

Generalization of GWAS Findings to Non-European Ancestry Groups

November 18, 2013

JM Kocarnik

Genome-wide association studies (GWAS) have successfully identified thousands of single nucleotide polymorphisms (SNPs) associated with a large variety of diseases and traits. The great majority of these studies, however, have primarily been performed in populations of European ancestry. Because many SNP allele frequencies differ by ancestral group, it is important to understand whether these European-ancestry based SNP-trait associations generalize to other ancestry populations. A recent report in *PLoS Biology* by Dr. Chris Carlson and colleagues in the Public Health Sciences Division provides evidence that while the effect of most SNPs generalize from European to non-European populations, a substantial portion have significantly different effects in at least one non-European population. These findings suggest caution when considering the application of genetic risk information to other ancestry groups.

To investigate generalization broadly, the researchers utilized the rich data available in the Population Architecture using Genomics and Epidemiology (PAGE) study, a consortium of multi-ancestry, population-based studies such as the Women's Health Initiative. "The exciting thing about this research was the opportunity to work in the context of some of the largest available multi-ethnic studies in the U.S.," says Dr. Carlson. These studies have rich phenotypic data on multiple traits in multiple populations, allowing the researchers to investigate 68 index SNPs associated with 5 phenotypes (body mass index, HDL cholesterol, LDL cholesterol, triglycerides, and type 2 diabetes) in six ancestry groups (European American, African American, Hispanic American, Asian American, Native American, and Pacific Islander). "Importantly," says Carlson, "this let us try to get away from a generalization analysis that is specific to a single trait, and instead focus explicitly on generalization."

Overall, the researchers found that the majority (at least 70%) of European ancestry GWAS findings can be expected to show an effect in the same direction for non-European ancestry populations. At the same time, however, the magnitude of the effect varied considerably, with a substantial portion of SNPs (25%) showing a significantly different (usually weaker) effect in at least one non-European ancestry population. These inconsistent effects were most frequent in the African American population, representing a trend of consistent direction but diluted effect size for these SNPs relative

to European ancestry (see figure). This dilution effect was consistent with the expectations of differential tagging, which refers to differences in the correlation between the index SNP and the functional variant across different populations. As was seen, populations with lower correlation between the index (tag) SNP and the functional SNP were expected to have diluted effects at the index SNP. The researchers further used this differential tagging information in different ancestry groups to perform fine-mapping analyses for ten loci, and found that at least half of European ancestry GWAS signals tag a common, functional variant. This suggests that trans-ethnic fine-mapping may be a powerful approach to identify both population-specific and common signals.

On a broader scale, these generalization findings have important implications for the translation and application of genetic information. "As the genetic epidemiology community works toward the laudable goal of genetic risk prediction models that could be clinically actionable," says Carlson, "these findings caution that the models need to be recalibrated for application in populations with different genetic ancestries." Such recalibrations will require genetic studies that include non-European ancestries, in order to ensure the equitable translation of genetic information into public health and clinical applications across populations. Furthermore, by including multiple ancestral groups in genetic studies, trans-ethnic analyses can help identify the functional genetic variation that is shared in common across ancestries. Says Carlson, "as the functional variants are identified at each locus, risk prediction models based directly on genotype at the functional variants may extrapolate robustly across populations." While the identification of functional variants remains challenging, the results from this paper suggest that the effort will be both necessary and rewarding.

Other PHS investigators contributing to this project were Drs. Ulrike Peters, Cara Carty, and Charles Kooperberg, with additional support from Alicia Young and members of the WHI/FHCRC component of the PAGE study.

[Carlson CS, Matise TC, North KE, Haiman CA, Fesinmeyer MD, Buyske S, Schumacher FR, Peters U, Franceschini N, Ritchie MD, Duggan DJ, Spencer KL, Dumitrescu L, Eaton CB, Thomas F, Young A, Carty C, Heiss G, Le Marchand L, Crawford DC, Hindorff LA, Kooperberg CL, PAGE Consortium.](#) 2013. Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study. *PLoS Biol* 11(9):e1001661.

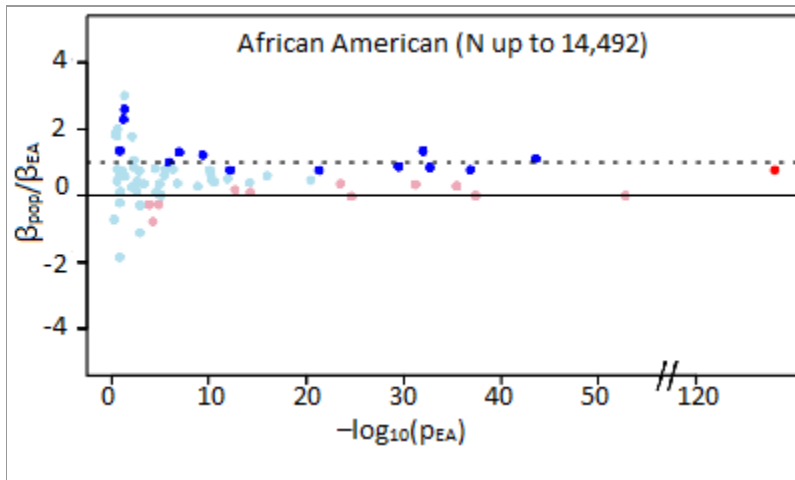


Image adapted from Carlson et al. *PLoS Biol* 2013.

Generalization analysis in the African American population, plotting the ratio of the effect size in African Americans to the effect size in European Americans (y axis) against the p-value for the trait association in European Americans (x axis). These ratios are generally between 0 and 1, suggesting two trends: the general consistency in direction of effect in both populations, and the general dilution of effect size in the African Americans.