Effect sizes of causal variants for gene expression and complex traits differ between populations

December 6, 2021

Roshni A. Patel*,[1] Shaila A. Musharoff*,[1,2] Jeffrey P. Spence,[1] Harold Pimentel,[3,4] Catherine Tcheandjieu,[2,5] Hakhamanesh Mostafavi,[1] Nasa Sinnott-Armstrong,[1,2] Shoa L. Clarke,[2,5] Courtney J. Smith,[1] VA Million Veteran Program, Peter P. Durda,[6] Kent D. Taylor,[7] Russell Tracy,[6] Yongmei Liu,[8] Craig W. Johnson,[9] Francois Aguet,[10] Kristin G. Ardlie,[10] Stacey Gabriel,[10] Josh Smith,[11] Deborah A. Nickerson,[11] Stephen S. Rich,[12] Jerome I. Rotter,[7] Philip S. Tsao,[2,5] Themistocles L. Assimes,[2,5] Jonathan K. Pritchard[1,13]

[1]Genetics, Stanford University School of Medicine, Stanford, CA, [2]VA Palo Alto Health Care System, Palo Alto, CA, [3]Computational Medicine, University of California Los Angeles, Los Angeles, CA, [4]Human Genetics, University of California Los Angeles, Los Angeles, CA, [5]Stanford University School of Medicine, Stanford, CA, [6]The Robert Larner, M.D. College of Medicine at The University of Vermont, Burlington, VT, [7]Institute for Translational Genomics and Population Sciences, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, [8]Duke University School of Medicine, Durham, NC, [9]Biostatistics, University of Washington, Seattle, WA, [10]Broad Institute of MIT and Harvard, Cambridge, MA, [11]Genome Sciences, University of Washington, Seattle, WA, [12]Public Health Sciences, University of Virginia, Charlottesville, VA, [13]Biology, Stanford University, Stanford, CA

* These authors contributed equally to this work.

Abstract

Despite the growing number of genome-wide association studies (GWAS) for complex traits, it remains unclear whether effect sizes of causal genetic variants differ between populations. In principle, effect sizes of causal variants could differ between populations due to gene-by-gene or gene-by-environment interactions. However, comparing causal variant effect sizes is challenging: it is difficult to know which variants are causal, and comparisons of variant effect sizes are confounded by differences in linkage disequilibrium (LD) structure between ancestries. Here, we develop a method to assess causal variant effect size differences that overcomes these limitations. Specifically, we leverage the fact that segments of European ancestry shared between European-American and admixed African-American individuals have similar LD structure, allowing for unbiased comparisons of variant effect sizes in European ancestry segments. We apply our method to two types of traits: gene expression and low-density lipoprotein cholesterol (LDL-C). We find that causal variant effect sizes for gene expression are significantly different between European-Americans and African-Americans; for LDL-C, we observe a similar point estimate although this is not significant, likely due to lower statistical power. Cross-population differences in variant effect sizes highlight the role of genetic interactions in trait architecture and will contribute to the poor portability of polygenic scores across populations, reinforcing the importance of conducting GWAS on individuals of diverse ancestries and environments.

Introduction

Human complex traits are influenced by tens to hundreds of thousands of causal variants, each with tiny marginal effect sizes. It remains unclear whether these causal variants have the same effect size in all populations (Brown *et al.*, 2016; Galinsky *et al.*, 2019; Mostafavi *et al.*, 2020). Causal variants could have different effect sizes between populations with different ancestries or environments, possibly due to epistatic (gene-by-gene) or gene-by-environment interactions. However, comparing causal variant effect sizes between populations is challenging. The causal variants underlying human complex traits are generally unknown and instead, genome-wide association studies (GWAS) typically identify single nucleotide polymorphisms (SNPs) that are statistically associated with the trait due to strong linkage disequilibrium (LD) with the causal variant(s). Due to differences in LD structure, these trait-associated SNPs may not be equally correlated with the same causal variant in two different populations, resulting in different marginal effect sizes. This is especially true if the causal variant is private, or only present in a single

population. Thus, although several studies have observed differences between populations in the marginal effect sizes of trait-associated SNPs (de Candia *et al.*, 2013; Mancuso *et al.*, 2016; Wojcik *et al.*, 2019; Bitarello and Mathieson, 2020), this could correspond both to differences in the effect sizes of causal variants themselves or to differences in LD structure.

Polygenic scores (PGS) are one application for which it is particularly important to understand whether the true effect sizes of causal variants differ between populations. PGS aim to predict the genetic component of an individual's phenotype and have recently been used to successfully predict risk for human complex diseases including type I diabetes, breast cancer, and coronary artery disease (Khera *et al.*, 2018; Mavaddat *et al.*, 2019; Sharp *et al.*, 2019). However, PGS constructed in one population have reduced accuracy in populations with different ancestries or environments, a problem referred to as poor PGS portability (Martin *et al.*, 2019; Berg *et al.*, 2019; Sohail *et al.*, 2019; Mostafavi *et al.*, 2020). Poor PGS portability could be partially due to differences in LD structure between populations, but recent work suggests there are additional factors responsible (Bitarello and Mathieson, 2020; Wang *et al.*, 2020). One of these additional factors could be that causal variant effect sizes differ between populations. Thus, understanding whether causal variant effect sizes differ between populations is critical for improving PGS portability in diverse populations and ensuring the equitable application of PGS.

Previous work comparing causal variant effect sizes between human populations has leveraged LD reference panels to account for differences in LD structure between populations (Brown *et al.*, 2016; Galinsky *et al.*, 2019). These studies found modest differences in causal variant effect sizes for both gene expression and complex traits. However, these existing methods are limited by their reliance on accurate LD reference panels and their difficulty in accounting for rare or population-specific causal variants. Furthermore, these methods are not suitable for application to recently admixed populations such as African-Americans and Latin Americans due to the complexities of long-range admixture LD.

In this paper, we compare the genetic architecture of gene expression and low-density lipoprotein cholesterol (LDL-C) between European-Americans and African-Americans. We first compare the marginal effect sizes of trait-associated SNPs for gene expression and LDL-C when estimated from European-Americans and African-Americans. We next quantify the contribution of local ancestry and local ancestry-by-genotype interactions to phenotypic variance. Lastly, we leverage the multiple ancestries in the genomes of admixed populations to test whether causal variant effect sizes differ between populations. Admixed African-American genomes contain regions of European ancestry that share the same LD structure as the genomes of European-Americans.
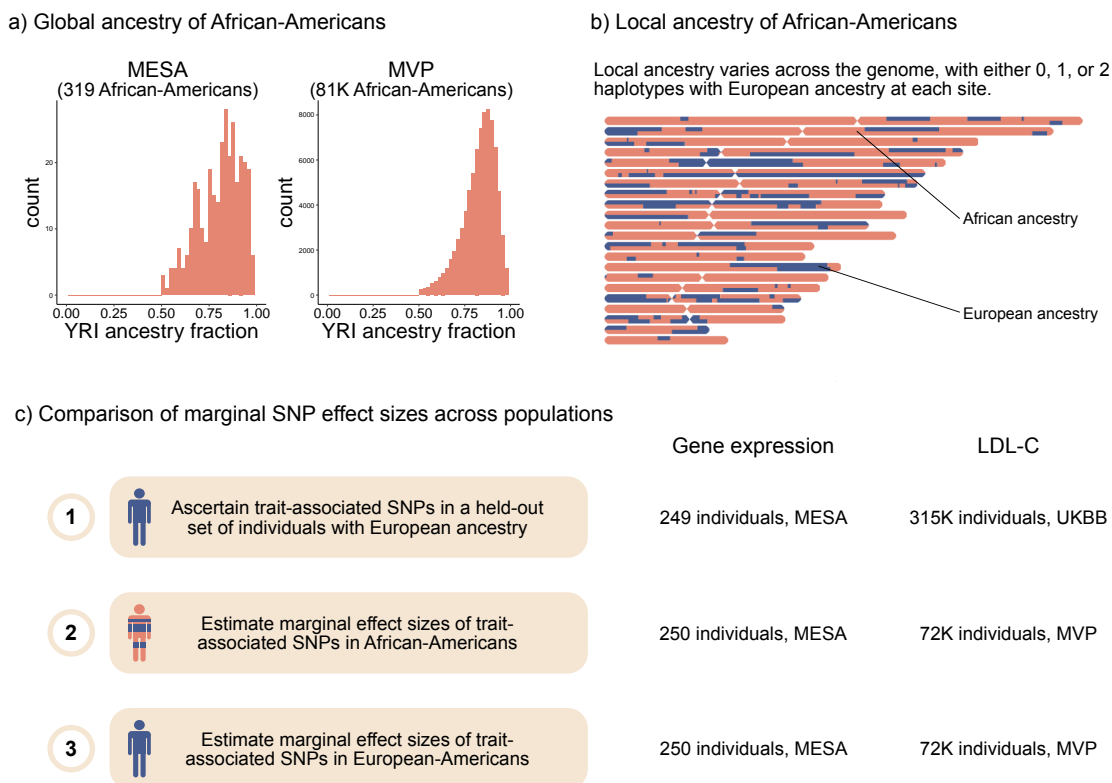
Figure 1: **Schematic of the analysis pipeline. a)** Global ancestry of African-Americans is predominantly African, with an average global African ancestry fraction of 0.79 in MESA and 0.82 in MVP. **b)** Local ancestry for one sample individual in MESA. Individuals have either 0, 1, or 2 haplotypes with European ancestry at each position. **c)** We compare marginal effect sizes of SNPs between African-Americans and European-Americans.

Within these regions of shared ancestry, we can therefore compare variant effect sizes between populations without bias from differences in LD structure.

Results

We performed analyses for gene expression and LDL-C, both of which are driven by a combination of genetic factors and environmental factors. We analyzed gene expression using the Multi-Ethnic Study of Atherosclerosis (MESA), a dataset with whole genome sequencing and bulk RNA-Seq in peripheral blood mononuclear cells for 319 African-Americans and 499 European-Americans. We analyzed LDL-C using the Million Veteran Program (MVP) dataset, which has dense SNP genotyping and LDL-C measurements for 87K African-Americans and 319K European-Americans. Of existing human genetic datasets, MESA and MVP have some of the largest cohorts of admixed individuals for their respective phenotypes.

Inferring global and local ancestry. We inferred global and local ancestry for the African-American individuals in MESA and MVP. In both cases, we modeled African-Americans as a two-way admixture between African and European populations that occurred 8 generations ago (Baharian *et al.*, 2016). We estimated global ancestry using supervised ADMIXTURE with 1000 Genomes populations (CEU as European and YRI as African) as our reference populations (Auton *et al.*, 2015; Alexander *et al.*, 2009). The average global African ancestry of African-American individuals is 0.79 in MESA and 0.82 in MVP, concordant with previous estimates from similar populations (Martin *et al.*, 2017) (Fig. 1a). We performed local ancestry inference with RFMix using the same 1000 Genomes reference populations (Maples *et al.*, 2013) (Fig. 1b). As expected based on their admixture history, the local ancestry of African-American individuals alternates between blocks of African and European ancestry along the genome and contains relatively large European blocks (mean size is 15 Mb in MESA, 14 Mb in MVP).

Comparing marginal SNP effect sizes between populations. We first sought to compare marginal effect sizes of trait-associated SNPs when estimated from European-Americans and from African-Americans (Fig. 1c). We expect that marginal effect sizes of trait-associated SNPs will differ between the two populations due to known differences in LD structure between African and European ancestries, as well as potential differences in epistatic or gene-by-environment interactions between the populations. However, the magnitude of this difference in marginal effect sizes remains unclear. The observed magnitude of differences may be inflated by sampling error, particularly if one population has a small sample size. Additionally, effect sizes are usually largest in a discovery sample due to Winner's Curse, which further exacerbates differences between discovery and replication datasets. To minimize these biases, we ascertained trait-associated SNPs in a held-out set of individuals and compared effect sizes in an equal number of African-Americans and European-Americans (Fig 1c).

To compare effect sizes between populations, we log-transformed phenotype measurements for variance stabilization. We did not perform quantile normalization given that phenotypic variance might differ between populations (Musharoff *et al.*, 2018). We ascertained unlinked, trait-associated SNPs in individuals of European ancestry. For gene expression, we restricted our analyses to putatively *cis*-acting variants (i.e. within 100 kb of TSS) because *cis*-acting variants have stronger effects than *trans*-acting variants and are more easily detected in modest sample sizes (GTEx Consortium, 2017; Võsa *et al.*, 2021). In the event that there were multiple SNPs associated with a gene, we chose the most significant SNP for downstream analyses (GTEx Consortium, 2017). We

ascertained trait-associated SNPs (false discovery rate < 0.01) in a held-out subset of 249 European-Americans in MESA, which resulted in 2,697 SNP-gene associations. For LDL-C, we ascertained trait-associated SNPs ($p < 5 \times 10^{-8}$) in 315K UK Biobank White British individuals, and clumped and pruned them, which resulted in 124 trait-associated SNPs. We performed all subsequent analyses on these trait-associated SNPs.

We next estimated the effect sizes of trait-associated SNPs separately in African-Americans (β_{AA}) and European-Americans (β_{EA}). For gene expression, β_{AA} and β_{EA} were each estimated from 250 individuals. For LDL-C, β_{AA} and β_{EA} were each estimated from 72K individuals. For each trait, we compared marginal effect sizes between the two populations by regressing effect sizes estimated from African-Americans ($\hat{\beta}_{AA}$) on effect sizes estimated from European-Americans ($\hat{\beta}_{EA}$) (Fig. 2). We used total least squares (TLS) to perform the regression because it is robust to statistical noise in the independent variable ($\hat{\beta}_{EA}$), while ordinary least squares is not. Because we used the same number of samples to estimate β_{AA} and β_{EA}, their standard errors will be comparable, as is necessary for TLS regression. We constructed 95% confidence intervals for the slope by bootstrapping over SNPs and concluded that marginal SNP effect sizes are significantly different if this confidence interval (CI) does not contain 1.

For gene expression, effect sizes estimated from African-Americans are significantly smaller in magnitude than the corresponding effect sizes estimated from European-Americans, with a slope of 0.84 (95% CI of 0.81-0.87) (Fig. 2). For LDL-C, we similarly observe a slope of 0.84, but this is not significantly different from 1 (95% CI of 0.66-1.01), likely due to lower statistical power for this trait. Our observation that marginal effect sizes estimated from African-Americans are smaller in magnitude can be at least partially explained by our ascertainment of trait-associated SNPs in individuals of European ancestry. Blocks of LD structure are smaller in populations of African ancestry than in populations of European ancestry, and the African-Americans in MESA and MVP have a mean African global ancestry of approximately 80%. Thus, the correlation between causal variants and trait-associated SNPs ascertained in European populations will generally be weaker in African-Americans than in European-Americans, meaning that marginal effect sizes estimated from African-Americans will have a smaller magnitude. Potential differences in epistatic and gene-by-environment interactions between populations likely also contribute to the observed differences in marginal effect sizes.

Quantifying role of local ancestry in phenotypic variance. Given that African-Americans are admixed with both African and European ancestries, we next sought to assess the contribution
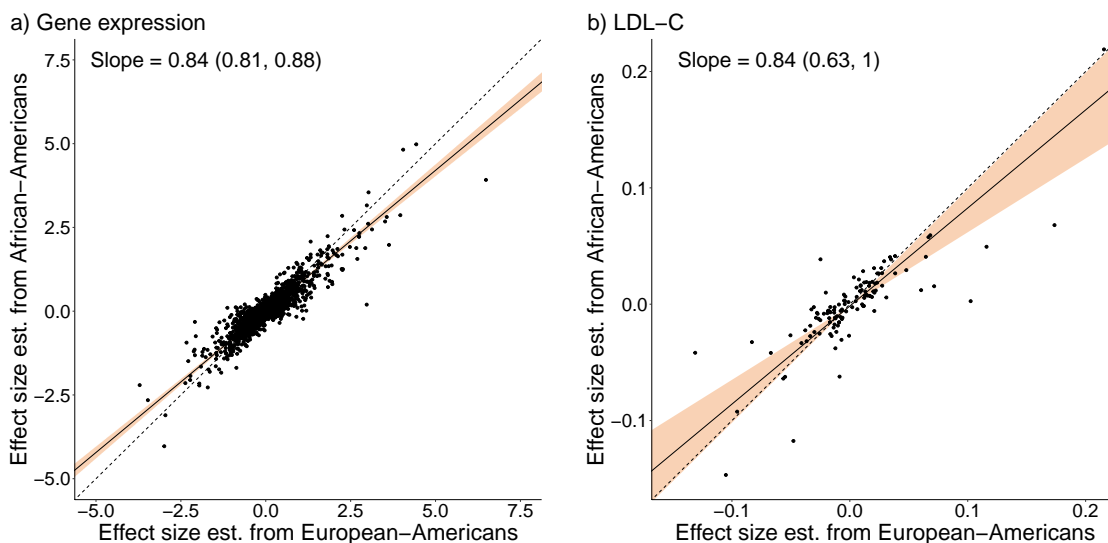
Figure 2: **Comparing marginal SNP effect sizes between populations.** For each trait-associated SNP, we estimated its effect size in African-Americans and European-Americans for both gene expression ($\log_2$ fold change) and LDL-C (mg/dL). For all traits, we regressed the effect size estimated from African-Americans on the effect size estimated from European-Americans. We represent the 95% bootstrap confidence interval with the orange shaded region. Effect sizes estimated from African-Americans are **a)** significantly smaller in magnitude than the corresponding effect sizes estimated from European-Americans for gene expression and **b)** smaller but not significantly so for LDL-C.

	Term added	Model	Total variance explained (%)	
			Gene expression	LDL-C
(1)	Technical covariates	$y_i = c_i\beta_c$	17.08	0.10
(2)	Race	$y_i = c_i\beta_c + r_i\beta_r$	19.51	0.13
(3)	Global ancestry & PCs	$y_i = c_i\beta_c + r_i\beta_r + \theta_i\beta_\theta$	20.31	0.16
(4)	Local ancestry	$y_i = c_i\beta_c + r_i\beta_r + \theta_i\beta_\theta + \gamma_i\beta_\gamma$	20.71	0.41
(5)	Genotype	$y_i = c_i\beta_c + r_i\beta_r + \theta_i\beta_\theta + \gamma_i\beta_\gamma + g_i\beta_g$	26.05	2.47
(6)	Local ancestry-by-genotype interaction	$y_i = c_i\beta_c + r_i\beta_r + \theta_i\beta_\theta + \gamma_i\beta_\gamma + g_{i,A}\beta_A + g_{i,E}\beta_E$	26.49	2.72

Table 1: **Quantifying role of local ancestry in phenotypic variance.** We constructed a series of models and computed the percentage of phenotypic variance explained. In general, each model adds one term to its predecessor, which we describe in the first column. We report the mean total phenotypic variance explained by each model for gene expression and LDL-C. The variables in the models are defined as follows: c_i is a vector of technical/batch covariates; r_i is a race indicator variable; θ_i is a vector of global African ancestry fraction and principal components; γ_i is a local ancestry covariate that counts the number of haplotypes with African ancestry at trait-associated SNPs; g_i is the genotype at trait-associated SNPs (i.e. number of alternate alleles); $g_{i,A}$ is the number of alternate alleles with African local ancestry and $g_{i,E}$ is the number of alternate alleles with European local ancestry.

of local ancestry to phenotypic variation. Specifically, we considered the role of both a local ancestry covariate term and a local ancestry-by-genotype interaction term. We quantified the contribution of both terms to phenotypic variation by constructing a series of phenotypic models and computing the amount of variance explained by each model. We fit each model to roughly 80% of our data allocated as a training set: 256 African-Americans and 206 European-Americans were in the training set for MESA, and 52K African-Americans and 52K European-Americans were in the training set for MVP. We computed the proportion of phenotypic variance explained by the model in a test set using the remaining 20% of our data. For MESA, the test set contained 63 African-Americans and 51 European-Americans; for MVP, the test set contained 22K African-Americans and 22K European-Americans. For gene expression, we report the average phenotypic variance explained across all genes.

We constructed six phenotypic models in total, where each model has an increasing number of terms relative to its predecessor. Our first phenotypic model (Table 1; Eq. 1) included only technical covariates (sex and batch for gene expression; sex and age for LDL-C) and explains 17.08% of phenotypic variance for gene expression and 0.10% of phenotypic variance for LDL-C. Most of the variance explained by these terms for gene expression is due to batch, which is characteristic of RNA-Seq assays. We next added race, global African ancestry fraction, and principal components to the model (Table 1; Eq. 2, 3). These covariates can capture unmeasured genetic and environmental factors relevant for trait variation. In particular, an indicator variable for race allows for race-specific phenotypic intercepts, which can capture trait-relevant differences in environment between African-American and European-American populations (Price *et al.*, 2008). Meanwhile, global African ancestry fraction and principal components capture additional population structure: global African ancestry fraction stratifies African-Americans, while the principal components we include (PC2 for gene expression and PC1 for LDL-C) stratify European-Americans (Supp. Fig. 3). In the context of gene expression, global ancestry and genotype principal components are known to be relevant for trait variation, potentially because they capture the effect of *trans* genetic variation on expression (Nédélec *et al.*, 2016; Randolph *et al.*, 2020; Price *et al.*, 2008). Compared to a model that only includes technical covariates, including race explains an additional 2.43% of variance in gene expression and 0.03% of variance in LDL-C. Including global ancestry and principal components explains an additional 0.81% of variance in gene expression and 0.03% of variance in LDL-C.

We next considered the importance of a local ancestry covariate that measures the number of haplotypes with African ancestry at each trait-associated SNP (Table 1; Eq. 4). Local ancestry

implicitly captures the effect of local genetic variation from SNPs which are not explicitly modeled; in the context of gene expression, these unmodeled, trait-associated SNPs are likely *cis*-acting variants. We find that including local ancestry explains an additional 0.39% of variance in gene expression and 0.25% of variance in LDL-C. For gene expression, local ancestry explains nearly half the variance explained by global ancestry and principal components, and for LDL-C, it explains nearly two-thirds.

Lastly, we considered the importance of a local ancestry-by-genotype interaction term (Table 1; Eq. 6). We quantified its contribution to phenotypic variation relative to a model which includes genotypes for trait-associated SNPs without an interaction with local ancestry (Table 1; Eq. 5). Trait-associated SNPs contribute considerably to trait variation, explaining an additional 5.33% of variance in gene expression and 2.06% of variance in LDL-C. We therefore sought to understand whether a local ancestry-by-genotype interaction could explain additional phenotypic variation relative to genotype alone. We modeled a local ancestry-by-genotype interaction by replacing the term for trait-associated SNPs, $g_i\beta_g$, with separate terms for African and European trait-associated SNPs, $g_{i,A}\beta_A$ and $g_{i,E}\beta_E$. g_i is defined as the number of alternate alleles at a given trait-associated locus, and we correspondingly define $g_{i,A}$ as the number of alternate alleles with African local ancestry and $g_{i,E}$ as the number of alternate alleles with European local ancestry. $g_{i,A}$ and $g_{i,E}$ therefore sum to g_i, and β_A is the effect size in African local ancestry while β_E is the effect size in European local ancestry. Replacing the genotype term in our model with a local ancestry-by-genotype interaction term explains an additional 0.45% of variance in gene expression and 0.25% of variance in LDL-C. We therefore find that local ancestry-by-genotype interactions are relevant for phenotypic variation. This is at least partly driven by the differences in LD structure between ancestries and the ways in which local LD structure impacts the correlation between trait-associated SNPs and causal variants.

Testing for differences in causal variant effect sizes between populations. Finally, we asked whether causal variant effect sizes differ between populations. This question is difficult to answer with standard approaches, due to the way in which LD structure biases comparisons of marginal effect sizes. We therefore developed a model that leverages the multiple ancestries within admixed genomes to indirectly test whether causal variant effect sizes differ between populations. Specifically, we test whether a genetic variant in a region of European ancestry has the same effect in African-Americans and European-Americans. Given the relatively short time since admixture in African-Americans (approximately 8 generations), we assume that regions of European ancestry in

African-Americans and European-Americans are identical with respect to causal variants and LD structure. Then, under the null hypothesis that causal variants have the same effect in all populations, trait-associated SNPs in regions of European ancestry will have the same marginal effect size in African-Americans and European-Americans. However, if causal variant effect sizes differ between populations due to epistatic or gene-by-environment interactions, trait-associated SNPs in regions of European ancestry will have different marginal effect sizes in African-Americans and European-Americans. Specifically, we hypothesize that, in African-Americans, the presence of epistatic or gene-by-environment interactions would drive the marginal effect size of SNPs in regions of European ancestry to be more similar to that of SNPs in regions of African ancestry. Importantly, this comparison is only possible because we compare marginal effect sizes between genomic regions with shared LD structure; without accounting for LD structure, we would expect marginal effect sizes of trait-associated SNPs to differ between populations regardless of whether causal variant effect sizes do (i.e. regardless of whether epistatic or gene-by-environment interactions exist).

To test this hypothesis, we extend Equation 6 in Table 1 and model the phenotype y_i for a single individual i as follows:

$$y_i = c_i\beta_c + r_i\beta_r + \theta_i\beta_\theta + \gamma_i\beta_\gamma + g_{i,A}\beta_A + g_{i,E}\beta_E + \delta r_i g_{i,E}(\beta_A - \beta_E) \quad (1)$$

The first four terms (c_i, r_i, θ_i, γ_i) are technical covariates: race; global ancestry and principal components; and local ancestry respectively. The next two terms ($g_{i,A}\beta_A$, $g_{i,E}\beta_E$), model a local ancestry-by-genotype interaction as described above.

In the final term, we introduce the parameter δ, which measures the extent to which marginal effect sizes of SNPs in regions of European ancestry in African-Americans deviate from those in European-Americans. Using the parameter δ, we can indirectly test whether causal variant effect sizes differ between African-Americans and European-Americans. When δ equals 0, the marginal effect size of a SNP in a region of European ancestry in an African-American is equal to β_E; as δ approaches 1, the marginal effect size approaches β_A. Thus, under the null hypothesis that causal effect sizes are not modified by the genome or environment, δ will be equal to 0. On the other hand, values of δ greater than 0 indicate that SNPs in regions of European ancestry in African-Americans and European-Americans have different marginal effect sizes, which provides evidence for a difference in causal variant effect sizes between populations. Importantly, because we estimate δ by comparing regions of shared ancestry in two different populations, our estimator is not biased

a) Because LD structure differs between ancestries, we model the marginal effect size of a SNP with different parameters for European and African local ancestry.
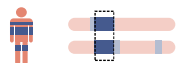
$$y_i = g_i \beta_E$$

$$y_i = g_i \beta_A$$

Suppose that an admixed African-American has a region of European ancestry at this SNP.

$$y_i = ?$$

Is the marginal effect size of the SNP determined strictly by the local European ancestry or modified by epistatic and gene-environmental interactions?
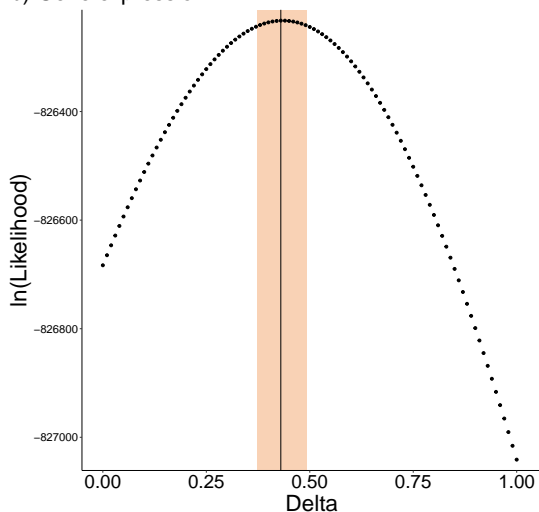
We model the marginal effect size of SNPs in regions of European ancestry in African-Americans as

$$\beta_E + \delta(\beta_A - \beta_E)$$

where δ ranges from 0 to 1. As δ increases, evidence for a difference in causal variant effect sizes across populations increases.

$$\delta = 0 \quad\quad y_i = g_i \beta_E$$

$$0 < \delta < 1 \quad\quad y_i = g_i\big(\beta_E + \delta(\beta_A - \beta_E)\big)$$

$$\delta = 1 \quad\quad y_i = g_i \beta_A$$

b) Gene expression

c) LDL–C
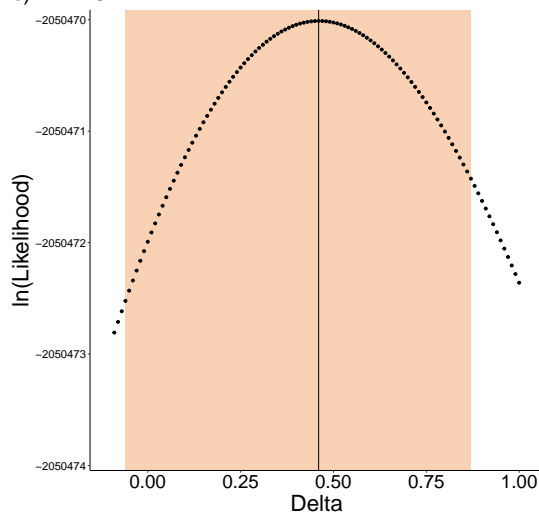
Figure 3: **Testing for differences in causal variant effect sizes between populations. a)** We tested for differences in causal variant effect sizes by leveraging regions of European ancestry shared between African-Americans and European-Americans. The parameter δ measures the extent to which the marginal effect sizes of SNPs in regions of European ancestry in African-Americans deviate from those in European-Americans. **b,c)** Likelihood surface for δ. Maximum likelihood estimates are 0.43 (0.37, 0.49) for gene expression and 0.46 (-0.06, 0.87) for LDL-C. We denote the MLE and 95% bootstrap confidence interval with the vertical line and shaded region, respectively.

11

by differences in LD structure between European and African ancestries, nor by the possibility of private causal variants.

For both traits, we fit this model to the trait-associated SNPs we previously ascertained. We expect that estimates of δ will be noisy at individual SNPs, so for each trait, we estimated a single shared value of δ across all SNPs. This results in one value of δ for gene expression, estimated from 2,697 SNP-gene associations, and one value for LDL-C, estimated from 124 LDL-associated SNPs. Because this model is non-linear, we iteratively optimized δ and all other coefficients, $\beta = (\beta_c, \beta_r, \beta_\theta, \beta_\gamma, \beta_A, \beta_E)$ with ordinary least squares until convergence. To construct a confidence interval for $\hat{\delta}$, we bootstrapped over SNPs.

We first assessed the bias in our estimator $\hat{\delta}$. We simulated genotypes and phenotypes for gene expression, with true values for δ ranging between 0 and 1. We applied our iterative optimization method to estimate δ and found that we recovered the true value of δ in simulations, demonstrating our estimator is both accurate and unbiased (Supp. Fig. 6a). We next performed an additional test of the robustness of our estimator using real data. Specifically, we adapted our model to test whether causal variants have the same effect size in two independent, randomly sampled subsets of European-Americans. We expect that SNPs have the same marginal effect size in both groups of individuals, which means δ should be approximately equal to 0. If δ is estimated to be nonzero between these sets of European-Americans, this would indicate that our method is biased. We applied our model to gene expression and LDL-C and find that $\hat{\delta}$ is not significantly different from 0 when comparing two sets of European-Americans: $\hat{\delta}$ is 0.00 for gene expression (95% CI is -0.03, 0.03) and -0.03 for LDL-C (95% CI is -0.09, 0.03) (Supp. Fig. 5). This demonstrates that our estimator $\hat{\delta}$ is unbiased under the null where the true parameter δ equals zero.

Finally, we used our model to test whether causal variants have the same effect size in African-Americans and European-Americans. For gene expression, $\hat{\delta}$ is significantly different from zero, with a maximum likelihood estimate (MLE) of 0.43 and a 95% CI of $(0.37, 0.49)$ (Fig. 3b). For LDL-C, we estimate a similar MLE of 0.46 with a 95% CI of $(-0.06, 0.87)$ (Fig. 3c). Our results indicate that SNPs in regions of European ancestry in African-Americans and European-Americans have different marginal effect sizes, providing evidence for different causal variant effect sizes between populations. To compute the phenotypic variance explained by this model, we used the same approach described above. In a test set, the δ term describes a modest amount of phenotypic variance: 0.02% for gene expression, 0.01% LDL-C. Thus, we find that although these differences explain a small amount of trait variance, causal variant effect sizes do differ between populations. Moreover, these effect sizes differ because they are modified by the genome or environment.

12

Discussion

We developed a model in which we introduce the parameter δ to test for differences in causal variant effect sizes between populations. Specifically, we leveraged regions of European ancestry shared between African-Americans and European-Americans to compare marginal effect sizes in a manner unbiased by LD structure. We applied our model to two traits, gene expression in MESA and LDL-C in MVP. For gene expression, we observe that $\hat{\delta}$ is significantly different from zero, implying that causal variant effect sizes differ between African-Americans and European-Americans. For LDL-C, we obtain a MLE for δ that is similar to that from gene expression but not significantly different from zero.

The larger confidence interval for LDL-C is likely due to differences in statistical power between the two traits. Though we used significantly associated SNPs for both traits, many fewer SNPs were used in LDL-C analyses (124 SNPs) than in gene expression analyses (2,697 SNPs). Moreover, trait-associated SNPs were ascertained within the same dataset (MESA) for gene expression but were ascertained from an external dataset (UK Biobank) for LDL-C. This should not bias the estimation of δ but may mean that trait-associated SNPs capture a larger proportion of phenotypic variance for gene expression relative to LDL-C.

Our observation that causal variant effect sizes differ between populations is relevant to previous work on quantifying cross-population genetic correlations (Brown *et al.*, 2016; Galinsky *et al.*, 2019). Though there is no straightforward analytical relationship between our parameter δ and genetic correlation, our results are intuitively consistent with a cross-population genetic correlation less than one. Because our method is not biased by differences in LD structure, differences in causal variant effect sizes between populations must be due to unmodeled epistatic or gene-by-environment interactions. The magnitude of genetic interactions in complex traits has been difficult to quantify because GWAS are generally underpowered to detect individual interactions of small effect. We therefore only estimate one parameter δ from all trait-associated SNPs to maximize power. Nevertheless, future areas of investigation include adapting our model to understand how the magnitude of genetic interactions varies across SNPs or functional regions of the genome.

Though we observe that causal variant effect sizes significantly differ between populations, we also find that the inclusion of the δ term in the model does not substantially increase the amount of phenotypic variance explained. This apparent discrepancy can be resolved by noting that we evaluate model performance on the full dataset of African-Americans and European-Americans, but the δ term will only improve the modeling of effect sizes in regions of European ancestry in

African-Americans, which only represents about 10% of the full dataset.

Our results have implications for the application of polygenic scores (PGS). PGS constructed in European populations are known to have poor portability to non-European populations. We find that trait-associated SNPs ascertained in Europeans have attenuated effect sizes in African-Americans, which is consistent with European-ascertained SNPs tagging causal variants more poorly in African ancestry. This suggests that a PGS constructed by ascertaining SNPs in Europeans could have poor accuracy in African-Americans even if SNP effect sizes are estimated from African-Americans. Furthermore, we find that local ancestry and local ancestry-by-genotype interactions both contribute to phenotypic variance, implying that PGS in admixed populations can be improved by modeling these terms. Lastly, we find that causal variant effect sizes differ between populations. Our results suggest that differences in causal variant effect sizes contribute to poor PGS portability but are not the main factors responsible. Nevertheless, differences in causal variant effect sizes that arise from gene-by-environment interactions could help explain the observed decrease in PGS performance when applied to populations of the same ancestry and different environments (Mostafavi *et al.*, 2020). Maximizing PGS accuracy for an individual therefore requires constructing the PGS in a population with the same ancestry and environment as the individual, underscoring the importance of conducting GWAS in diverse populations.

In summary, our method tests for causal variant effect size differences between populations by leveraging the multiple ancestries contained in admixed populations to control for differences in LD structure. This method can be applied to any quantitative or dichotomous trait and additionally describes the contribution of global and local ancestry to phenotypic variance. Ultimately, our finding that causal variant effect sizes differ by population gives insight into the importance of genetic interactions in human complex traits.

Methods

Data processing

For MVP, we used data processed and subject to quality control as described in Hunter-Zinck *et al.* (2020) and used GRCh37 genotype calls. Using KING coefficients (Manichaikul *et al.*, 2010), we removed relatives who were closer than 3rd degree cousins, which left ∼72K African-American and ∼298K European-American individuals. Data were imputed with IMPUTE using the 1000 Genomes Phase 3 reference panel (Auton *et al.*, 2015; Howie *et al.*, 2009).

For MESA, we obtained phased whole genome sequencing data and gene expression data from TOPMed Freeze 8. We restricted our analyses to individuals for whom there is available gene expression data in peripheral blood mononuclear cells, and who have self-reported race as either White/European-American or Black/African-American.

Inferring global and local ancestry

We performed global ancestry inference with supervised ADMIXTURE using default program parameters (Alexander *et al.*, 2009). We used 99 CEU individuals and 108 YRI individuals from 1000 Genomes Phase 3 as our reference populations. We filtered for biallelic SNPs with MAF > 0.05 in both the admixed population and the reference populations, and again filtered for MAF > 0.1 after merging the admixed and reference datasets. We pruned SNPs with an r^2 value > 0.1. As previously done for MVP (Hunter-Zinck *et al.*, 2020), the race of each individual (i.e. African-American or European-American) was determined by HARE (Fang *et al.*, 2019).

We performed local ancestry inference on admixed African-American individuals with RFMix v1.5.4, using no EM iterations and default program parameters (Maples *et al.*, 2013). We assumed 8 generations since the time of admixture between an African population and a European population (Baharian *et al.*, 2016). We again used 99 CEU individuals and 108 YRI individuals from 1000 Genomes Phase 3 as our reference populations. We used biallelic SNPs with MAF > 0.05 in both the admixed population and the reference populations, and removed SNPs with an r^2 value > 0.5.

Global ancestry fractions from ADMIXTURE are highly correlated with those implied by RFMix (MESA $\rho = 0.99$, MVP $\rho = 0.98$) (Supp. Fig. 1). In MESA, we excluded 18 African-American individuals from downstream analyses who either had < 0.5 global African ancestry or else had discordant RFMix and ADMIXTURE global ancestry estimates. We determined that an individual had discordant global ancestry estimates if the Euclidean distance between the two estimates and their average was greater than 0.05. We additionally excluded one European-

15

American individual from MESA who clustered with African-American individuals in principal components analysis of genotypes.

Comparing marginal SNP effect sizes between populations

LDL-C (MVP)

Principal components were calculated on all individuals in the MVP dataset with HARE (Hunter-Zinck *et al.*, 2020). We ascertained genome-wide significant SNPs in ∼315K UK Biobank White British individuals (application number 24983). After applying genomic filters (MAF $\geq$ 0.01, missing genotype rate $\leq$ 0.05, Hardy-Weinberg equilibrium), we tested for association with inverse-variance quantile normalized phenotypes using a linear model (–glm) in plink with the covariates age, sex, assessment center, and statin usage. Significant variants ($p < 5 \times 10^{-8}$) were clumped and pruned to leave independent SNPs.

To estimate effect sizes of these variants in MVP, we extracted variants from the imputed genotype set using 1000 Genomes Phase 3 as our reference panel. We filtered for MAF $\geq$ 0.003 in European-Americans and African-Americans, leaving 124 independent SNPs. Each individual had LDL-C measurements at multiple time points, and we therefore used the maximum LDL-C measurement across time points for all analyses. Additionally, we numerically adjusted LDL-C measurements based on statin usage. We inferred statin usage at the time of LDL-C measurement if a statin prescription was filled within the length of the prescription plus a buffer of 15 days within the measurement date. Our covariates included age, sex, global ancestry, and genetic PC1, which stratifies European-Americans and is the only principal component associated with LDL-C after residualizing on the other covariates.

To estimate effect sizes from the 72K African-Americans (β_{AA}), we used linear regression in plink (–glm) and included the covariates above. We likewise randomly sampled an equal number of European-Americans and estimated effect sizes (β_{EA}).

Gene expression (MESA)

After filtering for individuals based on ancestry, 499 European-American individuals and 319 African-American individuals with PBMC RNASeq remain. 380 of these individuals had gene expression data available at two time points, spaced 5 years apart (exam 1 and exam 5). For these individuals, we used a greedy algorithm to select which time point to use such that the proportions of potential confounders (sex, time of exam, sequencing center) were approximately balanced between European-Americans and African-Americans.

16

Gene-level read counts were obtained with RNA-SeQC v1.1.9 (DeLuca *et al.*, 2012). As done previously by GTEx Consortium (2017), we selected genes with expression thresholds of >0.1 TPM in at least 20% of samples and ≥ 6 reads in at least 20% of samples, thresholding separately for European-Americans and African-Americans in both cases. A total of 10,765 genes passed this filtering step. We log-transformed gene expression measurements and used these transformed phenotypes in all downstream analyses. We selected biallelic SNPs with a MAF $>$ .05 and minor allele sample count $>$ 5 in both European-Americans and African-Americans.

To identify SNPs affecting expression in *cis*, we filtered for SNPs within 100 kb of the TSS for each gene. We ascertained trait-associated SNPs in a randomly sampled subset of 249 European-Americans using ordinary least squares. We included four covariates that were significantly correlated with expression phenotypes: sequencing center, time of exam, sex, and genotype PC 2, which captures structure within European-Americans (Supp. Fig. 2, Supp. Fig. 3a).

We ascertained trait-associated SNPs with a false discovery rate of 0.01. If there were multiple significant SNPs per gene remaining after multiple testing correction, we chose the most significant SNP. We ascertained a total of 2,697 SNP-gene associations, and all downstream analyses are restricted to these significant associations. Furthermore, all downstream analyses exclude the European-Americans who were used to ascertain trait-associated SNPs.

For each significant SNP-gene association, we performed two separate regressions to estimate β_{AA}, the effect size in African-Americans, and β_{EA}, the effect size in European-Americans. To estimate β_{EA}, we used the same covariates described above. To estimate β_{AA}, we used sequencing center, time of exam, sex, and global African ancestry fraction. We did not include genotype PC 1 as a covariate despite its significant association with expression because this is highly correlated with global African ancestry fraction (Supp. Fig. 3b). We estimated β_{EA} in a held-out set of 250 European-Americans and randomly sampled an equal number of African-Americans to estimate β_{AA}.

Comparison of effect sizes

We used total least squares regression to assess the slope of the relationship between $\hat{\beta}_{AA}$ and $\hat{\beta}_{EA}$. Estimates of SNP effect sizes are statistically noisy, and unlike ordinary least squares, total least squares is robust to uncertainty in the x-axis variable. We created 1000 bootstrap replicates for each trait by sampling with replacement over SNPs and report the 95% CI of the slope as defined by the 0.025 and 0.975 quantiles.

Quantifying role of local ancestry in phenotypic variance

We subset our data into a training set of individuals and a test set of individuals using an $80\% - 20\%$ split. We used the training subset to fit model coefficients with ordinary least squares regression. We then used the test set to compute the proportion of variance explained by each model as the ratio of the variance of the predicted phenotype and the measured phenotype, $\frac{Var(\hat{y})}{Var(y)}$. Here, we performed a regression on the joint sample of African-Americans and European-Americans, which is distinct from our analyses above which perform a regression on each population separately. Models with genotype terms included the significant trait-associated SNPs only.

For the gene expression data, there were 256 African-Americans and 206 European-Americans in the training set; and 63 African-Americans and 51 European-Americans in the test set. We report the mean total phenotypic variance explained and the mean increase in variance explained, averaged across all significant genes. For the LDL-C, we used 51,650 training and 22,136 test individuals from each population, and averaged variance estimates over SNPs.

Testing for differences in causal variant effect sizes between populations

We constructed a phenotypic model in which the parameter δ measures differences in the effect size of trait-associated SNPs in regions of European ancestry in African-Americans versus European-Americans.

For the gene expression data, we estimated one value of δ from all SNP-gene associations to avoid overparameterization. For LDL-C (MVP), we estimated one value of δ across all trait-associated SNPs.

To fit this model, we began by initializing $\hat{\delta}$ to a random value on the interval $[0, 1]$, which is the most biologically intuitive range of values for δ (see Fig. 3). We next optimized $\hat{\beta} = (\hat{\beta}_c, \hat{\beta}_r, \hat{\beta}_\theta, \hat{\beta}_\gamma, \hat{\beta}_A, \hat{\beta}_E)$ conditional on this value of $\hat{\delta}$, and we then optimized $\hat{\delta}$ conditional on $\hat{\beta}$. For both gene expression and LDL-C, we performed this regression marginally on each SNP. We continued this iterative optimizing with ordinary least squares regression until $\hat{\delta}$ converged (i.e. did not change by $>.0001$). Though $\hat{\delta}$ was initialized on the interval $[0, 1]$, the optimization procedure itself was unconstrained. Additionally, we found that regardless of the initial value of $\hat{\delta}$, our optimization procedure converged to the same value. The optimization method converged quickly for both datasets (20 iterations for gene expression, 18 for LDL-C).

To construct 95% confidence intervals for $\hat{\delta}$, we bootstrapped over SNPs and reported the 0.025 and 0.975 quantiles. (For gene expression, this procedure is equivalent to bootstrapping over genes because each gene is modeled by exactly one SNP.)

To assess the amount of bias in our estimator $\hat{\delta}$ when the true parameter δ equals zero, we performed a negative control. We modified our model such that we could use δ to compare effect sizes between two randomly sampled, independent subsets of European-Americans. On average, these two groups of individuals have the same race, global ancestry, local ancestry, and environment. Thus, we expect that $\hat{\delta} = 0$ even in the presence of epistatic or gene-by-environment interactions. To modify our model, we first excluded any African-Americans with European ancestry at trait-associated SNPs. This ensured that β_A was estimated only from African-Americans with African ancestry at trait-associated SNPs. Next, we randomly assigned a subset of the European-Americans into a validation set. For gene expression, this was 100 individuals, and for LDL-C, this was 72K individuals. We then replaced the race indicator in the last term of the model with a validation set indicator. With this particular modification of the model, our estimator $\hat{\delta}$ tests whether trait-associated SNPs have the same effect size in the validation set and the remaining European-Americans. If $\hat{\delta}$ is an unbiased estimator, $\hat{\delta}$ should be 0 when we conduct the negative control.

Simulations

We constructed a simulation framework designed to emulate our analyses of gene expression in MESA. We used sample sizes for individuals and SNPs that were analogous to the real data. We sampled effect sizes of causal variants from a multivariate Normal distribution:

$$\beta_A, \beta_E \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_A^2 & \sigma_{A,E} \\ \sigma_{A,E} & \sigma_E^2 \end{bmatrix}\right).$$

For each variant, we simulate allele frequencies separately in African and European ancestry by randomly sampling from the empirical joint distribution of allele frequency for trait-associated SNPs in our dataset. We note that this decouples the relationship between effect size and allele frequency, but this relationship is known to be weak for gene expression (Glassberg *et al.*, 2019). We simulated ascertainment in a European-ancestry population with a rejection sampling approach. Specifically, we simulated estimates of effect sizes from a Normal distribution:

$$\hat{\beta}_E \sim N\left(\beta_E, \frac{\sigma_\epsilon^2}{n\sigma_g^2}\right).$$

We simulated residuals, $\hat{\epsilon}$ from a Normal distribution in order to compute a simulated standard error, $s_{\hat{\beta}}$:

$$\hat{\epsilon} \sim N\left(0, \sigma_{\epsilon}^2\right),$$
$$s_{\hat{\beta}}^2 = \frac{\frac{1}{n-2}\sum_i \hat{\epsilon}^2}{ns_g^2}.$$

We then computed t-statistics and rejected the simulated tag SNP if it was insignificant,

$$t = \frac{\hat{\beta}_{E,asc}}{s_{\hat{\beta}}}.$$

We simulated genotypes and phenotypes conditional on race and ancestry. Global African ancestry fraction, θ, is beta-distributed for African-American individuals and 0 for European-American individuals:

$$\theta \sim \begin{cases} Beta(\alpha, \beta) \text{ if African-American} \\ \\ 0 \text{ if European-American.} \end{cases}$$

Local ancestry, γ, is binomially-distributed conditional on global ancestry:

$$\gamma|\theta \sim Binom(2, \theta).$$

Genotypes are then binomially-distributed conditional on local ancestry, where p_A is the allele frequency in African ancestry and p_E is the allele frequency in European ancestry:

$$g_A \sim Binom(\gamma, p_A),$$
$$g_E \sim Binom(\gamma, p_E).$$

Lastly, phenotypes are normally distributed based on our specified model:

$$y \sim N\left(g_A\beta_A + g_E\beta_E + \delta r g_E(\beta_A - \beta_E), \sigma_{\epsilon}^2\right).$$

Hyperparameters were chosen such that the simulated and empirical distributions of $\hat{\beta}$ are as similar as possible (Supp. Fig. 6b-d). We simulated 11 values of δ ranging between 0 and 1, and used our iterative optimization procedure to estimate δ from simulations. To assess the accuracy of our estimator $\hat{\delta}$, we plotted the estimates $\hat{\delta}$ against the true value δ (Supp. Fig. 6a).

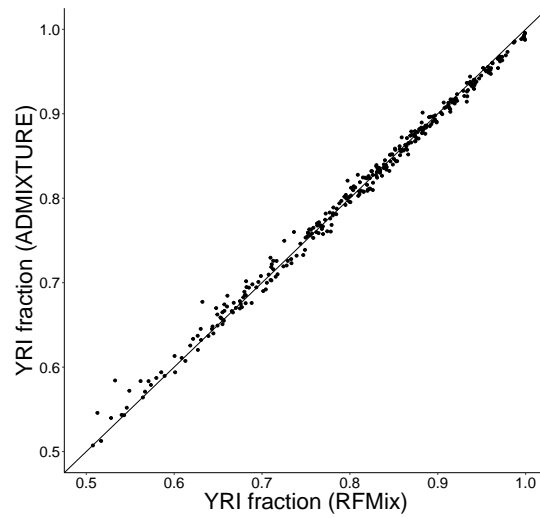# Acknowledgments

Supplementary Items

Figure S1: **Comparison of ancestry inference methods.** We observe a strong correlation between RFMix and ADMIXTURE estimates of global African ancestry fraction for African-American individuals in MESA.
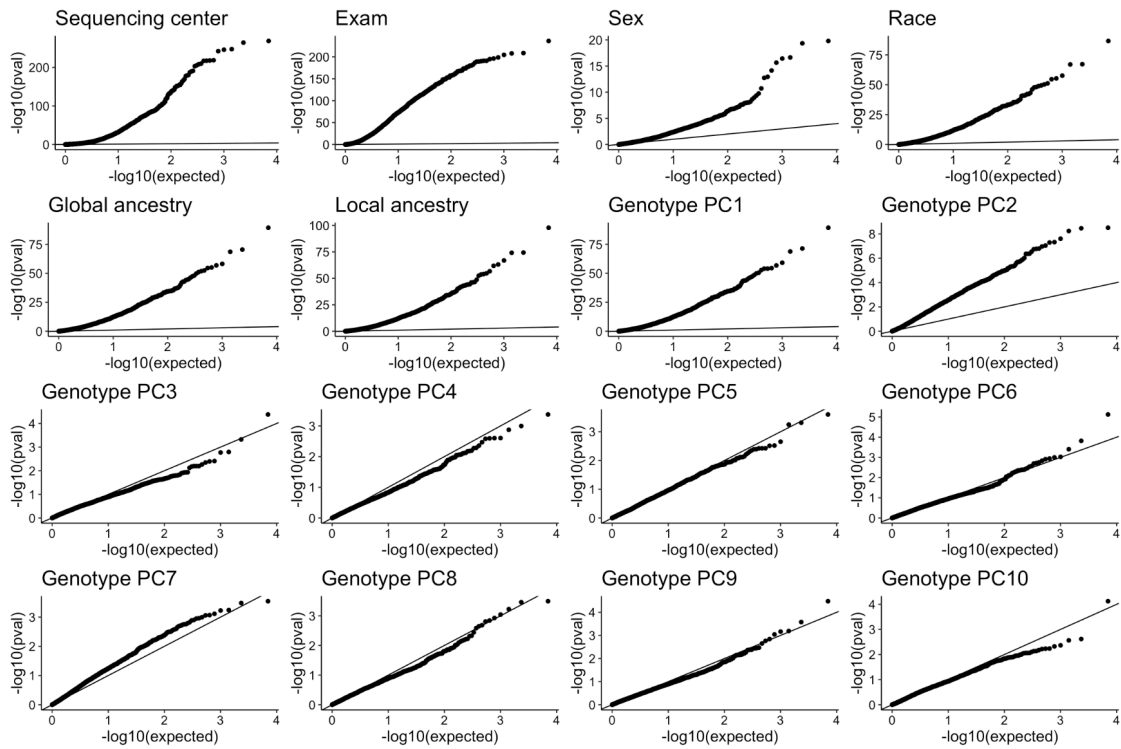
Figure S2: **Association between gene expression phenotypes and covariates.** We tested for statistical association between phenotypes of all 2,697 significant genes and 16 covariates, including 2 batch covariates (sequencing center and exam number), sex, race, global and local ancestry, and 10 genotype principal components (PCs). We show the resulting QQ plots of association p-values, demonstrating that expression phenotypes are significantly associated with both batch covariates, sex, race, global and local ancestry, and the first two PCs.
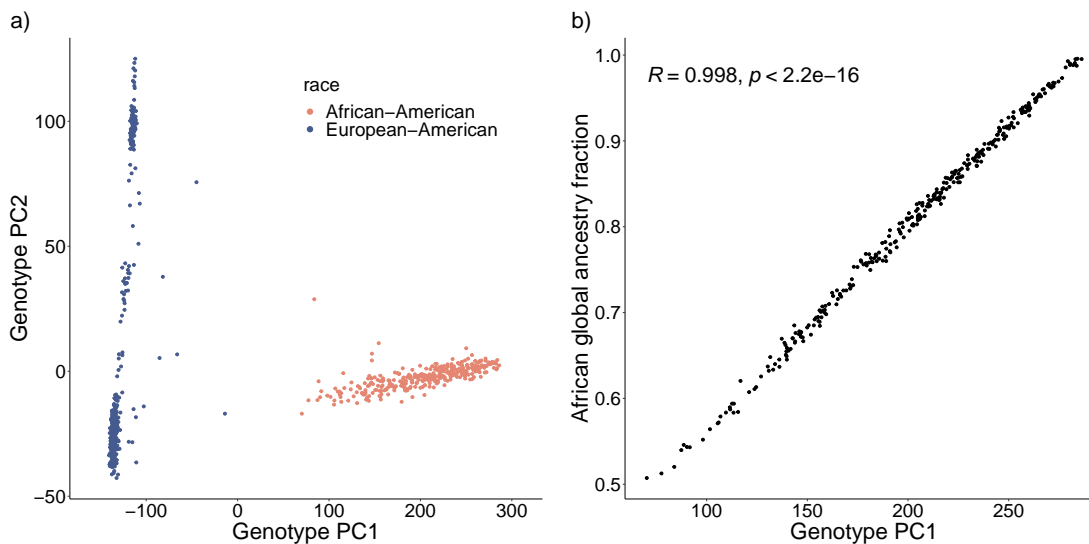
Figure S3: **Principal components analysis of MESA genotypes. a)** We computed principal components from the genotypes of 319 African-Americans and 499 European-Americans in MESA. The first genotype PC stratifies the African-Americans and the second genotype PC stratifies the European-Americans. **b)** Within African-Americans, the first genotype PC is highly correlated with African global ancestry fraction.
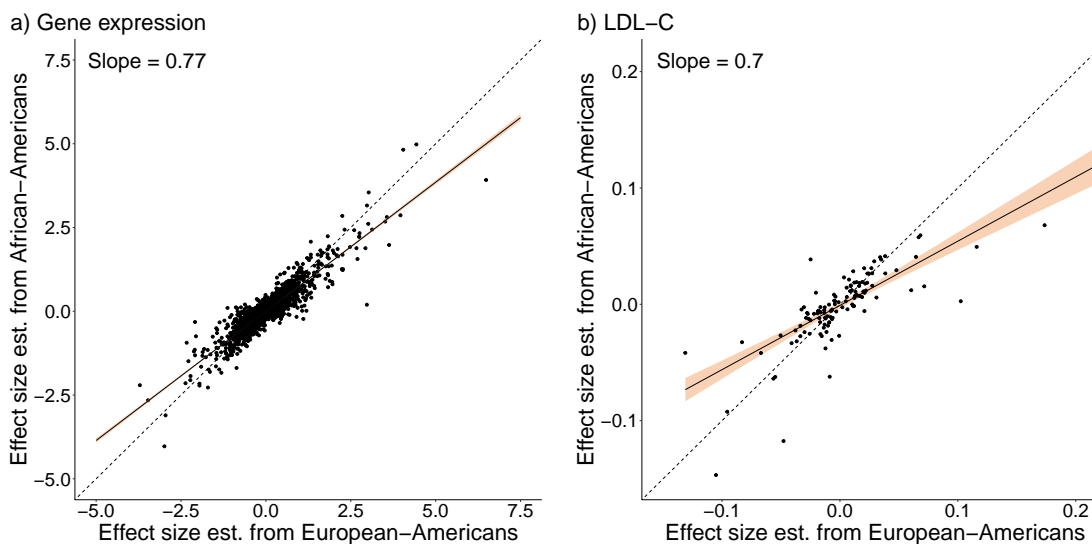

Figure S4: Ordinary least squares regression of β_{AA} on β_{EA} for **a)** gene expression and **b)** LDL-C. We represent the 95% confidence interval with the orange shaded region.
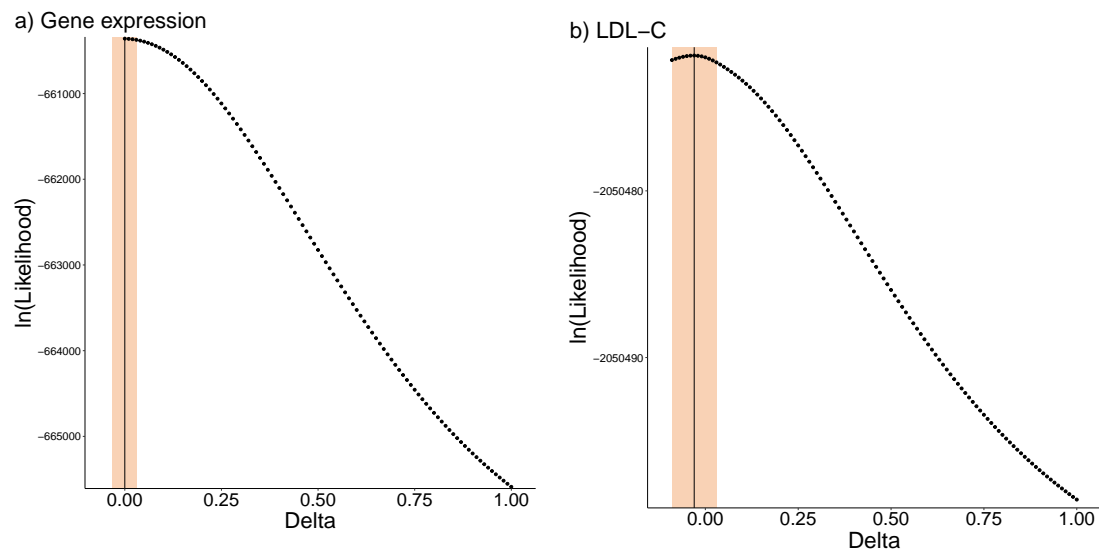
Figure S5: **Negative control.** Likelihood surface for δ. Maximum likelihood estimates are **a)** 0.00 (-0.03, 0.03) for gene expression and **b)** -0.03 (-0.09, 0.03) for LDL-C. We denote the MLE and 95% bootstrap confidence interval with the vertical line and shaded region, respectively.
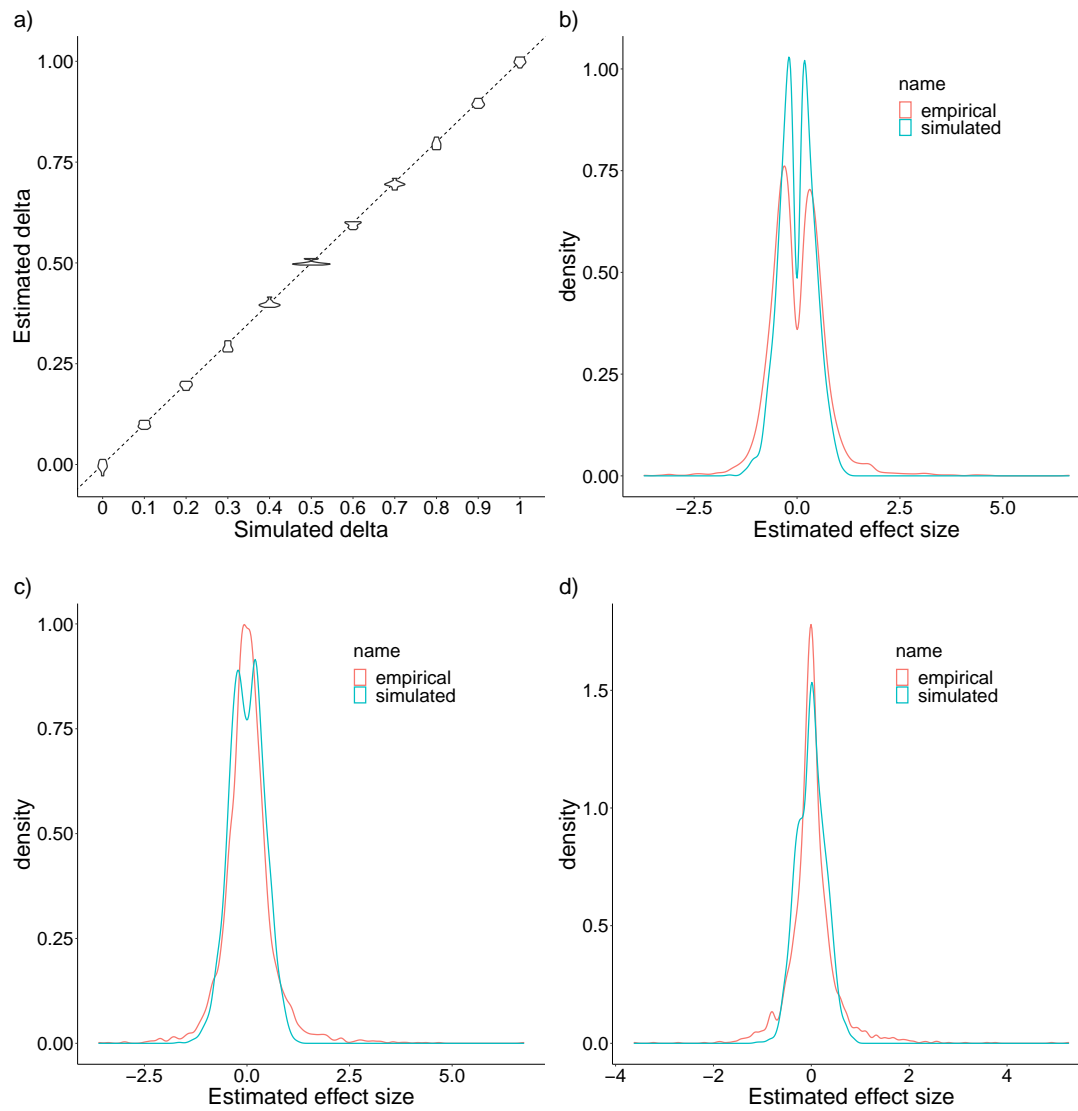
Figure S6: **Simulations.** **a)** Distribution of $\hat{\delta}$ values estimated for 11 different simulated values of δ. **b, c, d)** Comparison of empirical and simulated distributions of $\hat{\beta}$ in the ascertainment European-American population, the non-ascertainment European-American population, and the African-American population.

References

Alexander, D. H., J. Novembre, and K. Lange, 2009 Fast model-based estimation of ancestry in unrelated individuals. Genome Research **19**: 1655–1664.

Auton, A., G. R. Abecasis, D. M. Altshuler, R. M. Durbin, G. R. Abecasis, *et al.*, 2015 A global reference for human genetic variation. Nature **526**: 68–74.

Baharian, S., M. Barakatt, C. R. Gignoux, S. Shringarpure, J. Errington, *et al.*, 2016 The Great Migration and African-American Genomic Diversity. PLOS Genetics **12**: e1006059.

Berg, J. J., A. Harpak, N. Sinnott-Armstrong, A. M. Joergensen, H. Mostafavi, *et al.*, 2019 Reduced signal for polygenic adaptation of height in UK Biobank. eLife **8**.

Bitarello, B. D. and I. Mathieson, 2020 Polygenic Scores for Height in Admixed Populations. G3: Genes, Genomes, Genetics **10**: 4027–4036.

Brown, B. C., Asian Genetic Epidemiology Network Type 2 Diabetes Consortium, C. J. Ye, A. L. Price, and N. Zaitlen, 2016 Transethnic Genetic-Correlation Estimates from Summary Statistics. American Journal of Human Genetics **99**: 76–88.

DeLuca, D. S., J. Z. Levin, A. Sivachenko, T. Fennell, M.-D. Nazaire, *et al.*, 2012 RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics (Oxford, England) **28**: 1530–1532.

de Candia, T., S. Lee, J. Yang, B. Browning, P. Gejman, *et al.*, 2013 Additive Genetic Variation in Schizophrenia Risk Is Shared by Populations of African and European Descent. American Journal of Human Genetics **93**: 463–470.

Fang, H., Q. Hui, J. Lynch, J. Honerlaw, T. L. Assimes, *et al.*, 2019 Harmonizing Genetic Ancestry and Self-identified Race/Ethnicity in Genome-wide Association Studies. American Journal of Human Genetics **105**: 763–772.

Galinsky, K. J., Y. A. Reshef, H. K. Finucane, P.-R. Loh, N. Zaitlen, *et al.*, 2019 Estimating cross-population genetic correlations of causal effect sizes. Genetic epidemiology **43**: 180–188.

Glassberg, E. C., Z. Gao, A. Harpak, X. Lan, and J. K. Pritchard, 2019 Evidence for weak selective constraint on human gene expression. Genetics **211**: 757–772.

GTEx Consortium, 2017 Genetic effects on gene expression across human tissues. Nature **550**: 204–213.

Howie, B. N., P. Donnelly, and J. Marchini, 2009 A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLOS Genetics **5**: e1000529.

Hunter-Zinck, H., Y. Shi, M. Li, B. R. Gorman, S.-G. Ji, *et al.*, 2020 Genotyping Array Design and Data Quality Control in the Million Veteran Program. American Journal of Human Genetics **106**: 535–548.

Khera, A. V., M. Chaffin, K. G. Aragam, M. E. Haas, C. Roselli, *et al.*, 2018 Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics **50**: 1219–1224.

Mancuso, N., N. Rohland, K. A. Rand, A. Tandon, A. Allen, *et al.*, 2016 The contribution of rare variation to prostate cancer heritability. Nature genetics **48**: 30–35.

Manichaikul, A., J. C. Mychaleckyj, S. S. Rich, K. Daly, M. Sale, *et al.*, 2010 Robust relationship inference in genome-wide association studies. Bioinformatics **26**: 2867–2873.

Maples, B. K., S. Gravel, E. E. Kenny, and C. D. Bustamante, 2013 RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. American Journal of Human Genetics **93**: 278–288.

Martin, A. R., C. R. Gignoux, R. K. Walters, G. L. Wojcik, B. M. Neale, *et al.*, 2017 Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. American Journal of Human Genetics **100**: 635–649.

Martin, A. R., M. Kanai, Y. Kamatani, Y. Okada, B. M. Neale, *et al.*, 2019 Clinical use of current polygenic risk scores may exacerbate health disparities. Nature Genetics **51**: 584–591.

Mavaddat, N., K. Michailidou, J. Dennis, M. Lush, L. Fachal, *et al.*, 2019 Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. American Journal of Human Genetics **104**: 21–34.

Mostafavi, H., A. Harpak, I. Agarwal, D. Conley, J. K. Pritchard, *et al.*, 2020 Variable prediction accuracy of polygenic scores within an ancestry group. eLife **9**: e48376.

Musharoff, S., D. S. Park, A. Dahl, J. M. Galanter, X. Liu, *et al.*, 2018 Existence and implications of population variance structure. BioRxiv .

Nédélec, Y., J. Sanz, G. Baharian, Z. A. Szpiech, A. Pacis, *et al.*, 2016 Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. Cell **167**: 657–669.e21.

Price, A. L., N. Patterson, D. C. Hancks, S. Myers, D. Reich, *et al.*, 2008 Effects of cis and trans genetic ancestry on gene expression in African Americans. PLoS Genetics **4**: e1000294.

Randolph, H. E., Z. Mu, J. K. Fiege, B. K. Thielen, J.-C. Grenier, *et al.*, 2020 Single-cell RNA-sequencing reveals pervasive but highly cell type-specific genetic ancestry effects on the response to viral infection. preprint, Genomics.

Sharp, S. A., S. S. Rich, A. R. Wood, S. E. Jones, R. N. Beaumont, *et al.*, 2019 Development and Standardization of an Improved Type 1 Diabetes Genetic Risk Score for Use in Newborn Screening and Incident Diagnosis. Diabetes Care **42**: 200–207.

Sohail, M., R. M. Maier, A. Ganna, A. Bloemendal, A. R. Martin, *et al.*, 2019 Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. eLife **8**.

Võsa, U., A. Claringbould, H.-J. Westra, M. J. Bonder, P. Deelen, *et al.*, 2021 Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nature Genetics **53**: 1300–1310.

Wang, Y., J. Guo, G. Ni, J. Yang, P. M. Visscher, *et al.*, 2020 Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. Nature Communications **11**: 3865.

Wojcik, G. L., M. Graff, K. K. Nishimura, R. Tao, J. Haessler, *et al.*, 2019 Genetic analyses of diverse populations improves discovery for complex traits. Nature **570**: 514–518.