-0.61)	1.20×10^{-08}
75-89	1.77 (1.40-2.25)	2.07×10^{-06}	1.36 (1.06-1.74)	1.59×10^{-02}	1.74 (1.18-2.55)	4.90×10^{-03}
90-94	3.10 (2.07-4.63)	3.42×10^{-08}	1.06 (0.73-1.54)	7.59×10^{-01}	3.61 (1.66-7.86)	1.23×10^{-03}
95-100	2.69 (1.85-3.92)	2.14×10^{-07}	1.61 (1.12-2.32)	1.09×10^{-02}	5.14 (2.07-12.7)	4.13×10^{-04}
194 Independent Variants - Weighted						
0-5	0.17 (0.11-0.29)	1.16×10^{-11}	0.39 (0.26-0.58)	3.08×10^{-06}	0.19 (0.13-0.28)	$<2.00 \times 10^{-16}$
6-10	0.40 (0.27-0.60)	8.40×10^{-06}	0.56 (0.38-0.82)	2.66×10^{-03}	0.28 (0.19-0.42)	6.77×10^{-10}
11-25	0.55 (0.43-0.69)	6.47×10^{-07}	0.56 (0.44-0.71)	1.62×10^{-06}	0.48 (0.36-0.63)	1.01×10^{-07}
75-89	2.08 (1.64-2.63)	1.49×10^{-09}	1.23 (0.97-1.56)	8.83×10^{-02}	1.84 (1.27-2.67)	1.37×10^{-03}
90-94	3.01 (2.01-4.49)	7.58×10^{-08}	1.51 (1.02-2.23)	3.81×10^{-02}	4.00 (1.73-9.23)	1.15×10^{-03}
95-100	2.67 (1.80-3.95)	9.73×10^{-07}	1.78 (1.20-2.66)	4.56×10^{-03}	4.89 (1.97-12.1)	6.19×10^{-04}

[a]Europeans analyzed are the 1811 European individuals with MS and 477 control samples from UCSF

OR = Odds Ratio, L95 = lower 95% confidence bound, U95 = upper 95% confidence bound