Admixture Mapping Reveals Diverse Genetic Ancestry of HLA and non-HLA MS-Associated Alleles in Admixed Populations

C. Chi[1,2], X. Shao[1], B. Rhead[1,2], E. Gonzales[3], J.B. Smith[3], A.H. Xiang[3], J. Graves[4], A. Waldman[5], T. Lotze[6], T. Schreiner[7], B. Weinstock-Guttman[8], G. Aaen[9], J.M. Tillema[10], J. Ness[11], M. Candee[12], L. Krupp[13], M. Gorman[14], L. Benson[14], T. Chitnis[15], S. Mar[16], A. Belman[17], C. Casper[12], J. Rose[12], M. Moodley[18], M. Rensel[18], M. Rodriguez[10], B. Greenberg[19], L. Kahn[22], J. Rubin[23], C. Schaefer[20], E. Waubant[4], A.M. Langer-Gould[3,21], L.F. Barcellos[1]

1) Genetic Epidemiology and Genomics Laboratory, UC Berkeley, Berkeley, CA; 2) Computational Biology Graduate Group, UC Berkeley, Berkeley, CA; 3) Kaiser Permanente Southern California, Department of Research & Evaluation, Los Angeles, CA; 4) UC San Francisco, Department of Neurology, San Francisco, CA; 5) Children's Hospital of Philadelphia, Philadelphia, PA; 6) Texas Children's Hospital, Houston, TX; 7) University of Colorado School of Medicine, Aurora, CO; 8) SUNY Buffalo, Buffalo, NY; 9) Loma Linda University, Loma Linda, CA; 10) Mayo Clinic, Rochester, MN; 11) University of Alabama at Birmingham, Birmingham, AL; 12) University of Utah, Salt Lake City, UT; 13) SUNY Stony Brook, Stony Brook, NY; 14) Boston Children's Hospital, Boston, MA; 15) Brigham and Women's Hospital, Boston, MA; 16) Washington University St. Louis, St. Louis, MO; 17) NYU Langone Medical Center, New York, NY; 18) Cleveland Clinic, Cleveland, OH; 19) University of Texas Southwestern, Dallas, TX; 20) Kaiser Permanente Division of Research, Oakland, CA; 21) Southern California Permanente Medical Group, Los Angeles Medical Center, Neurology Department, Los Angeles, CA; 22) Children's National Medical Center, NW Washington, DC; 23) Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

Introduction

- Multiple Sclerosis (MS [MIM 126200]) is a complex demyelinating disease of the central nervous system.
- The prevalence of MS is highest in Europeans compared to African Americans, Hispanics, and Asian Americans.
- Established MS-associated alleles in Europeans:
  - 11 human leucocyte antigen (HLA) alleles[1].
  - 200 non-HLA alleles[2].
- In an admixture mapping study, local ancestry estimation can be used to characterize admixture of MS-associated alleles to test the hypothesis that MS risk alleles are European.
- Objectives:
  1. Characterize global ancestry of cases and controls in African Americans, Hispanics, and Asian Americans.
  2. Characterize local ancestry of established non-HLA MS alleles and HLA alleles associated with MS in African Americans, Hispanics, and Asian Americans.
  3. Search for genome-wide signals of European ancestry-MS association in African Americans, Hispanics, and Asians Americans.

Subjects

Number of Cases and Controls by Admixed Population

Ethnicity	Cases	Controls	Total
African American	1,081	2,611	3,692
Hispanic	326	3,451	3,777
Asian	64	4,851	4,915
Total	1,471	10,913	12,384

Number of MS cases and controls for African Americans, Hispanics, and Asians Americans, after excluding related individuals, and excluding European subjects using multi-dimensional scaling (MDS) and fastStructure.

Methods

Genotyping and Imputation

1. Genome-wide SNP genotypes assessed with the Human660W-Quad BeadChip, Infinium Human OmniExpress BeadChip, Infinium Human OmniExpress Exome BeadChip, and the Immunochip.
2. Genome-wide genotype imputation against 1,000 Genomes haplotypes using IMPUTE2.
3. HLA allele imputation using SNP2HLA with reference HLA haplotypes derived from additional sequencing of 1,000 Genomes samples[3].

Quality Control

1. Removed SNPs for minor allele frequency (< 0.01) and missingness of SNPs and samples (< 0.10).
2. Removed related individuals ($\hat{\pi} > 0.25$).
3. Removed SNPs for info score (< 0.30) and call rate (< 0.60) after genome-wide imputation.
4. Removed A/T and C/G SNPs prior to local ancestry inference.
5. Removed HLA alleles for imputation quality ($R^2 < 0.80$) and with allele frequency less than 0.005 after HLA allele imputation.

Analysis of Population Structure: Used MDS and fastStructure to identify Europeans for exclusion from the study and characterize global ancestry in admixed populations[4].

Local Ancestry Inference: Inferred local ancestry across the genome using RFMix, with 1,000 Genomes reference panel tailored to the target population[5].

Statistical Analysis

1. Significance of ancestry-MS association evaluated with non-parametric admixture mapping test statistic developed by Montana and Pritchard (2004)[6].

$$T(l,\ k) = \frac{(\bar{z}_{l,d}(k) - \bar{z}_{l,c}(k)) - (\bar{q}_d(k) - \bar{q}_c(k))}{SD(\bar{z}_{l,d}(k) - \bar{z}_{l,c}(k))}$$

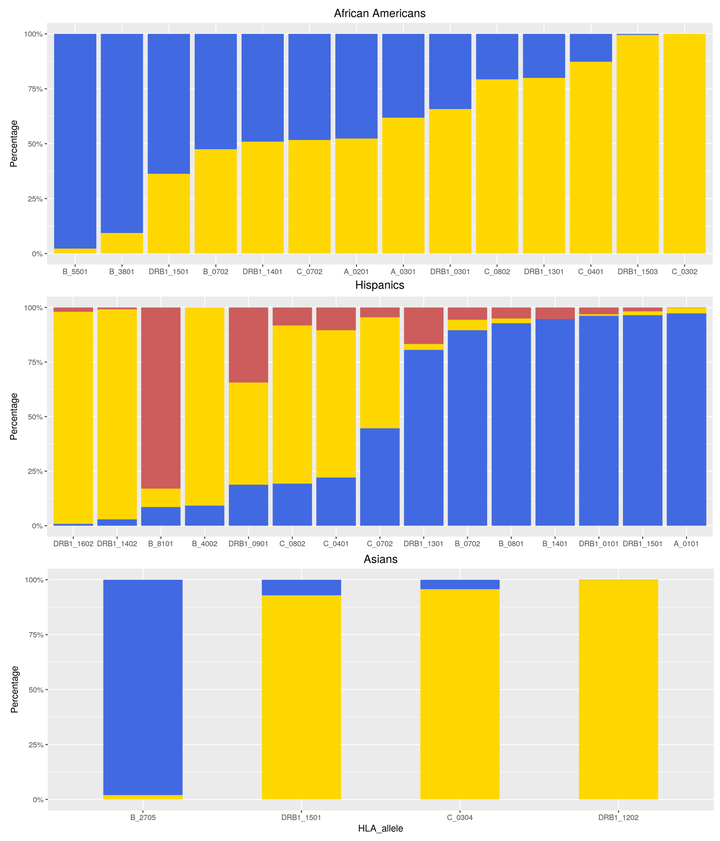
2. Significance of allele-MS association evaluated with multivariable logistic regression adjusting for top 3 MDS principal components.
3. Chi-square test of independence used to test for association between MS and European vs African *HLA-DRB1*15:01.*

References

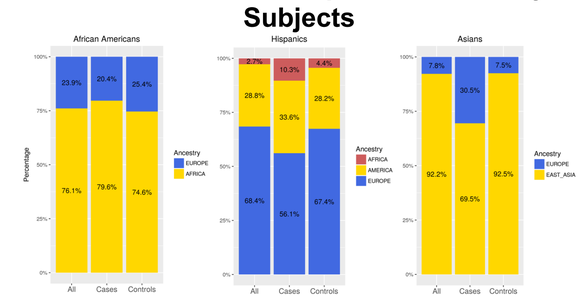
1. Patsopoulos, *et al.* (2013). Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 9, e1003926.
2. IMSGC *Science* in press.
3. Jia, X., *et al.* (2013). Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. PLoS One 8, e64683.
4. Raj, A., *et al.* (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 197, 573–589.
5. Maples, *et al.* (2013). RFMix: A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference. Am. J. Hum. Genet. 93, 278–288.
6. Montana, G., and Pritchard, J.K. (2004). Statistical Tests for Admixture Mapping with Case-Control and Cases-Only Data. Am. J. Hum. Genet. 75, 771–789.

Results

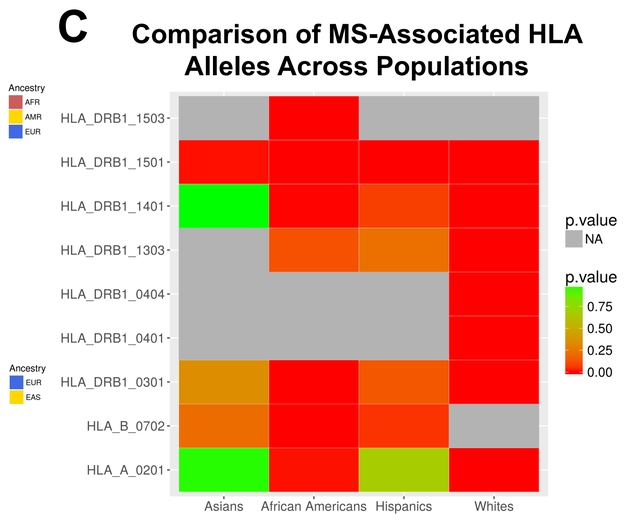
A



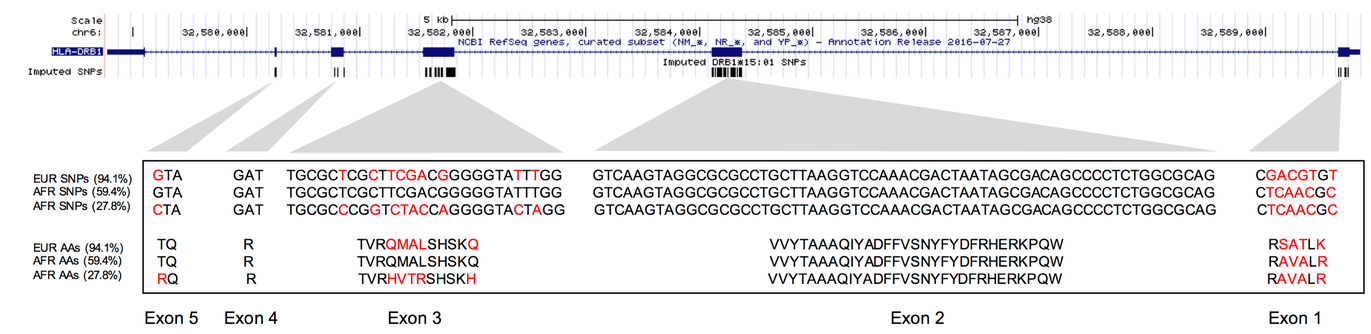
B



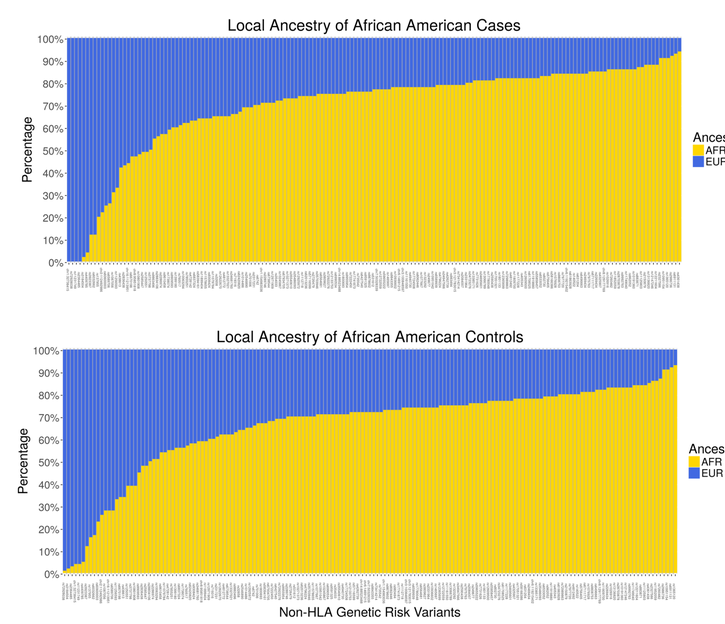
C



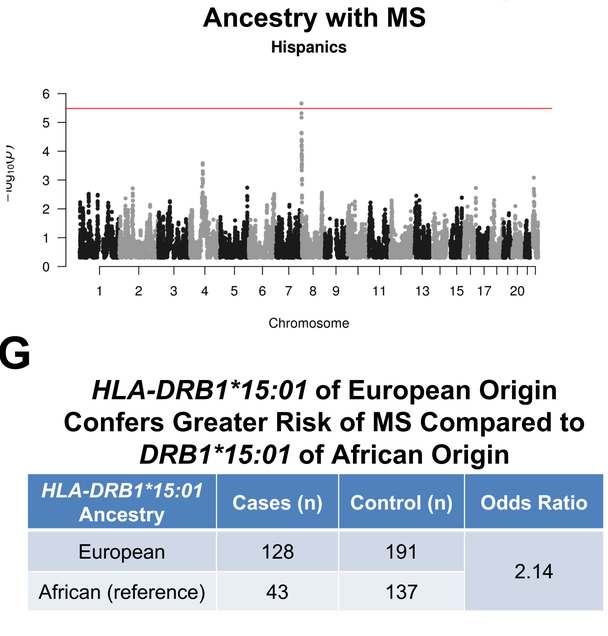
D



E



F



G

***HLA-DRB1*15:01* of European Origin Confers Greater Risk of MS Compared to *DRB1*15:01* of African Origin**

*HLA-DRB1*15:01* Ancestry	Cases (n)	Control (n)	Odds Ratio
European	128	191	2.14
African (reference)	43	137	

Chi *et al.* under review; ancestry proportions available upon request

Conclusions

- Most HLA alleles associated with MS ($p < 0.05$) are ancestrally diverse (Figure A).
- *HLA-DRB1*15:01* is consistently associated with MS across all populations (Figure B).
- European *HLA-DRB1*15:01* confers twice the risk of MS compared to *HLA-DRB1*15:01* of African origin (OR = 2.14, 95% CI: 1.42-3.22) (Table G).
- There is evidence that European and African *HLA-DRB1*15:01* allele sequences differ at exons 1, 3, and 5, but not at exon 2 (Figure D).
- Non-HLA MS risk variants do not have increased European ancestry in cases compared to controls in all populations, controlling for global ancestry (Figure E; Hispanic and Asian American results not shown).
- Whole-genome admixture mapping revealed significant evidence for increased European ancestry in Hispanic MS cases compared to controls on chromosome 8 (GRCh37 chr8:207,207-314,620) (Figure F).

This material is based upon work supported by NIH grants R01AI076544, R01NS049510, and R01ES017080, and the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1106400. Genotyping of the GERA cohort was supported by a grant (RC2 AG036607; C.S. and Neil Risch, PIs). Development of the Kaiser Permanente Research Program on Genes, Environment, and Health was supported by grants from the Wayne and Gladys Valley Foundation, the Ellison Medical Foundation, the Robert Wood Johnson Foundation, and the Kaiser Permanente Community Benefit Program.