A model for the genetic architecture of quantitative traits under stabilizing selection

Yuval B. Simons\textsuperscript{a,1}, Kevin Bullaughey\textsuperscript{b}, Richard R. Hudson\textsuperscript{b}, and Guy Sella\textsuperscript{a,1}

\textsuperscript{a} Department of Biological Sciences, Columbia University, New York, NY 10027;

\textsuperscript{b} Department of Ecology & Evolution, University of Chicago, Chicago, Illinois 60637;

\textsuperscript{1} To whom correspondence should be addressed: yuval.simons@columbia.edu, gs2747@columbia.edu

Abstract

Genome-wide association studies (GWAS) in humans are revealing the genetic architecture of biomedical, life history and anthropomorphic traits, i.e., the frequencies and effect sizes of variants contributing to heritable variation in a trait. To interpret these findings, we need to understand how genetic architecture is shaped by basic population genetics processes—notably, by mutation, natural selection and genetic drift. Because many quantitative traits are subject to stabilizing selection and genetic variation that affects one trait often affects many others, we model the genetic architecture of a focal trait that arises under stabilizing selection in a multi-dimensional trait space. We solve the model for the phenotypic distribution and allelic dynamics at steady state and derive robust, closed form solutions for summaries of genetic architecture. Our results suggest that the distribution of genetic variance among the loci discovered in GWAS take a simple form that depends on one evolutionary parameter, and provide a simple interpretation for missing heritability and why it varies among traits. We test our predictions against the results of GWAS for height and body mass index (BMI) and find that they fit the data well, allowing us to make inferences about the degree of pleiotropy and the mutational target size. Our findings help to understand why GWAS for height explain more of the heritable variance than similarly-sized GWAS for BMI, and to predict how future increases in sample size will translate into explained heritability.
Much of the phenotypic variation in human populations, including variation in morphological, life history and biomedical traits, is “quantitative”, meaning that heritable variation in the trait is largely due to small contributions from many genetic variants segregating in the population (1, 2). Quantitative traits have been studied since the birth of biometrics over a century ago (1-3), but only in the past decades have technological advances made it possible to systematically dissect their genetic basis (4-6). Notably, since 2007, genome-wide association studies (GWAS) in humans have led to the identification of thousands of variants reproducibly associated with hundreds of quantitative traits, including susceptibility to a wide variety of diseases (4). While still ongoing, these studies already provide important insights into the genetic architecture of quantitative traits, i.e., the number of variants that contribute to heritable variation and their frequencies and effect sizes.

Perhaps the most striking observation to emerge from these studies is that, despite the large sample size of many GWAS, all variants significantly associated with any given trait typically account for less (often much less) than 25% of the narrow sense heritability (4, 7, 8) (but see (9)). (Henceforth, we use “heritability” to refer to narrow sense heritability). While many factors have been hypothesized to contribute to the “missing heritability” (7, 8, 10-14), the most straightforward explanation and the emerging consensus is that much of the heritable variation derives from variants with frequencies that are too low or effect sizes that are too small for current studies to detect. Comparisons among traits also suggest that there are substantial differences in architectures. For example, recent meta-analyses GWAS uncovered seven times as many variants for height (697) than for body mass index (97), and together the variants for height account for more than four times the heritable variance for body mass index (~20% vs. ~3-5%, respectively), despite the smaller sample size of the GWAS (250k compared to 340k, respectively) (15, 16).

These first glimpses underscore the need for theory that explains how genetic architectures are shaped in the course of evolution and why they might differ among traits. Such theory should allow us to use GWAS findings to make inferences about underlying evolutionary parameters, helping to answer enduring questions about the processes that maintain phenotypic variation in quantitative traits (5, 17). From a practical perspective, it
will help to interpret GWAS findings and specifically it may inform solutions to the “missing heritability” problem (18-21). In so doing, it will inform the design of future mapping studies and phenotypic predictions (18, 20-22).

Development of such theory can be guided by empirical observations and first-principles considerations. First, we know that the architecture of a trait arises from genetic and population genetic processes. New mutations affecting a trait arise at a rate that depends on its “mutational target size” (i.e., the number of sites at which a mutation would affect the trait). Once they arise, the trajectories of variants through the population are determined by the interplay between genetic drift, demographic processes, and natural selection acting on them. These processes determine the number and frequencies of segregating variants underlying variation in the trait. The genetic architecture further depends on the relationship between the selection on variants and their effects on the trait. Notably, selection on variants depends not only on their effect on the focal trait but also on their pleiotropic effects on other traits. We would therefore expect both direct and pleiotropic selection to shape the joint distribution of allele frequencies and effect sizes.

Multiple lines of evidence suggest that many quantitative traits are subject to stabilizing selection, i.e., selection favoring an intermediate trait value (5, 23-27). For instance, a decline in fitness components (e.g., viability and fecundity) is observed with displacement from mean values for a variety of traits in human populations (28-30), in other species in the wild (31, 32) and in experimental manipulations (31, 33). While less is known about the selection acting on complex diseases, some likely reflect large deviations from optimal values of underlying traits under stabilizing selection. What remains unclear is the extent to which stabilizing selection is acting directly on variation in a given trait or results from pleiotropic effects of this variation on other traits.

Yet other lines of evidence suggest that pleiotropy is pervasive. For one, theoretical considerations about the variance in fitness in natural populations and its accompanying genetic load suggest that only a moderate number of independent traits can be effectively selected on at the same time (34). Thus, the aforementioned relationships between the value of a focal trait and fitness are likely heavily affected by the pleiotropic effects of genetic variation on other traits (26, 34-36). Second, many of the variants detected in
human GWAS have been found to be associated with more than one trait (37-41). For example, a recent analysis of GWAS revealed that variants that delay the age of menarche in women tend to delay the age of voice drop in men, decrease body mass index, increase adult height, and decrease risk of male pattern baldness (37). More generally, the extent of pleiotropy revealed by GWAS appears to be increasing rapidly with improvements in the power and methodology (37, 42-44). Such considerations point to the potential importance of pleiotropic selection on quantitative genetic variation for many traits.

Moreover, the discoveries emerging from human GWAS suggest that genetic variance is predominated by additive contributions from numerous variants with small effect sizes. Specifically, most or all of the heritability explained in GWAS of many traits derives from variants with squared effect sizes that are substantially smaller than the total genetic variance (e.g. (15, 16, 45, 46)). Moreover, statistical quantifications of the genetic variance tagged by genotyping suggest that such variants may account for the bulk of heritable variance in many traits (9, 47, 48). Other analyses of GWAS results suggest that non-additive interactions, e.g., dominance and epistasis, have only minor contributions to heritable variance, consistent with theory and other lines of evidence (49-55). Notably, considerable efforts to detect epistatic interactions in human GWAS have, by and large, came up empty-handed (9, 56, 57), with few counter-examples mostly involving variants in the MHC region (52, 56, 58, 59). Motivated by these considerations, we model how direct and pleiotropic stabilizing selection shape the genetic architecture of continuous, quantitative traits by considering additive variants of small effects and assuming that together they account for most of the heritable variance.

There has been relatively little theoretical work aimed at relating population genetics processes with the results emerging from GWAS. Moreover, the few existing models have reached divergent predictions about genetic architecture, largely because of making different assumptions about the effects of pleiotropy. Pritchard (20) considered the "purely pleiotropic" extreme, in which selection on variants is independent of their effect on the trait being considered (he focused on disease susceptibility). In this case, we would expect the largest contribution to genetic variance in a trait to come from mutations that have large effect sizes but are also weakly selected or neutral, allowing them to ascend to higher
frequencies than other mutations. Other studies have considered the opposite extreme, in which selection on variants stems entirely from their effect on the trait under consideration (27, 60-64), and have shown that the greatest contribution to genetic variance arises from strongly selected mutations (61, 62) (we return to this case below).

In practice, we would expect most traits to fall somewhere in between these two extremes. While there are compelling reasons to believe that quantitative genetic variation is highly pleiotropic, the effects of variants on different traits are likely to be correlated. If so, even if a given trait is not subject to selection, variants that have a large effect on it will also tend to have larger effects on traits that are under selection (e.g., by causing a large perturbation to pathways that affect multiple traits) (36). Motivated by such considerations, Eyre-Walker (2010) and Caballero et al. (2015) considered models in which the correlation between the strength of selection on an allele and its effect size can vary between the purely pleiotropic and direct selection extremes (also cf. 65). These models diverge in their predictions about architecture, however. Assuming, as seems plausible, an intermediate correlation between the strength of selection and effect size, Eyre-Walker finds that genetic variance should be dominated by strongly selected mutations, whereas Caballero et al. conclude the greatest contribution should arise from weakly selected ones (19). Their conclusions differ because of how they chose to model the relationship between selection and effect size, a choice based largely on mathematical convenience. We approach this problem by explicitly modeling stabilizing selection on multiple traits, thereby learning, rather than assuming, the relationship between selection and effect sizes.

The Model

We model an individual’s phenotype as a vector in an n-dimensional Euclidian space, in which each dimension corresponds to an additive, continuous, quantitative trait. We focus on the architecture of one of these traits (say, the 1st dimension), where the total number of traits parameterizes pleiotropy. Fitness is assumed to decline with distance from the optimal phenotype positioned at the origin, thereby introducing stabilizing selection into the model. Specifically, we assume that absolute fitness takes the form
\[ W(\bar{r}) = \exp \left( -\frac{r^2}{2w^2} \right), \]  

where \( \bar{r} \) is the \((n\text{-dimensional})\) phenotype, \( r = \|\bar{r}\| \) is the distance from the origin and \( w \) parameterizes the strength of stabilizing selection. However, we later show that the specific form of the fitness function doesn’t matter. Moreover, the additive environmental contribution to the phenotype can be absorbed into \( w \) (see (66) and SI Section 1.1); we therefore consider only the genetic contribution.

We further assume the standard population dynamic for a diploid, panmictic population of constant size \( N \), including Wright-Fisher sampling with viability selection (as described by Eq. (1)), mutation, free recombination and Mendelian segregation. Mutations affecting the phenotype are introduced according to an infinite sites model. The number per gamete, per generation follows a Poisson distribution with mean \( U \), reflecting both the mutational target size and the mutation rate per site. The size of mutations in the \( n\) dimensional trait space, \( a^2 (= \|\vec{a}\|^2) \), is drawn from some distribution, assuming only that \( a^2 \ll w^2 \). We later show that this requirement is equivalent to the standard assumption about selection coefficients satisfying \( s \ll 1 \) (also cf. Supplementary Section 4.3). The directions of mutations are assumed to be uniformly (isotropically) distributed, although we later show that our results are robust to relaxing this assumption. An individual’s phenotype can then be obtained adding up the phenotypic effects of his or her alleles, i.e.,

\[ \bar{r} = \frac{1}{2} \sum_{l=1}^{L} g_l \vec{a}_l, \]  

where \( L \) is the number of sites with fixed or segregating mutations, \( \vec{a}_l \) is the phenotypic difference between homozygotes for the ancestral and derived alleles at site \( l \) (because of the infinite sites assumption, there are at most two alleles) and \( g_l = 0, 1 \) or \( 2 \) is the number of derived alleles at site \( l \).

**Results**

**The phenotypic distribution.** In the first three sections, we develop the tools that we later use to study genetic architecture. We start by considering the equilibrium distribution of phenotypes in the population and generalize previous results for the case with a single trait.
Under biologically sensible conditions on the rate and size of mutations (i.e., when $1 \gg U \gg 1/2N$ and $a^2 \ll w^2$; SI Section 4), this distribution is well approximated by a tight multivariate normal centered at the optimum, namely by the probability density:

$$f(\mathbf{r}) = \frac{1}{(2\pi\sigma^2)^\frac{D}{2}} \exp \left( -\frac{\mathbf{r}^2}{2\sigma^2} \right)$$

(3)

where the variance of the distribution satisfies $\sigma^2 \ll w^2$ (see SI Section 4). Intuitively, the phenotypic distribution is normal because it derives from additive and (approximately) i.i.d. contributions from many segregating sites. The distribution stays tightly concentrated around the optimum because stabilizing selection becomes stronger with increasing displacement from it, and because the population responds to such selection by minor changes to allele frequencies at many segregating sites, rapidly offsetting the displacement.

With phenotypes close to the optimum, only the curvature of the fitness function at the optimum (i.e., the multi-dimensional second derivative) affects the selection acting on individuals. In addition, it is always possible to choose an orthonormal coordinate system centered at the optimum, in which the trait under consideration varies along the first coordinate and a unit change in other traits (in other coordinates) near the optimum have the same effect on fitness. These considerations suggest that the equilibrium behavior is insensitive to our choice of fitness function. Moreover, in the SI (Section 5), we show that perturbations of the population mean from the optimum are rapidly offset by minor changes to allele frequencies at numerous sites, and that this behavior lends robustness to the equilibrium dynamics with respect to the presence of major loci, changes in the optimal phenotype over time, and moderate asymmetries in the mutational distribution.

**Allelic dynamic.** Next, we consider the dynamics at a segregating site, and generalize previous results for the case with a single trait (62-64). The dynamics can be described in terms of the first two moments of change in allele frequency in a single generation (cf. (67)). For an allele with phenotypic effect $\bar{a}$ and frequency $\bar{q}$, the distribution of contributions to the phenotype from the genetic background, $\bar{R}$, follows from the
equilibrium phenotypic distribution (Eq. 3) and is well approximated by probability density

\[ f(R|\bar{a}, q) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{(R+q\bar{a})^2}{2\sigma^2}\right) \]  

(4)

By averaging the fitness of the three genotypes over the distribution of genetic backgrounds, we find that the first moment is well approximated by

\[ E(\Delta q) \approx \frac{a^2}{4w^2} pq \left(q - \frac{1}{2}\right), \]  

(5)

assuming that \(a^2\) and \(\sigma^2 \ll w^2\) (cf. SI Section 4). By the same token, we find that

\[ V(\Delta q) \approx \frac{pq}{2N}, \]  

(6)

which is the standard second moment with genetic drift.

The functional form of the first moment is equivalent to that of the standard viability selection model with under-dominance. This result is a hallmark of stabilizing selection on (additive) quantitative traits: with the population mean at the optimum, the dynamics at different sites are decoupled and selection at a given site reduces the phenotypic variance (i.e., \(\frac{1}{2}a^2pq\)), thereby pushing rare alleles to loss. Comparison with the standard viability selection model shows that the selection coefficient in our model is \(s=a^2/4w^2\), or \(S=2Ns=Na^2/2w^2\) in scaled units. In other words, the selection acting on an allele is proportional to its size-squared in the \(n\)-dimensional trait space (where \(w\) translates effect size into units of fitness).

**The relationship between selection and effect size.** The statistical relationship between the strength of selection acting on mutations and their effect on a given trait follows from the aforementioned geometric interpretation of selection. Specifically, all mutations with a given selection coefficient, \(s\), lie on a \((n-1)\) - dimensional hypersphere with radius \(a = 2w\sqrt{s}\), and any given mutation satisfies

\[ s = \frac{1}{4w^2} a^2 \approx \frac{1}{4w^2} \sum_{i=1}^{n} a_i^2, \]  

(7)
where $a_i$ is the allele's effect on the $i$-th trait (Fig. 1A). Our assumption that mutation is isotropic then implies that the probability density of mutations on the hypersphere is uniform.

![Diagram](image)

**Fig. 1.** The distribution of effect sizes corresponding to a given selection coefficient. (a) Mutations with selection coefficient, $s$, lie on a $(n-1)$-dimensional hypersphere with radius $r = 2w\sqrt{s}$; and the probability of such mutations with effect size $a_1$ is proportional to the volume of the $(n-2)$-dimensional cross section of that hyper-sphere with projection $a_1$. (b) The distribution of effect sizes corresponding to a given selection coefficient, measured in units of the distribution's standard deviation.

The distribution of effect sizes on a focal trait, $a_1$, corresponding to a given selection coefficient, $s$, follows. Given that mutation is symmetric in any given trait, $E(a_1|s) = 0$, and given that it is symmetric among traits,

$$E(a_1^2|s) = a^2/n = (2w^2/n)s. \quad (8)$$

More generally, the probability density corresponding to an effect size $a_1$ is proportional to the volume of the $(n-2)$-dimensional cross section of that hyper-sphere with projection $a_1$ (Fig. 1A). For a single trait, this implies that $a_1 = \pm a$ with probability $1/2$, and for $n > 1$, it implies the probability density

$$\varphi_n(a_1|a) = \frac{\Gamma(n/2)\Gamma((n-1)/2)}{\sqrt{n/2}} \frac{1}{\sqrt{2\pi(a^2/n)}} \left(1 - \frac{1}{n} \frac{a_1^2}{(a^2/n)}\right)^{n-3} \quad (9)$$

(cf. SI Section 1.2). Intriguingly, when the number of traits $n$ increases, this density approaches a normal distribution, i.e.,
\[ \frac{a_i}{\sqrt{a^2/n}} \sim N(0,1). \]  

(10)

This limit is already well approximated for a moderate number of traits (e.g., \( n=10 \); Fig. 1B). The limit behavior also holds when we relax the assumption of isotropic mutation. This generalization is an important one, because having chosen a parameterization of traits in which the fitness function near the optimum is isotropic, we can no longer assume that the distribution of mutations is also isotropic (68). Specifically, mutations might tend to have larger effects on some traits than on others and their effects on different traits might be correlated. In SI Section 5, we show, however, that excluding pathological cases, the limit distribution (Eq. 10) also holds for anisotropic mutation. To this end, we introduce the concept of an effective number of traits, which can take any real value \( \geq 1 \) and is defined as the number of traits, with expected fitness effect equal to that of the focal trait, required to satisfy Eq. 8. In the limit of high pleiotropy, the selection strength on a mutation and the magnitude of its effect on a focal trait are correlated, implying that the kind of "purely pleiotropic" extreme postulated in previous work cannot arise under our model (19-21). Mounting evidence that genetic variation is highly pleiotropic (see Introduction) alongside the robustness of our model raise the intriguing possibility that this limit form applies quite generally.

**Genetic architecture.** We can now derive closed forms for summaries of genetic architecture (cf. SI Section 2). For mutations with a given selection coefficient, the frequency distribution follows from the diffusion approximation based on the first two moments of change in allele frequency (Eq. 5 and 6; (67)), and the distribution of effect sizes follows from the geometric considerations of the previous section. Conditional on the selection coefficient, these distributions are independent and therefore the joint distribution of frequency and effect size equals their product. Summaries of architecture can be expressed as expectations over the joint distribution of frequencies and effect sizes for a given selection coefficient, and then weighted according to the distribution of selection coefficients. Given that we know little about the distribution of selection
coefficients of mutations affecting quantitative traits, we examine how summaries depend on the strength of selection.

**Expected variance per site.** We focus on the distribution of additive genetic variance among sites, a central feature of architecture that is key to connecting our model with GWAS results. We start by considering how selection affects the expected contribution of a site to additive genetic variance in a focal trait. We include monomorphic sites in the expectation, such that the expected total variance is given by the product of the expectation per-site and the population mutation rate (i.e., $2NU$). Under the infinite sites assumption, sites can be monomorphic or bi-allelic and their expected contribution is

$$E(\frac{1}{2}a_1^2pq|S) = E(a_1^2|S)E(\frac{1}{2}pq|S) = \frac{2w^2}{nN}S E(\frac{1}{2}pq|S)$$  \hspace{1cm} (11)

(expresssed in terms of the scaled selection coefficient $S$). Thus, the degree of pleiotropy only affects the expectation through a multiplicative constant. This effect of pleiotropy is not identifiable from data, because even if we could measure genetic variance in units of fitness (e.g., as opposed to units of the total phenotypic variance), we still wouldn’t be able to distinguish between $w$ and $n$. We instead focus on the proportional effect of selection on the contribution to variance, which is insensitive to the degree of pleiotropy.

The proportional contribution to genetic variance as a function of the selection coefficient was described by Keighley and Hill (in the one dimensional case; (62)) and is shown in Fig. 2A. When selection is strong ($S\gg1$), its effect on allele frequency (which scales with $1/S$) is canceled out by its relationship with the effect size, yielding a constant contribution to genetic variance per site, $\nu_s=2w^2/(nN)$, regardless of the selection coefficient (SI Section 3; Fig. 2A). Henceforth, we measure genetic variance relative to $\nu_s$. When selection is effectively neutral ($S \ll 1$) and thus too weak to affect the allele frequency, the expected contribution of a site to genetic variance scales with the effect size and is $\frac{1}{2}S$, which is considerably lower than under strong selection (SI Section 3; Fig. 2A). In between these selection regimes, allele frequency is affected by under-dominance, resulting in a more complex dependency on the selection coefficient (SI Section 3). The maximal contribution to genetic variance per site occurs within this range (at $S \approx 10$) and is slightly higher (by $\sim30\%$) than under strong selection (Fig. 2A). Henceforth, we refer to this selection regime
as intermediate, to distinguish it from the considerably weaker selection in the nearly neutral regime. These results suggest that effectively neutral sites should contribute much less to genetic variance than intermediate and strongly selected ones (61, 62).

**Fig. 2.** The distribution of additive genetic variance among sites. In (a), we plot the expected contribution as a function of the scaled selection coefficient. In (b) & (c), we show the proportion of genetic variance stemming from sites that contribute more than the value on the x-axis, for a single trait (b) and in the pleiotropic limit (c).

**Distribution of variance among sites.** Next, we consider how genetic variance is distributed among sites with a given selection coefficient. We focus on the distribution among segregating sites (including monomorphic effects would just add a point mass at 0). This distribution is especially relevant to interpreting the results of GWAS, because, to a first approximation, a study will detect only sites with contributions to variance exceeding a certain threshold (which decreases as the study size increases; see Discussion). We therefore represent the distribution in terms of the proportion of genetic variance, $G(v)$, arising from sites whose contribution to genetic variance exceeds a threshold $v$.

We begin with the case of a single trait ($n=1$), in which selection on an allele determines its effect size (Fig 2B). When selection is strong ($S>>1$), the proportion of genetic variance exceeding a threshold $v$ is also insensitive to the selection coefficient and takes a simple form, with

$$ G(v) = \exp(-2v) $$  \hspace{1cm} (12)

(SI Section 3). In contrast, in the effectively neutral range ($S \ll 1$),

$$ G(v) = \sqrt{1 - v/v_{max}} $$  \hspace{1cm} (13)
where the dependency on the selection coefficient enters through \( v_{max} = \frac{1}{8} S \), which is the maximal contribution to variance for a site with selection coefficient \( S \) that is obtained with an allele frequency of \( \frac{1}{2} \) (SI Section 3). In the intermediate selection regime, \( G(v) \) is also intermediate and takes a more elaborate functional form (SI Section 3). These results suggest how genetic variance would be distributed among sites given any distribution of selection coefficients (Fig 2B): starting from sites that contribute the most, the distribution would at first be dominated by contributions from strongly selected sites, then the intermediate selected sites would begin to contribute, whereas effectively neutral sites would enter only for \( v < \frac{1}{8} S \ll 1 \).

The effect of pleiotropy is to cause sites with a given selection coefficient to have a distribution of effect sizes on the focal trait, thereby increasing the contribution to genetic variance of some sites and decreasing it for others. In SI Section 3, we show that increasing the degree of pleiotropy, \( n \), increases the proportion of genetic variance, \( G(v) \), for any threshold, \( v \), regardless of the distribution of selection coefficients (Fig. S2). When variation in a trait is sufficiently pleiotropic for the distribution of effect sizes to attain the limit form (Eq. 10):

\[
G(v) = (1 + 2\sqrt{v}) \exp(-2\sqrt{v})
\]

for strongly selected sites and

\[
G(v) = \exp(-4v/S)
\]

for effectively neutral ones (Fig. 2C). For intermediately selected sites, \( G(v) \) is similar to the case of strong selection, with measurable differences only when \( v \gg v_s \) (see inset in Fig. 2C and SI Section 3). We would therefore expect that as the sample size of GWAS increase and the threshold contribution to variance decreases, intermediate and strongly selected sites will be discovered first, and effectively neutral sites will be discovered much later. In the SI (SI Section 3 and Fig. S13), we also derive corollaries of these results for the distribution of numbers of segregating sites that make a given contribution to genetic variance.
Discussion

In humans, GWAS for many traits display a similar behavior: when sample sizes are small, the studies discovered almost nothing, but once they exceeded a threshold sample size, both the number of associations discovered and the heritability explained suddenly began to increase rapidly (4, 69). Intriguingly though, both the threshold study size and rate of increase vary among traits. These observations raise several questions, including: How is the threshold study size determined? How should the number of associations and explained heritability increase with study size once this threshold is exceeded? With an order of magnitude increase in study sizes into the millions imminent, how much more of the genetic variance in traits should we expect to explain? The theory that we developed offers tentative answers to these questions.

To relate our theory to GWAS, we must first account for the power to detect loci that contribute to quantitative genetic variation. In studies of continuous traits, the power can be approximated by a step function in the contribution of loci to additive genetic variance (ignoring the effects of genotyping, which are considered below, and some other potential complications, e.g. (70)). In other words, loci that contribute more than a threshold value $v^*$ should be detected and those that contribute less should not. The threshold is affected by the study size, $m$, and by the total phenotypic variance in the trait, $V_p$, where $v^*(m, V_p) \propto V_p/m$ (SI Section 6) (69). Given a trait and study size, the number of associations discovered and heritability explained should then follow from the tails of the distributions that we have considered, corresponding to the threshold $v^*(m, V_p)$.

Our results therefore suggest that when genetic variation in a trait is sufficiently pleiotropic, the first loci to be discovered in GWAS will be intermediate or strongly selected, with correspondingly large effect sizes (i.e., $S \approx \frac{Nn}{2w^2}a_1^2 \gg 1$). The threshold study size for their discovery is proportional to $V_p/v_s$, i.e., to the total phenotypic variance measured in units of the expected contribution of a strongly selected site to variance. This result makes intuitive sense, as the total phenotypic variance can be viewed as the background noise for the discovery of individual strongly (and intermediate) selected sites. Once the threshold study size is exceeded, the increase in the number of associations and
explained heritability with study size directly follow from the functional forms that we derived for intermediate and strongly selected sites (Eq. 14 and S41 and Figs. 2C, S13B, 3A and S15A). The trait specific behavior boils down to the same parameter, $V_r/V_s$, relating the study size with the threshold $v^*$. Some results are modified when variation in a trait is only weakly pleiotropic, which is probably less common: notably, intermediate selected sites would begin to be discovered only after strongly selected ones (cf. Fig. 2B, S13A, and S14) and the threshold study size for strongly selected sites would be higher (cf. Eq. 12 and Fig. S14A).

**Fig. 3.** Predictions for GWAS of continuous traits, in the pleiotropic limit (cf. SI Section 3 for derivations). (a) The proportion of heritability explained as a function of study size. (b) The heritability explained in re-sequencing and genotyping studies as a function of the scaled selection coefficient. The study size was chosen such that a re-sequencing study would capture 25% of the strongly selected variance (corresponding to a study size of $\sim 16 V_r/V_s$). For the case without pleiotropy, see Fig. S14A-B.

Our results further suggest that for missing heritability to be mainly due to the limited power of genotyping to detect rare variants (14, 71), selection on the genetic variance in a trait would have to be very strong. GWAS impute allele frequencies at loci that are not included in the genotyping array, and the accuracy of this imputation declines rapidly when minor allele frequencies become too small to be well represented in the imputation panel (72, 73). This decline in power is dealt with in several ways, all of which are similar in effect to considering alleles above a threshold frequency that allows them to be imputed accurately (74). For example, a study using an Illumina 1M SNP array and the 1000
genomes phase III as an imputation panel would include only loci with MAF above ~1% (75). Even if loci below that frequency were imputed with perfect accuracy, however, they would only be detected if their contribution to variance exceeded the study’s threshold, implying that they would also have to have extremely large effect sizes; and under our model, extremely large effect sizes imply very strong selection (Fig. 3B). As an illustration, if a re-sequencing study identified 25% of the strongly selected (S>>1) variance, a genotyping study with the same sample size would suffer a 50% decrease in explained heritability only if $S \approx 200$ (assuming the same genotyping platform and imputation panel as above). If we further assuming a constant effective population size of $10^4$ this would imply a (heterozygote) selection disadvantage of 1%. We note, however, that recent population growth of the kind that has been inferred for European populations (76, 77) is expected to reduce the frequency of alleles under very strong selection (77, 78), leading to a greater loss of power due to genotyping (see below).

While someheritable variation may also arise from effectively neutral mutations, our results indicate that their expected contribution to genetic variance per site is much smaller than for strongly selected and intermediate ones (Fig. 2; (61, 62)). One implication is that for the contribution of effectively neutral variation to be comparable to that of intermediate and strongly selected variation, they would need to have a much larger mutational target size. It is difficult to evaluate whether this is plausible, because current GWAS are likely severely under powered to detect them (but see (79)). For example, sites with $S=0.1$ are expected to contribute only 1/20th of the genetic variance that arises from strongly selected ones; in the high pleiotropy limit, only 1% of that variance would be mapped when study sizes are large enough to map 95% of the strongly selected variance.

Importantly, these theoretical predictions can be used to make inferences based on GWAS results. As an illustration, we consider height and body mass index (BMI) in Europeans, two traits for which GWAS have discovered a sufficiently large number of genome-wide significant associations to allow for well powered tests and inferences (697 for height (16) and 97 for BMI (15)). We fit our theoretical predictions for the distribution of variance among loci to the distributions observed for the genome-wide significant associations reported for each of these traits (cf. Supplement Section 7 for details). In so doing, we assume that the mapped loci are under intermediate or strong selection, because our
theory suggests that they should be the first to be mapped. We also make the plausible assumption that genetic variance for these traits is highly pleiotropic (see Introduction; 37, 42) and therefore assume the limit form for the distribution of effect sizes. Under these assumptions, we expect the distribution of variance among loci to be well approximated by a simple form (Eq. S89), which depends on a single parameter \( v_3 \). When we estimate this parameter, we find that the theoretical distribution fits the data well (Fig. 4A). Specifically, we cannot reject our model based on the data, suggesting that with only one parameter, we obtain good fits to the data for both traits (by a Kolmogorov-Smirnov test, \( p = 0.14 \) for height and \( p = 0.54 \) for BMI; cf. Supplement Section 7). By comparison, without pleiotropy (\( n = 1 \)), our predictions provide a poor fit to these data (by a Kolmogorov-Smirnov test, \( p < 10^{-5} \) for height and \( p = 0.05 \) for BMI; Fig. S11).

![Fig. 4. Model fit and predictions for height and BMI. In (a), we show the fit for associated loci. In (b), we present our predictions for future increases in the heritability explained with GWAS size. The predicted number of associations is shown in Fig. S16. 95% CIs are based on bootstrap; see Supplement Section S7 for details.](image)

By fitting the model to GWAS results, we can also make inferences about evolutionary parameters underlying quantitative genetic variation. Estimating the degree of pleiotropy (\( n \)) as an additional parameter in the model, we find that for both height and BMI, \( n \) is sufficiently large for it to be indistinguishable from the high pleiotropy limit. Based on the shape of the fitted distributions in this limit and the number of loci that fall above the
threshold contribution to variance of the studies, we can also estimate the mutational target size and proportion of the heritable variance arising from mutations within the range of selection coefficients visible in these GWAS. These estimates suggest that, within this range, height has a target size of ~5Mb, which account for ~50% of the heritable variance, whereas BMI has a target size of ~1Mb, which explains ~15% of the variance (see SI Table S2).

These parameter estimates can help to interpret GWAS results and make predictions about future studies. They suggest that the GWAS for height succeeded in mapping a substantially greater proportion of the heritable variance (~20% compared to ~3-5%) despite their smaller sample size (~250K compared to ~340K), because the proportion of variance arising from mutations within the range of detectable selection effects is much greater for height than for BMI (~50% compared to ~15%). Furthermore, the estimates of target sizes and the relationship between sample size and threshold contribution to variance can be used to predict how the explained heritability and number of associations should increase with sample size (Figs. 4B and S16). These predictions are likely under-estimates as the range of selection effects itself should also increase with sample size.

To make these inferences and predictions more reliable, an important next step will be to account for the effects of recent human demography on the distributions of genetic variance (78, 80-82). Notably, the recent growth of European populations (76, 77) would have affected the distribution of genetic variance arising from very strongly selected mutations. Specifically, it will have reduced its expectation, because many of these mutations would have entered the population since the onset of growth and thus will be at lower frequencies (77, 78, 82). This consideration alongside the effect of genotyping would suggest that the range of selection effects revealed by current GWAS is in fact bound from above, and includes only moderately selected loci. However, moving beyond qualitative considerations would require incorporating the effects of demography into the model (77, 78, 82). Doing so may allow us to predict the outcome of future GWAS as well as to learn about the relative importance of different evolutionary and genetic forces in shaping quantitative genetic variation for different traits.
Acknowledgements. We have benefited hugely from discussions and comments from Guy Amster, Nick Barton, Jeremy Berg, Graham Coop, Laura Hayward, David Murphy, Joe Pickrell, Jonathan Pritchard and Molly Przeworski. The work was funded by NIH grant GM115889 to GS.

References:

1. Falconer DS & Mackay TFC (1996) *Introduction to quantitative genetics* (Benjamin Cummings, Essex, England) 4th Ed.
2. Lynch M & Walsh B (1998) *Genetics and analysis of quantitative traits* (Sinauer Sunderland, MA).
3. Provine WB (2001) *The origins of theoretical population genetics: With a new afterword* (University of Chicago Press).
4. Visscher PM, Brown MA, McCarthy MI, & Yang J (2012) Five years of gwas discovery. *Am J Hum Genet* 90:7-24.
5. Barton NH & Keightley PD (2002) Understanding quantitative genetic variation. *Nature Reviews Genetics* 3:11-21.
6. Manolio TA (2010) Genomewide association studies and assessment of the risk of disease. *New Engl J Med* 363(2):166-176.
7. Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, & Kruglyak L (2013) Finding the sources of missing heritability in a yeast cross. *Nature* 494(7436):234-237.
8. Manolio TA, et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461:747-753.
9. Zaitlen N, et al. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet* 9(5):e1003520.
10. Gibson G (2012) Rare and common variants: Twenty arguments. *Nat Rev Genet* 13(2):135-145.
11. Eichler EE, et al. (2010) Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* 11(6):446-450.
12. Zuk O, Hechter E, Sunyaev SR, & Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 109(4):1193-1198.
13. Lee SH, Wray NR, Goddard ME, & Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88(3):294-305.
14. Zuk O, et al. (2014) Searching for missing heritability: Designing rare variant association studies. *Proc Natl Acad Sci USA* 111:E455-E464.
15. Locke AE, et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518:197-206.
16. Wood AR, et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 46(11):1173–1186.
17. de Vladar HP & Barton N (2014) Stability and response of polygenic traits to stabilizing selection and mutation. *Genetics* 197(2):749-767.
18. Agarwala V, Flannick J, Sunyaev S, Consortium GD, & Altshuler D (2013) Evaluating empirical bounds on complex disease genetic architecture. *Nat Genet* 45:1418-1427.
19. Caballero A, Tenesa A, & Keightley PD (2015) The nature of genetic variation for complex traits revealed by gwas and regional heritability mapping analyses. *Genetics* 201(4):1601-1613.
20. Pritchard JK (2001) Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69:124–137.
21. Eyre-Walker A (2010) Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proc Natl Acad Sci USA* 107:1752–1756.
22. Speed D & Balding DJ (2015) Relatedness in the post-genomic era: Is it still useful? *Nat Rev Genet* 16(1):33-44.
23. Turelli M (1984) Heritable genetic variation via mutation-selection balance: Lerch’s zeta meets the abdominal bristle. *Theor Popul Biol* 25(2):138-193.
24. Barton NH & Turelli M (1989) Evolutionary quantitative genetics: How little do we know? *Annu Rev Genet* 23:337-370.
25. Hill WG & Kirkpatrick M (2010) What animal breeding has taught us about evolution. *Annu Rev Ecol Evol Syst* 41:1-19.
26. Johnson T & Barton N (2005) Theoretical models of selection and mutation on quantitative traits. *Philos Trans R Soc Lond B Biol Sci* 360(1459):1411-1425.
27. Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314-334.
28. Byars SG, Ewbank D, Govindaraju DR, & Stearns SC (2010) Natural selection in a contemporary human population. *Proc Natl Acad Sci USA* 107(suppl 1):1787-1792.
29. Stulp G, Pollet TV, Verhulst S, & Buunk AP (2012) A curvilinear effect of height on reproductive success in human males. *Behav Ecol Sociobiol* 66(3):375-384.
30. Frederick DA & Haselton MG (2007) Why is masculinity sexy? Tests of the fitness indicator hypothesis. *Personality and Social Psychology Bulletin* 33(8):1167-1183.
31. Endler JA (1986) *Natural selection in the wild* (Princeton University Press).
32. Kingsolver JG, *et al.* (2001) The strength of phenotypic selection in natural populations. *Am Nat* 157(3):245–261.
33. Charlesworth B, Lande R, & Slatkin M (1982) A neo-darwinian commentary on macroevolution. *Evolution* 36(3):474-498.
34. Barton NH (1990) Pleiotropic models of quantitative variation. *Genetics* 124(3):773-782.
35. Kondrashov AS & Turelli M (1992) Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics* 132(2):603-618.
36. Wagner GP (1996) Apparent stabilizing selection and the maintenance of neutral genetic variation. *Genetics* 143(1):617-619.
37. Pickrell JK, *et al.* (2016) Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* 48(7):709-717.
38. Sivakumaran S, *et al.* (2011) Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet* 89(5):607-618.
39. Solovieff N, Cotsapas C, Lee PH, Purcell SM, & Smoller JW (2013) Pleiotropy in complex traits: Challenges and strategies. *Nat Rev Genet* 14(7):483-495.
40. Cotsapas C, et al. (2011) Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* 7(8):e1002254.
41. Andreassen OA, et al. (2013) Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS Genet* 9(4):e1003455.
42. Bulik-Sullivan B, et al. (2015) An atlas of genetic correlations across human diseases and traits. *Nat Genet* 47(11):1236-1241.
43. Lee SH, Yang J, Goddard ME, Visscher PM, & Wray NR (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 28(19):2540-2542.
44. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide snps. *Nat Genet* 45(9):984-994.
45. Perry JRB, et al. (2014) Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* 514(7520):92-97.
46. Scott RA, et al. (2012) Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* 44(9):991-1005.
47. Shi H, Kichaev G, & Pasaniuc B (2016) Contrasting the genetic architecture of 30 complex traits from summary association data. *Am J Hum Genet* 99(1):139-153.
48. Yang J, et al. (2010) Common snps explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565-569.
49. Teslovich TM, et al. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466(7307):707-713.
50. Charlesworth B (1979) Evidence against fisher’s theory of dominance. *Nature* 278(5707):848-849.
51. Crow JF (2010) On epistasis: Why it is unimportant in polygenic directional selection. *Philos Trans R Soc Lond B Biol Sci* 365(1544):1241-1244.
52. Clayton DG (2009) Prediction and interaction in complex disease genetics: Experience in type 1 diabetes. *PLoS Genet* 5(7):e1000540.
53. Hill WG, Goddard ME, & Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* 4(2):e1000008.
54. Ávila V, et al. (2014) The action of stabilizing selection, mutation, and drift on epistatic quantitative traits. *Evolution* 68(7):1974-1987.
55. Wright S (1929) The evolution of dominance. *Am Nat* 63(689):556-561.
56. Wei W-H, Hemani G, & Haley CS (2014) Detecting epistasis in human complex traits. *Nat Rev Genet* 15(11):722-733.
57. Wood AR, et al. (2014) Another explanation for apparent epistasis. *Nature* 514(7520):E3-E5.
58. Evans DM, et al. (2011) Interaction between erap1 and hla-b27 in ankylosing spondylitis implicates peptide handling in the mechanism for hla-b27 in disease susceptibility. *Nat Genet* 43(8):761-767.
59. The International Multiple Sclerosis Genetics Consortium (2015) Class ii hla interactions modulate genetic risk for multiple sclerosis. *Nat Genet* 47(10):1107-1113.
60. Barton N & Turelli M (1987) Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genet Res 49(02):157-173.
61. Bulmer MG (1972) The genetic variability of polygenic characters under optimizing selection, mutation and drift. Genet Res 19(01):17-25.
62. Keightley PD & Hill WG (1988) Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. Genet Res 52(01):33-43.
63. Robertson A (1956) The effect of selection against extreme deviants based on deviation or on homozygosis. J Genet 54(2):236-248.
64. Wright S (1935) The analysis of variance and the correlations between relatives with respect to deviations from an optimum. J Genet 30(2):243-256.
65. Keightley PD & Hill WG (1990) Variation maintained in quantitative traits with mutation-selection balance: Pleiotropic side-effects on fitness traits. Proc R Soc Lond B Biol Sci 242(1304):95-100.
66. Lande R (1975) The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet Res 26(3):221-235.
67. Ewens WJ (2004) Mathematical population genetics 1: I. Theoretical introduction (Springer Science & Business Media).
68. Martin G & Lenormand T (2006) A general multivariate extension of fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60(5):893-907.
69. Sham PC & Purcell SM (2014) Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet 15(5):335-346.
70. Dickson SP, Wang K, Krantz I, Hakonarson H, & Goldstein DB (2010) Rare variants create synthetic genome-wide associations. PLoS Biol 18(1):e1000294.
71. Cirulli ET & Goldstein DB (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 11(6):415-425.
72. Li Y, Willer C, Sanna S, & Abecasis G (2009) Genotype imputation. Annu Rev Genomics Hum Genet 10:387-406.
73. Pritchard JK & Przeworski M (2001) Linkage disequilibrium in humans: Models and data. Am J Hum Genet 69(1):1-14.
74. Evangelou E & Ioannidis JPA (2013) Meta-analysis methods for genome-wide association studies and beyond. Nat Rev Genet 14(6):379-389.
75. The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. Nature 526(7571):68-74.
76. Tennessen JA, et al. (2012) Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337(6090):64-69.
77. Keinan A & Clark AG (2012) Recent explosive human population growth has resulted in an excess of rare genetic variants. Science 336(6082):740-743.
78. Simons YB, Turchin MC, Pritchard JK, & Sella G (2014) The deleterious mutation load is insensitive to recent population history. Nat Genet 46(3):220-224.
79. Loh P-R, et al. (2015) Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. Nat Genet 47(12):1385-1392.
80. Gazave E, Chang D, Clark AG, & Keinan A (2013) Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. Genetics 195(3):969-978.
81. Uricchio LH, Torres R, Witte JS, & Hernandez RD (2015) Population genetic simulations of complex phenotypes with implications for rare variant association tests. *Genet Epidemiol* 39(1):35-44.

82. Lohmueller KE (2014) The impact of population demography and selection on the genetic architecture of complex traits. *PLoS Genet* 10(5):e1004379.
A model for the genetic architecture of quantitative traits under stabilizing selection

**Supplementary Information**

Simons, Y. B., Bullaughey, K., Hudson, R. R. and G. Sella

**Table of Contents**

1. The model ................................................................................................................................................. 2
   1.1. Absorbing the environmental contribution into the fitness function ........................................... 2
   1.2. The distribution of mutational effect sizes on a given trait .......................................................... 2
   2. Solving for summaries of genetic architecture ......................................................................................... 4
       2.1. The first two moments of change in allele frequency ................................................................. 5
       2.2. Sojourn time ..................................................................................................................................... 6
   2.3. Calculating expectations of summaries of architecture ................................................................. 6
   3. Additive genetic variance and number of segregating sites ................................................................. 8
       3.1. Expectations ..................................................................................................................................... 8
       3.2. Densities ......................................................................................................................................... 10
       3.3. Comparing predictions against simulations .................................................................................... 14
   4. Justification for assumptions .................................................................................................................... 15
       4.1. Normal and isotropic phenotypic distribution around the optimum ........................................... 16
       4.2. The phenotypic variance satisfies $\sigma^2 \ll w^2$ ........................................................................ 16
       4.3. Mutational effect sizes satisfy $a^2 \ll w^2$ .................................................................................... 17
       4.4. Deviations of the mean phenotype from the optimum can be neglected .................................... 17
   5. Model robustness ..................................................................................................................................... 19
       5.1. Changes to the optimal phenotype ............................................................................................... 19
       5.2. Asymmetric mutational input ....................................................................................................... 22
       5.3. Major effect loci ............................................................................................................................. 27
       5.4. Anisotropic mutation ..................................................................................................................... 28
   6. The power to detect loci in GWAS ............................................................................................................ 33
       6.1. Re-sequencing studies .................................................................................................................... 34
       6.2. Genotyping ..................................................................................................................................... 36
   7. Inference .................................................................................................................................................... 38
       7.1. The composite likelihood ............................................................................................................... 38
       7.2. Determining $v^*$ and removing outliers ..................................................................................... 40
       7.3. Estimating target size and explained variance ............................................................................. 41
       7.4. Estimating confidence intervals .................................................................................................. 42
       7.5. Testing goodness of fit ................................................................................................................... 44
   8. Glossary of notation ................................................................................................................................. 46
   9. Bibliography ............................................................................................................................................. 47
   10. Additional figures ................................................................................................................................. 49
1. The model

1.1. Absorbing the environmental contribution into the fitness function

Here, we show that the additive environmental contribution to the phenotype can be absorbed into the fitness function, which justifies our considering only the additive genetic contribution in our analysis. This result has been derived multiple times for the one dimensional case (e.g., (1)). The argument in the multi-dimensional case is similar and included for completeness.

First, assume that the additive environmental contribution to the phenotype, \( r^*_e \), is distributed as a multi-normal with mean 0 and isotropic variance \( V_e \). The expected absolute fitness of an individual with additive genetic contribution to the phenotype, \( r_g \), is given by averaging fitness over the distribution of environmental contributions. Namely,

\[
\bar{W}(r_g) = \int r_e \frac{1}{(2\pi V_e)^{n/2}} e^{-\frac{|r_g|^2}{2V_e}} W(r_g + r_e) = \int r_e \frac{1}{(2\pi V_e)^{n/2}} e^{-\frac{|r_g|^2}{2V_e}} e^{-\frac{|r_g + r_e|^2}{2w^2}}
\]

\[
= \frac{1}{(1 + V_e/w^2)^{n/2}} e^{-\frac{|r_g|^2}{2(V_e + V_e)}}.
\]

(S1)

Given that absolute fitness is defined up to a multiplicative constant, we can therefore absorb the additive environmental contribution by using the Gaussian fitness function

\[
\bar{W}(r_g) = \exp \left( -\frac{|r_g|^2}{2\bar{w}^2} \right),
\]

(S2)

where \( \bar{w}^2 = w^2 + V_e \). Even when the environmental contribution is anisotropic, we can always choose a coordinate system in which the effective fitness function takes an isotropic form around the fitness peak (Eq. 1).

1.2. The distribution of mutational effect sizes on a given trait

In the main text, we define the distribution of phenotypic effects of newly arising mutations in the \( n \)-dimensional trait space, \( \tilde{a} \). Here, we consider the projection of
these effects on a given trait, $a_1$, taken without loss of generality to be on the 1st dimension. The distribution of effect sizes on a focal trait will depend on the degree of pleiotropy, $n$, and the form of this dependency becomes important when we consider how pleiotropy affects genetic architecture.

We want to calculate the distribution of effect sizes on the focal trait, $a_1$, conditional on their overall effect, $a = ||\vec{a}||$. We assume that the distribution of effects of de novo mutations is isotropic in trait space. The effect of a mutation, $\vec{a}$, therefore has equal probability to occupy any point on an n-dimensional sphere with radius $a$. Let $S_m(x)$ denote the surface area of an m-dimensional sphere of radius $x$ and $\theta$ denote the angle between the vector $\vec{a}$ and its projection $a_1$, i.e., $a_1 = a \cos \theta$. In these terms, the surface area element corresponding to angle $\theta$ is

$$S_{n-1}(a \sin \theta) \ a \ d\theta,$$

and by a change of variables, the surface area element corresponding to projection $a_1$ on the focal trait is

$$S_{n-1}(a \sin \theta) \ a \ d\theta = S_{n-1} \left( \frac{a^2 - a_1^2}{a^2 - a_1^2} \right) \ a \ d a_1,$$

since $d a_1 = \left| \frac{d a_1}{d \theta} \right| d \theta = a \sin \theta \ d\theta = \sqrt{a^2 - a_1^2} d\theta$. This result implies that the probability density of $a_1$ is

$$\varphi_n(a_1 | a) = \frac{S_{n-1} \left( \frac{a^2 - a_1^2}{a^2 - a_1^2} \right) \ a \ a_1^{n-1}}{S_n(a)} \ a \ a_1^{n-2} \ \frac{1}{a}$$

(for a similar derivation, see (2)).

Next, we consider the high pleiotropy limit form of this distribution. For any degree of pleiotropy, the symmetry of the mutational distribution implies that

$$E(a_1 | a) = 0$$

and the equivalence among traits implies that
\[ V(a_1|a) = a^2/n \]  
\text{(S7)}

(cf. main text for more details). It follows that when \( n \) becomes sufficiently large, \( a_1/a \ll 1 \) and therefore

\[ \left(1 - \frac{a_1^2}{a^2}\right)^{\frac{n-3}{2}} \approx \exp\left(-\frac{n}{2} \frac{a_1^2}{a^2}\right). \]  
\text{(S8)}

In addition, \( \Gamma\left(\frac{n}{2}\right)/\Gamma\left(\frac{n-1}{2}\right) \approx \sqrt{n/2}. \) Substituting these expressions into Eq. S5, we find that for sufficiently large \( n \) the distribution of effect sizes approaches the normal distribution

\[ \phi_n(a_1|a) \approx \frac{1}{\sqrt{2\pi(a^2/n)}} \exp\left(-\frac{1}{2} \frac{a_1^2}{a^2/n}\right). \]  
\text{(S9)}

As we elaborate in the main text, important implications about quantitative genetic variation follow from this high pleiotropy limit. The limit also holds quite generally when the distribution of effect sizes is anisotropic (see Section S5.4).

2. **Solving for summaries of genetic architecture**

Here, we derive closed forms for summaries of genetic architecture under our model. We begin by deriving the first two moments of change in allele frequency in a single generation. With these moments at hand, we use the diffusion approximation to calculate the sojourn time for alleles that contribute to quantitative genetic variation (3). Together with the distribution of effect sizes derived in the previous section, the sojourn time allows us to obtain closed forms for summaries of genetic architecture. Specifically, we can obtain a closed form for any summary that can be described as a function of allele frequencies and effect sizes at sites contributing to quantitative genetic variation. We use these expressions to calculate the summaries used in the main text, for example the expected additive genetic variance and its distribution across sites.
2.1. The first two moments of change in allele frequency

We assume that:

- The phenotypic distribution at steady state is well approximated by an isotropic multivariate normal distribution centered at the optimum, namely by the probability density

\[
f(r) = \frac{1}{(2\pi \sigma^2)^{n/2}} \exp \left(-\frac{r^2}{2\sigma^2}\right).\tag{S10}
\]

- That both \(a^2\) and \(\sigma^2 \ll w^2\).

These assumptions are justified in Section S4.

We rely on these assumptions to calculate the first two moments of change in frequency in a single generation for an allele with phenotypic effect \(\bar{a}\) and frequency \(q\). The fitnesses of the three genotypes at the site depend on its distribution of genetic backgrounds, i.e., on the total phenotypic contribution of sites other than the focal one \(\bar{R}\). Following Eq. S10 and assuming every allele contributes only a small proportion of the genetic variance, the distribution of \(\bar{R}\) is well approximated by

\[
f(\bar{R}|\bar{a}, q) = \frac{1}{(2\pi \sigma^2)^{n/2}} \exp \left(-\frac{|\bar{R}+q\bar{a}|^2}{2\sigma^2}\right).\tag{S11}
\]

The expected fitnesses of the three genotypes then follow from integrating over backgrounds:

\[
W_{00} = \int_{\bar{R}} f(\bar{R}|\bar{a}, q)W(\bar{R}) = \left(\frac{w}{\sqrt{w^2+\sigma^2}}\right)^n \exp \left(-\frac{a^2q^2}{2(w^2+\sigma^2)}\right),\tag{S12}
\]

\[
W_{01} = \int_{\bar{R}} f(\bar{R}|\bar{a}, q)W(\bar{R} + \bar{a}/2) = \left(\frac{w}{\sqrt{w^2+\sigma^2}}\right)^n \exp \left(-\frac{a^2(q-1/2)^2}{2(w^2+\sigma^2)}\right)\tag{S13}
\]

and

\[
W_{11} = \int_{\bar{R}} f(\bar{R}|\bar{a}, q)W(\bar{R} + \bar{a}) = \left(\frac{w}{\sqrt{w^2+\sigma^2}}\right)^n \exp \left(-\frac{a^2q^2}{2(w^2+\sigma^2)}\right).\tag{S14}
\]

The first moment of change in allele frequency is then

\[
E(\Delta q) = -pq \frac{p(W_{00}-W_{01})+q(W_{01}-W_{11})}{w} \approx -\frac{a^2}{4w^2} pq \left(\frac{1}{2} - q\right),\tag{S15}
\]
relying on our assumptions that \( a^2 \) and \( \sigma^2 \ll w^2 \). The functional form of the first moment is equivalent to that of the standard viability selection model with under-dominance and selection coefficient \( s = \frac{a^2}{4w^2} \) or scaled selection coefficient

\[
S = \frac{Na^2}{2w^2}
\]  

(S16)

Similarly, we find that

\[
V(\Delta q) \approx \frac{pq}{2N},
\]

(S17)

which is the standard second moment with genetic drift.

2.2. Sojourn time

Based on the first two moments, we can use the diffusion approximation to calculate the sojourn time as a function of allele frequency, i.e., the density of the time that an allele spends at a given frequency \( q \) before it fixes or is lost (3). For a mutant allele with initial frequency \( 1/2N \) and scaled selection coefficient \( S \), the sojourn time is

\[
\tau(q|S) = \begin{cases} 
\frac{N\sqrt{\pi/S}}{\text{erf}(\sqrt{S}/2)} \frac{e^{S(1-2x)^2/4}}{x(1-x)} f_-(S, q)f_+(S, 1/2N) & 0 \leq q \leq 1/2N \\
\frac{N\sqrt{\pi/S}}{\text{erf}(\sqrt{S}/2)} \frac{e^{S(1-2x)^2/4}}{x(1-x)} f_+(S, q)f_-(S, 1/2N) & 1/2N \leq q \leq 1 
\end{cases}
\]  

(S18)

where erf is the error function and \( f_{\pm}(S, y) \equiv \text{erf}(\sqrt{S}/2) \pm \text{erf}(\sqrt{S}(1-2y)/2) \).

The sojourn time takes simple limiting forms when selection is effectively neutral \( (S \ll 1) \) or strong \( (S \gg 1) \). In the effectively neutral range, it is well approximated by \( \tau(q|S) \approx \frac{2}{q} \) and in the strongly selected range, it is well approximated by

\[
\tau(q|S) \approx \frac{2}{q} \exp(-Sq).
\]

2.3. Calculating expectations of summaries of architecture

Many summaries of interest can be expressed as sums over segregating sites of some function \( c(q, a_1) \), where \( q \) is the derived allele frequency and \( a_1 \) is the effect size on the trait. For example, the additive genetic variance in a trait is given by the
sum of $n(q, a_1) = \frac{1}{2} a_1^2 q(1 - q)$ over sites. The expectation over such summaries can be expressed as

$$E(C) = 2NU \int_q \int_{a_1} c(q, a_1) \rho(q, a_1),$$ (S19)

where $C$ is the summery summed over all sites, $2NU$ is the population mutation rate per generation and $\rho(q, a_1)$ is the density of sites with the corresponding frequency and effect size per unit mutational input.

The density $\rho(q, a_1)$ can be broken down into contributions from sites with different selection coefficients, i.e.,

$$\rho(q, a_1) = \int_S f(S) (\tau(q|S) \eta(a_1|S)),$$ (S20)

where $f(S)$ is the distribution of selection coefficients and $\tau(q|S)$ is the sojourn time of a mutation with selection coefficient $S$ (Eq. S18). The probability density $\eta(a_1|S)$ of effect sizes given selection coefficient $S$ follows from Eqs. S5 and S16

$$\eta(a_1|S) = \varphi_n(a_1|a(S)) = \varphi_n \left( a_1 \left| \sqrt{\frac{2w^2}{N}} S \right. \right).$$ (S21)

This allows us to break down our summaries into contributions from sites with different selection coefficients

$$E(C) = 2NU \int_S f(S) E(C|S)$$ (S22)

with

$$E(C|S) = \int_q \int_{a_1} c(q, a_1) \tau(q|S) \eta(a_1|S).$$ (S23)

We use Eq. S23 to study how summaries of architecture depend on the strengths of selection, and how these summaries will depend on different distributions of selection coefficients. This allows us to draw general implications about genetic architecture despite our limited knowledge about this distribution.
3. Additive genetic variance and number of segregating sites

The distributions of additive genetic variance and of the number of segregating sites are critical to understanding genetic architecture and specifically to interpreting results of GWAS. Here we derive closed forms for both distributions as well as simple approximations under strong and effectively neutral selection.

3.1. Expectations

We begin by considering the expected contribution of a site to additive genetic variance. Substituting the contribution to variance from a single site \( v(q, a_1) = \frac{1}{2} a_1^2 q (1 - q) \) into Eq. S23, we find that

\[
E(V|S) = \int_q \int_{a_1} v(q, a_1) \tau(q|S) \eta(a_1|S)
\]

\[
= \int_q \frac{1}{2} q (1 - q) \tau(q|S) \int_{a_1} a_1^2 \eta(a_1|S)
\]

\[
= \int_q \frac{1}{2} q (1 - q) \tau(q|S) \frac{2w^2}{nN} S = \frac{2w^2}{nN} \int_q \frac{1}{2} S q (1 - q) \tau(q|S)
\]

(S24).

The total additive genetic variance is

\[
\sigma^2 = 2NU \int_S f(S) E(V|S).
\]

(S25)

The closed form for \( E(V|S) \) (Eq. S24) was integrated numerically to obtain Fig. 2a in the main text. We can use the results of (4) to obtain an analytic approximation for \( E(V|S) \):

\[
E(V|S) = \frac{2w^2}{nN} \sqrt{\frac{\pi S}{4}} \text{erfi} \left( \frac{\sqrt{S}}{4} \right) \exp \left( -\frac{S}{4} \right) + O \left( \frac{1}{2N} \right).
\]

(S26)

with \( \text{erfi}(x) = \text{erf}(ix) / i \).

In the effectively neutral and strong selection limits, we can use simple limiting forms for the sojourn time to derive simple forms for \( E(V|S) \). In the effectively neutral limit, \( \tau(q|S) \approx 2/q \) and so

\[
E(V|S)_{S \ll 1} \approx \frac{2w^2 S}{nN^2 2}.
\]

(S27)
In the strong selection limit, \( \tau(q|S) \approx 2 \exp(-S q) / q \) and so

\[
E(V|S)_{S \gg 1} \approx \frac{2w^2}{nN}, \quad (S28)
\]

The constant \( \frac{2w^2}{nN} \), which recurs in our derivations (e.g., Eq. S24), thus has a simple interpretation: it is the expected contribution of strongly selected sites to additive genetic variance (per unit mutational input). We therefore denote it by

\[
v_s = \frac{2w^2}{nN} \quad (S29)
\]

and henceforth measure variance in these units.

We next consider how the expected number of segregating sites depends on the strength of selection. This expectation (per unit mutational input) is simply the mean sojourn time of a newly arising mutation. Formally, it follows from substituting \( c(q, a_1) = 1 \) into Eq. S23, i.e.,

\[
E(K|S) = \int_q \int_{a_1} \tau(q|S) \eta(a_1|S) = \int_q \tau(q|S) \int_{a_1} \eta(a_1|S) = \int_q \tau(q|S). \quad (S30)
\]

In Fig. S1, we have calculated this integral numerically for different values of \( S \) to show that the number of segregating sites depends only weakly on \( S \).

**Figure S1.** The number of segregating sites per unit mutational input (or equivalently the expected sojourn time of a newly arising mutation), Eq. S30, is only weakly dependent on the strength of selection. Calculated for a population size of 20,000.
3.2. Densities

Here, we consider how the additive genetic variance is distributed among sites. The density of segregating sites with a given contribution to variance \( v \) can be calculated by substituting Dirac’s delta function \( \delta(v - \frac{1}{2} a_1^2 q(1 - q)) \) into Eq. S23, yielding

\[
\rho(v|S) = \mathbb{E}(\delta(v - \frac{1}{2} a_1^2 q(1 - q))|S) = \int_q \int_{a_1} \delta(v - \frac{1}{2} a_1^2 q(1 - q)) \tau(q|S) \eta(a_1|S) = \]

\[
= \int_{a_1} \left[ \tau(q_+(v, a_1)|S) \left[ \frac{\partial q_+(v, a_1)}{\partial v} \right] + \tau(q_-(v, a_1)|S) \left[ \frac{\partial q_-(v, a_1)}{\partial v} \right] \right] \eta(a_1|S) \]

\[
= \int_{a_1} \left[ \tau(q_+(v, a_1)|S) + \tau(q_-(v, a_1)|S) \right] \frac{2}{a_1^2 \sqrt{1 - 8v/a_1^2}} \eta(a_1|S), \tag{S31}
\]

where \( q_\pm(v, a_1) = \frac{1}{2} \left( 1 \pm \sqrt{1 - 8v/a_1^2} \right) \) are the two frequencies that yield \( v = \frac{1}{2} a_1^2 q(1 - q) \). This integral can be calculated numerically for any \( S \) and \( \eta(a_1|S) \) (e.g., for different degrees of pleiotropy).

As we discuss in the main text and in Section S6, the tails of this distribution are especially important for the interpretation of GWAS. Notably, to a first approximation, the loci captured in a given GWAS would be those whose contribution to additive variance exceeds some threshold contribution \( v^* \).

We are particularly interested in the proportion of additive genetic variance captured in GWAS, which we approximate in terms of the additive variance carried by sites in the \( v > v^* \) tail. The expected variance in such a tail is

\[
E(V|v^*, S) = \int_{v>v^*} v \rho(v|S), \tag{S32}
\]

and its proportion of the total additive genetic variance is

\[
G(v^*|S) \equiv \frac{E(V|v^*, S)}{E(V|S)} = \frac{\int_{v>v^*} v \rho(v|S)}{\int_v v \rho(v|S)}. \tag{S33}
\]

For a distribution of selection coefficients, the proportion of variance becomes

\[
G(v^*) = \frac{\int_S G(v^*|S) E(V|S)f(S)}{\int_S E(V|S)f(S)} = \int_S G(v^*|S) \frac{E(V|S)f(S)}{\int_S E(V|S)f(S)}. \tag{S34}
\]
For cases without pleiotropy or with extensive pleiotropy, we can greatly simplify these expressions in the limits of effectively neutral and strong selection. We illustrate how these approximations can be obtained for the proportion of additive variance in a tail, $G(v^*|S)$ (Table S1). When selection is effectively neutral, then $\tau(q|S) = 2/q$ and thus

$$G(v^*|S) = \frac{\int_{q > v^*} v \rho(v|S)}{\int_v \rho(v|S)}$$

$$= \frac{2}{S} \int_{v > v^*} \int_{a_1} v \left( \frac{1}{q_{a_1}} - \frac{2}{\sqrt{1 - 8v/a_1^2}} \right) \eta(a_1|S)$$

$$= \frac{2}{S} \int_{v > v^*} \int_{a_1} v \left( \frac{1}{1 + \sqrt{1 - 8v/a_1^2}} + \frac{2}{\sqrt{1 - 8v/a_1^2}} \right) \eta(a_1|S)$$

$$= \frac{2}{S} \int_{v > v^*} \int_{a_1} \frac{2}{\sqrt{1 - 8v/a_1^2}} \eta(a_1|S)$$

with variance measured in units of $v_s$ and effect size measured in units of $\sqrt{v_s}$.

Without pleiotropy ($n \approx 1$), we know that $a_1 = \pm \sqrt{S}$ and the expression for the proportion of variance simplifies to

$$G(v^*|S) = \sqrt{1 - 8v^*/S}, \quad (S36)$$

where $S/8$ is the maximal contribution to variance for a mutation with selection coefficient $S$. When there is a high degree of pleiotropy ($n \gg 1$), $a_1$ is approximately normally distributed with mean 0 and variance $S$ and the expression for the proportion of variance simplifies to

$$G(v^*|S) = \int_{a_1 > \sqrt{8v^*}} \frac{a_1^2}{S \sqrt{1 - 8v^*/a_1^2}} \frac{1}{\sqrt{2\pi}} \exp \left( -a_1^2 / 2S \right) = \exp \left( -4v^*/S \right). \quad (S37)$$

When selection is strong, derived alleles are rare ($q \ll 1$), implying that $v \ll a_1^2$ and $q \approx 2v/a_1^2$, and the sojourn time is well approximated by $\tau(q|S) \approx 2 \exp(-Sq)/q$. The proportion of variance then simplifies to
\[ G(v^*|S) \approx \int_{v>v^*} \int_{a_1} v \tau(q(v,a_1)|S) \frac{2}{a_1^2} \eta(a_1|S) \]

\[ \approx \int_{v>v^*} \int_{a_1} v \frac{a_1^2}{v} \exp\left(-2Sv/a_1^2\right) \frac{2}{a_1^2} \eta(a_1|S) \]

\[ = \int_{v>v^*} \int_{a_1} 2 \exp\left(-2Sv/a_1^2\right) \eta(a_1|S) \]

\[ = \int_{a_1} \frac{a_1^2}{S} \exp\left(-2Sv^*/a_1^2\right) \eta(a_1|S). \]  

(S38)

Without pleiotropy, this expression further simplifies to

\[ G(v^*|S) \approx \exp(-2v^*), \]  

(S39)

and when there is a high degree of pleiotropy \((n \gg 1)\), then

\[ G(v^*|S) \approx \int_{a_1} \frac{a_1^2}{S} \exp\left(-\frac{2Sv^*}{a_1^2}\right) \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{a_1^2}{2S}\right) \]

\[ = (1 + 2\sqrt{v^*}) \exp(-2\sqrt{v^*}). \]  

(S40)

| Selection | Effectively neutral \((S \ll 1)\) | Strongly selected \((S \gg 1)\) |
|-----------|----------------------------------|-------------------------------|
| # of traits | \(n \approx 1\) | \(n \gg 1\) | \(n \approx 1\) | \(n \gg 1\) |
| E\((V|S)\) | \(S/2\) | \(S/2\) | 1 | 1 |
| G\((v^*|S)\) | \(\sqrt{1 - 8v^*/S}\) | \(\exp(-4v^*/S)\) | \(\exp(-2v^*)\) | \((1 + 2\sqrt{v^*})\exp(-2\sqrt{v^*})\) |

Table S1. Simple limits for the expected proportion of variance arising from sites with additive contribution greater than \(v^*\).

A noteworthy property of the proportion of variance from sites with contributions \(v > v^*\), \(G(v^*|S)\), is that it always appears to increase with the degree of pleiotropy, \(n\) (Figure S2). We do not have a proof for this property but can suggest an intuitive explanation. Without pleiotropy \((n=1)\), the selection coefficient determines the effect size, such that any contribution \(v^*\) to genetic variance corresponds to a specific minor allele frequency \(q^*\). The sites with contributions \(v > v^*\) are therefore
those with minor allele frequencies \( q > q^\ast \). Pleiotropy causes sites with a given selection coefficient to have a distribution of effect sizes on the trait under consideration. As a result, some sites with frequencies above \( q^\ast \) end up with contributions to variance below \( v^\ast \) while others exceed \( v^\ast \). To understand how this affects \( G(v^\ast| S) \), recall that for any selection coefficient, the density of variants always rapidly increases as \( q^\ast \) decreases. As long as the contribution \( v^\ast \) and the corresponding frequency without pleiotropy \( q^\ast \) are not close to 0, we may therefore expect that introducing pleiotropy would result in pushing more sites above \( v^\ast \) than those pushed below \( v^\ast \), resulting in a net increase to the proportion \( G(v^\ast| S) \).

**Figure S2.** The effect of pleiotropy on the proportion of the variance explained by sites contributing more than \( v^\ast \) to the variance, \( G(v^\ast| S) \) (see Eq. S33). For all selection coefficients, the proportion of variance explained increases as the number of traits, i.e. the degree of pleiotropy, increases.

Lastly, we consider how the number of segregating sites depends on their contribution to genetic variance vary as a function of this contribution (Fig. S13). The number of segregating sites \( K \) with \( v > v^\ast \) per unit mutational input is

\[
E(K|v^\ast, S) = \int_{v>v^\ast} \rho(v|S),
\]

and therefore the distribution of variance at these sites is

\[
f(v) = \frac{\rho(v|S)}{\int_{v>v^\ast} \rho(v|S)}.
\]

We use these results below, when we estimate model parameters from GWAS results (Section S7). They also allow us to translate the properties of the density of
genetic variance into properties of the density of sites (Fig. S13). Notably, by applying the same reasoning one can show that the density of sites is insensitive to the distribution of strong selection coefficients and pleiotropy always increases the number of sites above \( v^* \) (i.e., for any \( v^*>0 \)).

### 3.3. Comparing predictions against simulations

We tested our theoretical derivations for the total genetic variance and its distribution among sites against computer simulations. To this end, we ran forward simulation of the full model (the code is available at https://github.com/sellalab) with a range of parameter values chosen to balance between biologically plausibility and maintaining manageable running times. Notably, we used a population size of \( N = 1000 \), with a burn-in time of 10,000 generations. We varied the number of traits to include \( n = 1, 3, \) and 10, and varied the mutation rate per haploid genome per generation within the range \( 1/2N \leq U \leq 1 \). Selection coefficients were chosen from an exponential distribution with mean \( E(S) = 0.1, 10, \) and 50. For simplicity, we took \( w^2 = 1 \), which is tantamount to choosing the units that we use to measure effect sizes. For the sake of computational efficiency, the simulations were performed with fecundity selection, but we ran simulations with viability selection for a subset of the parameters to ensure that this choice does not affect our results (not shown).

Our theoretical derivations for both the total genetic variance and its distribution among sites agree well with results from simulations (Fig. S3). Notably, we found that as long as \( 1/2N \ll \sigma^2/w^2 \ll 1 \) (cf. Section S3.3), our theoretical predictions for the total genetic variance (Eq. S25) are indistinguishable from the simulation results (Fig. S3A). Theoretical and simulation results seem to agree even when \( \sigma^2/w^2 \leq 1/2N \), although we consider this range to be of lesser biological relevance (cf. Section S4.4). We tested our predictions for the distribution of variance across sites in terms of \( G(v^*) \) (Eq. S34), i.e., the proportion of the variance arising from sites that contribute more than \( v^* \) (Fig. S3B), and also found them to be highly accurate.
Figure S3. Testing theoretical predictions for genetic variance against simulations. (a) Total genetic variance (in units of $w^2$) as a function of the mutation rate. The range of relevant mutation rates, $1/2N \ll U \ll 1$, and the expected range for the validity of our derivation, $w^2/2N \ll \sigma^2 \ll w^2$, are marked by a grey box. (b) The distribution of variance among sites, for $U = 0.01$; $G(v^*)$ is the proportion of variance from sites contributing more than $v^*$ (Eq. S34). Error bars represent one standard deviation. For each set of parameters, the number of simulations was chosen to obtain standard deviations below 10%. In practice, we often reached standard deviations $\ll 10\%$, which is why error bars are mostly too small to be visible.

4. Justification for assumptions

Here, we justify the assumptions that we relied upon in deriving the first two moments of change in allele frequency (cf. Section S2.1; modeling assumptions are motivated in the introduction to the main text). We rely in part on self-consistency arguments, which should not be mistaken for being circular: specifically, we make assumptions about the behavior of the system and show that the solution to which we arrive satisfies these assumptions.
4.1. Normal and isotropic phenotypic distribution around the optimum

The assumption that the phenotypic distribution is well approximated by a normal distribution stems from an additive model of quantitative traits. By assuming that the phenotype arises from many additive contributions and that these additive contributions arise from some underlying distribution, normality follows from the law of large numbers. In terms of model parameters, we would expect normality to hold if the rate of mutations affecting the trait is sufficiently large, i.e., when $2NU \gg 1$.

We further assume that the phenotypic distribution is isotropic and its mean is at the optimum. Isotropy of the phenotypic distribution follows from assuming isotropy in the mutational input. In Section S5.4, we explore the consequences of anisotropy in the mutational input. In Section S4.4, we further show that the fluctuations of the mean phenotype around the optimum over time have negligible effects on allelic dynamics; a similar argument applies to fluctuations in the variance.

4.2. The phenotypic variance satisfies $\sigma^2 \ll w^2$

With the mean phenotype centered at the optimum, requiring that $\sigma^2 \ll w^2$ is equivalent to assuming that moving a standard deviation away from the mean phenotype entails only a minor reduction in fitness. This seems plausible for many phenotypes: if, for example, this assumption did not hold for human height, then individuals whose height is a standard deviation or more away from the population mean would suffer from a substantial reduction in fitness. Arguably, deviations from the mean height would then be recognized as a very common and severe disease.

Another line of argument that it is likely that $\sigma^2 \ll w^2$ is based on our results. If we assume that mutations are strongly selected, then our results suggest that

$$\sigma^2 = 2NU \cdot \nu_3 = \frac{4U}{n}w^2.$$  \hspace{1cm} (S43)
It follows that if the rate of mutations affecting the phenotype under consideration satisfies $U \ll 1$ then $a^2 \ll w^2$. The number of mutations per diploid human genome per generation is estimated to be $\sim 60$ (5), and less than 10% of the genome is assumed to be functional (6), suggesting that the number of de novo mutations with any effect on function is less than 3 per haploid per generation. It then seems plausible that the (haploid) mutation rate affecting a specific trait satisfies $U \ll 1$. Assuming that mutations are weakly selected increases the variance in Eq. S43 only moderately and assuming the mutations are effectively neutral would suggest it is much smaller, leaving the above argument intact.

4.3. **Mutational effect sizes satisfy $a^2 \ll w^2$**

As we argued in the introduction of the main text, variants for which the stronger condition $a^2 \ll \sigma^2$ holds account for most or all of the heritability explained in GWAS for many traits (e.g. (7-10)). Moreover, evidence for many traits suggests that the same is true for the variants that underlie the heritability that remains to be explained (11-14). Indeed, for this assumption to be violated, much of the genetic variance would have to arise from mutations that have a very large impact on fitness (i.e., with $s$ on the order of 1). While this may be the case for some diseases (e.g., autism (15)), it does not appear to be the case for most phenotypes that have been examined.

4.4. **Deviations of the mean phenotype from the optimum can be neglected**

In reality, the mean phenotype of the population fluctuates around the optimum. Here, we derive equations for the dynamic of the mean phenotype in order to estimate the magnitude and timescale of these fluctuations. We then show that these fluctuations have a negligible effect on the first two moments of change in allele frequency and thus on the results that follow from these moments.

We begin by deriving the first and second moment of change in mean phenotype. To this end, we assume the distribution of phenotypes is a multivariate normal centered around a mean phenotype, $\bar{r}$, i.e. that
\[ f(\tau^\prime) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{-|\tau^\prime - \tau|^2}{2\sigma^2}}. \] (S44)

The expected change in mean phenotype due to selection in one generation is therefore

\[ E(\Delta \bar{\tau}) = \frac{\int f(\tau^\prime) W(\tau^\prime) \tau^\prime}{\int f(\tau^\prime) W(\tau^\prime)} - \bar{\tau} = -\frac{\sigma^2}{w^2 + \sigma^2} \bar{\tau} \approx -\frac{\sigma^2}{w^2} \bar{\tau}. \] (S45)

By the same token, the variance in \(\Delta \bar{\tau}\) is simply the sampling variance

\[ V(\Delta \bar{\tau}) \approx \frac{\sigma^2}{N}, \] (S46)

where in both cases we relied on the assumption that \(\sigma^2 \ll w^2\).

These two moments define an Ornstein-Uhlenbeck process in \(\bar{\tau}\), allowing us to rely on well-known results (16). Notably, when the mean phenotype \(\bar{\tau}\) starts far from the optimum, it decays exponentially to the optimum with exponent \(\sigma^2/w^2\) (see Section S5.1 below). At steady state, \(\bar{\tau}\) will fluctuate with mean zero and \(E(\bar{\tau}^2) = n w^2/2N\) over a time scale of \(w^2/\sigma^2\) generations. The typical displacement of \(\bar{\tau}\) in any given direction will be \(\sqrt{w^2/2N}\), reflecting a balance between drift and the pull of selection toward the optimum.

Next we show that these fluctuations of the mean have negligible effects on allelic trajectories. To this end, we derive the first two moments of change in allele frequency, but this time, we include the effect of the displacement of \(\bar{\tau}\) from the optimum. While the second moment remains the same, the first moment becomes

\[ E(\Delta q) \approx -\frac{\bar{\tau}_a}{2w^2} pq - \frac{\sigma^2}{4w^2} pq \left(\frac{1}{2} - q\right) = -\frac{\bar{\tau}_a}{\sqrt{w^2/2N}} \frac{\sqrt{3}}{2N} pq - \frac{S}{2N} pq \left(\frac{1}{2} - q\right), \] (S47)

where \(\bar{\tau}_a\) is \(\bar{\tau}\)'s component in the direction of \(\bar{\alpha}\). However, our analysis establishes that \(\frac{\bar{\tau}_a}{\sqrt{w^2/2N}}\) is a scalar of the order of 1, which fluctuates around zero on a timescale of \(w^2/\sigma^2\). We can therefore compare the first term in the above equation, which represents directional selection, and the second term, which represents stabilizing selection. When stabilizing selection is strong, \(S \gg 1\), the stabilizing selection term
dominates over the directional selection term. In contrast, when selection is weak, i.e., $S \approx 1$ or smaller, then in any given generation, the directional term is not necessarily negligible. However, in this case, both terms affect substantial change in allele frequency only over a timescale of $2N$ generations; on this timescale, if $2N \gg w^2/\sigma^2$, the directional effect would average to zero. The directional term will become important only when $2N \leq w^2/\sigma^2$, that is $\sigma^2 \leq w^2/2N$. For $\sigma^2$ to be that small, virtually all alleles must have $S \ll 1$, such that their trajectories will be determined by drift, not selection. In summary, regardless of the selection acting on an allele, fluctuations of the mean phenotype around the optimum will have a negligible effect on its trajectory.

5. Model robustness

In this section, we consider the sensitivity of our results to relaxing some of the simplifying modeling assumptions about selection and mutation. Specifically, we show our results to be robust to moderate changes to the optimal phenotype; small asymmetry in the mutational input; the presence of major loci maintained at high frequency by selection on traits that are not included in the model; as well as to most forms of anisotropic mutation.

5.1. Changes to the optimal phenotype

We first consider how changes to the optimal phenotype over time would affect our results. It is easy to imagine how events such as migration from Africa to Europe or the onset of agriculture may have introduced rapid changes in beneficial phenotypes. In order to evaluate the potential impact of such events, we consider how an instantaneous change to the optimal phenotype would affect the allelic dynamics.

We begin by considering how such an instantaneous change to the optimum would affect the mean phenotype. If the shift to the optimum is small, on the order of the fluctuations in the mean phenotype at steady state or smaller, then the arguments provided in Section S4.4 will still hold and the shift would have a negligibly small effect on our results. We therefore assume that the shift in optimum, $\tilde{z}$, is large
compared to the scale of fluctuations \((z^2 \gg w^2/2N)\). This assumption means that we can use a deterministic approximation (based on Eq. S45) and describe the change in mean phenotype in a single generation by

\[
\Delta \bar{r} \approx E(\Delta \bar{r}) = -\frac{\sigma^2}{w^2} (\bar{r} - \bar{z})
\]  

(S48)

(neglecting higher moments). Further assuming that the mean phenotype was at the optimum, \(\bar{0}\), before the optimum shifted (at time \(t = 0\)) and neglecting changes to the genetic variance \(\sigma\), we find that

\[
\bar{r}(t) = \bar{z} \left(1 - \exp \left(-\frac{\sigma^2}{w^2} t\right)\right)
\]  

(S49).

Thus, the mean \(\bar{r}\) adapts to the new optimum on a timescale of \(w^2/\sigma^2\) generations (cf. (17) for a similar derivation).

We can rely on this approximation to learn when a shift in optimum will have negligible effects on allele trajectories. Recalling Eq. S47, the first moment of change in allele frequency is given by

\[
E(\Delta q) \approx -\frac{(\bar{r}(t) - \bar{z}) d}{2w^2} pq - \frac{\alpha^2}{4w^2} pq \left(\frac{1}{2} - q\right).
\]  

(S50)

where, based on our approximation (Eq. S49), the directional selection term introduced by the shift in optimum takes the time-dependent form

\[
E(\Delta_D q) = -\frac{(\bar{r}(t) - \bar{z}) d}{2w^2} pq \approx \frac{\bar{z} d}{2w^2} \exp \left(-\frac{\sigma^2}{w^2} t\right) pq.
\]  

(S51)

The effect of this directional term over the entire adaptive trajectory can be quantified by comparing the expected allele frequency after adaptation to the shift, \(q_D\), with initial frequency before the shift, \(q_0\), i.e.,

\[
\ln(q_D/q_0) = \int_t \frac{E(\Delta_D q)}{q(t)} = \frac{\bar{z} d}{2w^2} \int_t p(t) \exp \left(-\frac{\sigma^2}{w^2} t\right) < \frac{\bar{z} d}{2w^2} \int_t \exp \left(-\frac{\sigma^2}{w^2} t\right) = \frac{\bar{z} d}{2\sigma^2}.
\]  

(S52)

This result suggests that the relative change in allele frequency will be negligible so long as
This condition suggests that mutations with smaller effects would be less affected by the shift in the optimum. It further suggests that alleles that satisfy \( a^2 \ll \sigma^2 \), as appears to be the case for most loci discovered in GWAS (e.g., (7-10)), will be negligibly affected by shifts in optimum on the order of the total genetic variation (i.e., \( z \leq \sigma \)). These analytic predictions are confirmed by simulations (Fig. S4).

**Figure S4.** Distribution of the contributions of sites to variance after a shift in the optimum. The y-axis is the proportion of the variance explained by sites that contribute more than \( v^* \) to the variance. The dashed black is the theoretical prediction without adaptation and in color simulation results with adaptation. When the root mean square of \( \frac{\sqrt{\frac{E(a^2)}{2\sigma^2}}}{2\sigma^2} \) becomes larger than 1, directional selection substantially affects allele frequencies and therefore the contributions of sites to variance, as predicted by Eq. S53. (Since mutation is symmetric the mean of \( \frac{\sqrt{\frac{E(a^2)}{2\sigma^2}}}{2\sigma^2} \) is zero and we quantify its characteristic value by its root mean square \( \frac{\sqrt{\frac{E(a^2)}{2\sigma^2}}}{2\sigma^2} \).) Simulations were run with an exponential distribution of selection coefficients with \( E(S) = 25, N = 1,000, n = 1, U = 0.01 \), and a burn-in time of 10,000 generations. Results were taken 50 generations after the shift in optimum, which, for these parameters, is just after the population mean has reached the new optimum.
5.2. Asymmetric mutational input

In this section, we consider the sensitivity of our results to asymmetries in the mutational input, i.e., to the case in which mutations in a given direction in trait space, \( \tilde{b} \), are more likely to arise than mutations in the opposite direction, \(-\tilde{b}\) (see (18) for treatment of this problem in the limit of high per-site mutation rate).

An asymmetric mutational input introduces a shift in the mean phenotype every generation. With new mutations arising at frequency \(1/2N\), the expected shift is

\[
\Delta_M \tilde{r} = 2NU \ E_M (\tilde{d}) \cdot 1/2N = UE_M (\tilde{d}),
\]

(S54)

where \(E_M\) is the expectation over newly arising mutations. For each trait, effects have a characteristic size \(\sqrt{E(a^2)/n} = \sqrt{v_S} \sqrt{E(S)}\). The characteristic effect size sets the scale for the maximal shift in any direction, that is \(||\Delta_M \tilde{r}||\) is of the order of \(U \sqrt{E(a^2)/n}\) or smaller. We therefore parameterize the shift in mean phenotype due to new mutations by

\[
\Delta_M \tilde{r} = UE_M (\tilde{d}) = U \sqrt{E(a^2)/n} \tilde{b},
\]

(S55)

where the vector \(\tilde{b}\) parameterizes the strength and direction of the bias and \(b = ||\tilde{b}||\) is assumed to be \(<< 1\).

At steady state, the mutational shift must be offset by selection, such that

\[
\Delta_M \tilde{r} + \Delta_D \tilde{r} + \Delta_S \tilde{r} = 0,
\]

(S56)

where \(\Delta_D \tilde{r}\) and \(\Delta_S \tilde{r}\) are the expected shifts due to directional and stabilizing selection, respectively. We previously found that the expected directional shift is

\[
\Delta_D \tilde{r} = -\frac{\sigma^2}{w^2} \tilde{r},
\]

(S57)

where \(\tilde{r}\) denotes the mean phenotype (cf. Eq. S45). As we show next, when mutations are strongly selected, stabilizing selection offsets the mutational shift to maintain the mean phenotype at the optimum, implying that directional selection is negligible. In contrast, when mutations are effectively neutral, stabilizing selection is
negligible and a directional term might not be negligible by comparison. However, as long as asymmetry is small, $b \ll 1$, we show that this directional term is not large enough to change the allele dynamics, both when all mutations are effectively neutral and when some mutations are strongly selected.

First, we consider the shift in mean phenotype due to stabilizing selection. This shift arises because, with asymmetric mutational input, the distribution of phenotypes becomes skewed. Therefore, even if the mean phenotype is at the optimum, individuals with a given fitness may have an asymmetric distribution of phenotypes around the optimum, leading stabilizing selection to change the mean phenotype. We have already shown (Eq. S15) that the expected change in allele frequencies per generation due to stabilizing selection at any given site $i$ is

$$
E(\Delta q_i) = -\frac{a_i^2}{4w^2}p_iq_i\left(\frac{1}{2} - q_i\right).
$$

(S58)

The expected change in mean phenotype can then be calculated by adding up the contributions over sites

$$
\Delta_s\tilde{r} = -E\left(\sum_i \tilde{a}_i \frac{a_i^2}{4w^2}p_iq_i\left(\frac{1}{2} - q_i\right)\right).
$$

(S59)

The right-hand side of this equation reflects the skewness of the phenotypic distribution. Indeed, in one dimension, it can be shown that

$$
\Delta_s\tilde{r} = -\frac{\mu_3(r)}{2w^2},
$$

(S60)

with $\mu_3(r)$ being the third central moment of the phenotypic distribution. In $n$-dimensions, for every direction $x$,

$$
\Delta_s\tilde{r}_x = -\frac{1}{2w^2}E((\tilde{r} - \tilde{r})_x((\tilde{r} - \tilde{r})^2)) = -\frac{1}{2w^2}\sum_k \mu_3(\tilde{r})_{xkk},
$$

(S61)

with $\mu_3(\tilde{r})_{jkl} = E((\tilde{r} - \tilde{r})_j(\tilde{r} - \tilde{r})_k(\tilde{r} - \tilde{r})_l)$.

When sites are under strong selection, $\Delta_s\tilde{r}$ takes a simple form. Assuming the asymmetry is small, the shift due to stabilizing selection can be expanded in orders of $b$. The leading term in the frequency distribution takes the same form as it does
without the bias. For strongly selected alleles with no bias, \( q \ll 1 \) and therefore the frequency dependence in this term can be approximated by \( pq \left( \frac{1}{2} - q \right) \approx \frac{1}{2} q \).

Moreover, \( q \) scales with \( 1/a^2 \), implying that the distribution of \( a^2 q \) is independent of \( \tilde{a} \) and that \( E(a^2 q) = 4w^2/N \) (cf. Section S3.1). Therefore, when all sites are strongly selected, the leading term in the shift due to stabilizing selection is

\[
\Delta^S a = -E \left( \sum_i \tilde{a}_i \frac{a_i^2}{4w^2} q_i \right) = -\frac{E(a^2 q)}{8w^2} E(\sum_i \tilde{a}_i) = -\frac{1}{2N} 2NUE_M(\tilde{a}) = -U \sqrt{\nu_s \sqrt{E(S)}} b^* \\
= -\Delta_M \tilde{F}.
\]

(S62)

Thus, to a first order in \( b \), the shift of the mean phenotype due to stabilizing selection offsets the mutational shift, implying that there will be no directional term and that the allele dynamics will not be affected by asymmetry.

When alleles are instead effectively neutral, then \( a^2/4w^2 \ll 1/2N \) (cf. Section S2.2) and allele frequencies are well approximated by the neutral sojourn time, \( \tau(q) \approx 2/q \). The shift due to stabilizing selection then satisfies

\[
\Delta_S \tilde{F} = -E \left( \sum_i \tilde{a}_i \frac{a_i^2}{4w^2} p_i q_i \left( \frac{1}{2} - q_i \right) \right) \approx -E \left( pq \left( \frac{1}{2} - q \right) \right) E \left( \sum_i \tilde{a}_i \frac{a_i^2}{4w^2} \right) = \\
-\frac{1}{6} E \left( \sum_i \tilde{a}_i \frac{a_i^2}{4w^2} \right) \ll -\frac{1}{2N} E \left( \sum_i \tilde{a}_i \right) = -\Delta_M \tilde{F},
\]

(S63)

implying that it makes a negligible contribution to offsetting the mutational shift. In this case, the mutational effect on the mean phenotype is therefore offset by directional selection, where

\[
\Delta_D \tilde{F} = -\frac{\sigma^2}{w^2} \tilde{F} \approx -\Delta_M \tilde{F},
\]

(S64)

indicating a displacement of the mean phenotype from the optimum

\[
\tilde{F} = \frac{w^2}{\sigma^2} \Delta_M \tilde{F}.
\]

(S65)

This displacement introduces a directional selection term into the first moment of change in allele frequency that, if large enough, could alter allele dynamics (cf. Section S4.4).

However, when all alleles are effectively neutral, we have
\[
\tilde{r} = \frac{w^2}{\sigma^2} \Delta_M \approx \frac{w^2}{2Nuv_s \sqrt{E(S)}} U \sqrt{v_s \sqrt{E(S)}} \bar{b} = \frac{1}{2N} \frac{2w^2}{\sqrt{v_s \sqrt{E(S)}}} \bar{b},
\]
and therefore the scaled directional selection coefficient, for an allele with effect size \( \bar{d} \) and scaled stabilizing selection coefficient \( S = \frac{N\sigma^2}{2w^2} \) is of the order of

\[
2N \frac{\tilde{r} \bar{d}}{2w^2} \approx \frac{b_a a}{\sqrt{v_s \sqrt{E(S)}}} \sim (b/\sqrt{n})a = \frac{b \sqrt{v_s \sqrt{E(S)}}}{\sqrt{v_s \sqrt{E(S)}}} \approx \frac{S}{E(S)} b,
\]
with \( b_a \approx b/\sqrt{n} \) being the projection of \( \tilde{r} \) in the direction of \( \bar{d} \). Since \( b \ll 1 \), for all alleles other than those with unusually large selection coefficients, the scaled directional selection coefficient will be much smaller than 1 and the trajectories will still be determined by drift and not selection. Even in this case, therefore, we do not expect asymmetry to affect allele dynamics.

Next, we consider the case where there is a mix of effectively neutral and strongly selected mutations. The existence of strongly selected mutations in addition to effectively neutral ones reduces the deviation of the mean phenotype from the optimum. Denoting the proportion of strongly selected mutations by \( p_s \), we have

\[
\tilde{r} = \frac{w^2}{\sigma^2} U (1 - p_s) \sqrt{v_s \sqrt{E(S e.n.)}} \bar{b},
\]
where \( E(S e.n.) \leq 1 \) is the mean scaled stabilizing selection coefficient for effectively neutral mutations. Since \( \sigma^2 > 2NU p_s v_s \), we can then obtain an upper bound to the magnitude of scaled directional selection coefficient for an allele with effect size \( \bar{d} \) and scaled stabilizing selection coefficient \( S = \frac{N\sigma^2}{2w^2} \):

\[
2N \frac{\tilde{r} \bar{d}}{2w^2} = 2N \frac{1}{2 \sigma^2} U (1 - p_s) \sqrt{v_s \sqrt{E(S e.n.)}} \bar{b} \cdot \bar{d} \leq \frac{1}{2p_s} \frac{1 - p_s}{v_{S e.n.}} \sqrt{E(S e.n.)} \sqrt{S} \bar{b}.
\]
With a substantial proportion of strongly selected sites, \( \frac{1 - p_s}{2p_s} \) is of the order of 1, and therefore \( \frac{1 - p_s}{2p_s} \sqrt{E(S e.n.)} \bar{b} \ll 1 \). This condition implies that for effectively neutral alleles (i.e., \( S \leq 1 \)), the scaled directional selection coefficient is \( \ll 1 \) and allele trajectories will be determined by genetic drift, whereas for strongly selected alleles...
(i.e., when $S \gg 1$), the scaled directional selection coefficient is $\ll S$ and therefore negligible compared to the scaled stabilizing selection coefficient.

Weakly selected alleles (with $S \sim 10$) behave largely like strongly selected alleles except that stabilizing selection on them only partially cancels out the mutational bias (for example, for $S = 10$ only 85% of the bias is canceled). The rest of the bias is canceled by directional selection and therefore induces a small shift in the mean phenotype. It is straightforward to repeat the arguments given above and show that the shift in the mean phenotype for a trait with only weakly selected alleles or a mixture that includes weakly selected alleles negligibly affects allele trajectories.

Thus, we conclude that small asymmetry in mutation will not affect the allelic dynamic (see Fig. S5).

**Figure S5.** The effect of asymmetric mutational input on mean phenotype and the contribution of sites to variance. (a) The mean phenotype $\bar{r}$, in units of $\sqrt{v_s}$, as a function of mutational bias $b$. (b) Proportion of genetic variance as a function of the threshold contribution to variance $v^*$, i.e., $G(v^*)$, for different bias strengths. Simulations were run with $N = 1,000$, $n = 1$ and with different mixtures of effectively neutral (distributed around $S = 0.1$) and strong (distributed around $S = 50$) selection coefficients. Asymmetry was simulated by having more trait increasing
than trait decreasing mutations; if $\beta$ is the proportion of trait increasing mutations then the asymmetry coefficient is $b = 2\beta - 1$. As expected, only for large biases (when $b \sim 1$), are there substantial changes in the distribution of the contribution of sites to variance. Simulations were run with a 10,000 generations burn-in period without asymmetry and then 10,000 generations with asymmetry and averaged over many runs (>300), with the number of runs varied across plots keep errors in (b) below 1%.

### 5.3. Major effect loci

In this section, we show that our results are insensitive to the presence of major loci, i.e., individual loci that contribute substantially to quantitative genetic variation. We have in mind, for example, loci whose alleles are maintained at high frequencies by balancing selection on a Mendelian trait but have pleiotropic effects on the quantitative traits under consideration (e.g., MHC loci; (19, 20)). While such loci violate our assumptions, we show that they do not affect the dynamics at other loci that fulfill them.

To this end, we calculate the first two moments of change in allele frequency in the presence of a major locus. We denote the frequency and effect size of the focal allele by $q$ and $\bar{a}$, and the frequency and effect size of the major allele by $q_M$ and $\bar{a}_M$, respectively. As in our previous derivations (Section S2), the distribution of background phenotypic contribution from all other loci, $\bar{R}$, is well approximated by the normal distribution

$$f(\bar{R} | \bar{a}_M, q_M, \bar{a}, q) = \frac{1}{(2\pi(\sigma^2 - \sigma^2_M))^{n/2}} \exp\left(-\frac{(\bar{R} + q \bar{a} + q_M \bar{a}_M)^2}{2(\sigma^2 - \sigma^2_M)}\right),$$

(S70)

where $\sigma^2_M$ is the contribution to genetic variance from the major locus. The population mean remains close to the optimum because any shift caused by the major locus is quickly compensated for by the other loci (see Section S4.4). We then average over both this distribution and the three genotypes at the major locus to calculate the mean fitness associated with each genotype at the focal locus. Namely,
\[ W_{00} = (1 - q_M)^2 \int_{\mathbf{R}} f(\mathbf{R}|\tilde{a}_M, q_M, \tilde{a}, q)W(\mathbf{R}) \]
\[ + 2q_M (1 - q_M) \int_{\mathbf{R}} f(\mathbf{R}|\tilde{a}_M, q_M, \tilde{a}, q)W(\mathbf{R} + \tilde{a}_M/2) \]
\[ + q_M^2 \int_{\mathbf{R}} f(\mathbf{R}|\tilde{a}_M, q_M, \tilde{a}, q)W(\mathbf{R} + \tilde{a}_M), \tag{S71} \]

and similarly for the other genotypes. In this way, we obtain the first moment of the change in allele frequency

\[ E(\Delta q) = -pq \frac{p(W_{00} - W_{01}) + q(W_{01} - W_{11})}{W} \approx -\frac{a^2}{4w^2}pq \left( q - \frac{1}{2} \right), \tag{S72} \]

which is the same as we derived in the absence of a major locus (Eq. S15). Similarly, we find the second moment to be unaffected.

### 5.4. Anisotropic mutation

In this section, we consider how relaxing the assumption that the distribution of newly arising mutations is isotropic in trait space would affect our results. As noted, we can always choose an orthonormal coordinate system centered at the optimum, in which the trait under consideration varies along the first coordinate and a unit change in other traits (i.e., in other coordinates) near the optimum have the same effect on fitness. There is, however, no obvious reason for the distribution of newly arising mutations to be isotropic in this coordinate system (see (21) for generalizations of Fisher’s Geometric Model along similar lines).

Anisotropy in mutation does not affect the moments of change in allele frequency, as these depend only on the selection on an allele or equivalently on its effect size but not on its direction in trait space. Anisotropy could affect the distribution of allelic effect sizes on the focal trait conditional on the selection acting on them. Here, we provide heuristic arguments suggesting that, barring extreme cases, we can define an effective number of traits \( n_e \) and an effective strength of selection \( w_e^2 \) for which the relationship between selection and effect size in anisotropic models is well approximated by the relationship found for isotropic ones (Section S1.2).
We focus on a family of anisotropic mutational distributions that can be described as a projection of a multivariate normal distribution on the unit sphere in trait space. Namely, we draw the size of a mutation $a = ||\vec{a}||$ from some distribution and to obtain its direction, we draw a vector $\vec{a}$ from a multi-variate normal distribution $\text{MVN}(0, \Sigma)$ and normalize it, i.e.,

$$\vec{a} = a \frac{\vec{a}}{a},$$  \hspace{1cm} (S73)$$
and therefore

$$a_1 = a \frac{\alpha_1}{a_1}. \hspace{1cm} (S74)$$

This family of mutational distributions gives us a mathematically tractable framework with which to examine the behavior of our model with anisotropy.

With anisotropy, the behavior of our model greatly depends on the relative contribution of the focal trait to selection, which we parameterize by

$$\gamma_1 = \frac{\mathbb{E}(\alpha_1^2)}{\mathbb{E}(\alpha^2)} = \frac{\Sigma_{11}}{\text{tr}(\Sigma)}. \hspace{1cm} (S75)$$

When selection acts mainly on our focal trait, i.e. when $\gamma_1 \approx 1$, then $|a_1| \approx a$ and therefore $a_1 \approx \pm a$. Such a relationship between the strength of selection and effect size is well approximated by an isotropic model with $n_e = 1$. We therefore focus on cases in which there is a significant pleiotropic contribution to selection, i.e., $\gamma_1$ is substantially less than 1. Anisotropy then has two effects: the first is to introduce heterogeneity in the strength of selection on different traits and the second is to introduce correlations in the effects of a mutation on different traits, notably between the focal trait and others.

We first consider the case in which the strength of selection differs among traits, but traits are uncorrelated, corresponding to a diagonal covariance matrix, $\Sigma$. When many traits have a non-negligible contribution to selection, $\alpha^2 = \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2$ would have a small coefficient of variation, i.e., $C_V^2(\alpha^2) = \text{V}(\alpha^2)/\mathbb{E}(\alpha^2)^2 \ll 1$, because of the law of large numbers. In this case,
\[ a_1 = a \frac{\alpha_1}{\|a\|} = a \frac{\alpha_1}{\sqrt{E(\alpha^2)}} \left( 1 + O\left(C_Y^2(\alpha^2)\right) \right) \approx a \frac{\alpha_1}{\sqrt{E(\alpha^2)}} \]

\begin{equation}
= \frac{a}{\sqrt{1/\gamma_1}} \frac{\alpha_1}{\sqrt{E(\alpha^2)}},
\end{equation}

(S76)

Since \( \alpha_1/\sqrt{E(\alpha^2)} \sim N(0,1) \) and \( S = \frac{N}{2w^2} a^2 \) this implies that, conditional on the selection coefficient \( S \), the effect size on the focal trait will be distributed as

\[ a_1 \sim N \left( 0, \frac{2w^2}{n_eN} S \right) \quad \text{(S77)} \]

with \( n_e = 1/\gamma_1 \). This is the same relationship between selection and effect size as the high pleiotropy isotropic model with \( n = n_e \) (Eq. 8 & 10 of main text). This result suggests the concept of an effective number of traits, which can be thought of as the number of traits that have the same effect on fitness as the focal one and are required to produce the same strength of selection on alleles. The effective number of traits describes the distribution of effect sizes both in the limit of high pleiotropy \( n_e \gg 1 \) and low pleiotropy \( n_e \approx 1 \) and simulations show that it describes the distribution, at least qualitatively, also for intermediate values of \( n_e \) (Fig. S6).

However, there is an extreme scenario in which an effective number of traits cannot describe the distribution of effect sizes. This happens when \( C_Y^2(\alpha^2) \geq 1 \), that is when selection acts mainly on a small number of traits but our focal trait contributes very little to selection \( (\gamma_1 \ll 1) \). In this case, we might be tempted to use \( n_e = 1/\gamma_1 \gg 1 \) but, as Eq. S76 suggests, the high pleiotropy limit would be inadequate. In fact, the variance in selection on newly-arising mutations (due to the contribution of the selected traits) will result in a long-tailed distribution of effect sizes on the focal trait, which is not well-approximated by any isotropic model. In summary, except for these extreme cases, isotropic models provide a good approximation for the relationship between selection and effect size, even when there is heterogeneity in the strength of selection on different traits.

To illustrate the effect of heterogeneity in the strength of selection among traits, we consider a simple example in which all non-focal traits make the same contribution to selection and therefore can be modeled by
\[ \Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c^2 & 0 \\ 0 & 0 & \vdots \end{pmatrix} \]  \hspace{1cm} (S78)

with \( c^2 \) being the ratio between the expected contribution of a non-focal trait to selection and the expected contribution of the focal trait to selection. It is easy to see that in this model \( n_e = 1/\gamma_1 = 1 + (n - 1)c^2 \). Numerical results for this model are shown in Fig. S6.

Figure S6. The effects of heterogeneity in the strength of selection on different traits on the distribution of effect sizes in the focal trait. Numerical results for models with the correlation matrix defined in Eq. S78 are shown in blue and the corresponding isotropic model in black dashes. When there are many selected traits, an isotropic model with \( n_e = 1/\gamma_1 = 1 + (n - 1)c^2 \) provides a good approximation of the distribution of effect sizes, both when the focal trait contributes substantially to selection (a) and when it does not (b). When there are few traits, an isotropic model
with \( n_e = 1/y_1 = 1 + (n - 1)c^2 \) provides a good approximation only when the focal trait contributes substantially to selection (c & d).

Next, we consider the case in which the effect sizes on different traits are correlated, i.e., when the covariance matrix \( \Sigma \) has off-diagonal terms. \( a_1^2 \propto a_1^2/\alpha^2 \) and therefore we parameterize the effect of these terms using the correlation between \( a_1^2 \) and \( \alpha^2 \), \( \rho^2 \equiv \text{corr}(a_1^2, \alpha^2) \). If the correlation is small, \( \rho^2 \ll 1 \), then our previous reasoning holds. In the other extreme, when all selected traits are highly correlated with the focal trait, i.e. \( \rho^2 \approx 1 \), then the proportional contribution of the focal trait to selection is constant, \( \frac{a_1^2}{\alpha^2} \approx \frac{E(a_1^2)}{E(\alpha^2)} = \gamma_1 \), and the effect size is \( a_1 = \pm \sqrt{\gamma_1} a \). This model is therefore equivalent to an isotropic one with \( n_e = 1 \) and \( w_e^2 = \gamma_1 w^2 \); the latter change corresponds to increasing the strength of selection on the focal trait to account for selection on the other traits which are highly correlated with the focal trait. Intermediate cases are more complex: while effect sizes are still of the order of \( \sqrt{\gamma_1} a \), the shape of the distribution of effect sizes is intermediate between the single trait and high pleiotropy limits. Isotropic models with an effective number of traits, \( n_e < 1/\gamma_1 \), and increased selection \( w_e^2 = \gamma_1 n_e w^2 \) can qualitatively describe these cases but may not completely capture the distribution of effect sizes. The value of \( n_e \) would change from 1 when \( \rho^2 \rightarrow 1 \) to \( 1/\gamma_1 \) when \( \rho^2 \rightarrow 0 \). Note that with a large number of traits, very strong correlations among many of the traits will be necessary in order to create a large enough \( \rho^2 \) to have a significant effect on \( n_e \) (see Fig. S7).

To illustrate the effect of correlations among traits, we consider the following simple example (Figure S7). We assume the correlation matrix \( \Sigma \) takes the form

\[
\Sigma = \begin{pmatrix}
1 & r^2 & r^2 & \cdots \\
r^2 & 1 & r^2 & \cdots \\
r^2 & r^2 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

meaning that that all traits contribute equally to the fitness and every pair of traits has the same correlation coefficient \( r^2 \). When \( r^2 = 0 \) this becomes an isotropic
model. When \( r^2 = 1 \), effect sizes are always identical for every trait; thus, this case is equivalent to having only one trait with selection that is increased \( n \)-fold.

Intermediate cases can be approximated by finding an effective number of traits \( n_e < 1/\gamma_1 = n \), such that an isotropic model with \( n_e \) and \( w_e^2 = \gamma_1 n_e w^2 = w^2 n_e / n \) qualitatively describes the distribution of effect sizes. Numerical results of this model are shown in Fig. S7.

Figure S7. Effects of correlations among traits on the distribution of effect sizes.

Numerical results for our model with the correlation matrix defined in Eq. S79 and \( n = 50 \) traits are shown in blue and the corresponding isotropic model in black dashes. (a) When correlations are low, the isotropic model approximates the distribution of effect sizes well. (b) With large correlations, we need to use an effective number of traits, in this example \( n_e = 5 \), and rescale selection, in this case to \( w_e^2 = w^2 n_e / n = w^2 / 10 \), in order to approximate the distribution of effect sizes. (c) When the correlations approach 1, the distribution of effect sizes becomes singular and approaches the distribution for an isotropic model with \( n_e = 1 \) and \( w_e^2 = w^2 / n = w^2 / 50 \).

6. The power to detect loci in GWAS

In this section, we summarize the results that we rely on in connecting our theoretical results with the observations in GWAS (see Discussion in main text). These results provide a first approximation to the power to detect loci in GWAS in re-sequencing and genotyping studies. They neglect some potential complications, which lie beyond the scope of this study (e.g., synthetic associations (22, 23)).
6.1. Re-sequencing studies

First, we consider how the power to detect a locus in a GWAS depends on its contribution to genetic variance. To this end, we follow Sham and Purcell (22) in assuming a simplified model for a GWAS in which loci are detected using a linear regression of the phenotype against the genotype at individual loci, and the dependence of phenotype on genotype follows an additive model. The slope of the regression (the regression coefficient), which is also our estimate of the effect size, \( \hat{a}_1 \) is then approximately normally distributed as

\[
\hat{a}_1 \sim N \left( a_1, \frac{\nu_p}{mx(1-x)/2} \right),
\]

where \( a_1 \) is the true effect size and \( x \) is the minor allele frequency at the locus (which, due to the large study sizes, we assume is estimated without error), \( \nu_p \) is the total phenotypic variance, and \( m \) the study size (which in reality may be an effective size reflecting study design, e.g., when the sample was split into discovery and validation panels) (22).

Under the null hypothesis, the effect size is 0, meaning that

\[
\hat{a}_{1\text{null}} \sim N \left( 0, \frac{\nu_p}{mx(1-x)/2} \right)
\]

and therefore the estimated contribution to variance has a chi-squared distribution

\[
\frac{\hat{\vartheta}_{\text{null}}}{\nu_p/m} = \frac{\hat{a}_{1\text{null}}^2 x(1-x)}{\nu_p/m} \sim \chi_1^2.
\]

The power to identify a locus as significant with p-value \( p^* \) is the probability that the estimated contribution of the locus to variance, \( \hat{\vartheta} \), is large enough that

\[
\Pr(\hat{\vartheta}_{\text{null}} > \vartheta) < p^*.
\]

This condition can be translated into a threshold contribution to variance \( \nu^* \) for which loci with \( \hat{\vartheta} > \nu^* \) are considered significant, i.e. \( \Pr(\hat{\vartheta}_{\text{null}} > \nu^*) = p^* \), with \( \nu^* \) given by
\[ \frac{v^*}{V_p/m} = 2 \left( \text{erf}^{-1}(1 - p^*) \right)^2, \quad (S84) \]

and \( \text{erf} \) denoting the error function. The power to identify a locus as significant would then be \( \Pr(\theta > v^*) \), and the distribution \( \theta \) given by

\[ \frac{\theta}{V_p/m} = \frac{1}{2} \chi^2_x \left( \frac{v}{V_p/m} \right), \quad (S85) \]

with \( \chi^2_1 \) denoting a non-central chi-squared distribution with one degree of freedom. Therefore, power is given by

\[ H(v, p^*) = \Pr(\theta > v^*) = h_+ \left( \frac{v}{V_p/m}, p^* \right) + h_- \left( \frac{v}{V_p/m}, p^* \right), \quad (S86) \]

with \( h_\pm(y, p^*) = \frac{1}{2} \left( 1 \pm \text{erf} \left( \frac{\sqrt{y/2} \mp \text{erf}^{-1}(1 - p^*)} \right) \right) \) and the two terms correspond to the estimated and true effect sizes having the same or opposite sign.

The form of the power function carries important implications (Eq. S86 and Fig. S8). Notably, it shows that (in this approximation) power depends only on the contribution of a locus to variance, and this contribution should be measured relative to, or in units of, \( V_p/m \). This scale makes intuitive sense, because the total phenotypic variance generates the background noise for detecting any individual locus, and the background noise decreases is inverse proportional to the study size. In particular, the threshold contribution to variance \( v^* \), as defined above, is proportional to \( V_p/m \) and is also the contribution to variance at which power is 50%, i.e.,

\[ H(v^*, p^*) = 1/2. \quad (S87) \]

The power function can then be approximated by a step function (see Fig. S8)

\[ H(v) \approx \Theta(v - v^*) = \begin{cases} 1 & v > v^* \\ 0 & v < v^* \end{cases}. \quad (S88) \]
This will be a good approximation when the number of loci that fall at intermediate range (e.g., with power between 0.1 and 0.9) is negligible compared to the number that falls outside this range.

**Figure S8.** The power to detect loci as a function of their contribution to genetic variance (given in units of $V_P/m$). Shown are the exact power function (Eq. S86) and its step function approximation (Eq. S88) for $p = 5 \cdot 10^{-8}$.

Further insights come from considering this power function in conjunction with our theoretical results (Section S3). Notably, our results suggest that the first loci to be detected, those that contribute the most to variance, are weakly and strongly selected, and that their contributions to variance are on the scale of $v_s$. We therefore expect GWAS to begin to identify loci (and account for genetic variance) when the study size is such that $v^* \propto V_P/m$ is on the order of $v_s$, i.e., when $m \sim V_P/v_s$. We would further expect the rate of increase in identifying new loci (and accounting for the variance) to be similar for different traits when the variance is measured in units of $v_s$.

### 6.2. Genotyping

Most current GWAS rely on genotyping instead of re-sequencing, resulting in an additional loss of power (24). Specifically, these studies impute the alleles at loci
that are not included in the genotyping platform (25), and the imputation becomes imprecise when the imputed alleles are rare (Fig. S9). If causal loci with rare alleles are included in GWAS, this imprecision leads to an under-estimation of their effect size, resulting in a loss of power (24). For loci with MAF $x$ and effect size $a$, the expected estimate of the effect size would be reduced by a factor of $r(x)$, where $r^2(x)$ is the mean correlation between the imputed and real alleles (26), and the distribution of estimates can be approximated by

$$\hat{a} \sim N \left( r(a) \frac{v_p}{m x (1-x)/2} \right).$$ (S89)

Employing the reasoning of the previous subsection, we can therefore approximate the power to detect a locus by $H(r^2 v, p^*)$, where $H$ is the power function defined in Eq. S86.

**Figure S9.** The precision of imputation decreases with MAF. Specifically we show the mean correlation between imputed and real genotypes as function of minor allele frequency, for a study using an Illumina 1M SNP array and the 1000 genomes phase III as an imputation panel (based on Extended Fig. 9A in (27)). We approximate the effect on power by excluding loci with MAF < 1% and assuming that loci with greater MAFs are imputed correctly.

In practice, GWAS often include only loci with MAF above a threshold, which is chosen to ensure precise imputation. We therefore approximate the effect of
genotyping on power by excluding loci below a threshold MAF and assume that loci that exceed this threshold are imputed correctly.

7. **Inference**

In this section, we describe how we used our model to make inferences based on GWAS results for height and body mass index (BMI). As we note in the Discussion, these inferences are meant as an illustration and do not incorporate the effects of demography and a few other factors (e.g., genotyping and errors in the estimation of effect sizes (22, 24)), which lie beyond the scope of this study.

7.1. **The composite likelihood**

Our inferences are based on a composite-likelihood approach. We begin by describing the composite-likelihood function and its maximization, when the loci detected by GWAS are strongly selected and can be described by the high-pleiotropy limit. In this case, we have shown that the distribution of variance among loci is insensitive to the distribution of selection coefficients, depends on a single parameter $\nu$, and is well approximated by the probability density

$$
\rho(\nu) = \frac{2 \exp(-2\sqrt{\nu/v_s})}{v} \quad (S90)
$$

(Section S3.2). Further approximating the power in GWAS as a step function (cf. Section S6), we find that the probability density of sites that exceed a threshold $\nu^*$ can be approximated by

$$
f(\nu|\nu_s, \nu^*) = \frac{\rho(\nu)}{\int_{\nu > \nu^*} \rho(\nu)} = \frac{\exp(-2\sqrt{\nu/v_s})}{2\nu \Gamma(2\sqrt{\nu/v_s})}, \quad (S91)
$$

where $I(x) \equiv \int_{t>0} \exp(-t)/t$ (cf. Eq. S42). We therefore approximate the log-composite-likelihood of $\nu_s$ given the contributions to variance of the $K$ loci detected in a GWAS, $\{v_i\}_{i=1}^K$, by

$$
\text{LCL}(\nu_s|\{v_i\}_{i=1}^K, \nu^*) = \sum_{i=1}^K \log(f(v_i|\nu_s)) = -\left(\frac{2}{\sqrt{\nu_s}}\right) \sum_{i=1}^K \sqrt{v_i} - K \log \left(1\left(2\sqrt{\nu^*/\nu_s}\right)\right) - \sum_{i=1}^K \log(\sqrt{v_i}). \quad (S92)
$$
It follows that the composite-likelihood is maximized when

\[ \tilde{v}_s = \arg\min_{v_s} \left\{ 2\sqrt{v}/\sqrt{v_s} + \log \left( 1 \left( 2v^*/v_s \right) \right) \right\}, \quad (S93) \]

where \( \sqrt{v} \equiv \frac{1}{K} \sum_{i=1}^{K} \sqrt{v_i} \).

We also consider the models without pleiotropy and in which the degree of pleiotropy is a parameter. In the case without pleiotropy,

\[ \rho(v) = \frac{2\exp(-2v/v_s)}{v} \quad (S94) \]

(cf. Section S3.2). By following the same steps, we find that the composite-likelihood is then maximized when

\[ \tilde{v}_s = \arg\min_{v_s} \left\{ 2\sqrt{v}/v_s + \log \left( 1 \left( 2v^*/v_s \right) \right) \right\}, \quad (S95) \]

where \( \sqrt{v} \equiv \frac{1}{K} \sum_{i=1}^{K} v_i \). When the degree of pleiotropy \( n \) is a parameter of the model, we found that

\[ \rho_n(v) = \int_{a_1}^{a_2} \frac{2}{v} \exp \left( -\frac{2v/v_s}{a_2/(a^2/n)} \right) \phi_n(a_1|a) \quad (S96) \]

(cf. Section S6). Again, following the same steps, we find that the probability density of sites that exceed a threshold \( v^* \) is

\[ f_n(v|v_s) = \frac{\rho_n(v)}{\int_{v^*}^{\infty} \rho_n(v)} \quad (S97) \]

and the log-composite-likelihood is

\[ \text{LCL}(v_s, n | \{v_i\}_{i=1}^{K}, v^*) = \sum_{i=1}^{K} \log \left( \rho_n(v_i) \right) - K \log \left( \int_{v^*}^{\infty} \rho_n(v) \right). \quad (S98) \]

In the latter case, we used numerical maximization to show that the composite-likelihood estimates for height and BMI converge to the high pleiotropy limit. Specifically, we maximized the composite-likelihood specifying an interval of \([1, 1000]\) for \( n \), where for both traits the estimates converged to the upper limit of 1000. While numerical optimization does not allow us to specify an infinite interval, the likelihood function and maximal value for \( n=1000 \) are indistinguishable from those in the high-pleiotropy limit.
7.2. Determining \(v^*\) and removing outliers

Our likelihood maximization requires us to specify the value of the threshold \(v^*\). We choose this threshold based on the empirical distributions of the contributions to variance among genome-wide significant associations (Fig. S10A and B). Specifically, when the contributions to variance approach the lower boundary for discovery, we observe a decline in the density of loci. This is likely due to a gradual reduction in power and suggests that our approximation for power (as a step function) breaks down for these values of variance. We therefore choose thresholds that appear to be above this decline (\(v^* = 1.4 \cdot 10^{-4}V_P\) for height and \(v^* = 1.35 \cdot 10^{-4}V_P\) for BMI; Fig. S10A and B), resulting in the removal of 53 loci for height and 11 for BMI. We also examine how our estimates of \(v_s\) depend on the choice of \(v^*\), and find that they are much more sensitive to reducing the threshold than to increasing it; in fact, the estimates we obtain by increasing the threshold are within the confidence intervals of the estimate with the chosen thresholds (Fig. S10C and D). This analysis further supports our choice to exclude the loci with the lowest contribution to variance. For BMI, we also dropped the locus with the largest contribution to variance (near FTO), which appears to be an outlier (Fig. S10B) and has been suggested to be under balancing selection (28).
Figure S10. Determining $\nu^*$ and removing outliers. The total variance from significant associations as a function of the threshold contribution to variance, for height (a) and BMI (b). The insets show a close up of the lower range of contributions to variance, highlighting the decline in the density of discovered loci. Our chosen thresholds are shown by the dashed vertical line (in all graphs). Our estimates of $\nu_s$ as a function of the chosen threshold, for height (c) and BMI (d). When we increase the threshold, the estimates remain within the 95% CI of the estimate with our chosen threshold.

### 7.3. Estimating target size and explained variance

We estimate the target size and the variance explained, both for varying study size and total, based on our estimates of $\nu_s$. The population-scaled mutational input per generation from strongly selected loci, $2NU_s$, is estimated by

$$2NU_s = K/\int_{\nu>\nu^*} \rho(\nu|\theta_s),$$

(cf. Eq. S41) and the corresponding estimate for the target size is
\[ L_s = \frac{2NU_s}{2Nu}, \]  
(S100)

where the estimate for the population scaled mutation rate per site per generation \( 2Nu \approx 0.5 \) is based on heterozygosity (27). The explained variance corresponding to GWAS with study size \( m \) is estimated by

\[ \sigma^2 \sim (m) = \frac{2NU}{\int_{v>V*} v\rho(v|\hat{v})} = K \left( \frac{\int_{v>V*} v\rho(v|\hat{v})}{\int_{v>V*} \rho(v|\hat{v})} \right), \]  
(S101)

where we approximate the threshold corresponding to study size \( m \) based on the study size, \( m_0 \), and threshold, \( v^* \), in current GWAS, by

\[ v^*(m) = v^* \cdot \left( m_0 / m \right). \]  
(S102)

To estimate the total variance arising from strongly selected loci, we simply set the threshold in Eq. (S101) to 0.

### 7.4. Estimating confidence intervals

We use a combination of non-parametric and parametric bootstrap to estimate confidence intervals (CI). We use non-parametric bootstrap to estimate the CI for the model parameters \( v_s \) and \( L_s \): specifically, we perform 10,000 iterations, in which we resample the loci identified by GWAS and repeat the estimation of \( v_s \). We use parametric bootstrap to estimate the confidence intervals in Fig. 5A, describing the explained variance as a function of threshold based on our model. To that end, we rely on our model with the point estimates for \( v_s \) and \( L_s \), to generate 10,000 samples from GWAS with the specified threshold, and then calculate the total variance explained by these samples. We use a combination of non-parametric and parametric bootstrap to calculate the CI for model predictions, including the total variance, \( \sigma^2 \), and the explained variance, \( \sigma^2 (m) \), and number of loci as a function of study size (Figs. 5B and S16). In this case, we generate 10,000 samples by: i) estimating \( v_s \) based on a resampled set of GWAS loci (similar to the non-parametric procedure), and ii) using the estimated \( v_s \) and corresponding \( L_s \) to generate a GWAS sample based on our model (similar to the parametric procedure); we then calculate the appropriate summary based on the latter samples. This two stage procedure is intended to capture the uncertainty generated by both the errors in estimating our
basic model parameters and the noise generated by the stochastic processes underlying the number and variance at segregating loci that are yet to be discovered. The resulting estimates and CI are summarized in Table S2.

| Contribution to variance per strongly selected locus (in units of the total phenotypic variance) $\frac{\theta_s}{V_p}$ | Height | BMI |
|---|---|---|
| $1.8 \times 10^{-4}$ | $1 \times 10^{-4}$ |

| Expected study size required to describe 50% of the strongly selected variance $\frac{m_{50\%}}{43V_p/\theta_s}$ | Height | BMI |
|---|---|---|
| $230 \times 1000$ K | $420 \times 250$ K |

| Number of newly arising strongly selected mutations per generation in the population $2NU_s$ | Height | BMI |
|---|---|---|
| $2300 \times 1800$ | $600 \times 3000$ |

| Mutational target size for strongly selected mutations $L_s$ | Height | BMI |
|---|---|---|
| $4.6 \times 3.6$ Mbp | $1.3 \times 0.6$ Mbp |

| % contribution to phenotypic variance from strongly selected loci $\frac{\sigma_s^2}{V_p}$ | Height | BMI |
|---|---|---|
| $42 \times 39, 45$ % | $7 \times 5, 10$ % |

| Proportion of heritability from strongly selected loci $\frac{\sigma_s^2}{V_G} = \frac{\sigma_s^2}{h^2 V_p}$ | Height | BMI |
|---|---|---|
| $53 \times 49, 57$ % | $13 \times 10, 21$ % |

**Table S2.** Parameter estimates and their confidence intervals for height and BMI based on GWAS results; the heritability was assumed to be 0.8 for height and 0.5 for BMI (8, 10).
7.5. Testing goodness of fit

We use the Kolmogorov-Smirn D-statistic (29, 30) to test the goodness of fit of our models without pleiotropy and in the high pleiotropy limit. Since our parameter estimates are inferred from the data that we are testing against, we cannot rely on the standard tables for the p-values. We therefore generate null distributions for the D-statistic using parametric bootstrap based on our models. Specifically: i) we generate $10^5$ samples of $K$ significant loci based on the model under consideration, with the corresponding estimate of $\nu_s$, ii) we infer $\nu_s$ based on each sample, and iii) calculate the Kolmogorov-Smirnov D-statistic between the distribution of variance for the $K$ loci in each sample and the corresponding theoretical distribution based on the $\nu_s$ inferred from that sample. The resulting distribution of D-statistics corresponds to our null hypothesis, i.e., that the loci detected in GWAS arose according to our model, and specifically to the way we calculate the D-statistic between the observed distribution of variance for the $K$ detected loci and the theoretical distribution that we inferred based on these observations. We then calculate the D-statistic, $D_r$, based on the real data and corresponding theoretical distribution, and estimate the one-sided p-value by

$$\hat{p}_{K-s} = \frac{\# \text{ simulated datasets with } D > D_r}{\# \text{ simulated datasets}}.$$  \hspace{1cm} (S103)

Note that unlike the common case, here the inability to reject the null indicates that the data is consistent with our model.
Figure S11. Q-Q plots comparing the distribution of variance among significant loci taken from the GWAS of height (10) and BMI (8) with the theoretical distributions inferred from these data, based on the models without pleiotropy (a) and in the high pleiotropy limit (b). These plots show that the model assuming high pleiotropy cannot be rejected for either trait and fits these data much better than the model assuming no-pleiotropy.
8. Glossary of notation

| Symbol | Description |
|--------|-------------|
| $i^2$  | Phenotype   |
| $W(i^2)$ | Fitness     |
| $w$    | Scale of selection |
| $n$    | Number of traits (dimensions) |
| $\bar{d}$ | $n$-dimensional effect size |
| $a_1$  | Effect size on focal trait |
| $U$    | Haploid mutation rate per generation |
| $\sigma^2$ | Phenotypic variance in a trait |
| $K$    | Number of segregating sites |
| $\varphi_n(a_1|a)$ | Distribution of focal trait effect sizes conditional on overall effect size |
| $\eta(a_1|S)$ | Distribution of focal trait effect sizes conditional on the scaled selection coefficient |
| $S = \frac{Na^2}{2w^2}$ | Scaled selection coefficient |
| $q$    | Derived allele frequency |
| $p$    | Ancestral allele frequency, $p = 1 - q$ |
| $\tau(q|S)$ | The sojourn time for a mutation with scaled selection coefficient $S$ |
| $v$    | Contribution to variance from a site ($v = \frac{1}{2}a_1^2q(1 - q)$) |
| $v_s$  | Expected contribution of a strongly selected site to variance ($v_s = \frac{2w^2}{nN}$). |
| $E(V|S)$ | Expected contribution to genetic variance from sites under selection $S$ |
| $E(K|S)$ | Expected contribution to the number of segregating sites from sites under selection $S$ |
| $\rho(v)$ | Density of segregating sites contributing variance $v$ |
| $G(v^*)$ | Fraction of variance explained by sites contributing more than $v^*$ to the variance |
| $m$    | GWAS study size |
| $H$    | Power to identify a locus in GWAS |
9. Bibliography

1. Lande R (1975) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet Res* 26(3):221-235.
2. Poon A & Otto SP (2000) Compensating for our load of mutations: Freezing the meltdown of small populations. *Evolution* 54(5):1467-1479.
3. Ewens WJ (2004) *Mathematical population genetics 1: I. Theoretical introduction* (Springer Science & Business Media).
4. Keightley PD & Hill WG (1988) Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. *Genet Res* 52(01):33-43.
5. Kong A, et al. (2012) Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 488(7412):471-475.
6. Ward LD & Kellis M (2012) Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science* 337(6102):1675-1678.
7. Perry JRB, et al. (2014) Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* 514(7520):92-97.
8. Locke AE, et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518:197-206.
9. Scott RA, et al. (2012) Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* 44(9):991-1005.
10. Wood AR, et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 46(11):1173–1186.
11. Schork NJ, Murray SS, Frazer KA, & Topol EJ (2009) Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 19(3):212-219.
12. Hunt KA, et al. (2013) Negligible impact of rare autoimmunity locus coding-region variants on missing heritability. *Nature* 498(7453):232-235.
13. Do R, et al. (2015) Exome sequencing identifies rare ldlr and apoa5 alleles conferring risk for myocardial infarction. *Nature* 518(7537):102-106.
14. Purcell SM, et al. (2014) A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* 506(7487):185-190.
15. Chen JÅ, Peñagarikano O, Belgard TG, Swarup V, & Geschwind DH (2015) The emerging picture of autism spectrum disorder: Genetics and pathology. *Annu Rev Pathol Mech Dis* 10:111-144.
16. Uhlenbeck GE & Ornstein LS (1930) On the theory of the brownian motion. *Phys Rev* 36(5):823.
17. Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314-334.
18. Charlesworth B (2013) Stabilizing selection, purifying selection, and mutational bias in finite populations. *Genetics* 194(4):955-971.
19. Leffler EM, et al. (2013) Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science* 339(6127):1578-1582.
20. Denny JC, et al. (2013) Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotech* 31(12):1102-1111.
21. Martin G & Lenormand T (2006) A general multivariate extension of fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 60(5):893-907.
22. Sham PC & Purcell SM (2014) Statistical power and significance testing in large-scale genetic studies. *Nat Rev Genet* 15(5):335-346.
23. Dickson SP, Wang K, Krantz I, Hakonarson H, & Goldstein DB (2010) Rare variants create synthetic genome-wide associations. *PLoS Biol* 8(1):e1000294.

24. Evangelou E & Ioannidis JPA (2013) Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* 14(6):379-389.

25. Visscher PM, Brown MA, McCarthy MI, & Yang J (2012) Five years of gwas discovery. *Am J Hum Genet* 90:7-24.

26. Pritchard JK & Przeworski M (2001) Linkage disequilibrium in humans: Models and data. *Am J Hum Genet* 69(1):1-14.

27. The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526(7571):68-74.

28. Liu X, *et al.* (2015) Signatures of natural selection at the fto (fat mass and obesity associated) locus in human populations. *PLoS One* 10(2):e0117093.

29. Genest C & Remillard B (2008) Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann Inst H Poincare Probab Statist* 44(6):1096-1127.

30. Stute W, Manteiga WG, & Quindimil MP (1993) Bootstrap based goodness-of-fit-tests. *Metrika* 40(1):243-256.
10. Additional figures

Figure S12. The distribution of variance among sites as a function of minor allele frequency. This density is independent of the extent of pleiotropy because, the expected variance corresponding to a given selection coefficient and frequency depends only on the expected squared effect size, which is not affected by the degree of pleiotropy.
Figure S13. The number of loci per Mb contributing more than $v$ to the variance, as a function of $v$. Calculated for a one-dimensional model without pleiotropy (a) and for the high pleiotropy limit (b). This figure corresponds to Fig. 2 but shows the number of loci discovered per Mb instead of the proportion of the heritable variance explained. Calculated for a constant population size of 20,000, with a mutation rate of $1.2 \times 10^{-8}$ (5).
Figure S14. Theoretical predictions for GWAS of continuous traits, in the case without pleiotropy ($n = 1$). This figure is equivalent to Fig. 3 from the main text, which describes the case with pleiotropy ($n \gg 1$). (a) Proportion of heritability explained as a function of study size. (b) Proportion of heritability explained as a function of scaled selection coefficient for studies based on re-sequencing and genotyping. Calculated for a study size of $\sim 43 \frac{V_p}{v_s}$, corresponding to having 25% of strongly selected ($S>>1$) variance explained with re-sequencing. Note, that this study size is $\sim 2.5$ times larger than the one required to explain 25% of strongly selected variance with pleiotropy (see Fig. 3).
Figure S15. Prediction of the number per Mb of loci associated with a continuous trait in GWAS as a function of study size, in the limit of high pleiotropy. Calculated for a constant population size of 20,000, with a mutation rate of $1.2 \times 10^{-8}$ (5).
Figure S16. Model predictions for the number of discovered variants in future GWAS for height and BMI. The corresponding predictions for the % heritability explained appear in Fig. 4.