Received: 17 May 2020 | Revised: 22 July 2020 | Accepted: 20 August 2020

DOI: 10.1002/ajpa.24144

RESEARCH ARTICLE

American Journal of PHYSICAL ANTHROPOLOGY WILEY

Leveraging genetic ancestry to study health disparities

Ken Batai[1] | Stanley Hooker[2] | Rick A. Kittles[2]

[1]Department of Urology, University of Arizona, Tucson, Arizona

[2]Division of Health Equities, Department of Population Sciences, City of Hope Comprehensive Cancer Center, Duarte, California

Correspondence
Rick A. Kittles, Division of Health Equities, Department of Population Sciences, City of Hope Comprehensive Cancer Center, 1500 E. Duarte Rd, Duarte, CA 91010-3000.
Email: rkittles@coh.org

Funding information
National Cancer Institute, Grant/Award Numbers: 1R21CA248361, U54CA143924; National Institute on Minority Health and Health Disparities, Grant/Award Number: R01MD007105

Abstract

Research to understand human genomic variation and its implications in health has great potential to contribute in the reduction of health disparities. Biological anthropology can play important roles in genomics and health disparities research using a biocultural approach. This paper argues that racial/ethnic categories should not be used as a surrogate for sociocultural factors or global genomic clusters in biomedical research or clinical settings, because of the high genetic heterogeneity that exists within traditional racial/ethnic groups. Genetic ancestry is used to show variation in ancestral genomic contributions to recently admixed populations in the United States, such as African Americans and Hispanic/Latino Americans. Genetic ancestry estimates are also used to examine the relationship between ancestry-related biological and sociocultural factors affecting health disparities. To localize areas of genomes that contribute to health disparities, admixture mapping and genome-wide association studies (GWAS) are often used. Recent GWAS have identified many genetic variants that are highly differentiated among human populations that are associated with disease risk. Some of these are population-specific variants. Many of these variants may impact disease risk and help explain a portion of the difference in disease burden among racial/ethnic groups. Genetic ancestry is also of particular interest in precision medicine and disparities in drug efficacy and outcomes. By using genetic ancestry, we can learn about potential biological differences that may contribute to the heterogeneity observed across self-reported racial groups.

KEYWORDS

genetic ancestry, health disparities, precision medicine

1 | INTRODUCTION

Race is a sociocultural concept, and genetic and biological evidence does not support racial classification for human populations. When race and ethnicity is used in biomedical research, it is often self-reported and used as a proxy for measurable indicators of group differences, such as socioeconomic status, cultural and behavioral lifestyle, and biology. Reducing each of these contributors into a composite called "race" precludes independent analysis of important factors, such as genetics and the physical and social environments, which vary significantly within populations.

Human genetic variation is structured by the evolutionary history of our species. The pattern of this population structure, however, is not bounded or discrete, but continuous, resulting from the demographic history of populations which includes such forces as natural and social "mate" selection, genetic drift, gene flow and mutations

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2020 The Authors. *American Journal of Physical Anthropology* published by Wiley Periodicals LLC.

(Cavalli-Sforza, Menozzi, & Piazza, 1994; Livingstone, 1962; Serre & Pääbo, 2004). Our knowledge of human genetic variation has grown enormously over the past several decades. Single-nucleotide polymorphisms (SNPs) are the most common form of DNA variation in the human genome. These are nucleotide changes which can contribute to variation in genetically controlled traits, such as skin and hair color, height, physical composition, susceptibility to disease and therapeutic response. At present, there are more than 84 million SNPs in the human genome (The 1000 Genomes Project Consortium et al., 2015). A large fraction of these SNPs is found at a frequency less than 5% and thus is private or common in only a single population (The 1000 Genomes Project Consortium et al., 2012; The 1000 Genomes Project Consortium et al., 2015). Genetic polymorphisms, such as SNPs, have been used to explore how genetic variation is structured within and between human populations. Allocating individuals into clusters based on genotypes which reflect shared ancestry is possible depending on which genetic markers are used (Novembre & Peter, 2016; Rosenberg et al., 2002; Shriver & Kittles, 2004; Tang, Coram, Wang, Zhu, & Risch, 2006; Tishkoff et al., 2009).

The existence of genomic regions that differ significantly among human groups raises two important questions. Does human genetic variation explain differences in health status or health disparities? Will genomics technology and applications help us move away from the use of race as a proxy for biology in biomedical research? In this paper, we show that racial/ethnic categories should not be used as a proxy for sociocultural factors or global genomic clusters in biomedical research or clinical settings, because of the high genetic heterogeneity that exists within traditional racial/ethnic groups. In many cases, this has hampered biomedical research in diverse populations. Thus, there is a clear need for scientists to advance the understanding of the concepts of self-identified race and ethnicity (SIRE) and ancestry informative markers (AIMs) in biomedical research. We will illustrate this using breast cancer (Bca) and prostate cancer (Pca) health disparities research as examples, including an analysis of variants identified in genome-wide association studies (GWASs) for Pca.

2 | SIRE AND GENETIC ANCESTRY

Many studies continue to utilize socially and politically constructed racial/ethnic categories that are not appropriate for investigating genetic contributions to the etiology of complex diseases, drug response, and more importantly, health disparities. Race is a social construct, and racial categorization in the United States has largely been based on skin color and ancestry and it lacks biological integrity. Descriptions of African Americans in much of the biomedical literature are predicated on the assumption that people of African descent, no matter where they reside, constitute a biological race. Hispanic/Latino Americans, who also are biologically and culturally heterogeneous, sadly are classified as an ethnic group based on shared language. Because inequality based on race/ethnicity are manifested in health and health care in many different ways, racial/ethnic categories are still useful in biomedical research. Race and ethnicity may be an important determinant to monitor health status (e.g., incidence, prevalence, and mortality), health care access and utilization, and health care quality (Burchard et al., 2003; LaVeist, 1994). Race and ethnicity are also useful to assess impacts of racism on health and health care access and quality (Gee, Hing, Mohammed, Tabor, & Williams, 2019; Krieger, 2012). However, the use of race to identify groups may also confound biomedical studies (Caulfield et al., 2009).

This is because race reflects deeply confounded sociocultural as well as biological factors, especially when one examines differences between African Americans and European Americans. In 2015, age-standardized mortality rates for African Americans exceeded European Americans by 40% for cerebrovascular disease including stroke, 22% for heart disease, and 13% for cancer (Cunningham et al., 2017). In the U.S. population in 2017 and 2018, African Americans had the highest rates of obesity at 50% compared to 45% in Hispanic/Latino Americans and 42% in European Americans (Hales, Carroll, Fryar, & Ogden, 2020). Data on sociocultural factors, such as educational attainment, unemployment, and household below poverty rate, as well as health care access factors, such as health insurance status, also showed inequalities between African Americans and European Americans (Cunningham et al., 2017).

However, the traditional paradigm of using race as a proxy for ancestral background in biomedical research is slowly shifting given the heterogeneity within racial/ethnic groups that exists in U.S. populations. This is especially the case for research on African Americans and Hispanic/Latino Americans. In these recently admixed populations, continental genetic contribution or biogeographic ancestry may be estimated using AIMs. AIMs are genetic markers, typically SNPs, which are found across the human genome and have large allele frequency differences between continental groups such as Western Europeans and West Africans and are powerful for estimating biogeographic ancestry (Shriver & Kittles, 2004). Genetic ancestry for each individual is estimated by comparing individual's AIM genotypes to that of a reference panel consist of samples from continental ancestral populations using a statistical probability modeling approach (Pritchard, Stephens, & Donnelly, 2000). Continental ancestral genetic contributions estimated using AIMs illustrate the fluidity and variation of genetic ancestry within traditional U.S. "racial" groups. Levels of European admixture in African American populations vary across the U.S. African Americans from the southern states tend to have lower levels of European admixture, while much higher estimates of European ancestry is observed in the Pacific North West (Bryc, Durand, Macpherson, Reich, & Mountain, 2015; Kittles, Santos, Oji-Njideka, & Bonilla, 2007). Recent studies have investigated the genetic evidence of very recent migrations and admixture in African Americans and Hispanic Americans using identity-by-descent approach (Baharian et al., 2016; Dai et al., 2020). These studies examined spatial distributions of shared haplotype and reveal heterogeneity of admixture and the contribution from diverse recent ancestors among African Americans and Hispanic Americans across the United States. The geographic distribution of genetic ancestry and haplotype sharing should be interpreted in terms of well-known historical and demographic events that have played an important role in African

American history (Torres, Doura, Keita, & Kittles, 2012). This illustrates the limitation with the use of traditional SIRE grouping ignoring within-group heterogeneity for biomedical research.

3 | UTILITY OF GENETIC ANCESTRY IN HEALTH DISPARITIES RESEARCH

The use of genetic ancestry in biomedical research has become popular among biomedical researchers who understand that self-reported race is not a strong proxy for biology (Shriver & Kittles, 2004; Winkler, Nelson, & Smith, 2010). Genetic ancestry estimates can be used to correct for confounding effects of population stratification in genetic association studies of admixed populations (Halder & Shriver, 2003). Adjusting for population stratification in statistical models reduces chance of false positive or negative associations, when risk allele frequency is different between the ancestral populations and prevalence of disease is also different in two groups.

Genetic ancestry can also be used to dissect the complex relationship between the potentially ancestry-related genetic factors or sociocultural factors that are associated with health disparities. For example, investigators have long been interested in the relationship between skin color and socioeconomic status and hypothesized that psychosocial stress associated with having dark pigmentation was the reason for higher prevalence of hypertension in African Americans compared to European Americans (Harburg, Gleibermann, Roeper, Schork, & Schull, 1978; Klag, Whelton, Coresh, Grim, & Kuller, 1991). A study also found evidence of an interaction between skin color and socioeconomic factors (Sweet, McDade, Kiefe, & Liu, 2007). More recently, a study that included a large number of African American and Hispanic American women from Women's Health Initiative showed a significantly positive association between West African ancestry and hypertension in both African American and Hispanic American women, even after accounting for risk factors and a summary measurement of neighborhood socioeconomic characteristics derived from participants' census tract information (Kosoy et al., 2012). However, the Family Blood Pressure Program study did not support the finding, revealing that education, not genetic ancestry, was an important factor associated with blood pressure in African Americans (Non, Gravlee, & Mulligan, 2012). Given the lack of consensus in studies of genetic ancestry and hypertension, it is likely that ancestry-related genetic factors may contribute a small portion of risk of high blood pressure in African Americans. However, socioeconomic factors and psychosocial stress likely account for a larger portion of the differences in risk observed in across populations.

Cancer also disproportionately affects recently admixed populations. Anthropologists have studied many complex diseases but have not contributed much to cancer disparities research. Anthropology and anthropological perspectives can be integrated with oncology, population genetics, molecular epidemiology, and behavioral sciences to address cancer health disparities (Newman & Kaljee, 2017). The conceptual models used in population studies of health disparities generally include various factors from individual (biology/genetics and behavior), to community and societal levels, and further to structural levels (Alvidrez, Castille, Laude-Sharp, Rosario, & Tabor, 2019; Warnecke et al., 2008). Because numerous biological and sociocultural factors, including ancestral genomic background, contribute to cancer disparities, such a transdisciplinary approach is necessary to understand the etiology and variation of clinical and pathological characteristics in effort to reduce the cancer burden in racial/ethnic minority groups.

Genetic ancestry has been used to explore whether ancestral genetic background or behavioral and social factors can explain the difference in Bca incidence and clinical features and outcomes between Latinas and European American women (Al-Alem et al., 2014; Fejerman et al., 2008; Fejerman et al., 2010; Fejerman et al., 2013). Bca is the most common type of cancer among women (Siegel, Miller, & Jemal, 2020). Latinas have lower Bca incidence rate than European American women, but they are more likely to be diagnosed with advanced stage and have lower Bca specific survival rate than European American women (Iqbal, Ginsburg, Rochon, Sun, & Narod, 2015; Jemal et al., 2017). Studies among Latinas have shown that both genetic and social/behavioral factors contribute to the difference in Bca incidence between Latinas and European American women. Higher European ancestry was associated with increased risk of Bca in Latinas from California (Fejerman et al., 2008) and Mexican women (Fejerman et al., 2010). Without adjusting for socioeconomic status and risk factors, the association between European ancestry and Bca risk was very strong. However, genetic ancestry is also associated with educational attainment, socioeconomic status, and Bca risk factors (Ziv et al., 2006). Indigenous American ancestry was positively associated with obesity, especially among foreign-born Latinas. The inclusion of education, socioeconomic status, and Bca risk factors in the regression models weakened the association between genetic ancestry and Bca risk in these studies, illustrating that both genetic ancestry and sociocultural/socioeconomic factors increase Bca risk.

Similar approaches have been used to investigate associations of genetic ancestry with cardiovascular disease and atherosclerosis characteristics (Gebreab et al., 2015; Wassel et al., 2009), Type 2 diabetes (Cheng, Reich, et al., 2012; Qi et al., 2012), bone mineral density (Noel et al., 2017), obesity (Klimentidis, Arora, Zhou, Kittles, & Allison, 2016), and other diseases and traits. However, there are some limitations with the use of genetic ancestry. First, there has been a limitation in the number of socioeconomic, lifestyle, and health care access variables collected and included in the statistical analyses. Second, genetic ancestry estimates may be strongly correlated with sociocultural and behavioral risk factors (Florez et al., 2009; Ziv et al., 2006). Both genetic ancestry and traditional race/ethnicity groups may be a proxy for socioeconomic status or other sociocultural factors, and there may be residual confounding factors or unmeasured sociocultural factors that influenced the association between genetic ancestry and disease phenotypes that were not included in the statistical modeling (Drake, Galanter, & Burchard, 2008). Moreover, overemphasis of genetic ancestry's effect on health should be avoided because it may reinforce misconceptions on race and takes attentions away from sociocultural and behavioral factors that can be modified

to reduce health disparities (Sankar et al., 2004). Nonetheless, genetic ancestry can capture the heterogeneity in ancestral genomic contributions in African American and Hispanic/Latino American populations, and the use of genetic ancestry allows one to investigate the relative effect of ancestry-related genetic factors on disease phenotypes compared to sociocultural factors, before investigators even attempt to identify specific genomic regions or polymorphisms that are associated with diseases that disproportionately impact racial/ethnic groups.

4 | GENOMICS AND HEALTH DISPARITIES

What are those ancestry-related genetic risk factors? Can we localize the areas of human genomes that may explain the health disparities? Population genetics models show that population bottlenecks after the out-of-African migrations and other events, such as population isolation (through isolation-by-distance and geographic barriers), changes in population size, admixture, and gene flow from archaic human, shaped human genomic variation (Bergström et al., 2020; The 1000 Genomes Project Consortium et al., 2015). African populations have deep demographic histories evidenced by a greater number of genetic variants than non-African populations, and many of which are not observed in non-African populations. After the out-of-African event, population size grew rapidly outside Africa accumulating population-specific low frequency variants. Population subdivision at a continental level generally suggests limited gene flows between continental groups after initial dispersals (Li et al., 2008). Finer-scale population structure exists within African resulting from multiple factors, such as geographic barriers, climates, and language differences (Campbell, Hirbo, Townsend, & Tishkoff, 2014; Gomez, Hirbo, & Tishkoff, 2014).

Natural selection also greatly influenced the pattern of genetic variation across human populations depending on geographic distribution of the selective pressure (Casto & Feldman, 2011; Novembre & di Rienzo, 2009; Quintana-Murci, 2016). An evolutionary framework for common disease suggests that old genetic variants reflect ancient adaptations to the lifestyle of old-world populations. With changing environment and lifestyle, these same variants now increase risk for common disease in modern populations. DNA sequence variants in the *APOE* and *PPARG* genes, which influence risk for Alzheimer's disease and Type 2 diabetes respectively, are some examples of this (di Rienzo & Hudson, 2005). For both genes, ancestral alleles (ε4 for *APOE* and P12 or rs1801282 C allele for *PPARG*) are risk variants suggesting these ancestral alleles may be "thrifty" alleles that were beneficial to populations in the past (prewesternized lifestyles). Interestingly, rs1801282 allele frequencies are correlated with latitude (Hancock et al., 2008). Other well-known examples of natural (and cultural) selection that shaped human genomic variation are gene for lactase persistence and adaptation to environment that malaria is common (Gomez et al., 2014). More recently, *APOL1* was identified as a gene that may explain high prevalence of chronic kidney disease and end stage renal disease in African Americans. Two nonsynonymous variants (rs73885319 and rs60910145) that are in perfect linkage disequilibrium (r^2 = 1.0) are significantly associated with kidney disease in African Americans (Langefeld et al., 2018; Lin et al., 2019). Risk alleles of these SNPs along with a six-base pair deletion at the nearby location (rs71785313) are common in sub-Saharan Africa, but they are not observed in European and East Asian populations. It is hypothesized that these alleles provide a protective effect against *Trypanosoma brucei* infection and became common in sub-Saharan African possibly through natural selection (Genovese et al., 2010).

To localize areas of genomes that contribute health disparities, admixture mapping and GWAS are often used. Admixture mapping is an approach to identify genomic regions enriched with ancestral genomic contributions that are associated with a disease, when the prevalence of the disease varies significantly between two or more populations (Seldin, Pasaniuc, & Price, 2011; Winkler et al., 2010). Admixture mapping takes advantage of long-range linkage disequilibrium in recently admixed populations, because not enough time is accumulated to break down the long-range haplotypes unique to their ancestral populations through recombination. In this approach, a couple of thousands AIMs scattered across genome or over 100,000 SNPs generated for GWAS are used to infer local ancestry (genetic ancestry estimated for each locus) and then association between local ancestry and the disease trait are tested (Price et al., 2009; Seldin et al., 2011).

This admixture mapping approach has been used to study Pca in African Americans, revealing that West African genetic background within Chromosome 8 (8q24) was strongly associated with Pca (Bensen et al., 2014; Freedman et al., 2006). Pca is another cancer type that disproportionately affects African Americans. Pca is the most common type of cancer among men (Siegel et al., 2020). This cancer appears to have a strong influence of genetic ancestral risk factors (Haiman et al., 2011a). The incidence rate for Pca is higher in men of African descent than men from other populations (Odedina et al., 2009), and the highest Pca incidence rate has been observed among Afro-Caribbean living in the United Kingdom (Chinegwundoh et al., 2006). The associations of variants in the 8q24 region with Pca risk have also been demonstrated in many other studies of African Americans (Chang et al., 2011; Haiman et al., 2007; Haiman et al., 2011a; Robbins et al., 2007; Wang et al., 2011; Xu et al., 2009; Xu, Bensen, Smith, Mohler, & Taylor, 2011), Caribbeans (Okobia, Zmuda, Ferrell, Patrick, & Bunker, 2011), and West Africans (Murphy et al., 2012).

Admixture mapping of Bca among Latinas reveals a different pattern. A region on Chromosome 6 (6q25) has been identified where genetic ancestry was significantly associated with Bca risk (Fejerman et al., 2012). The investigators show that increasing Indigenous American chromosomal copy at 6q25 reduced Bca risk, while increasing European chromosomal copy increased Bca risk. This finding is consistent with previous studies showing increased Bca risk with high European ancestry among U.S. Latinas and Mexican women (Fejerman et al., 2008; Fejerman et al., 2010). Although admixture mapping is a powerful approach, fine mapping is necessary to locate the causative locus, after potential regions are identified through admixture mapping.

The GWAS approach uses densely genotyped and imputed data including over millions of markers or whole-exome sequence data to localize causative locus. Recent GWASs of cancer risk have identified many highly differentiated population-specific variants that are also associated with cancer risk. These cancer risk variants may explain the differences in cancer burden in racial/ethnic groups. For example, there are several subtypes of Bca, and triple-negative (estrogen-receptor, progesterone-receptor, and human epidermal growth factor receptor 2 negative) Bca is an aggressive type common among women of African descent and women with *BRCA1* mutations (Brewster, Chavez-MacGregor, & Brown, 2014; Dietze, Sistrunk, Miranda-Carboni, O'Regan, & Seewaldt, 2015). African descent women are also diagnosed with Bca at a younger age (Der et al., 2015). Pathogenic germline mutations on *BRCA1* and *BRCA2* are highly penetrant predisposing carriers to early-onset Bca. The prevalence of pathogenic *BRCA1* and *BRCA2* mutations in African descent populations is not well understood, but some studies suggest a high prevalence in African descent populations (Adedokun et al., 2020; Akbari et al., 2014; Fackenthal et al., 2012; Pal et al., 2015). GWAS, on the other hand, have identified many common SNPs that contribute to Bca susceptibility with much smaller effect. One SNP (rs10069690) on the *TERT* gene on chromosome 5 is associated with triple-negative Bca in European populations (Purrington et al., 2014) and estrogen-receptor negative Bca in African Americans (Huo et al., 2016). The frequency of the rs10069690 risk allele (T) is higher in African populations compared to other populations in 1000 Genomes Project (63.9% in YRI, 52.5% in ASW, 27.6% in EUR, 16.9% in EAS, 22.0% in AMR).

Contrary to a genetic variant that increases Bca risk in African American women, GWAS among the U.S. Latinas identified population-specific variants in 6q25 region, minor allele of rs140068132 and rs147157845, show strong protective effects, particularly against estrogen negative Bca (Fejerman et al., 2014). Indigenous American ancestry in 6q25 was previously shown to be associated with reduced risk (Fejerman et al., 2012). These variants were common in 1000 Genome Project Latin American population ranging between 5% Puerto Ricans and 14% in Mexicans but were absent in European or African populations. The peak of signal from the admixture mapping encompasses the estrogen receptor 1 (*ESR1*) gene, and the top hits from the GWAS were also located upstream of the *ESR1* gene.

Similar to differences in Bca subtypes across racial/ethnic groups, frequencies of renal cell carcinoma histologic subtypes vary across racial/ethnic groups. Renal cell carcinoma is the most common type of kidney cancer, and African Americans have higher kidney cancer incidence compared to European Americans (Siegel et al., 2020). There are three major histologic subtypes of renal cell carcinoma (clear cell, papillary, and chromophobe). The frequency of papillary type is higher in African Americans than European Americans (Batai et al., 2019; Olshan et al., 2013; Wang et al., 2017). Among Hispanic/Latino subgroups, Dominicans exhibit higher West African genetic ancestry estimates (Bryc et al., 2010), and papillary renal cell carcinoma is more common in Dominicans than other Hispanic/Latino subgroups or European Americans (Batai et al., 2019). There may be ancestry-related genetic (or sociocultural) factors that explain this observation, but differences in genetic risk factors across racial/ethnic groups have yet to be explored for renal cell carcinoma.

The most significantly associated SNPs for Pca have been found in the 8q24, the region of Chromosome 8 that was initially identified in African Americans using an admixture mapping approach. There are now a number of SNPs on 8q24 region that are strongly associated with Pca risk in multiple populations (Eeles et al., 2008; Thomas et al., 2008; Wang et al., 2015; Xu et al., 2012). The first Pca GWAS in African American men identified a SNP, rs10505483 in this region that was most strongly associated with Pca (Haiman et al., 2011b). This SNP was also strongly associated with Pca in other global populations (Cheng, Chen, et al., 2012; Murabito et al., 2007). The frequency of risk allele (A) is higher in West African populations than non-African populations (53.7% in YRI, 32.0% in ASW, 3.6% in EUR, 21.3% in EAS, 5.0% in AMR from 1000 Genomes Project). Later studies among men of African ancestry found several population-specific risk variants (rs114798100 and rs72725854) (Conti et al., 2017; Han et al., 2016). These population-specific risk SNPs exist at low frequency (<10%) in West African populations and African Americans but are not found in European or East Asian populations.

The higher frequency of risk alleles in human populations with a greater disease burden likely accounts for a portion of the differences we see in Pca risk and outcomes. It is important to explore more of these variants for diseases that are disproportionately affecting racial/ethnic groups.

5 | ANALYSIS OF PCA GWAS IDENTIFIED RISK ALLELES

We further examined the relationship between genetic ancestry and risk alleles identified using GWAS in global populations (Supplementary Methods). GWAS have identified that many of these Pca risk allele frequencies vary considerably across populations (Figure 1). When Pca risk alleles are ancestral alleles, the risk allele frequency is higher in African descent populations compared to non-African populations (Lachance et al., 2018). It has been noted that some disease associated variants show moderate allele frequency differentiation between continental populations, but generally, the mean F_{ST} [$F_{ST} = \frac{\sigma_p^2}{(1-)}$, where σ_p^2 is the variance in allele frequency between subpopulations, and (1−) is the mean frequency of an allele in the total population] of disease associated SNPs is not higher than estimates from whole genome data or randomly selected SNPs (Adeyemo & Rotimi, 2010; Lohmueller, Mauney, Reich, & Braverman, 2006; Marigorta et al., 2011; Myles, Davison, Barrett, Stoneking, & Timpson, 2008). Adeyemo and Rotimi (2010) used HapMap Phase 3 data and performed population genetic analyses of GWAS reported SNPs for 26 diseases and traits. They accessed the GWAS catalogue in March 3, 2009 and found 25 Pca associated SNPs. The Pca GWAS SNPs had the second largest mean F_{ST} (0.132) after coronary artery disease associated SNPs (F_{ST} = 0.164) among the 26 phenotypes. Many Pca SNPs that we examined also had high F_{ST} (average F_{ST} = 0.161, Supplementary Figure S1).

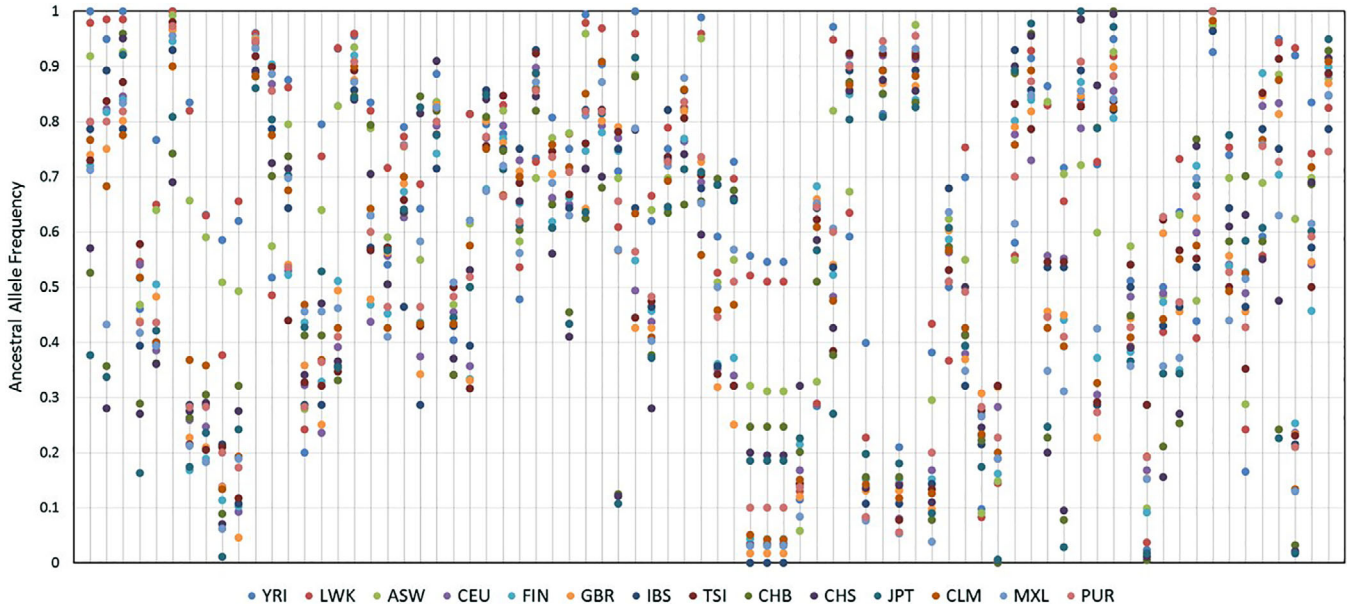

FIGURE 1 A great ancestral allele frequency observed for 76 prostate cancer genome-wide association study (GWAS) single-nucleotide polymorphisms (SNPs) (X-axis) in 1000 Genomes Project populations

TABLE 1 Washington, DC Pca GWAS associated SNPs

Chromosome	SNP	Physical position (BP)[a]	MA[b]	OR (95% CI)[c]	*p*-Value[d]
2	rs10187424	85794297	T	0.73 (0.55–0.95)	.02
8	rs10505483	128125195	A	1.43 (1.10–1.85)	.008
8	rs4242384	128518554	C	1.51 (1.07–2.13)	.02
11	rs10896449	68994667	A	0.74 (0.56–0.99)	.04

Abbreviations: BP, base pair; CI, confidence interval; GWAS, genome-wide association study; MA, minor allele; OR, odds ratio; Pca, prostate cancer; SNP, single-nucleotide polymorphism.
[a]BP position in GRCh37.
[b]MA.
[c]OR and CI.
[d]*p*-Values are from logistic regression analysis adjusting for age and first five principal components.

In order to assess genetic ancestral effect on Pca risk in African American men, we tested associations of 61 Pca GWAS reported SNPs in men from Washington, DC, who were part of Pca GWAS in AA men (Haiman et al., 2011b). Four SNPs (rs10187424, rs10505483, rs4242384, and rs10896449) were associated with Pca risk in this dataset (Table 1). Two 8q24 region SNPs, rs10505483 and rs4242384, that were weakly linked (r^2 = .01), were independently associated with Pca. Risk allele frequency of the four SNPs in a population of AA men from Washington, DC were similar in frequency to that of 1000 Genomes Project African populations and was higher than in non-African populations (Supplementary Figure S2). The Yoruba samples from Nigeria had the highest risk allele frequency for three of the SNPs, while the risk allele frequency of rs10896449 was the highest in African Americans. The allele frequency difference between Yoruba and European populations for rs10505483 was 51.1% which was the largest of four SNPs. SNP rs10896449 had the greatest risk allele frequency difference between Yoruba and Asian populations (64.8%).

Using these four SNPs, we investigated if the number of risk alleles were associated with West African ancestry and Pca. We saw that the number of risk alleles were significantly positively correlated with West African ancestry (p < .001) (Figure 2) and was also significantly associated with Pca (p_{TREND} < .001) as well as aggressive Pca (p_{TREND} = .008) (Table 2). African American men who had five risk alleles had on average 7.1% higher West African ancestry than African American men with less than two risk alleles, and they also had about twofold increased odds of Pca diagnosis (OR = 2.13, 95% CI; 1.11–4.11). A small proportion of African American men (12.9%) had six to seven risk alleles and their mean West African ancestry was 9.2% higher than African American men with less than two risk alleles.

We also performed the same analysis for two 8q24 SNPs and observed a similar pattern. The number of Pca risk alleles was positively correlated with 8q24 local West African ancestry (p < .001) and was also significantly associated with Pca diagnosis (p_{TREND} < .016) as well as high risk Pca (p_{TREND} = .008). African American men with two risk alleles had 22.6% higher 8q24 local West African ancestry than

men with one or no risk alleles, and they had more than twofold increased odds of Pca diagnosis (OR = 2.17, 95% CI; 1.23–3.85) and about fivefold increased odds of high risk Pca (OR = 5.07, 95% CI; 1.78–14.45). Men with three or four risk alleles had even higher 8q24 local West African ancestry, and they had almost threefold increased odds of Pca (OR = 2.75, 95% CI: 1.31–5.79) and over fourfold increased odds of aggressive Pca (OR = 4.23 95% CI; 1.23–14.56).

It is clear that African Americans have an elevated risk of Pca compared to European Americans, but Pca risk varies among African Americans, depending on how many predisposing risk alleles they have. Individuals within traditional racial/ethnic groups have varying degrees of disease risk. The use of racial/ethnic categories as a proxy for genomic geographic clusters ignores the genetic diversity within a group and is inappropriate in biomedical research or clinical settings. It is more appropriate to assess each individual's genetic risk factors. This precision medicine approach is increasingly being incorporated in clinical settings.

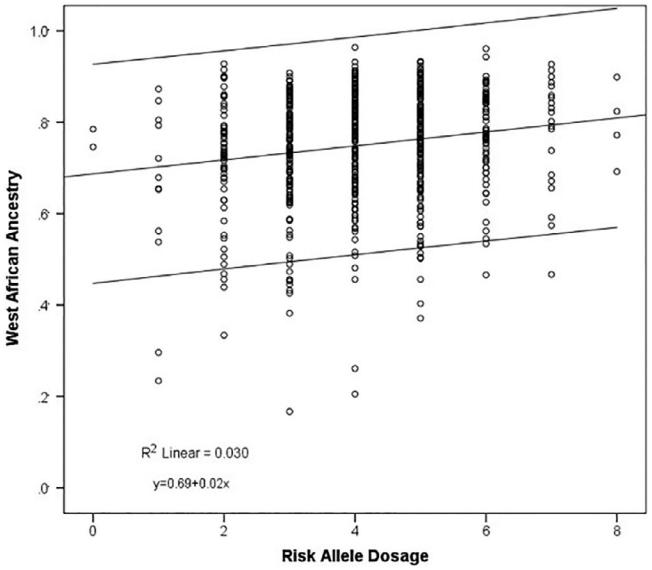

FIGURE 2 Positive correlation between the number of prostate cancer (Pca) risk allele and West African ancestry

6 | RACE/ETHNICITY AND PRECISION MEDICINE

There are wide and varied applications of genomic technology for precision medicine or genomic medicine (Collins & Varmus, 2015; Denny et al., 2019; Green et al., 2011). While recognizing that social/cultural, behavioral, and health care access factors are important causes of health disparities, many believe that genomic research can play important roles in reducing health disparities by understanding population differences in genetic risk factors, treatment response, and also interactions between gene and environment (social/cultural and lifestyle factors) and epigenomic variation. At the same time, some raised concerns about how genomics study findings may re-enforce the common misconception about race as a biological classification, as our understanding of genetic basis of health disparities increases (Bonham, Green, & Pérez-Stable, 2018).

It should be noted that the goals of precision medicine or genomic medicine include an understanding of the genomic variation in diverse populations and its interaction with sociocultural factors that influences disease risk or treatment response in order to improve clinical care by improving accuracy of genetic testing or assessment of therapeutic response (Hindorff et al., 2018). Patients' demographic,

TABLE 2 Correlation between WAA and risk allele dosage and association of risk allele dosage with Pca risk

Number of risk alleles	*n* (%)	Mean global and 8q24 local WAA (*SD*)[a]	Overall risk		Aggressive Pca	
			OR (95% CI)	p_{TREND}[b]	OR (95% CI)	p_{TREND}[b]
Four replicated SNPs[c]						
0–2	172 (12.7)	0.728 (0.157)		<.001		.008
3–4	687 (50.7)	0.770 (0.140)	1.42 (0.78–2.58)		1.79 (0.64–5.07)	
5	321 (23.7)	0.799 (0.117)	2.13 (1.11–4.11)		2.30 (0.74–7.16)	
6–7	175 (12.9)	0.820 (0.101)	4.46 (2.12–9.37)		5.90 (1.83–19.08)	
2 8q24 SNPs[d]						
0	278 (20.5)	0.643 (0.285)		.02		.005
1	574 (42.3)	0.816 (0.234)	1.46 (0.86–2.46)		2.13 (0.77–5.91)	
2	373 (27.5)	0.881 (0.183)	2.15 (1.21–3.84)		5.31 (1.85–14.48)	
3–4	131 (9.7)	0.942 (0.081)	2.72 (1.29–5.76)		4.16 (1.19–14.97)	

Abbreviations: AIMs, ancestry informative markers; Pca, prostate cancer; SNP, single-nucleotide polymorphism; WAA, West African ancestry.

[a]Kruskal–Wallis test $p < .001$.

[b]p-Value for linear tread was estimated treating risk allele categories as an ordinal variable.

[c]n = 1,355, one individual was removed due to one missing genotype. Logistic regression model adjusting for age and global WAA estimated using 1,911 AIMs across the genome.

[d]n = 1,356, logistic regression model adjusting for age, global WAA, and 8q24 local WAA estimated using 70 AIMs on 8q24 region.

medical, and genetic information all together can be used for clinical decision-making or genetic counseling. For example, dosage of a commonly used anticoagulant, warfarin, varies based on patients' age, race/ethnicity, genetic variation, co-medications, and other factors. Studies have shown that race/ethnicity stratified dosage algorithms, at times incorporating population-specific variants, is necessary to accurately determine appropriate warfarin dosage (Bress et al., 2012; Johnson et al., 2017; Perera et al., 2013).

Given the role of genetic variation on drug efficacy and response, it should not be surprising that a relationship may exist between ancestral genetic background and therapeutic response (Bress, Kittles, Wing, Hooker Jr., & King, 2015). Naltrexone is a U.S. Food and Drug Administration approved drug for opioid and alcohol dependence. Differences in efficacy of this drug for treatment of alcohol dependence and smoking cessation have been demonstrated between European Americans and African Americans (Bress et al., 2015; Ray & Oslin, 2009). When AAs in a randomized placebo-controlled clinical trial of naltrexone for smoking cessation were stratified based on high and low West African ancestry, the naltrexone treatment group had a higher quit rate compared to the placebo group among AAs with low West African ancestry, while there was no difference in quite rate among AAs with high West African ancestry. Allele frequency differences of functional variants in genes that naltrexone targets may explain the observed difference in efficacy. Future studies are warranted to identify variants that explain the difference in efficacy.

Polygenic risk scores (PRSs) calculated from GWAS identified risk variants also have many clinical utilities, including prediction of risk, age of onset, disease subtype, and disease prognosis as well as risk stratification for population based screening (Lambert, Abraham, & Inouye, 2019). However, heterogeneity in genetic associations between European and non-European populations exists due to difference in LD and allele frequency, population-specific risk variants (or different causative variants among populations), and other factors, including differences in study design and environmental exposure (Ioannidis, Thomas, & Daly, 2009; Sirugo, Williams, & Tishkoff, 2019). Studies in non-European groups often fail to replicate variants identified in the GWAS or show decreased effect size in non-European populations (Wojcik et al., 2019). Thus, the PRS calculated using SNPs identified in mainly GWAS in European populations may not be transferrable to other racial/ethnic groups. Accuracy of prediction is lower in non-European populations, and PRS is not as useful in understudied populations as in European populations (Kim, Patel, Teng, Berens, & Lachance, 2018; Martin et al., 2019). For example, European American and Hispanic/Latino American men have a greater increase in Pca risk with increasing PRS compared to African American and East Asian American men (Hoffmann et al., 2015). Within the GWAS identified genomic regions in European populations, different variants are more strongly associated with Pca in African Americans (Haiman et al., 2011a). When using variants identified in African Americans, Pca risk prediction in African American men improved (Haiman et al., 2011a; Han et al., 2014). Before PRS is used in clinical settings in racially/ethnically diverse populations, inclusion of diverse populations in genomics studies and use of a population-specific (or ancestry-specific) PRS is necessary.

The studies that utilize array-based and imputed genotype data relies on LD between genotyped and causal variants and is limited by the populations included in the array developments and in the reference panels used for imputation (Need & Goldstein, 2009). On the other hand, whole genome sequencing is an unbiased approach to discover disease associated variants. Whole genome sequencing approach will be available at lower cost in the near future. There are also advancements in statistical approaches to identify variants in large population-based studies. A new approach was developed to account for relatedness in admixed populations and reduce chance of false positive associations due to underlying population structure (Conomos, Miller, & Thornton, 2015; Conomos, Reiner, Weir, & Thornton, 2016; Lin et al., 2014). Moreover, identity-by-descent mapping approaches can be used in population-based studies to detect smaller genomic regions of ancestry associated with disease traits than the admixed LD approach used in admixture mapping and estimates of local genomic ancestry (Liu, Fazio, Hu, & Paterson, 2016).

Finally, racial/ethnic minority groups are underrepresented in genomics studies (Bustamante, Burchard, & de la Vega, 2011; Popejoy & Fullerton, 2016; Spratt, Chan, Waldron, et al., 2016), clinical trials (Chen Jr, Lara, Dang, Paterniti, & Kelly, 2014), biobanks (Cohn, Hamilton, Larson, & Williams, 2017), and National Institute of Health funded studies (Burchard, Oh, Foreman, & Celedón, 2015). Underrepresentation of racial/ethnic minority groups has severely hindered our progress in understanding the genetic architecture of complex traits and causes and risk factors of diseases in various racial/ethnic groups. Underrepresentation in research studies and the lack of genomic technology availability in racial/ethnic minority groups (e.g., lack of appropriate genetic testing for racial/ethnic minority groups) may widen health disparities (Smith et al., 2016). The National Human Genome Research Institute has recognized the importance of prioritizing diversity in genomics medicine and initiated several research programs, such as 1000 Genomes Project (Hindorff et al., 2018). Recently, the "All of US" research program, aims to further increase diversity of study participants and improve our understanding of health disparities from conception of the program (Denny et al., 2019). The program works collaboratively with community partiers to engage diverse populations. The enrolled participants can donate blood, urine, and/or saliva samples for biobanking. Increasing study participations from racial/ethnic minority groups and biobanking in these national research programs likely will provide an infrastructure for future genomics-based health disparities research.

In conclusion, research to understand human genomic variation and its implications in health has a great potential to reduce health disparities and move toward building health equity. Biological anthropology can play important roles in genomics and health disparities research by using a biocultural approach or holistic approach. Population studies of health disparities generally investigate interactions of various factors from multiple levels, including individual, community,

societal, and structural levels. Anthropologists are uniquely trained to understand this complex relationship among various factors influencing health, while appreciating the diversity that exist across and within human populations rather than treating racial/ethnic grouping as a surrogate for sociocultural and biological backgrounds.

ACKNOWLEDGMENTS

K. B. is supported by Urology Care Foundation Research Scholar Award, the Partnership for Native American Cancer Prevention (NACP) (funding to University of Arizona Cancer Center, U54CA143924), and a grant from National Institute of Health (1R21CA248361). RAK is supported by a grant from the National Institute of health (R01MD007105).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Ken Batai: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; validation; visualization; writing-original draft. **Stanley Hooker**: Data curation; formal analysis; methodology; writing-original draft. **Rick A. Kittles**: Conceptualization; data curation; funding acquisition; investigation; methodology; supervision; writing-review and editing.

DATA AVAILABILITY STATEMENT

Data used in this manuscript is available upon request.

ORCID

Rick A. Kittles https://orcid.org/0000-0002-5004-4437

REFERENCES

Adedokun, B., Zheng, Y., Ndom, P., Gakwaya, A., Makumbi, T., Zhou, A. Y., … Huo, D. (2020). Prevalence of inherited mutations in breast cancer predisposition genes among women in Uganda and Cameroon. *Cancer Epidemiology, Biomarkers & Prevention*, *29*(2), 359–367.

Adeyemo, A., & Rotimi, C. (2010). Genetic variants associated with complex human diseases show wide variation across multiple populations. *Public Health Genomics*, *13*(2), 72–79.

Akbari, M. R., Donenberg, T., Lunn, J., Curling, D., Turnquest, T., Krill-Jackson, E., … Hurley, J. (2014). The spectrum of BRCA1 and BRCA2 mutations in breast cancer patients in The Bahamas. *Clinical Genetics*, *85*(1), 64–67.

Al-Alem, U., Rauscher, G., Shah, E., Batai, K., Mahmoud, A., Beisner, E., … Kittles, R. (2014). Association of genetic ancestry with breast cancer in ethnically diverse women from Chicago. *PLoS One*, *9*(11), e112916.

Alvidrez, J., Castille, D., Laude-Sharp, M., Rosario, A., & Tabor, D. (2019). The National Institute on Minority Health and Health Disparities research framework. *American Journal of Public Health*, *109*(S1), S16–S20.

Baharian, S., Barakatt, M., Gignoux, C. R., Shringarpure, S., Errington, J., Blot, W. J., … Gravel, S. (2016). The great migration and African-American genomic diversity. *PLoS Genetics*, *12*(5), e1006059.

Batai, K., Harb-de la Rosa, A., Zeng, J., Chipollini, J. J., Gachupin, F. C., & Lee, B. R. (2019). Racial/ethnic disparities in renal cell carcinoma: Increased risk of early-onset and variation in histologic subtypes. *Cancer Medicine*, *8*(15), 6780–6788.

Bensen, J. T., Xu, Z., McKeigue, P. M., Smith, G. J., Fontham, E. T. H., Mohler, J. L., & Taylor, J. A. (2014). Admixture mapping of prostate cancer in African Americans participating in the North Carolina-Louisiana Prostate Cancer Project (PCaP). *Prostate*, *74*(1), 1–9.

Bergström, A., McCarthy, S. A., Hui, R., Almarri, M. A., Ayub, Q., Danecek, P., Tyler-Smith, C. (2020). Insights into human genetic variation and population history from 929 diverse genomes. *Science*, *367*(6484), eaay5012.

Bonham, V. L., Green, E. D., & Pérez-Stable, E. J. (2018). Examining how race, ethnicity, and ancestry data are used in biomedical research. *JAMA*, *320*(15), 1533–1534.

Bress, A., Kittles, R., Wing, C., Hooker, S. E., Jr., & King, A. (2015). Genetic ancestry as an effect modifier of naltrexone in smoking cessation among African Americans: An analysis of a randomized controlled trial. *Pharmacogenetics and Genomics*, *25*(6), 305–312.

Bress, A., Patel, S. R., Perera, M. A., Campbell, R. T., Kittles, R. A., & Cavallari, L. H. (2012). Effect of NQO1 and CYP4F2 genotypes on warfarin dose requirements in Hispanic-Americans and African-Americans. *Pharmacogenomics*, *13*(16), 1925–1935.

Brewster, A. M., Chavez-MacGregor, M., & Brown, P. (2014). Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry. *The Lancet Oncology*, *15*(13), e625–e634.

Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D., & Mountain, J. L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *American Journal of Human Genetics*, *96*(1), 37–53.

Bryc, K., Velez, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., … Ostrer, H. (2010). Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(Suppl 2), 8954–8961.

Burchard, E. G., Oh, S. S., Foreman, M. G., & Celedón, J. C. (2015). Moving toward true inclusion of racial/ethnic minorities in federally funded studies. A key step for achieving respiratory health equality in the United States. *American Journal of Respiratory and Critical Care Medicine*, *191*(5), 514–521.

Burchard, E. G., Ziv, E., Coyle, N., Gomez, S. L., Tang, H., Karter, A. J., … Risch, N. (2003). The importance of race and ethnic background in biomedical research and clinical practice. *The New England Journal of Medicine*, *348*(12), 1170–1175.

Bustamante, C. D., Burchard, E. G., & de la Vega, F. M. (2011). Genomics for the world. *Nature*, *475*(7355), 163–165.

Campbell, M. C., Hirbo, J. B., Townsend, J. P., & Tishkoff, S. A. (2014). The peopling of the African continent and the diaspora into the new world. *Current Opinion in Genetics & Development*, *29*, 120–132.

Casto, A. M., & Feldman, M. W. (2011). Genome-wide association study SNPs in the Human Genome Diversity Project Populations: Does selection affect unlinked SNPs with shared trait associations? *PLoS Genetics*, *7*(1), e1001266.

Caulfield, T., Fullerton, S., Ali-Khan, S., Arbour, L., Burchard, E., Cooper, R., et al. (2009). Race and ancestry in biomedical research: Exploring the challenges. *Genome Medicine*, *1*(1), 8.

Cavalli-Sforza, L. L., Menozzi, R., & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Chang, B.-L., Spangler, E., Gallagher, S., Haiman, C. A., Henderson, B., Isaacs, W., … Rebbeck, T. R. (2011). Validation of genome-wide prostate cancer associations in men of African descent. *Cancer Epidemiology, Biomarkers & Prevention*, *20*(1), 23–32.

Chen, M. S., Jr., Lara, P. N., Dang, J. H. T., Paterniti, D. A., & Kelly, K. (2014). Twenty years post-NIH Revitalization Act: Enhancing minority participation in clinical trials (EMPaCT): Laying the groundwork for improving minority clinical trial accrual. *Cancer*, *120*(S7), 1091–1096.

Cheng, C.-Y., Reich, D., Haiman, C. A., Tandon, A., Patterson, N., Elizabeth, S., … Kao, W. H. L. (2012). African ancestry and its correlation to type 2 diabetes in African Americans: A genetic admixture analysis in three U.S. population cohorts. *PLoS One*, *7*(3), e32840.

Cheng, I., Chen, G. K., Nakagawa, H., He, J., Wan, P., Laurie, C. C., … Haiman, C. A. (2012). Evaluating genetic risk for prostate cancer among Japanese and Latinos. *Cancer Epidemiology, Biomarkers & Prevention*, *21*(11), 2048–2058.

Chinegwundoh, F., Enver, M., Lee, A., Nargund, V., Oliver, T., & Ben-Shlomo, Y. (2006). Risk and presenting features of prostate cancer amongst African-Caribbean, South Asian and European men in North-East London. *BJU International*, *98*(6), 1216–1220.

Cohn, E. G., Hamilton, N., Larson, E. L., & Williams, J. K. (2017). Self-reported race and ethnicity of US biobank participants compared to the US census. *Journal of Community Genetics*, *8*(3), 229–238.

Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *The New England Journal of Medicine*, *372*(9), 793–795.

Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genetic Epidemiology*, *39*(4), 276–293.

Conomos, M. P., Reiner, A. P., Weir, B. S., & Thornton, T. A. (2016). Model-free estimation of recent genetic relatedness. *American Journal of Human Genetics*, *98*(1), 127–148.

Conti, D. V., Wang, K., Sheng, X., Bensen, J. T., Hazelett, D. J., Cook, M. B., … for the PRACTICAL/ELLIPSE Consortium. (2017). Two novel susceptibility loci for prostate cancer in men of African ancestry. *Journal of the National Cancer Institute*, *109*(8) 1–5.

Cunningham, T. J., Croft, J. B., Liu, Y., Lu, H., Eke, P. I., & Giles, W. H. (2017). Vital signs: Racial disparities in age-specific mortality among Blacks or African Americans—United States, 1999-2015. *MMWR. Morbidity and Mortality Weekly Report*, *66*(17), 444–456.

Dai, C. L., Vazifeh, M. M., Yeang, C. H., Tachet, R., Wells, R. S., Vilar, M. G., … Martin, A. R. (2020). Population histories of the United States revealed through fine-scale migration and haplotype analysis. *American Journal of Human Genetics*, *106*(3), 371–388.

Denny, J. C., Rutter, J. L., Goldstein, D. B., Philippakis, A., Smoller, J. W., Jenkins, G., & Dishman, E. (2019). The "All of Us" research program. *The New England Journal of Medicine*, *381*(7), 668–676.

Der, E. M., Gyasi, R. K., Tettey, Y., Edusei, L., Bayor, M. T., Jiagge, E., … Newman, L. A. (2015). Triple-negative breast cancer in Ghanaian women: The Korle Bu teaching hospital experience. *The Breast Journal*, *21*(6), 627–633.

di Rienzo, A., & Hudson, R. R. (2005). An evolutionary framework for common diseases: The ancestral-susceptibility model. *Trends in Genetics*, *21*(11), 596–601.

Dietze, E. C., Sistrunk, C., Miranda-Carboni, G., O'Regan, R., & Seewaldt, V. L. (2015). Triple-negative breast cancer in African-American women: Disparities versus biology. *Nature Reviews. Cancer*, *15*(4), 248–254.

Drake, K. A., Galanter, J. M., & Burchard, E. G. (2008). Race, ethnicity and social class and the complex etiologies of asthma. *Pharmacogenomics*, *9*(4), 453–462.

Eeles, R. A., Kote-Jarai, Z., Giles, G. G., Olama, A. A. A., Guy, M., Jugurnauth, S. K., et al. (2008). Multiple newly identified loci associated with prostate cancer susceptibility. *Nature Genetics*, *40*(3), 316–321.

Fackenthal, J. D., Zhang, J., Zhang, B., Zheng, Y., Hagos, F., Burrill, D. R., … Olopade, O. I. (2012). High prevalence of BRCA1 and BRCA2 mutations in unselected Nigerian breast cancer patients. *International Journal of Cancer*, *131*(5), 1114–1123.

Fejerman, L., Ahmadiyeh, N., Hu, D., Huntsman, S., Beckman, K. B., Caswell, J. L., et al. (2014). Genome-wide association study of breast cancer in Latinas identifies novel protective variants on 6q25. *Nature Communications*, *5*, 5260.

Fejerman, L., Chen, G. K., Eng, C., Huntsman, S., Hu, D., Williams, A., … Ziv, E. (2012). Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Human Molecular Genetics*, *21*(8), 1907–1917.

Fejerman, L., Hu, D., Huntsman, S., John, E. M., Stern, M. C., Haiman, C. A., … Ziv, E. (2013). Genetic ancestry and risk of mortality among U.S. Latinas with breast cancer. *Cancer Research*, *73*(24), 7243–7253.

Fejerman, L., John, E. M., Huntsman, S., Beckman, K., Choudhry, S., Perez-Stable, E., … Ziv, E. (2008). Genetic ancestry and risk of breast cancer among U.S. Latinas. *Cancer Research*, *68*(23), 9723–9728.

Fejerman, L., Romieu, I., John, E. M., Lazcano-Ponce, E., Huntsman, S., Beckman, K. B., … Torres-Mejía, G. (2010). European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiology, Biomarkers & Prevention*, *19*(4), 1074–1082.

Florez, J. C., Price, A. L., Campbell, D., Riba, L., Parra, M. V., Yu, F., … Reich, D. (2009). Strong association of socioeconomic status with genetic ancestry in Latinos: Implications for admixture studies of type 2 diabetes. *Diabetologia*, *52*(8), 1528–1536.

Freedman, M. L., Haiman, C. A., Patterson, N., McDonald, G. J., Tandon, A., Waliszewska, A., et al. (2006). Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(38), 14068–14073.

Gebreab, S. Y., Riestra, P., Khan, R. J., Xu, R., Musani, S. K., Tekola-Ayele, F., … Davis, S. K. (2015). Genetic ancestry is associated with measures of subclinical atherosclerosis in African Americans: The Jackson Heart Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *35*(5), 1271–1278.

Gee, G. C., Hing, A., Mohammed, S., Tabor, D. C., & Williams, D. R. (2019). Racism and the life course: Taking time seriously. *American Journal of Public Health*, *109*(S1), S43–s47.

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., … Pollak, M. R. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science*, *329*(5993), 841–845.

Gomez, F., Hirbo, J., & Tishkoff, S. A. (2014). Genetic variation and adaptation in Africa: Implications for human evolution and disease. *Cold Spring Harbor Perspectives in Biology*, *6*(7), a008524–a008524.

Green, E. D., Guyer, M. S., & National Human Genome Reseasrch Institute. (2011). Charting a course for genomic medicine from base pairs to bedside. *Nature*, *470*(7333), 204–213.

Haiman, C. A., Chen, G. K., Blot, W. J., Strom, S. S., Berndt, S. I., Kittles, R. A., … Henderson, B. E. (2011a). Characterizing genetic risk at known prostate cancer susceptibility loci in African Americans. *PLoS Genetics*, *7*(5), e1001387.

Haiman, C. A., Chen, G. K., Blot, W. J., Strom, S. S., Berndt, S. I., Kittles, R. A., … Henderson, B. E. (2011b). Genome-wide association study of prostate cancer in men of African ancestry identifies a susceptibility locus at 17q21. *Nature Genetics*, *43*(6), 570–573.

Haiman, C. A., Patterson, N., Freedman, M. L., Myers, S. R., Pike, M. C., Waliszewska, A., et al. (2007). Multiple regions within 8q24 independently affect risk for prostate cancer. *Nature Genetics*, *39*(5), 638–644.

Halder, I., & Shriver, M. D. (2003). Measuring and using admixture to study the genetics of complex diseases. *Human Genomics*, *1*(1), 52–62.

Hales, C. M., Carroll, M. D., Fryar, C. D., & Ogden, C. L. (2020). Prevalence of obesity and severe obesity among adults: United States, 2017-2018. *NCHS Data Brief*, *288*, 1–8.

Han, Y., Rand, K. A., Hazelett, D. J., Ingles, S. A., Kittles, R. A., Strom, S. S., … Haiman, C. A. (2016). Prostate cancer susceptibility in men of African ancestry at 8q24. *Journal of the National Cancer Institute*, *108*(7), djv431.

Han, Y., Signorello, L. B., Strom, S. S., Kittles, R. A., Rybicki, B. A., Stanford, J. L., et al. (2014). Generalizability of established prostate cancer risk variants in men of African ancestry. *International Journal of Cancer*, *136*, 1210–1217.

Hancock, A. M., Witonsky, D. B., Gordon, A. S., Eshel, G., Pritchard, J. K., Coop, G., & Di Rienzo, A. (2008). Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genetics*, *4*(2), e32.

Harburg, E., Gleibermann, L., Roeper, P., Schork, M. A., & Schull, W. J. (1978). Skin color, ethnicity, and blood pressure I: Detroit blacks. *American Journal of Public Health*, *68*(12), 1177–1183.

Hindorff, L. A., Bonham, V. L., Brody, L. C., Ginoza, M. E. C., Hutter, C. M., Manolio, T. A., & Green, E. D. (2018). Prioritizing diversity in human genomics research. *Nature Reviews. Genetics*, *19*(3), 175–185.

Hoffmann, T. J., van Den Eeden, S. K., Sakoda, L. C., Jorgenson, E., Habel, L. A., Graff, R. E., et al. (2015). A large multiethnic genome-wide association study of prostate cancer identifies novel risk variants and substantial ethnic differences. *Cancer Discovery*, *5*(8), 878–891.

Huo, D., Feng, Y., Haddad, S., Zheng, Y., Yao, S., Han, Y.-J., et al. (2016). Genome-wide association studies in women of African ancestry identified 3q26.21 as a novel susceptibility locus for oestrogen receptor negative breast cancer. *Human Molecular Genetics*, *25*(21), 4835–4846.

Ioannidis, J. P. A., Thomas, G., & Daly, M. J. (2009). Validating, augmenting and refining genome-wide association signals. *Nature Reviews. Genetics*, *10*(5), 318–329.

Iqbal, J., Ginsburg, O., Rochon, P. A., Sun, P., & Narod, S. A. (2015). Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *JAMA*, *313*(2), 165–173.

Jemal, A., Ward, E. M., Johnson, C. J., Cronin, K. A., Ma, J., Ryerson, A. B., et al. (2017). Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *Journal of the National Cancer Institute*, *109*(9), djx030.

Johnson, J. A., Caudle, K. E., Gong, L., Whirl-Carrillo, M., Stein, C. M., Scott, S. A., … Wadelius, M. (2017). Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clinical Pharmacology and Therapeutics*, *102*(3), 397–404.

Kim, M. S., Patel, K. P., Teng, A. K., Berens, A. J., & Lachance, J. (2018). Genetic disease risks can be misestimated across global populations. *Genome Biology*, *19*(1), 179.

Kittles, R. A., Santos, E. R., Oji-Njideka, N., & Bonilla, C. (2007). Race, skin color and genetic ancestry: Implications for biomedical research on health disparities. *Californian Journal of Health Promotion*, *5*, 9–23.

Klag, M. J., Whelton, P. K., Coresh, J., Grim, C. E., & Kuller, L. H. (1991). The association of skin color with blood pressure in us blacks with low socioeconomic status. *JAMA*, *265*(5), 599–602.

Klimentidis, Y. C., Arora, A., Zhou, J., Kittles, R., & Allison, D. B. (2016). The genetic contribution of west-African ancestry to protection against central obesity in African-American men but not women: Results from the ARIC and MESA studies. *Frontiers in Genetics*, *7*(89), 1–6.

Kosoy, R., Qi, L., Nassir, R., Garcia, L., Allison, M., Shigeta, R., … Seldin, M. F. (2012). Relationship between hypertension and admixture in post-menopausal African American and Hispanic American women. *Journal of Human Hypertension*, *26*(6), 365–373.

Krieger, N. (2012). Methods for the scientific study of discrimination and health: An ecosocial approach. *American Journal of Public Health*, *102*(5), 936–944.

Lachance, J., Berens, A. J., Hansen, M. E. B., Teng, A. K., Tishkoff, S. A., & Rebbeck, T. R. (2018). Genetic hitchhiking and population bottlenecks contribute to prostate cancer disparities in men of African descent. *Cancer Research*, *78*(9), 2432–2443.

Lambert, S. A., Abraham, G., & Inouye, M. (2019). Towards clinical utility of polygenic risk scores. *Human Molecular Genetics*, *28*(R2), R133–R142.

Langefeld, C. D., Comeau, M. E., Ng, M. C. Y., Guan, M., Dimitrov, L., Mudgal, P., … Freedman, B. I. (2018). Genome-wide association studies suggest that APOL1-environment interactions more likely trigger kidney disease in African Americans with nondiabetic nephropathy than strong APOL1-second gene interactions. *Kidney International*, *94*(3), 599–607.

LaVeist, T. A. (1994). Beyond dummy variables and sample selection: What health services researchers ought to know about race as a variable. *Health Services Research*, *29*(1), 1–16.

Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., … Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, *319*(5866), 1100–1104.

Lin, B. M., Nadkarni, G. N., Tao, R., Graff, M., Fornage, M., Buyske, S., … Franceschini, N. (2019). Genetics of chronic kidney disease stages across ancestries: The PAGE study. *Frontiers in Genetics*, *10*, 494.

Lin, D. Y., Tao, R., Kalsbeek, W. D., Zeng, D., Gonzalez, F., II, Fernández-Rhodes, L., … Heiss, G. (2014). Genetic association analysis under complex survey sampling: The Hispanic community health study/study of Latinos. *American Journal of Human Genetics*, *95*(6), 675–688.

Liu, X. Q., Fazio, J., Hu, P., & Paterson, A. D. (2016). Identity-by-descent mapping for diastolic blood pressure in unrelated Mexican Americans. *BMC Proceedings*, *10*(Suppl 7), 263–267.

Livingstone, F. B. (1962). On the non-existence of human races. *Current Anthropology*, *3*, 279–281.

Lohmueller, K. E., Mauney, M. M., Reich, D., & Braverman, J. M. (2006). Variants associated with common disease are not unusually differentiated in frequency across populations. *American Journal of Human Genetics*, *78*(1), 130–136.

Marigorta, U., Lao, O., Casals, F., Calafell, F., Morcillo-Suarez, C., Faria, R., et al. (2011). Recent human evolution has shaped geographical differences in susceptibility to disease. *BMC Genomics*, *12*(1), 55.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*(4), 584–591.

Murabito, J. M., Rosenberg, C. L., Finger, D., Kreger, B. E., Levy, D., Splansky, G. L., … Hwang, S. J. (2007). A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham heart study. *BMC Medical Genetics*, *1*(Suppl 1), S6.

Murphy, A. B., Ukoli, F., Freeman, V., Bennett, F., Aiken, W., Tulloch, T., … Kittles, R. A. (2012). 8q24 risk alleles in West African and Caribbean men. *The Prostate*, *72*(12), 1366–1373.

Myles, S., Davison, D., Barrett, J., Stoneking, M., & Timpson, N. (2008). Worldwide population differentiation at disease-associated SNPs. *BMC Medical Genomics*, *1*(1), 22.

Need, A. C., & Goldstein, D. B. (2009). Next generation disparities in human genomics: Concerns and remedies. *Trends in Genetics*, *25*(11), 489–494.

Newman, L. A., & Kaljee, L. M. (2017). Health disparities and triple-negative breast cancer in African American women: A review. *JAMA Surgery*, *152*(5), 485–493.

Noel, S. E., Arevalo, S., Smith, C. E., Lai, C.-Q., Dawson-Hughes, B., Ordovas, J. M., & Tucker, K. L. (2017). Genetic admixture and body composition in Puerto Rican adults from the Boston Puerto Rican Osteoporosis Study. *Journal of Bone and Mineral Metabolism*, *35*(4), 448–455.

Non, A. L., Gravlee, C. C., & Mulligan, C. J. (2012). Education, genetic ancestry, and blood pressure in African Americans and Whites. *American Journal of Public Health*, *102*(8), 1559–1565.

Novembre, J., & di Rienzo, A. (2009). Spatial patterns of variation due to natural selection in humans. *Nature Reviews. Genetics*, *10*(11), 745–755.

Novembre, J., & Peter, B. M. (2016). Recent advances in the study of fine-scale population structure in humans. *Current Opinion in Genetics & Development*, *41*, 98–105.

Odedina, F., Akinremi, T., Chinegwundoh, F., Roberts, R., Yu, D., Reams, R. R., … Kumar, N. (2009). Prostate cancer disparities in black men of African descent: A comparative literature review of prostate cancer burden among black men in the United States, Caribbean, United Kingdom, and West Africa. *Infectious Agents and Cancer*, *4*(Suppl 1), S2.

Okobia, M. N., Zmuda, J. M., Ferrell, R. E., Patrick, A. L., & Bunker, C. H. (2011). Chromosome 8q24 variants are associated with prostate cancer risk in a high risk population of African ancestry. *Prostate*, *71*(10), 1054–1063.

Olshan, A. F., Kuo, T.-M., Meyer, A.-M., Nielsen, M. E., Purdue, M. P., & Rathmell, W. K. (2013). Racial difference in histologic subtype of renal cell carcinoma. *Cancer Medicine*, *2*(5), 744–749.

Pal, T., Bonner, D., Cragun, D., Monteiro, A. N. A., Phelan, C., Servais, L., … Vadaparampil, S. T. (2015). A high frequency of BRCA mutations in young black women with breast cancer residing in Florida. *Cancer*, *121* (23), 4173–4180.

Perera, M. A., Cavallari, L. H., Limdi, N. A., Gamazon, E. R., Konkashbaev, A., Daneshjou, R., … Johnson, J. A. (2013). Genetic variants associated with warfarin dose in African-American individuals: A genome-wide association study. *Lancet*, *382*(9894), 790–796.

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature*, *538*(7624), 161–164.

Price, A. L., Tandon, A., Patterson, N., Barnes, K. C., Rafaels, N., Ruczinski, I., … Myers, S. (2009). Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genetics*, *5* (6), e1000519.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Purrington, K. S., Slager, S., Eccles, D., Yannoukakos, D., Fasching, P. A., Miron, P., … Couch, F. J. (2014). Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer. *Carcinogenesis*, *35*(5), 1012–1019.

Qi, L., Nassir, R., Kosoy, R., Garcia, L., Curb, J. D., Tinker, L., … Seldin, M. F. (2012). Relationship between diabetes risk and admixture in postmenopausal African-American and Hispanic-American women. *Diabetologia*, *55*(5), 1329–1337.

Quintana-Murci, L. (2016). Understanding rare and common diseases in the context of human evolution. *Genome Biology*, *17*(1), 225.

Ray, L. A., & Oslin, D. W. (2009). Naltrexone for the treatment of alcohol dependence among African Americans: Results from the COMBINE study. *Drug and Alcohol Dependence*, *105*(3), 256–258.

Robbins, C., Torres, J. B., Hooker, S., Bonilla, C., Hernandez, W., Candreva, A., … Carpten, J. (2007). Confirmation study of prostate cancer risk variants at 8q24 in African Americans identifies a novel risk locus. *Genome Research*, *17*(12), 1717–1722.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic structure of human populations. *Science*, *298*(5602), 2381–2385.

Sankar, P., Cho, M. K., Condit, C. M., Hunt, L. M., Koenig, B., Marshall, P., … Spicer, P. (2004). Genetic research and health disparities. *JAMA: The Journal of the American Medical Association*, *291*(24), 2985–2989.

Seldin, M. F., Pasaniuc, B., & Price, A. L. (2011). New approaches to disease mapping in admixed populations. *Nature Reviews. Genetics*, *12*(8), 523–528.

Serre, D., & Pääbo, S. (2004). Evidence for gradients of human genetic diversity within and among continents. *Genome Research*, *14*(9), 1679–1685.

Shriver, M. D., & Kittles, R. A. (2004). Genetic ancestry and the search for personalized genetic histories. *Nature Reviews. Genetics*, *5*(8), 611–618.

Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, *70*(1), 7–30.

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell*, *177*(1), 26–31.

Smith, C. E., Fullerton, S. M., Dookeran, K. A., Hampel, H., Tin, A., Maruthur, N. M., … Ordovás, J. M. (2016). Using genetic technologies to reduce, rather than widen, health disparities. *Health Affairs (Project Hope)*, *35*(8), 1367–1373.

Spratt, D. E., Chan, T., Waldron, L., et al. (2016). Racial/ethnic disparities in genomic sequencing. *JAMA Oncology*, *2*(8), 1070–1074.

Sweet, E., McDade, T. W., Kiefe, C. I., & Liu, K. (2007). Relationships between skin color, income, and blood pressure among African Americans in the CARDIA Study. *American Journal of Public Health*, *97*(12), 2253–2259.

Tang, H., Coram, M., Wang, P., Zhu, X., & Risch, N. (2006). Reconstructing genetic ancestry blocks in admixed individuals. *American Journal of Human Genetics*, *79*(1), 1–12.

The 1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., de Pristo, M. A., Durbin, R. M., … McVean, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65.

The 1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., et al. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74.

Thomas, G., Jacobs, K. B., Yeager, M., Kraft, P., Wacholder, S., Orr, N., et al. (2008). Multiple loci identified in a genome-wide association study of prostate cancer. *Nature Genetics*, *40*(3), 310–315.

Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., … Williams, S. M. (2009). The genetic structure and history of Africans and African Americans. *Science*, *324*, 1035–1044.

Torres, J. B., Doura, M. B., Keita, S. O. Y., & Kittles, R. A. (2012). Y chromosome lineages in men of west African descent. *PLoS One*, *7*(1), e29687–e29687.

Wang, J. C., Lai, W. R., Traore, E. J., Liu, J., Sholl, A. B., Mandava, S. H., et al. (2017). Racial disparities in histology and short-term renal functional outcomes following robotic nephron-sparing surgery. *Clinical Genitourinary Cancer 15*(2), 203–206.

Wang, M., Takahashi, A., Liu, F., Ye, D., Ding, Q., Qin, C., … Xu, J. (2015). Large-scale association analysis in Asians identifies new susceptibility loci for prostate cancer. *Nature Communications*, *6*, 8469.

Wang, Y., Ray, A. M., Johnson, E. K., Zuhlke, K. A., Cooney, K. A., & Lange, E. M. (2011). Evidence for an association between prostate cancer and chromosome 8q24 and 10q11 genetic variants in African American men: The flint men's health study. *Prostate*, *71*(3), 225–231.

Warnecke, R. B., Oh, A., Breen, N., Gehlert, S., Paskett, E., Tucker, K. L., … Hiatt, R. A. (2008). Approaching health disparities from a population perspective: The National Institutes of Health Centers for Population Health and Health Disparities. *American Journal of Public Health*, *98*(9), 1608–1615.

Wassel, C. L., Pankow, J. S., Peralta, C. A., Choudhry, S., Seldin, M. F., & Arnett, D. K. (2009). Genetic ancestry is associated with subclinical cardiovascular disease in African Americans and Hispanics from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation. Cardiovascular Genetics*, *2*(6), 629–636.

Winkler, C. A., Nelson, G. W., & Smith, M. W. (2010). Admixture mapping comes of age. *Annual Review of Genomics and Human Genetics*, *11*(1), 65–89.

Wojcik, G. L., Graff, M., Nishimura, K. K., Tao, R., Haessler, J., Gignoux, C. R., … Carlson, C. S. (2019). Genetic analyses of diverse populations improves discovery for complex traits. *Nature*, *570*(7762), 514–518.

Xu, J., Kibel, A. S., Hu, J. J., Turner, A. R., Pruett, K., Zheng, S. L., … Isaacs, W. B. (2009). Prostate cancer risk associated loci in African Americans. *Cancer Epidemiology, Biomarkers & Prevention*, *18*(7), 2145–2149.

Xu, J., Mo, Z., Ye, D., Wang, M., Liu, F., Jin, G., et al. (2012). Genome-wide association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4. *Nature Genetics*, *44*(11), 1231–1235.

Xu, Z., Bensen, J. T., Smith, G. J., Mohler, J. L., & Taylor, J. A. (2011). GWAS SNP replication among African American and European American men in the North Carolina–Louisiana prostate cancer project (PCaP). *Prostate*, *71*(8), 881–891.

Ziv, E., John, E. M., Choudhry, S., Kho, J., Lorizio, W., Perez-Stable, E. J., & Burchard, E. G. (2006). Genetic ancestry and risk factors for breast

cancer among Latinas in the San Francisco Bay Area. *Cancer Epidemiology, Biomarkers & Prevention*, *15*(10), 1878–1885.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Batai K, Hooker S, Kittles RA. Leveraging genetic ancestry to study health disparities. *Am J Phys Anthropol*. 2021;175:363–375. https://doi.org/10.1002/ajpa.24144