Low generalizability of polygenic scores in African populations due to genetic and environmental diversity

Lerato Majara[1,2], Allan Kalungi[1,3,4,5], Nastassja Koen[1,6,7], Heather Zar[8], Dan J. Stein[6,7], Eugene Kinyanda[5], Elizabeth G. Atkinson[9,10,11], Alicia R. Martin[9,10,11]

[1] Global Initiative for Neuropsychiatric Genetics Education in Research (GINGER), Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA, USA
[2] MRC Human Genetics Research Unit, Division of Human Genetics, Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Observatory 7925, South Africa
[3] Department of Psychiatry, College of Health Sciences, Makerere University, Kampala, Uganda.
[4] Department of Psychiatry, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.
[5] Mental Health Project, Medical Research Council/Uganda Virus Research Institute (MRC/UVRI) & London School of Hygiene and Tropical Medicine (LSHTM), Uganda Research Unit, Entebbe, Uganda
[6] Department of Psychiatry and Neuroscience Institute, University of Cape Town, Cape Town, South Africa
[7] South African Medical Research Council (SAMRC) Unit on Risk and Resilience in Mental Disorders, Cape Town, South Africa
[8] Department of Paediatrics and Child Health, Red Cross Children's Hospital and Medical Research Council Unit on Child and Adolescent Health, University of Cape Town, Cape Town, South Africa.
[9] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA
[10] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[11] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

Correspondence: armartin@broadinstitute.org

Key words: polygenic scores, Africa, GWAS, health disparities, global health, population genetics

Abstract

African populations are vastly underrepresented in genetic studies but have the most genetic variation and face wide-ranging environmental exposures globally. Because systematic evaluations of genetic prediction had not yet been conducted in ancestries that span African diversity, we calculated polygenic risk scores (PRS) in simulations across Africa and in empirical data from South Africa, Uganda, and the UK to better understand the generalizability of genetic studies. PRS accuracy improves with ancestry-matched discovery cohorts more than from ancestry-mismatched studies. Within ancestrally and ethnically diverse South Africans, we find that PRS accuracy is low for all traits but varies across groups. Differences in African ancestries contribute more to variability in PRS accuracy than other large cohort differences considered between individuals in the UK versus Uganda. We computed PRS in African ancestry populations using existing European-only versus ancestrally diverse genetic studies; the increased diversity produced the largest accuracy gains for hemoglobin concentration and white blood cell count, reflecting large-effect ancestry-enriched variants in genes known to influence sickle cell anemia and the allergic response, respectively. Differences in PRS accuracy across

African ancestries originating from diverse regions are as large as across out-of-Africa continental ancestries, requiring commensurate nuance.

Introduction

Genome-wide association studies (GWAS) have yielded important biological insights into the heritable basis of many complex traits and diseases (Visscher et al., 2017). However, the vast majority of studies have been conducted in populations of European descent, raising questions about their utility across diverse populations (Manrai et al., 2016; Martin et al., 2019; Morales et al., 2018; Popejoy and Fullerton, 2016; Sirugo et al., 2019). Previous studies have evaluated the generalizability of GWAS by using polygenic risk scores (PRS) to compare the association between genetically predicted versus measured phenotypes in diverse populations. These studies have found that PRS accuracy decreases with increasing genetic distance between the GWAS discovery and PRS target cohorts (Martin et al., 2017, 2019; Scutari et al., 2016). Since the earliest applications of PRS in human genetics, these concepts--coupled with Eurocentric study biases--have resulted in PRS that are most accurate in European ancestry populations and least accurate in African ancestry populations (International Schizophrenia Consortium et al., 2009). These study biases and phenomena continue to replicate a decade later, with several-fold differences in prediction accuracy of many traits between European and non-European ancestry populations (Martin et al., 2019).

Quantifying PRS generalizability within and among African populations requires considerable nuance as they represent the most genetically diverse populations globally, with more than a million more genetic variants per person than out-of-Africa populations (1000 Genomes Project Consortium et al., 2015). Populations collected even within the same geographic regions of Africa have complex demographic histories with complicated patterns of admixture and population structure (Busby et al., 2016; Choudhury et al., 2020; Pagani et al., 2015; Uren et al., 2016). Further, African ancestry populations experience vastly different environments within versus outside continental Africa as well as more locally among diverse communities, countries, and regions of Africa. These differences provide unique epidemiological opportunities to query the impacts of vastly differing environments on PRS accuracy. Previous empirical analyses and theoretical work fundamentally informs how demographic history and environmental variation interplay to produce PRS heterogeneity in traditionally underserved populations (Mostafavi et al., 2020; de Vlaming et al., 2017; Wang et al., 2020; Wray et al., 2013; Zaidi and Mathieson, 2020).

The inclusion of African ancestry participants in large-scale genetic studies is uniquely important for many reasons. They have the lowest life expectancies globally (Hero et al., 2017; Roser, 2013), receive the lowest access to and quality of medical care in the US (of Health et al., 2017), and are the most underserved by genetic technologies (Martin et al., 2018; Sirugo et al., 2019). A more nuanced understanding of PRS transferability will critically inform which populations are currently the most underserved and thus where building genetic studies and resources will have the biggest benefits globally.

There are also clear benefits to including African populations in statistical genetics efforts. Because humans originated in Africa, populations from Africa have the most genetic diversity among global populations (1000 Genomes Project Consortium et al., 2015; Campbell and Tishkoff, 2008; Henn et al., 2012a), such that more genotype-phenotype associations are expected in Africa than can be found elsewhere. African Americans have been shown to contribute disproportionately to GWAS findings (Morales et al., 2018), making up 2.8% of GWAS participants but contributing 7% of trait associations. African ancestry populations also have shorter blocks of linkage disequilibrium, which improves resolution to fine-map causal variants (Genovese et al., 2010). PRS accuracy is lowest in African ancestry populations due to GWAS study biases (Martin et al., 2019), but when GWAS include these and other diverse populations, PRS predict traits such as schizophrenia more accurately across all populations compared to single-ancestry GWAS (Bigdeli et al., 2019).

In this study, we have investigated how PRS generalize within and among diverse African populations in simulations and with empirical genotype-phenotype data for dozens of quantitative traits. We first simulated causal effects and computed genetic risk prediction accuracy using data from the African Genome Variation Project. We then calculated PRS using publicly available GWAS summary statistics from predominantly European ancestry populations to: 1) quantify PRS accuracy for 5 physical and psychosocial traits among populations in the Drakenstein Child Health Study (DCHS) of South Africa, a birth cohort study; and 2) compare PRS accuracy for 34 quantitative traits across the Ugandan General Population Cohort (GPC) versus ancestrally diverse UK Biobank participants. Our results highlight the disproportionate benefits of genetic studies in diverse African populations to improve trait prediction. Further, while PRS hold promise as biomarkers in precision medicine, a critical prerequisite is equitable accuracy in diverse populations to avoid exacerbating existing health disparities.

Results

Our study uses both simulation-based and empirical approaches to evaluate the generalizability of PRS across diverse African ancestry populations. An overview of the study design is shown in **Figure 1**, abbreviations are in **Table S1**, and a summary of datasets used in this study are shown in **Table S2**.

Simulated generalizability within and across diverse African populations

We simulated several quantitative traits with varying numbers of causal variants (N = 5; 20; 100; 2,000; 10,000; and 50,000) and heritabilities (h^2 = 0.1, 0.2, 0.4, and 0.8), then conducted independent GWAS for each scenario in East and West African ancestry populations (**Methods, Figures S1-4**). We calculated the prediction accuracy for PRS derived from the GWAS summary statistics considering ten different p-value thresholds within and across independent target populations from East, West, and South Africa. In general, ancestry-matched results with the sparsest and most heritable genetic architectures produced the highest prediction accuracy. Prediction accuracy was highest with trait h^2 = 0.8 and fewer than 100 causal variants (**Figure 2A-C**), as indicated by the highest R^2 and the identification of genome-wide significant associations. Conversely, when the number of causal variants exceeded 100, prediction accuracy was negligible (**Figure S4**) because of the small discovery cohort sample sizes, as evidenced by no variants meeting genome-wide significance in these simulations.

Prediction accuracy was highest with 5 and 20 causal variants (**Figure 2C**). The within-ancestry prediction at p-value threshold < 5e-08 and five causal variants were: R^2 = 0.86, p = 1.74 X 10^{-74} for East discovery - East target scores; R^2 = 0.85, p = 9.9e-74 for West discovery - West target scores. We observed lower prediction accuracy with ancestry mismatched discovery versus target cohorts at five causal variants and p-value threshold = 1e-6 (R^2 = 0.66, p = 1.79e-42 for West discovery - West target scores, compared to R^2 = 0.53, p =1.29e-74 for East discovery - West target scores). The scores in the South target sample were comparable when using East- or West-derived summary statistics (R^2 = 0.86, p = 5.19e-84 for West-derived summary statistics, and R^2 = 0.86, p = 1.35e-83 for East-derived summary statistics).

PRS accuracies in South African populations

While our simulations have shown that PRS generalize poorly across Africa due to substantial genetic diversity and differences across the continent, there is also considerable genetic and environmental diversity within regions and countries. We quantified PRS accuracy for a range of measured phenotypes in mothers genotyped in the DCHS cohort in South Africa, including several sociodemographic, physical/biomedical, and psychosocial risk traits (**Table S3**). The DCHS cohort consists of participants with multiple ancestry groups that include an admixed population with ancestry from multiple continents as well as a population with almost exclusively African population. These ancestry groups correlate with self-reported “Mixed” and “Black/African” ethnicities, respectively (**Figure S5**). We computed PRS for maternal height, depression, psychological distress, alcohol consumption, and smoking in DCHS overall, by ethnic group, and by ancestry within the Mixed ethnic group (**Methods**).

Across all genetically predicted phenotypes, only height was significantly predicted (**Figure S6**). We predicted height more accurately in the Mixed versus Black/African ethnic groups (R^2 = 0.099, 95% bootstrapped CI = [0.012, 0.18], p =1.5e-7 versus R^2 = 0.021, 95% CI = [-0.031, 0.043], p = 5.27e-3, respectively). We also expect that PRS accuracy increases with decreasing African ancestry within the Mixed ethnic group as has been shown previously in admixed African populations (Bitarello and Mathieson, 2020); we find suggestive evidence consistent with this trend when partitioning the Mixed group into two bins along PC1 (R^2 = 0.091, 95% CI = [-0.04, 0.17], p = 6.4e-4 in lower half of PC1 with more African ancestry vs R^2 = 0.12, 95% CI = [-9.0e-4, 0.21], p=5.7e-5 with more out-of-Africa ancestry), although small sample sizes limit definitive comparisons (N = 137 in each PC1 bin). Our results are consistent with variable prediction accuracy among diverse African ancestry groups within South Africa and insignificant prediction in African populations for all but the most heritable and accurately predicted traits elsewhere.

Variable phenotypic and genetic similarities across the Uganda General Population Cohort (GPC) and UK Biobank

Lower phenotypic correlations in Uganda GPC suggest higher contributing environmental effects

We next investigated phenotypic similarities within and across the Uganda GPC and UK Biobank participants because these are two of the largest cohorts with dozens of traits measured in African ancestry individuals. We first considered overall cohort differences between these cohorts--the Uganda GPC enrolled participants using a house-to-house study design and generated genetic data on 5,000 adults from rural villages in southwestern Uganda (Asiki et al., 2013), while the UK Biobank enrolled 500,000 people aged between 40-69 years in 2006-2010 from across the country (**Methods** (Bycroft et al., 2018)). Previous studies have reported higher rates of infectious diseases (e.g. HIV, hepatitis B and C) in the Uganda GPC than would be expected in the UK Biobank (Asiki et al., 2013). There are many additional potential environmental explanations for mean shifts in phenotypes, such as dietary, food security, and age differences contributing to considerable BMI differences across cohorts (μ = 21.3 and σ = 3.8 in Uganda GPC versus μ = 27.4 and σ = 4.8 in UK Biobank, $p < 2.2e-16$). To quantify comparisons while controlling for demographic differences for each of the 34 quantitative traits measured in both cohorts, we first mean centered each phenotype and regressed out the effects of age and sex within each cohort. Next, we then compared the distributions and variances of each phenotype across cohorts via Kolmogorov-Smirnov and F-tests, respectively (**Table S4**). Given the large sample sizes, all K-S tests were significantly different, with several phenotypes showing distributional and variance differences of considerable magnitude (**Figure S7** and **Table S4**, e.g. Bilirubin, BASO, HbA1c, ALP, EOS, TG, and NEU).

We next analyzed how similar the relationships are between phenotypes across datasets. Similar trends emerge overall, with distances across variance-covariance matrices for these cohorts showing evidence of significant correlation (Mantel test Z-statistic = 0.73, p < 1e-4).The correlations among phenotypes are slightly higher overall in the Uganda GPC than in UK Biobank, both among related and unrelated individuals, as expected from a household versus volunteer-based design (**Figure 3B, Figure S8**). More specifically, we see consistent correlations among combinations of phenotypes including SBP and DBP; RBC, Hb, and HCT; Cholesterol and LDL; WC, BMI, WT, and HC; MCHC, MCH, and MCV; GGT, ALT, AST, and ALP; and MONO, NEU, and WBC with high overall correlations across these datasets for these traits (**Figure 3A-B,** see abbreviations in **Table S1**). Some pairs of traits, however, have significantly different correlations across datasets. The largest difference in phenotypic correlations across datasets is between ALP and WT (ρ = 0.11, p < 2.2e-16 in UK Biobank versus ρ = -0.36, p < 2.2e-16 in Uganda GPC).

Our next goal was to compare trait heritability estimates in the UK Biobank versus Uganda GPC data (**Methods**). However, the sample size and study design differences between these cohorts required the application of different methods that limit comparability. Specifically, the household design of Uganda GPC included smaller sample sizes with more relatives in which family-based heritability estimates are most appropriate, whereas the large sample size and volunteer design in UK Biobank makes SNP-based heritability estimates from unrelated individuals most appropriate. **Figure S9** compares heritability estimates across traits in the UK Biobank versus Uganda GPC using these approaches (Gurdasani et al., 2019). As expected from the differences in the methods, study designs, and sample sizes, we find higher but noisier estimates in Uganda GPC for most traits, consistent with expectation from family-based versus unrelated heritability estimates across these two studies.

African genetic risk predictions from European ancestry GWAS data are remarkably inaccurate

To understand baseline trans-ethnic PRS accuracy using a typical approach, we predicted 32 traits in the Uganda GPC using GWAS summary statistics from the UK Biobank European ancestry individuals. While several traits were significantly predicted across ancestries, prediction accuracy was low for most traits (**Figure S10**); the most accurate PRS was for MPV, (R^2 = 0.036, 95% CI = [0.0069, 0.063], p = 5.73e-7) while the average variance explained across all traits was less than 1% (mean R^2 = 0.007). To assess the relative effects of ancestry versus cohort differences on decreases in prediction accuracy across populations, we next withheld 10,000 European ancestry individuals from UK Biobank for use as a target cohort, reran all GWAS, then used individuals with diverse continental ancestries in the UK Biobank as target populations (EUR = Europeans withheld from the GWAS, AMR = admixed American, MID = Middle Eastern, CSA = Central/South Asian, EAS = East Asian, and AFR = African, **Figure S11**), subcontinental African ancestries in the UK Biobank (Ethiopian, Admixed, South, East, West African ancestries, **Figure S12**), as well as the Uganda GPC (**Figure 4A, Table S4**).

Among continental ancestries, we computed R^2 and 95% confidence intervals for each trait (**Figure S13**), then computed median relative accuracy (RA) compared to Europeans and median absolute deviation (MAD) across all traits. We predict these traits most accurately in EUR (RA = 1, MAD = 0), followed by AMR (RA = 0.784, MAD = 0.023), MID (RA = 0.643, MAD = 0.034), CSA (RA = 0.621, MAD = 0.031), EAS (RA = 0.477, MAD = 0.024), and AFR (RA = 0.219, MAD = 0.014) (**Figure 4A**). We next compared prediction accuracy within African ancestry populations. Because some PRS accuracy estimates were noisy due to small sample sizes in UK Biobank Africans (especially Ethiopian and South African ancestry individuals, **Table S4**), we restricted analyses to those traits predicted with a 95% confidence interval < 0.08. Among these traits, we predicted most accurately those with Ethiopian ancestry (RA = 0.511, MAD = 0.059), followed by recently admixed individuals with West African and European ancestry (RA = 0.276, MAD = 0.016), East African ancestry (RA = 0.193, MAD = 0.023), West African ancestry (RA = 0.150, MAD = 0.012), and South African ancestry (RA = 0.083, MAD = 0.014) (**Figure 4A**). These results track with genetic distance and population history; the highest prediction accuracy identified in Ethiopians is expected given closer genetic proximity to European populations relative to other Africans due to back-to-Africa migrations influencing population structure there (Henn et al., 2012b; Hodgson et al., 2014; Pagani et al., 2015). The lowest prediction accuracy is in populations with southern African ancestry, consistent also with higher genetic divergence from European populations and more genetic diversity overall (Busby et al., 2016; Choudhury et al., 2020; Henn et al., 2011).

Lower prediction accuracy across ancestries than across cohorts

To compare prediction accuracy among similar ancestry participants from different cohorts, we next computed PRS for 32 traits using GWAS summary statistics from UK Biobank Europeans in two target populations: UK Biobank participants with East African ancestry versus Uganda GPC. As expected, prediction accuracy in these populations is very low across all traits in both cohorts and only slightly higher in the UK East African ancestry individuals than in the Uganda GPC individuals (mean R^2 = 0.017, sd = 0.013 versus mean R^2 = 0.012, sd = 0.010, respectively, **Figure S14**). Across traits, the differences in PRS accuracy across cohorts but within the same ancestry are much smaller than the differences across ancestries but within the UK Biobank, indicating that ancestry has a larger impact on genetic risk prediction than cross-cohort differences analyzed here. Smaller effects on genetic prediction accuracy differences across cohorts may be attributable to environmental differences, such as higher rates of malnutrition and infectious diseases previously reported in Uganda and in the GPC (Asiki et al., 2013; Nalwanga et al., 2020).

Improved African genetic risk prediction accuracy with multi-ethnic GWAS summary statistics

We next maintained the target populations but varied the discovery cohort to determine how more diverse GWAS impacts PRS accuracy for these phenotypes in diverse populations. Specifically, we computed PRS accuracy in diverse target populations in the UK Biobank (**Table S5**) using one of two discovery cohorts: the UKB European-only cohort versus diverse discovery cohorts combined via meta-analysis (**Table S6**). Meta-analyzed GWAS summary statistics come from several cohorts, including the UK Biobank (UKB), Biobank Japan (BBJ) (Nagai et al., 2017), Population Architecture Using Genomics and Epidemiology (PAGE) Consortium (Wojcik et al., 2019), and Uganda Genome Resource (UGR) (Gurdasani et al., 2019). For each trait, discovery cohort, and target cohort combination, we normalized the PRS R^2 values from the p-value threshold that explained the maximum phenotypic variance with respect to the prediction accuracy in the European target cohort using UK Biobank summary statistics only, then computed relative accuracies as before.

We find that prediction accuracy improves the most across populations when using a discovery cohort consisting of GWAS summary statistics meta-analyzed across the UKB, BBJ, and PAGE cohorts (**Figure 4B**), but not the UGR data (**Figure S15**). Instead, meta-analyzing the UGR data with UKB did not improve prediction accuracy for any population and most notably decreased accuracy in African ancestry target populations (discovery UKB median RA = 0.22, UGR+UKB median RA = 0.15, **Figure S16**). We hypothesize that this can be explained by the relatively small sample size of UGR adding more noise than signal compared to the other relatively large discovery datasets, but another explanation could come from environmental heterogeneity. When predicting traits using the UKB, BBJ, and PAGE meta-analysis as a discovery cohort, we find that prediction accuracy increases most for the AMR, EAS, and AFR target populations, which more closely resemble the ancestry patterns of PAGE and BBJ (**Figure 4B**). These findings are consistent with ancestry-matched discovery data disproportionately improving prediction accuracy in the corresponding target population (Bigdeli et al., 2019; Lam et al., 2019; Martin et al., 2019).

Large-effect population-enriched genetic variants drive heterogeneity in polygenic score accuracy for blood panel traits

We find that PRS accuracy improvements from higher diversity in the discovery cohorts vary across traits, with the largest increases seen in MCHC and WBC. We searched for specific genetic loci that could explain this pattern by comparing the significance of genetic associations in UKB alone versus the meta-analysis of UKB, BBJ, and PAGE (**Table S6**). For MCHC and WBC in particular, the genetic variants contributing to these improved PRS consist of several well-known population-enriched variants (**Figure 4C** and **4D**). For example, genetic variants that disproportionately explain population-specific risk for MCHC include variants previously associated with hemoglobin concentration, including rs9399137 upstream of *HBS1L* and *MYB* in a study of sickle cell anemia ($p = 5.24e-249$ and $\beta = 0.0783$ in the meta-analysis) (Lettre et al., 2008), rs855791 in *TMPRSS6* ($p = 3.49e-241$, $\beta = 0.0692$) (Benyamin et al., 2009; Chambers et al., 2009), and rs551118 upstream of *PIEZO1* and *CDT1* ($p = 5.18e-100$, $\beta = -0.0451$) (Astle et al., 2016) (**Table S7**). Associations with WBC tend to show more population-enriched associations as shown in the meta-analysis (**Figure 4D**), including rs3936197 in *MED24* ($p = 5.18e-289$, $\beta = -0.0772$), rs58650325 near the high affinity IgE receptor *FCER1A* that initiates the allergic response (1.57e-163, $\beta = -0.097$, also close to *OR10J3*), and rs11533993 in *CDK6* ($p = 1.55e-84$, $\beta = -0.0799$). Thus, genetic architecture and population genetic considerations are important to bear in mind when considering the generalizability of polygenic scores.

Discussion

PRS have been proposed as genetic biomarkers for use in preventative medicine (Khera et al., 2018; Knowles and Ashley, 2018), but are currently limited by low accuracy across populations especially in African ancestry populations (Martin et al., 2019; Sirugo et al., 2019). This study has enabled unique insights into PRS transferability within and among diverse continental African populations as well as among African ancestry populations living in considerably different environments. We demonstrate looming challenges for applying current PRS in African ancestry populations; because relatively few genetic studies have been conducted in African populations coupled with their uniquely deep population histories, PRS accuracy is low but widely variable. Differences in PRS accuracy across diverse African ancestries from different regions can be larger than across out-of-Africa continents. This is particularly problematic as widely-used algorithms that guide health decisions already have ingrained racial biases (Obermeyer et al., 2019), warning of compounding challenges with implementation. We demonstrate that there are clear steps the field can take to work against these biases. Specifically, including ancestrally diverse populations in GWAS discovery cohorts improves accuracy for all populations and especially underrepresented populations more than conducting similarly sized studies with only European ancestry cohorts.

Another advantage of using GWAS from globally diverse populations to compute PRS is the routine inclusion of population-enriched variants. Clear examples such as African-enriched variants in *APOL1* and *G6PD* have been shown to contribute especially high risk of chronic kidney disease and to missed diabetes diagnosis, respectively (E et al., 2018; Rotimi et al., 2017). These examples highlight the importance of studying diverse populations to predict genetic risk of disease equitably by aggregating variants across the spectrum of allele frequencies and effect sizes in different populations. Relevant to the traits studied in genetic analyses here, hematological differences such as anemia are more common in lower income countries in Africa and in African ancestry populations elsewhere compared to European ancestry populations in high income countries, particularly among older individuals. These hematological differences potentially arise in part due to genetic variation as well as the higher prevalence of infectious diseases and pathogens, poorer nutritional status, and altitude (Mugisha et al., 2013, 2016). Here, we show that variants influencing risk of beta thalassemia disproportionately increase PRS accuracy for hemoglobin variation particularly in African ancestry populations. The inclusion of population-enriched variants in PRS could eliminate genetic justifications for race-based medicine, which problematically reinforces implicit racial biases by overemphasizing the link between genetics and race despite the fact that there is more genetic variation within than between populations (Cerdeña et al., 2020).

In addition to reduced PRS accuracy with ancestral distance from GWAS cohorts, genetic nurture, social genetic, and environmental effects can also contribute to low portability of PRS across populations (He et al., 2019; Mostafavi et al., 2020), with some interventions modulating health along PRS strata (Barcellos et al., 2018). In this study, however, ancestry appears to have a larger effect on portability than cohort differences overall. An important distinction when comparing the magnitude of these and other non-genetic effects in other studies is that the traits most accurately genetically predicted here were primarily anthropometric and blood panel traits. When analyzing traits with more sociodemographic influences in increasingly diverse populations, population stratification, confounding, and study design considerations are thornier issues (Kerminen et al., 2019; Novembre and Barton, 2018; Zaidi and Mathieson, 2020). PRS accuracy comparisons across ancestrally similar but environmentally diverse populations are especially important for medically actionable traits. For example, particularly low PRS portability for triglycerides (TG) from European to the Uganda GPC resulted at least in part from effect size heterogeneity that has previously been connected to pleiotropic and gene * environment effects; specifically, most non-transferable genome-wide significant associations with TG showed pleiotropic associations with BMI in Europeans but not Ugandans (Kuchenbaecker et al., 2019).

While PRS currently have limited portability, increased diversity in genetic studies is already decreasing prediction accuracy gaps across populations (Bigdeli et al., 2019; Kuchenbaecker et al., 2019). This is consistent with causal genetic effects tending to be similar across populations but with LD and allele frequency differences modifying marginal effect size estimates (Martin et al., 2019). This is also consistent with trans-ethnic genetic correlations tending to be close to or not significantly different from 1 (Brown et al., 2016; Shi et al., 2020). The most rapid path to closing gaps in PRS transferability is to increase the inclusion of GWAS participants from populations most divergent from those already routinely studied. As empirically demonstrated here, when comparing PRS accuracy calculated from diverse cohort meta-analysis versus data from Europeans only, large-scale GWAS with diverse African populations will most rapidly reduce portability gaps across global populations because they have the most genetic diversity, most rapid linkage disequilibrium decay, and highest genetic divergence from the best studied populations. Major efforts underway such as the Human Hereditary and Health in Africa Initiative, PAGE, All of Us, and NeuroGAP programs (All of Us Research Program Investigators et al., 2019; Hindorff et al., 2018; Mulder et al., 2018; Stevenson et al., 2019; Wojcik et al., 2019) are especially promising for rectifying current PRS gaps and missed scientific opportunities by increasing inclusion of diverse African participants.

Beyond expanding on diversity by increasing the number of study participants in large-scale studies, it is equally important to diversify researchers working on genomics studies. Currently, the vast majority of researchers in genomics studies are of European ancestry (Ginther et al., 2011; Hamrick, 2019; Hoppe et al., 2019), paralleling the over-representation of European-ancestry individuals in genomic studies. The exclusion of African researchers leads to the disparity in research leadership and reduced scientific output from African researchers (Bentley et al., 2020). Efforts such as the NeuroGAP Global Initiative for Neuropsychiatric Genetics Education and Research (GINGER) program (van der Merwe et al., 2018), which provides mentorship and training for early-career investigators on the African continent (particularly in Uganda, Kenya, Ethiopia and South Africa, including several of this study's authors), are important in moving toward a more inclusive and representative research community.

Conclusion

Previous studies that have examined PRS accuracy across globally diverse ancestry groups have demonstrated that accuracy is lowest in African ancestry samples. However, the extent to which this accuracy varies within African-ancestry populations has not been previously investigated. Our findings that prediction accuracy varies by African-ancestry populations is a clear reflection of the vast genetic diversity of the continent. It is therefore critically important to create well-powered GWAS that reflect the full range of diversity within Africa.

Materials and Methods

Genetic and Phenotypic Data

Total counts of individuals by population and/or study are shown in **Table S2**.

1000 Genomes Project

1000 Genomes Project data from the phase 3 integrated call set was accessed and used as a reference panel and for phasing and imputation. (1000 Genomes Project Consortium et al., 2015)

Human Genome Diversity Project (HGDP)

Genotype data for samples from HGDP was publicly available on the Illumina HumanHap650K GWAS array on hg18 (Li et al., 2008). We lifted over the genotype data to the hg19 genome build using hail (http://hail.is).

African Genome Variation Project (AGVP)

As described previously (Gurdasani et al., 2015), the AGVP data consists of dense genotype data from 1,481 individuals from 18 ethno-linguistic groups from Eastern, Western, and Southern Africa when including the Luhya and Yoruba from the 1000 Genomes Project (1000 Genomes Project Consortium et al., 2015). When accessed from the European Genome-Phenome Archive (EGA), "Ethiopian" is the provided population label encompassing the Oromo, Amhara, and Somali groups. After collapsing these groups and counting the 1000 Genomes data separately, 1,307 individuals from 14 populations are uniquely represented in AGVP, and 2,504 individuals from 26 populations are represented in the 1000 Genomes Project data (661 individuals from 7 populations are in the AFR super population grouping).

Drakenstein Children's Health Study (DCHS) in South Africa

The DCHS is an ongoing, multidisciplinary population-based birth cohort study in the Drakenstein area in Paarl (outside Cape Town, South Africa) (Stein et al., 2015; Zar et al., 2015, 2019). After providing informed consent, pregnant women were enrolled during their second trimester (20–28 weeks gestation); maternal-child dyads were then followed through childbirth and longitudinally thereafter. Enrollment occurred from March 2012 to March 2015 at two primary health care clinics - TC Newman (serving a predominantly mixed ancestry population) and Mbekweni (serving a predominantly Black African population). Women were eligible to participate in the DCHS if they attended one of the study clinics, were at least 18 years of age and intended to remain residing in the study area.

Uganda General Population Cohort (GPC)

The rural Uganda GPC of MRC/UVRI & LSHTM Uganda Research Unit was set up in 1989 initially to monitor the HIV epidemic among adults, children, and adolescents, but its mandate has since expanded to include other medical conditions (Asiki et al., 2013). The 'original GPC' is located in the sub-county of Kyamulibwa in rural south-western Uganda with activities having recently been expanded to the neighbouring two peri-urban townships of Lwabenge and Lukaya. The 'original GPC' includes about 10,000 adults and about 10,000 children and adolescents. In 2011, genotype data was generated on more than 5,000 adult participants from nine ethnolinguistic groups using the Illumina HumanOmni2.5 BeadChip at the Sanger Wellcome Trust Institute (Asiki et al., 2013; Heckerman et al., 2016).

UK Biobank (UKB)

The UK Biobank enrolled 500,000 people aged between 40-69 years in 2006-2010 from across the country, as described previously (Bycroft et al., 2018). A more detailed description of the cohort is available on their website: https://www.ukbiobank.ac.uk/. We analyzed phenotypes that overlapped with those studied in the Uganda GPC.

Ancestry analysis in the UK Biobank

As described previously (Bycroft et al., 2018), the UK Biobank consists of approximately 500,000 participants of primarily European ancestry who have thousands of measured or reported phenotypes. To assess polygenic score accuracy across diverse ancestries, we identified populations of ancestral groups at two levels: 1) among continental groups, and 2) among regions in Africa. To define continental ancestries, we first combined reference data from the 1000 Genomes Project and HGDP. We combined these reference datasets into continental ancestries according to their corresponding meta-data (**Table S5**). We then ran PCA on unrelated individuals from the reference dataset. To partition individuals in the UK Biobank based on their continental ancestry, we used the PC loadings from the reference dataset to project UK Biobank individuals into the same PC space. We trained a random forest classifier given continental ancestry meta-data (AFR = African, AMR = admixed American, CSA = Central/South Asian, EAS = East Asian, EUR = European, and MID = Middle Eastern) based on the top 6 PCs from the reference training data. We applied this random forest to the projected UK Biobank PCA data and assigned initial ancestries if the random forest probability was >50% (similar results obtained for $p > 0.9$), otherwise individuals were dropped from further analysis.

We next further partitioned African ancestry individuals using the same random forest approach as above but without further probability thresholding using African ancestry reference data from AGVP, HGDP, and the 1000 Genomes Project. We partitioned these reference data into UN regional codes with an additional region for Ethiopian populations given their unique population history and collapsing in African Genome Variation Project data (Admixed, Central, East, Ethiopia, South, and West Africa), as shown in **Table S5**. PCA with reference data at the continental and subcontinental level within Africa are shown in **Figures S10-11**.

Phasing and imputation

We used the Ricopili pipeline to conduct pre-imputation QC and perform phasing and imputation for AGVP and the Uganda GPC (Lam et al., 2020). This pipeline was also used on the DCHS data, as described previously (Duncan et al., 2018). Briefly, we phased the data using Eagle 2.3.5 and imputed variants using minimac3 in chunks ≥ 3 Mb. The 1000 Genomes phase 3 haplotypes were used as the reference panel for phasing and imputation. For the AGVP, we used strict best guess genotypes where a variant was called if it had a probability of $p > 0.8$ and a missing rate less than 0.01 and MAF > 5%. Then, variants with MAF < 0.001 were excluded from the dataset. For Uganda GPC, we used combined best guess genotypes where a variant was called if it had a probability $p > 0.8$ or set to missing otherwise. Then, SNPs were filtered to keep sites with missingness < 0.01 and MAF > 0.05. We used genotype dosages when computed PRS.

PCA

Only SNPs with high imputation quality (INFO>0.8) were considered for principal component analysis. We computed the first 20 principal components using plink with the --pca flag for autosomal SNPs MAF > 0.05 and individual missingness < 0.05.

Simulation setup

To test the PRS prediction accuracy within and across African populations, we simulated four quantitative traits while varying heritabilities ($h^2 = 0.1$, 0.2, 0.4 and 0.8) as follows:

We randomly assigned an effect size to 5, 20, 100, 2,000, 10,000 and 50,000 causal variants, respectively. The causal effect was calculated based on the relationship between effect size and minor allele frequency as shown by (Schoech et al., 2019). We then calculated an individual's 'true' polygenic risk as the sum of all causal effects using the --score flag in PLINK v1.07B (Chang et al., 2015). True polygenic scores were standardized to a mean of zero and standard deviation of 1. To account for the contribution of environmental risk factors, we assigned environmental effects from a normal random distribution (mean = 0 and sd = 1). The phenotype was generated according to its heritability as the weighted sum of the true polygenic risk and a random environmental effect as below:

$$phenotype \; = \; true \; polygenic \; risk \; + \; (\,1 - h^2) \; \times environmental \; effect$$

We then conducted GWAS for the simulated phenotype by splitting the AGVP dataset into three groups broadly representing the three geographical areas where samples were obtained from: East (n = 589), West (n = 517) and South Africa (n = 186, **Figure 2A**). To allow for the quantification of PRS prediction accuracy across the geographical regions, each group was further split into discovery and target cohorts. The size of the target cohorts was maintained at n = 186 across all groups, while the discovery cohort consisted of all remaining individuals (East n = 403, West n = 331, and no South Africans). We conducted a linear regression for all the simulated traits for the East and West discovery datasets, controlling for the first 20 principal components.

In PLINK v1.07, independent SNP sets were obtained for each discovery cohort by clumping SNPs from corresponding summary statistics files with an R^2 value greater than 0.1 using in-sample LD and within 500 kb of each other. The effect sizes from the SNP set was used as weights to compute PRS for all three of our target datasets for a range of *P*-values (5e-08, 1e-06, 1e-04, 1e-03, 1e-02, 0.05, 0.1, 0.2, 0.5 and all). PRS was calculated as the sum of all SNPs multiplied by their effect sizes. We calculated PRS for each of the target datasets using the summary statistics from the discovery dataset GWAS (**Figure 2B**).

Heritability estimation

For the Ugandan GPC, we relied on heritability estimates of 34 quantitative traits computed previously (Gurdasani et al., 2019). For UK Biobank, we computed heritability estimates for the same traits using LD score regression with the default model (i.e. without any functional annotations) (Bulik-Sullivan et al., 2015) and using used population-matched LD score references from European populations downloaded from the authors' website (https://data.broadinstitute.org/alkesgroup/LDSCORE/).

Polygenic score calculation

All PRS were calculated using a pruning and thresholding approach implemented either in plink2 or in hail using custom scripts. All clumping was done in plink2 using an LD threshold of r^2 = 0.1 and a window size of 500 kb with discovery cohort population-specific reference panels. We calculated PRS using plink2 with the --score and --q-score-range flags for AGVP simulations and DCHS. We wrote custom scripts in hail (http://hail.is) to calculate PRS in the Uganda GPC and UK Biobank data due to the larger sample sizes (see **Web resources**). For imputed genotypes, we used SNP dosages in PRS calculations. We computed 10 PRS for each analysis using the following p-value thresholds: 1, 0.5, 0.2, 0.1, 0.05, 0.01, 1e-3, 1e-4, 1e-6, 5e-8. The PRS that explained the most phenotypic variance is shown in most figures.

We calculated PRS accuracy for continuous traits computed with custom scripts in R (**Web resources**). For AGVP simulations and DCHS (because all participants were mothers of a similar age), we included the first 10 PCs as covariates when computing the partial R^2 specifically attributable to the PRS. For Uganda GPC data, we included age, sex, and the first 10 PCs when computing partial R^2 of the PRS. For consistency with the GWAS that were run in UK Biobank previously (Howrigan, 2017) and here with a holdout target set, we included, age, sex, age^2, age*sex, age^2*sex, and the first 10 PCs as covariates when computing the PRS partial R^2. (The UKB European GWAS included 20 PCs, but fewer were used here due to the particularly small sample sizes of some other target ancestry groups, **Table S5**, coupled with minimal population structure observed in PCs lower than PC10).

Meta-analysis

We used plink2 to conduct inverse variance-weighted meta-analysis across GWAS summary statistics with the --meta-analysis option.

LD reference panels and clumping

All PRS calculations required an LD panel for clumping. Our analyses used in-sample LD where feasible and reference panel data as a proxy with ancestry matching from the 1000 Genomes Project phase 3 data when individual-level data was unavailable. We weighted the ancestral representation of each population per trait matching at the continental level. We matched individuals as follows:

Cohort	**1000 Genomes phase 3 reference data**
BBJ	East Asian (EAS)
UKB	European (EUR)
UGR	African (AFR)
PAGE	Proportional weighting of AFR, EAS, AMR (depending on trait, see **Table S6** description for more detail)

We then used the maximal number of individuals available when weighting proportionally to construct this reference panel. For example, in the meta-analysis of height across the UKB, BBJ, and PAGE cohorts, UKB has the largest sample size in the discovery cohort (N = 350,353), so all Europeans from 1000 Genomes were included in the reference panel (N = 503), then a random sampling of EAS, AFR, and AMR individuals were included proportionally to the overall diversity of the discovery cohorts in the meta-analysis.

Acknowledgements

We thank Lori Chibnik, Bizu Gelaye, Kristi Post, and Courtney White for facilitating the GINGER program and making this work possible. This work was supported by funding from the National Institutes of Health (K99MH117229 to A.R.M.; K01MH121659 and T32MH017119 to E.G.A.). UK Biobank analyses were conducted via application 31063. The DCHS is funded by the Bill and Melinda Gates Foundation (OPP 1017641). Additional support for HJZ and DJS, and for the research reported in this publication, was provided by the South African Medical Research Council (SAMRC). The SAMRC provided additional support through its Division of Research Capacity Development under the National Health Scholarship programme from funding received from the Public Health Enhancement Fund/South African National Department of Health. The views and opinions expressed are those of the authors and do not necessarily represent the official views of the SAMRC. We thank the mothers and their children for participating in the DCHS and the study staff, the clinical and administrative staff of the Western Cape Government Health Department at Paarl Hospital and at the clinics for support of the study. We also thank all research participants in the UK Biobank, BioBank Japan, PAGE study, UGR and Uganda GPC, and AGVP studies.

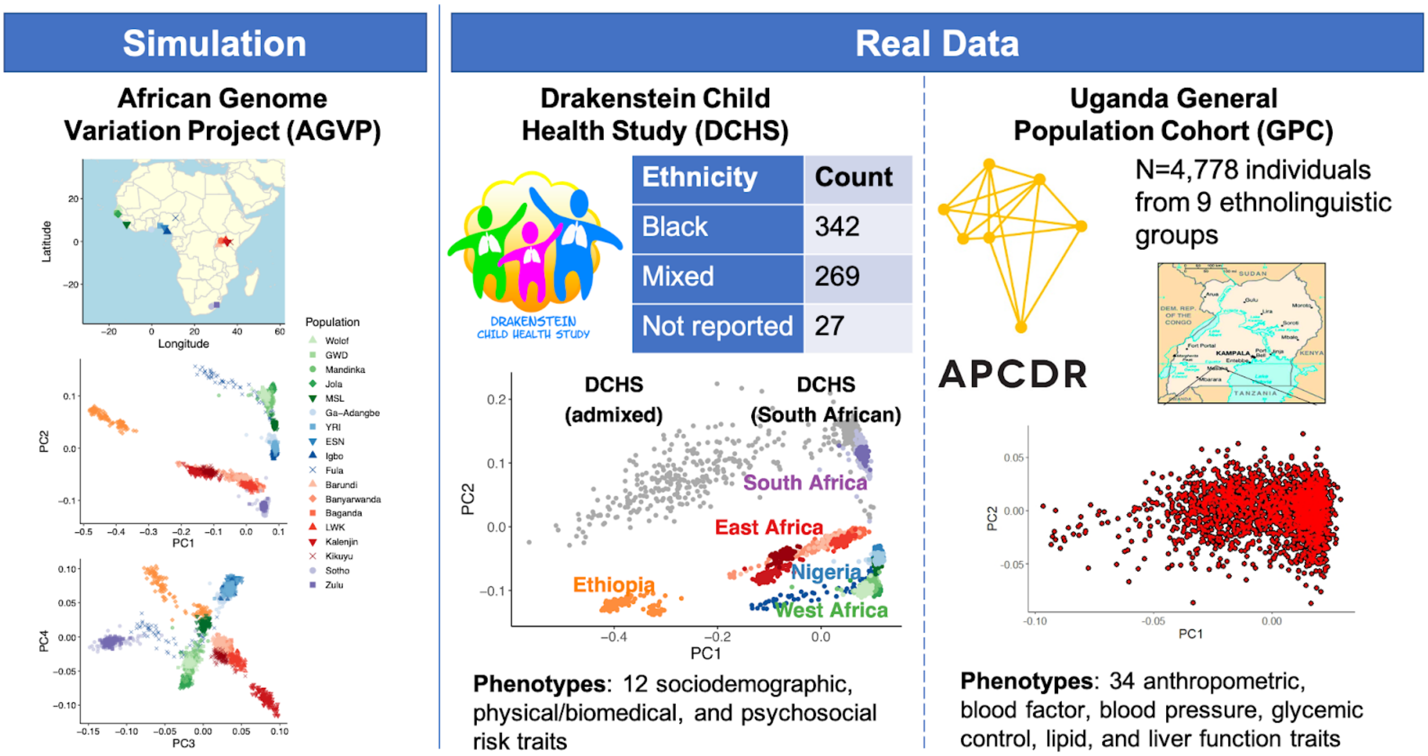
Data and Code Availability

All data used in this study are publicly available. Data from the African Genome Variation Project was accessed by combining EGAD00010001045, EGAD00010001046, EGAD00010001047, EGAD00010001048, EGAD00010001049, EGAD00010001050, EGAD00010001051, EGAD00010001052, EGAD00010001053, EGAD00010001054, EGAD00010001055, EGAD00010001056, EGAD00010001057, and

EGAD00010001058. The Drakenstein Child Health Study is committed to the principle of data sharing. De-identified data will be made available to requesting researchers as appropriate. Requests for collaborations to undertake data analysis are welcome. More information can be found on our website (http://www.paediatrics.uct.ac.za/scah/dclhs). Uganda GPC genetic data used in this paper were accessed through EGAD00010000965 and phenotype data was accessed via sftp from EGA (reference: DD_PK_050716 gwas_phenotypes_28Oct14.txt). We accessed data from the UK Biobank with application 31063. BioBank Japan summary statistics were accessed from http://jenger.riken.jp/en/result. GWAS summary statistics for the Population Architecture using Genomics and Epidemiology (PAGE) study were accessed through the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/downloads/summary-statistics).

All code used in analysis is available here: https://github.com/armartin/africa_prs.

Figures



Ethnicity	Count
Black	342
Mixed	269
Not reported	27

Phenotypes: 12 sociodemographic, physical/biomedical, and psychosocial risk traits

N=4,778 individuals from 9 ethnolinguistic groups

Phenotypes: 34 anthropometric, blood factor, blood pressure, glycemic control, lipid, and liver function traits

Figure 1 - Project overview of genetic and phenotypic datasets used to assess polygenic score generalizability within and across diverse African populations. Using publicly available GWAS data from primarily Eurocentric populations, we measure how polygenic scores perform in Africa. In simulations, we use AGVP genetic data and simulated phenotypes to assess polygenic score generalizability within Africa. In real data, we use two datasets to measure polygenic score accuracy: the South African DCHS cohort data and the Ugandan GPC cohort data.

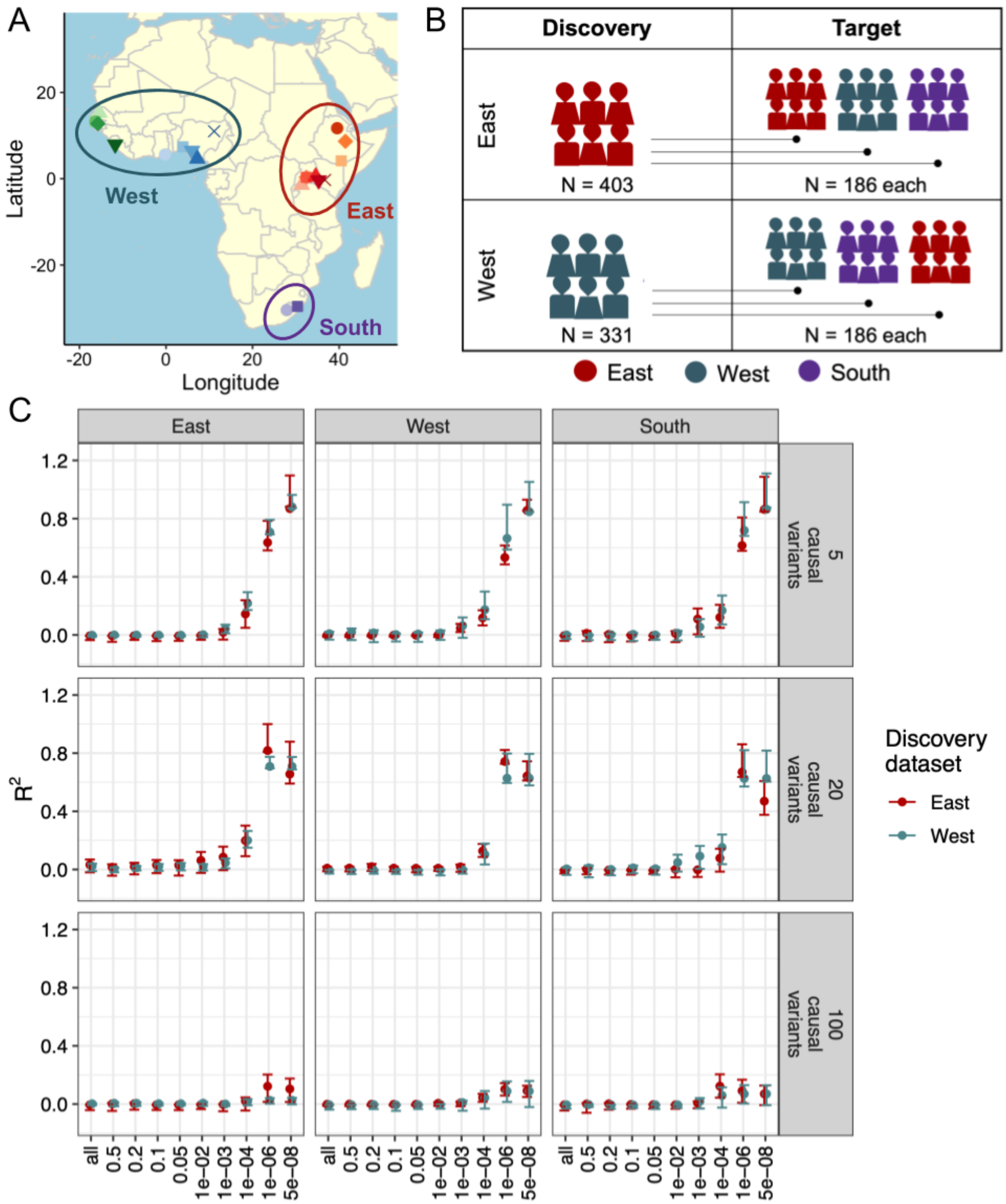


Figure 2 - Simulated GWAS and polygenic scores indicate differential prediction accuracy across diverse regions of Africa using genetic data from the AGVP. A) Populations were grouped into East, West, and South based on the United Nations geoscheme groupings. B) GWAS discovery cohorts included East (N = 403) and West (N=331) African individuals, which were independent of each target cohort (N = 186 individuals per region). South Africans were excluded from the discovery population due to the limited total sample size (2 populations and 186 individuals total). C) Predictive accuracy of the simulated quantitative trait at the heritability of 0.8. The predictive accuracy was calculated for six categories of causal variants for the West and East discovery cohorts, across ten *p*-value thresholds.

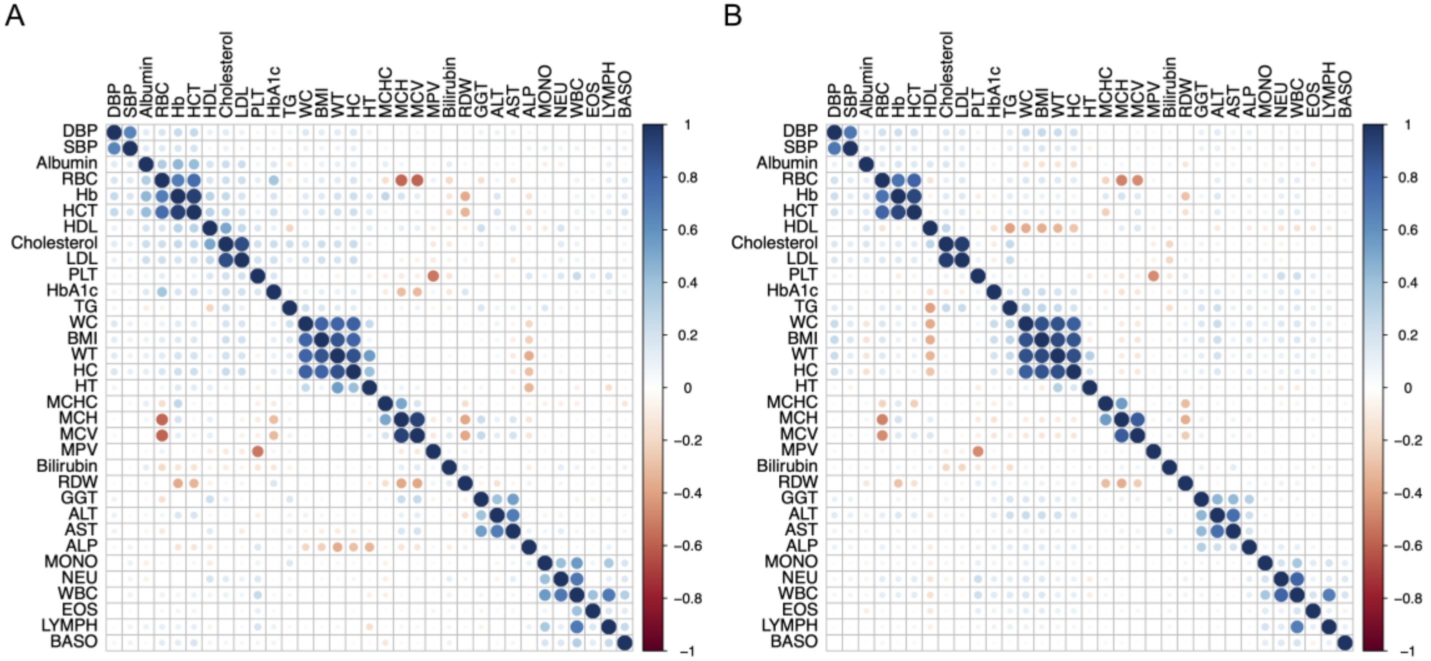


Figure 3 - Phenotype and genotype correlations among 33 quantitative traits measured in the Uganda GPC data and the UK Biobank. A) Phenotypic correlations measured in traits in the Uganda GPC among unrelated individuals. B) Phenotypic correlations in the UK Biobank European ancestry unrelated individuals. A-B) Phenotypes were mean centered and adjusted for age and sex within each cohort prior to correlation analysis. The order of each phenotype correlation is determined by hierarchical clustering in the Uganda GPC.

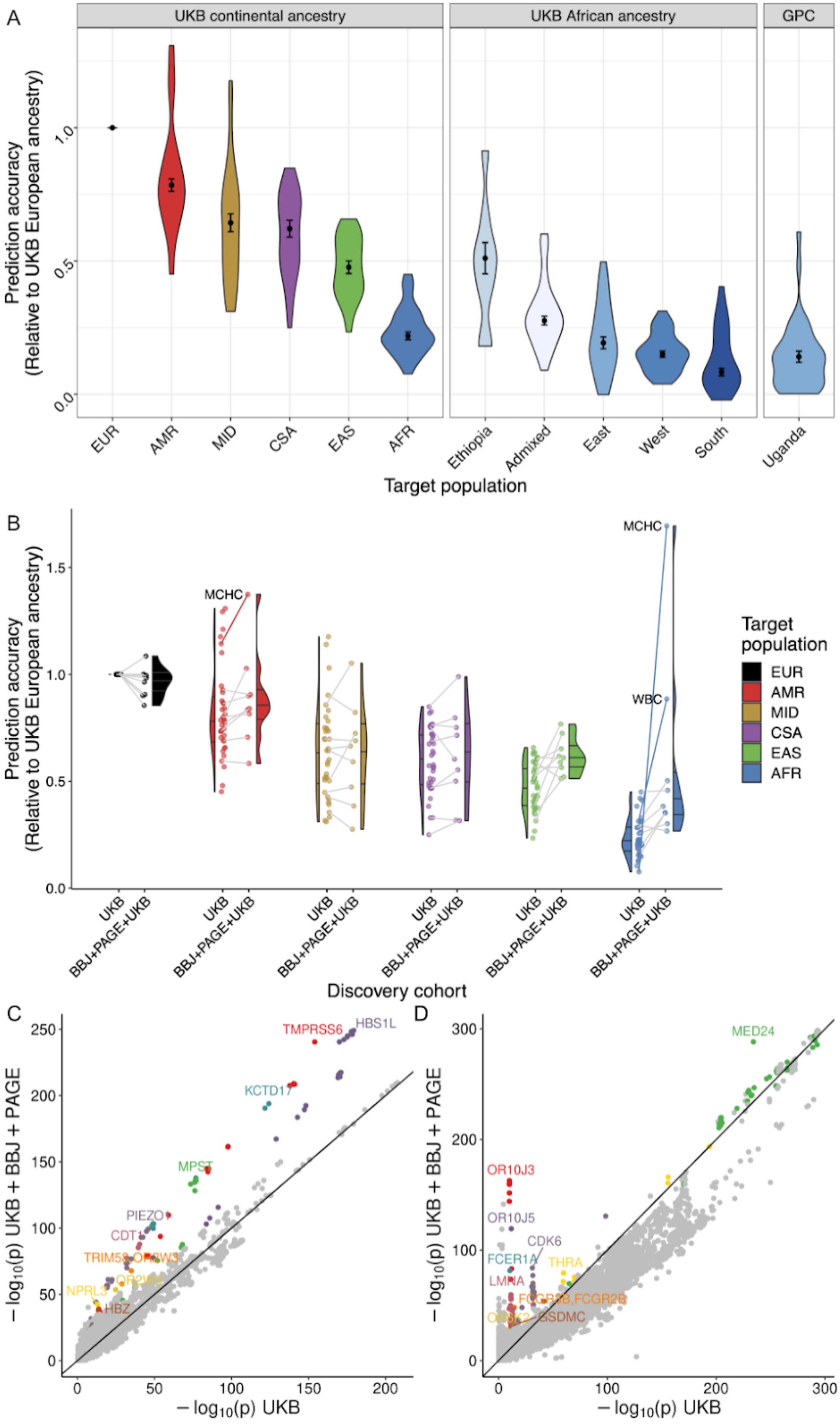

A
UKB continental ancestry
UKB African ancestry
GPC
Prediction accuracy
(Relative to UKB European ancestry)
0.0
0.5
1.0
EUR
AMR
MID
CSA
EAS
AFR
Ethiopia
Admixed
East
West
South
Uganda
Target population
B
Prediction accuracy
(Relative to UKB European ancestry)
0.0
0.5
1.0
1.5
MCHC
MCHC
WBC
UKB
BBJ+PAGE+UKB
Discovery cohort
Target population
EUR
AMR
MID
CSA
EAS
AFR
C
$-\log_{10}(p)$ UKB + BBJ + PAGE
$-\log_{10}(p)$ UKB
0
50
100
150
200
250
TMPRSS6
HBS1L
KCTD17
MPST
PIEZO1
CDT1
TRIM58,OR2W3
OR2W3
NPRL3
HBZ
D
$-\log_{10}(p)$ UKB + BBJ + PAGE
$-\log_{10}(p)$ UKB
0
100
200
300
MED24
OR10J3
OR10J5
CDK6
FCER1A
THRA
LMNA
FCGR3B,FCGR2B
OR6K2
GSDMC

Figure 4 - PRS accuracy and corresponding genetic variant contributions for up to 34 traits within and across diverse ancestries. A) PRS accuracy relative to European ancestry individuals in diverse target ancestries. Discovery data consisted of GWAS summary statistics from UK Biobank (UKB) European ancestry data. Target data consisted of globally diverse continental ancestries (including withheld European target individuals) and regional African ancestry participants from UKB, or unrelated individuals from the Uganda GPC cohort. Traits were filtered to those with a 95% confidence interval range in PRS accuracy < 0.08. B) PRS accuracy from a homogeneous versus multi-ancestry discovery dataset. GWAS discovery data consisted of summary statistics from UKB European ancestry data only or from the meta-analysis of UKB, BioBank Japan (BBJ), and Population Architecture using Genomics and Epidemiology (PAGE). Target populations are from the UKB. Lines connect the 10 traits available in both discovery cohorts to indicate how accuracy changed for the same trait in the UKB only versus meta-analyzed discovery data, while half violin plots show the distribution across all phenotypes in each discovery cohort. When lines are missing, the trait is absent in PAGE. Trait outliers are labeled in text and with solid lines. A-B) Relative PRS accuracies are compared to the maximum for each trait in target samples withheld from discovery consisting of UKB European ancestry individuals. To simplify comparisons, only the polygenic scores with the highest prediction accuracy are shown here. Colors in these two panels correspond to the same continental ancestries. C-D) Trait-specific genetic outlier plots. QQ-like plot showing p-values in UKB only versus multi-cohort meta-analysis of UKB, BBJ, and PAGE. The ten regions that are genome-wide significant in both dataset and show the most significant differences are colored and labeled for: C) MCHC, and D) WBC.

References

1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., et al. (2015). A global reference for human genetic variation. Nature *526*, 68–74.

All of Us Research Program Investigators, Denny, J.C., Rutter, J.L., Goldstein, D.B., Philippakis, A., Smoller, J.W., Jenkins, G., and Dishman, E. (2019). The “All of Us” Research Program. N. Engl. J. Med. *381*, 668–676.

Asiki, G., Murphy, G., Nakiyingi-Miiro, J., Seeley, J., Nsubuga, R.N., Karabarinde, A., Waswa, L., Biraro, S., Kasamba, I., Pomilla, C., et al. (2013). The general population cohort in rural south-western Uganda: a platform for communicable and non-communicable disease studies. Int. J. Epidemiol. *42*, 129–141.

Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell *167*, 1415–1429.e19.

Barcellos, S.H., Carvalho, L.S., and Turley, P. (2018). Education can reduce health differences related to genetic risk of obesity. Proc. Natl. Acad. Sci. U. S. A. *115*, E9765–E9772.

Bentley, A.R., Callier, S.L., and Rotimi, C.N. (2020). Evaluating the promise of inclusion of African ancestry populations in genomics. Npj Genomic Medicine *5*.

Benyamin, B., Ferreira, M.A.R., Willemsen, G., Gordon, S., Middelberg, R.P.S., McEvoy, B.P., Hottenga, J.-J., Henders, A.K., Campbell, M.J., Wallace, L., et al. (2009). Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. Nat. Genet. *41*, 1173–1175.

Bigdeli, T.B., Genovese, G., Georgakopoulos, P., Meyers, J.L., Peterson, R.E., Iyegbe, C.O., Medeiros, H., Valderrama, J., Achtyes, E.D., Kotov, R., et al. (2019). Contributions of common genetic variants to risk of schizophrenia among individuals of African and Latino ancestry. Mol. Psychiatry.

Bitarello, B.D., and Mathieson, I. (2020). Polygenic Scores for Height in Admixed Populations. G3 *10*, 4027–4036.

Brown, B.C., Asian Genetic Epidemiology Network Type 2 Diabetes Consortium, Ye, C.J., Price, A.L., and Zaitlen, N. (2016). Transethnic Genetic-Correlation Estimates from Summary Statistics. Am. J. Hum. Genet. *99*, 76–88.

Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. *47*, 291–295.

Busby, G.B., Band, G., Si Le, Q., Jallow, M., Bougama, E., Mangano, V.D., Amenga-Etego, L.N., Enimil, A., Apinjoh, T., Ndila, C.M., et al. (2016). Admixture into and within sub-Saharan Africa. Elife *5*.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O’Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209.

Campbell, M.C., and Tishkoff, S.A. (2008). African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu. Rev. Genomics Hum. Genet. *9*, 403–433.

Cerdeña, J.P., Plaisime, M.V., and Tsai, J. (2020). From race-based to race-conscious medicine: how anti-racist uprisings call us to act. Lancet *396*, 1125–1128.

Chambers, J.C., Zhang, W., Li, Y., Sehmi, J., Wass, M.N., Zabaneh, D., Hoggart, C., Bayele, H., McCarthy, M.I., Peltonen, L., et al. (2009). Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat. Genet. *41*, 1170–1172.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.

Choudhury, A., Aron, S., Botigué, L.R., Sengupta, D., Botha, G., Bensellak, T., Wells, G., Kumuthini, J., Shriner, D., Fakim, Y.J., et al. (2020). High-depth African genomes inform human migration and health. Nature *586*, 741–748.

Duncan, L.E., Ratanatharathorn, A., Aiello, A.E., Almli, L.M., Amstadter, A.B., Ashley-Koch, A.E., Baker, D.G., Beckham, J.C., Bierut, L.J., Bisson, J., et al. (2018). Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. Mol. Psychiatry *23*, 666–673.

E, W., Wheeler, E., Leong, A., Liu, C.T., Hievert, M.F., Strawbridge, R., and Podmore, C. (2018). Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. Yearbook of Paediatric Endocrinology.

Genovese, G., Friedman, D.J., Ross, M.D., Lecordier, L., Uzureau, P., Freedman, B.I., Bowden, D.W., Langefeld, C.D., Oleksyk, T.K., Uscinski Knob, A.L., et al. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science *329*, 841–845.

Ginther, D.K., Schaffer, W.T., Schnell, J., Masimore, B., Liu, F., Haak, L.L., and Kington, R. (2011). Race, ethnicity, and NIH research awards. Science *333*, 1015–1019.

Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., Karthikeyan, S., Iles, L., Pollard, M.O., Choudhury, A., et al. (2015). The African Genome Variation Project shapes medical genetics in Africa. Nature *517*, 327–332.

Gurdasani, D., Carstensen, T., Fatumo, S., Chen, G., Franklin, C.S., Prado-Martinez, J., Bouman, H., Abascal, F., Haber, M., Tachmazidou, I., et al. (2019). Uganda Genome Resource Enables Insights into Population History and Genomic Discovery in Africa. Cell *179*, 984–1002.e36.

Hamrick, K. (2019). Women, minorities, and persons with disabilities in science and engineering: 2019. National Science Foundation, National Center for Science and Engineering Statistics (NCSES), Alexandria, VA, Special Report NSF 19–304.

He, Y., Lakhani, C.M., Manrai, A.K., and Patel, C.J. (2019). Poly-Exposure and Poly-Genomic Scores Implicate Prominent Roles of Non-Genetic and Demographic Factors in Four Common Diseases in the UK.

of Health, U.S.D., Services, H., and Others (2017). 2016 National Healthcare Quality and Disparities Report.

Heckerman, D., Gurdasani, D., Kadie, C., Pomilla, C., Carstensen, T., Martin, H., Ekoru, K., Nsubuga, R.N., Ssenyomo, G., Kamali, A., et al. (2016). Linear mixed model for heritability estimation that explicitly addresses environmental variation. Proc. Natl. Acad. Sci. U. S. A. *113*, 7377–7382.

Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodríguez-Botigué, L., Ramachandran, S., Hon, L., Brisbin, A., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. Proc. Natl. Acad. Sci. U. S. A. *108*, 5154–5162.

Henn, B.M., Cavalli-Sforza, L.L., and Feldman, M.W. (2012a). The great human expansion. Proc. Natl. Acad. Sci. U. S. A. *109*, 17758–17764.

Henn, B.M., Botigué, L.R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J.K., Fadhlaoui-Zid, K., Zalloua, P.A., Moreno-Estrada, A., Bertranpetit, J., et al. (2012b). Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS Genet. *8*, e1002397.

Hero, J.O., Zaslavsky, A.M., and Blendon, R.J. (2017). The United States Leads Other Nations In Differences By Income In Perceptions Of Health And Health Care. Health Aff. *36*, 1032–1040.

Hindorff, L.A., Bonham, V.L., Brody, L.C., Ginoza, M.E.C., Hutter, C.M., Manolio, T.A., and Green, E.D. (2018). Prioritizing diversity in human genomics research. Nat. Rev. Genet. *19*, 175–185.

Hodgson, J.A., Mulligan, C.J., Al-Meeri, A., and Raaum, R.L. (2014). Early back-to-Africa migration into the Horn of Africa. PLoS Genet. *10*, e1004393.

Hoppe, T.A., Litovitz, A., Willis, K.A., Meseroll, R.A., Perkins, M.J., Hutchins, B.I., Davis, A.F., Lauer, M.S., Valantine, H.A., Anderson, J.M., et al. (2019). Topic choice contributes to the lower rate of NIH awards to African-American/black scientists. Sci Adv *5*, eaaw7238.

Howrigan, D. (2017). Details and Considerations of the UK Biobank GWAS.

International Schizophrenia Consortium, Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., and Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature *460*, 748–752.

Kerminen, S., Martin, A.R., Koskela, J., Ruotsalainen, S.E., Havulinna, A.S., Surakka, I., Palotie, A., Perola, M., Salomaa, V., Daly, M.J., et al. (2019). Geographic Variation and Bias in the Polygenic Scores of Complex Diseases and Traits in Finland. Am. J. Hum. Genet. *104*, 1169–1181.

Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P., Lander, E.S., Lubitz, S.A., Ellinor, P.T., et al. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat. Genet. *50*, 1219–1224.

Knowles, J.W., and Ashley, E.A. (2018). Cardiovascular disease: The rise of the genetic risk score. PLoS Med. *15*, e1002546.

Kuchenbaecker, K., Telkar, N., Reiker, T., Walters, R.G., Lin, K., Eriksson, A., Gurdasani, D., Gilly, A., Southam, L., Tsafantakis, E., et al. (2019). The transferability of lipid loci across African, Asian and European cohorts. Nat. Commun. *10*, 4330.

Lam, M., Chen, C.-Y., Li, Z., Martin, A.R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M., Benyamin, B., Brown, B.C., et al. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. Nat. Genet. 1–9.

Lam, M., Awasthi, S., Watson, H.J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., Karlsson, R., Frei, O., Fan, C.-C., De Witte, W., et al. (2020). RICOPILI: Rapid Imputation for COnsortias PIpeLIne. Bioinformatics *36*, 930–933.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N., et al. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and β-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc. Natl. Acad. Sci. U. S. A. *105*, 11869–11874.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., et al. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. Science *319*, 1100–1104.

Manrai, A.K., Funke, B.H., Rehm, H.L., Olesen, M.S., Maron, B.A., Szolovits, P., Margulies, D.M., Loscalzo, J., and Kohane, I.S. (2016). Genetic Misdiagnoses and the Potential for Health Disparities. N. Engl. J. Med. *375*, 655–665.

Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D., and Kenny, E.E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. Am. J. Hum. Genet. *100*, 635–649.

Martin, A.R., Teferra, S., Möller, M., Hoal, E.G., and Daly, M.J. (2018). The critical needs and challenges for genetic architecture studies in Africa. Curr. Opin. Genet. Dev. *53*, 113–120.

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

van der Merwe, C., Mwesiga, E.K., McGregor, N.W., Ejigu, A., Tilahun, A.W., Kalungi, A., Akimana, B., Dubale, B.W., Omari, F., Mmochi, J., et al. (2018). Advancing neuropsychiatric genetics training and collaboration in

Africa. Lancet Glob Health *6*, e246–e247.

Morales, J., Welter, D., Bowler, E.H., Cerezo, M., Harris, L.W., McMahon, A.C., Hall, P., Junkins, H.A., Milano, A., Hastings, E., et al. (2018). A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. Genome Biol. *19*, 21.

Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J.K., and Przeworski, M. (2020). Variable prediction accuracy of polygenic scores within an ancestry group. Elife *9*.

Mugisha, J.O., Baisley, K., Asiki, G., Seeley, J., and Kuper, H. (2013). Prevalence, types, risk factors and clinical correlates of anaemia in older people in a rural Ugandan population. PLoS One *8*, e78394.

Mugisha, J.O., Seeley, J., and Kuper, H. (2016). Population based haematology reference ranges for old people in rural South-West Uganda. BMC Res. Notes *9*, 433.

Mulder, N., Abimiku, A. ’le, Adebamowo, S.N., de Vries, J., Matimba, A., Olowoyo, P., Ramsay, M., Skelton, M., and Stein, D.J. (2018). H3Africa: current perspectives. Pharmgenomics. Pers. Med. *11*, 59–66.

Nagai, A., Hirata, M., Kamatani, Y., Muto, K., Matsuda, K., Kiyohara, Y., Ninomiya, T., Tamakoshi, A., Yamagata, Z., Mushiroda, T., et al. (2017). Overview of the BioBank Japan Project: Study design and profile. J. Epidemiol. *27*, S2–S8.

Nalwanga, D., Musiime, V., Kizito, S., Kiggundu, J.B., Batte, A., Musoke, P., and Tumwine, J.K. (2020). Mortality among children under five years admitted for routine care of severe acute malnutrition: a prospective cohort study from Kampala, Uganda. BMC Pediatr. *20*, 182.

Novembre, J., and Barton, N.H. (2018). Tread Lightly Interpreting Polygenic Tests of Selection. Genetics *208*, 1351–1355.

Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science *366*, 447–453.

Pagani, L., Schiffels, S., Gurdasani, D., Danecek, P., Scally, A., Chen, Y., Xue, Y., Haber, M., Ekong, R., Oljira, T., et al. (2015). Tracing the route of modern humans out of Africa by using 225 human genome sequences from Ethiopians and Egyptians. Am. J. Hum. Genet. *96*, 986–991.

Popejoy, A.B., and Fullerton, S.M. (2016). Genomics is failing on diversity. Nature News *538*, 161.

Roser, M. (2013). Life expectancy. Our World in Data.

Rotimi, C.N., Bentley, A.R., Doumatey, A.P., Chen, G., Shriner, D., and Adeyemo, A. (2017). The genomic landscape of African populations in health and disease. Hum. Mol. Genet. *26*, R225–R236.

Schoech, A.P., Jordan, D.M., Loh, P.-R., Gazal, S., O’Connor, L.J., Balick, D.J., Palamara, P.F., Finucane, H.K., Sunyaev, S.R., and Price, A.L. (2019). Quantification of frequency-dependent genetic architectures in 25 UK Biobank traits reveals action of negative selection. Nat. Commun. *10*, 790.

Scutari, M., Mackay, I., and Balding, D. (2016). Using Genetic Distance to Infer the Accuracy of Genomic Prediction. PLoS Genet. *12*, e1006288.

Shi, H., Burch, K.S., Johnson, R., Freund, M.K., Kichaev, G., Mancuso, N., Manuel, A.M., Dong, N., and Pasaniuc, B. (2020). Localizing Components of Shared Transethnic Genetic Architecture of Complex Traits from GWAS Summary Data. Am. J. Hum. Genet. *106*, 805–817.

Sirugo, G., Williams, S.M., and Tishkoff, S.A. (2019). The Missing Diversity in Human Genetic Studies. Cell *177*, 26–31.

Stein, D.J., Koen, N., Donald, K.A., Adnams, C.M., Koopowitz, S., Lund, C., Marais, A., Myers, B., Roos, A., Sorsdahl, K., et al. (2015). Investigating the psychosocial determinants of child health in Africa: The Drakenstein Child Health Study. J. Neurosci. Methods *252*, 27–35.

Stevenson, A., Akena, D., Stroud, R.E., Atwoli, L., Campbell, M.M., Chibnik, L.B., Kwobah, E., Kariuki, S.M., Martin, A.R., de Menil, V., et al. (2019). Neuropsychiatric Genetics of African Populations-Psychosis (NeuroGAP-Psychosis): a case-control study protocol and GWAS in Ethiopia, Kenya, South Africa and Uganda. BMJ Open *9*, e025469.

Uren, C., Kim, M., Martin, A.R., Bobo, D., Gignoux, C.R., van Helden, P.D., Möller, M., Hoal, E.G., and Henn, B.M. (2016). Fine-Scale Human Population Structure in Southern Africa Reflects Ecogeographic Boundaries. Genetics *204*, 303–314.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., and Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. Am. J. Hum. Genet. *101*, 5–22.

de Vlaming, R., Okbay, A., Rietveld, C.A., Johannesson, M., Magnusson, P.K.E., Uitterlinden, A.G., van Rooij, F.J.A., Hofman, A., Groenen, P.J.F., Thurik, A.R., et al. (2017). Meta-GWAS Accuracy and Power (MetaGAP) Calculator Shows that Hiding Heritability Is Partially Due to Imperfect Genetic Correlations across Studies. PLoS Genet. *13*, e1006495.

Wang, Y., Guo, J., Ni, G., Yang, J., Visscher, P.M., and Yengo, L. (2020). Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations.

Wojcik, G.L., Graff, M., Nishimura, K.K., Tao, R., Haessler, J., Gignoux, C.R., Highland, H.M., Patel, Y.M., Sorokin, E.P., Avery, C.L., et al. (2019). Genetic analyses of diverse populations improves discovery for complex traits. Nature.

Wray, N.R., Yang, J., Hayes, B.J., Price, A.L., Goddard, M.E., and Visscher, P.M. (2013). Pitfalls of predicting complex traits from SNPs. Nat. Rev. Genet. *14*, 507–515.

Zaidi, A.A., and Mathieson, I. (2020). Demographic history impacts stratification in polygenic scores.

Zar, H.J., Barnett, W., Myer, L., Stein, D.J., and Nicol, M.P. (2015). Investigating the early-life determinants of illness in Africa: the Drakenstein Child Health Study. Thorax *70*, 592–594.

Zar, H.J., Pellowski, J.A., Cohen, S., Barnett, W., Vanker, A., Koen, N., and Stein, D.J. (2019). Maternal health and birth outcomes in a South African birth cohort study. PLoS One *14*, e0222399.