PLOS BIOLOGY

RESEARCH ARTICLE

Simple scaling laws control the genetic architectures of human complex traits

Yuval B. Simons[1,2,3*], **Hakhamanesh Mostafavi**[3,4,5], **Huisheng Zhu**[6], **Courtney J. Smith**[3], **Jonathan K. Pritchard**[3,6*], **Guy Sella**[7,8*]

1 Section of Genetic Medicine, University of Chicago, Chicago, Illinois, United States of America, **2** Department of Human Genetics, University of Chicago, Chicago, Illinois, United States of America, **3** Department of Genetics, Stanford University, Stanford, California, United States of America, **4** Center for Human Genetics and Genomics, New York University School of Medicine, New York, New York, United States of America, **5** Department of Population Health, New York University School of Medicine, New York, New York, United States of America, **6** Department of Biology, Stanford University, Stanford, California, United States of America, **7** Department of Biological Sciences, Columbia University, New York, New York, United States of America, **8** Program for Mathematical Genomics, Columbia University, New York, New York, United States of America

* yuval.simons@uchicago.edu (YBS); pritch@stanford.edu (JKP); gs2747@columbia.edu (GS)

Citation: Simons YB, Mostafavi H, Zhu H, Smith CJ, Pritchard JK, Sella G (2025) Simple scaling laws control the genetic architectures ofhuman complex traits. PLoS Biol 23(10): e3003402. https://doi.org/10.1371/journal.pbio.3003402

Editor: Greg Gibson, Georgia Institute of Technology, UNITED STATES OF AMERICA

Received: August 10, 2025

Accepted: September 5, 2025

Published: October 13, 2025

Data availability statement: The authors confirm that all data underlying the findings are fully available without restriction. Code available at https://github.com/YuvalSimons/Scaling. Figure data available at https://doi.org/10.5281/zenodo.17041176.

Funding: This work was supported by the National Institutes of Health (https://grants.nih.gov/): R01 HG008140 and R01HG011432 to JKP; R01GM115889 to GS; and F32HG011202 to YBS. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Genome-wide association studies have revealed that the genetic architectures of complex traits vary widely, including in terms of the numbers, effect sizes, and allele frequencies of significant hits. However, at present we lack a principled way of understanding the similarities and differences among traits. Here, we describe a probabilistic model that combines the effects of mutation, drift, and stabilizing selection at individual sites with a genome-scale model of phenotypic variation. In this model, the architecture of a trait arises from the distribution of selection coefficients of mutations and from two scaling parameters. We fit this model for 95 highly polygenic quantitative traits of different kinds from the UK Biobank. Notably, we infer that all these traits have fairly similar, though not identical, distributions of selection coefficients. This similarity suggests that differences in architectures of highly polygenic traits arise mainly from the two scaling parameters: the mutational target size and heritability per site, which vary by orders of magnitude among traits. When these two scale factors are accounted for, we find that the architectures of all 95 traits are very similar.

Introduction

A central goal of genetics is to understand how genetic variation maps to phenotypic variation. Starting in the late 20th century, there was huge progress toward identifying the genes for Mendelian traits. But most phenotypic variation in humans is genetically complex, and it is only in the last 15 years that genome-wide association studies (GWAS) have started to reveal the genetic basis of variation in a wide array of complex traits [1]. These studies have now identified tens of thousands of robust associations between genetic variants and a wide array of traits and diseases.

Competing interests: The authors have declared that no competing interests exist.

One intriguing observation from this work is the striking variation in genetic architecture among complex traits [2–5]. (Here, we use the term *architecture* to refer to the numbers of causal variants and their joint distribution of allele frequencies and effect sizes.) Traits have been found to vary in all aspects of genetic architecture, including: the number and magnitude of significant signals found at a given sample size [6]; the fraction of heritability explained by lead GWAS signals [3]; the allele frequency distributions of significant variants [4,7]; the estimated numbers of causal variants [8,9]; and the SNP-based heritability [10,11].

Nonetheless, diverse traits do show important similarities. First, most complex traits are influenced by large numbers of variants with small effects, only a small fraction of which can be confidently detected at current sample sizes [12–14]. Indeed, even relatively "simple" complex traits such as molecular biomarkers are highly polygenic with $\sim 10^4$ causal variants spread widely across the genome, compared to $\sim 10^5$ or more variants for traits such as height or BMI [2,9,15–17].

Second, the distributions of effect sizes of causal variants are not fit well using standard modeling assumptions such as normal distributions. Instead, effect sizes typically span several orders of magnitude, much like power-law distributions [18,19].

Third, trait-associated variants are often highly pleiotropic: i.e., they influence many traits simultaneously. Many pairs of traits show significant genetic correlations, indicating that allelic impacts are often shared [16,20,21]; moreover, whenever different traits are mediated through overlapping cell types or pathways, we can expect that they will share many of the same regulatory variants even if the directions of effects are uncorrelated [5,14,16,22].

Fourth, selection plays a central role in shaping complex trait architecture. Evolutionary theory and empirical evidence indicate that variants with phenotypic effects would usually be under selection and, in particular, that selection is usually stronger for larger-effect variants [23]. Consistent with this, variants with larger effect sizes tend to be at lower frequencies, suggesting that selection prevents such variants from reaching high frequencies [24–26]. Since heritability depends on both effect sizes and allele frequencies, an important consequence is that the genes that are most important for a trait contribute less to heritability than would be expected in the absence of selection, thus flattening the heritability distribution across genes [17,27,28].

Here we develop a principled approach for understanding similarities and differences in genetic architecture. Specifically, we want to understand how the population genetic processes of mutation, selection, and drift alongside properties of individual traits determine the numbers of variants, as well as the joint distributions of allele frequencies and effect sizes. *What features of these processes are shared across traits? And which are different? And, consequently, what features of the genetic architecture are shared or differ among traits?*

To answer these questions, we require a model for how population genetic processes shape complex trait architecture. Current models differ primarily in their assumptions about the relationship between selection on alleles (or alternatively their frequencies), and the effect sizes of those alleles on a trait of interest [23]. The heuristic 'α-model', developed for estimating SNP heritability, assumes a particular parametric relationship between allele frequencies and effect sizes [25,29,30]. While the α-model is motivated by the observed inverse relationship between effect size and frequency, the precise functional form is arbitrary. In turn, several evolutionary models postulate particular parametric relationships between the strength of selection on alleles and their effect sizes, and then rely on explicit population genetics models to derive the relationship between allele frequencies and effect sizes and other aspects of genetic architecture [31–36]. These models, however, differ in their predictions about architecture, owing to the various ad-hoc parametric relationships they assume.

Simons et al. (2018) introduced an evolutionary model that moves beyond ad-hoc choices by deriving the relationship between selection on alleles and their effects on a trait under an explicit, interpretable, biological model [27]. Motivated by extensive evidence that many quantitative traits are subject to stabilizing selection, where fitness declines with displacement from an optimal trait value [23,37,38], and that genetic variation affecting one trait often affects many others [5,16,22], they modeled selection on alleles that arises from stabilizing selection in a multi-dimensional trait space. They then used an explicit population genetic model to derive the genetic architecture of a focal trait with mutation, genetic drift, and stabilizing selection in a multi-dimensional trait space.

As we will show in the next section, the Simons et al. model can be reframed as a generative (statistical) model for the genetic architecture of a continuous complex trait, which depends on a few biologically interpretable parameters. We next describe how to infer these parameters and test the model fit based on data from GWAS. Applying our inference to 95 highly-polygenic quantitative traits of different kinds from UK Biobank, we show that this model provides an excellent fit to the data. Surprisingly, we find that most variation in architecture among traits is explained by differences in just two scaling parameters: the mutational target size and heritability per site.

Results

A population genetic model of complex traits. As a starting point, we assume that phenotypic variation exists in a high-dimensional trait space under stabilizing selection. Here we outline key elements of the model; further details and biological motivation can be found in S1 Note and in [23,27].

We model each person's phenotype as a point in an n-dimensional trait space, and assume that this dimension is high ($n \geq 10$). To model stabilizing selection we assume that there is an optimal phenotype, and that fitness decreases with Euclidean distance from the optimum (Fig 1A).

The phenotypic effect of each variant is represented by a random vector in the n-dimensional trait space, namely $\vec{b} = (b_1, b_2, \ldots b_n)$, where b_i is the additive effect of the derived allele on the i-th trait. We assume that a person's phenotype arises from their genotype according to the standard additive model in n-dimensions: it is a vector sum over the effects of all variants plus a random vector representing the environmental effects [39].

The model thus far is mathematically similar to Fisher's Geometric model [40], which Fisher and others used to study adaptive processes [41], but we consider a different question and a different evolutionary setting. We focus on the genetic architecture of a single highly polygenic trait that arises in the balance between mutation, stabilizing selection in the multi-dimensional trait space, and genetic drift.

Mutation, selection, and drift at individual sites. Each generation, mutation introduces new trait-affecting variants into the population at a rate μ per site, per gamete, per generation. The long-term fate of variants is determined by the combined action of selection and drift.

Under stabilizing selection, at equilibrium, selection holds the population's phenotypic mean very close to the optimal phenotype, and thus acts against mutations (and against variation in general). The strength of selection, s, acting against a variant is proportional to its squared magnitude in the n-dimensional trait space (Fig 1B):

$$s = ||\vec{b}||^2 / V_S = \sum_{j=1}^{n} b_j^2 / V_S, \quad (1)$$

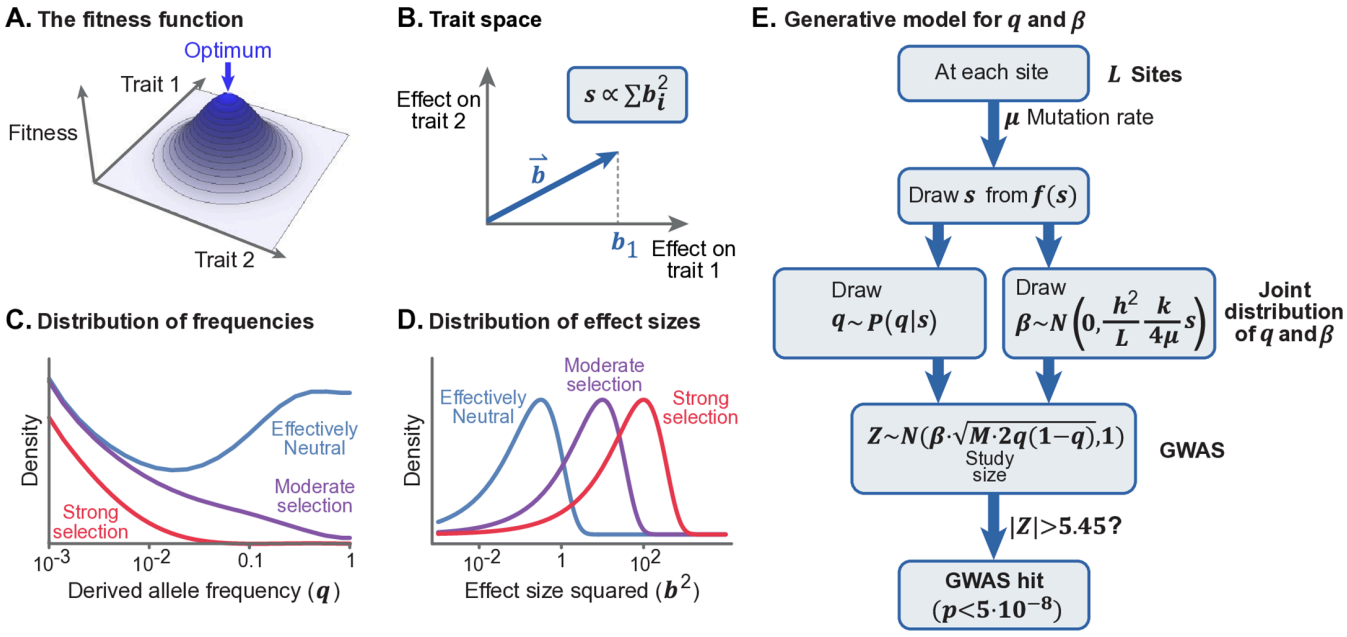


Fig 1. The model. **(A)** We use Fisher's concept of a multi-dimensional trait space. Under stabilizing selection, an individual's fitness declines with distance from the optimal phenotype [40]. **(B)** The selection coefficient experienced by a variant is proportional to the sum of squared effects on all traits. **(C)** We compute the distribution of derived allele frequencies (q) conditional on s and demography. **(D)** The distribution of effect sizes for trait 1 (b_1) is normally distributed given s. **(E)** The generative model for q, β, and the observed Z-score at any given site. In C and D, the curves for effectively neutral, moderate and strong selection correspond to $s = 10^{-4.5}, 10^{-3}$ and 10^{-2} respectively.

https://doi.org/10.1371/journal.pbio.3003402.g001

with V_S reflecting the width of the fitness function around the optimum.

Given s we can compute the present day allele frequency distribution, as follows. When traits are subject to stabilizing selection, selection at individual sites is under-dominant, meaning that selection acts against minor alleles, regardless of the direction of effect [27,42,43]. At strongly selected sites, this approximates the standard model of selection against deleterious alleles. Specifically, the expected change in allele frequency at an autosomal site in a single generation, given current derived allele frequency q is

$$\mathrm{E}[\Delta q] = -sq(1-q)\left(\frac{1}{2} - q\right). \quad (2)$$

Meanwhile, the variance in the change in allele frequencies, i.e., drift, scales inversely with population size [44]. Hence, the distribution of present-day allele frequencies is the result of a stochastic process including past mutations, selection, and drift – which depends on the history of population sizes. For our analysis here, we computed the distribution of present day allele frequencies under the stabilizing selection model using a demographic model estimated for the British population [45]. As expected, strongly selected variants (large s) tend to be rare, while nearly-neutral variants (small s) can drift to high frequencies (Fig 1C).

The relationship between selection and effect sizes. Next we need to understand how selection in the multi-dimensional trait space relates to the genetic architecture of a single focal trait of interest. Without loss of generality, we focus on the first dimension in the n-dimensional trait space, and to simplify the notation we denote the effect size b_1 of a variant on trait 1 simply as b.

Causally, effect sizes determine selection coefficients. However, since we want to describe the co-distribution of frequencies and effect sizes we need to invert this relationship: specifically, we need the conditional distribution of the effect size on trait 1 given s. This conditional distribution reflects uncertainty about the projection of $\vec{b}$ onto the first dimension if all we know is s (or equivalently $||\vec{b}||^2$). Fortunately, when the number of traits is sufficiently large, this conditional distribution is well approximated by a Normal distribution (see [27] and Sect 1.3 in S1 Note), namely:

$$b|s \sim N(0, c \cdot s), \tag{3}$$

where $c = V_S/n$. Intuitively, variants under weak selection (small s) tend to have small squared effect sizes (b^2) and variants under strong selection (large s) tend to have larger squared effect sizes (Fig 1D).

With the distributions for b and q given s, we can now compute the expected per site contribution to phenotypic variance as a function of s, given by $\mathrm{E}[2b^2q(1-q)|s]$. Under non-equilibrium demography, the expectation does not have a simple form but it is plotted in Fig B in S1 Note. At sites where selection is weak, b^2 is small, and these sites contribute little to the additive genetic variance, V_A. When selection is strong, b^2 is large, but selection holds $q(1–q)$ low, and these effects cancel out, so these sites are capped in terms of how much they can contribute to V_A [27].

Single-site dynamics and heritability. Moving from single sites to a genome-wide model, let L be the number of sites in the genome at which mutations can affect trait 1 (more precisely, we require mutations to have effects that exceed some predefined small value); we refer to L as the *mutational target size*. We use $f(s)$ to denote the unknown distribution of selection coefficients of mutations at these L sites. Then the expected total additive genetic variance for trait 1 is given by

$$V_A = L \int \mathrm{E}[2b^2q(1-q)|s] \cdot f(s)ds. \tag{4}$$

Next, we rescale b from the original but arbitrary measurement units into units of standard deviations of the trait value: we define $\beta = b/\sqrt{V_P}$, where V_P is the phenotypic variance. Dividing both sides by V_P, and noting that V_A/V_P is the (narrow sense) heritability h^2, we can relate heritability to the site-level parameters:

$$h^2 = L \int \mathrm{E}[2\beta^2q(1-q)|s] \cdot f(s)ds. \tag{5}$$

This equation expresses the key relationship between heritability (h^2), mutational target size (L), and the expected contribution to variance per site.

Finally, Eq 3 can be rewritten in terms of β and population genetics parameters (for details see Sect 1.5 in S1 Note):

$$\beta|s \sim N\left(0, \frac{h^2}{L} \cdot \frac{k}{4\mu}s\right), \tag{6}$$

where k is a constant that depends on $f(s)$ and demography and is approximately 1, and μ is the mutation rate. Crucially, Eq 6 shows that the trait's heritability per site, h^2/L, is a fundamental scaling factor that relates selection on alleles to their effects on the trait.

Together, these results provide a generative model for the genetic architecture of a complex trait (Fig 1E). Assuming that the demographic history and mutation rate per site are known in advance, this model is fully specified in terms of three unknowns: the mutational target size, L; the heritability per site, h^2/L; and the distribution of selection coefficients, $f(s)$. We now describe how we estimate these from GWAS data.

Inference of model parameters from GWAS data. In principle we would want to perform inference using all causal variants, but this is technically challenging since most causal sites have very small effect sizes; hence there is great uncertainty about which sites are causal and their true effect sizes. As a tractable alternative, we restricted our inference to the independent genome-wide significant hits for each trait. We account for this restriction in the inference by noting that we only observe the subset of sites for which the absolute GWAS z-score exceeds 5.45, corresponding to the conventional significance threshold of $p < 5 \times 10^{-8}$, and the minor allele frequency (MAF) exceeds 1%, corresponding to the imputation threshold (see Sect 2.1 in S1 Note). We performed simulations to illustrate the changes in architecture at GWAS hits as a function of each of the main model components (Fig 2; Sect 5 in S1 Note). As expected from theory:

A. When selection is weak, causal variants can drift to high frequencies, and most significant hits are at common variants. Conversely, when selection is strong, there is a greater fraction of rare variants among the significant hits, and an inverse relationship between effect size and allele frequency.

B. When traits have high heritability per site h^2/L, the squared effect sizes and z-scores tend to be larger, there are more genome-wide significant hits and they explain a greater proportion of heritability, compared with traits with low h^2/L.

C. When traits have a large mutational target size L (holding h^2/L and $f(s)$ constant), there are more causal variants, and more genome-wide significant hits, but the distribution of allele frequencies and effect sizes, and the proportion of heritability explained by hits, are unaffected.

We implemented a maximum likelihood method that estimates the components of our model from the joint distribution of q and $|z|$ across significant hits (Fig 2D–2F; Sect 5 in S1 Note). We fit $f(s)$ using a spline function with four knots (see Sect 4.8 in S1 Note for how the number and position of knots were chosen), thus our full model includes six parameters per trait: four for $f(s)$, as well as h^2/L, and L.

We tested this method using simulated GWAS data under a variety of parameter values. We find that even with modest numbers of hits (∼100) the method provides accurate estimates, while the estimates are noisy for traits with fewer hits (e.g., panel 2E). It may seem surprising that we can estimate $f(s)$ from relatively few observations, but each variant carries considerable information about the strength of selection: the allele frequency bounds s for that variant from above and the effect size bounds s both from below and above, such that jointly they are highly informative (Fig D in S1 Note). Given these results we analyzed traits with at least 100 hits.

We also observed that the data are less informative at both ends of the range of possible selection coefficients: GWAS has low power to detect very strongly selected variants ($s \gtrsim 10^{-2}$) as their allele frequencies are too low, and low power to detect effectively-neutral variants ($s \lesssim 10^{-5}$) as their effect sizes are too small. We therefore implemented a regularization penalty to constrain $f(s)$ to sensible values at these extremes (Sect 3.8 in S1 Note).

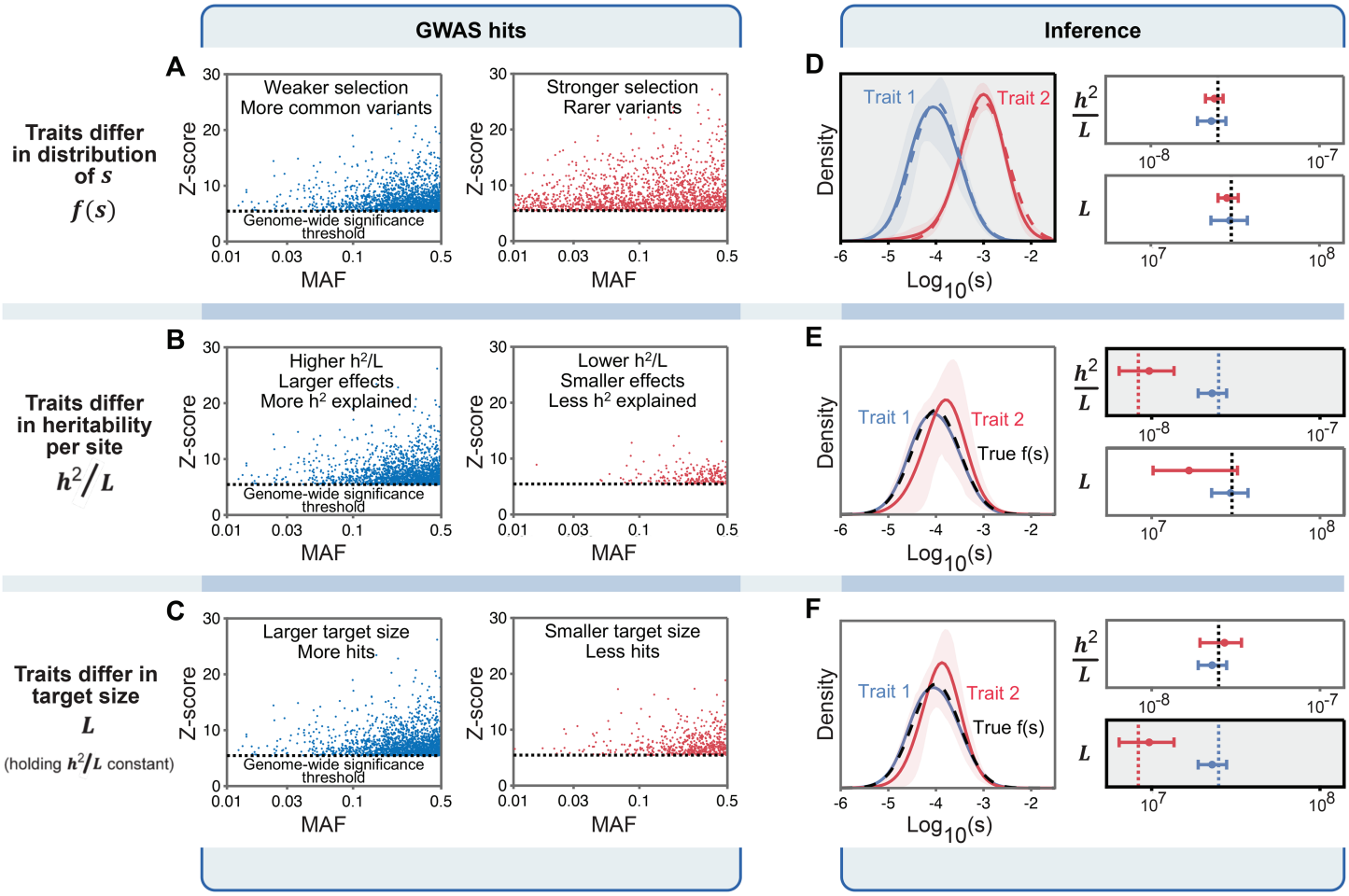


Fig 2. Inference of model parameters. **(A)–(C)** The joint distributions of minor allele frequencies (MAFs), z-scores, and numbers of hits per trait depend on model parameters as illustrated here. Each graph shows simulated distributions of genome-wide significant hits, with the graphs in each row differing in one of the main axes of our model. **(D)–(F)** True values of the distributions of $f(s)$ as well as h^2/L and L are indicated by the dashed lines; inferences are indicated by solid lines or by point estimates with sleeves and bars indicating 90% bootstrap CIs. The inferred parameters differing between the pair of traits in each row are highlighted in gray. See Sect 9 in S1 Note for parameter values used. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g002

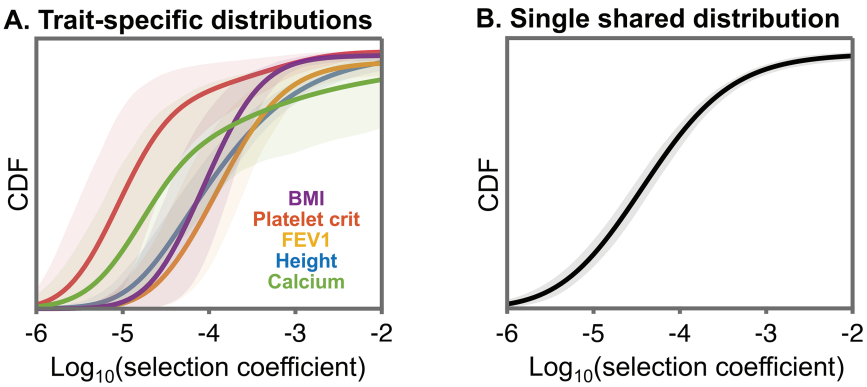
Dataset of 95 quantitative traits from the UK BioBank. We selected traits from the UK Biobank for analysis, as follows (Sect 2.4 in S1 Note). Since our model is most directly applicable to quantitative continuous traits, we restricted our analysis to such traits. We identified independent lead variants for each trait using COJO [46]. Since low-frequency variants are often poorly imputed, we removed hits with MAF<1% (see Sect 6.3 in S1 Note for how we account for this in the inference). We excluded traits for which more than 10% of hits fell within a single LD block, as well as hits in regions of extremely high LD (LD score >500). As noted above, we restricted ourselves to traits with at least 100 independent hits; doing so implies that the traits are among the more polygenic and heritable traits in UKBB. For each trait we recorded the number of hits, and the estimated allele frequency, z-score, and effective sample size for each hit.

Our inference uses the architecture of GWAS hits as a proxy for the architecture of the causal variants they tag. In Sect 7 in S1 Note we tested the validity of this approximation using data analysis and simulations. For example, we started with the genotypes from UKB,

picked variants to be causal and assigned them effect sizes in various ways, assigned phenotypes to individuals based on the additive model, and then performed GWAS and ran COJO, allowing us to compare the inferred number of hits and their distribution of frequencies and effect sizes with the known underlying architecture. Our results indicate that under sensible assumptions the GWAS hits picked by COJO faithfully reflect the architecture of the underlying causal variants.

Our criteria resulted in a list of 95 traits, with a range of 100 to 1,760 hits per trait (mean=495). These traits include 40 morphometric traits, of which 26 are related to body weight or adiposity (e.g., BMI, waist circumference and birth weight) as well as 14 others (e.g., height, bone mineral density and hand grip strength). The traits also include 27 blood phenotypes (e.g. platelet traits, lymphocyte count, and hemoglobin measurements), and 12 molecular traits sampled from blood or urine (e.g., IGF-1, triglycerides and calcium levels). Additionally, we have 9 cardiovascular traits, including pulse rate, blood pressure measurements, and pulmonary function traits. Lastly, we include 6 ophthalmologic traits, and 1 behavioral trait (age at first sexual intercourse).

Distributions of trait parameters. We applied our inference to all 95 traits. Fig 3A shows the estimated distributions of selection coefficients that vary in polygenicity and kind, from calcium levels to BMI, and are representative of the range spanned by all 95 traits (see Fig F in S1 Note). Note that although our inference is based on significant hits, $f(s)$ represents the distribution of s among *new* mutations.



C. Inferred heritability and target size for 9 traits

Trait	No. of hits	Heritability (h^2)	Target size (L)	h^2/L
BMI	541	0.1 0.3 1	10^6 10^7 10^8 10^9	10^{-9} 10^{-8} 10^{-7} 10^{-6}
sBP	321			
FEV1	394			
Height	1760			
Pulse rate	331			
Spherical power (eye)	125			
Platelet crit	724			
Calcium level	285			
Urea	231			

Fig 3. Parameter estimates for example traits. (A) CDF of the distributions of selection coefficients for newly-arising mutations, $f(s)$, estimated for each trait separately, using trait-specific distributions, with 90% confidence envelopes. **(B)** CDF of the single shared distribution (SSD) of selection coefficients for newly-arising mutations, estimated using all 95 traits together. **(C)** Properties of example traits. h^2 and L are estimated using $f(s)$ from the SSD. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g003

Our point estimates of the distribution (solid lines) vary considerably among traits, but they share important features. Most trait-affecting mutations are under weak selection, with modes ranging between 10^{-4} – $10^{-5.5}$. In this range, the strength of selection is roughly comparable to genetic drift ($\sim 10^{-4.5}$ per generation), consistent with the observation that many GWAS variants are common, and more generally that much of the heritable variation in complex traits arises from common variants [3,4,10,12]. However, the distributions have a substantial tail in the strong selection range ($s > 10^{-3}$), and therefore span multiple orders of magnitude. Since selection coefficients span multiple orders of magnitude, so too should effect sizes. This echoes recent results showing that distributions of effect sizes for complex traits do indeed span multiple orders of magnitude (in contrast to the normal distribution which has often been assumed in statistical genetics models) [18,19].

Although our point estimates of the distribution (solid lines) vary considerably among traits, their confidence envelopes largely overlap (shades). With our inferences being based on the architecture of significant hits, the overlap among confidence envelopes suggests that this architecture is largely insensitive to the differences reflected in our point estimates (also see Sect 5.2 in S1 Note). We thus conjectured that we could build a unified model by assuming a Single Shared Distribution for $f(s)$, which we refer to as the SSD, instead of assuming separate Trait-Specific Distributions (TSDs). The SSD is shown in Fig 3B. We also inferred the SSD for a subset of the traits that were chosen to minimize genetic correlations among traits (15 traits with $|r_g| < 0.2$ for any pair) and found that the estimated SSD is insensitive to correlations among traits in the full dataset (Sect 6.4 in S1 Note). As we will show, the SSD provides a useful approximation for the architecture of individual traits, while greatly cutting down the number of model parameters and highlighting important shared features of trait architecture.

In contrast to $f(s)$, our estimates of heritability h^2 and of the mutational target size L span a wide range that far exceeds the CIs for individual traits (Fig 3C and S1 Table). For example, among the traits in Fig 3C, BMI (541 hits) has an estimated target size of 220 MB, one of the highest estimated target sizes, in contrast to urea (231 hits) with a target size of 4 MB, one of the lowest estimated target sizes. These results are broadly consistent with expectations from previous studies of polygenicity showing that morphological traits including height and BMI have many more contributing variants than do molecular traits, illustrated here by urea and calcium [2,9]. Our estimates of heritability vary by over one order of magnitude among traits, and are concordant with previous estimates (S1 Table, Fig J in S1 Note; [10]). Our estimates of the ratio h^2/L vary less than h^2 and L alone, plausibly because conditioning on traits with ≥ 100 hits biases us toward traits with greater (and thus a more restricted range of) h^2/L (see Fig 2B and below).

Quantifying model fit. Next, we assessed the fit of our models to the genetic architecture observed in GWAS, including how the fit is affected by using the SSD approximation (Sect 4 in S1 Note). To do so, we computed a measure of model fit using the predicted distribution of z-scores given the allele frequencies and study size. For each variant i, we computed what we refer to as a *residual p-value*: $\Pr(|z| > z_i \mid z_i > 5.45, q_i, \ Model)$, where z_i and q_i are the observed z-score and frequency of SNP i, respectively, and *Model* indicates the SSD or TSD model. The SSD model fit depends on only one trait-specific parameter, h^2/L, as $f(s)$ is shared across traits. In contrast, the TSD model fit depends on five parameters per trait: h^2/L and four parameters to fit $f(s)$. The estimated target size, L, does not affect the fit by this measure and more generally (because given the other model parameters, L is estimated by matching the expected and observed number of hits).

The residual p-value has a simple interpretation: If we correctly model the distribution of z-scores among significant hits, then the distribution of residual p-values will be uniform between 0 and 1. If the observed z-scores are too small then the residual p-values will skew

toward 1, and if the z-scores are too high they will skew toward 0. To avoid overfitting, we split the genome into approximately independent blocks [47], each time inferring the model on 90% of the blocks and computing residual p-values for the held-out 10%.

We first considered the fit of our model for height, the trait with the greatest number of hits in our dataset (Fig 4A). For height, the distribution of z-scores across the 1760 hits is fit essentially perfectly by the SSD model (Fig 4A). In contrast, two simpler heuristic models provide a poor fit to the distribution of z-scores (Fig 4A). First, when we assumed that effect sizes are normally distributed, the resulting residual p-values deviate greatly from uniformity with many extremely small p-values. We also considered a version of the α-model with a normal density of effect sizes conditional on allele frequencies [30]. By fitting the inverse relationship between allele frequencies and effect sizes, the alpha model improves the overall fit, but still has an excess of tiny p-values. In both cases, the underlying normal distribution is too narrow to accommodate the wide variation in observed effect sizes, including many hits close to the significance threshold and a minority of much stronger hits.

For BMI, with 541 hits, our SSD model fits most of the distribution well, but the top hits are larger than expected (Fig 4B). In particular, the residual p-value of one SNP is significant even after Bonferroni correction. This outlier represents the well-known FTO signal that was detected even in very early GWAS studies [48,49].

We find 14 additional outlier SNPs for a variety of other traits. The 14 outliers include both missense and noncoding variants, and are all found near genes previously implicated in the biology of the relevant traits (Table A in S1 Note and S2 Table). Our inferences are insensitive to the exclusion of these outliers (see Sect 4.5 in S1 Note). We hypothesize that these outliers violate our model assumptions in some way that allows them to be common despite having large effects. For example, they might have much smaller pleiotropic effects than most other variants affecting those traits, leading to weaker selection than expected given their effect sizes [50]. Alternatively, they may have been targets of strong positive or balancing selection that allowed them to reach high frequencies despite their large effect sizes.

We next performed goodness-of-fit tests for each trait to determine whether the overall distributions of residual p-values match the expected uniform distribution, using Kolmogorov-Smirnov statistics (Fig 4C; Sect 4 in S1 Note). We find that the α-model fits

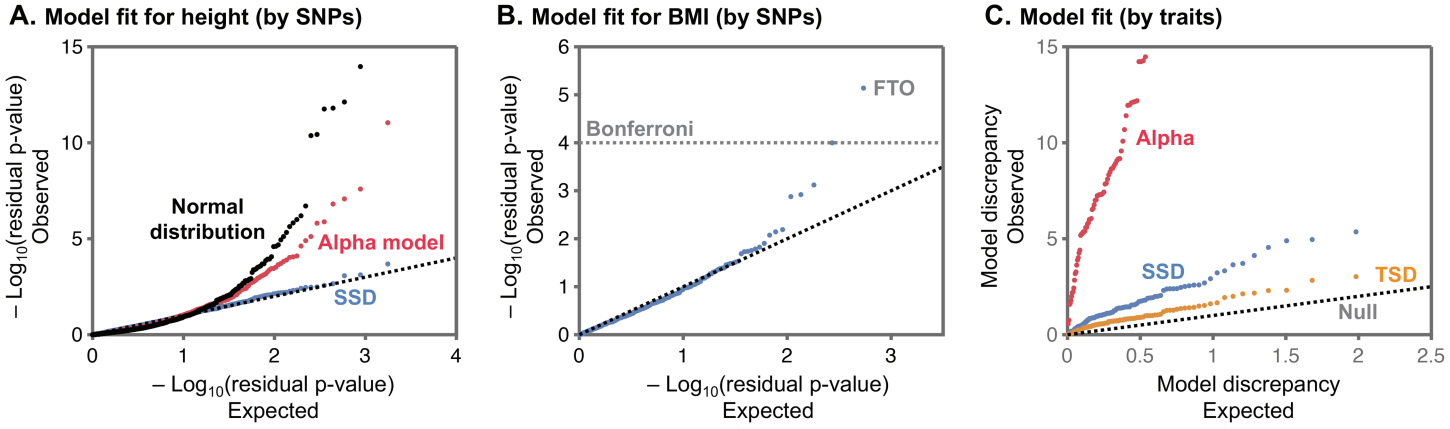


Fig 4. Model fit. (A) QQ-plot of residual p-values for height (each data point is a SNP) under three models: the SSD model provides a good fit to the distribution of z-scores, while two other models fit poorly. **(B)** For BMI, the SSD model fits most of the z-score distribution, but a few hits are more significant than expected, notably at FTO. **(C)** QQ-plots for model fits (by trait) for TSD, SSD, and α-models. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g004

the data far worse than either the TSD and SSD models; notably, we can reject it for 93 out of 95 traits at a FDR of 0.05 (Sect 4 in S1 Note). In contrast, both the TSD and SSD models appear to fit the data well, with a modest inflation of p-values, perhaps relating to simplifying assumptions of the model and inference (Fig 4C). For most traits, the TSD model fits slightly better than the SSD model, with no trait rejected for the TSD and 26 traits for the SSD at a FDR of 0.05 (Sect 4 in S1 Note), indicating that the TSD correctly identifies some degree of trait specific signal. Nonetheless, the SSD model fits the data quite well and is far more parsimonious, with one trait specific parameter compared to five for the TSD model.

Prediction of allele ages. The previous results show that the model provides a good empirical fit to the data. We next wanted to evaluate whether it can also predict the evolutionary processes underlying the genetic architecture. To this end, we turned to an entirely different type of predictions from our model. Conditional on allele frequency, deleterious alleles tend to be younger than neutral alleles [24,51]; and our parameter estimates can be used to predict the extent of this effect for GWAS hits. We compared our model's predictions to allele ages estimated from a reconstruction of the ancestral recombination graph using Relate [45] (see Sect 8 in S1 Note for details, including a correction for the estimator of allele ages).

Looking at the allele ages of GWAS hits for all 95 traits as estimated by Relate, we see that they are much younger than frequency-matched, putatively neutral alleles (Fig 5). We estimate that the median age of a GWAS hit variant is 137,000 years, compared to 594,000 years for matched neutral variants. These observations highlight the competing influences of selection and drift on GWAS hits: due to selection GWAS hits are much younger than matched neutral variants, yet at the same time selection is weak enough that most GWAS hits are fairly old, predating the out-of-Africa bottleneck. This finding echoes previous ones showing that many common GWAS hits are shared among African and non-African populations [52,53].

Most importantly, the shift in ages of GWAS hits compared to matched control SNPs is predicted well by the SSD model, indicating that this model captures the right magnitude

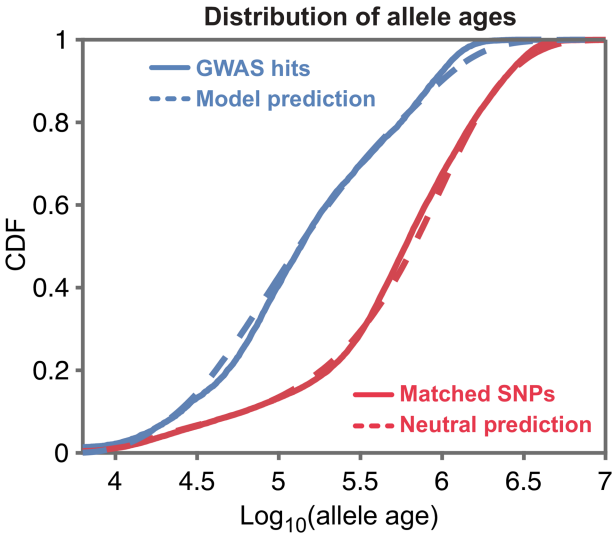


Fig 5. Allele ages. The distribution allele ages of GWAS hits for all 95 traits (solid blue), estimated using Relate, compared to the distribution of allele ages predicted by our model (dashed blue). Also shown, the distribution of allele ages for neutral frequency-matched SNPs (solid red) and the distribution predicted by a neutral model (dashed red). Allele ages were converted to years by assuming 28 years per generation [54]. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g005

of selection coefficients. (For results about our ability to distinguish selection coefficients with this analysis see Fig X in S1 Note.) Since the allele ages inferred by Relate are estimated from local haplotype structure, information that isn't used by our inference, this concordance provides an external validation of the SSD model.

Simple scaling rules control differences in trait architectures. The fit of the SSD model suggests an intriguing prediction: that the differences in genetic architecture among traits are primarily due to just two trait-specific parameters: the heritability per site (h^2/L) and the mutational target size (L).

To see why, consider the full genetic architecture of a trait, of *all* variants regardless of whether they can be detected by GWAS. First, for a given demographic history, the distribution of allele frequencies depends only on the distribution of selection coefficients, $f(s)$. Hence, under the SSD approximation, the allele frequency distribution is shared across traits and can be predicted from our estimate of $f(s)$ (black line, Fig 6A).

Next, while $f(s)$ represents the distribution of selection coefficients among new mutations, strongly selected variants are less likely to reach high frequencies. Hence, the distribution of selection coefficients shifts to be smaller with increasing allele frequencies, as shown in Fig 6B.

Given the distributions of selection coefficients at different allele frequencies from Fig 6B, we can compute the distribution of squared effect sizes as a function of allele frequencies, by integrating Eq 6 over s. Crucially, under the SSD model, these distributions are identical across traits, if the effect sizes on the x-axis are scaled in terms of h^2/L (black lines, Fig 6C).

How do these distributions affect GWAS hits? Unlike the underlying distributions, the power to detect significant variants in GWAS depends on the actual squared effect sizes β^2, not scaled by h^2/L (and it depends on allele frequency and sample size). Consequently, there is more power to detect variants for traits with higher h^2/L – this is intuitive, because higher h^2/L implies that on average each site explains more variance in the trait. This is illustrated in Fig 6C: for traits with high h^2/L, variants within both light and dark blue regions are genome-wide significant, but for traits with low h^2/L, only variants within the dark blue regions are detected.

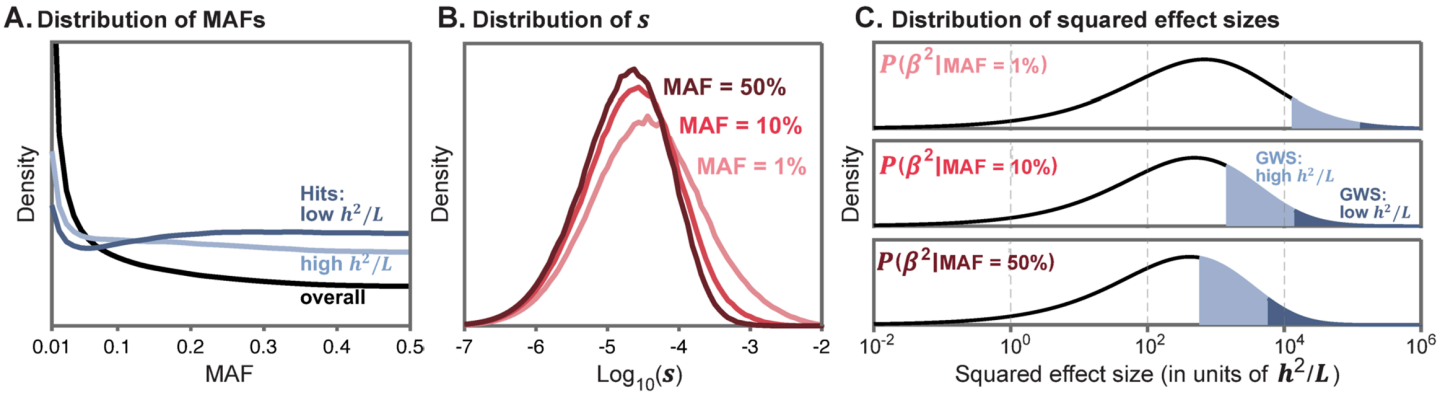


Fig 6. Shared genetic architecture under the SSD model. (A) Distribution of allele frequencies for all causal variants (black), and for genome-wide significant hits (light/dark blue), for our inferred $f(s)$ given British population history. **(B)** Distributions of selection coefficients at causal variants with different minor allele frequencies. **(C)** Distributions of squared effect sizes β^2, shown here for three example minor allele frequencies; note that effect sizes are scaled by the natural units of h^2/L. For traits with high h^2/L, variants within both the light and dark blue regions are genome-wide significant (GWS); for traits with low h^2/L, only the dark blue regions are significant. See Sect 9 in S1 Note for parameter values used. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g006

Moreover, even though the underlying distribution of causal variant allele frequencies is shared among traits, the frequency distribution of hits is predicted to vary. For traits with high h^2/L there is relatively more power to detect low frequency variants than for traits with low h^2/L (Fig 6A).

Lastly, the second scaling parameter, L, represents the mutational target size. Conditional on h^2/L, changing L only changes the *numbers* of causal variants (and numbers of hits) but does not change any of the distributions.

We first tested these predictions for three traits: height, platelet crit (a blood phenotype), and FEV1 (a measure of lung function), where we reduced the sample size to 330,000 so it is identical for all traits (Sect 6.7 in S1 Note). Fig 7A shows height and platelet crit. These two traits differ greatly in their number of hits (1533 vs 648), estimated heritability (77% vs 32%) and mutational target size (30 MB vs 10 MB). However, we estimate that they have very similar values of h^2/L ($\sim 3 \times 10^{-8}$). Consistent with our prediction, the distributions of z-scores, effect sizes, and MAFs for significant hits are nearly identical for the two traits.

In contrast, height and FEV1 have similar mutational target sizes, but h^2/L for height is about 6-fold higher than for FEV1 (Fig 7B). As expected, this results in greater power to detect variants associated with height than with FEV1 (40% vs. 14% of h^2 explained by significant hits). Consequently, genome-wide significant hits for height have larger effect sizes and larger z-scores than for FEV1. We also see slightly fewer low-MAF hits for FEV1 ($p = 0.01$, KS test), due to the lower power compared to height.

We predicted that after rescaling the z-scores (and effect sizes) for these traits by $\sqrt{h^2/L}$, their architectures should become nearly identical. On this scale, the significance cutoff is $5.45/\sqrt{h^2/L}$, which is higher for FEV1 than for height (illustrated by the light and dark blue regions in Fig 6C). We therefore compared the architecture of genome-wide significant hits for both traits using the higher threshold in the scaled units. After doing so, we only have 95 hits for height, but it is apparent that the summary properties for both traits are indeed highly similar (Fig 7C).

We repeated this scaling procedure for the 57 other traits in our dataset whose scaled threshold is below that of FEV1. After doing so, the distributions for all hits are highly similar to the scaled distributions for height and FEV1 (dotted line, Fig 7C). We repeated similar analyses for all traits, and found that after applying these scaling laws the architectures of all 95 traits are nearly identical (Sect 10 in S1 Note).

Discussion

What determines the genetic architecture of complex traits? Does genetic variation in a given trait primarily reflect the idiosyncrasies of its biology–or alternatively, does it reflect processes that are shared among different traits?

Here we describe a principled approach to tackle these questions. Our point of departure is an evolutionary model of genetic architecture based on empirically motivated and interpretable biological assumptions. Given arguments and evidence that many quantitative (continuous) traits are subject to stabilizing selection [23,37,38], and that genetic variation affecting one trait often affects many others [5,20,21], we model selection on alleles that arises from stabilizing selection in multi-dimensional trait space. Otherwise, we assume the standard population genetic model incorporating the effects of mutation, selection, genetic drift and demographic history. This model gives rise to a family of genetic architectures, where the architecture of a given trait is determined by the distribution of selection coefficients at trait-affecting sites as well as two scaling parameters: the mutational target size and the heritability per site.

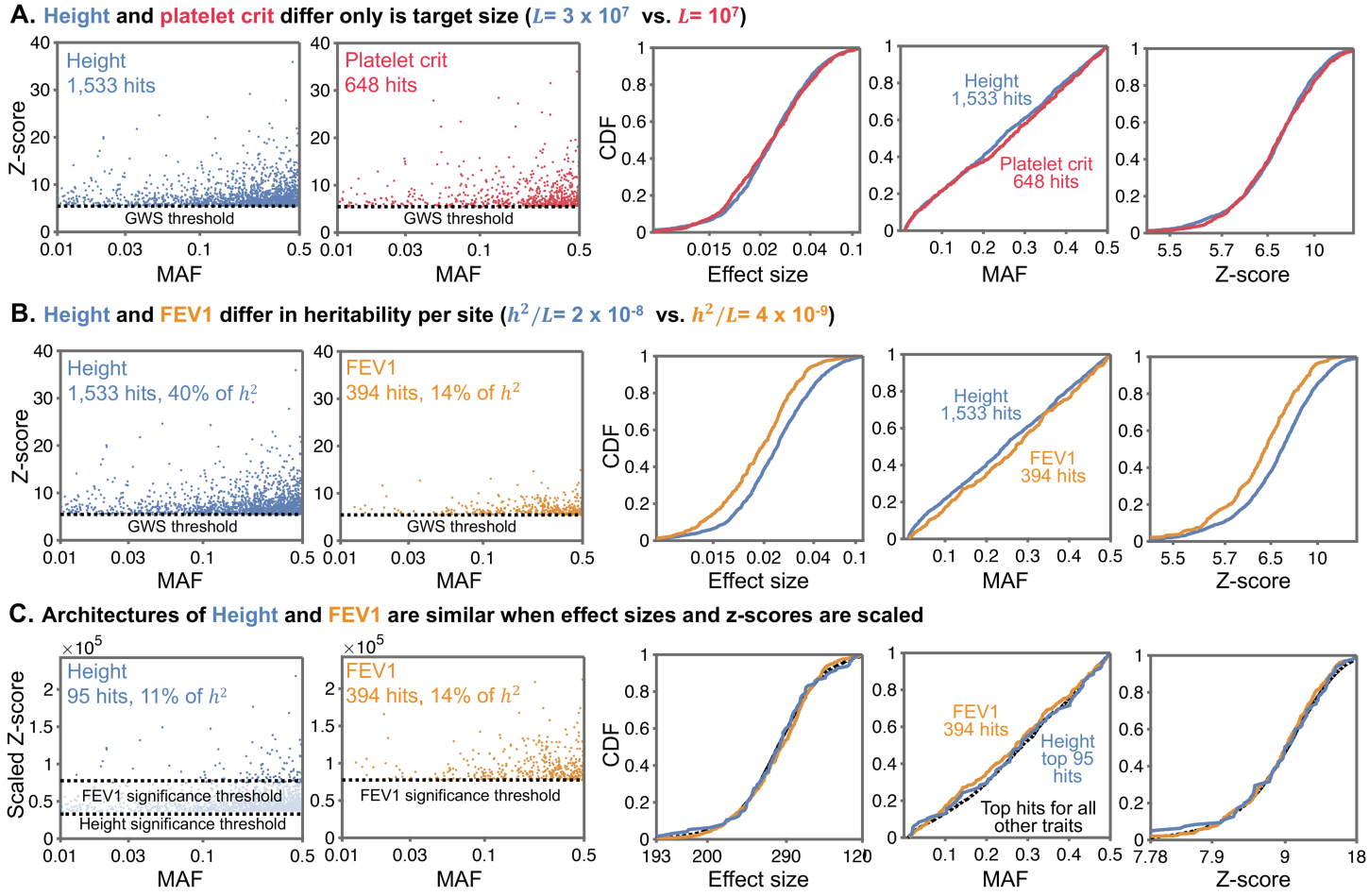


Fig 7. Heritability and target size underlie differences between trait architectures: examples for three traits. **(A)** Height (blue) and platelet crit (red) have the same heritability per site h^2/L, but height has a much higher mutational target size L. This results in many more hits for height (1533) than for platelet crit (648) (2 left panels). However, the marginal distributions of effect sizes, MAFs, and z-scores of hits are nearly identical for the two traits (3 right panels). **(B)** Height (blue) and FEV1 (gold) differ in h^2/L, but have similar L. Consequently, the joint distribution of z-scores and MAFs of their hits are markedly different (2 left panels), as are the marginal distributions of hit effect sizes, MAFs and z-scores (right). **(C)** After scaling by their respective $\sqrt{h^2/L}$, and imposing the more stringent scaled significance threshold (corresponding to FEV1) for both traits, the joint distribution of z-scores and MAFs of their hits (2 left panels) and the corresponding marginal distributions (3 right panels) are highly similar. Figure data available at: https://doi.org/10.5281/zenodo.17041176.

https://doi.org/10.1371/journal.pbio.3003402.g007

We performed inference allowing all these parameters to vary among traits. We found that the model provides a good fit to the joint distribution of allele frequencies and effect sizes at genome-wide significant hits for the 95 quantitative traits in our dataset. Intriguingly, we also found that the confidence envelopes for the distribution of selection coefficients largely overlap among these traits, suggesting that their genetic architecture can be approximated by a Single Shared Distribution (SSD) of selection coefficients and the two additional scaling parameters per trait. We then estimated the SSD using the data from all traits jointly, showing that it fits the architecture of individual traits well, and validated our estimate of the SSD by showing that it accurately predicts the distribution of allele ages at GWAS hits for all traits.

The fit of the SSD model implies that, aside from the two scaling parameters, the genetic architecture of genome-wide significant hits is similar among all 95 traits. Cross-sections of the estimated shared architecture are visualized in Fig 6. Indeed, as predicted, after we rescale the effect sizes by the estimated heritability per site, we find that the joint distributions

of effect sizes and allele frequencies for GWAS hits are similar among traits. Meanwhile the number of hits for a given trait is proportional to estimated target size (Fig 7).

These findings delineate the attributes of genetic architecture that are shared among traits and those that are trait-specific. In doing so, they raise new insights and questions:

How similar are the distributions of selection coefficients underlying heritable variation in different complex traits? Our results indicate that the distributions of selection effects among the 95 traits we examined were similar enough for the architectures of GWAS hits to be insensitive to the differences among them. Previous work has hinted at this similarity: for example, work using the α-model reported broadly similar relationships between MAF and effect size among a variety of traits [25,26,36], though it should be noted that there is no straightforward interpretation of α in terms of selection coefficients [36]. Previous work also hinted at the shared features of the distributions that we found, notably about selection on many GWAS hits being weak enough for them to precede the Out-of-Africa bottleneck and be shared among populations [52,53], and about their effect sizes spanning several orders of magnitude [18,19].

We might expect some similarity in the distribution of selection coefficients among complex traits. In a hypothetical extreme in which all functional variation affects all traits, the distribution would necessarily be shared among traits. This logic may well explain similarities among traits whose mutational target sizes encompass much of the functional portion of the genome; for example, we estimate the target size for BMI at ∼7% of the genome compared to ∼8% estimated to be functional [55–57]. It could also apply to traits that are mediated through the same tissues [14]. However, this logic does not explain the similarity among traits whose target sizes are substantially smaller and are primarily mediated through different tissues [58]. The biomarkers in our dataset, for example, have target sizes that are more than an order of magnitude smaller than BMI (e.g., calcium level with a target size of ∼0.1% of the genome), and are mediated through distinct cell types or tissues [9].

Shared features in the distribution of selection coefficients may reflect similarity in the biological systems in which variation affecting complex traits arises, notably in gene-regulatory networks. Heritable variation in complex traits is spread across most of the genome and is enriched in regulatory regions near most genes that are expressed in the tissues that affect these traits [14]. The lead variants typically explain a tiny fraction of the heritability [3], and the most relevant biological pathways are usually only modestly enriched for heritability [9,14]. In other words, most heritable variation is mediated through the regulation of genes and pathways that are not closely connected to the trait's biology [9,14,59]. Perhaps, the essential logic of gene regulatory networks and their evolution are sufficiently similar across tissues to generate the shared features we found, even if the specific pathways, genes, enhancers, and variants differ.

The similarity among distributions has important qualifications, however. Even among the 95 traits we examined and with our inference restricted to GWAS hits, we still identified statistically significant differences among the distributions of selection effects for different traits (see above and Sect 4 in S1 Note). As we already noted, restricting our inference to GWAS hits with MAF > 1%, delimits our power to infer selection effects at the weak and strong ends of the range of selection coefficients (e.g., for $s \lesssim 10^{-6}$ and $\gtrsim 10^{-2}$). We therefore cannot rule out more substantial differences in the distribution of, e.g., strongly selected, rare variants, even among the traits we examined (see, e.g., [28]). Additionally, the traits we examined were ascertained based on having ≥ 100 hits. Traits with vastly different distributions of selection coefficients may fail to pass this threshold, e.g. traits for which most heritability arises from strongly selected variants.

Limitations and future analyses. These limitations and others can be addressed in the future. For example, whole genome sequencing [60] would enable greater power for strongly selected variants. Meanwhile, methodologies that integrate over the full distribution of causal variants accounting for LD [11,19,36] may allow us to relax the reliance on genome-wide significant hits, thus increasing the power to identify weakly selected variants, more generally increasing the precision of our inference for both the TSD and SSD models, and allowing for the analysis of traits with fewer GWAS hits.

Future increases in sensitivity and applications of our inference to different kinds of traits may also warrant extensions of our evolutionary model. Notably, our current model is restricted to quantitative traits and is not immediately applicable to binary traits including diseases. One could imagine that variants that affect the risk of some complex diseases are selected to minimize that risk and are thus subject to purifying (directional) selection, but recent studies indicate that, at least common variation affecting them is predominated by the kind of pleiotropic stabilizing selection we assumed here [61,62]. With an appropriate adjustment of the model to fit data from case-control GWAS, we may find that complex diseases have similar architectures to those of the quantitative traits that we examined.

Our model and inference rely on the simplifying assumption that all the genetic variation in a given trait is affected by the same degree of pleiotropic selection (this degree was reflected in the dimension of the trait space; see, e.g., Eq 3 and Simons et al. 2018). We now have strong evidence that this assumption is violated in ways that affect the kinds of variants that are identified in GWAS [50]. Future extensions of the model may incorporate variation in the degree of pleiotropic selection, where this variation may also differ among complex traits.

Why do trait-specific scaling parameters vary? Future refinements notwithstanding, our evolutionary model with a shared distribution of selection effects fits the data from all 95 traits in our dataset remarkably well, indicating, as we have also confirmed, that differences in genetic architecture of genome-wide significant hits among highly polygenic traits are largely determined by the two trait-specific scaling factors: the mutational target size and heritability per site.

The mutational target size varies over 2 orders of magnitude among our 95 traits. While the estimates are novel, the variation among them is hardly surprising given the vast differences in the biology of these traits. These traits vary in being affected by few to many tissues and by the number and properties of genes and pathways that directly affect the trait in these tissues [59,63]. Moreover, the pathways that directly affect these traits plausibly differ in how buffered they are against genetic (and environmental) variation and in their modularity, plausibly reflective of the kinds of traits and of selection pressures over longer evolutionary timescales than the turnover time of heritable variation [64,65]. Our estimates of heritability are less variable, but still span an order of magnitude, and vary among different kinds of traits (S1 Table). Both of these observations have been known for almost a century, and yet the question about their causes remains largely open [23,37,38]. Our results indicate that other differences in architecture among traits are dwarfed by the variation caused by these two scaling parameters.

Outlook. Taking a step back, our results highlight that evolutionary thinking is essential to understanding of the findings emerging from human GWASs, and more generally, heritable variation in complex traits. This insight is consistent with long-standing thinking in the field, given that heritable variation in complex traits reflects the outcome of numerous genetic perturbations to the phenotype that passed through the sieve of evolution [23,38]. Specifically, the signal measured in GWASs reflects causal variants' contribution to heritability, which

depends on their effect on the trait under consideration, but also on their minor allele frequency, where the relationship between the two is mediated by natural selection. What is surprising, at least to us, is how far an analysis based on evolutionary modeling can go, in this case, showing that the genetic architectures of many highly polygenic quantitative traits are largely shared. This finding carries many implications about human GWASs and their applications, some of which we plan to explore elsewhere. Alongside other evidence, it also hints at underlying biology that largely remains to be discovered, plausibly relating to properties of gene regulatory networks. We think that a combination of evolutionary reasoning alongside a systems approach to gene regulation would move us closer to answering questions tracing back to the very beginning of the field of genetics about the mapping from genetic variation to phenotypes.

Acknowledgments

We thank Nick Barton, Graham Coop, Arbel Harpak, Magnus Nordborg, Roshni Patel, Molly Przeworski, Jeff Spence, Carl Veller, Laura Hayward, Himani Sachdeva and the Coop, Pritchard and Sella labs for conversations and comments.

Supporting information

S1 Note. Including:

Fig A. Geometric relationship between the *n*-dimensional effect and its projection on a single trait.

Fig B. Mutation-selection balance approximation.

Fig C. Histogram of *d* values for all GWAS hits used.

Fig D. MAF and z-score bound the selection coefficient.

Fig E. Demographic model.

Fig F. CDFs of the inferred distribution of selection coefficients for all 95 traits.

Fig G. Tuning the strength of the penalty.

Fig H. Choosing the number of knots.

Fig I. Lowered sample sizes.

Fig J. Heritability estimates.

Fig K. Proportion of rare variants.

Fig L. Inference on simulated datasets.

Fig M. Inference on a non-SSD dataset.

Fig N. SSD results for 15 nearly independent traits.

Fig O. Summaries of GWAS hits for the 95 traits.

Fig P. Relative Uncertainty.

Fig Q. Maximal correlation.

Fig R. Effect of assortative mating.

Fig S. The genetic architecture and inference for simulated architectures.

Fig T. Simulation with 9,834 hits.

Fig U. Schematic of mutation origin on terminal branch.

Fig V. Distribution of allele ages.

Fig W. Heuristic estimators of allele ages.

Fig X. Prediction of allele ages with shifted distributions of selection coefficients.

Table A. Outlier SNPs.

Table B. Model and inference assumptions.
(PDF)

S1 Table. Inferred target size and heritability for all 95 traits.
(CSV)

S2 Table. Expanded table of outlier SNPs.
(XLSX)

Author contributions

Conceptualization: Yuval B. Simons, Jonathan K. Pritchard, Guy Sella.

Data curation: Yuval B. Simons, Courtney J. Smith.

Formal analysis: Yuval B. Simons.

Funding acquisition: Yuval B. Simons, Jonathan K. Pritchard, Guy Sella.

Investigation: Yuval B. Simons, Hakhamanesh Mostafavi, Huisheng Zhu.

Methodology: Yuval B. Simons.

Software: Yuval B. Simons, Hakhamanesh Mostafavi, Huisheng Zhu.

Supervision: Jonathan K. Pritchard, Guy Sella.

Validation: Yuval B. Simons.

Visualization: Yuval B. Simons.

Writing – original draft: Yuval B. Simons, Guy Sella.

Writing – review & editing: Jonathan K. Pritchard, Guy Sella.

References

1. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, et al. A brief history of human disease genetics. Nature. 2020;577(7789):179–89. https://doi.org/10.1038/s41586-019-1879-7 PMID: 31915397
2. Loh P-R, Bhatia G, Gusev A, Finucane HK, Bulik-Sullivan BK, Pollack SJ, et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. Nat Genet. 2015;47(12):1385–92. https://doi.org/10.1038/ng.3431
3. Shi H, Kichaev G, Pasaniuc B. Contrasting the genetic architecture of 30 complex traits from summary association data. Am J Hum Genet. 2016;99(1):139–53. https://doi.org/10.1016/j.ajhg.2016.05.013 PMID: 27346688
4. Timpson NJ, Greenwood CMT, Soranzo N, Lawson DJ, Richards JB. Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat Rev Genet. 2018;19(2):110–24. https://doi.org/10.1038/nrg.2017.101 PMID: 29225335

5. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet. 2019;51(9):1339–48. https://doi.org/10.1038/s41588-019-0481-0 PMID: 31427789
6. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat Genet. 2018;50(11):1593–9. https://doi.org/10.1038/s41588-018-0248-z PMID: 30349118
7. Yang J, Bakshi A, Zhu Z, Hemani G, Vinkhuyzen AAE, Lee SH, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. Nat Genet. 2015;47(10):1114–20. https://doi.org/10.1038/ng.3390 PMID: 26323059
8. Zhang Y, Qi G, Park J-H, Chatterjee N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. Nat Genet. 2018;50(9):1318–26. https://doi.org/10.1038/s41588-018-0193-x PMID: 30104760
9. Sinnott-Armstrong N, Naqvi S, Rivas M, Pritchard JK. GWAS of three molecular traits highlights core genes and pathways alongside a highly polygenic background. Elife. 2021;10:e58615. https://doi.org/10.7554/eLife.58615 PMID: 33587031
10. Ge T, Chen C-Y, Neale BM, Sabuncu MR, Smoller JW. Correction: phenome-wide heritability analysis of the UK Biobank. PLoS Genet. 2018;14(2):e1007228. https://doi.org/10.1371/journal.pgen.1007228 PMID: 29425192
11. Hou K, Burch KS, Majumdar A, Shi H, Mancuso N, Wu Y, et al. Accurate estimation of SNP-heritability from biobank-scale data irrespective of genetic architecture. Nat Genet. 2019;51(8):1244–51. https://doi.org/10.1038/s41588-019-0465-0 PMID: 31358995
12. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42(7):565–9. https://doi.org/10.1038/ng.608 PMID: 20562875
13. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291–5. https://doi.org/10.1038/ng.3211 PMID: 25642630
14. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. Cell. 2017;169(7):1177–86.
15. Zhang Y, Qi G, Park J-H, Chatterjee N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. Nat Genet. 2018;50(9):1318–26. https://doi.org/10.1038/s41588-018-0193-x PMID: 30104760
16. Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S, et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. Nat Commun. 2019;10(1):2417. https://doi.org/10.1038/s41467-019-10310-0 PMID: 31160569
17. O'Connor LJ, Schoech AP, Hormozdiari F, Gazal S, Patterson N, Price AL. Extreme polygenicity of complex traits is explained by negative selection. Am J Hum Genet. 2019;105(3):456–76. https://doi.org/10.1016/j.ajhg.2019.07.003 PMID: 31402091
18. O'Connor LJ. The distribution of common-variant effect sizes. Nat Genet. 2021;53(8):1243–9. https://doi.org/10.1038/s41588-021-00901-3 PMID: 34326547
19. Spence JP, Sinnott-Armstrong N, Assimes T, Pritchard JK. A flexible modeling and inference framework for estimating variant effect sizes from GWAS summary statistics. bioRxiv. 2022.
20. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236–41. https://doi.org/10.1038/ng.3406 PMID: 26414676
21. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48(7):709–17. https://doi.org/10.1038/ng.3570 PMID: 27182965
22. Smith CJ, Sinnott-Armstrong N, Cichońska A, Julkunen H, Fauman EB, Würtz P, et al. Integrative analysis of metabolite GWAS illuminates the molecular basis of pleiotropy and genetic correlation. eLife. 2022:e79348. https://doi.org/10.7554/eLife.79348
23. Sella G, Barton NH. Thinking about the evolution of complex traits in the era of genome-wide association studies. Annu Rev Genomics Hum Genet. 2019;20:461–93. https://doi.org/10.1146/annurev-genom-083115-022316 PMID: 31283361
24. Gazal S, Finucane HK, Furlotte NA, Loh PR, Palamara PF, Liu X, et al. Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. Nature Genetics. 2017;49(10):1421–7.
25. Zeng J, de Vlaming R, Wu Y, Robinson MR, Lloyd-Jones LR, Yengo L, et al. Signatures of negative selection in the genetic architecture of human complex traits. Nat Genet. 2018;50(5):746–53. https://doi.org/10.1038/s41588-018-0101-4 PMID: 29662166

26. Speed D, Holmes J, Balding DJ. Evaluating and improving heritability models using summary statistics. Nat Genet. 2020;52(4):458–62. https://doi.org/10.1038/s41588-020-0600-y PMID: 32203469

27. Simons YB, Bullaughey K, Hudson RR, Sella G. A population genetic interpretation of GWAS findings for human quantitative traits. PLoS Biol. 2018;16(3):e2002985. https://doi.org/10.1371/journal.pbio.2002985

28. Weiner DJ, Nadig A, Jagadeesh KA, Dey KK, Neale BM, Robinson EB, et al. Polygenic architecture of rare coding variation across 394,783 exomes. Nature. 2023;614(7948):492–9. https://doi.org/10.1038/s41586-022-05684-z PMID: 36755099

29. Speed D, Hemani G, Johnson MR, Balding DJ. Improved heritability estimation from genome-wide SNPs. Am J Hum Genet. 2012;91(6):1011–21. https://doi.org/10.1016/j.ajhg.2012.10.010 PMID: 23217325

30. Speed D, Cai N, UCLEB Consortium, Johnson MR, Nejentsev S, Balding DJ. Reevaluation of SNP heritability in complex human traits. Nat Genet. 2017;49(7):986–92. https://doi.org/10.1038/ng.3865 PMID: 28530675

31. Keightley PD, Hill WG. Variation maintained in quantitative traits with mutation–selection balance: pleiotropic side-effects on fitness traits. Proceedings of the Royal Society of London Series B: Biological Sciences. 1990;242(1304):95–100.

32. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases?. Am J Hum Genet. 2001;69(1):124–37. https://doi.org/10.1086/321272 PMID: 11404818

33. Eyre-Walker A. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. Proceedings of the National Academy of Sciences. 2010;107(suppl_1):1752–6.

34. Agarwala V, Flannick J, Sunyaev S, GoT2D Consortium, Altshuler D. Evaluating empirical bounds on complex disease genetic architecture. Nat Genet. 2013;45(12):1418–27. https://doi.org/10.1038/ng.2804 PMID: 24141362

35. Mancuso N, Rohland N, Rand KA, Tandon A, Allen A, Quinque D, et al. The contribution of rare variation to prostate cancer heritability. Nat Genet. 2016;48(1):30–5. https://doi.org/10.1038/ng.3446 PMID: 26569126

36. Zeng J, Xue A, Jiang L, Lloyd-Jones LR, Wu Y, Wang H, et al. Widespread signatures of natural selection across human complex traits and functional genomic categories. Nat Commun. 2021;12(1):1164. https://doi.org/10.1038/s41467-021-21446-3 PMID: 33608517

37. Johnson T, Barton N. Theoretical models of selection and mutation on quantitative traits. Philos Trans R Soc Lond B Biol Sci. 2005;360(1459):1411–25. https://doi.org/10.1098/rstb.2005.1667 PMID: 16048784

38. Walsh B, Lynch M. Evolution and selection of quantitative traits. Oxford University Press; 2018.

39. Lynch M, Walsh B, et al. Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer; 1998.

40. Fisher RA. The genetical theory of natural selection. Oxford: Clarendon Press; 1930.

41. Orr HA. The genetic theory of adaptation: a brief history. Nat Rev Genet. 2005;6(2):119–27. https://doi.org/10.1038/nrg1523 PMID: 15716908

42. Wright S. The analysis of variance and the correlations between relatives with respect to deviations from an optimum. Journ of Genetics. 1935;30(2):243–56. https://doi.org/10.1007/bf02982239

43. Robertson A. The effect of selection against extreme deviants based on deviation or on homozygosis. J Genet. 1956;54(2):236–48. https://doi.org/10.1007/bf02982779

44. Charlesworth B, Charlesworth D. Elements of evolutionary genetics. Roberts and Company; 2010.

45. Speidel L, Forest M, Shi S, Myers SR. A method for genome-wide genealogy estimation for thousands of samples. Nat Genet. 2019;51(9):1321–9. https://doi.org/10.1038/s41588-019-0484-x PMID: 31477933

46. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44(4):369–75, S1-3. https://doi.org/10.1038/ng.2213 PMID: 22426310

47. Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in human populations. Bioinformatics. 2016;32(2):283–5. https://doi.org/10.1093/bioinformatics/btv546 PMID: 26395773

48. Claussnitzer M, Dankel SN, Kim K-H, Quon G, Meuleman W, Haugen C, et al. FTO obesity variant circuitry and adipocyte browning in humans. N Engl J Med. 2015;373(10):895–907. https://doi.org/10.1056/NEJMoa1502214 PMID: 26287746

49. Sobreira DR, Joslin AC, Zhang Q, Williamson I, Hansen GT, Farris KM, et al. Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of IRX3 and IRX5. Science. 2021;372(6546):1085–91. https://doi.org/10.1126/science.abf1008 PMID: 34083488

50. Spence JP, Mostafavi H, Ota M, Milind N, Gjorgjieva T, Smith CJ. Specificity, length, and luck: how genes are prioritized by rare and common variant association studies. bioRxiv. 2024. https://www.biorxiv.org/content/early/2024/12/16/2024.12.12.628073

51. Kiezun A, Pulit SL, Francioli LC, van Dijk F, Swertz M, Boomsma DI, et al. Deleterious alleles in the human genome are on average younger than neutral alleles of the same frequency. PLoS Genet. 2013;9(2):e1003301. https://doi.org/10.1371/journal.pgen.1003301 PMID: 23468643

52. Marigorta UM, Navarro A. High trans-ethnic replicability of GWAS results implies common causal variants. PLoS Genet. 2013;9(6):e1003566. https://doi.org/10.1371/journal.pgen.1003566 PMID: 23785302

53. Li YR, Keating BJ. Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. Genome Med. 2014;6(10):91. https://doi.org/10.1186/s13073-014-0091-5 PMID: 25473427

54. Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. Proc Natl Acad Sci U S A. 2016;113(20):5652–7. https://doi.org/10.1073/pnas.1514696113 PMID: 27140627

55. Ward LD, Kellis M. Evidence of abundant purifying selection in humans for recently acquired regulatory functions. Science. 2012;337(6102):1675–8. https://doi.org/10.1126/science.1225057 PMID: 22956687

56. Rands CM, Meader S, Ponting CP, Lunter G. 8.2% of the Human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. PLoS Genet. 2014;10(7):e1004525. https://doi.org/10.1371/journal.pgen.1004525 PMID: 25057982

57. Gulko B, Hubisz MJ, Gronau I, Siepel A. A method for calculating probabilities of fitness consequences for point mutations across the human genome. Nat Genet. 2015;47(3):276–83. https://doi.org/10.1038/ng.3196 PMID: 25599402

58. Finucane HK, Reshef YA, Anttila V, Slowikowski K, Gusev A, Byrnes A, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat Genet. 2018;50(4):621–9. https://doi.org/10.1038/s41588-018-0081-4 PMID: 29632380

59. Liu X, Li YI, Pritchard JK. Trans effects on gene expression can drive omnigenic inheritance. Cell. 2019;177(4):1022-1034.e6. https://doi.org/10.1016/j.cell.2019.04.014 PMID: 31051098

60. Wainschtein P, Jain D, Zheng Z, TOPMed Anthropometry Working Group, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, Cupples LA, et al. Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. Nat Genet. 2022;54(3):263–73. https://doi.org/10.1038/s41588-021-00997-7 PMID: 35256806

61. Koch E, Connally NJ, Baya N, Reeve MP, Daly M, Neale B, et al. Genetic association data are broadly consistent with stabilizing selection shaping human common diseases and traits. bioRxiv. 2024.

62. Berg JJ, Li X, Riall K, Hayward LK, Sella G. Mutation-selection-drift balance models of complex diseases. bioRxiv. 2025.

63. Udler MS. Type 2 diabetes: multiple genes, multiple diseases. Curr Diab Rep. 2019;19(8):55. https://doi.org/10.1007/s11892-019-1169-7 PMID: 31292748

64. de Visser JAGM, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: evolution and detection of genetic robustness. Evolution. 2003;57(9):1959–72. https://doi.org/10.1111/j.0014-3820.2003.tb00377.x PMID: 14575319

65. Wagner GP, Pavlicev M, Cheverud JM. The road to modularity. Nat Rev Genet. 2007;8(12):921–31. https://doi.org/10.1038/nrg2267 PMID: 18007649