Causal effects on complex traits are similar for common variants across segments of different continental ancestries within admixed individuals

Kangcheng Hou [1] ✉, Yi Ding[1], Ziqi Xu[2], Yue Wu[2], Arjun Bhattacharya [3], Rachel Mester [4], Gillian M. Belbin[5,6,24], Steve Buyske [7,24], David V. Conti[8,24], Burcu F. Darst[9,24], Myriam Fornage[10,24], Chris Gignoux[11,24], Xiuqing Guo [12,24], Christopher Haiman[8,24], Eimear E. Kenny[5,13,14,24], Michelle Kim[9,24], Charles Kooperberg [9,24], Leslie Lange[15,24], Ani Manichaikul[16,24], Kari E. North [7,17,24], Ulrike Peters [9,24], Laura J. Rasmussen-Torvik[18,24], Stephen S. Rich [16,24], Jerome I. Rotter[12,24], Heather E. Wheeler [19,20,24], Genevieve L. Wojcik [21,24], Ying Zhou[9,24], Sriram Sankararaman[1,2,22,23] & Bogdan Pasaniuc [1,3,22,23] ✉

Individuals of admixed ancestries (for example, African Americans) inherit a mosaic of ancestry segments (local ancestry) originating from multiple continental ancestral populations. This offers the unique opportunity of investigating the similarity of genetic effects on traits across ancestries within the same population. Here we introduce an approach to estimate correlation of causal genetic effects (r_{admix}) across local ancestries and analyze 38 complex traits in African-European admixed individuals ($N = 53,001$) to observe very high correlations (meta-analysis $r_{admix} = 0.95$, 95% credible interval 0.93–0.97), much higher than correlation of causal effects across continental ancestries. We replicate our results using regression-based methods from marginal genome-wide association study summary statistics. We also report realistic scenarios where regression-based methods yield inflated heterogeneity-by-ancestry due to ancestry-specific tagging of causal effects, and/or polygenicity. Our results motivate genetic analyses that assume minimal heterogeneity in causal effects by ancestry, with implications for the inclusion of ancestry-diverse individuals in studies.

Large-scale genotype–phenotype studies are increasingly analyzing diverse sets of individuals of various continental and subcontinental ancestries[1–4]. A fundamental open question in these studies is to what extent the genetic basis of common human diseases and traits are shared/distinct across different ancestry populations and its impact to genetic discovery and prediction[5–9]. For example, it is unclear how much of the low polygenic score portability can be attributed to differences in genetic causal effects across ancestries[5,10,11]. Hence, understanding the role of ancestry in variability of causal effect sizes has tremendous implications for understanding the genetic basis of disease
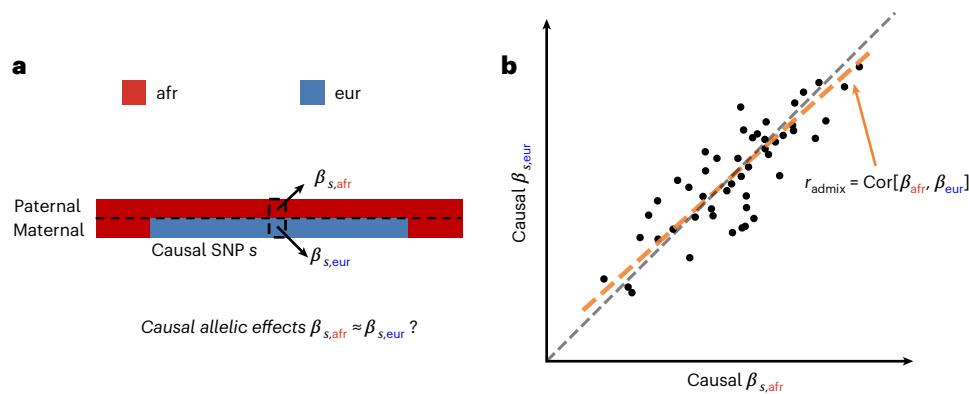
Fig. 1 | Concepts of estimating similarity in the causal effects across local ancestries. a, For a given trait, with phased genotype (paternal haplotype at the top and maternal haplotype at the bottom) and inferred local ancestry (denoted by color), we investigate whether $\beta_{s,\text{afr}} \approx \beta_{s,\text{eur}}$ across each causal SNP s. **b**, We

focus on estimating the genome-wide correlation of genetic effects across ancestries $r_{\text{admix}} = \text{Cor}[\beta_{\text{afr}}, \beta_{\text{eur}}]$, which is the regression slope (orange line) of ancestry-specific causal effects. For reference, the gray dashed line corresponds $\beta_{\text{afr}} = \beta_{\text{eur}}$.

and portability of genetic risk scores in personalized and equitable genomic medicine[1,10–13].

The standard approach to estimating similarity in causal effects across ancestries has focused on cross-population analyses (typically at continental level) in which effect sizes estimated by large-scale genome-wide association studies (GWAS) are compared across continental-level ancestry groups[5–8,14,15]. Such studies have found significant differences, albeit with modest magnitude, of causal effects in cross-continental comparisons. However, a main drawback of such studies is the differences in definition of environment/phenotype across such broad units of ancestry that can reduce the observed similarity; for example, the low estimated similarity in causal genetic effects for major depressive disorder across Europeans and East Asians may be attributed to different diagnostic criteria in the two populations[8,16].

As an alternative to studying populations across different continents, causal effects similarity by ancestry can also be studied within recently admixed populations. Recently admixed individuals have the unique feature of having their genomes as mosaic of ancestry segments (local ancestry) originating from the ancestral populations within the past few dozen generations; for example, African American genomes are composed of segments of African and European ancestries within the past 5–15 generations[17]. Unfortunately, admixed populations are vastly underrepresented in genomic studies[18], partly because of the lack of understanding of how the genetic causal effects vary across ancestries[10,17,19–22]. For example, heterogeneity of marginal effects (which is estimated in GWAS single variant scan and can tag effects from nearby variants due to linkage disequilibrium (LD)) for a few traits and loci has been reported[23–26], but it remains unknown whether this reflects true difference in causal genetic effects or confounding due to different allele frequencies and/or LD by ancestry. Recent work[15] has reported evidence of causal effect heterogeneity for single nucleotide polymorphisms (SNPs) in regions of European ancestries comparing individuals of European versus African American ancestries; however, these studies focused on cross-population comparisons instead of comparing effects across local ancestries within admixed populations. Estimating the magnitude of similarity in causal effects across ancestries is important for all genotype–phenotype studies in admixed populations from mapping to polygenic prediction, particularly within methods that allow for effects to vary across local ancestry segments[19–22].

In this Article, we quantify the similarity in the causal effects (that is, change in phenotype per allele substitution) across local ancestries within admixed populations; such similarity can be defined as the correlation of ancestral causal genetic effects $r_{\text{admix}} = \text{Cor}[\beta_{\text{afr}}, \beta_{\text{eur}}]$ across African (β_{afr}) and European (β_{eur}) local ancestries. We develop a method that leverages the polygenic architecture of complex traits to

model all variants (GWAS-significant and non-significant); this approach is accurate and robust across a wide range of realistic simulated genetic architectures. We also investigate regression-based approaches that use marginal effects of SNPs prioritized in GWAS risk regions. Through simulation studies, we find that regression-based methods can yield deflated estimates of similarity (that is, inflated heterogeneity) especially for highly polygenic traits.

We analyze complex traits in African-European admixed individuals in Population Architecture using Genomics and Epidemiology (PAGE)[1] (24 traits, average $N = 9{,}296$), UK Biobank (UKBB)[2] (26 traits, average $N = 3{,}808$), and All of Us (AoU)[3] (10 traits, average $N = 20{,}496$); there are 38 unique traits in total. We find causal effects are largely consistent across local ancestries within admixed individuals (through meta-analysis across 38 traits, estimated correlation of $r_{\text{admix}} = 0.95$, 95% credible interval 0.93–0.97). In addition, we find that the heterogeneity in marginal effects exhibited at several trait–locus pairs can be explained by multiple nearby causal variants within a region, consistent with our simulation studies. Our results suggest that the causal effects are largely consistent across local ancestries within African-European admixed individuals, and this motivates future genetic analysis in admixed populations that assume similar effects across ancestries for improved power.

Results

Overview

We start by describing the statistical model we use to relate genotype to phenotypes in two-way admixed individuals; we focus on two-way African-European admixture because their local ancestries can be accurately inferred (Methods; for extension to other admixed populations, see Discussion). For a given individual, at each SNP s, we denote number of minor alleles from maternal and paternal haplotypes as $x_{s,M}, x_{s,P} \in \{0,1\}$ and local ancestries as $\gamma_{s,M}, \gamma_{s,P} \in \{\text{afr}, \text{eur}\}$. Denoting $\mathbb{I}(\cdot)$ as the indicator function, we define the local ancestry dosage as allele counts from each of ancestries; for example, $\ell_s = \mathbb{I}(\gamma_{s,M} = \text{afr}) + \mathbb{I}(\gamma_{s,P} = \text{afr})$ for African (similarly for European). For modeling convenience, we use variables that encode the genotypes conditional on local ancestries $g_{s,\text{afr}}, g_{s,\text{eur}}$ as the allele counts specific to each of local ancestries: $g_{s,\text{afr}} := x_{s,M}\mathbb{I}(\gamma_{s,M} = \text{afr}) + x_{s,P}\mathbb{I}(\gamma_{s,P} = \text{afr})$ (similarly for $g_{s,\text{eur}}$). The phenotype of an admixed individual is modeled as a function of allelic effect sizes that are allowed to vary across ancestries:

$$y = \sum_{s=1}^{S} \left(g_{s,\text{afr}}\beta_{s,\text{afr}} + g_{s,\text{eur}}\beta_{s,\text{eur}} \right) + \mathbf{c}^{\top}\boldsymbol{\alpha} + \epsilon, \tag{1}$$
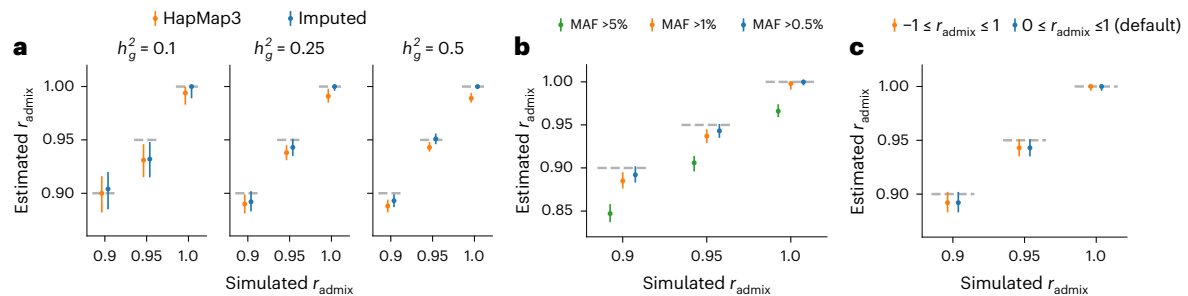
Fig. 2 | Results of genetic correlation r_{admix} estimation in genome-wide simulations. Simulations were based on 17,299 PAGE individuals and 6.9 million genome-wide imputed variants with MAF > 0.5% in both ancestries. We fixed the proportion of causal variants P_{causal} as 0.1% and varied genetic correlation r_{admix} = 0.90, 0.95 and 1.0. **a**, Impact of using HapMap3 or imputed variants in estimation. We varied simulated genome-wide heritability h_g^2 = 0.1, 0.25 and 0.5. **b**, Impact of selecting common variants at different MAF thresholds in estimation. h_g^2 was fixed to 0.25, and imputed variants at different MAF thresholds were used in estimation. **c**, Impact of prior assumption in estimation. h_g^2 was fixed to 0.25, and imputed variants were used in estimation. For each simulated genetic architecture, we plot the mode and 95% credible interval based on the meta-analysis across 100 simulations (Methods). Numerical results are reported in Supplementary Tables 1–4 (including results for other P_{causal}, r_{admix}).

where $\beta_{s,afr}$, $\beta_{s,eur}$ are the causal effects at SNP s, S is the total number of causal SNPs in the genome, **c** and **α** are other covariates (for example, age, sex and genome-wide ancestries) and their effects, and ϵ is the environmental noise. $\beta_{s,afr}$, $\beta_{s,eur}$ are usually referred as allelic effects: change in phenotype with each additional allele. This is in contrast with standardized effects defined as change in phenotype per standard deviation increase of genotype where genotypes at each SNP s are standardized to have unit variance[5,27]. We refrain from using standardized effects in this work due to complexities arising from different ancestries yielding different ancestry-specific frequencies for the same SNP[5] (Methods).

Our goal is to estimate the similarity in the causal effects across local ancestries in admixed populations (Fig. 1); the similarity can be evaluated across all genome-wide causal SNPs that are common across ancestries in a form of cross-ancestry genetic correlation[5,8] (for consistency with previous works we use 'genetic correlation' to refer to correlation of genetic effects across ancestries): $\beta_{s,afr}$, $\beta_{s,eur}$ are modeled as random variables following a bivariate Gaussian distribution parametrized by σ_g^2, ρ_g, denoting the variance and covariance of the effects:

$$\begin{bmatrix} \beta_{s,afr} \\ \beta_{s,eur} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tau_s^2 \cdot \begin{bmatrix} \frac{\sigma_g^2}{S} & \frac{\rho_g}{S} \\ \frac{\rho_g}{S} & \frac{\sigma_g^2}{S} \end{bmatrix} \right), s = 1, \dots, S, \qquad (2)$$

where τ_s are variant-specific parameters determined by the genetic architecture assumption (Methods). Under this model, the genome-wide causal effects correlation is defined as $r_{admix} := \frac{\rho_g}{\sigma_g^2}$; $r_{admix} = 1$ indicates same causal effects across local ancestries, while $r_{admix} < 1$ indicates differences across ancestries. To estimate r_{admix}, given the genotype and phenotype data for a trait, we calculate the profile likelihood curve of r_{admix}, obtained by maximizing the likelihood of model defined by equations (1) and (2) with regard to parameters σ_g^2 and environmental variance for each fixed $r_{admix} \in [0, 1]$. We assume $r_{admix} > 0$ a priori both because causal effects are unlikely to be negatively correlated across ancestries and to reduce r_{admix} search space for reducing computational cost; we have also performed real data analyses to verify this assumption (see below). We obtain the point estimate, credible interval and perform hypothesis testing $H_0 : r_{admix} = 1$ either for each individual trait using the trait-specific profile likelihood curve, or for meta-analysis across multiple traits using the multiplication of the likelihood curves across multiple traits (analogous to inverse variance weighted meta-analysis; Methods).

We organize next sections as follows. First, we show that our proposed approach provides accurate estimation of r_{admix} in extensive simulations. Second, we show r_{admix} is very close to 1 in real data of African-European admixed individuals from PAGE, UKBB and AoU. Third, we replicate our findings using methods that use GWAS summary data (marginal SNP effects at GWAS significant loci). Finally, we investigate pitfalls of methods[4,14,15,28] that use marginal SNP effects showing inflated heterogeneity; we find that Deming regression is the only approach robust enough to quantify r_{admix} from marginal GWAS effects in admixed individuals.

Polygenic method for r_{admix} is accurate in simulations

We performed simulations to evaluate our proposed polygenic method using real genome-wide genotypes. We simulated phenotypes using genotypes and inferred local ancestries with N = 17,299 individuals and S = 6.9 million SNPs (with MAF >0.5% in both ancestries in PAGE dataset; we omitted population-specific rare SNPs to reduce estimation variance; Methods). Phenotypes were simulated under a range of genetic architectures with a frequency-dependent causal effects distribution[29,30], and varying proportion of causal variants P_{causal}, heritability h_g^2 and true r_{admix} (Methods). We used P_{causal} = 0.1% in our main simulations (to simulate a typical polygenic complex trait[31]). When estimating r_{admix}, we either used all SNPs in the imputed genotypes that were used to simulate phenotypes, or restricted to HapMap3 (HM3) SNPs[32] to simulate scenarios where causal variants are not perfectly typed in the data (Methods).

Our method produced accurate point estimates and well-calibrated credible intervals of r_{admix} across a range of simulation settings (Fig. 2a and Supplementary Tables 1 and 2). We first evaluated our method in simulations with a realistic range of h_g^2 = 0.1, 0.25 and 0.5 and r_{admix} = 0.9, 0.95 and 1.0. When using the imputed SNPs for estimation, results were approximately unbiased (average and maximal relative biases across simulation settings were −0.42% and −1.8% respectively). Credible intervals of r_{admix} meta-analyzed across simulations approximately cover true r_{admix}: for the most biased setting (h_g^2 = 0.1, P_{causal} = 0.1%, r_{admix} = 0.95), 95% credible interval 0.915–0.948. When using the HM3 SNPs for estimation, there was a consistent but small downward bias (Fig. 2a; average and maximal relative biases were −1.0% and −2.0%, respectively). This small downward bias was due to imperfect tagging that some of the causal SNPs were not included in the HM3 SNPs. Nonetheless, the magnitude of bias using either imputed or HM3 SNPs was small, indicating our method was accurate and robust to imperfect tagging. We next performed simulations to investigate the potential bias in estimating r_{admix} due to omitting population-specific rare
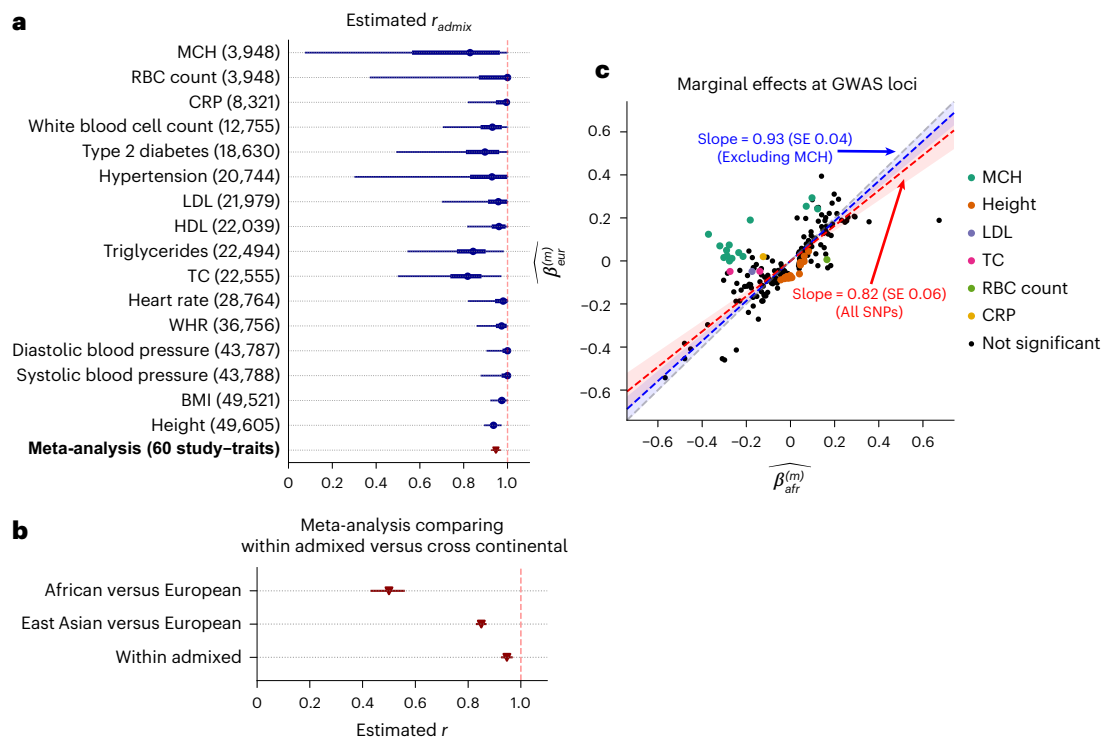
Fig. 3 | Similarity of causal effects and marginal effects across local ancestries meta-analyzed across PAGE, UKBB and AoU. a, We plot the trait-specific estimated r_{admix} for 16 traits. For each trait, dots denote the estimation modes; bold lines and thin lines denote 50%/95% highest density credible intervals, respectively. Traits are ordered according to total number of individuals included in the estimation (shown in parentheses). These traits are selected to be displayed either because they have the largest total sample sizes, or because the associated SNPs of these traits exhibit heterogeneity in marginal effects (see the panel on the right). We also display the meta-analysis results across 60 study–trait pairs (38 unique traits). Numerical results are provided in Table 1. **b,** Comparison of r_{admix} ($n = 38$ traits) to meta-analysis results from transcontinental genetic correlation of African versus European ($n = 26$ traits) and East Asian versus European ($n = 31$ traits). Point estimates and 95% confidence intervals are denoted using triangles and lines. **c,** We plot the

ancestry-specific marginal effects for 217 GWAS significant clumped trait–SNP pairs across 60 study–trait pairs. Trait–SNP pairs with significant heterogeneity in marginal effects by ancestry ($p_{HET} < 0.05/217$ via HET test) are denoted in color (non-significant trait–SNP pairs denoted as black dots; some black dots with large differences across ancestries were not significant because of the large standard errors in estimated effects). Numerical results are reported in Supplementary Table 11. Point estimates and 95% confidence intervals for Deming regression slopes of $\widehat{\beta_{s,\text{eur}}^{(m)}} \sim \widehat{\beta_{s,\text{afr}}^{(m)}}$ are provided either for all 217 SNPs (red), or for 193 SNPs after excluding 24 MCH-associated SNPs (blue). RBC, red blood cell; CRP, C-reactive protein; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TC, total cholesterol; BMI, body mass index; WHR, waist-to-hip ratio.

rare variants. We re-applied our methods using SNPs with MAF >1% and MAF >5% in both populations (in addition to the default MAF >0.5%) to the same simulated data. We observed downward bias in estimated r_{admix} as more stringent MAF threshold was used and more SNPs were filtered out in estimation procedure. For example, the mode of the estimation was 0.966 when methods were applied with MAF >5% in simulation of $r_{admix} = 1.0$ (Fig. 2b and Supplementary Table 3). This indicates omitting population-specific rare variants can lead to downward bias (Discussion). We also investigated the impact of prior assumption of r_{admix}: we applied a revised methodology that allows for $-1 \leq r_{admix} \leq 1$ and we found that estimated r_{admix} were highly consistent when assuming $0 \leq r_{admix} \leq 1$ (default method) versus when assuming $-1 \leq r_{admix} \leq 1$ (Fig. 2c).

We performed several secondary analyses. We determined our method remained accurate at other simulated P_{causal} (Supplementary Table 2; P_{causal} ranging from 0.001% to 1%) and broader range of simulated r_{admix} (Supplementary Table 4; r_{admix} ranging from −0.5 to 1). In null simulations ($r_{admix} = 1$), we determined the false positive rate of hypothesis test $H_0 : r_{admix} = 1$ was properly controlled for most simulation settings, and was only slightly inflated when HM3 SNPs were used, and/ or extremely low P_{causal} was simulated. In simulations with $r_{admix} < 1$, power to detect $r_{admix} < 1$ increased with increasing h_g^2 and decreasing r_{admix} (Supplementary Tables 1 and 2). In addition, we found heritability

can be accurately estimated in these simulations (Supplementary Tables 5 and 6, and Methods). In summary, our method can be reliably used to estimate r_{admix}.

Causal effects are similar across local ancestries

We applied our polygenic method to estimate r_{admix} within African-European admixed individuals in PAGE[1] (24 traits, average $N = 9,296$, average fraction of African ancestries 78%), UKBB[2] (26 traits, average $N = 3,808$, average fraction of African ancestries 59%) and AoU[3] (10 traits, average $N = 20,496$, average fraction of African ancestries 74%) (Methods). Meta-analyzing across 38 traits from PAGE, UKBB and AoU (60 study–trait pairs), we observed a high similarity in causal effects across ancestries ($\hat{r}_{admix} = 0.95$, 95% credible interval 0.93–0.97). Results were highly consistent across datasets despite different ancestry compositions (PAGE: $\hat{r}_{admix} = 0.90$, 0.85–0.94; UKBB: $\hat{r}_{admix} = 0.98$, 0.91–1; AoU: $\hat{r}_{admix} = 0.97$, 0.94–1) as well as across traits (Fig. 3a, Table 1 and Supplementary Table 7). Height was the only trait that had significant $\hat{r}_{admix} < 1$ (after Bonferroni correction; nominal $P = 4.3 \times 10^{-4} < 0.05/38$; meta-analyzed across three datasets; Table 1) albeit with high estimated $\hat{r}_{admix} = 0.936$, 0.89–0.97. Estimates of the same traits across datasets were only weakly correlated (Extended Data Fig. 1), suggesting similar causal effects by ancestry consistently across traits (true $r_{admix} \approx 1$ for all traits).

Table 1 | Genome-wide genetic correlation across 38 complex traits for African-European admixed individuals in PAGE, UKBB and AoU

Trait	N	$\hat{r}_{admix}$ mode	95% credible interval(s)	P value	$\widehat{h_g^2}$
BMD	1,668	0.000	0.00–0.78	0.012	0.34±0.16
Neuroticism	3,044	1.000	0.36–1.00	1	0.36±0.11
Education years	3,324	0.000	0.00–0.94	0.4	0.055±0.075
MCHC	3,650	0.228	0.00–0.87	0.061	0.21±0.092
Type 1 diabetes	3,767	0.381	0.00–0.95	0.77	−0.033±0.016
HLR count	3,852	1.000	0.07–1.00	1	0.12±0.086
RBC distribution width	3,925	1.000	0.27–1.00	1	0.28±0.087
Lymphocyte count	3,935	1.000	0.00–0.60, 0.66–1.00	1	0.13±0.086
Monocyte count	3,935	0.972	0.26–1.00	0.82	0.3±0.087
MCH	3,948	0.829	0.07–1.00	0.36	0.2±0.076
RBC count	3,948	1.000	0.37–1.00	1	0.31±0.09
Hypothyroidism	4,063	1.000	0.05–1.00	1	0.046±0.07
PR interval	4,071	0.844	0.08–1.00	0.36	0.22±0.084
QRS interval	4,078	1.000	0.07–1.00	1	0.12±0.082
Asthma	4,079	1.000	0.15–1.00	1	0.21±0.087
Ever smoked	4,083	0.764	0.04–0.98	0.31	0.17±0.082
QT interval	4,089	0.920	0.07–1.00	0.69	0.16±0.083
HbA1c	5,353	0.954	0.08–1.00c	0.77	0.19±0.078
Cigarettes per day	6,995	0.999	0.08–1.00	1	0.097±0.047
Fasting insulin	7,753	1.000	0.21–1.00	1	0.13±0.044
eGFR	7,978	0.805	0.16–1.00	0.09	0.19±0.046
C-reactive protein	8,321	0.995	0.82–1.00	0.94	0.28±0.046
Fasting glucose	9,646	0.695	0.00–0.93	0.27	0.064±0.035
Coffee consumption	11,587	0.982	0.10–1.00	0.9	0.074±0.03
Platelet count	12,545	0.783	0.20–0.98	0.025	0.19±0.038
White blood cell count	12,755	0.931	0.70–1.00	0.26	0.23±0.036
Type 2 diabetes	18,630	0.897	0.49–1.00	0.23	0.12±0.024
Hypertension	20,744	0.929	0.30–1.00	0.45	0.08±0.027
LDL	21,979	0.958	0.70–1.00	0.55	0.14±0.046
HDL	22,039	0.961	0.82–1.00	0.46	0.22±0.057
Triglycerides	22,494	0.843	0.54–0.98	0.012	0.18±0.027
Total cholesterol	22,555	0.818	0.50–0.97	0.007	0.18±0.039
Heart rate	28,764	0.980	0.82–1.00	0.74	0.099±0.015
WHR	36,756	0.973	0.86–1.00	0.55	0.12±0.015
Diastolic blood pressure	43,787	1.000	0.90–1.00	1	0.077±0.024
Systolic blood pressure	43,788	1.000	0.88–1.00	1	0.071±0.013
BMI	49,521	0.974	0.92–1.00	0.33	0.22±0.02
Height	49,605	0.936	0.89–0.97	0.00043	0.4±0.014
Meta-analysis		0.947	0.93–0.97	8.7×10^{-7}	

For each trait, we report number of individuals, posterior mode and 95% credible interval(s) for estimated r_{admix}, nominal one-sided P value for rejecting the null hypothesis of H_0:r_{admix}=1 (unadjusted for multiple testing; Methods), and estimated heritability and standard error. Meta-analysis results performed across 38 traits are shown in the last row. Traits are ordered according to number of individuals. For each trait, we perform meta-analysis across studies if the trait is in multiple studies (Methods). Lymphocyte count has two credible intervals because of the non-concave profile likelihood curve, as a result of small sample size. BMD, bone mineral density; HLR, high light scattering reticulocytes; MCHC, mean corpuscular hemoglobin concentration.

We performed several secondary analyses. Similar to previous simulation studies, we determined prior assumption of r_{admix} had minimal impact to results: estimated r_{admix} of 24 traits in PAGE were highly consistent when assuming $0 \leq r_{admix} \leq 1$ (default method) versus when assuming $-1 \leq r_{admix} \leq 1$ (Extended Data Fig. 2). Such consistency between the two methods again indicates similar genetic causal effects across local ancestries ($r_{admix} \approx 1$) and that estimation is robust to choices of statistical prior on r_{admix}. Our results were robust to different assumption of effects distribution (Extended Data Fig. 3 and Supplementary Table 8), consistent with previous work[33]. Results were also robust to the SNP set used in the estimation (Extended Data Fig. 3 and Supplementary Table 8), and criterion of the included admixed
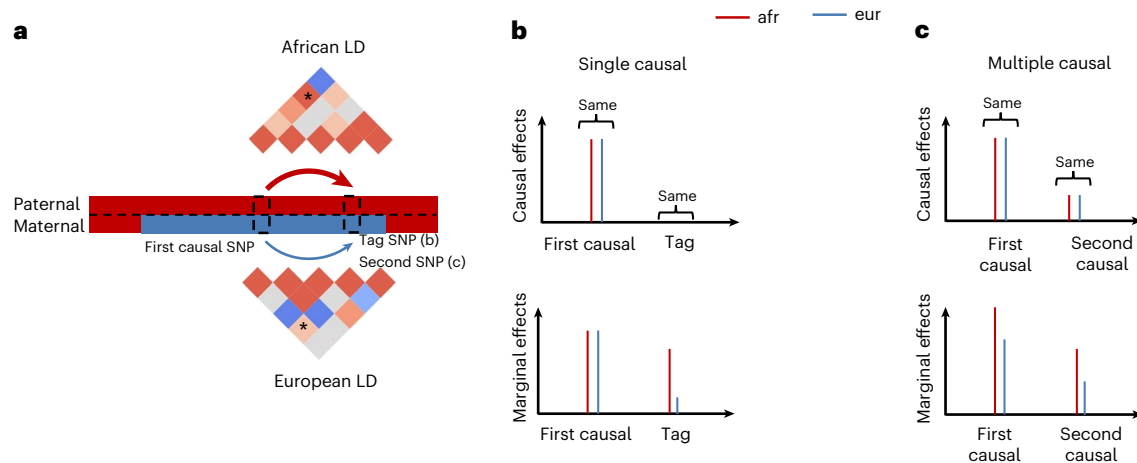
Fig. 4 | Induced heterogeneities in marginal effects across local ancestries.
a, Illustrations that different LD patterns across local ancestries can induce differential tagging between a causal SNP and a tag SNP in **b** or another causal SNP in **c**. LD strengths between the two SNPs are indicated both in the thickness of arrows and in the color shades of '*' elements in LD matrices. **b**, Example of single causal SNP with no heterogeneity. Causal effects are the same across local ancestries, and the estimated marginal effects at causal SNP will be also very similar with sufficient sample size. However, because of differential tagging across local ancestries, the estimated marginal effects evaluated at the tag SNP are different. **c**, Example of multiple causal SNPs with no heterogeneity. Causal effects for both SNPs are the same across local ancestries. In this example, the correlation between the two causal variants is higher for genotypes in African local ancestries than those in European local ancestries. Therefore, African ancestry-specific genotypes tag more effects, creating different ancestry-specific marginal effects at each causal SNP.

individuals (Extended Data Fig. 4). Additionally, an alternative formulation of method assuming different variance component by ancestry did not outperform our default method assuming same variance component by ancestry (Extended Data Fig. 5, Supplementary Table 9 and Supplementary Note).

Next, we contrasted r_{admix} to transcontinental genetic correlations of (1) European versus African and (2) European versus East Asian (Fig. 3b and Methods). We determine a much higher similarity across local ancestries within admixed populations ($\hat{r}_{admix}$ = 0.95, 95% credible interval 0.93–0.97) as compared with transcontinental correlations of African versus European within UKBB ($\hat{r}_{eur-afr}$ = 0.50, meta-analysis across 26 traits, 95% confidence interval 0.43–0.56) and East Asian (Biobank Japan) versus European (UKBB)[8] ($\hat{r}_{eur-eas}$ = 0.85, meta-analysis across 31 traits, 95% confidence interval 0.83–0.87) (Supplementary Table 10). Overall, our results are consistent with r_{admix} being less susceptible to heterogeneity due to differences in phenotyping/environment in transcontinental comparisons.

We sought to replicate high r_{admix} using regression-based methods that leverage estimated ancestry-specific marginal effects at GWAS loci (Methods). Specifically, we used the following marginal regression equation (restricting equation (1) to each GWAS SNP s):

$$y = g_{s,eur}\beta_{s,eur}^{(m)} + g_{s,afr}\beta_{s,afr}^{(m)} + \mathbf{c}^{\top}\boldsymbol{\alpha} + \epsilon$$ (we distinguish marginal effects

$\beta^{(m)}$ from causal effects β; Methods). Across 60 study–trait pairs, we detected 217 GWAS significant clumped trait–SNP pairs and we estimated the ancestry-specific marginal effects for each SNP (Fig. 3c and Supplementary Table 11). We determined the estimated marginal effects are largely consistent by local ancestry at these GWAS clumped SNPs via Deming regression slope[34] of 0.82 (standard error 0.06) (applied to $\widehat{\beta_{s,eur}^{(m)}} \sim \widehat{\beta_{s,afr}^{(m)}}$; Deming regression properly accounts for uncertainty in both dependent and independent variables; Methods). Mean corpuscular hemoglobin (MCH)-associated SNPs at 16p13.3 drove most of the differences by ancestry: Deming regression slope was 0.93 (standard error 0.04) on the rest of 193 SNPs after excluding 24 MCH-associated SNPs; MCH-associated SNPs also have the strongest heterogeneity in marginal effects by ancestry (using heterogeneity score test (HET) for testing effects heterogeneity at each

SNP[35]; Supplementary Table 11 and Methods). By performing statistical fine-mapping analysis, we found there are multiple conditionally independent association signals at MCH-associated and other loci with heterogeneity by ancestry (Extended Data Fig. 6 and Supplementary Note). In fact, the MCH-associated loci locate at a region harboring alpha-globin gene cluster (*HBZ–HBM–HBA2–HBA1–HBQ1*) known to contain multiple causal variants[36]. These results suggest that, similar to causal effects, marginal effects at GWAS loci are also largely consistent by local ancestry across multiple traits, with the exception of 16p13.3 loci for MCH in our study, where multiple large-effect causal variants drive some extent of heterogeneity by ancestry in marginal effects.

Pitfalls of using marginal effects to estimate heterogeneity

Next, we focused on thoroughly evaluating methods that use marginal effects at GWAS significant variants to estimate heterogeneity. Marginal effects are frequently used to compare effect sizes across populations or across studies[4,14,15,28] and enjoy popularity for their simplicity and requirement of only GWAS summary statistics (estimated effect sizes and standard errors).

We first note that heterogeneities in marginal effects can be induced due to different LD patterns across ancestries even when the underlying causal effects are identical, especially when multiple causal variants are nearby in the same LD block (Fig. 4). We investigate the extent of heterogeneity by ancestry that can be induced in simulations with identical causal effects across ancestries, due to (1) local ancestry adjustment; (2) unknown causal variants coupled with ancestry-specific LD patterns; (3) highly polygenic genetic architectures with multiple causal SNPs within the same LD block; (4) standard errors in estimated marginal effects across ancestries. Our following simulations were based on real imputed genotypes from African-European individuals in PAGE data (17,299 individuals, average fraction of African ancestries 78%).

Regressing out local ancestry can deflate the observed similarity in causal effects across ancestries. We first discuss the use of local ancestry in the heterogeneity estimation, which is a unique and important component to consider when studying admixed populations. We used simulations to investigate the role of local ancestry adjustment
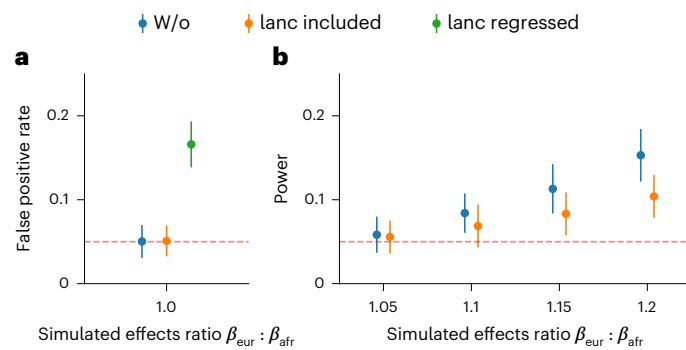
Fig. 5 | Pitfalls of including local ancestry in estimating heterogeneity. In each simulation, we selected a single causal variant and simulated quantitative phenotypes where these causal variants explain heritability $h_g^2 = 0.6\%$; we also varied ratios of effects across ancestries $\beta_{eur} : \beta_{afr}$. **a**, False positive rate in null simulation $\beta_{eur} : \beta_{afr} = 1.0$. **b**, Power to detect $\beta_{eur} \neq \beta_{afr}$ in power simulations with $\beta_{eur} : \beta_{afr} > 1$. We did not include 'lanc regressed' because it is not well-calibrated in null simulations. We plot the mean and 95% confidence intervals, calculated via 100 random subsamplings with each sample consisting of 500 SNPs (Methods). Numerical results are reported in Supplementary Table 12.

using three main approaches: (1) ignoring local ancestry altogether ('w/o'); (2) including local ancestry as covariate in the model ('lanc-included'); (3) regressing out the local ancestry from phenotype followed by heterogeneity estimation on residuals ('lanc-regressed') (Methods). First, in null simulations with identical causal effects (ratio of $\beta_{eur} : \beta_{afr} = 1$), we observed that ignoring local ancestry or including local ancestry as covariate yielded well-calibrated HET tests; in contrast, regressing out the local ancestry effect induced inflated HET test statistics (Fig. 5 and Supplementary Table 12). Next, in power simulations with varying amount of heterogeneity (defined as ratio of $\beta_{eur} : \beta_{afr}$), including local ancestry in the covariate significantly reduced the power of HET test of up to 50% at high magnitude of heterogeneity (Fig. 5 and Supplementary Table 12) (see more details in Supplementary Note). Thus, with respect to local ancestry, we recommend either not using it or including it as a covariate in the model and not regressing out its effect before heterogeneity estimation as that will bias heterogeneity estimation.

Having investigated the role of local ancestry adjustment, we next turn to heterogeneity estimation for GWAS SNPs. We focused on investigating properties of HET test and Deming regression in null simulations with identical causal effects across ancestries ($\beta_{eur} : \beta_{afr} = 1$). Since the true causal variants are usually uncertain, we investigated each method either at the true simulated causal variants or at the LD-clumped variants (Methods).

Uncertainty in which variants are causal can deflate the observed similarity in effects by ancestry. We first performed simulations with single causal variant: we randomly selected one SNP as causal in each simulation. Evaluated at the causal SNPs (Methods), we found that HET test and Deming slope were well-calibrated (Fig. 6a–c, Extended Data Fig. 7 and Supplementary Table 13). However, evaluated at the clumped variants, as a more realistic setting (because causal variants need to be inferred), we found HET test became increasingly miscalibrated with increased h_g^2, while Deming slope remained relatively robust (with an upward but not statistically significant trend with increasing h_g^2). Ordinary least squares (OLS) slope had bias even when evaluated at causal variants because of its ignorance of the standard errors in the estimated effects (Methods and Supplementary Note); such bias became smaller with increased h_g^2.

High polygenicity can deflate the observed similarity in effects by ancestry. Next, we performed simulations where multiple causal

variants locate nearby within the same LD block (typical for polygenic complex traits[37,38]; Methods). In this scenario, marginal GWAS effects could tag multiple causal effects, thus potentially inflating the observed heterogeneity (Fig. 4c). In simulations, we varied the number of causal SNPs from 0.25 to 4.0 per Mb to span most polygenic architectures. In contrast to simulations with a single causal variant, all three methods (HET test, Deming slope and OLS slope) were biased in the presence of multiple nearby causal variants; the miscalibration/bias increased with number of causal variants per region, and LD clumping did not alleviate the miscalibration/bias (Fig. 6d–f). Such miscalibrations occurred irrespective of sample size (Extended Data Fig. 8), or simulated heritability h_g^2 (Supplementary Table 14).

In summary, we find that methods for heterogeneity-by-ancestry estimation based on marginal GWAS SNP effects are susceptible to inflated estimates of heterogeneity. HET test is susceptible to false positives when causal variants are unknown. Deming regression was robust in scenarios with low polygenicity, but was still susceptible to inflated estimates of heterogeneity for highly polygenic traits; the inflated estimates can be explained by differential tagging of causal effects across ancestries among causal SNPs. OLS slope had bias because it did not account for uncertainty in estimated effects. We also performed additional simulations with less than identical causal effects ($\beta_{eur} : \beta_{afr} \neq 1$) and broader range of per-SNP h_g^2 and we determined Deming regression was robust to quantify the heterogeneity level at the marginal effects in simulations of different $\beta_{eur} : \beta_{afr}, h_g^2$ (Extended Data Fig. 9 and Supplementary Table 15).

Discussion

In this work, we developed a polygenic method that model genome-wide causal effects to complex traits of admixed individuals. We determined causal effects are largely similar across local ancestries in analysis of 53,001 African-European admixed individuals across 38 complex traits in PAGE, UKBB and AoU. In addition to causal effects, we also replicated such consistency-by-ancestry for marginal effects at GWAS loci. We highlighted realistic simulation scenarios where regression-based methods using marginal effects can report false heterogeneity when causal effects are identical across ancestries.

Our study has several implications for future genetic study of admixed populations, and more broadly of ancestrally diverse individuals. First, reduced accuracy of polygenic score has been observed in African-European admixed populations with increasing proportion of non-European ancestries[21]; our results suggest the causal effects difference has limited contribution to such reduced accuracy. Second, there has been recent work on incorporating local ancestry in statistical modeling of admixed populations, for example, in association testing[19] and polygenic score[21,22], based on the hypothesis that effects may differ across ancestries. Our results indicate the largely consistent causal effects across local ancestries (and also marginal effects at most GWAS loci). The robustness of our results to imperfect tagging also suggests that imperfect tagging induce limited effects heterogeneity across local ancestries, once SNPs are properly modeled in a polygenic model. The small heterogeneity-by-ancestry at causal effects or marginal effects suggest that association tests that do not model heterogeneity-by-ancestry should be preferred in most cases[19,20] for improved statistical power for association. On the other hand, including local ancestry in association models could be useful in correcting for LD induced by admixture[39] and lead to improved causal effect estimation. Full consideration of incorporating local ancestry in statistical models should also take into account the extent of confounding and heterogeneity in the data[40]. Third, our study further motivates studies of ancestrally diverse individuals to identify population-specific risk variants that cannot be investigated due to being rare in European individuals; for example, inclusion of individuals with diverse populations could further disentangle causal from tagging effects, thus increasing the power of heterogeneity-by-ancestry estimation. More importantly,
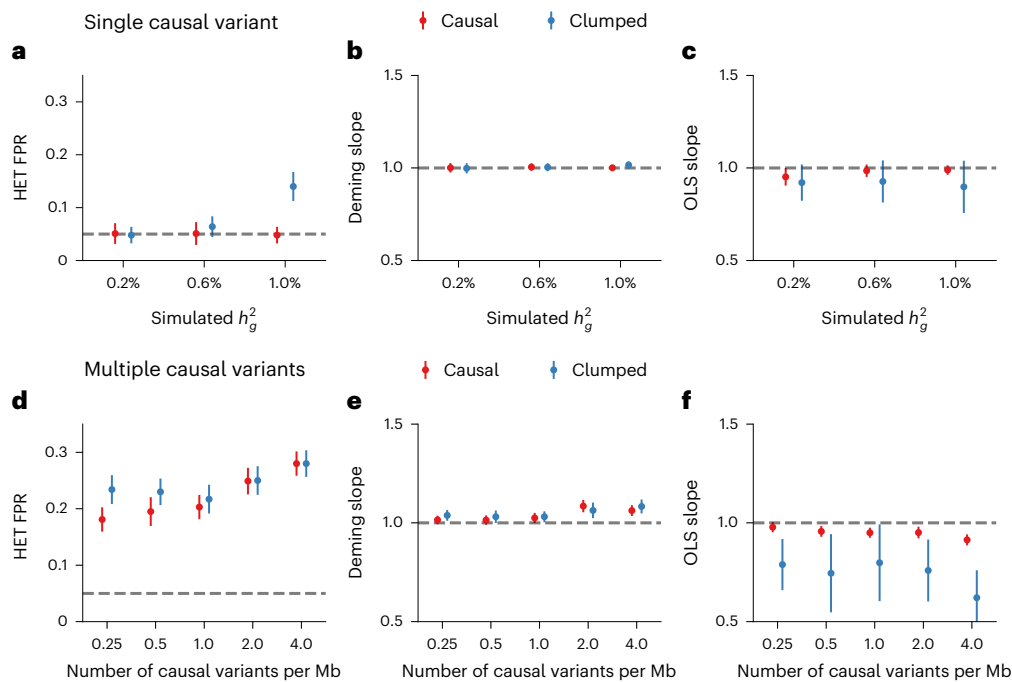
Fig. 6 | Miscalibration of HET test/Deming regression/OLS regression in simulations with $r_{admix} = 1$. a–c, Simulations with single causal variant. Each causal variant had the same causal effects across local ancestries and each causal variant explained a fixed amount of heritability (0.2%, 0.6% and 1.0%): false positive rate (FPR) of HET test (**a**); Deming regression slope (**b**) and of OLS regression slope (**c**) of $\widehat{\beta_{eur}^{(m)}} \sim \widehat{\beta_{afr}^{(m)}}$. Numerical results are reported in Supplementary Table 13. **d–f**, Simulation with multiple causal variants, where we simulated different levels of polygenicity, such that on average there were approximately 0.25, 0.5, 1.0, 2.0 and 4.0 causal variants per Mb; causal variants had the same causal effects across local ancestries, and the heritability explained by all causal variants was fixed at $h_g^2 = 10\%$: FPR of HET test (**d**); Deming regression slope (**e**) and OLS regression slope (**f**) of $\widehat{\beta_{eur}^{(m)}} \sim \widehat{\beta_{afr}^{(m)}}$. The 95% confidence intervals were based on 100 random subsamplings with each sample consisting of 1,000 SNPs (Methods). Results for other number of SNPs used for subsampling are shown in Extended Data Fig. 8. Numerical results are reported in Supplementary Table 14.

larger and robust trans-ancestry studies may allow for the examination of differential causal effects on a locus-by-locus basis, in addition to the genome-wide approach as presented in this work.

Our results add to the existing literature to further delineate sources of causal effects differences. Previous works have shown moderate causal effects differences across transcontinental populations[5,6,8,28], with part of differences being induced by heterogeneity in the definition of environment/phenotype across continental ancestries. Similarly, a recent work[15] concluded differences between causal effects in European local ancestries within African American admixed individuals and that in European American individuals. Our results showcase that, if environments are well controlled (as is the case for genetic variants across local ancestries within admixed populations), causal effects are highly similar across genetic ancestries, agreeing with a recent study finding similar effects across ancestries at level of gene expression in controlled environments[41]. Moreover, our results suggest that local epistatic interaction, if any, does not lead to large causal effects differences across genetic ancestries. By contrasting the high genetic correlation within admixed populations and the low genetic correlation across continental populations, our results support the hypothesis that different environments modify the genetic effects to complex traits (gene-by-environment interaction) across populations.

We note several limitations and future directions of our work. First, we have analyzed SNPs with MAF ≥0.5% in both ancestries. We excluded population-specific SNPs (with MAF <0.5% in one of the ancestries) because these SNPs provide little information for estimating r_{admix}, since effects for these SNPs are estimated with large noises. We used simulations to show that omitting these rare variants could lead to downward bias in r_{admix} estimation because of population-specific tagging of shared causal variants (Supplementary Note). However, it remains possible that causal variants themselves are rare and population-specific,

and upward bias in the estimation of r_{admix} may be present. While in this work we focused on estimating r_{admix} for common variants, future work with larger sample sizes is needed to further investigate the impact of population-specific causal SNPs to r_{admix} estimation. Second, we have considered two-way African-European admixed individuals. Several practical considerations remain before applying this method to other admixed populations such as three-way admixture: local ancestries are typically inferred with larger errors[42], and this should be accounted for in statistical modeling (it may be possible to incorporate posterior probabilities in estimated local ancestries to obtain calibrated estimates); additional parameters need to be estimated (for example, three pairwise correlation parameters across ancestries for three-way admixture populations). We note that our methods can be readily applied to these populations when reliable local ancestry calls can be obtained. Third, our modeling can be extended to estimate correlations in causal effects stratified by functional annotation categories and we leave that as future work. Fourth, our polygenic method requires individual-level genotype and phenotype; if not available, we found Deming regression may be applied to evaluate heterogeneity with caution: in our simulation, Deming regression was the only method robust to most scenarios except for high polygenicity. In our analysis of marginal effects, we found LD clumping can produce cluster of SNPs that were nearby and probably dependent with each other, as a combined result of multiple causal variants within a region and long-range LD in admixed populations. Such dependence may induce bias for methods like Deming regression, highlighting the need for improved methods of identifying conditionally independent SNPs in admixed populations. Fifth, we have meta-analyzed three publicly available studies of PAGE, UKBB and AoU with large cohort of African-European admixed individuals. Such meta-analysis with greatly increased total sample size enabled us to derive the conclusion of the high similarity

in causal effects by local ancestry across a broad range of traits. However, our estimates for each individual trait were still associated with large standard errors and can be further improved by analyzing more individuals. Additional limitations are discussed in Supplementary Note. Despite these limitations, our study has shown that causal effects to complex traits are highly similar across local ancestries, and this knowledge can be used to guide future genetic studies of ancestrally diverse populations.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-023-01338-6.

References

1. Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).
2. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
3. Ramirez, A. H. et al. The All of Us Research Program: data quality, utility, and diversity. *Patterns* **3**, 100570 (2022).
4. Zhou, W. et al. Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. *Cell Genomics* **2**, 100192 (2022).
5. Brown, B. C., Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic genetic-correlation estimates from summary statistics. *Am. J. Hum. Genet.* **99**, 76–88 (2016).
6. Galinsky, K. J. et al. Estimating cross-population genetic correlations of causal effect sizes. *Genet. Epidemiol.* **43**, 180–188 (2019).
7. Shi, H. et al. Localizing components of shared transethnic genetic architecture of complex traits from GWAS summary data. *Am. J. Hum. Genet.* **106**, 805–817 (2020).
8. Shi, H. et al. Population-specific causal disease effect sizes in functionally important regions impacted by selection. *Nat. Commun.* **12**, 1098 (2021).
9. Kanai, M. et al. Insights from complex trait fine-mapping across diverse populations. Preprint at *medRxiv* https://doi.org/10.1101/2021.09.03.21262975 (2021).
10. Wang, Y. et al. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. *Nat. Commun.* **11**, 3865 (2020).
11. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
12. Gurdasani, D., Barroso, I., Zeggini, E. & Sandhu, M. S. Genomics of disease risk in globally diverse populations. *Nat. Rev. Genet.* **20**, 520–535 (2019).
13. Sirugo, G., Williams, S. M. & Tishkoff, S. A. The missing diversity in human genetic studies. *Cell* **177**, 1080 (2019).
14. Marigorta, U. M. & Navarro, A. High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS Genet.* **9**, e1003566 (2013).
15. Patel, R. A. et al. Genetic interactions drive heterogeneity in causal variant effect sizes for gene expression and complex traits. *Am. J. Hum. Genet.* **109**, 1286–1297 (2022).
16. Cai, N. et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
17. Seldin, M. F., Pasaniuc, B. & Price, A. L. New approaches to disease mapping in admixed populations. *Nat. Rev. Genet.* **12**, 523–528 (2011).
18. Mills, M. C. & Rahal, C. The GWAS Diversity Monitor tracks diversity by disease in real time. *Nat. Genet.* **52**, 242–243 (2020).
19. Atkinson, E. G. et al. Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nat. Genet.* **53**, 195–204 (2021).
20. Hou, K., Bhattacharya, A., Mester, R., Burch, K. S. & Pasaniuc, B. On powerful GWAS in admixed populations. *Nat. Genet.* **53**, 1631–1633 (2021).
21. Bitarello, B. D. & Mathieson, I. Polygenic scores for height in admixed populations. *G3* **10**, 4027–4036 (2020).
22. Marnetto, D. et al. Ancestry deconvolution and partial polygenic score can improve susceptibility predictions in recently admixed individuals. *Nat. Commun.* **11**, 1628 (2020).
23. Bentley, A. R. et al. Gene-based sequencing identifies lipid-influencing variants with ethnicity-specific effects in African Americans. *PLoS Genet.* **10**, e1004190 (2014).
24. Rajabli, F. et al. Ancestral origin of ApoE ε4 Alzheimer disease risk in Puerto Rican and African American populations. *PLoS Genet.* **14**, e1007791 (2018).
25. Blue, E. E., Horimoto, A. R. V. R., Mukherjee, S., Wijsman, E. M. & Thornton, T. A. Local ancestry at APOE modifies Alzheimer's disease risk in Caribbean Hispanics. *Alzheimers Dement.* **15**, 1524–1532 (2019).
26. Naslavsky, M. S. et al. Global and local ancestry modulate APOE association with Alzheimer's neuropathology and cognitive outcomes in an admixed sample. *Mol. Psychiatry* **27**, 4800–4808 (2022).
27. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
28. Sakaue, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
29. Zeng, J. et al. Signatures of negative selection in the genetic architecture of human complex traits. *Nat. Genet.* **50**, 746–753 (2018).
30. Schoech, A. P. et al. Quantification of frequency-dependent genetic architectures in 25 UK Biobank traits reveals action of negative selection. *Nat. Commun.* **10**, 790 (2019).
31. Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nat. Genet.* **50**, 1318–1326 (2018).
32. The International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
33. Speed, D. & Balding, D. J. SumHer better estimates the SNP heritability of complex traits from summary statistics. *Nat. Genet.* **51**, 277–284 (2019).
34. Deming, W. E. Statistical adjustment of data. Wiley. (1943).
35. Pasaniuc, B. et al. Enhanced statistical tests for GWAS in admixed populations: assessment using African Americans from CARe and a Breast Cancer Consortium. *PLoS Genet.* **7**, e1001371 (2011).
36. Hodonsky, C. J. et al. Ancestry-specific associations identified in genome-wide combined-phenotype study of red blood cell traits emphasize benefits of diversity in genomics. *BMC Genomics* **21**, 228 (2020).
37. Loh, P.-R. et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat. Genet.* **47**, 1385–1392 (2015).
38. Johnson, R. et al. Estimation of regional polygenicity from GWAS provides insights into the genetic architecture of complex traits. *PLoS Comput. Biol.* **17**, e1009483 (2021).
39. Zhang, J. & Stram, D. O. The role of local ancestry adjustment in association studies using admixed populations. *Genet. Epidemiol.* **38**, 502–515 (2014).

40. Liu, J., Lewinger, J. P., Gilliland, F. D., Gauderman, W. J. & Conti, D. V. Confounding and heterogeneity in genetic association studies with admixed populations. *Am. J. Epidemiol.* **177**, 351–360 (2013).
41. Saitou, M., Dahl, A., Wang, Q. & Liu, X. Allele frequency differences of causal variants have a major impact on low cross-ancestry portability of PRS. Preprint at *medRxiv* https://doi.org/10.1101/2022.10.21.22281371 (2022).
42. Pasaniuc, B. et al. Analysis of Latino populations from GALA and MEC studies reveals genomic loci with biased local ancestry estimation. *Bioinformatics* **29**, 1407–1415 (2013).

[1]Bioinformatics Interdepartmental Program, UCLA, Los Angeles, CA, USA. [2]Department of Computer Science, UCLA, Los Angeles, CA, USA. [3]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. [4]Graduate Program in Biomathematics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. [5]Institute for Genomic Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [6]The Charles Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [7]Department of Statistics, Rutgers University, Piscataway, NJ, USA. [8]Center for Genetic Epidemiology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [9]Division of Public Health Science, Fred Hutchinson Cancer Center, Seattle, WA, USA. [10]Brown Foundation Institute for Molecular Medicine, The University of Texas Health Science Center, Houston, TX, USA. [11]Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Denver, CO, USA. [12]Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Lundquist Institute at Harbor-UCLA Medical Center, Torrance, CA, USA. [13]Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [14]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [15]Department of Medicine, University of Colorado, Aurora, CO, USA. [16]Center for Public Health Genomics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA. [17]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [18]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [19]Department of Biology, Loyola University Chicago, Chicago, IL, USA. [20]Program in Bioinformatics, Loyola University Chicago, Chicago, IL, USA. [21]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. [22]Department of Computational Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. [23]Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. [24]These authors contributed equally: Gillian M. Belbin, Steve Buyske, David V. Conti, Burcu F. Darst, Myriam Fornage, Chris Gignoux, Xiuqing Guo, Christopher Haiman, Eimear E. Kenny, Michelle Kim, Charles Kooperberg, Leslie Lange, Ani Manichaikul, Kari E. North, Ulrike Peters, Laura J. Rasmussen-Torvik, Stephen S. Rich, Jerome I. Rotter, Heather E. Wheeler, Genevieve L. Wojcik, Ying Zhou. ✉e-mail: houkc@ucla.edu; pasaniuc@ucla.edu

Methods

Ethical approval

This research complies with all relevant ethical regulations. Ethics committee/institutional review board (IRB) of PAGE gave ethical approval for collection of PAGE data. Ethics committee/IRB of UKBB gave ethical approval for collection of UKBB data (https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics). Approval to use UKBB individual level in this work was obtained under application 33297 at http://www.ukbiobank.ac.uk. Ethics committee/IRB of AoU gave ethical approval for collection of AoU data (https://allofus.nih.gov/about/who-we-are/institutional-review-board-irb-of-all-of-us-research-program). Approval to use AoU controlled tier data in this work was obtained through application at https://www.researchallofus.org.

Statistical model of phenotype for admixed individuals

For individual $i = 1, \ldots, N$ and SNP $s = 1, \ldots, S$, we denote $x_{i,s,M}, x_{i,s,P}$ as number of minor alleles at maternal and paternal haplotypes, respectively. We denote corresponding local ancestries as $\gamma_{i,s,M}, \gamma_{i,s,P} \in \{1, 2\}$ (we focus on two-way admixture here, for example, '1' and '2' denote African and European ancestries for African-European admixture). Then we use $g_{i,s,1}, g_{i,s,2}$ to encode allele counts that are specific to each local ancestry:

$$g_{i,s,1} := x_{i,s,M}\mathbb{I}\left(\gamma_{i,s,M} = 1\right) + x_{i,s,P}\mathbb{I}\left(\gamma_{i,s,P} = 1\right);$$

$$g_{i,s,2} := x_{i,s,M}\mathbb{I}\left(\gamma_{i,s,M} = 2\right) + x_{i,s,P}\mathbb{I}\left(\gamma_{i,s,P} = 2\right),$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. Denoting causal allelic effects as $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^S$ for two ancestries, we model the phenotype of each individual y_i as

$$y_i = \mathbf{c}_i^\top \boldsymbol{\alpha} + \sum_{s=1}^{S} \left(g_{i,s,1}\beta_{s,1} + g_{i,s,2}\beta_{s,2}\right) + \epsilon_i, i = 1, \ldots, N$$

where $\mathbf{c}_i \in \mathbb{R}^C, \boldsymbol{\alpha} \in \mathbb{R}^C$ denote C covariates (including all '1' intercepts) and their effects. ϵ_i denotes environmental noise. By further aggregating $g_{i,s,1}, g_{i,s,2}$ into matrices $\mathbf{G}_1 \in \{0, 1, 2\}^{N \times S}$ and $\mathbf{G}_2 \in \{0, 1, 2\}^{N \times S}$ for ancestry 1 and 2, and $\mathbf{c}_i$ into $\mathbf{C} \in \mathbb{R}^{N \times C}$, equation (1) becomes

$$\mathbf{y} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{G}_1\boldsymbol{\beta}_1 + \mathbf{G}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon} \tag{3}$$

We pose the following distribution assumptions $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ and $\boldsymbol{\epsilon}$

$$\begin{bmatrix} \beta_{s,1} \\ \beta_{s,2} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tau_s^2 \cdot \begin{bmatrix} \sigma_g^2/S & \rho_g/S \\ \rho_g/S & \sigma_g^2/S \end{bmatrix}\right), s = 1, \ldots, S, \epsilon_i \sim N(0, \sigma_e^2), i = 1, \ldots, N \tag{4}$$

where σ_g^2 denotes variance of effects for both populations, ρ_g denotes covariance for similarity of effect sizes by ancestry, and σ_e^2 denotes the variance for environments. τ_s denote SNP-specific parameters (fixed a priori) for effect sizes distribution (see 'Specifying τ_s under different heritability models' below). We define correlation of causal genetic effects as $r_{\text{admix}} = \frac{\rho_g}{\sigma_g^2}$. $r_{\text{admix}} = 1$ indicates $\beta_{s,1} = \beta_{s,2}$ for all variants $s = 1, \ldots, S$, that is, causal effects are the same across ancestries; $r_{\text{admix}} < 1$ indicates differences in causal effects across ancestries.

Calculating and filtering by ancestry-specific allele frequencies.

For each SNP s, we calculated MAF as $f_s := \frac{\sum_{i=1}^{N}(g_{i,s,1}+g_{i,s,2})}{2N}$. We also calculated ancestry-specific MAF as $\frac{\sum_{i=1}^{N} g_{i,s,1}}{\sum_{i=1}^{N}[\mathbb{I}(\gamma_{i,s,M}=1)+\mathbb{I}(\gamma_{i,s,P}=1)]}, \frac{\sum_{i=1}^{N} g_{i,s,2}}{\sum_{i=1}^{N}[\mathbb{I}(\gamma_{i,s,M}=2)+\mathbb{I}(\gamma_{i,s,P}=2)]}$

for ancestry 1 and 2. For a SNP s with close-to-zero frequency for either of the ancestry, its effect β_s will be estimated with very large noise. Therefore, we used SNPs with MAF >0.5% in both ancestries in analyses.

Specifying τ_s under different heritability models.

τ_s parameters model the coupling of SNP effects variance with MAF, local LD or other functional annotations. Commonly used heritability models include GCTA[43], frequency-dependent[29,30], LDAK[44] and S-LDSC[45] models. While heritability model is important to estimate heritability and functional enrichment of heritability[33,46,47], genetic correlation estimation, the main focus of this study, has shown to be robust to different heritability models[33]. In this work, we mainly used the frequency-dependent model for both simulations and real data analyses (where $\tau_s^2 \propto [f_s(1-f_s)]^\alpha; f_s$ is the MAF of the SNP s and $\alpha = -0.38$ is estimated in a meta-analysis across 25 UKBB complex traits[30]). For real data analysis, we additionally used GCTA model for estimation and found results are robust to heritability models (Extended Data Fig. 3).

Alternative choice of genotype normalization by ancestry.

We discuss an alternative choice of normalization by ancestry, in which we have two parameters $\tau_{s,1}$ and $\tau_{s,2}$ separately for two ancestries for each SNP. For example, $\tau_{s,1}^2 \propto \frac{1}{f_{s,1}(1-f_{s,1})}, \tau_{s,2}^2 \propto \frac{1}{f_{s,2}(1-f_{s,2})}$ parametrizing effects distribution

$$\begin{bmatrix} \beta_{s,1} \\ \beta_{s,2} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{s,1}^2 \cdot \sigma_g^2/S & \tau_{s,1}\tau_{s,2} \cdot \rho_g/S \\ \tau_{s,1}\tau_{s,2} \cdot \rho_g/S & \tau_{s,2}^2 \cdot \sigma_g^2/S \end{bmatrix}\right), s = 1, \ldots, S$$

This implies that effects per genotype standard deviation is being modeled (ref. [5] termed this as correlation of allelic impact). While genetic correlation estimation is robust to genotype standardization (Supplementary Table 8; refs. [5,33]), we recommend modeling allelic effects via same τ_s across ancestries (as used in our default method).

Evaluation of genome-wide genetic effects consistency

We discuss parameter estimation and hypothesis testing in equations (3) and (4). Marginalizing over random effects $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ in equation (3), the distribution of $\mathbf{y}$ is

$$\mathbf{y} \sim N\left(\mathbf{C}\boldsymbol{\alpha}, \sigma_g^2 \frac{\mathbf{G}_1\mathbf{T}\mathbf{G}_1^\top + \mathbf{G}_2\mathbf{T}\mathbf{G}_2^\top}{S} + \rho_g \frac{\mathbf{G}_1\mathbf{T}\mathbf{G}_2^\top + \mathbf{G}_2\mathbf{T}\mathbf{G}_1^\top}{S} + \sigma_e^2 \mathbf{I}\right).$$

where $\mathbf{T}$ is a diagonal matrix with $(\mathbf{T})_{ss} = \tau_s^2$. By denoting $\mathbf{K}_1 = \frac{\mathbf{G}_1\mathbf{T}\mathbf{G}_1^\top + \mathbf{G}_2\mathbf{T}\mathbf{G}_2^\top}{S}, \mathbf{K}_2 = \frac{\mathbf{G}_1\mathbf{T}\mathbf{G}_2^\top + \mathbf{G}_2\mathbf{T}\mathbf{G}_1^\top}{S}$, and $\rho_g = \sigma_g^2 \cdot r_{\text{admix}}$, the distribution of $\mathbf{y}$ is simplified as

$$y \sim N\left(\mathbf{C}\boldsymbol{\alpha}, \sigma_g^2(\mathbf{K}_1 + r_{\text{admix}}\mathbf{K}_2) + \sigma_e^2\mathbf{I}\right). \tag{5}$$

The maximum likelihood estimates of $(\boldsymbol{\alpha}, \sigma_g^2, r_{\text{admix}}, \sigma_e^2)$ can be found by directly maximizing the corresponding likelihood function $L\left(\boldsymbol{\alpha}, \sigma_g^2, r_{\text{admix}}, \sigma_e^2\right)$. However, the constraint that the correlation parameter r_{admix} should be small than 1 cannot be easily incorporated here. Instead, we use the profile likelihood $L_p(r_{\text{admix}}) := \max_{(\boldsymbol{\alpha}, \sigma_g^2, \sigma_e^2)} L\left(\boldsymbol{\alpha}, \sigma_g^2, r_{\text{admix}}, \sigma_e^2\right)$

and perform grid search of r_{admix} to maximize profile likelihood (similar to ref. [30]): for each candidate r_{admix}, we compute $\mathbf{K}_1 + r_{\text{admix}}\mathbf{K}_2$, and solve $(\boldsymbol{\alpha}, \sigma_g^2, \sigma_e^2)$ for the single variance component model in equation (5) using GCTA[27] (v1.94.0beta). In practice, we calculate profile likelihood $L_p(r_{\text{admix}})$ for a predefined set of $r_{\text{admix}} = 0.00, 0.05, \ldots, 1.00$ ($r_{\text{admix}} \in [0, 1]$ is a reasonable prior assumption here; we alternatively used an extended range of $r_{\text{admix}} = -1, -0.95, \ldots, 0.95, 1.0$ in simulation studies (Supplementary Table 4) and real data analyses (Extended Data Fig. 2)). We use natural cubic spline to interpolate pairs of $(r_{\text{admix}}, L_p(r_{\text{admix}}))$ to get a likelihood curve of r_{admix}. Then we obtain the estimated $\hat{r}_{\text{admix}}$ using the value that maximize the likelihood curve, and credible interval by combining the likelihood curve with a uniform prior of $r_{\text{admix}} \sim \text{Uniform}[0, 1]$ and calculating the highest posterior density interval as credible interval. To perform the meta-analysis across independent estimates, we obtain the joint likelihood by calculating the product of likelihood curves across estimates (or

equivalently, the sum of log-likelihood curves), and similarly calculate the estimate and credible interval.

Evaluation of genetic effects consistency at individual variant with marginal effects

Parameter estimation and hypothesis testing. We use a model between individual SNP and phenotype by restricting equation (1) to the SNP of interest s, as

$$y_i = \mathbf{c}_i^\top \boldsymbol{\alpha} + \left(g_{i,s,1}\beta_{s,1}^{(m)} + g_{i,s,2}\beta_{s,2}^{(m)} \right) + \epsilon_i, \, i = 1, \dots, N,$$

or in vector form,

$$y = \mathbf{C}\boldsymbol{\alpha} + \mathbf{g}_{s,1}\beta_{s,1}^{(m)} + \mathbf{g}_{s,2}\beta_{s,2}^{(m)} + \epsilon \qquad (6)$$

where $\mathbf{C}, \mathbf{g}_{s,1}, \mathbf{g}_{s,2}, \epsilon$ contain $\mathbf{c}_i, g_{i,s,1}, g_{i,s,2}, \epsilon_i$ for all individuals $i = 1, \dots, N$, respectively. We distinguish marginal effects $\beta_{s,1}^{(m)}, \beta_{s,2}^{(m)}$ in equation (6) from causal effects $\beta_{s,1}, \beta_{s,2}$ in Eq. (1): marginal effects tag effects from nearby causal SNPs with taggability as a function of ancestry-specific correlation between the focal SNP and nearby causal SNPs. Therefore, heterogeneity in marginal effects by local ancestry can be induced even if causal effects are the same (see extensive simulations in Results and more details in Supplementary Note). We estimate $\beta_{s,1}^{(m)}, \beta_{s,2}^{(m)}$ using least squares (jointly for $\beta_{s,1}^{(m)}, \beta_{s,2}^{(m)}$) and perform hypothesis testing of $H_0 : \beta_{s,1}^{(m)} = \beta_{s,2}^{(m)}$ with a likelihood ratio test by comparing Eq. (6) to a restricted model where the allelic effects are the same $\beta_s^{(m)} = \beta_{s,1}^{(m)} = \beta_{s,2}^{(m)}$:

$$y = \mathbf{C}\boldsymbol{\alpha} + \left(\mathbf{g}_{s,1} + \mathbf{g}_{s,2} \right) \beta_s^{(m)} + \epsilon \qquad (7)$$

Marginal effects-based methods for estimating heterogeneity. We describe details of marginal effects-based methods to estimate heterogeneity with input from a set of estimated effect sizes $\widehat{\beta_{s,1}^{(m)}}, \widehat{\beta_{s,2}^{(m)}}$ and corresponding estimated standard errors $\widehat{se(\beta_{s,1}^{(m)})}, \widehat{se(\beta_{s,2}^{(m)})}$ for a set of SNPs.

- Pearson correlation: by calculating the Pearson correlation of $\widehat{\beta_{s,1}^{(m)}}, \widehat{\beta_{s,2}^{(m)}}$ across SNPs. Pearson correlation does not model errors in estimated effects, therefore is expected be smaller than 1 and decreases with increasing error magnitude.
- OLS regression slope: by regressing $\widehat{\beta_{s,1}^{(m)}} \sim \widehat{\beta_{s,2}^{(m)}}$ ($\widehat{\beta_{s,1}^{(m)}}$ as dependent variable, $\widehat{\beta_{s,2}^{(m)}}$ as independent variable) or $\widehat{\beta_{s,2}^{(m)}} \sim \widehat{\beta_{s,1}^{(m)}}$. It does not model errors in independent variable. Moreover, it assumes homogeneous errors in dependent variable across SNPs. Therefore, it is susceptible to these error terms and notably results can vary when one exchange the regression orders[48] ($\widehat{\beta_{s,1}^{(m)}} \sim \widehat{\beta_{s,2}^{(m)}}$ versus $\widehat{\beta_{s,2}^{(m)}} \sim \widehat{\beta_{s,1}^{(m)}}$; for example, $\widehat{\beta_{s,1}^{(m)}}$ and $\widehat{\beta_{s,2}^{(m)}}$ are associated with different standard errors when being estimated in an admixed population with different ancestry proportion).
- Deming regression slope: obtained with Deming regression[34] of $\widehat{\beta_{s,1}^{(m)}}, \widehat{\beta_{s,2}^{(m)}}$ and estimated standard errors $\widehat{se(\beta_{s,1}^{(m)})}, \widehat{se(\beta_{s,2}^{(m)})}$. Deming regression models heterogeneous error terms in both independent and dependent variables, therefore is more robust than Pearson correlation and OLS regression. Specifically, given a set of data and estimated standard errors $(x_i, y_i, \sigma_{x,i}, \sigma_{y,i}), i = 1, \dots, n$ (we use a different set of notations for simplicity), Deming regression optimizes the following objective function to obtain estimated intercept α and slope β:

$$\min_{\alpha, \beta} \quad \sum_{i=1}^{n} \left[\frac{\epsilon_i^2}{\sigma_{y,i}^2} + \frac{\delta_i^2}{\sigma_{x,i}^2} \right],$$

$$\delta_1, \dots, \delta_n$$

$$\epsilon_i, \dots, \epsilon_n$$

$$\text{subject to} : y_i + \epsilon_i = \alpha + \beta \left(x_i + \delta_i \right), i = 1, \dots, n.$$

Standard errors of α, β can be obtained with bootstrapping. Notably, Deming regression slope produce symmetric results with different regression orders (the obtained slope β will be reciprocal to each other). However, Deming regression can still produce biased results when the standard errors $\sigma_{x,i}, \sigma_{y,i}$ are misspecified[48].

- False positive rate of the HET test, as described above in 'Parameter estimation and hypothesis testing'. It is expected to be well calibrated under the null, because its derivation as a likelihood ratio test. Similar to Deming regression, HET test properly models heterogeneous standard errors.

Genotype data processing

PAGE genotype. We analyzed 17,299 genotyped individuals self-identified as African American in PAGE study[1]. These individuals were from three studies: Women's Health Initiative ($N = 6,820$), Multi-ethnic Cohort ($N = 5,325$) and the Icahn School of Medicine at Mount Sinai BioMe biobank in New York City (BioMe) ($N = 5,154$). See more details in ref. [1]. The genotypes were imputed to the TOPMed reference panel and we retained well-imputed SNPs with imputation $R^2 > 0.8$ and MAF >0.5%. We further retained variants with ancestry-specific MAF > 0.5% in both ancestries. This resulted in ~6.9 million variants and 17,299 individuals in our analysis.

UKBB genotype. We analyzed individuals with African-European admixed ancestries in UKBB. We first inferred the proportion of ancestries for each individual in UKBB using SCOPE[49] (https://github.com/sriramlab/SCOPE; version 6 December 2021) supervised using 1,000 Genomes Phase 3 allele frequencies (AFR, EUR, EAS and SAS). We retained 4,327 African-European admixed individuals with more than 5% of both AFR and EUR ancestries, and with less than 5% of both EAS and SAS ancestries. We retained well-imputed SNPs with imputation $R^2 > 0.8$ and MAF >0.5%. We further retained variants with ancestry-specific MAF >0.5% in both ancestries. This resulted in ~6.6 million variants and 4,327 individuals in our analysis.

AoU genotype. We analyzed individuals with African-European admixed ancestries in AoU. We first performed principal component analysis of all 165,208 individuals in AoU microarray data (release v5) joint with 1,000 Genomes Phase 3 reference panel. Then we identified 31,375 individuals with African-European admixed ancestries (with at least both 10% European ancestries and 10% African ancestries, and who was within 2× normalized distance from the line connecting individuals of European ancestries and African ancestries in 1,000 Genomes reference panel; Supplementary Note). For these individuals, we performed quality control using PLINK2 (ref. [50]) (v2.0a3) with `--geno 0.05 --max-alleles 2 --maf 0.001`, and statistical phasing using Eagle2 (ref. [51]) (v2.4.1) with default settings. We retained variants with ancestry-specific MAF >0.5% in both ancestries. This resulted in ~0.65 million variants and 31,375 individuals in our analysis. For AoU, we chose to use microarray data instead of whole genome sequencing data because microarray data of AoU contained more individuals and analyzing microarray data reduced the computational cost.

Local ancestry inference. We performed local ancestry inference using RFMix[52] (https://github.com/slowkoni/rfmix; v2) with default parameters (eight generations since admixture). We used 99 CEU individuals (Utah residents with Northern and Western European ancestry) and 108 YRI individuals (individuals from Yoruba in Ibadan, Nigeria) from unrelated individuals in 1,000 Genome Project Phase 3 (ref. [53]) as our reference populations, similar to previous works[52,54]. We used HapMap3 SNPs[32] in inference, and then interpolated the inferred local ancestry results to other variants in both PAGE and UKBB data sets. The accuracy of RFMix for local ancestry inference has been validated for African-European admixed individuals[19] (for example,

~98% accuracy for simulations with a realistic demographic model for African American individuals). We performed additional analyses using PAGE African American individuals to assess the robustness of local ancestry inference using an alternative set of reference data. We used all European and African individuals in 1,000 Genomes project (excluding African Caribbean in Barbados and African Ancestry in SW USA because they were admixed). We determined a high consistency of 98.9% for the inferred local ancestry using reference data of CEU/YRI or all European/African individuals. We used the inferred local ancestry for both simulation study and real data analysis described below.

Simulation study

We describe methods for simulations that corresponds to each section of Results.

Pitfalls of including local ancestry in estimating heterogeneity.

We first describe strategies of including local ancestry in estimating heterogeneity.

- For 'lanc included', we follow common practices[17,19,39,55] to use a local ancestry term ℓ_s (defined above) in equation (1):

$$y = \ell_s \beta_{s,\text{lanc}}^{(m)} + g_{s,1} \beta_{s,1}^{(m)} + g_{s,2} \beta_{s,2}^{(m)} + \mathbf{c}^\top \boldsymbol{\alpha} + \epsilon,$$

where $\beta_{s,\text{lanc}}^{(m)}$ denotes the effect of local ancestry.

- For 'lanc regressed', we use $y = \ell_s \beta_{s,\text{lanc}}^{(m)} + g_{s,1} \beta_{s,1}^{(m)} + g_{s,2} \beta_{s,2}^{(m)} + \epsilon$. We first estimate $\widehat{\beta_{s,\text{lanc}}^{(m)}}$ in the regression of $y \sim \ell_s \beta_{s,\text{lanc}}^{(m)}$, and then estimate $\beta_{s,1}^{(m)}, \beta_{s,2}^{(m)}$ in regression of $\left(y - \ell_s \widehat{\beta_{s,\text{lanc}}^{(m)}}\right) \sim g_{s,1} \beta_{s,1}^{(m)} + g_{s,2} \beta_{s,2}^{(m)}$.

To assess the impact of including local ancestry term when applying HET test, we randomly selected 1,000 SNPs on chromosome 1 from PAGE genotype. We simulated traits with single causal SNP. For each SNP, we simulated quantitative trait with the given single causal SNP with varying $\beta_{\text{eur}} : \beta_{\text{afr}} = 1.0, 1.05, 1.1, 1.15, 1.2$. We scaled $\beta_{\text{eur}}, \beta_{\text{afr}}$ such that the causal SNP explained the given amount of h_g^2. For each SNP, simulations of $\beta_{\text{eur}}, \beta_{\text{afr}}$ and environmental noises were repeated 30 times. We then applied different strategies of including local ancestry to these simulations and obtained p-value of HET testing $H_0 : \beta_{\text{eur}} = \beta_{\text{afr}}$. We additionally included the top principal component as a covariate throughout. We evaluated the distribution of FPR or power of HET test by subsampling without replacement: we drew 100 random samples, each sample consisted of 500 SNPs, randomly drawn from the pool of 1,000 SNPs and 30 simulations; such sampling accounts for the randomness from both the environmental noises and SNP MAF. We calculated FPR or power for each sample of 500 SNPs, obtained empirical distributions of FPR or power (100 points each), and then calculated the mean and SE (using empirical standard deviation) from the empirical distribution.

Simulations with single causal variant.

We performed simulations with single causal variant to assess the properties of methods based on estimated marginal effects. We randomly selected 100 regions each spanning 20 Mb on chromosome 1 (approximately 120,000 SNPs per region on average, standard deviation 6,000). For each region, the causal variant located at the middle of the region; it had same causal effects across local ancestries and was expected to explain a fixed amount of heritability (0.2%, 0.6% and 1.0%); the sign of the causal effect and environmental noises were randomly drawn 100 times. We evaluated four metrics at both causal variants and clumped variants; clumped variants were obtained with regular LD clumping (index $P < 5 \times 10^{-8}$, $r^2 = 0.1$, window size 10 Mb) using PLINK (v1.90b6.24): `--clump--clump-p1 5e-8 --clump-p2 1e-4 --clump-r2 0.1 --clump-kb 10000`. We used a 10 Mb clumping window to account for the larger LD window within admixed individuals; other parameters

were adopted from ref. [56]. We found that, when the simulated h_g^2 was large, LD clumping can result in multiple SNPs because the secondary SNPs can reach $P < 5 \times 10^{-8}$ when we applied a commonly-used $r^2 = 0.1$ threshold. Therefore, for each region, we either retained only the SNP with strongest association (matching the simulation setup of a single simulated causal variant), or retained all the SNPs from clumping results. Similar as above, we evaluated the distribution of four metrics by subsampling without replacement: we drew 100 random samples, each sample consisted of 500 regions (each region has one causal SNP), randomly drawn from the pool of 100 regions and 100 simulations; such sampling accounted for the randomness from both the environmental noises and SNP MAF. We then calculated the mean and SE from the 100 random samples.

Simulation with multiple causal variants.

We performed simulations with multiple causal variants. We simulated multiple causal variants randomly distributed on chromosome 1 (515,087 SNPs). We drew $n_{\text{causal}} = 62, 125, 250, 500$ and $1,000$ causal variants to simulate different levels of polygenicity, such that on average there were approximately $0.25, 0.5, 1.0, 2.0$ and 4.0 causal variants per Mb. We fixed the heritability explained by all variants on chromosome 1 as $h_g^2 = 2.5\%, 5\%, 10\%$ and 20%. We performed subsampling without replacement to estimate the average and standard errors of four metrics (each sample consisted of 1,000 SNPs, randomly drawn from SNPs across 500 simulations). We found that when the simulated h_g^2 was small ($h_g^2 = 2.5\%, 5\%$), because of the limited sample size in our data ($n = 17,299$) for PAGE data, very few SNPs reach $P < 5 \times 10^{-8}$ in these simulations and consequently standard errors are very large and results cannot be reliably reported. Therefore, we chose to report results only from $h_g^2 = 10\%$ and 20% in Supplementary Table 14.

Genome-wide simulation for evaluating our polygenic method.

We performed simulations to evaluate our polygenic method in terms of parameter estimation of r_{admix} and hypothesis testing $H_0 : r_{\text{admix}} = 1$ using real genome-wide genotypes. We simulated quantitative phenotypes using genotypes and inferred local ancestries from PAGE dataset. The phenotypes were simulated under a wide range of genetic architectures varying proportion of causal variants P_{causal}, heritability h_g^2 and true correlation r_{admix}, and a frequency-dependent effects distribution for causal variants: in each simulation, we randomly drew P_{causal} proportion of causal variants. Given the set of causal variants, we simulated quantitative phenotypes on the basis of equations (3) and (4). The environmental noises were then simulated according to the desired heritability h_g^2.

Real data analysis

Phenotype processing.

For PAGE, we analyzed 24 heritable traits in PAGE based on ref. [1]. For UKBB, we analyzed 26 heritable traits based on heritability and number of individuals with non-missing phenotype values, following ref. [57]. For AoU, we analyzed ten heritable traits, including physical measurement and lipid phenotypes, which were straightforward to phenotype and have large sample sizes. Physical measurement phenotypes were extracted from Participant Provided Information in AoU dataset. Lipid phenotypes (including LDL, HDL, TC and TG) were extracted following https://github.com/all-of-us/ukb-cross-analysis-demo-project/tree/main/aou_workbench_siloed_analyses, including extracting most recent measurements per person, and correcting value with statin usage. These traits included both quantitative and binary traits and it was previously shown that genetic correlation methodology can be directly applied to binary traits[58]. For each trait, we quantile normalized phenotype values. We included age, sex, age*sex and top ten in-sample principal components (and 'study center' for PAGE) as covariates. We quantile normalized each covariate and used the average of each covariate to imputed missing values in covariates.

Genome-wide genetic correlation estimation. We calculated $\mathbf{K}_1$, $\mathbf{K}_2$ matrices in equation (5) using either imputed SNPs and HapMap3 SNPs (for PAGE and UKBB), or microarray SNPs (for AoU). We used either frequency-dependent or GCTA heritability models via specifying τ_s^2. $\mathbf{K}_1$, $\mathbf{K}_2$ matrices were separately calculated for individuals within PAGE, UKBB and AoU studies. For each given r_{admix}, we used GCTA[27] (v1.94.0beta) to fit a single variance component model with the calculated $\mathbf{K}_1 + r_{admix}\mathbf{K}_2$ using `gcta64 --reml --reml-no-constrain`. We additionally included the causal signals at Duffy SNP (rs2814778) in 1q23.2 as covariates for analysis of white blood cell count and C-reactive protein because of the known strong admixture peak[59,60]. Specifically, we used the local ancestries of SNP closest to Duffy SNP in our data as proxies for Duffy SNP (Duffy SNP itself is not typed or imputed in our data). The local ancestries are valid proxies of Duffy SNP because Duffy SNP is known to be highly differentiated across ancestries (alternate allele frequency is 0.006 versus 0.964 in ref. [53]) and therefore local ancestries are highly correlated with the Duffy SNP. We excluded closely related individuals in the analysis (<3rd-degree relatives; using ref. [61] with `plink2 --king-cutoff 0.0884`). We note that our meta-analysis credible interval across traits can be anti-conservative (that is, the actual coverage probability is less than the nominal coverage probability) because we did not account for the genetic correlation across traits.

Individual trait–SNP analysis. We evaluated effects consistency at individual SNPs that were significantly associated with each trait. First, we performed GWAS and LD clumping with the same parameters described above. Even though LD clumping was performed using stringent parameters, we found cluster of clumped SNPs that were probably dependent with each other as a combined result of multiple causal variants within a region the long-range LD in admixed populations (Supplementary Table 11 and Discussion). For each clumped trait–SNP pair, we estimated ancestry-specific effects and standard errors.

Statistical fine-mapping analysis. We performed fine-mapping analysis to each trait–SNP pair with significant heterogeneity by ancestry using SuSiE[62] (v0.12) (for PAGE and UKBB, for which we used genotype data with high SNP density). For each trait–SNP, we included all imputed SNPs in a 3 Mb window. We ran SuSiE with individual-level genotype and phenotype (covariates were regressed out of genotype and phenotype), using default settings with maximum number of ten non-zero effects. We obtained posterior inclusion probability and credible sets.

Statistics and reproducibility
We analyzed three publicly available datasets of PAGE, UKBB and AoU, and sample sizes were determined in these studies. We did not use randomization or blinding. We focused on analyzing individuals with admixed African-European ancestries, and individuals with other genetic ancestries were not included in analyses of this work. We replicate our findings across these three independent datasets.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

PAGE individual-level genotype and phenotype data are available through dbGaP https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000356.v2.p1. UKBB individual-level genotype and phenotype data are available through application at https://www.ukbiobank.ac.uk/. AoU individual-level genotype and phenotype are available through application at https://www.researchallofus.org/. The set of preprocessed HapMap3 variants used in this manuscript is retrieved from https://ndownloader.figshare.com/files/25503788.

Code availability

Software implementing genome-wide genetic correlation estimation method: https://github.com/kangchenghou/admix-kit (ref. https://doi.org/10.5281/ZENODO.7482679) Code for replicating analyses: https://github.com/kangchenghou/admix-genet-cor (ref. https://doi.org/10.5281/ZENODO.7482683).

References

43. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
44. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **91**, 1011–1021 (2012).
45. Gazal, S. et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).
46. Gazal, S., Marquez-Luna, C., Finucane, H. K. & Price, A. L. Reconciling S-LDSC and LDAK functional enrichment estimates. *Nat. Genet.* **51**, 1202–1204 (2019).
47. Hou, K. et al. Accurate estimation of SNP-heritability from biobank-scale data irrespective of genetic architecture. *Nat. Genet.* **51**, 1244–1251 (2019).
48. Linnet, K. Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clin. Chem.* **44**, 1024–1031 (1998).
49. Chiu, A. M., Molloy, E. K., Tan, Z., Talwalkar, A. & Sankararaman, S. Inferring population structure in biobank-scale genomic data. *Am. J. Hum. Genet.* **109**, 727–737 (2022).
50. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
51. Loh, P.-R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
52. Maples, B. K., Gravel, S., Kenny, E. E. & Bustamante, C. D. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.* **93**, 278–288 (2013).
53. The 1000 Genomes Project Consortium. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
54. Schubert, R., Andaleon, A. & Wheeler, H. E. Comparing local ancestry inference models in populations of two- and three-way admixture. *PeerJ* **8**, e10090 (2020).
55. Gay, N. R. et al. Impact of admixture and ancestry on eQTL analysis and GWAS colocalization in GTEx. *Genome Biol.* **21**, 233 (2020).
56. Pardiñas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
57. Schoech, A. P. et al. Negative short-range genomic autocorrelation of causal effects on human complex traits. Preprint at *medRxiv* https://doi.org/10.1101/2020.09.23.310748 (2020).
58. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
59. Reich, D. et al. Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene. *PLoS Genet.* **5**, e1000360 (2009).
60. Reiner, A. P. et al. Genome-wide association and population genetic analysis of C-reactive protein in African American and Hispanic American women. *Am. J. Hum. Genet.* **91**, 502–512 (2012).
61. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
62. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J. R. Stat. Soc. Ser. B* **82**, 1273–1300 (2020).
63. Cook, J. P. & Morris, A. P. Multi-ethnic genome-wide association study identifies novel locus for type 2 diabetes susceptibility. *Eur. J. Hum. Genet.* **24**, 1175–1180 (2016).

Author contributions

K.H. and B.P. conceived and designed the experiments. K.H. performed the experiments and statistical analyses with assistance from Y.D., Z.X., Y.W., A.B., R.M., S.S. and B.P. G.M.B., S.B., D.V.C., B.F.D., M.F., C.G., X.G., C.H., E.E.K., M.K., C.K., L.L., A.M., K.E.N., U.P., L.J.R.-T., S.S.R., J.I.R., H.E.W., G.L.W. and Y.Z. provided data and feedback on analysis. K.H. and B.P. wrote the manuscript with feedback from all authors.

Competing interests

E.E.K. has received personal fees from Regeneron Pharmaceuticals, 23&Me and Illumina, and serves on the advisory boards for Encompass Biosciences, Foresite Labs and Galateo Bio. The remaining authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-023-01338-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-023-01338-6.

Correspondence and requests for materials should be addressed to Kangcheng Hou or Bogdan Pasaniuc.

Peer review information *Nature Genetics* thanks Loïc Yengo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.
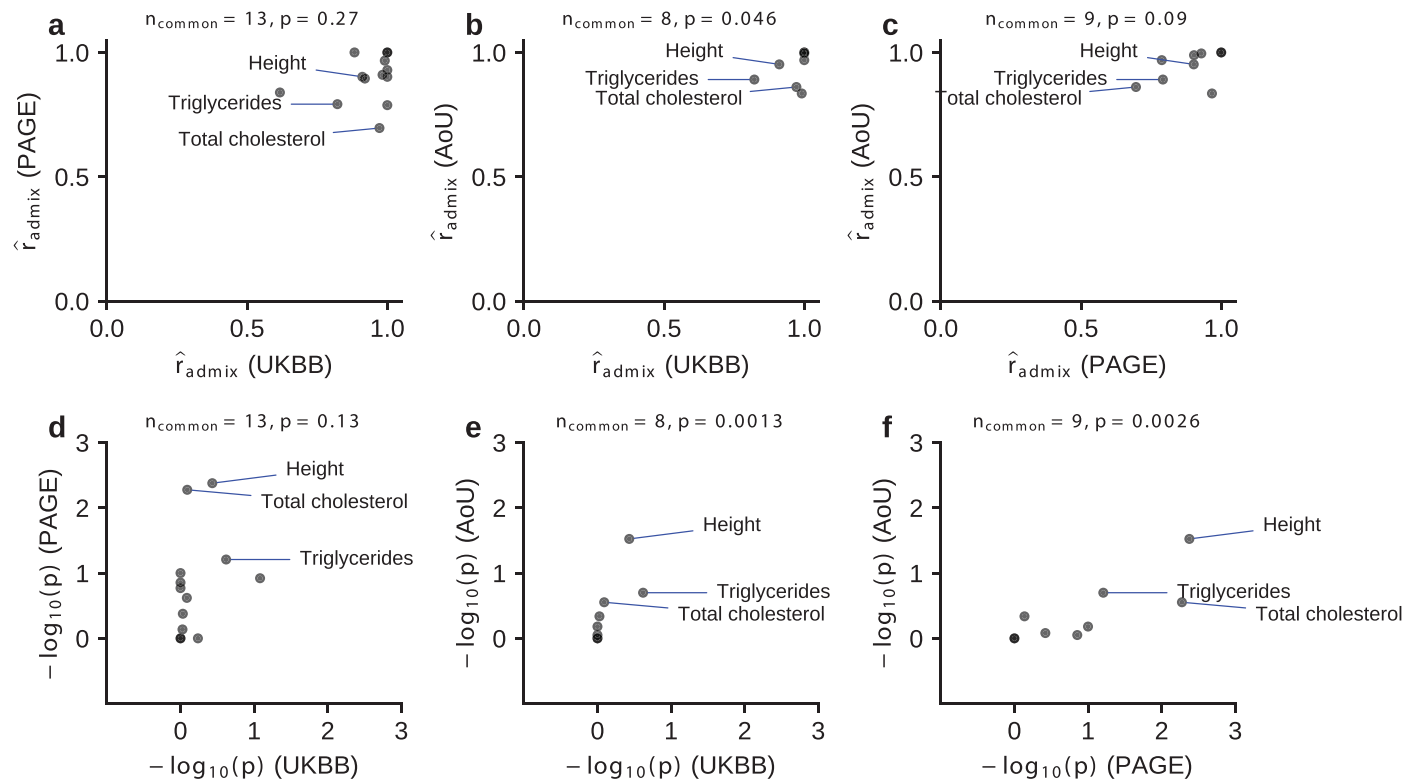
Extended Data Fig. 1 | Consistency of r_{admix} for shared traits across studies. We compared estimated r_{admix} for shared traits across studies. We compared both $\hat{r}_{admix}$ **(a-c)** and $-\log_{10}(p)$ (for one-sided test of $H_0 : r_{admix} = 1$; Methods) **(d-f)**. Three traits (Height, Triglycerides, Total cholesterol) with the most significant p-values for $H_0 : r_{admix} = 1$ were annotated. Number of common traits shared across studies (n_{common}) and Spearman correlation p-value were shown in the title for each panel. Overall, there were weak consistency of estimated $\hat{r}_{admix}$ for shared traits across studies (although p-values for $H_0 : r_{admix} = 1$ were consistent significantly). Numerical results are reported in Supplementary Table 7.
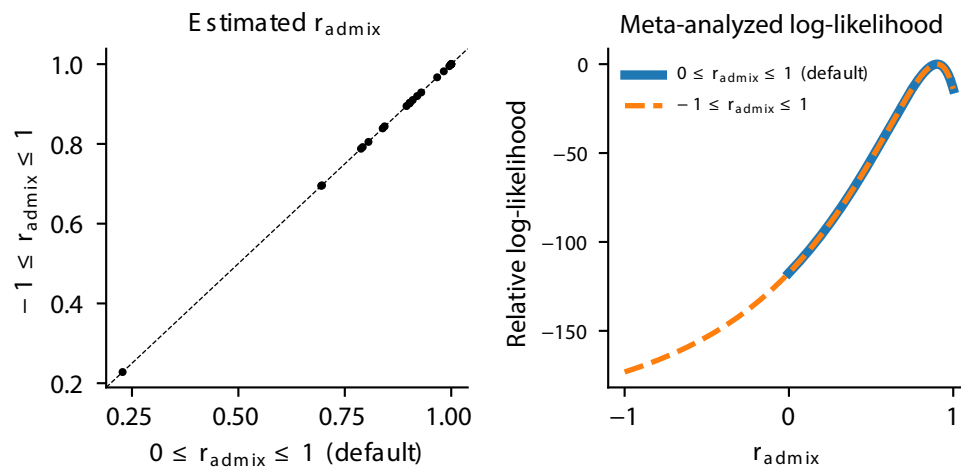
Extended Data Fig. 2 | r_{admix} estimation is robust to the assumption of $r_{admix} > 0$.
We performed r_{admix} estimation using alternative assumption of $-1 \le r_{admix} \le 1$ in real trait analysis in PAGE in light of potential scenarios of effect sizes in opposite directions[36,63]. We compared estimated r_{admix} when assuming $0 \le r_{admix} \le 1$ (default method) and when assuming $-1 \le r_{admix} \le 1$. Left: comparing point estimates of r_{admix} across 24 traits in PAGE. Right: comparing the meta-analyzed log-likelihood. Results obtained from two methods are highly consistent.
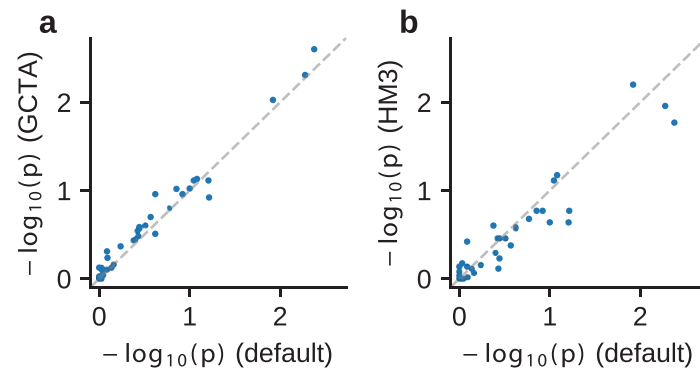
Extended Data Fig. 3 | r_{admix} estimation is robust to genetic architecture and SNP set. We performed r_{admix} estimation under the assumption of alternative genetic architecture and SNP set on real trait analysis across PAGE and UKBB. We compared p-values (for one-sided test of $H_0 : r_{admix} = 1$) of our default setting (using frequency-dependent genetic architecture and imputed SNPs; Table 1) to those obtained using GCTA genetic architecture and imputed SNPs **(a)**, and to those obtained using frequency-dependent genetic architecture and HM3 SNPs **(b)**. Numerical results are reported in Supplementary Table 8.
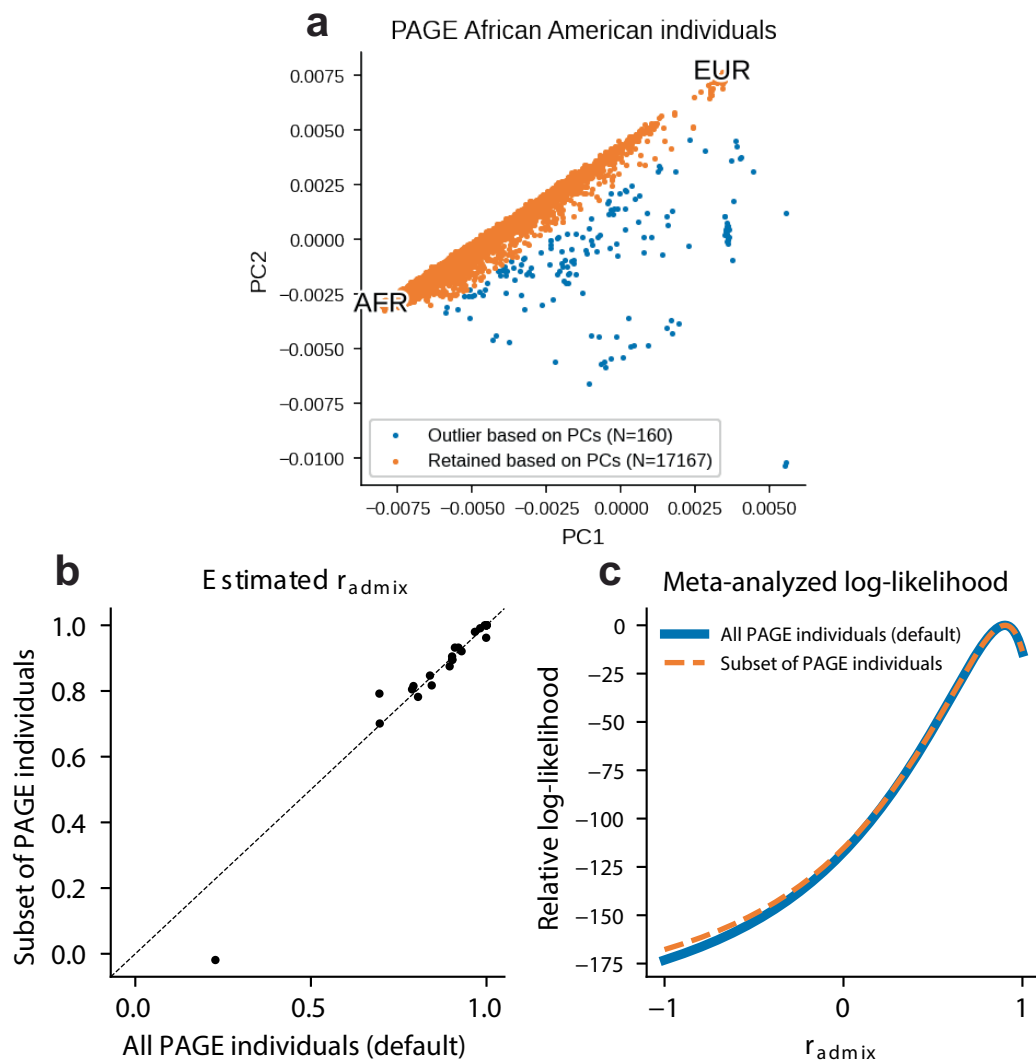
Extended Data Fig. 4 | r_{admix} estimation is robust to subsetting PAGE African American individuals based on genotype PCs. We subsetted PAGE individuals with self-identified race/ethnicity label of 'African American' (total $N = 17,327$) based on genotype PCs and retained $N = 17,167$ individuals **(a)**. We found that the estimated r_{admix} were highly consistent between using all PAGE African American individuals (default) and using subset of PAGE African American individuals based on genotype PCs. **(b)** comparing point estimates of r_{admix} across 24 traits in PAGE. (Dot on the bottom left of the figure corresponds to MCHC trait, with a small sample size of 3,650.) **(c)** comparing the meta-analyzed log-likelihood. Results obtained from two sets of individuals are highly consistent.
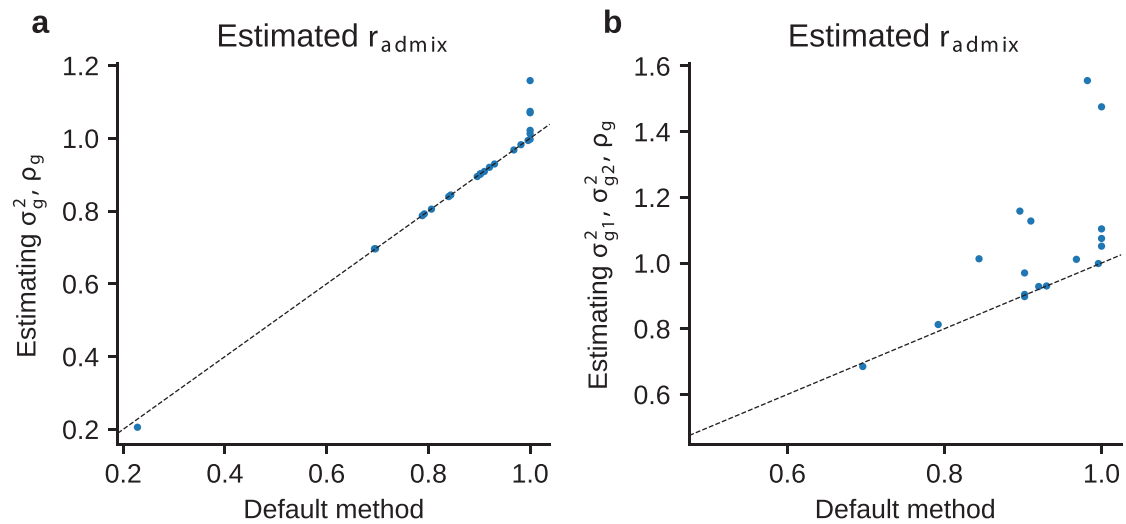
Extended Data Fig. 5 | Comparing estimated r_{admix} between alternative method formulations and default method. Each dot corresponds to a trait. **(a)** Comparing results of default method and of directly optimizing and estimating σ_g^2, ρ_g. **(b)** Comparing results of default method and of directly optimizing and estimating $\sigma_{g,1}^2, \sigma_{g,2}^2$ (different variance components per ancestry) and ρ_g. See Supplementary Table 9 and Supplementary Note for details.
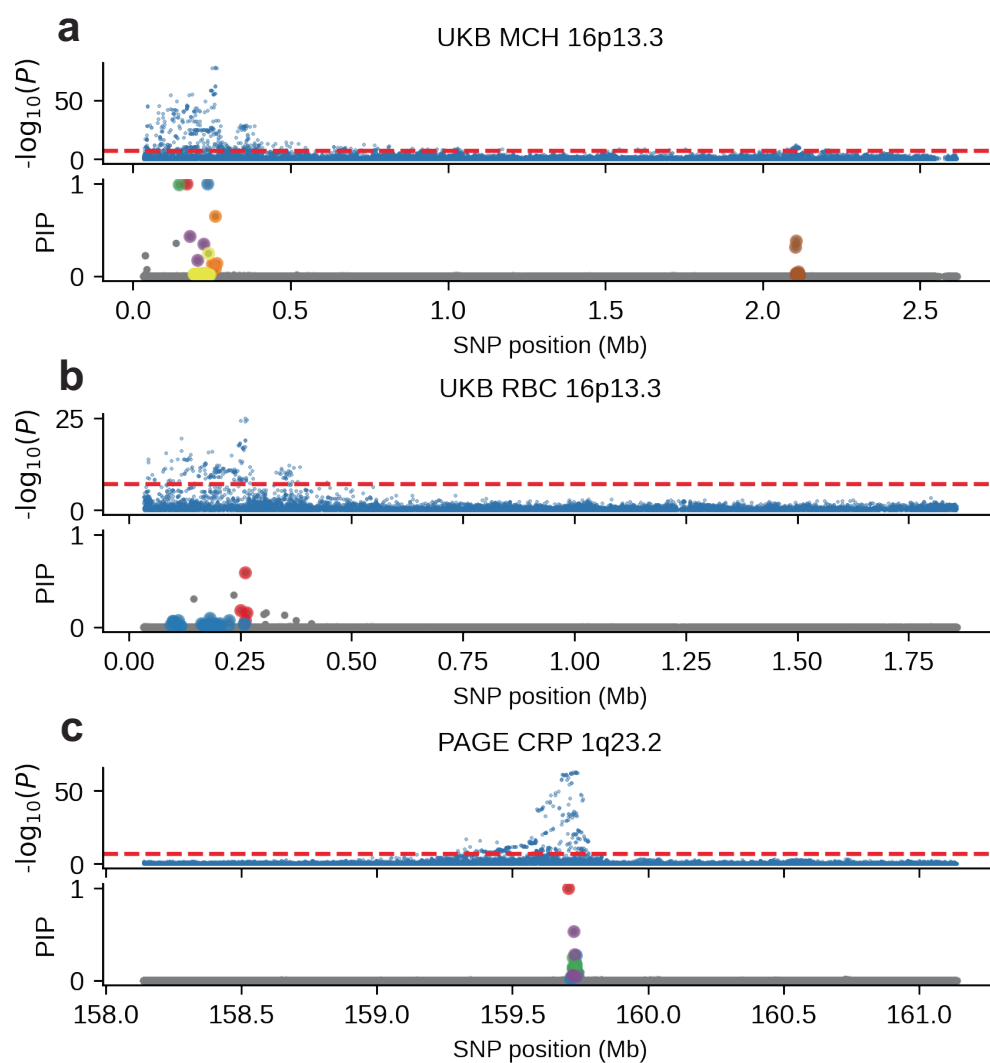
Extended Data Fig. 6 | Multiple conditionally independent association signals for loci with heterogeneity by ancestry. Upper panel corresponds to the two-sided association *p*-values and lower panel corresponds to the fine-mapping PIP. Different colors in the PIP plot corresponds to different credible sets. **(a)** MCH at 16p13.3 for UK Biobank European-African admixed individuals. **(b)** RBC at 16p13.3 for UK Biobank European-African admixed individuals. **(c)** CRP at 1q23.2 for PAGE European-African admixed individuals.
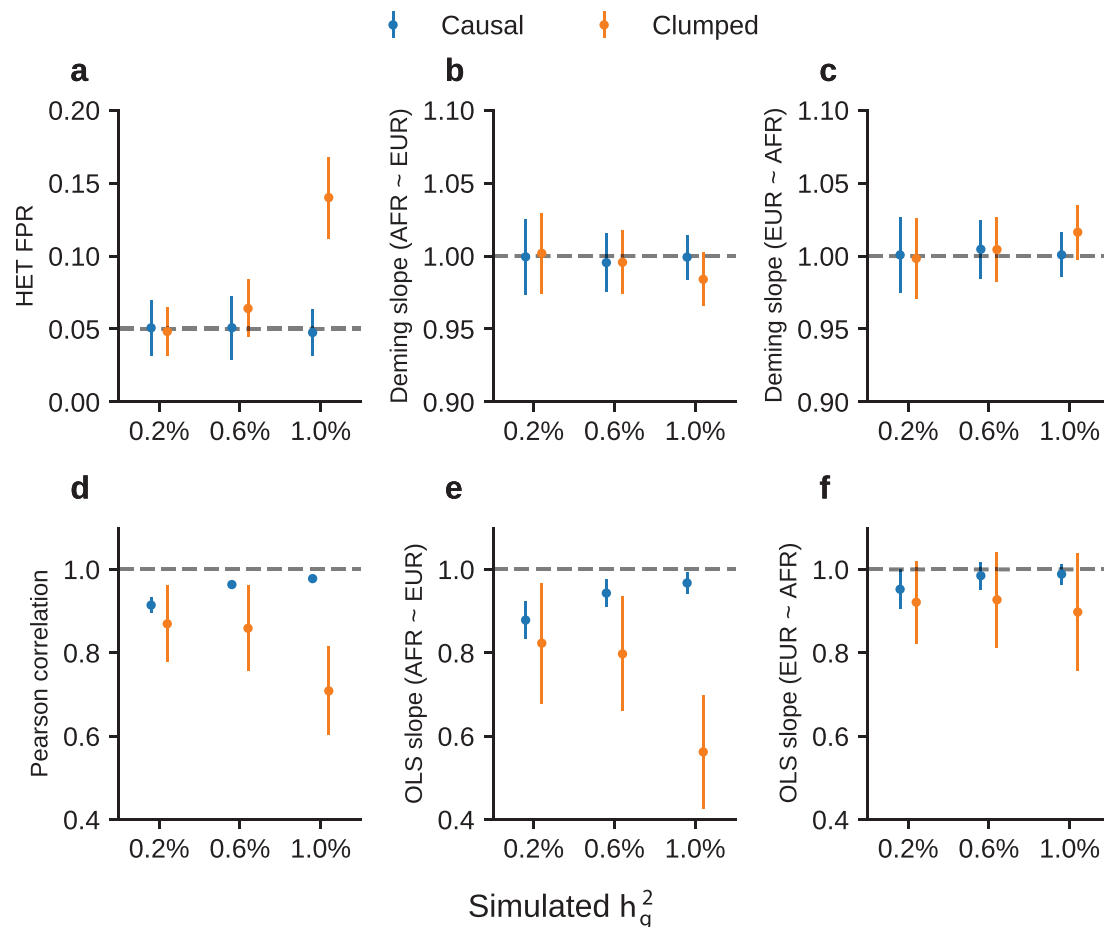
Extended Data Fig. 7 | Simulations with single causal variant. Simulations were based on 100 regions each spanning 20 Mb on chromosome 1 and 17,299 PAGE individuals. In each simulation, we randomly selected single causal variant and simulated quantitative phenotypes where these causal variants had same causal effects across ancestries and each causal variant was expected to explain a fixed amount of heritability (0.2%, 0.6%, 1.0%). Each panel corresponds to one metric for both causal and clumped variants. **(a)** False positive rate (FPR) of HET test. **(b)** Deming regression slope with $\beta_{afr} \sim \beta_{eur}$. **(c)** Deming regression slope with $\beta_{eur} \sim \beta_{afr}$. **(d)** Pearson correlation. **(e)** OLS regression slope with $\beta_{afr} \sim \beta_{eur}$. **(f)** OLS regression slope with $\beta_{eur} \sim \beta_{afr}$. 95% confidence intervals were based on 100 random sub-samplings with each sample consisted of 500 SNPs (Methods). Numerical results are reported in Supplementary Table 13.
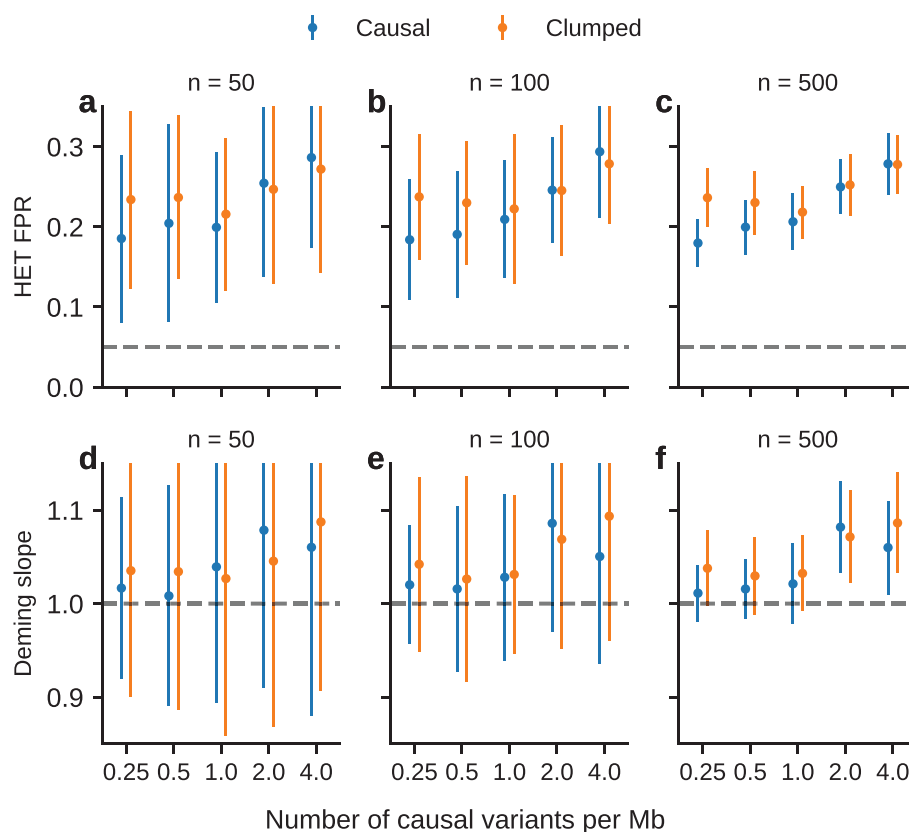
Extended Data Fig. 8 | Simulation with multiple causal variants at other sample sizes (Fig. 6d–f). Simulations were based on chromosome 1 (515,087 SNPs) and 17,299 PAGE individuals. We drew 62, 125, 250, 500, 1000 causal variants to simulate different level of polygenicity, such that on average there were approximately 0.25, 0.5, 1.0, 2.0, 4.0 causal variants per Mb. The heritability explained by all causal variants was fixed at $h_g^2 = 10\%$. **(a-c)** False positive rate of HET test for the causal variants and clumped variants. **(d-f)** Deming regression slope of estimated ancestry-specific effects (β_{eur} - β_{af}) for the causal variants and clumped variants. 95% confidence intervals were based on 100 random sub-samplings with each sub-sample consisted of $n = 50, 100, 500$ SNPs (instead of $n = 1,000$ SNPs in Fig. 6c, d) (Methods).

(i) Results for causal variants
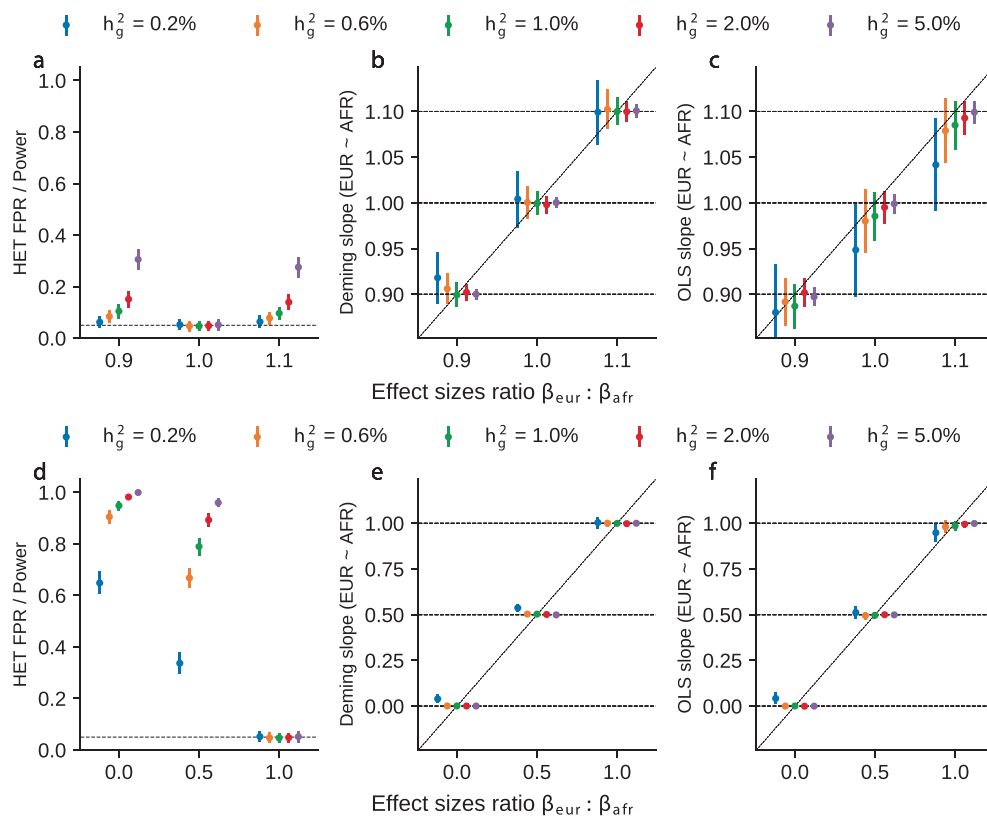
(ii) Results for clumped variants
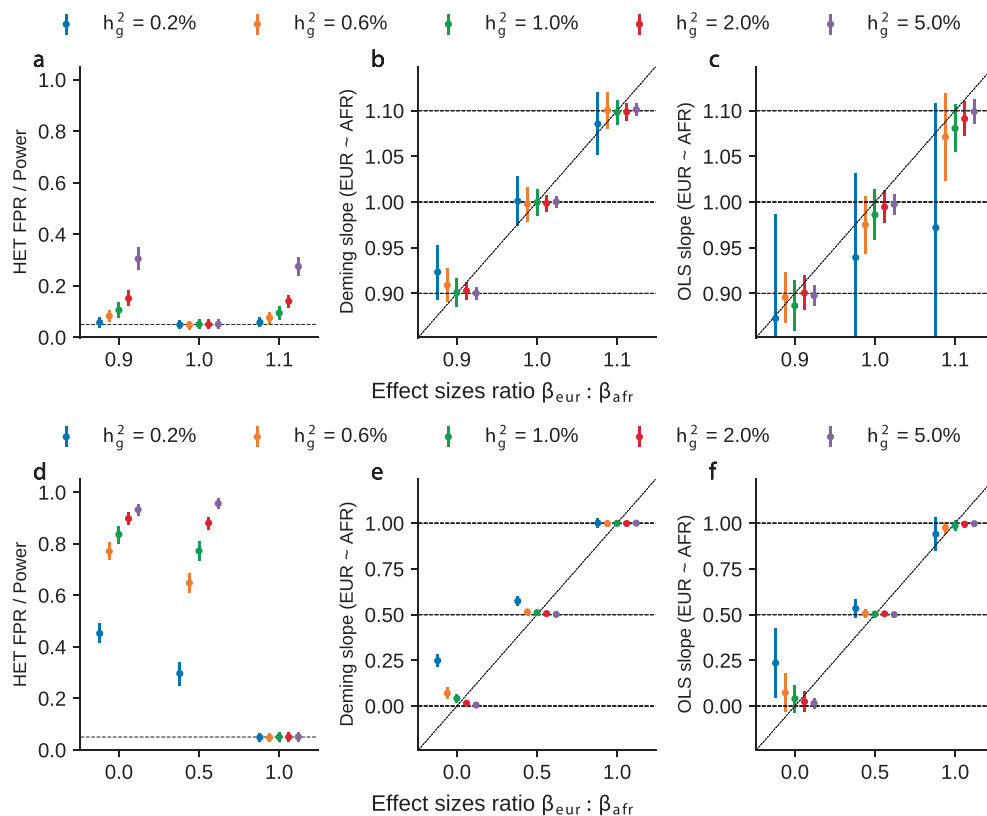
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Additional results for simulations with single causal variant with varying β_{eur}:β_{afr}, and h_g^2. Simulations were based on 100 regions each spanning 20 Mb on chromosome 1 from 17299 PAGE individuals. In each simulation, we randomly selected single causal variant and simulated quantitative phenotypes where these causal variants had varying causal effects across ancestries and each causal variant was expected to explain a fixed amount of heritability (0.2%, 0.6%, 1.0%, 2.0%, 5.0%). We provide results for both causal variants and LD-clumped variants. We separate results into two rows for better visualization: upper row (**a**-c): β_{eur}:β_{afr} = 0.9, 1.0, 1.1; lower row (**d-f**): β_{eur}:β_{afr} = 0.0, 0.5, 1.0. We show results for False positive rate (FPR) of HET test, Deming regression slope with $\beta_{eur} \sim \beta_{afr}$, and OLS regression slope with $\beta_{eur} \sim \beta_{afr}$. 95% confidence intervals were based on 100 random sub-samplings with each sample consisted of 500 SNPs (Methods). Numerical results and further discussions are provided in Supplementary Table 15.

Corresponding author(s): Kangcheng Hou

Last updated by author(s): Jan 1, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Software implementing genome-wide genetic correlation estimation method: https://github.com/kangchenghou/admix-kit. All code (including customized python and R scripts) for replicating analyses will be available: https://github.com/kangchenghou/admix-genet-cor. Other software used in the manuscript: GCTA (v1.94.0beta), pandas (v1.3.3), PLINK (v1.90b6.24), PLINK2 (v2.0a3), Eagle2 (v2.4.1), RFMix (v2), SuSiE (v0.12), scipy (v1.9.2), numpy (v1.23.5), statsmodels (v0.13.2), matplotlib (v3.4.2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

PAGE individual-level genotype and phenotype data are available through dbGaP https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	We used biological sex as covariates in statistical modeling to model possible differences of phenotype mean across sex (for example, height). We used the sex attribute provided by UK Biobank (Bycroft et al. Nature 2018), PAGE (Wojcik et al. Nature 2019), and used "Sex at Birth" provided in All of Us dataset (https://aousupporthelp.zendesk.com/hc/en-us/articles/360039728751-Sex-gender-and-sexual-orientation-generalizations)
Population characteristics	Population characteristics can be found in UK Biobank, PAGE, All of Us studies.
Recruitment	Participants were not recruited in this study. Details can be found in UK Biobank, PAGE, All of Us studies.
Ethics oversight	No new data is collected in this work. Ethics committee/IRB of Population Architecture using Genomics and Epidemiology (PAGE) gave ethical approval for collection of PAGE data (described in Wojcik et al. 2019). PAGE data are available through dbGaP (phs000356.v2.p1). Ethics committee/IRB of UK Biobank gave ethical approval for collection of UK Biobank data (https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics). Approval to use UK Biobank individual-level in this work was obtained under application 33297 at http://www.ukbiobank.ac.uk. Ethics committee/IRB of All of Us gave ethical approval for collection of All of Us data (https://allofus.nih.gov/about/who-we-are/institutional-review-board-irb-of-all-of-us-research-program). Approval to use All of Us controlled tier data in this work was obtained through application at https://www.researchallofus.org.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We did not collect new data and sample size corresponds to the those from existing datasets. These three datasets are used because they are the largest publicly available datasets for genotypes and phenotypes from African-European admixed individuals. We determine these datasets provide sufficient power for our analysis through simulation studies.
Data exclusions	We focused on analyzing individuals with admixed African-European ancestries and individuals with other genetic ancestries were not included in analyses of this work.
Replication	We replicate our findings across three large-scale datasets: PAGE, UK Biobank, All of Us.
Randomization	Randomization is not relevant for this study as we did not perform randomization for individuals.
Blinding	The investigators are blinded to group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging