Race/ethnicity is a social construct incorporating biological, sociocultural, psychological, and behavioral components, and there is no genetic definition of race. In contrast, genetic ancestry is a genetics concept that describes the architecture of genome variation between populations. All genetic variation begins locally, as a new mutation in an individual, and thus, all new mutations are initially geographically localized. By virtue of this shared population history, there is a correlation between genetic variants found in one geographic region as compared with those found elsewhere. It follows that variation that is common across human populations separated by geography tends to be old, and many of these variants were likely present at the time of human expansion out of Africa ∼100,000 years ago. In contrast, more recent mutations have not had the time to spread between populations and tend to be of lower frequency, unless under unusual selection pressure and localized to geographic populations. When the recombination history has been short (<20 generations or so), these recent variants will tend to remain in linkage disequilibrium with other markers located on ancestral haplotype on which the new mutation occurs.

The primary utility of estimates of continental ancestry using AIMs is to control for confounding from population stratification in genetic association studies among unrelated individuals from admixed populations. It has led to the recognition that within African Americans residing in the United States, the degree of European admixture varies by geographic region. This has the potential to lead to confounding from population stratification in genetic association studies leading to both false-positive (type I error) and false-negative (type II error) results. The study by Wassel et al in this edition of Circulation: Cardiovascular Genetics further highlights the existence of population substructure among African Americans recruited from multiple clinical centers across the United States. Using a set of 96 AIMs selected from the Multi-Ethnic Study of Atherosclerosis (MESA) genetic repository data, while using the HapMap Yoruban Nigerians and MESA self-identified whites as reference ancestral groups, mean European ancestry for African Americans was 20.1%; however, significant (P<0.001) differences in European ancestry were identified in African Americans across geographically distinct MESA field sites. These findings underscore the importance of accounting for population substructure even within analyses limited to groups categorized by self-reported race/ethnicity, especially in multicenter studies.

Other investigators have used assessment of continental ancestry to make inferences regarding the potential role of genetics in the causal pathways for complex disease within populations. In this study, Wassel et al examined the association of genetic ancestry with subclinical cardiovascular disease (coronary artery calcium [CAC] and carotid intimal medial thickness) in 712 African American and 705 Hispanic participants in the MESA candidate gene substudy. In the self-identified African Americans, estimates of European ancestry were independently associated with an increased risk for CAC but inversely associated with the risk for carotid intimal medial thickness, consistent with disparities in these phenotypes noted between racial/ethnic groups at the population level. Among the Hispanics, results were more difficult to interpret and suggested a “threshold effect” for European ancestry; the highest tertile of European ancestry was associated with a 34% increased CAC prevalence. The authors conclude that the linear association of ancestry with CAC prevalence and common carotid intimal medial thickness in African American participants in MESA suggests that “genetic effects may indeed be important for African Americans.”

Although this conclusion seems justified, it is problematic to conclude that the association of genetic ancestry with disease risk conclusively indicates that there are likely underlying genetic contributions to the disease of interest. Race
as a social construct devoid of genetic meaning incorporates many social environmental influences that influence disease susceptibility. Moreover, estimates of continental ancestry are correlated with socially defined race. Given the strong correlations between ancestry, race, environmental, and social factors, and our imprecision in measuring and adjusting for these factors, the potential remains for residual confounding from both measured and unmeasured nongenetic factors. Therefore, we must be cautious to avoid overinterpreting or frankly misinterpreting the results of studies that demonstrate associations between genetic ancestry and disease.

It should also be noted that others have advanced arguments from a genetic perspective for tempered enthusiasm for the relative importance (or existence) of a genetic explanation for health disparities between populations. In a sense, the “common variant-common disease” paradigm for the genetic contribution to complex diseases is by definition incongruous with the notion of a significant influence of population-specific genetic factors in the development of common complex diseases; common variants are by definition old—likely present in the human genome at the time of African outmigration 100,000 years ago—and therefore global. This argument is supported by a recent study that examined the total known allelic variation at several genomic loci in 3 continental populations. They demonstrated that the majority of total allelic variation (known) is captured by 3 to 5 haplotypes shared by at least 2 of the 3 continental populations. The counterargument is that the “genetic component” of population differences in disease risk (if it exists) is the result of accumulated small differences in common alleles and an effect of population-specific alleles. There is a paucity of examples of functional, population-specific alleles that contribute to common diseases. Recently, we have reported on 2 nonsynonymous polymorphisms in the human corin gene that are in complete linkage disequilibrium and only found in persons of African ancestry. The dysfunctional corin I555(P568) minor allele prevalence is 6.5% in African Americans and is associated with risk for hypertension and cardiac hypertrophy. However, the aggregate attributable risk associated with population-specific alleles as related to population-level differences in common disease incidence is likely to be very small, and the existence of a significant influence of accumulated small differences in common alleles in explaining population differences in disease risk seems implausible unless one postulates that disease risk is the result of interactions between multiple genetic loci. The truth of the matter is that we simply do not know enough at the present time about the genetic components of complex disease in the human population, much less within specific populations.

Nonetheless, the pursuit of the genetic factors that impact the health of specific populations continues and is not without potential merit and benefit. One of the most powerful direct methods available to use population admixture as a tool to identify genetic factors that contribute to complex disease risk is mapping by admixture linkage disequilibrium, hereafter referred to as admixture mapping. On average, African American genomes are a mixture of 15% to 20% European-derived segments and 80% to 85% African-derived segments. Because admixture occurred ~200 years ago, there has been minimal recombination between chromosomes of African and European ancestry. Therefore, admixture results in long-range linkage disequilibrium between genomic markers that have different frequencies in the 2 populations. The underlying assumption of admixture mapping for disease loci is simple. If there are genetic contributions to the differences in disease prevalence between groups, the risk in each group may be associated with ancestrally shared allelic variants because admixture creates linkage disequilibrium between the disease causing variants and nearby markers on the ancestral chromosome. Using a “case-only” approach, genomic regions in persons with the disease or phenotype of interest are genotyped for AIMs and the average degree of African or European ancestry for the group is calculated across genomic loci across each chromosome. The presence of “overtransmission” of African or European ancestry at a genomic locus in cases suggests the presence of a potential disease susceptibility allele that is cosegregating with the other ancestral markers. In subsequent case-control analysis, the finding that this locus-specific overtransmission of African or European ancestry is absent in the control population strengthens the conclusion that the locus might harbor alleles related to disease risk. The genomic resolution of admixture mapping can identify genomic loci associated with disease in the range of ~30 cM, harboring an average of 100 potential candidate genes. Therefore, the next step of a successful admixture study is the fine mapping of the identified regions to identify the specific alleles associated with disease. One might identify the genes in the region and proceed with a classic case-control association study using a candidate gene–tagged single-nucleotide polymorphism/haplotype-based approach. However, this is not without significant challenge in the African population, because they have smaller haplotype blocks and less linkage disequilibrium than non-African populations making tagged single-nucleotide polymorphism approaches more difficult. For example, in one study that used extensive resequencing of 76 genes as a reference found than even with the dense coverage provided by the Affymetrix version 6.0 gene chip, on ~45% of single-nucleotide polymorphisms were tagged (r²>0.8) in the YRI sample.

There have been several recent successes using the methodology of admixture mapping to identify loci associated with complex diseases, in particular, those diseases that demonstrate disparities in prevalence among racial/ethnic groups. Admixture mapping has been used to successfully identify a locus associated with increased risk for prostate cancer in African American men, a disease that previous epidemiological studies consistently demonstrated to have both an increased incidence rate and mortality rate compared with European Americans. Multiple sclerosis is another disease with a higher prevalence in European populations than African American populations. Using DNA samples from 605 African Americans with multiple sclerosis (case-only approach), Reich et al successfully identified a genomic locus with increased European ancestry in cases, and subsequent case-control studies demonstrated that it was associated with a 44% increased relative risk for multiple
sclerosis in African Americans given heterozygosity for European ancestry at this locus. Zhu et al.\textsuperscript{36} carried out an admixture mapping study using genome scan microsatellite markers in the US National Heart, Lung, and Blood Institute’s Family Blood Pressure Program and identified 2 genomic loci in chromosomes 6q24 and 21q21 that are associated with hypertension African American men.

The role of admixture mapping relative to the merit of genome-wide association studies in multiethnic samples remains to be determined.\textsuperscript{37} At present there are only a few reports of successful genome-wide studies in multiethnic samples. Of note, admixture mapping can also be a useful tool to be used in conjunction with genome-wide association studies. Limiting genome-wide association study analysis to genomic regions first identified by admixture mapping has the potential to reduce the total number of comparisons and reduce type II (false-negative) error.

In summary, this study by Wassel et al demonstrates that continental ancestry is associated with the risk for prevalent CAC and carotid intimal medial thickness. Given the population differences in the prevalence of these intermediate cardiovascular phenotypes, this suggests that admixture mapping may be a useful approach to identify the genetic determinants of coronary calcification and internal carotid intimal medial vascular remodeling. However, the ultimate public health importance of the information we obtain remains to be determined. In the meantime, continued attention to modifiable nongenetic social and environmental risk factors remains of paramount importance at both the individual and population level.

Disclosures

None.

References


**KEY WORDS:** genetics  ■  Editorial
Genetic Ancestry, Population Admixture, and the Genetic Epidemiology of Complex Disease
Daniel L. Dries

doi: 10.1161/CIRCGENETICS.109.922898

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/2/6/540

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org/subscriptions/