Different genetic architectures of complex traits and their relevance to polygenic score performance

Nuno R. G. Carvalho[1,2], Adrian M. Harris[1,2], and Joseph Lachance[1*]

[1]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA

[2]These authors contributed equally

*Corresponding author: joseph.lachance@biology.gatech.edu

Keywords: complex traits, genetic architecture, Gini coefficients, polygenic scores, population genetics

Running title: Gini coefficients and polygenic traits

Abstract

Complex traits differ in their genetic architectures, and these differences can affect polygenic score performance. Examining 177 complex traits from the UK Biobank, we first identified pairs of traits that have trait-associated SNPs in shared genomic regions. We then compared and contrasted three aspects of genetic architecture (SNP heritability, trait-specific recombination rates, and a novel metric of polygenicity) with three aspects of polygenic score performance (correlations between predicted and actual trait values, portability of genetic predictions, and divergence across populations). Although highly heritable traits tended to be easier to predict, heritability was largely uninformative with respect to the portability of genetic predictions. By contrast, there was a positive relationship between trait-specific recombination rates and the portability of genetic predictions. Analyzing 100kb bins, we used Gini coefficients to quantify the extent that SNP heritability is unequally distributed across the genome. Polygenic score performance was largely independent of Gini – traits with more Mendelian architectures need not be easier to predict. By contrast, Gini coefficients were negatively correlated with the prevalence of binary traits. We also found that binary traits were more difficult to predict than quantitative traits. Interestingly, lifestyle and psychological traits tend to have low heritability, low Gini coefficients, as well as poor predictability and portability across populations. Because of this, our results caution against the application of polygenic scores to traits like general happiness, alcohol frequency, and average income, especially when polygenic scores are applied to individuals who have an ancestry that differs from the original source population.

Introduction

Recent years have seen an explosive growth in our understanding of the genetics of complex traits.[1-3] Large datasets, such as the UK Biobank,[4] have facilitated the genetic analysis of complex traits, and to date over 6000 genome-wide association studies (GWAS) have yielded over 425,000 associations.[5] Despite these discoveries, major knowledge gaps still exist.[6] How can the field go beyond mere catalogs of statistical associations? One way that GWAS results can be leveraged is to infer the genetic architecture of complex traits. A second downstream application of GWAS involves generating polygenic predictions of complex traits and hereditary disease risks.

Genetic architecture refers to the distribution of allelic effects, their interactions, and how segregating genetic variation contributes to differences in traits across individuals.[7] The genetic architectures of complex traits can be described by lists of trait-associated loci, their frequencies, and allele-specific effect sizes. The relative importance of genetic and environmental effects varies by trait, and this can be quantified via estimates of SNP heritability (h^2_{SNP}).[8; 9] An additional aspect of genetic architecture involves the extent to which trait-associated SNPs are found in high or low recombination regions of the genome. These differences are particularly relevant to traits which are due to a small (i.e., < 100) number of genes, since patterns of linkage disequilibrium impact the ability of SNPs to tag causal variants and for GWAS findings to replicate across populations.[10] Traits can also vary in their polygenicity.[11; 12] On one extreme are traits which have a Mendelian genetic architecture, like cystic fibrosis.[13] On the other extreme are highly polygenic traits like height.[14] Importantly, polygenicity can act as a confounding factor in GWAS analyses.[15] Recently, an omnigenic model has been proposed - whereby traits have a set of core genes, but most of the heritability can be explained via indirect effects that are due to gene regulatory networks.[16] Consistent with the omnigenic model, causal variants for anthropometric and blood pressure traits have been found throughout the human genome.[17] Despite an awareness of the

multiple ways that traits can differ, many aspects of genetic architecture have yet to be quantified in a comprehensive way.

GWAS findings can be used to generate polygenic scores (PGS), called polygenic risk scores in the context of hereditary diseases.[18] These scores enable traits to be predicted from genetic information, and they are commonly calculated by summing allele doses across all trait-associated loci and weighting by effect sizes.[19-21] Although the clinical utility of risk scores has received a significant amount of attention during the past few years, many PGS yield only modest case prediction accuracy and show weak correlations between predicted and actual trait values ($R^2 < 0.1$).[22-24] Even within the same ancestry, PGS accuracy can vary due to socio-economic status and genotype-environment interactions.[25; 26] These issues are even more pronounced when genetic predictions are applied to populations that have different ancestries than the original discovery population.[27-29] For example, predictions of anthropomorphic and blood-related traits generated from UK Biobank data perform better when applied to British individuals than Japanese individuals, while predictions generated from Biobank Japan data perform better when applied to Japanese individuals than British individuals.[23] In a landmark study of over 200 traits from the UK Biobank, Privé et al. found that the portability of genetic predictions is reduced in proportion to the genetic distance from the original discovery population.[30] In addition, the predicted values of complex traits can vary between populations. These shifts in PGS distributions can either be due to ascertainment bias[31-33] or due to actual differences in traits.[34; 35] Although thousands of PGS have been generated to date,[24] multiple knowledge gaps exist: Are there particular types of traits that are hard to predict from genetic data? Which traits have PGS that differ the most across populations?

Here, we leveraged polygenic score weights of 177 traits from the UK Biobank to quantify how different aspects of genetic architecture affect PGS performance. We first identified pairs of traits that are influenced by shared genomic regions. We then compared and contrasted three

aspects of genetic architecture (SNP heritability, trait-specific recombination rates, and polygenicity) with three aspects of PGS performance (correlations between predicted and actual trait values, portability of genetic predictions, and how much PGS distributions diverge across populations). Finally, we identified suites of traits that have similar genetic architectures and PGS performance. Notably, binary as well as lifestyle and psychological traits were difficult to predict with genetic data, while also having PGS that generalize poorly across populations.

Materials and methods

PGS weights for 177 complex traits

Our paper builds upon the PGS weights previously generated by Privé et al.[30] To our knowledge, this prior study contains the largest number of traits with multi-ancestry PGS performance metrics. After correcting for sex, age, deprivation index, and 16 principal components (PCs), Privé et al. used lasso penalized regression[36] to generate PGS from 391,124 British-ancestry individuals of European descent from the UK Biobank.[30] We restricted our analyses to 177 traits that had publicly available PGS weights, SNP heritability, and PGS accuracy statistics. We also required that binary traits have a prevalence above 1% in the UK Biobank. These traits were grouped into four categories: 55 biological measures (including blood phenotypes), 50 diseases, 24 lifestyle/psychological traits, and 48 physical measures (such as height and weight). A total of 68 of binary and 109 quantitative traits were analyzed here. A full list of traits, as well as summary statistics of genetic architecture and PGS performance, can be found in Supplemental Table S1 of our paper.

Genetic variance contributions

We leveraged effect size and allele frequency information to identify the most important trait-associated SNPs. Specifically, alleles with large effect sizes and minor allele frequencies close to 50% contribute more to the heritable variation of a trait than alleles with small effect sizes and minor allele frequencies close to zero. As described in earlier work,[37; 38] the contribution of a SNP to the total genetic variance of a trait under Hardy-Weinberg equilibrium and an additive polygenic model is given by:

$$gvc = 2 \times \beta^2 \times p \times (1 - p)$$

where gvc refers to the genetic variance contribution, β is effect size per allele copy, and p is the frequency of the reference allele in a given population. We recreated eight of the ancestry groups found in the UK Biobank using the procedure described in Note A of Privé et al.[30]: UK (i.e., British), Poland, Italy, Iran, India, China, Caribbean, and Nigeria. This data was them used to calculate ancestry-specific allele frequencies and gvc values for each trait-associated SNP.

Traits with shared genetic architectures

To enable comparisons between the genetic architectures of different traits, we divided the human genome into non-overlapping 100kb bins and summed the gvc for every SNP in each bin. We opted for a bin-based approach because pleiotropic loci can have different lead SNPs for different traits, and these lead SNPs need not be in linkage disequilibrium. We forced-ranked 100kb bins to find the top 100 genomic bins with the largest summed gvc for the UK ancestry group, i.e., the top 100 genomic gins for each trait. We then generated a 177x177 similarity matrix, where each element is the count of the number of top 100 bins shared by a pair of traits. This similarity matrix was then used to generate undirected network graphs using the *igraph* package in R. Graphs were generated for multiple thresholds (5, 10, 15, and 20 overlapping bins). Argo Lite[39] was used

135 to generate an interactive visualization of the graph resulting from a threshold of 10 overlapping
136 bins.
137

138 **Quantifying three aspects of genetic architecture: heritability, recombination rates, and**
139 **polygenicity**

140 SNP heritability (h^2_{SNP}) estimates for each trait were previously calculated by Privé et al.[30] using
141 LDpred2-auto.[40]

142 We used the high-resolution recombination map from deCODE[41] to calculate trait-specific
143 recombination rates. *Liftover* was used to convert this recombination map from GRCh38 to
144 GRCh37 coordinates, as PGS weights used the GRCh37 build. For each trait, we focused on
145 SNPs located in the 100 bins with the largest summed gvc for the UK ancestry group. We
146 then calculated the local arithmetic mean recombination rate for 100kb genomic windows
147 centered around each trait-associated SNP (50kb to either side). We then weighted each
148 SNP by gvc to calculate the arithmetic mean recombination rate for each trait, which we
149 denote as R (units: cM per Mb).

150 We quantified whether traits have Mendelian or polygenic architectures using a novel
151 application of Gini coefficients. These coefficients have typically been used in economics to
152 calculate wealth or income inequality,[42] and they range between zero (maximum equality) and
153 one (maximum inequality). Here, we used Gini coefficients to quantify the extent that summed
154 gvc is evenly distributed among the top 100 genomic bins for each trait. Focusing on the top 100
155 genomic bins maximizes the dynamic range of Gini coefficients for the 177 traits analyzed in this
156 study (see Supplemental Information). For traits that had fewer than 100 bins with significant
157 SNPs, the remaining bins were padded with values of zero so that Gini coefficient calculations

always consisted of 100 elements. Gini coefficients were calculated using the following equation, which requires that bins are sorted by gvc in ascending order:

$$G_{100,UK} = \frac{\sum_{i=1}^{n}(2i - n - 1) \times gvc_{i,UK}}{n \times \sum_{i=1}^{n} gvc_{i,UK}}$$

where n = 100 and $gvc_{i,UK}$ is the summed genetic variance contribution of the ith bin using allele frequencies from the UK ancestry group. Using https://github.com/oliviaguest/gini as a guide, we implemented a computationally efficient R script that uses the above equation to compute Gini coefficients for each trait. An examination of the robustness of Gini coefficients to the source of PGS weights, different numbers of 100kb bins, and choice of ancestry group can be found in the Supplemental Information.

Quantifying three aspects PGS performance: accuracy, portability, and divergence

Here, PGS accuracy refers to how well genetic predictions work when individuals are ancestry-matched to the original training set (i.e., British individuals from the UK). For each trait, partial correlations between predicted and actual trait values for Privé et al.'s UK ancestry group (ρ_{UK}) were used to quantify PGS accuracy.[30] These partial correlations were generated using the residuals of actual trait values vs. PGS after correcting for the following covariates: age, sex, birth date, deprivation index, and population structure (16 PCs).[30] This statistic was calculated using genome-wide PGS weights.

Partial correlations between predicted and actual trait values were obtained for the seven other ancestries,[30] which were then used to derive a portability index that quantifies how PGS accuracy diminishes with increased genetic distance from the original training population. For each trait and ancestry group, PGS accuracy relative to the UK ancestry group was found by dividing the partial correlation between predicted and actual trait values for each ancestry group by the partial correlation for the UK ancestry group. Similarly, as per Privé et al.,[30] the geometric

mean position of each ancestry group in 16-dimensional PC space was used to obtain the Euclidean genetic distance between each ancestry group and the UK ancestry. For each trait, relative PGS accuracy was plotted against genetic distance to the UK ancestry group, and linear regression was used to quantify PGS portability. We define the slope of the regression line for each trait (m) as the portability index of that trait. Regression lines were required to pass through the UK datapoint (0,1). Noisy PGS accuracy statistics can cause some traits to have slopes above 0. When this occurred, we manually set m = 0, i.e., perfect portability. Because estimates of m are noisy for individual traits, we restricted our analyses of trait portability to sets of traits.

We also developed a summary statistic that quantifies how much PGS distributions have diverged across populations. For computational efficiency, SNPs in the top 100 genomic bins with the largest summed gvc were used to calculate PGS distributions for each of the eight UK Biobank ancestries described above. Plink 1.9 was used to convert genetic data into an R-readable matrix of the number of effect alleles. Numbers of individuals for each ancestry group in the UK Biobank were down-sampled to 1,234 - the smallest number of samples in any one ancestry group. PGS scores were calculated by summing across all L trait-associated SNPs in the top 100 genomic bins for each trait, and weighting by allele dose ($d_{j,k}$) and effect size (B_k):

$$PGS_j = \sum_{k=1}^{L} d_{j,k}\beta_k$$

where j indexes each individual and k indexes each trait-associated SNP. We then generated PGS distributions for each ancestry group and all 177 traits. Because PGS are calculated by summing the effects of multiple independent SNPs, these distributions tend to normally distributed. We then log-transformed the F-statistic from a one-way ANOVA to derive a metric (D) which quantifies population-level shifts in PGS distributions for each trait:

$$D = log_{10}F$$

where F refers to the ratio of between-ancestry group variability to within-ancestry group variability. We note that meaningful comparisons of D statistics for different traits require that same number of individuals from each ancestry group were analyzed for each trait, as was the case in our study design.

Comparisons between different aspects of genetic architecture and PGS performance

Our manuscript focuses on three summary statistics of genetic architecture (h^2_{SNP}, R, $G_{100,UK}$) and three summary statistics of PGS performance (ρ_{UK}, m, and D) for each trait. To compare different aspects of genetic architecture and PGS performance, we obtained a linear best fit for each pair of summary statistics, generating correlation coefficients and p-values. Because 15 pairwise comparisons were made, a False Discovery Rate (FDR)[43] adjustment was applied to each p-value using the p.adjust() command in R. Note that the Benjamini-Hochberg FDR procedure can cause p-values to clump.

We also assessed how distributions of six summary statistics (h^2_{SNP}, R, $G_{100,UK}$, ρ_{UK}, m, and D) vary for different types of traits. First, we compared the summary statistic distributions of binary traits with the distributions of quantitative traits. We then compared the summary statistic distributions of quantitative lifestyle/psychological traits with the distributions of other quantitative traits. Wilcoxon rank sum tests[44] were used for these summary statistic comparisons, and FDR-adjusted p-values were used to correct for multiple comparisons. For binary traits, we also compared summary statistics of genetic architecture and PGS performance to the $\log_{10}$ prevalence of each trait in the UK Biobank, generating Pearson's correlation coefficients and FDR-adjusted p-values for each comparison.

Principal component analysis (PCA) was used to identify traits with similar summary statistics. Specifically, we applied the prcomp() function in R on a 177x6 array containing genetic

architecture and PGS performance summary statistics for all traits. This approach was repeated for all quantitative traits using a 109x6 array of summary statistics. PCA plots used a modified version of the *ggbiplot* package in R, with 68% probability ellipses (+/- one standard deviation) shown for different types of traits.

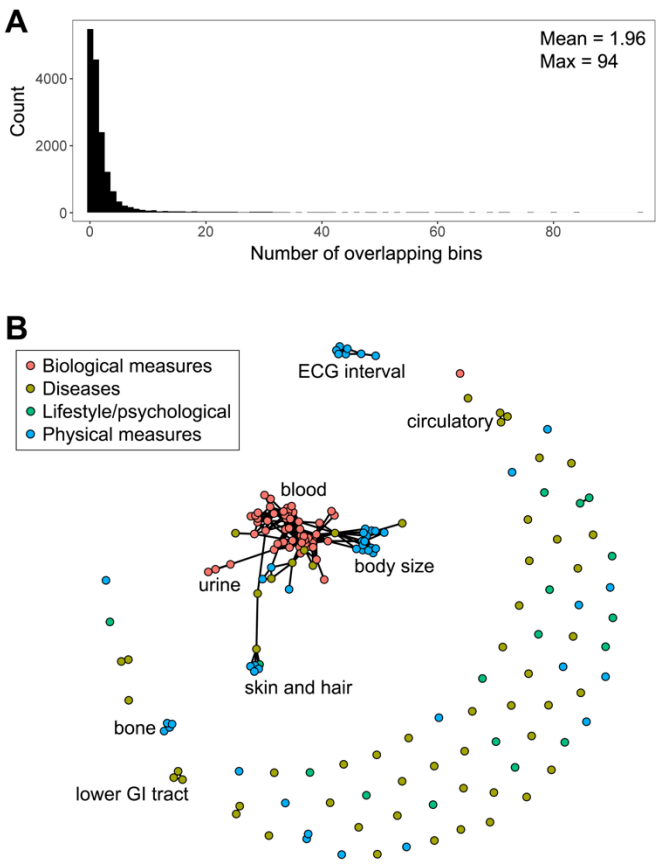

Figure 1. Constellations of traits with shared genetic architectures. Here, we focused on the top 100 genomic bins for each trait. (A) Histogram of the number of 100kb bins shared between all possible pairs of traits. (B) Networks of traits with shared genetic architectures. Each node is a different trait, and coloring indicates trait group. Edges link pairs of traits that have at least 10 overlapping bins out of the top 100 bins for each trait, as inferred by summed gvc. Distances between separate networks are arbitrary.

Results

Overlap between the genetic architectures of different traits

Which traits have heritable variation in the same regions of the genome? To answer this question, we divided the genome into 100kb bins and examined whether the top 100 bins for each trait

257 overlapped with other traits. In general, most pairs of traits share a small number of 100kb bins
258 (1.96 mean shared bins in Figure 1A). Overlap tended to be greater for pairs of traits in the same
259 trait group (Figure S1). For example, traits associated with biological measures shared a mean of
260 7.03 bins. The traits with the most overlap between one another were two body weight traits (log
261 water mass and log fat free mass, sharing 94 of their top 100 bins). The 100kb bin associated
262 with the most traits was chr2:27,700,001-27,800,000 (hg19) with overlap in 48 of the 177 traits.
263 This genomic region includes *GCKR* (Glucokinase Regulator), a gene integral to glucose
264 metabolism and abundantly expressed in the liver.[45]

265 Overlapping 100kb bins reveal networks of connected traits. Here, we focused on pairs of
266 traits that overlap in at least 10 of their top 100 genomic bins (Figure 1B). Networks generated
267 from other overlap thresholds can be seen in Figure S2. An interactive data visualization that
268 includes the names of all traits can be found at: https://tinyurl.com/y2pmpx8a. In Figure 1B, the
269 network with the greatest number of nodes contains 86 traits. Subnetworks within this large
270 structure correspond to urine, blood pressure, erythrocyte, platelet, leukocyte, inflammation, skin
271 and hair, body size, and blood pressure traits. Interestingly, the nodes at which these subnetworks
272 connect to the central mass are biologically intuitive. For instance, hypertension, diastolic blood
273 pressure and systolic blood pressure have deep interconnectedness with edges extending to
274 coronary atherosclerosis, which in turn, connects to blood metabolite traits like cholesterol,
275 apolipoprotein B, and lipoprotein A. Figure 1B also reveals smaller networks of interconnected
276 traits. For example, diseases of the lower GI tract cluster together, as do traits associated with
277 ECG intervals. We note that our bin-based approach was able to reconstitute similar clusters of
278 correlated traits as found via cross-trait LD-Score regression[46] and linear mixed models.[47]
279 However, while trait networks reveal the shared genetic architecture between traits, we note that
280 they do not capture other aspects of the genetic architecture of complex traits, such as heritability,
281 polygenicity, and local recombination rates.

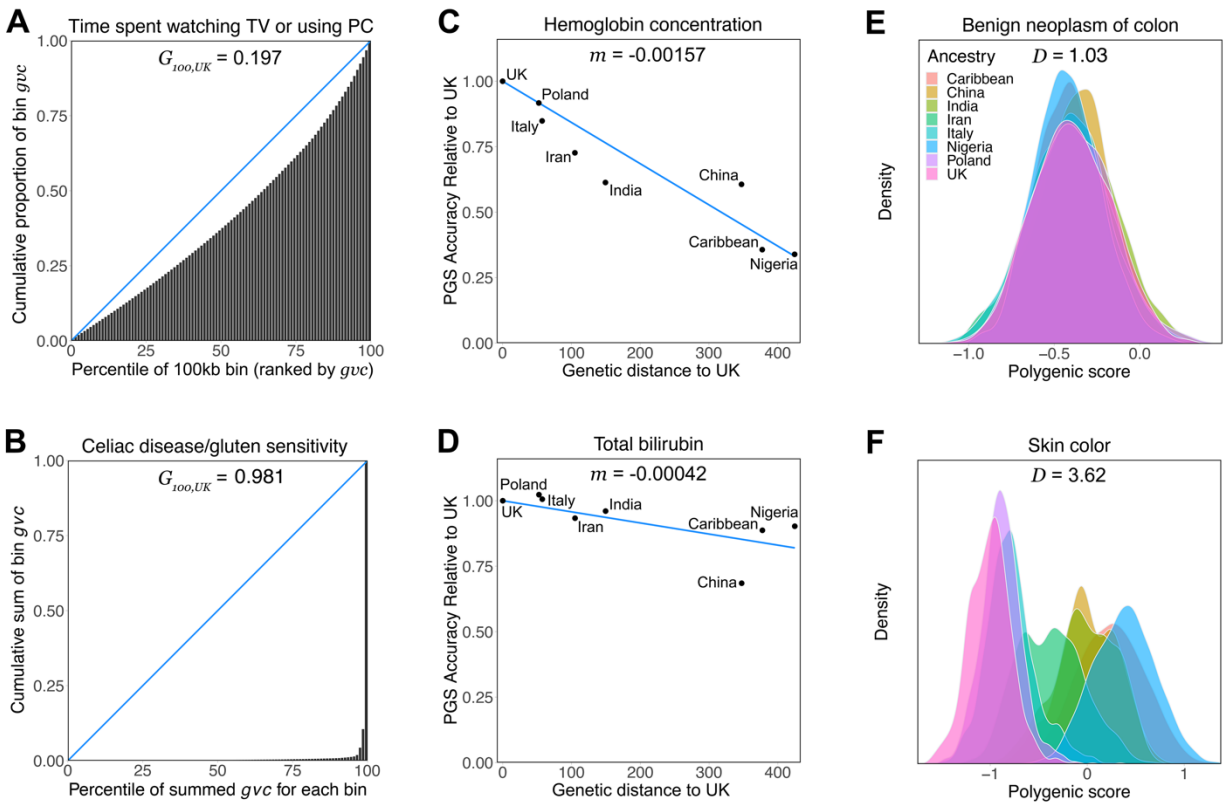

Figure 2. Examples of trait-level differences in polygenicity, portability, and divergence. Gini coefficients, represented here by $G_{100,UK}$, quantify the relative polygenicity of different traits. Lorenz curves of the top 100 genomic bins for two traits are shown here. (A) Time spent watching television or using a computer has a highly polygenic architecture. (B) Celiac disease/gluten sensitivity has a more Mendelian architecture. Portability statistics, represented here by m, quantify how well genetic predictions generalize to different ancestry groups. (C) Predictions of hemoglobin concentration have low portability. (D) Predictions of total bilirubin have high portability. Divergence statistics, represented here by D, quantify the extent that PGS distributions differ across ancestry groups. (E) PGS distributions for benign neoplasms of the colon have a high amount of overlap. (F) PGS distributions for skin color are highly divergent.

Novel metrics of genetic architecture and polygenic score performance

We used a novel application of Gini coefficients to quantify the polygenicity of complex traits. A low Gini coefficient indicates that the genetic variance of a trait tends to be equally distributed

among different SNPs, with fully polygenic traits having Gini coefficients of zero. By contrast, a high Gini coefficient indicates that a large portion of a trait's genetic variance is explained by a small set of SNPs, with fully Mendelian traits having Gini coefficients of one. Across the 177 traits analyzed here, Gini coefficients ($G_{100,\mathrm{UK}}$) ranged from 0.149 to 0.988. Traits with the lowest and highest Gini coefficients are shown in Table 1. Lorenz curves can be used to visualize how Gini coefficients summarize differences in gvc distributions for complex traits. An example of a more polygenic trait is "time spent watching television or using the computer" ($G_{100,\mathrm{UK}}$ = 0.197, Figure 2A). An example of a more Mendelian trait is celiac disease/gluten sensitivity ($G_{100,\mathrm{UK}}$ = 0.981, Figure 2B). We note that traits with large Gini-coefficients also tend to have a larger proportion of their summed gvc in the top 100 bins (Figure S3). In addition, the rank orders of Gini coefficients were largely robust to: the method used to obtain PGS weights (Figure S4), whether Gini coefficients were generated from the top SNPs or the top genomic bins (Figure S5), the number of genomic bins that were examined (Figure S6), and choice of ancestry group (Figure S7); see Supplemental Information for additional details.

Another aspect of PGS performance is the portability of results across different populations. For each trait, we plotted the relative PGS accuracy for different ancestry groups, applied a linear model, and used the slope (m) to quantify the portability of genetic predictions. If m = 0, genetic predictions work equally well for each ancestry group. By contrast, strongly negative slopes ($m < 0$) indicate increasingly poor predictive power relative to the UK ancestry group. Here, we use two examples to illustrate how genetic predictions of complex traits can differ in their portability. Hemoglobin concentration has a low portability statistic (m = -0.00157, Figure 2C). By contrast, total bilirubin concentration has a high portability statistic (m = -0.00042, Figure 2D). For the 177 traits analyzed in this study, m ranged from -0.00489 to 0 (see Table S1 for a full list).

Our paper also includes a novel metric of PGS divergence across populations (D). This metric was calculated by examining PGS distributions for eight different ancestries and converting F statistics from a one-way ANOVA to a $\log_{10}$ scale. D statistics near zero arise when ancestry-specific PGS distributions have similar means; higher values of D statistics arise when ancestry-specific PGS distributions have different means. Because of this, D statistics can be used to quantify ancestry-specific shifts in PGS distributions. Lists of the most and least divergent traits are shown in Table 2 and Table S1. An example of a trait with minimal PGS divergence is benign neoplasm of the colon (D = 1.03, Figure 2E). An example of a trait with substantial divergence between ancestries is skin color (D = 3.62, Figure 2F). Values of D statistics range between 1.03 to 3.62 for the 177 traits analyzed here.

Table 1. Traits with the lowest and highest Gini coefficients

Rank	Trait	$G_{100,UK}$
1	Overall health rating	0.149
2	Neuroticism score	0.168
3	Ever smoked	0.179
4	Plays computer games	0.187
5	Time spent watching television or using computer	0.197
...	...	...
173	Superficial cellulitis and abscess	0.951
174	Hair color (natural before greying)	0.959
175	Other biliary tract disease	0.977
176	Diagnosed with celiac disease or gluten sensitivity	0.981
177	Allergy/adverse effect of penicillin	0.988

Higher values of $G_{100,UK}$ are indicative of traits with more mendelian architectures.

Table 2. Traits with the lowest and highest divergence statistics

Rank	Trait	D
1	Benign neoplasm of the colon	1.03
2	Urinary calculus	1.21
3	Sodium in urine	1.22
4	Myopia diagnosis	1.24
5	High light scatter reticulocyte count	1.26
...	...	...
173	Albumin	3.12
174	Use of sun/UV protection	3.28
175	Hair color (natural, before greying)	3.51
176	Ease of skin tanning	3.57
177	Skin color	3.62

Higher values of D are indicative of traits that have PGS distributions that differ more across between groups.

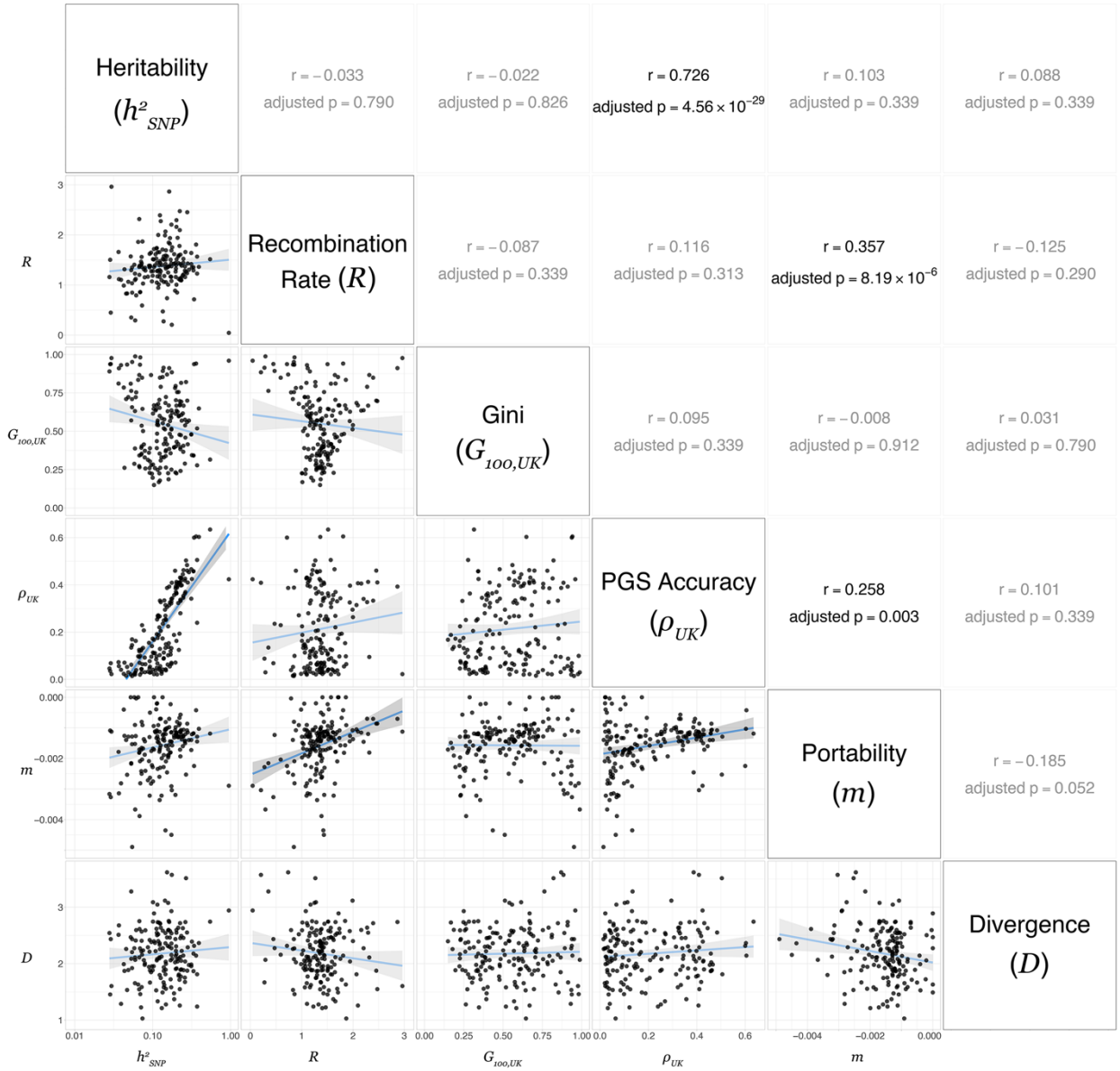

Figure 3. Comparisons between different aspects of genetic architecture and PGS performance. Each scatterplot corresponds to a different pair of summary statistics, and each datapoint corresponds to a different complex trait. All comparisons and p-values use FDR-adjusted Wilcoxon rank sum tests. Statistically significant results are shown in bold. Summary statistics for all 177 traits can be found in Supplemental Table S1.

352 **Comparisons between different aspects of genetic architecture**

353 SNP heritabilities, Gini coefficients, and trait-specific recombination rates capture different
354 aspects of genetic architecture (upper left part of Figure 3). We found that highly heritable traits
355 were not any more or less likely to have SNPs in high recombination regions of the genome (r =
356 -0.033, p = 0.790). Similarly, h^2_{SNP} was largely independent of polygenicity, as quantified by $G_{100,UK}$
357 (r = -0.022, p = 0.826). Although there was no linear relationship between trait-specific
358 recombination rates and Gini coefficients (r = -0.087, p = 0.339), we note that a plot of R vs $G_{100,UK}$
359 yields a distinctive triangular pattern (Figure 3). This pattern arises because highly polygenic traits
360 tend to have recombination rates that resemble the genome-wide mean, since those traits'
361 recombination rates are averaged across the entire genome, while traits with more Mendelian
362 architectures can potentially be enriched for SNPs in either high or low recombination regions of
363 the genome. For example, genetic variants associated with Apolipoprotein B tend to be found in
364 high recombination regions of the genome (R = 2.87 cM/Mb), and genetic variants associated
365 with allergy/adverse effect of penicillin tend to be found in low recombination regions (R = 0.293
366 cM/Mb). The low correlations seen here underscore the fact that no single metric can summarize
367 the genetic architecture, while each metric yields important insights in understanding the genetics
368 of complex traits.

369

370 **Comparisons between different aspects of polygenic score performance**

371 Complex traits vary in multiple aspects of PGS performance: accuracy, portability, and divergence
372 (lower right part of Figure 3). We found that PGS accuracy and portability were positively
373 correlated (r = 0.258, p = 0.003). An example of a trait with high PGS accuracy and high portability
374 is mean platelet volume (ρ_{UK} = 0.603, m = -00103). An example of a trait with low PGS accuracy
375 and low portability is general happiness (ρ_{UK} = 0.0697, m = -0.00388). Portability statistics are

noisier for traits that are hard to predict in British individuals from the UK. Because of this, values of m are highly variable when ρ_{UK} is close to zero. We note that divergence and portability statistics are negatively correlated, although this trend does not quite reach statistical significance (r = -0.185, p = 0.052). This pattern can arise from a combination of ascertainment bias and allele frequency differences between populations.[31] There were no broad trends when trait divergence was compared to PGS accuracy (r = 0.101, p = 0.339). Collectively, these results indicate that PGS performance is largely trait-specific.

Relevance of genetic architecture to polygenic score performance

How do SNP heritabilities, trait-specific recombination rates, and polygenicity relate to PGS accuracy, portability and divergence? Although not all combinations of genetic architecture and PGS performance yielded clear associations, a few notable patterns can be seen in Figure 3. As expected, traits with a high SNP heritability tended to have a high PGS accuracy (r = 0.726, p = 4.56×10^{-29}). We note that heritability refers to the proportion of phenotypic variance that is due to genetic effects in a single population (i.e., it is a population-specific concept). This suggests that h^2_{SNP} estimates may not be that informative about how well predictions generalize across populations. Indeed, SNP heritabilities were largely uninformative about the portability of polygenic predictions (r = 0.103, p = 0.339). SNP heritabilities were also non-informative about the divergence of predicted trait values (r = 0.088, p = 0.339). By contrast, traits with SNPs in high recombination regions of the genome tended to have genetic predictions that were more portable across populations, i.e., there was a positive correlation between trait-specific recombination rates and portability (r = 0.357, p = 8.19×10^{-6}). However, trait-specific recombination rates were largely uninformative with respect to PGS accuracy (r = 0.116, p = 0.313) and divergence (r = -0.125, p = 0.290). Gini coefficients were largely uncorrelated with PGS accuracy, portability, and trait divergence (Figure 3). This means that traits with more Mendelian, i.e., simpler, genetic

architectures need not be easier to predict. An additional complication is that rare binary traits tend to have large $G_{100,UK}$ statistics, while lifestyle and psychological traits tend to have small $G_{100,UK}$ statistics (see below). Rare binary traits are difficult to predict, as are lifestyle/psychological traits, which helps explain why Gini coefficients and PGS performance appear to be independent in our study.

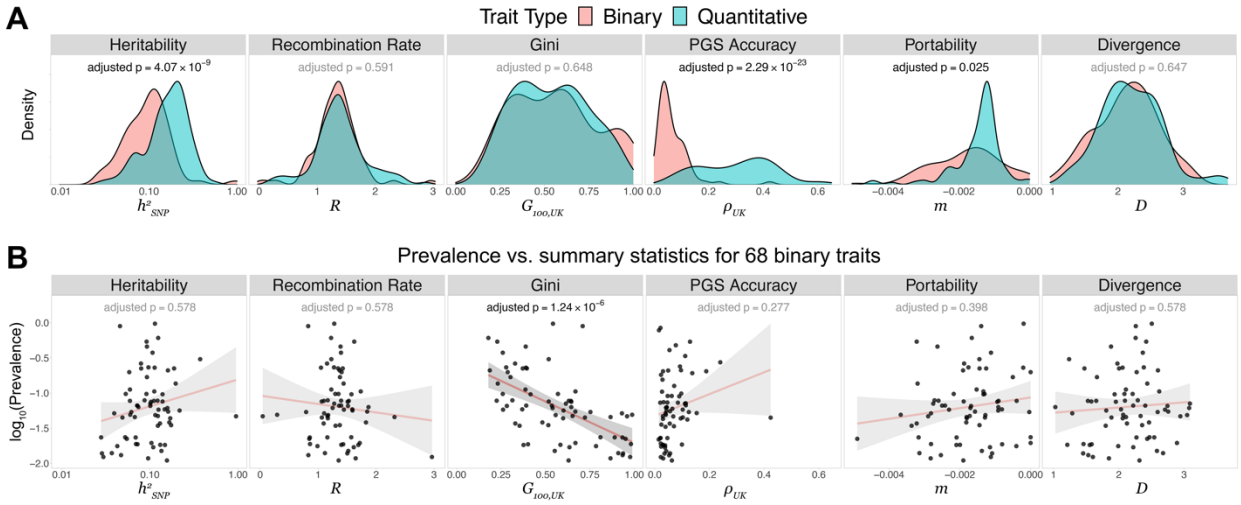

Figure 4. Characteristics of binary traits. All p-values use FDR-adjusted Wilcoxon rank sum tests. Statistically significant results are shown in bold. (A) Distributions of summary statistics for binary vs. quantitative traits. (B) Prevalence of binary traits in the UK Biobank vs. six different summary statistics of genetic architecture and PGS performance.

Summary statistics differ for binary and quantitative traits

One natural way to classify traits is whether they are binary or quantitative (the latter including both ordinal and continuous traits). We note that all the disease traits analyzed in our study are binary. In addition, the statistical power to detect genetic associations differs for binary and quantitative traits.[48] Both of these details can affect estimates of genetic architecture and PGS performance. Indeed, we observed differences in the summary statistics of binary and quantitative

traits (Figure 4A). On average, binary traits tended to have a lower SNP heritability than quantitative traits ($p = 4.07 \times 10^{-9}$). However, binary traits were not any more likely than quantitative traits to have Mendelian architectures ($p = 0.648$ for $G_{100,UK}$ comparisons). This pattern can occur if binary traits arise from polygenic effects that exceed a threshold.[49] Focusing on different aspects of PGS performance, we found that PGS accuracy was much lower for binary traits than quantitative traits ($p = 2.23 \times 10^{-23}$). For example, partial correlations between predicted and actual trait values were lower for hypertension ($\rho_{UK} = 0.188$) than for systolic blood pressure ($\rho_{UK} = 0.255$). Binary traits also had a wider range of portability statistics than quantitative traits (Figure 4A). Finally, we note that divergence statistics were similar for binary and quantitative traits. A PCA plot generated from six summary statistics of genetic architecture and PGS performance also demonstrates that binary traits have different profiles than quantitative traits (Figure S8).

To further explore the genetic architectures and PGS performance of binary traits, we examined whether rare or common traits have different properties than quantitative traits. We plotted each summary statistic against the $\log_{10}$-transformed prevalence of binary traits within the UK Biobank dataset (Figure 4B). For five of these summary statistics there was no linear trend. By contrast, we found that Gini coefficients were negative correlated with prevalence ($r = -0.581$, $p = 1.24 \times 10^{-6}$), indicating that rarer complex traits appear to have more Mendelian architectures. This pattern aligns with previously published findings that rare diseases tend to have a Mendelian architecture.[50; 51] The low $G_{100,UK}$ statistics of rare traits may be due to limited statistical power to identify trait-associated alleles.[52] However, we note that ρ_{UK} statistics are largely independent of the prevalence of binary traits (Figure 4B), i.e., PGS accuracy is not appreciably different for rare or common traits. This suggests that the relationship between Gini coefficients and prevalence is due to the underlying biology, rather than a statistical artifact.

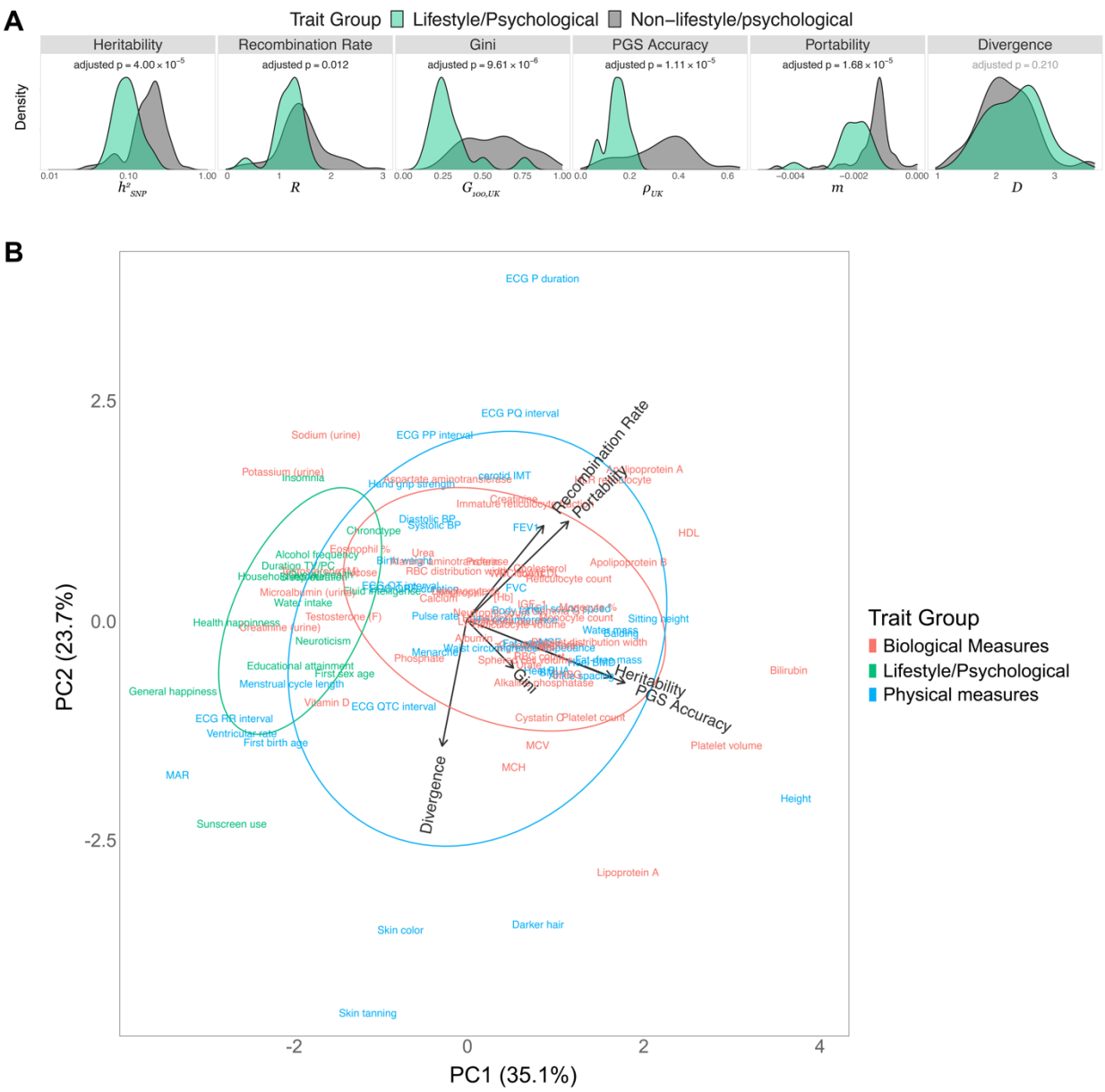

Figure 5. Lifestyle and psychological traits have different genetic profiles than other quantitative traits. (A) Distributions of summary statistics for quantitative lifestyle and psychological traits vs. other quantitative traits. (B) PCA plot generated from six summary statistics of genetic architecture and PGS performance (h^2_{SNP}, R, $G_{100,UK}$, ρ_{UK}, m, and D). Arrows indicate higher values of each summary statistic in PCA space. All traits shown here are quantitative, and trait names are colored by trait group. Ellipses indicate 68% (+/- one standard deviation) confidence intervals.

452 **Limitations of polygenic scores for lifestyle and psychological traits**

453 We also examined whether summary statistics of lifestyle and psychological traits differed from
454 other quantitative traits. Examples of these complex behavioral traits include alcohol frequency,
455 chronotype, educational attainment, happiness, neuroticism, and water intake. Overall, there are
456 noticeable differences in the genetic architecture and PGS performance of different types of
457 quantitative traits (Figure 5A). On average, lifestyle and psychological traits have lower SNP
458 heritabilities than other quantitative traits ($p = 4.00 \times 10^{-5}$). This pattern is consistent with the
459 importance of environmental effects for behavioral traits. Low heritability also has a knock-on
460 effect of reducing the effectiveness of polygenic predictions. Indeed, we found that PGS accuracy
461 was much lower for lifestyle and psychological traits ($p = 1.11 \times 10^{-5}$). Furthermore, we note that
462 there was a clear lack of lifestyle and psychological traits with high ρ_{UK} statistics (Figure 5A).
463 Genetic predictions of lifestyle and psychological traits were also less portable than other
464 quantitative traits ($p = 1.68 \times 10^{-5}$). We note that lifestyle and psychological traits also tended to
465 have low $G_{100,UK}$ statistics, a result which can only occur if trait-associated loci have been found
466 in many parts of the genome. This suggest that the poor PGS performance of these traits is not
467 simply a byproduct of a lack of GWAS hits. Other comparisons yield only modest differences:
468 lifestyle and psychological traits were slightly more likely to have SNPs in low recombination
469 regions of the genome ($p = 0.012$), and no appreciable differences were observed when we
470 compared D statistics for different types of quantitative traits. Taken together, Figure 5A reveals
471 that genetic predictions of lifestyle and psychological traits are severely limited given our present
472 knowledge of the genetic basis of these traits.

473 A PCA plot generated from six summary statistics of genetic architecture and PGS
474 performance further demonstrates that lifestyle and psychological traits have different profiles
475 than other quantitative traits (Figure 5B). Arrows in this plot indicate regions of PCA space that
476 are associated with higher values of each summary statistic, recapitulating our earlier findings:

477 higher heritability is associated with higher PGS accuracy, and traits that are due to SNPs in high
478 recombination regions of the genome tend to have more portable predictions. We also note that
479 the arrows for portability and divergence point in different directions. Lifestyle and psychological
480 traits form a noticeable cluster in the small region of PCA space pertaining to lower Gini,
481 heritability, prediction, and portability statistics. By contrast, biological measures and physical
482 measures occupy larger and more centralized regions of PCA space, which reflects that these
483 two trait groups are more varied in their genetic architecture and PGS performance. An outlier
484 among lifestyle and psychological traits is sunscreen use, most likely due to its overlap with
485 physical measures like skin color and skin tanning.

486

487

488 ## Discussion

489 Overall, we found that complex traits have a broad range of genetic architectures, which
490 contributes to differences in PGS performance. Our results indicate that highly heritable traits are
491 easier to predict when individuals are ancestry-matched to the original GWAS cohort – a finding
492 that consistent with expectations from statistical genetics.[53] However, SNP heritability is largely
493 uninformative when it comes to the portability of genetic predictions. By contrast, we found that
494 traits with SNPs in high recombination regions of the genome tend to have genetic predictions
495 that generalize well across populations. This suggests that linkage disequilibrium and the ability
496 of PGS variants to tag narrow genomic regions may be important for PGS portability.

497 Shifts in PGS distributions are due to allele frequency differences between populations.
498 Because of this, one might expect to find greater divergence for high Gini traits than for low Gini
499 traits. The reasoning here is that allele frequency differences at different SNPs can average out
500 if traits are highly polygenic. However, summary statistics of trait divergence (D) were largely
501 independent of Gini coefficients ($G_{100,UK}$). This suggests that other phenomena like natural

502 selection[54] and ascertainment bias[31] are drivers of ancestry-specific shifts in PGS distributions.
503 Indeed, hair and skin color, which are among the most divergent traits in our study, have
504 previously been implicated in scans of selection.[55] We also note that natural selection can erode
505 the portability of polygenic predictions.[56; 57]

506 Our polygenicity analyses shed light on the omnigenic model. Gini coefficients varied
507 substantially across traits, and many traits had high Gini-coefficients. Furthermore, this pattern
508 was not an artifact of focusing on the top 100 genomic bins (see Figure S6). Our results
509 demonstrated that genetic architectures are often trait-specific, and that core genes can
510 potentially make outsized contributions to SNP heritability. That said, the omnigenic model
511 proposes that genetic effects cascade through cellular regulatory networks, as expression of core
512 genes ends up affecting gene expression at other genes.[16] Because existing GWAS have
513 generally yielded low-hanging fruit,[6] these indirect effects might be one reason why many PGS
514 poorly predict complex traits. It also might explain why transcriptional risk scores can potentially
515 outperform genetic risk scores.[58] When large well-powered GWAS are conducted on diverse
516 cohorts, rare family-specific or ancestry-specific variants are more likely to identified. Genetic
517 associations involving these private alleles are unlikely yield portable predictions. Indeed, a height
518 GWAS of 5.4 million individuals found that SNP heritability clusters in genomes, and that out-of-
519 sample prediction accuracy was lower for individuals who did not have European ancestry.[59]
520 There is also evidence that pruning sets of trait-associated SNPs can lead to improved PGS
521 performance among diverse populations.[60]

522 Finally, we mention that PGS are not immune to controversy – especially when it comes
523 to lifestyle and psychological traits. Some have envisioned a world where PGS for educational
524 attainment might be used inform the allocation of resources to those who have the most need.[61]
525 This has spurred intense debate about both efficacy of polygenic predictions for behavioral traits
526 and whether they should be used in a public policy setting.[62; 63] Others have gone a step further

527 and advocated using PGS to screen embryos for cognitive traits,[64] a position that has received
528 well-warranted criticism.[65; 66] Regardless of the specific trait, there are major challenges to
529 polygenic screening of embryos.[67-69] Polygenic predictions of complex behavioral traits are
530 particularly problematic. As seen in Figure 5A, lifestyle and psychological traits are difficult to
531 predict, which means that any downstream applications of PGS for these traits would be deeply
532 flawed. This issue is particularly acute when PGS are applied to populations that have ancestries
533 that differ from the original GWAS population, given their low portability. Ultimately, genetic
534 predictions of traits like alcohol intake, general happiness, income, or educational attainment in
535 non-European populations should be treated with extreme skepticism; racist claims about the
536 supposed intellectual superiority of any particular ancestry are genetically untenable.
537

538 **Conclusion**

539 We note that the summary statistics examined here are not exhaustive. Going forward, future
540 studies will be able to explore additional aspects of genetic architecture and PGS performance.
541 For example, some traits are highly canalized, while others show evidence of substantial PGS-
542 by-environment interactions.[70] Epistatic interactions also contribute to the genetic architecture of
543 complex traits, and this information can be incorporated into predictive models.[71] Finally, we note
544 that PGS generated from multi-ancestry cohorts are more likely to yield portable predictions.[72]
545 Nevertheless, we still expect there to be a significant limitations to the genetic prediction of
546 complex behavioral traits.

Declaration of interests

The authors declare no competing interests.

Acknowledgements

We thank study participants from the UK Biobank. In addition, we thank Greg Gibson, Sini Nagpal, King Jordan, Aaron Pfennig, Mimi Holness, and other members of the Center for Integrative Genomics at Georgia Institute of Technology for their helpful suggestions and feedback. This work was funded by an NIGMS MIRA grant (R35GM133727) to J.L. This research has been conducted using the UK Biobank Resource under application number 17984.

Author contributions

N.C.: methodology, formal analysis, data curation, visualization, and writing; A.H.: methodology, formal analysis, data curation, visualization, and writing; J.L.: conceptualization, funding acquisition, methodology, supervision, visualization, and writing.

Data and code availability

UK Biobank data can be requested via: https://www.ukbiobank.ac.uk/enable-your-research.

All code used for this paper is available at https://github.com/LachanceLab/gini/. Additional details about how PGS were generated can be found in Privé et al.[28]

Supplemental information

Supplemental information, including eight supplemental figures and one supplemental table, can be found online at URL.

Web resources

UK Biobank, https://www.ukbiobank.ac.uk/

PGS weights and ancestry-specific partial correlations (PGS accuracy statistics) from Privé et al.,[28] https://figshare.com/articles/dataset/Effect_sizes_for_215_polygenic_scores/14074760/2

SNP heritabilities from Privé et al.,[28] https://github.com/privefl/UKBB-PGS/blob/main/phenotype-info.csv

Argo Lite, https://poloclub.github.io/argo-graph-lite/

References

1. 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.
2. Agarwala, V., Flannick, J., Sunyaev, S., Go, T.D.C., and Altshuler, D. (2013). Evaluating empirical bounds on complex disease genetic architecture. Nat Genet 45, 1418-1427.
3. Watanabe, K., Stringer, S., Frei, O., Umicevic Mirkov, M., de Leeuw, C., Polderman, T.J.C., van der Sluis, S., Andreassen, O.A., Neale, B.M., and Posthuma, D. (2019). A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet 51, 1339-1348.
4. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12, e1001779.
5. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47, D1005-D1012.
6. Tam, V., Patel, N., Turcotte, M., Bosse, Y., Pare, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. Nat Rev Genet 20, 467-484.

599 7. Timpson, N.J., Greenwood, C.M.T., Soranzo, N., Lawson, D.J., and Richards, J.B. (2018).
600 Genetic architecture: the shape of the genetic contribution to human traits and disease.
601 Nat Rev Genet 19, 110-124.
602 8. Yang, J., Zeng, J., Goddard, M.E., Wray, N.R., and Visscher, P.M. (2017). Concepts, estimation
603 and interpretation of SNP-based heritability. Nat Genet 49, 1304-1310.
604 9. Tang, M., Wang, T., and Zhang, X. (2022). A review of SNP heritability estimation methods.
605 Brief Bioinform 23.
606 10. Scheinfeldt, L.B., Schmidlen, T.J., Gerry, N.P., and Christman, M.F. (2016). Challenges in
607 Translating GWAS Results to Clinical Care. Int J Mol Sci 17.
608 11. Lander, E.S., and Schork, N.J. (1994). Genetic dissection of complex traits. Science 265,
609 2037-2048.
610 12. Glazier, A.M., Nadeau, J.H., and Aitman, T.J. (2002). Finding genes that underlie complex
611 traits. science 298, 2345-2349.
612 13. Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online
613 Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic
614 disorders. Nucleic Acids Res 33, D514-517.
615 14. Guo, M.H., Hirschhorn, J.N., and Dauber, A. (2018). Insights and Implications of Genome-
616 Wide Association Studies of Height. J Clin Endocrinol Metab 103, 3155-3168.
617 15. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working
618 Group of the Psychiatric Genomics, C., Patterson, N., Daly, M.J., Price, A.L., and Neale,
619 B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-
620 wide association studies. Nat Genet 47, 291-295.
621 16. Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From
622 Polygenic to Omnigenic. Cell 169, 1177-1186.
623 17. Johnson, R., Burch, K.S., Hou, K., Paciuc, M., Pasaniuc, B., and Sankararaman, S. (2021).
624 Estimation of regional polygenicity from GWAS provides insights into the genetic
625 architecture of complex traits. PLoS Comput Biol 17, e1009483.
626 18. Sugrue, L.P., and Desikan, R.S. (2019). What are polygenic scores and why are they
627 important? Jama 321, 1820-1821.
628 19. Torkamani, A., Wineinger, N.E., and Topol, E.J. (2018). The personal and clinical utility of
629 polygenic risk scores. Nat Rev Genet 19, 581-590.
630 20. Lambert, S.A., Abraham, G., and Inouye, M. (2019). Towards clinical utility of polygenic risk
631 scores. Hum Mol Genet 28, R133-R142.

632 21. Choi, S.W., Mak, T.S., and O'Reilly, P.F. (2020). Tutorial: a guide to performing polygenic risk
633 score analyses. Nat Protoc 15, 2759-2772.
634 22. Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P.,
635 Lander, E.S., Lubitz, S.A., Ellinor, P.T., et al. (2018). Genome-wide polygenic scores for
636 common diseases identify individuals with risk equivalent to monogenic mutations. Nat
637 Genet 50, 1219-1224.
638 23. Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019). Clinical
639 use of current polygenic risk scores may exacerbate health disparities. Nat Genet 51, 584-
640 591.
641 24. Lambert, S.A., Gil, L., Jupp, S., Ritchie, S.C., Xu, Y., Buniello, A., McMahon, A., Abraham,
642 G., Chapman, M., Parkinson, H., et al. (2021). The Polygenic Score Catalog as an open
643 database for reproducibility and systematic evaluation. Nat Genet 53, 420-425.
644 25. Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J.K., and Przeworski, M. (2020).
645 Variable prediction accuracy of polygenic scores within an ancestry group. Elife 9.
646 26. Abdellaoui, A., Dolan, C.V., Verweij, K.J.H., and Nivard, M.G. (2022). Gene-environment
647 correlations across geographic regions affect genome-wide association studies. Nat
648 Genet 54, 1345-1354.
649 27. Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J.,
650 Bustamante, C.D., and Kenny, E.E. (2017). Human Demographic History Impacts Genetic
651 Risk Prediction across Diverse Populations. Am J Hum Genet 100, 635-649.
652 28. Ruan, Y., Lin, Y.-F., Feng, Y.-C.A., Chen, C.-Y., Lam, M., Guo, Z., He, L., Sawa, A., Martin,
653 A.R., and Qin, S. (2022). Improving polygenic prediction in ancestrally diverse populations.
654 Nature Genetics 54, 573-580.
655 29. Cavazos, T.B., and Witte, J.S. (2021). Inclusion of variants discovered from diverse
656 populations improves polygenic risk score transferability. HGG Adv 2.
657 30. Prive, F., Aschard, H., Carmi, S., Folkersen, L., Hoggart, C., O'Reilly, P.F., and Vilhjalmsson,
658 B.J. (2022). Portability of 245 polygenic scores when derived from the UK Biobank and
659 applied to 9 ancestry groups from the same cohort. Am J Hum Genet 109, 12-23.
660 31. Kim, M.S., Patel, K.P., Teng, A.K., Berens, A.J., and Lachance, J. (2018). Genetic disease
661 risks can be misestimated across global populations. Genome Biol 19, 179.
662 32. Novembre, J., and Barton, N.H. (2018). Tread Lightly Interpreting Polygenic Tests of
663 Selection. Genetics 208, 1351-1355.

664 33. Berg, J.J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A.M., Mostafavi, H., Field, Y.,
665 Boyle, E.A., Zhang, X., Racimo, F., Pritchard, J.K., et al. (2019). Reduced signal for
666 polygenic adaptation of height in UK Biobank. Elife 8.
667 34. Lachance, J., Berens, A.J., Hansen, M.E.B., Teng, A.K., Tishkoff, S.A., and Rebbeck, T.R.
668 (2018). Genetic Hitchhiking and Population Bottlenecks Contribute to Prostate Cancer
669 Disparities in Men of African Descent. Cancer Res 78, 2432-2443.
670 35. Ju, D., and Mathieson, I. (2021). The evolution of skin pigmentation-associated variation in
671 West Eurasia. Proc Natl Acad Sci U S A 118.
672 36. Prive, F., Aschard, H., and Blum, M.G.B. (2019). Efficient Implementation of Penalized
673 Regression for Genetic Risk Prediction. Genetics 212, 65-74.
674 37. Park, J.H., Wacholder, S., Gail, M.H., Peters, U., Jacobs, K.B., Chanock, S.J., and Chatterjee,
675 N. (2010). Estimation of effect size distribution from genome-wide association studies and
676 implications for future discoveries. Nat Genet 42, 570-575.
677 38. Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M.,
678 and Stephens, M. (2015). A multivariate genome-wide association analysis of 10 LDL
679 subfractions, and their response to statin treatment, in 1868 Caucasians. PLoS One 10,
680 e0120758.
681 39. Li, S., Zhou, Z., Upadhayay, A., Shaikh, O., Freitas, S., Park, H., Wang, Z.J., Routray, S.,
682 Hull, M., and Chau, D.H. (2020). Argo lite: Open-source interactive graph exploration and
683 visualization in browsers. In Proceedings of the 29th ACM International Conference on
684 Information & Knowledge Management. pp 3071-3076.
685 40. Prive, F., Arbel, J., and Vilhjalmsson, B.J. (2020). LDpred2: better, faster, stronger.
686 Bioinformatics.
687 41. Halldorsson, B.V., Palsson, G., Stefansson, O.A., Jonsson, H., Hardarson, M.T., Eggertsson,
688 H.P., Gunnarsson, B., Oddsson, A., Halldorsson, G.H., and Zink, F. (2019). Characterizing
689 mutagenic effects of recombination through a sequence-level genetic map. Science 363,
690 eaau1043.
691 42. Gini, C. (1912). Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni
692 statistiche.[Fasc. I.].(Tipogr. di P. Cuppini).
693 43. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and
694 powerful approach to multiple testing. Journal of the Royal statistical society: series B
695 (Methodological) 57, 289-300.

696 44. Wilcoxon, F. (1992). Individual comparisons by ranking methods. In Breakthroughs in
697 statistics. (Springer), pp 196-202.
698 45. Irwin, D.M., and Tan, H. (2014). Evolution of glucose utilization: glucokinase and glucokinase
699 regulator protein. Mol Phylogenet Evol 70, 195-203.
700 46. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., ReproGen, C.,
701 Psychiatric Genomics, C., Genetic Consortium for Anorexia Nervosa of the Wellcome
702 Trust Case Control, C., Duncan, L., et al. (2015). An atlas of genetic correlations across
703 human diseases and traits. Nat Genet 47, 1236-1241.
704 47. Canela-Xandri, O., Rawlik, K., and Tenesa, A. (2018). An atlas of genetic associations in UK
705 Biobank. Nat Genet 50, 1593-1599.
706 48. Yang, J., Wray, N.R., and Visscher, P.M. (2010). Comparing apples and oranges: equating
707 the power of case-control and quantitative trait association studies. Genet Epidemiol 34,
708 254-257.
709 49. Falconer, D.S. (1967). The inheritance of liability to diseases with variable age of onset, with
710 particular reference to diabetes mellitus. Ann Hum Genet 31, 1-20.
711 50. Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. Trends Genet
712 17, 502-510.
713 51. Wright, C.F., FitzPatrick, D.R., and Firth, H.V. (2018). Paediatric genomics: diagnosing rare
714 disease in children. Nat Rev Genet 19, 325.
715 52. Hong, E.P., and Park, J.W. (2012). Sample size and statistical power calculation in genetic
716 association studies. Genomics Inform 10, 117-122.
717 53. Ma, Y., and Zhou, X. (2021). Genetic prediction of complex traits with polygenic scores: a
718 statistical review. Trends Genet 37, 995-1011.
719 54. Racimo, F., Berg, J.J., and Pickrell, J.K. (2018). Detecting Polygenic Adaptation in Admixture
720 Graphs. Genetics 208, 1565-1584.
721 55. Quillen, E.E., Norton, H.L., Parra, E.J., Lona-Durazo, F., Ang, K.C., Illiescu, F.M., Pearson,
722 L.N., Shriver, M.D., Lasisi, T., Gokcumen, O., et al. (2019). Shades of complexity: New
723 perspectives on the evolution and genetic architecture of human skin. Am J Phys
724 Anthropol 168 Suppl 67, 4-26.
725 56. Yair, S., and Coop, G. (2022). Population differentiation of polygenic score predictions under
726 stabilizing selection. Philos Trans R Soc Lond B Biol Sci 377, 20200416.
727 57. Durvasula, A., and Lohmueller, K.E. (2021). Negative selection on complex traits limits
728 phenotype prediction accuracy between populations. Am J Hum Genet 108, 620-631.

58. Marigorta, U.M., Denson, L.A., Hyams, J.S., Mondal, K., Prince, J., Walters, T.D., Griffiths, A., Noe, J.D., Crandall, W.V., Rosh, J.R., et al. (2017). Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease. Nat Genet 49, 1517-1521.
59. Yengo, L., Vedantam, S., Marouli, E., Ferreira, T., Karaderi, T., Malden, D., Fairhurst-Hunter, Z., Goel, A., Lin, K., and Liu, J. (2022). A saturated map of common genetic variants associated with human height. Nature.
60. Graham, B.E., Plotkin, B., Muglia, L., Moore, J.H., and Williams, S.M. (2021). Estimating prevalence of human traits among populations from polygenic risk scores. Human genomics 15, 1-16.
61. Harden, K.P. (2021). The genetic lottery: why DNA matters for social equality.(Princeton University Press).
62. Harden, K.P., Patterson, N., Reus, V.I., Schlinger Jr., H.D., Felman, M.W., and Riskin, J. (2022). 'Why Biology is Not Destiny': An Exchange. In The New York Review of Books. (
63. Coop, G., and Przeworski, M. (2022). Lottery, luck, or legacy. A review of "The Genetic Lottery: Why DNA matters for social equality". Evolution 76, 846-853.
64. Hsu, S.D. (2014). On the genetic architecture of intelligence and other quantitative traits. arXiv preprint arXiv:14083421.
65. Kaiser, J. (2019). Screening embryos for complex genetic traits called premature. Science 366, 405-406.
66. Turley, P., Meyer, M.N., Wang, N., Cesarini, D., Hammonds, E., Martin, A.R., Neale, B.M., Rehm, H.L., Wilkins-Haug, L., and Benjamin, D.J. (2021). Problems with using polygenic scores to select embryos. New England Journal of Medicine 385, 78-86.
67. Karavani, E., Zuk, O., Zeevi, D., Barzilai, N., Stefanis, N.C., Hatzimanolis, A., Smyrnis, N., Avramopoulos, D., Kruglyak, L., Atzmon, G., et al. (2019). Screening Human Embryos for Polygenic Traits Has Limited Utility. Cell 179, 1424-1435 e1428.
68. Goldstein, D.B. (2021). The End of Genetics : designing humanity's dna.(New Haven: Yale University Press).
69. Lencz, T., Backenroth, D., Granot-Hershkovitz, E., Green, A., Gettler, K., Cho, J.H., Weissbrod, O., Zuk, O., and Carmi, S. (2021). Utility of polygenic embryo screening for disease depends on the selection strategy. Elife 10.
70. Nagpal, S., Tandon, R., and Gibson, G. (2022). Canalization of the Polygenic Risk for Common Diseases and Traits in the UK Biobank Cohort. Mol Biol Evol 39.

761 71. Moore, J.H., and Williams, S.M. (2009). Epistasis and its implications for personal genetics.
762 Am J Hum Genet 85, 309-320.
763 72. Kullo, I.J., Lewis, C.M., Inouye, M., Martin, A.R., Ripatti, S., and Chatterjee, N. (2022).
764 Polygenic scores in biomedical research. Nat Rev Genet 23, 524-532.

765