

Polygenicity of complex traits is explained by negative selection

Luke J. O'Connor^{1,2*}, Armin P. Schoech¹, Farhad Hormozdiani¹, Steven Gazal¹, Nick Patterson³ and Alkes L. Price^{1,3*}

[1] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

[2] Program in Bioinformatics and Integrative Genomics, Harvard Graduate School of Arts and Sciences, Cambridge, MA

[3] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA

* Correspondence should be addressed to L.J.O. (loconnor@g.harvard.edu) or A.L.P. (aprice@hsph.harvard.edu)

Complex traits and common disease are highly polygenic: thousands of common variants are causal, and their effect sizes are almost always small. Polygenicity could be explained by negative selection, which constrains common-variant effect sizes and may reshape their distribution across the genome. We refer to this phenomenon as *flattening*, as genetic signal is flattened relative to the underlying biology. We introduce a mathematical definition of polygenicity, the *effective number of associated SNPs*, and a robust statistical method to estimate it. This definition of polygenicity differs from the number of causal SNPs, a standard definition; it depends strongly on SNPs with large effects. In analyses of 33 complex traits (average $N=361k$), we determined that common variants are $\sim 4x$ more polygenic than low-frequency variants, consistent with pervasive flattening. Moreover, functionally important regions of the genome have increased polygenicity in proportion to their increased heritability, implying that heritability enrichment reflects differences in the number of associations rather than their magnitude (which is constrained by selection). We conclude that negative selection constrains the genetic signal of biologically important regions and genes, reshaping genetic architecture.

Introduction

Genome-wide association studies (GWAS) have revealed that common diseases and complex traits are heritable and highly polygenic¹⁻¹². There are usually no large-effect common SNPs, and hundreds or thousands of small-effect SNPs are required to explain a large proportion of heritability. This aspect of polygenicity presents a challenge for geneticists, as small-effect SNPs are difficult to detect (leading to “missing heritability”¹³) and difficult to interpret^{14,15}. The large number of causal SNPs could be explained by extraordinary biological complexity, e.g. if thousands of genes affect a trait under an “omnigenic model”¹⁵. However, biological complexity does not explain the absence of large-effect SNPs.

Here, we investigate a complementary hypothesis: due to negative selection, large-effect SNPs are prevented from becoming common in the population while small-effect SNPs are unaffected, resulting in increased polygenicity for common variants (Figure 1). We refer to this phenomenon as *flattening*, as the distribution of heritability across the genome is flattened relative to the underlying biology. Although it is widely known that negative selection constrains common-

variant effect sizes on average^{10,16-20}, the extent to which negative selection leads to increased polygenicity is unknown.

Flattening would lead to increased polygenicity for common vs. low-frequency SNPs and especially high polygenicity within functionally important regions of the genome (Figure 1). In order to compare polygenicity across allele frequencies and functional categories, we introduced a mathematical definition of polygenicity and developed a method, stratified LD fourth moments regression (S-LD4M), to estimate it using summary association statistics. We applied S-LD4M to summary statistics for 33 diseases and complex traits (average N=361k).

Overview of methods

In order to quantify the effects of flattening, we introduce a mathematical definition of polygenicity, the *effective number of associated SNPs* (M_a). This quantity describes the number of genetic associations that explain the heritability of a trait. It differs from the total number of causal SNPs (M_c): if a small number of SNPs explain a large proportion of heritability, and a large number of SNPs have extremely small effect sizes, then M_a will be much smaller than M_c .

If there is no linkage disequilibrium (LD) between causal SNPs, then M_a is inversely proportional to the normalized fourth moment, or kurtosis κ , of the causal effect size distribution:

$$M_a = \frac{3M}{\kappa}, \quad \kappa = \frac{E[\beta^4]}{E[\beta^2]^2}, \quad (1)$$

where M is the number of SNPs and β is the causal effect size of a SNP in per-normalized-genotype units (i.e. per-allele effect size times $\sqrt{2p(1-p)}$, where p is the allele frequency). SNPs with large effects contribute strongly to $E[\beta^4]$, leading to decreased M_a . In the presence of LD between causal SNPs, M_a is defined using mixed fourth moments of causal and marginal effect sizes, ensuring that M_a is identifiable (Methods). Under an infinitesimal (Gaussian) architecture, M_a is equal to the effective number of independent SNPs²¹.

We also define the *average unit of heritability*, denoted $E_{h^2}(\alpha^2)$, as the marginal per-SNP heritability averaged across components of heritability (not uniformly across SNPs). $E_{h^2}(\alpha^2)$ multiplied by M_a is equal to h^2 . For example, in the special case that causal SNPs are independent and their effect sizes follow a $N(0, \sigma^2)$ distribution, $E_{h^2}(\alpha^2)$ is equal to σ^2 and M_a is equal to the number of causal SNPs. More generally, $E_{h^2}(\alpha^2)$ can be visualized as the area of the shaded region in Figure 2a, and it depends strongly on large-effect SNPs (Figure 2b). For example, if 4 causal SNPs each contribute $\frac{1}{6}$ of heritability ($\frac{2}{3}$ of heritability in total) and 96 other causal SNPs have much smaller effects, then the average unit of heritability is approximately $\frac{1}{6} \times \frac{2}{3} = \frac{1}{9}$, and $M_a \approx 9$ (Figure 2a bottom).

An alternative definition of polygenicity, which is widely used, is the total number of causal SNPs (M_c)^{1,3,5-7,10-12}. We chose to estimate M_a rather than M_c for three reasons. First, M_c is difficult to estimate robustly, as it is impossible to distinguish SNPs with zero effect from SNPs with arbitrarily small effect (see Simulations). Second, because negative selection affects large-effect SNPs more strongly than small-effect SNPs, it is expected to influence M_a much more

strongly than M_c (Figure 2b). Third, M_a (but not M_c) is closely related to missing heritability and polygenic prediction accuracy (Methods).

M_a can be defined for categories of SNPs, such as low-frequency SNPs or coding SNPs. To compare categories of different size, we divide M_a by the number of SNPs in each category. We refer to differences in per-SNP M_a simply as differences in polygenicity. We define *polygenicity enrichment* as the per-SNP M_a of all SNPs in a category divided by the per-SNP M_a of all SNPs; analogously, we define *heritability enrichment* as the per-SNP heritability of all SNPs in a category divided by the per-SNP heritability of all SNPs (similar to previous work²²). “All SNPs” refers to common ($MAF \geq 0.05$) and low-frequency ($0.005 \leq MAF < 0.05$) SNPs. Enrichment can be either >1 or <1 (i.e. depletion).

We developed a method, stratified LD fourth moments regression (S-LD4M), to estimate M_a from summary association statistics and LD from a reference panel. S-LD4M regresses squared χ^2 statistics (i.e. fourth powers of signed Z scores) on *LD fourth moments*, defined as sums of r^4 values to each category of SNPs. This approach is analogous to stratified LD score regression (S-LDSC), which regresses χ^2 statistics on LD scores (LD second moments)²². We use a slightly modified version of S-LDSC to estimate heritability enrichment. We run S-LD4M and S-LDSC using the baseline-LD model¹⁷, which includes 75 coding, conserved, regulatory, MAF- and LD-related annotations. When estimating the polygenicity and polygenicity enrichment of a category of SNPs, we restrict to well-powered traits. Further details are provided in the Methods section. We have released open-source software implementing S-LD4M (see URLs).

Simulations

First, we performed simple simulations with no LD to evaluate both the S-LD4M estimator of M_a (the effective number of associated SNPs), and a maximum-likelihood estimator of M_c (the number of causal SNPs) under a point-normal model (PN) (see Methods). This estimator is expected to produce similar results as previous estimators of M_c under a point-normal model^{3,7,9-12}. We simulated 400 large-effect SNPs and 9,600 small-effect SNPs, with $M = 50k$ total SNPs. The mixture of different causal effect sizes, violating the point-normal model, is consistent with evidence for real traits^{10,11}. When there was a large difference between the causal effect sizes of large- and small-effect SNPs, PN produced biased and sample size-dependent estimates of M_c (Figure 2c and Supplementary Table 1), consistent with recent work¹¹. In general, any method that estimates M_c will only detect causal SNPs with effect sizes greater than some threshold, where that threshold depends on modeling assumptions and power. In contrast, S-LD4M does not depend on parametric modeling assumptions, and it produces unbiased estimates of M_a that do not depend on power (Figure 2d). Power-dependent bias would be especially problematic for comparing common vs. low frequency polygenicity, due to lower power for low-frequency SNPs.

Next, we performed a series of simulations using real LD patterns to determine whether S-LD4M produces reliable estimates of polygenicity in realistic settings. We simulated summary association statistics from the asymptotic sampling distribution²³ for UK Biobank imputed SNPs on chromosome 1 ($M \approx 1.0M$ SNPs). We used $N = 50k$ samples and $h^2 = 0.2$, to approximately match Nh^2/M (and hence the expected χ^2 statistic) for UK Biobank traits. We

included MAF-dependent genetic architectures^{10,20} and causal effect size heterogeneity. Details of each simulation are provided in the Methods section.

First, we assessed the ability of S-LD4M to estimate polygenicity for the set of all common and low-frequency SNPs. We determined that median estimates of M_a were approximately unbiased across a wide range of true values, although slight bias was observed for very low values of M_a (Figure 3a and Supplementary Table 1). We report medians instead of means due to noise in the denominator of the estimator, which leads to instability in the mean. In general, S-LD4M is expected to produce approximately unbiased median estimates given sufficient power; the bias that we observed at low values of M_a results from lower power when polygenicity is low (see below for simulations at higher or lower power).

Second, we fixed the polygenicity of common SNPs ($\text{MAF} > 5\%$) at $M_a \approx 1,000$ and progressively reduced the polygenicity of low-frequency SNP ($\text{MAF} = 0.5 - 5\%$). Similar to Figure 3a, median estimates of low-frequency M_a were approximately unbiased except for very low values of M_a (Figure 3b). These estimates indicate that S-LD4M can be used to compare the polygenicity of common and low-frequency SNPs.

Third, we simulated equal heritability enrichment and polygenicity enrichment in four functional categories: coding ($\sim 8x$ enriched), enhancer ($\sim 5x$), DNase I hypersensitivity sites (DHS, $\sim 2x$), and repressed ($\sim 0.75x$), similar to our results on real traits (see below). We determined that S-LD4M produces approximately unbiased estimates of polygenicity enrichment (Figure 3c). We also considered an alternative model of heritability enrichment, under which polygenicity was approximately constant across functional categories. Estimates were approximately unbiased for enriched functional categories but biased for the repressed annotation (Supplementary Figure 1). Despite the fact that this genetic architecture is less realistic (see below), we avoid reporting estimates of polygenicity for depleted functional annotations on real traits.

Fourth, we performed simulations to assess whether genetic architectures with non-random clustering of causal SNPs would bias our estimates of polygenicity enrichment for SNPs in functional categories. Such clustering is expected in real data (e.g. due to biologically important genes), and it could potentially lead to bias, as S-LD4M uses an LD approximation whose accuracy depends on LD between causal SNPs (however, we do not assume that linked SNPs have independent causal effect sizes; see Methods). We simulated clusters of either 5 or 50 causal SNPs (on average) across three genomic length scales (10, 100, 1000 SNPs; roughly 3kb, 30kb, 300kb on average). We determined that median enrichment estimates were approximately unbiased in each case, indicating that our LD approximation is robust to non-random clustering of causal SNPs (Supplementary Figure 2).

Finally, we performed simulations at different sample sizes and included a filtering step (Methods) to exclude trait-annotation pairs with inadequate power. We performed simulations at $N=10k$, $N=50k$ and $N=250k$. For each SNP set (four functional categories + low-frequency SNPs), polygenicity enrichment estimates are reported in Supplementary Figure 3, and proportions of traits retained after the filtering step for each category are reported in Supplementary Table 2. At the default setting of $N=50k$, 61% of simulated traits were retained on average across categories, and polygenicity enrichment estimates were approximately

unbiased. Similar results were obtained at $N=250k$. At $N=10k$, only 8% of simulated traits were retained on average (including ~0% of simulated traits for coding SNPs and for low-frequency SNPs), and polygenicity enrichment estimates were downward biased. Our analyses of real traits more closely correspond to $N=50k$ both in terms of average χ^2 statistic and in terms of the proportion of traits retained (49% on average across categories). Nonetheless, we avoid reporting estimates of polygenicity enrichment for annotations having well-powered enrichment estimates for fewer than 10 out of the 33 traits that we analyzed. We determined that our confidence intervals were approximately well-calibrated or conservative in these simulations (Supplementary Table 2).

Polygenicity of common and low-frequency SNPs across 33 complex traits

We applied S-LD4M to publicly available summary association statistics for 33 diseases and complex traits (average $N = 361k$; Supplementary Table 3), including 29 UK Biobank traits²⁴⁻²⁶ (see URLs) and 4 additional common diseases. LD scores and LD fourth moments were computed using LD estimated from UK10K²⁷ ($M = 8.5$ million SNPs after QC, $MAF > 0.5\%$). As in previous work^{17,22,28}, we excluded the major histocompatibility complex, which has unusually large effect sizes and long-range LD due to balancing selection. (This choice increases M_a estimates for immune-related traits.) Details of our analyses are provided in the Methods section.

For most traits, M_a for common SNPs ($MAF > 5\%$) ranged between 500 and 20,000 ($10^{2.7}$ and $10^{4.3}$; Table 1). Fecundity- and brain-related traits were remarkably polygenic, with M_a estimates greater than 10,000. For Number of children, the most polygenic trait, M_a was almost as large as effective number of independent SNPs, corresponding to an infinitesimal trait ($\log_{10} M_a = 4.52$ (0.13) vs. $\log_{10} M_e = 4.69$). This estimate implies that most common SNPs are associated with this trait (though a much smaller fraction may be causal). Schizophrenia was also extremely polygenic ($\log_{10} M_a = 4.14$ (0.04)), consistent with previous work^{1,4}. Particularly high polygenicity for these traits could result from particularly strong negative selection (see below), although it is also possible that they have greater biological complexity. Red hair pigmentation and sunburn were the least polygenic traits, consistent with known large-effect common SNPs for pigmentation traits^{29,30} (however, we emphasize that the total number of causal SNPs may be much larger than M_a for these traits).

We compared per-SNP M_a estimates for common and low-frequency SNPs across 15 well-powered traits (Figure 4a and Table 1). Polygenicity was 3.9x (95%CI: 2.9 – 5.2x) smaller for low-frequency SNPs than for common SNPs on average, with substantial variation across traits. This difference was similar to the 3.9x (95%CI: 3.5 – 4.4x) smaller per-SNP heritability of low-frequency vs. common SNPs on average (consistent with previous estimates^{10,20}); this concordance was consistent across traits (Figure 4b). This concordance suggests that the upper bound on effect size imposed by negative selection is approximately constant in units of per-SNP heritability, consistent with evolutionary modeling (see below).

For 6 of the 33 traits, summary association statistics from independent cohorts were available. S-LD4M produced concordant M_a estimates for common SNPs on these data sets, despite the smaller sample size and much smaller number of regression SNPs (Supplementary Table 4). (We did not estimate low-frequency M_a for these data sets because summary statistics were not

available for low-frequency SNPs.) M_a provides an upper bound on the proportion of heritability explained by genome-wide significant loci for a trait (Methods). We compared this bound with a direct estimate of this proportion (which may be upwardly biased by winner's curse), and determined that the predicted bound corresponded fairly closely with the estimate (Spearman $r^2 = 0.61$; Supplementary Figure 4a). It also provided a conservative upper bound on the number of genome-wide significant SNPs (Supplementary Figure 4b).

Polygenicity of functional categories across 33 complex traits

We compared estimates of polygenicity enrichment with estimates of heritability enrichment across 25 main functional categories from the baseline-LD model¹⁷ and meta-analyzed results across well-powered traits for 21 categories with at least 10 well-powered traits (Figure 5, Supplementary Table 5 and Supplementary Table 6; 49% of trait-annotation pairs were well powered). For most annotations, polygenicity enrichment was approximately equal to heritability enrichment (regression slope = 0.93; $r^2 = 0.88$). For example, SNPs in conserved regions were 13x enriched for heritability and 14x enriched for polygenicity, and coding SNPs were 9.4x enriched for heritability and 6.6x enriched for polygenicity. (These functional enrichments for the union of common and low-frequency SNPs are larger than the corresponding enrichments for common SNPs, due to larger functional enrichment for low-frequency SNPs¹⁹.) Thus, heritability enrichment in functional categories is predominantly driven by differences in polygenicity, rather than differences in effect-size magnitude. In contrast, *de novo* SNP effect sizes are expected to be much larger in functionally important regions. Thus, genetic signals of important functional regions are constrained by negative selection (Figure 1).

We compared functional enrichment between groups of related traits (8 brain-related, 6 blood-related, and 5 immune-related traits; Supplementary Figure 5). Brain-related traits had smaller functional enrichments both for heritability and for polygenicity, consistent with previous findings^{17,19,22}. Smaller functional enrichment could be explained by stronger negative selection for these traits¹⁹, which may strongly limit the enrichment of any functional category. Stronger negative selection would also be consistent with greater genome-wide polygenicity for these traits (Table 1).

To investigate the relationship between functional enrichment and genome-wide polygenicity, we quantified the sparsity explained by functional annotations, where sparsity is inversely proportional to M_a (Methods). As a proportion of total sparsity, the sparsity explained by functional annotations from the baseline-LD model ranged between 7% and 42% (Supplementary Figure 6a). Traits with smaller M_a (greater sparsity) had greater sparsity explained ($r^2 = 0.78$; Supplementary Figure 6b).

GWAS signals of biologically important genes are constrained by negative selection

Flattening may lead to increased polygenicity not only at the level of SNPs, but also at the level of genes, particularly for genes with large effects on a trait. As a result, GWAS effect sizes may be similar for SNPs near small-effect genes and for SNPs near large-effect genes, and top GWAS SNPs may often implicate small-effect genes (Figure 1).

In order to investigate the impact of flattening on the distribution of heritability across genes, we explored two evolutionary fitness models: a realistic model with SNP-level flattening, and a

model without SNP-level flattening that is realistic in other respects. Under both models, SNPs affect genes and genes affect the trait. Each gene has a trait effect size and a selection coefficient. In the first model (Figure 6a and Supplementary Table 7), there are large-effect genes, and these genes are always strongly constrained. We refer to this model as the “direct selection” model because it could arise from direct selection acting on the trait (but see below). In the second model (Figure 6b and Supplementary Table 7), there are no large-effect genes and different genes have different levels of constraint. We refer to this model as the “pleiotropic selection” model. We did not consider a model with unconstrained large-effect genes, which would lead to a non-polygenic architecture. Further details of the two models are provided in the Methods section.

We analytically inferred the genetic architecture that arises under each model (Methods). Under both models, common variants have smaller per-allele effect sizes than low-frequency variants (Figure 6c-d), concordant with real traits^{10,20}; moreover, both models produced a highly polygenic common-variant architecture (Supplementary Table 7). However, polygenicity differed for the two models at lower allele frequencies (Figure 6e-f). SNP-level flattening (i.e. lower polygenicity at lower allele frequencies) was only observed under the direct selection model; polygenicity was ~30x lower for *de novo* SNPs than for common SNPs (vs. ~1.1x under the pleiotropic selection model). Similar to our results on real traits (Figure 4), low-frequency polygenicity was ~4x smaller, and low-frequency heritability was also ~4x smaller (Figure 6e). The concordance between low-frequency polygenicity enrichment and heritability enrichment is expected. If the selection coefficient of a SNP scales with its squared per-allele effect size, then the effect size bound will be constant, as a function of allele frequency, in units of per-SNP heritability. In units of per-allele effect size, the bound is higher for low-frequency SNPs, so fewer variants reach the bound, leading to lower heritability and proportionally lower polygenicity. However, this concordance is not expected to hold for very rare SNPs, whose effect sizes are limited by biological constraints rather than selection.

For a real trait, the difference in polygenicity between common and *de novo* SNPs could be much larger than the ~30x difference we observed under the direct-selection model, if there are genes with extremely large effect sizes. It could also be slightly smaller; however, the close correspondence between polygenicity and heritability at allele frequencies above 0.5% suggests that the trend (Figure 6e) will not plateau until much smaller allele frequencies. Thus, we expect that there is a $\gg 4x$ difference between the polygenicity of common and low-frequency SNPs for most real traits. We caution that the difference between common vs. low-frequency polygenicity (and heritability) may be a poor proxy for the difference between common vs. *de novo* polygenicity when selection is extremely strong. The effects of flattening may saturate, leading to a small difference in polygenicity between low-frequency and common SNPs (despite the large difference between common and *de novo* SNPs). Such saturation may explain why Number of children, though extremely polygenic, has a small difference in polygenicity between common and low-frequency SNPs (Figure 4).

We analytically computed the heritability explained by a gene (gene-heritability) as a function of its effect size, for SNPs at different allele frequencies. Under the direct selection model, common variant gene-heritability was approximately constant as a function of gene effect size (except for genes with near-zero effect), illustrating that flattening can act at the level of genes (Figure 6g);

the effect was weaker for rare SNPs. Under the pleiotropic selection model (which has no SNP-level flattening), there was also no gene-level flattening (Figure 6h). We computed the proportion of heritability explained by the top 10% of genes (ranked by effect size), at different MAF strata. This proportion was strongly frequency dependent under the direct selection model, but not under the pleiotropic selection model (Figure 6i-j).

These results suggest that large-effect disease genes are always constrained, and that one way that this constraint could arise is direct selection acting on the disease itself. However, some forms of pleiotropic selection may produce similar effects as direct selection; for example, in the case of schizophrenia, pleiotropic selection on neurodevelopment broadly may mimic the effects of direct selection on schizophrenia specifically. Therefore, our results do not imply that direct selection is more important than pleiotropic selection, and do not contradict models of selection that are primarily pleiotropic^{16,18}.

If flattening occurs at the level of genes, then polygenicity should be increased near strongly constrained genes. We estimated the heritability and polygenicity of SNPs within 50kb of 2,990 loss of function-intolerant genes from ExAC (“ExAC genic SNPs”)^{31,32}. These SNPs were more strongly enriched for polygenicity (~2.9x) than for heritability (~1.7x) (Supplementary Table 8). Compared with all genic SNPs (\pm 50kb), ExAC genic SNPs had 1.9x (95% CI: 1.7 – 2.0x) larger polygenicity enrichment but only 1.3x (95%CI: 1.3 – 1.4x) larger heritability enrichment, implying 0.71x (95%CI: 0.66 – 0.76x) smaller average effect sizes (Supplementary Table 8). These estimates suggest that ExAC genes are more likely to be causal but also are more strongly constrained relative to their effect size; they confirm that negative selection at the level of genes affects polygenicity at the level of SNPs.

If the GWAS signal of large-effect genes is constrained, then top GWAS loci should include a mixture of weak perturbations to large-effect, strongly constrained genes (like a “canary in a coal mine”³³) and strong perturbations to small-effect, weakly constrained genes. In particular, common coding variants, representing strong perturbations, may be less likely to implicate large-effect, strongly constrained genes (consistent with increased polygenicity for coding SNPs; Figure 4). We tested this prediction for 37 fine-mapped IBD GWAS loci³⁴, comparing the probability of loss-of-function intolerance (pLI)³¹ between genes harboring coding or noncoding causal risk SNPs. Indeed, 0/8 candidate genes containing fine-mapped coding variants had high pLI (\geq 0.9), compared to 12/29 candidate genes near fine-mapped noncoding variants (rank-sum test $p = 0.006$ for difference; Supplementary Figure 7 and Supplementary Table 9).

Discussion

Using a new definition of polygenicity and a new method to estimate it, we compared the polygenicity of 33 complex traits across allele frequencies and functional categories. We determined that low-frequency variants have lower polygenicity than common variants and that biologically important functional categories have higher polygenicity in proportion to their higher heritability. These results demonstrate that negative selection not only constrains common-variant effect sizes but flattens their distribution across the genome, explaining the high polygenicity of complex traits. Furthermore, negative selection constrains the genetic signals of biologically important regions and genes.

A recent study¹⁵ proposed an “omnigenic model” of complex traits: due to densely connected cellular networks, nearly every gene expressed in a relevant cell type contributes to heritability. As a result, most heritability is explained by “peripheral genes” rather than biologically important “core genes”. On the one hand, our results suggest that the omnigenic model is incomplete, and that negative selection can help explain why core genes explain a limited proportion of trait heritability. On the other hand, our results support a potential distinction between core and peripheral genes (which has been debated^{15,35}): among a large set of genes with similarly strong common-variant associations, a small subset may have much larger phenotypic effects, and it may be useful to label these as core genes.

Polygenicity has not precluded GWAS from producing biological insights, and our results provide guidelines to accelerate their progress. First, GWAS follow-up studies should prioritize genes with evidence of constraint, since all large-effect genes are constrained. Loss-of-function variants provide a useful metric of constraint³¹; for GWAS, a maximally informative constraint metric might incorporate noncoding variation³⁶, gene expression data³⁷, and functional predictions³⁸. Second, counter-intuitively, follow-up studies should prioritize associations that do not map to coding regions or large-effect regulatory elements, since SNPs with strongly deleterious effects on their target gene are less likely to implicate strongly pathogenic genes. As our ability to interrogate the gene-regulatory effects of GWAS SNPs improves³⁸⁻⁴⁰, a potential pitfall would be to prioritize GWAS SNPs with the largest regulatory effects. Third, rare-variant based evidence from exome sequencing studies, even if underpowered, can be used to prioritize GWAS genes. Indeed, exome sequencing studies⁴¹⁻⁴³ have been viewed as an attractive complement to GWAS¹⁵, and our results support this perspective. However, moderately rare coding variants may suffer the same limitations as common regulatory variants.

This study has several limitations. First, we have not quantified the polygenicity of rare or *de novo* variants. An exome-sequencing study of schizophrenia⁴² showed that rare-variant signals are sufficiently polygenic that there was greater power to detect significant pathways than to detect individual genes. Our comparison of common vs. low-frequency variant polygenicity provides a conservative lower bound ($\sim 4x$) on the effect of flattening, and our evolutionary modeling suggests a much stronger overall effect ($\sim 30x$), although we caution that this number may be slightly smaller or much larger, depending on the distribution of gene effect sizes. Second, our definition of M_a is dependent on the amount of LD between causal SNPs, which has both advantages and disadvantages. On the one hand, it makes M_a identifiable; on the other hand, M_a would be easier to interpret if it only depended on the causal effect size distribution. Third, the nonparametric approach that we used to define and estimate polygenicity may not be optimal for every application. A recent study¹¹ fit a parametric model involving a mixture of normal distributions; this approach may provide more accurate estimates of missing heritability as a function of sample size, and it may more accurately predict the performance of risk prediction methods that make similar parametric assumptions⁴⁴. Fourth, S-LD4M can produce biased estimates for depleted functional annotations, albeit in unrealistic settings (Supplementary Figure 1). However, this bias does not affect enriched annotations or low-frequency variants, which are not in strong LD with common variants, and we have avoided reporting polygenicity estimates for depleted functional annotations. Fifth, S-LD4M produces noisy estimates for some annotations and traits, making it necessary to perform meta-analyses across well-powered traits, which may not be representative of all traits. Comparisons of heritability enrichment and

polygenicity enrichment are not biased by this filtering process, because the same traits are used to estimate both types of enrichment. Sixth, S-LD4M can potentially be biased due to population stratification, and it should only be applied to data sets where stratification is well controlled. In particular, S-LD4M assumes that any stratification leads to uniform inflation of χ^2 statistics; this assumption could be violated at loci under positive selection, such as at the LCT locus for height⁴⁵. However, we have only applied S-LD4M to datasets that were corrected for population stratification (including UK Biobank, a relatively homogenous study). Seventh, although our evolutionary modeling supports the hypothesis that flattening affects the distribution of heritability across genes, an alternative explanation is that SNP-level flattening results from increased allelic heterogeneity for common variants near large-effect genes. However, this explanation would require that there be many independent associations per gene. Despite evidence of allelic heterogeneity, it is not a common phenomenon that multiple independent, similarly strong associations implicate the same gene⁴⁶. Eighth, inferences about components of heritability can potentially be biased by failure to account for LD-dependent architectures^{17,47-49}. All of our analyses used the baseline-LD model, which includes 6 LD-related annotations¹⁷. The baseline-LD model is supported by formal model comparisons using likelihood and polygenic prediction methods, as well as analyses using a combined model incorporating alternative approaches^{50,51}; however, there can be no guarantee that the baseline-LD model perfectly captures LD-dependent architectures. Despite these limitations, this study advances our understanding of genetic architecture and the evolutionary processes that shape it.

URLs

Open-source software implementing our method is available at:

<https://github.com/lukejoconnor/SLD4M>. UK Biobank summary statistics are available at: <https://data.broadinstitute.org/alkesgroup/UKBB/>.

Acknowledgements

We thank Dr. Benjamin Neale, Dr. Hilary Finucane, Dr. Omer Weissbrod and Dr. Yakir Reshef for helpful comments. This research was funded by NIH grants U01 HG009379, R01 MH101244, R01 MH107649, U01 HG009088 and R01 MH109978.

- 1 Consortium, I. S. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-752, doi:10.1038/nature08185 (2009).
- 2 Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-569, doi:10.1038/ng.608 (2010).
- 3 Stahl, E. A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* **44**, 483-489, doi:10.1038/ng.2232 (2012).
- 4 Loh, P. R. *et al.* Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat Genet* **47**, 1385-1392, doi:10.1038/ng.3431 (2015).
- 5 Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* **47**, 1114-1120, doi:10.1038/ng.3390 (2015).
- 6 Moser, G. *et al.* Simultaneous discovery, estimation and prediction analysis of complex traits using a bayesian mixture model. *PLoS Genet* **11**, e1004969, doi:10.1371/journal.pgen.1004969 (2015).

- 7 Palla, L. & Dudbridge, F. A Fast Method that Uses Polygenic Scores to Estimate the Variance Explained by Genome-wide Marker Panels and the Proportion of Variants Affecting a Trait. *Am J Hum Genet* **97**, 250-259, doi:10.1016/j.ajhg.2015.06.005 (2015).
- 8 Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5-22, doi:10.1016/j.ajhg.2017.06.005 (2017).
- 9 Yang, J., Fritsche, L. G., Zhou, X., Abecasis, G. & International Age-Related Macular Degeneration Genomics, C. A Scalable Bayesian Method for Integrating Functional Information in Genome-wide Association Studies. *Am J Hum Genet* **101**, 404-416, doi:10.1016/j.ajhg.2017.08.002 (2017).
- 10 Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nat Genet* **50**, 746-753, doi:10.1038/s41588-018-0101-4 (2018).
- 11 Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits *Nature genetics*, doi:<https://doi.org/10.1038/s41588-018-0193-x> (2018).
- 12 Zhu, X. & Stephens, M. A large-scale genome-wide enrichment analysis identifies new trait-associated genes, pathways and tissues across 31 human phenotypes. *bioRxiv*, 160770 (2017).
- 13 Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753, doi:10.1038/nature08494 (2009).
- 14 Hirschhorn, J. N. Genomewide association studies--illuminating biologic pathways. *N Engl J Med* **360**, 1699-1701, doi:10.1056/NEJMp0808934 (2009).
- 15 Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177-1186, doi:10.1016/j.cell.2017.05.038 (2017).
- 16 Eyre-Walker, A. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences*, 200906182 (2010).
- 17 Gazal, S. *et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nature genetics* **49**, 1421 (2017).
- 18 Simons, Y. B., Bullaughey, K., Hudson, R. R. & Sella, G. A population genetic interpretation of GWAS findings for human quantitative traits. *PLoS biology* **16**, e2002985 (2018).
- 19 Gazal, S. *et al.* Low-frequency variant functional architectures reveal strength of negative selection across coding and non-coding annotations. *bioRxiv; in press, Nature Genetics*, 297572 (2018).
- 20 Schoech, A. *et al.* Quantification of frequency-dependent genetic architectures and action of negative selection in 25 UK Biobank traits. *bioRxiv*, 188086 (2017).
- 21 Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS genetics* **2**, e190 (2006).
- 22 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* **47**, 1228 (2015).
- 23 Conneely, K. N. & Boehnke, M. So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. *The American Journal of Human Genetics* **81**, 1158-1168 (2007).

- 24 Bycroft, C. *et al.* Genome-wide genetic data on ~ 500,000 UK Biobank participants. *BioRxiv*, 166298 (2017).
- 25 Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nature genetics*, 1 (2018).
- 26 Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* **12**, e1001779 (2015).
- 27 consortium, U. K. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82 (2015).
- 28 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291 (2015).
- 29 Han, J. *et al.* A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS genetics* **4**, e1000074 (2008).
- 30 Sulem, P. *et al.* Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature genetics* **39**, 1443 (2007).
- 31 Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285 (2016).
- 32 Hormozdiari, F. *et al.* Leveraging molecular QTL to understand the genetic architecture of diseases and complex traits. *bioRxiv*, 203380 (2017).
- 33 Hunter, D. J., Altshuler, D. & Rader, D. J. From Darwin's finches to canaries in the coal mine—mining the genome for new biology. *New England Journal of Medicine* **358**, 2760-2763 (2008).
- 34 Huang, H. *et al.* Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* **547**, 173 (2017).
- 35 Wray, N. R., Wijmenga, C., Sullivan, P. F., Yang, J. & Visscher, P. M. Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. *Cell* **173**, 1573-1580 (2018).
- 36 Di Iulio, J. *et al.* The human noncoding genome defined by genetic diversity. *Nature genetics* **50**, 333 (2018).
- 37 Consortium, G. Genetic effects on gene expression across human tissues. *Nature* **550**, 204 (2017).
- 38 Zhou, J. *et al.* Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. *Nature genetics*, 1 (2018).
- 39 Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nature genetics* **48**, 245 (2016).
- 40 Ernst, J. *et al.* Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. *Nature biotechnology* **34**, 1180 (2016).
- 41 Zuk, O. *et al.* Searching for missing heritability: designing rare variant association studies. *Proceedings of the National Academy of Sciences* **111**, E455-E464 (2014).
- 42 Purcell, S. M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185 (2014).
- 43 Do, R. *et al.* Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature* **518**, 102 (2015).

- 44 Vilhjálmsón, B. J. *et al.* Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics* **97**, 576-592 (2015).
- 45 Campbell, C. D. *et al.* Demonstrating stratification in a European American population. *Nature genetics* **37**, 868 (2005).
- 46 Hormozdiari, F. *et al.* Widespread allelic heterogeneity in complex traits. *The American Journal of Human Genetics* **100**, 789-802 (2017).
- 47 Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. *The American Journal of Human Genetics* **91**, 1011-1021 (2012).
- 48 Speed, D. *et al.* Reevaluation of SNP heritability in complex human traits. *Nature genetics* **49**, 986 (2017).
- 49 Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature genetics* **47**, 1114 (2015).
- 50 Gazal, S., Finucane, H. K. & Price, A. L. Reconciling S-LDSC and LDAK functional enrichment estimates. *bioRxiv*, 256412 (2018).
- 51 Price, A. L., Gazal, S., & Finucane, H. K. Reconciling S-LDSC and LDAK models and functional enrichment estimates. Abstract to be presented at the 68th annual meeting of the American Society of Human Genetics (ASHG 2018), October 17, 2018, San Diego, CA.
- 52 O'Connor, L. J. & Price, A. L. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *bioRxiv*, 205435 (2018).

Methods

The effective number of associated SNPs

Here, we provide three definitions of the effective number of associated SNPs (M_a). First, we define M_a in terms of the mixed fourth moments of the distribution of causal and marginal effect sizes. This generalizes the definition in Equation (1) (see Overview of methods), which corresponds to the special case of no LD between causal SNPs. Second, we provide an alternative definition (under a random effects model) that formalizes the illustration in Figure 2a, involving the average unit of heritability explained by a causal SNP (see Overview of methods). Third, we provide a concise definition that provides less intuition. We show that these definitions are equivalent in the Supplementary Note.

First, let $\boldsymbol{\beta}$ denote the random vector of causal effect sizes for common and low-frequency SNPs. Let R be the fixed LD matrix, and let $\boldsymbol{\alpha} = R\boldsymbol{\beta}$ denote the random vector of marginal effect sizes. We use the notation α, β to denote randomly chosen entries of $\boldsymbol{\alpha}, \boldsymbol{\beta}$. The heritability is $h^2 = E(\boldsymbol{\beta}^T R \boldsymbol{\beta}) = ME(\alpha\beta)$. M_a is defined as:

$$M_a = \frac{3M}{\kappa_a}, \quad \kappa_a = \frac{3E(\alpha^2\beta^2) - 2E(\beta^4)}{E(\alpha\beta)^2}. \quad (2)$$

Another possible definition of polygenicity is the effective number of *causal* SNPs, denoted M_b :

$$M_b = \frac{3M}{\kappa}, \quad \kappa = \frac{E(\beta^4)}{E(\beta^2)^2}. \quad (3)$$

However, this definition is unidentifiable without making a strong assumption (roughly, that linked SNPs have independent causal effect sizes; see below). This definition is equivalent to equation (1) in the Overview of methods section, which defined M_a (rather than M_b) in the special case of no LD between causal SNPs; in that case, $M_a = M_b$ (because $\alpha = \beta$ whenever $\beta \neq 0$). M_b is different from M_c , the number of causal SNPs, except when causal effect sizes follow a point-normal distribution. Because it is implausible that causal effect sizes follow a normal distribution, we view the distinction between M_a and M_c as having greater importance than the distinction between M_a and M_b .

Second, we consider a non-i.i.d. normal model:

$$\boldsymbol{\beta} \sim N(0, \Sigma) \quad (4)$$

where Σ is a fixed diagonal matrix with diagonal entries $\sigma_1^2, \dots, \sigma_M^2$. This flexible model generalizes the point-normal model. Because Σ is diagonal, h^2 is equal to $\text{Tr}(\Sigma)$. Note that if $\Sigma \propto I$, then κ is equal to 3 and M_b is equal to M . Above, we used the notation $E(\cdot)$ to denote a uniform average across SNPs. We use the notation $E_{h^2}(\cdot)$ to denote an average across components of heritability, i.e. a weighted average where the probability of choosing SNP i is equal to σ_i^2/h^2 . Using this notation, we refer to $E_{h^2}(\alpha^2)$ as the *average unit of heritability*, and M_a is equal to:

$$M_a = \frac{h^2}{E_{h^2}(\alpha^2)}. \quad (5)$$

M_b can be defined in a similar manner:

$$M_b = \frac{h^2}{E_{h^2}(\beta^2)}. \quad (6)$$

Third, M_a can also be defined without specifying that Σ is a diagonal matrix. (The assumption that Σ is diagonal is similar to the common assumption that $E(\beta^T R \beta) = E(\beta^T \beta)^{2,22}$.) Let $S = \Sigma^{-1/2} R \Sigma^{-1/2}$. In the diagonal case, S is a weighted LD matrix, where the rows and columns are weighted by the expected effect size of each SNP. The heritability is equal to $\text{Tr}(S)$ (in the diagonal case, $\text{Tr}(S) = \text{Tr}(\Sigma)$). M_a is equal to:

$$M_a = \frac{\text{Tr}(S)^2}{\text{Tr}(S^2)}. \quad (7)$$

This definition, though concise and natural, provides little intuition.

Properties of M_a

M_a has notable mathematical properties.

- **Identifiability.** If there are two SNPs in perfect LD, there is no way to tell based on GWAS data whether one or both of them are causal. If a definition of polygenicity differs depending on whether one or both are causal, then it is unidentifiable. Indeed, M_b (as well as M_c) is larger if both SNPs are causal than if only one SNP is causal. However, because both SNPs have identical values of $E(\alpha^2)$, the value of $E_{h^2}(\alpha^2)$ (and therefore M_a) is unaffected.
- **Heritability explained by genome-wide significant SNPs.** M_a gives an upper bound on the proportion of heritability explained by SNPs whose effect sizes exceed a specified threshold T , such as the genome-wide significance threshold. This proportion can be denoted $P_{h^2}(\alpha^2 > T)$. Because α^2 is nonnegative,

$$P_{h^2}(\alpha^2 > T) \leq \frac{E_{h^2}(\alpha^2)}{T}. \quad (8)$$

This bound is relatively tight when $T \gg E_{h^2}(\alpha^2)$. When T is a significance threshold (for example, $T = \frac{30}{N}$ corresponds to genome-wide significance), this bound is most relevant at low sample size.

- **Polygenic prediction accuracy.** Intuitively, increased polygenicity makes prediction more difficult. If Σ is given, there is a simple expression for the optimal risk prediction accuracy; this expression provides an upper bound on prediction accuracy in the case that Σ is not given:

$$E(r^2) = h^2 E_{h^2} \left(\frac{\alpha^2}{\alpha^2 + \frac{1}{N}} \right). \quad (9)$$

(See Supplementary Note for derivation). At large N , prediction accuracy converges to h^2 . At small N , it is approximately a linear function of sample size with slope inversely proportional to M_a :

$$E(r^2) \approx N h^2 E_{h^2}(\alpha^2). \quad (10)$$

In practice, polygenic prediction is usually performed using large datasets for which this approximation is not appropriate.

- **Effective number of independent SNPs.** Under an infinitesimal model, where every SNP is causal with a normally distributed causal effect size, M_a is equal to the effective number of independent SNPs²¹ (M_e). M_e is defined as the number of SNPs divided by the average LD score, or in notation similar to equation (7), as

$$M_e = \frac{\text{Tr}(R)^2}{\text{Tr}(R^2)}. \quad (11)$$

We note that M_a might be close to M_e even when M_b is much smaller than M . For example, if the genome comprises perfect LD blocks of 100 SNPs each, then M_a can be equal to M_e even if only 1% of SNPs (1 per LD block) is causal. We also note that the value of M_e is strongly dependent on allele frequency, as rare SNPs (which have less LD) contribute strongly to M_e . In contrast, M_a will not diverge if many rare SNPs are included, as these SNPs explain little heritability.

- **Symmetry.** M_a is symmetric with respect to the two fixed parameters in the random effects model, R and Σ : $M_a(R, \Sigma) = M_a(\Sigma, R)$. This property is shared by the LD-dependent definition of heritability, $E(\beta^T R \beta) = \text{Tr}(S)$. It is not shared by M_e or by the LD-independent definition of heritability, $E(\beta^T \beta)$.

Regression equation used in stratified LD fourth moments regression

S-LD4M is justified by an approximate regression equation, which states that the expected value of α^4 for SNP i is approximately proportional to the LD fourth moment of SNP i , with a proportionality constant that can be used to estimate M_a .

Let $\ell^{(2)}, \ell^{(4)}$ denote the LD second moment (LD score²⁸) and LD fourth moment, respectively, for a randomly chosen SNP:

$$\ell_i^{(p)} = \sum_j r_{ij}^p \quad (12)$$

The regression equation is:

$$E(\alpha^4 | \ell^{(2)}, \ell^{(4)}) \approx 3E(\alpha^2 | \ell^{(2)}, \ell^{(4)})^2 + \ell^{(4)} K. \quad (13)$$

In the first term, $E(\alpha^2 | \ell^{(2)}, \ell^{(4)}) = \ell^{(2)} \tau$, where the coefficient τ is the variance of causal effect sizes^{22,28}. In the second term, K is related to M_a :

$$K = E_{h^2}(\alpha^2) - 3E(\alpha^2)E(\beta^2). \quad (14)$$

Note that there are three kinds of expectations. First, $E(\alpha^2 | \ell^{(2)}, \ell^{(4)})$ and $E(\alpha^4 | \ell^{(2)}, \ell^{(4)})$ are conditioned on LD. Second, $E(\alpha^2)$ and $E(\beta^2)$ are not conditioned on LD; rather, they represent unweighted averages over all reference SNPs. Third, $E_{h^2}(\alpha^2)$ is the average across components of heritability (not uniformly across SNPs).

In the case that there are P functional annotations, $\ell_i^{(4)}$ and K are vectors of size $1 \times P$ and $P \times 1$ respectively. Similar to S-LDSC, we make an implicit additivity assumption for the fourth moments of SNPs in the intersection of annotations. (Our simulations violate this assumption, but it does not appear to result in bias).

This regression equation relies on an LD approximation. Roughly, the LD approximation states that if SNP i is in LD with causal SNP j , then the expected marginal effect size of SNP i is approximately proportional to the marginal effect size of j . This approximation is exact in important special cases, suggesting that it will be robust in practice. First, we consider the stronger approximation that would be needed to estimate M_b . We would need to assume, roughly, that linked SNPs have independent causal effect sizes. More precisely, we would need

to assume that the contribution of causal SNP j to the marginal effect size of SNP i , not conditional on other SNPs, is proportional to the causal effect size of SNP j :

$$\text{cov}(\alpha_i^2, \beta_j^2) \approx r_{ij}^2 \text{var}(\beta_j^2). \quad (15)$$

We weaken assumption (15) to obtain our LD approximation:

$$\text{cov}(\alpha_i^2, \beta_j^2) \approx r_{ij}^2 \text{cov}(\alpha_j^2, \beta_j^2). \quad (16)$$

Now, we are assuming that the contribution of causal SNP j to the marginal effect size of SNP i is proportional to the *marginal* effect size of SNP j . Note that (15) implies (16) by applying (15) to the case $i = j$. We show that (16) implies (13) in the Supplementary Note. (16) is not violated when other causal SNPs k are in strong LD with SNP j ($r_{jk}^2 \approx 1$), since they would contribute to α_i^2 in proportion to their contribution to α_j^2 (because $r_{ik}^2 \approx r_{ij}^2 r_{jk}^2$). However, it could be violated if there were other causal SNPs in weak LD with SNPs i and j (violating $r_{ik}^2 \approx r_{ij}^2 r_{jk}^2$).

Equations (15-16) have second-moment analogues that can be used to justify LD score regression^{22,28}. LDSC has been justified by assuming that correlated SNPs have uncorrelated effect sizes; this assumption can be stated as:

$$E(\alpha_i \beta_j) = r_{ij} E(\beta_j^2), \quad (17)$$

which is analogous to (15). This assumption allows LDSC to estimate a non-LD-dependent definition of heritability, $ME(\beta^2)$. However, a weaker assumption is also possible, analogous to (16):

$$E(\alpha_i \beta_j) \approx r_{ij} E(\alpha_j \beta_j). \quad (18)$$

This weaker assumption still holds when perfectly correlated SNPs have correlated causal effect sizes. It allows LDSC to estimate $E(\alpha\beta)$, which is proportional to $\beta^T R \beta$, as follows:

$$E(\alpha_i^2) = \sum_j r_{ij} E(\alpha_i \beta_j) \approx \sum_j r_{ij}^2 E(\alpha_j \beta_j) = \ell_i^{(2)} E(\alpha\beta). \quad (19)$$

Thus, although LDSC has previously been justified using the assumption that correlated SNPs have uncorrelated effect sizes, only this weaker assumption is strictly necessary in order to estimate an LD-dependent definition of heritability.

Stratified LD fourth moments regression

In practice we do not observe α , but rather a noisy estimate $\hat{\alpha}$. We can correct for sampling noise using:

$$E(\hat{\alpha}^2 | \alpha) = \alpha^2 + \frac{1}{N}, E(\hat{\alpha}^4 | \alpha) = \alpha^4 + 6\alpha^2/N + 3/N^2,$$

where in practice we use the LD score regression intercept divided by N , denoted $1/\hat{N}$, instead of $1/N$. Let $\widehat{\alpha}^2 = \hat{\alpha}^2 - 1/\hat{N}$ and $\widehat{\alpha}^4 = \hat{\alpha}^4 - 6\widehat{\alpha}^2/\hat{N} - 3/\hat{N}^2$.

We perform a two-step inference procedure. First, we use a slightly modified version of S-LDSC²² (see below) to estimate $E(\alpha_i^2)$ and (β_i^2) for each SNP, conditional on their respective LD scores and annotation values. Second, we regress the $M \times 1$ vector $\widehat{\alpha}^4 - 3E(\alpha^2)^2$ on the $M \times P$ matrix $\mathbf{L}^{(4)}$ to obtain an estimate of the $1 \times P$ vector K . The regression is weighted: the weight of SNP i is 1 divided by the LD fourth moment of SNP i (to all common and low-frequency SNPs). For each annotation A_p , we subtract $3E(\alpha^2)E(\beta^2)|A_p$ from K_p to obtain

an estimate of L , which we combine with the S-LDSC heritability estimate for each category to estimate M_a for each category.

Our modified version of S-LDSC uses a slightly modified weighting scheme and does not exclude large-effect SNPs. The regression weight of each SNP is 1 divided by the LD score for that SNP to all common and low-frequency SNPs; this choice is similar to the original version of S-LDSC²², but slightly modified for consistency with the weights used in S-LD4M. (We do not exclude large-effect SNPs because these SNPs are important for estimating fourth moments; their exclusion would lead to upwardly biased estimates of M_a .)

Simulations

We simulated summary statistics directly using the asymptotic sampling distribution²³, rather than by simulating individual-level data. Specifically, the distribution of the summary statistic vector, $\hat{\alpha}$, as a function of the true causal effect size vector β , was:

$$\hat{\alpha} \sim N\left(\hat{R}\beta, \frac{1}{N}\hat{R}\right).$$

The estimated LD matrix \hat{R} was equal to I in simulations with no LD. In simulations with LD, \hat{R} was computed from UK Biobank data (N=460k). Sample correlations were computed for all typed and imputed SNPs (M=1.0M) within 0.1cM of each other on chromosome 1. Blocks of 5,000 SNPs were used, and R was set to zero outside of the blocks. These choices were necessary for \hat{R} to be stored in memory, due to the large number of SNPs. To ensure that \hat{R} was positive semidefinite, negative eigenvalues were discarded, as previously described⁵².

Simulations with no LD. We simulated 50,000 total SNPs, 9600 small-effect SNPs, and 400 large-effect SNPs. For small- and large-effect SNPs, effect sizes were drawn from a normal distribution with mean zero and variance σ_1^2 and σ_2^2 respectively; in the three simulations (corresponding to the three example architectures in Figure 2a), σ_1^2 was equal to 1/10,000, 1/1,200, and 1/600 respectively. σ_2^2 was equal to $\frac{1-400\sigma_1^2}{9600}$, resulting in a heritability of 1. (We note that in all simulations, only Nh^2 affects the results; for example, identical results are obtained at $h^2 = 1$ and $N = 20k$ and at $h^2 = 0.2$ and $N = 100k$.) We performed simulations at $N = 5,000, 25,000, 125,000$ and $625,000$.

Simulations with LD. We two sets of simulations with real LD (see above). First, we performed simulations with no functional annotations. We specified that a certain percentage of common SNPs (MAF>0.05) and a certain percentage of low-frequency SNPs (0.05>MAF>0.005) were causal. Conditional on being causal, effect sizes were drawn from a mixture of two normal distributions, with probability $p_1 = 0.1$ and $p_2 = 0.9$, and variance $\sigma_1^2 = 4\sigma_2^2$. The variance parameter was frequency-dependent; it was proportional to $[p(1-p)]^{0.25}$, where p is the allele frequency.

Second, we performed simulations with 5 real functional annotations (coding, enhancer, promoter, DHS, repressed). We specified a different probability of being causal for SNPs in each category: 1/2, 1/4, 3/20, 3/80, and 1/20, for SNPs in each annotation and SNPs in no annotation at all, respectively. When a SNP was in multiple annotations, the probability of being causal was the maximum of the respective values. Conditional on being causal, SNPs had independent and

identically distributed effect sizes; their effects were drawn from a mixture of two normal distributions, with probability $p_1 = 0.1$ and $p_2 = 0.9$, and variance $\sigma_1^2 = 4\sigma_2^2$. The variances were scaled so that the total (expected) heritability was 0.2.

Evolutionary modeling

In our evolutionary modeling, each SNP affected one gene and each gene affected both the trait and fitness. We specified an effect size distribution of effect sizes for SNPs on genes (identical for every gene) and a joint distribution for the effect size and selection coefficient of each gene. The effect of a SNP on a trait was its effect on its gene times the gene effect size; the selection coefficient of a SNP was its effect on its gene times the gene selection coefficient. Using an analytical formula for the distribution of allele frequencies conditional on selection coefficients, we obtained the joint distribution of trait effects and allele frequencies for each gene. In detail:

Under each model, we specify a joint distribution of gene effect sizes, denoted β_{gene} , and selection coefficients, denoted s_{gene} . For the direct selection model, s_{gene} was equal to β_{gene}^2 , and β_{gene} followed a mixture of normal distributions; 95% of genes had variance 10^{-3} , and 5% of genes had variance 10^{-1} . As a result of this choice, the *de novo* effect size distribution is dominated by large-effect genes. For the pleiotropic selection model, s_{gene} was independent of β_{gene} ; β_{gene} followed a normal distribution with variance 10^{-3} (so that there were no large-effect genes), and s_{gene} followed a gamma distribution with parameters $k = 5/2$ and $\theta = 1/1,250$.

Under each model, we specify a distribution of effect sizes for SNPs on each gene, denoted $\beta_{SNP \rightarrow gene}$. Each SNP affected one gene, and the distribution was identical for every gene. For the direct selection model, we specified an inverse-gamma distribution with parameters $k = 100$ and $\theta = 1$. This is a heavy-tailed distribution; it causes the lines in Figure 6g to plateau, rather than decreasing after attaining a maximum. For the pleiotropic selection model, we specified a mixture of normal distributions; 75% of SNPs had variance 0.2, and 25% of SNPs had variance 0.02. This choice causes the effect size distribution to be highly polygenic, despite the lack of flattening in this model; it also leads to an appropriate relationship between allele frequency and per-SNP heritability.

The effect size of a SNP on the trait, denoted β_{SNP} , was equal to $\beta_{SNP \rightarrow gene}\beta_{gene}$. The selection coefficient of a SNP, denoted s_{SNP} , was equal to $\beta_{SNP \rightarrow gene}s_{gene}$. For each gene, we computed the joint distribution of β_{SNP} and allele frequency p , using the formula for the probability density of p conditional on s_{SNP} :

$$f(p|s_{SNP}) = \frac{1}{Z} p^{4N\mu-1} (1 - p^{4N\mu-1}) \exp(-4Ns_{SNP}p),$$

where Z is a constant, $N = 100$, and $\mu = 1/200$ (so that the exponent $4N\mu - 1$ is equal to 1). We note that the model is overparameterized; for example, identical results would be obtained with 100x larger N and 100x smaller s_{gene} and μ .

We simulated 20,000 genes. For each gene, we computed the second and fourth moments of β_{SNP} for SNPs at allele frequencies $p = 0.25, 0.05, 0.01, 0.002, 0.004$, by multiplying the probability density function of β_{SNP} by $f(p|s_{SNP})$ and approximating the integral using a sum

over $\beta_{SNP \rightarrow gene}^2 = \frac{1}{5000}, \frac{2}{5000}, \dots, 8$. Averaging over genes, we computed the variance in per-allele effect sizes (Figure c-d), the per-SNP heritability (variance times heterozygosity; Figure 6e-f), and the polygenicity (inverse of kurtosis; Figure 6 e-f). We also computed the proportion of heritability explained by the 10% of genes with largest β_{gene}^2 (Figure 6i-j).

We also computed the heritability explained by a gene as a function of β_{gene}^2 (Figure 6g-h). In Figure 6g, s_{gene} was equal to β_{gene}^2 ; in Figure 6h, s_{gene} was fixed at 2×10^{-4} . The heritability explained by the gene at a given allele frequency was equal to the variance of β_{SNP}^2 at that allele frequency, times the heterozygosity.

Tables

Trait	$\log_{10}M_a$ common	$\log_{10}M_a$ LF
Infinitesimal trait	4.69	4.98
Number of children	4.50(0.13)	3.95(0.23)
Schizophrenia	4.14(0.04)	3.10(0.16)
Smoking status	4.11(0.02)	NA
College	4.08(0.05)	2.86(0.20)
Morning person	4.02(0.02)	2.16(0.25)
Age at first birth – F	3.99(0.04)	NA
Neuroticism	3.99(0.23)	NA
FVC	3.87(0.16)	NA
CVD/HT	3.81(0.03)	NA
BP – systolic	3.81(0.02)	2.31(0.42)
BMI	3.78(0.12)	2.53(0.21)
IBD	3.60(0.04)	3.19(0.10)
Height	3.56(0.02)	2.91(0.05)
WHR	3.50(0.04)	2.54(0.09)
Age at menarche	3.47(0.09)	2.47(0.26)
FEV1/FVC	3.38(0.05)	3.01(0.16)
WBC count	3.30(0.10)	2.42(0.18)
Eczema	3.20(0.05)	NA
Asthma	3.12(0.03)	NA
Eosinophil count	3.05(0.04)	NA
RA	3.03(0.10)	NA
Platelet count	2.97(0.04)	2.51(0.10)
AID	2.93(0.05)	NA
BMD – heel	2.90(0.08)	NA
Alzheimer’s	2.90(0.11)	NA
Type II diabetes	2.85(0.14)	NA
RBC distribution width	2.70(0.05)	NA
RBC count	2.67(0.18)	NA
Platelet distribution width	2.64(0.08)	1.36(0.54)
Balding	2.59(0.18)	2.16(0.20)
Age at menopause	2.55(0.06)	NA
Sunburn	1.99(0.08)	NA
Red hair	1.02(0.22)	NA

Table 1. Estimates of polygenicity for common and low-frequency SNPs across 33 complex traits. We report common variant estimates for all traits, and low-frequency estimates for well-powered traits (see Methods). The first row reports the effective number of independent SNPs²¹. M_a is close to this value when marginal effect sizes approximately follow a Normal distribution, which does not imply that every SNP is causal.

Figures

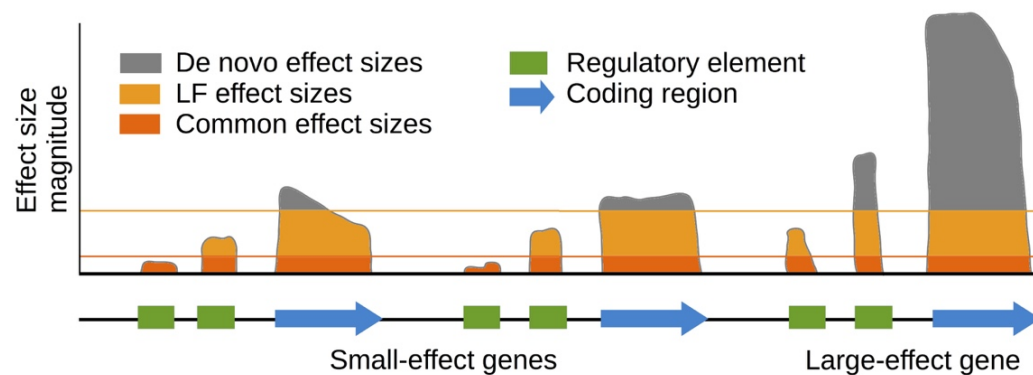


Figure 1. Illustration of flattening due to natural selection. We display the maximum per-allele effect size of a SNP at each site for a toy example of three genes and nearby regulatory regions. Here, the distribution of *de novo* effects is not highly polygenic; it is dominated by coding variants in a single large-effect gene (although other genes also harbor small effects). Negative selection imposes an upper effect size bound (possibly soft), resulting in increased polygenicity. The effect is stronger for common SNPs than for low-frequency SNPs. Within functionally important regions (e.g. coding), a larger proportion of variants have effect sizes near the bound, leading to especially large polygenicity. In practice, this bound may vary across the genome; we expect that the resulting distribution will be more flat than the distribution of *de novo* variants, but not perfectly flat.

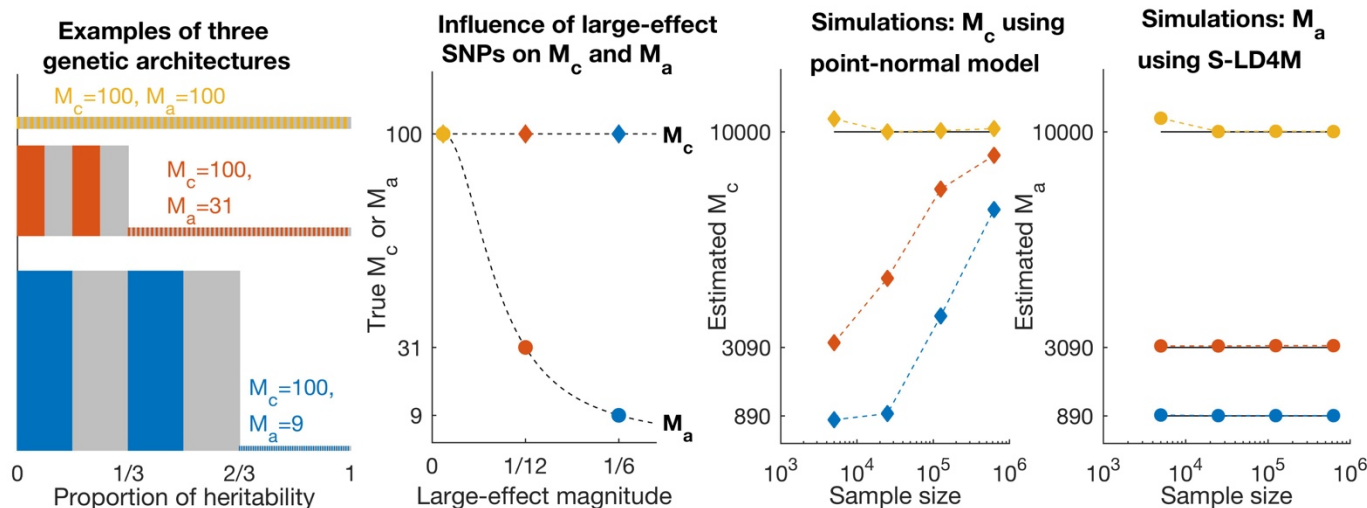


Figure 2. Comparison of the effective number of associated SNPs (M_a) with the number of causal SNPs (M_c). (a),(b) Examples of three genetic architectures with $M_c = 100$. (a) Each colored or gray block corresponds to one SNP; both height and width are proportional to the SNP's effect size variance. The average unit of heritability, denoted $E_{h^2}(\alpha^2)$, is the average height (equal to the total area) of the colored and gray regions. M_a is equal to $h^2/E_{h^2}(\alpha^2)$. (b) M_c and M_a as a function of the effect size magnitude of the 4 large-effect SNPs. (c),(d) Simulations of the same three genetic architectures with the number of SNPs (and causal SNPs) scaled up by 100x. (c) Estimates of M_c under a point-normal model, at different sample sizes. (d) Estimates of M_a using S-LD4M, at different sample sizes. Error bars denote 95% confidence intervals (based on 1,000 simulations), but are smaller than the data points. Numerical results are reported in Supplementary Table 1.

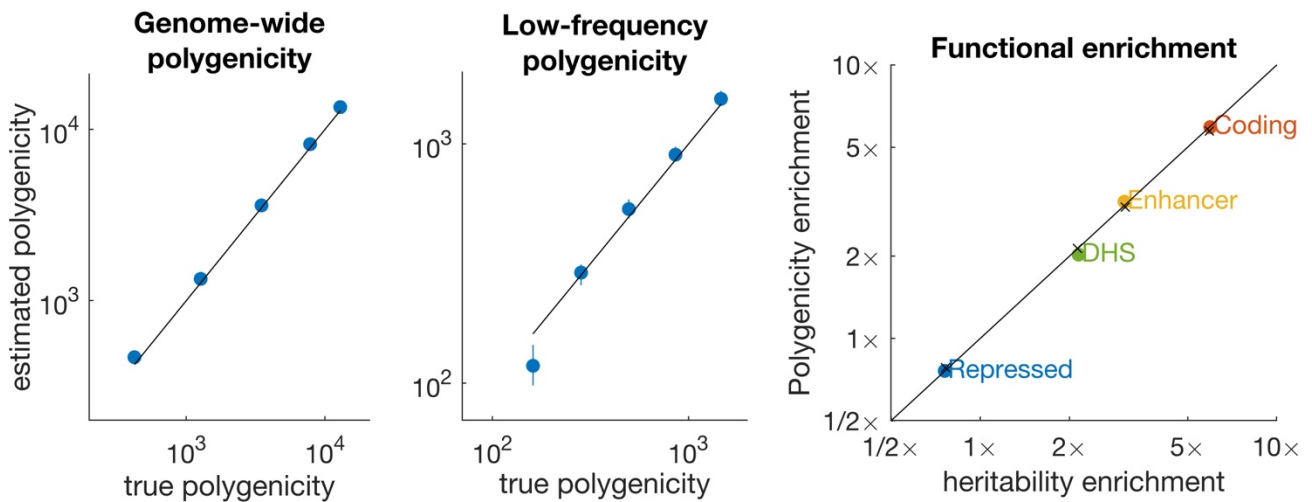


Figure 3. Accuracy of S-LD4M estimates in simulations with LD. (a) Estimates of M_a for all SNPs ($MAF = 0.5 - 50\%$). (b) Estimates of M_a for low-frequency SNPs ($MAF = 0.5 - 5\%$); common-SNP M_a is fixed at $\sim 1,000$ in these simulations. (c) Estimates of polygenicity enrichment and heritability enrichment in simulations with four functional categories. Black lines denotes $y=x$, and colored points denote estimates. In panel (c), \times denotes true values. Error bars denote 95% confidence intervals (based on 1,000 simulations), but are smaller than the data points in most cases. Numerical results are reported in Supplementary Table 1.

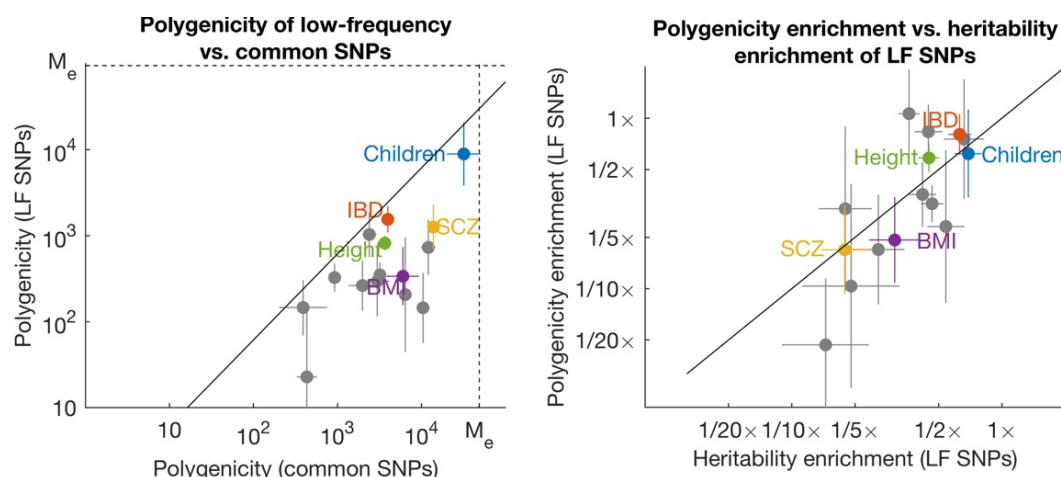


Figure 4. Comparison of common and low-frequency polygenicity across 15 complex traits. (a) Estimates of M_a for common and low-frequency SNPs. Estimates are meta-analyzed across well-powered traits. Common-variant polygenicity was $\sim 4\times$ greater on average than low-frequency polygenicity. Dotted lines denote the effective number of independent SNPs (M_e) for common and low-frequency SNPs respectively, corresponding to an infinitesimal (Gaussian) architecture. The solid line denotes equal per-SNP M_a . (b) Estimates of polygenicity enrichment and heritability enrichment for low-frequency SNPs (compared to all common and low-frequency SNPs). The solid line denotes equal enrichment. Error bars denote 95% confidence intervals. Numerical results are reported in Table 1 and Supplementary Table 5.

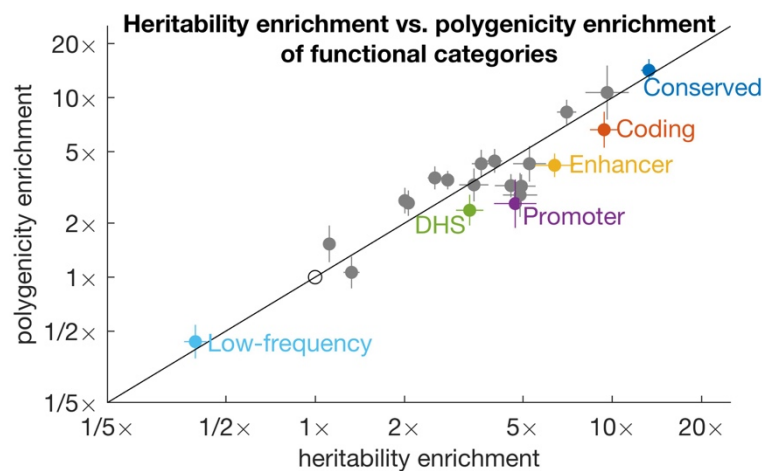


Figure 5. Estimates of polygenicity enrichment and heritability enrichment of functional categories. We report estimates for 20 functional categories plus low-frequency SNPs. Estimates are meta-analyzed across well-powered traits. Error bars denote 95% confidence intervals. Complete results for each trait are reported in Supplementary Table 5, and meta-analyzed results are reported in Supplementary Table 6.

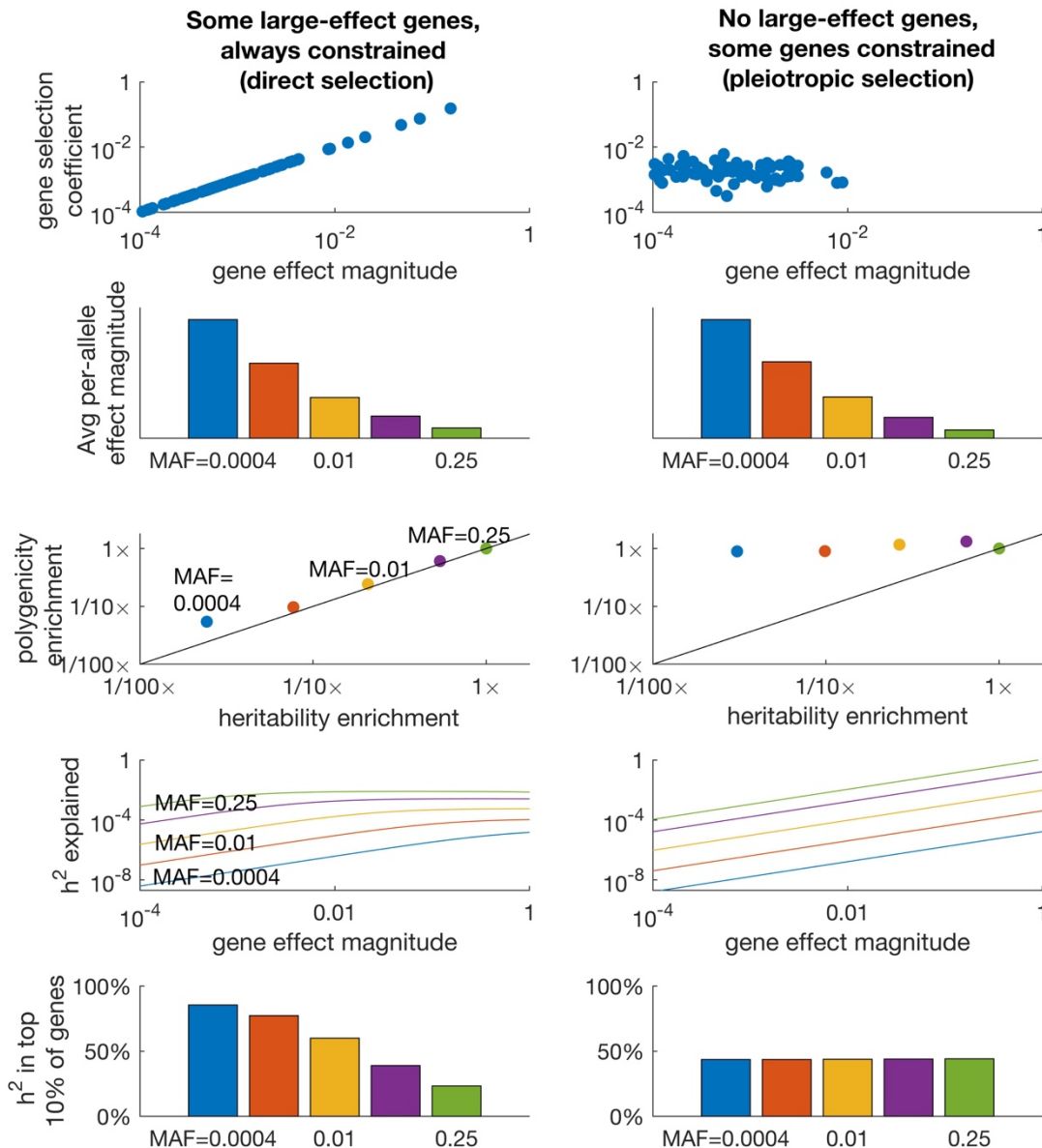


Figure 6 Gene-level flattening under an evolutionary model. In the left column (panels a,c,e,g,i), there are some large-effect genes, but direct stabilizing selection acting on the phenotype strongly constrains these genes. In the right column (panels b,d,f,h,j), there are no large-effect genes; pleiotropic stabilizing selection has varying effects on each gene, limiting common-SNP effect sizes on average. (a-b) Joint distribution of gene effect size magnitudes and selection coefficients. (c-d) Average squared per-allele effect sizes at different allele frequencies. The strength of selection was chosen to produce similar common-variant effect sizes in both columns. (e-f) Heritability and polygenicity enrichment at different allele frequencies (relative to MAF=0.25). Polygenicity at MAF=0.25 is approximately equal for the two columns, due to the different distributions of gene effect sizes. (g-h) Expected heritability explained by a gene as a function of its effect size for SNPs at different frequencies. (i-j) Proportion of heritability explained by the top 10% of largest-effect genes for SNPs at different allele frequencies.

Numerical results are reported in Supplementary Table 7.

Supplementary Tables

Supplementary Table 1 (See Excel file) Numerical results of simulations. Sheets (a)-(f) correspond to Figure 1a-c and Supplementary Figures 1-3, respectively.

<i>Category</i>	<i>Fraction retained</i>			<i>Normalized residuals</i>		
	N=10k	N=50k	N=250k	N=10k	N=50k	N=250k
<i>Coding</i>	0.007	0.39	0.5	NA	0.66	0.66
<i>Enhancer</i>	0.02	0.36	0.57	0.5	0.5	0.68
<i>DHS</i>	0.11	0.72	0.77	0.36	0.59	0.65
<i>Repressed</i>	0.24	0.69	0.81	0.43	0.61	0.64
<i>Low-frequency</i>	0.01	0.88	0.97	NA	0.7	0.81

Supplementary Table 2 Ascertainment and standard errors in simulations at different sample sizes. For each annotation, simulation runs were discarded if heritability for that annotation was not significantly different from zero ($Z < 2$) or if the standard error of the M_a estimate for the annotation was larger than two times the median M_a point estimate. For each annotation, the fraction of simulation runs that were ascertained is reported, as well as the standard deviation of the normalized residuals for the ascertained simulations. The normalized residuals are defined as $\log M_a - \log \widehat{M}_a$ divided by the jackknife standard error, and they should be equal to one if the jackknife standard errors are perfectly calibrated. Values less than 1 indicate conservative standard errors. Jackknife standard errors for polygenicity enrichment are nearly proportional to standard errors for M_a , since there is relatively little noise in the denominator (the estimate of M_a for all common and LF SNPs).

Supplementary Table 3 (see Excel file) Datasets analyzed. 29 UK Biobank traits were selected to have low pairwise genetic correlations and high power, as measured by the significance of the S-LDSC heritability estimate; 28 of these were analyzed in ref.²⁰, and red hair pigmentation was added. Four additional diseases were selected on the basis of availability of summary statistics for low-frequency SNPs. Seven additional datasets were used for replication.

<i>Trait</i>	<i>Reference</i>	<i>N</i>	<i>M</i>	<i>Common $\log_{10} M_a$</i>
<i>Height</i>	UKBB	458k	10M	3.56(0.02)
<i>Height</i>	Lango Allen et al. 2010	131k	1.0M	3.49(0.11)
<i>BMI</i>	UKBB	458k	10M	3.78(0.12)
<i>BMI</i>	Speloites et al. 2010	122k	1.0M	3.51(0.33)
<i>College</i>	UKBB	455k	10M	4.08(0.05)
<i>Years of education</i>	Rietveld et al. 2013	127k	1.0M	4.26(0.18)
<i>Years of education</i>	Okbay et al. 2016	329k	1.0M	4.23(0.09)
<i>Type II Diabetes</i>	UKBB	459k	10M	2.85(0.14)
<i>Type II Diabetes</i>	Morris et al. 2012	61k	1.0M	2.85(0.38)
<i>Neuroticism</i>	UKBB	372k	10M	3.99(0.23)
<i>Neuroticism</i>	Okbay et al. 2016	171k	1.0M	4.26(0.10)
<i>Rheumatoid arthritis</i>	Okada et al. 2014	38k	4.6M	3.03(0.10)
<i>Rheumatoid arthritis</i>	UKBB	459k	10M	3.08(0.20)

Supplementary Table 4 Comparison of common-SNP M_a estimates for traits with multiple available datasets. Standard errors are also reported. No M_a estimates were significantly different ($p < 0.05$ assuming independent errors) for any pair of datasets.

Supplementary Table 5 (See Excel file) Complete results of S-LD4M and S-LDSC on 33 traits.

<i>Category</i>	<i>Heritability enrichment</i>	<i>Polygenicity enrichment</i>	<i>Number of traits</i>	<i>Proportion of SNPs</i>
<i>Conserved (LindbladToh)</i>	13.29(0.51)	14.23(1.28)	21	0.03
<i>TSS (Hoffman)</i>	9.61(1.00)	10.70(2.33)	10	0.02
<i>Coding (UCSC)</i>	9.38(0.64)	6.63(0.97)	16	0.02
<i>Weak Enhancer (Hoffman)</i>	8.56(1.21)	4.86(0.62)	9	0.02
<i>Super Enhancer (Vahedi)</i>	7.03(0.33)	8.32(0.83)	18	0.02
<i>Typical Enhancer (Vahedi)</i>	6.40(0.60)	4.20(0.40)	13	0.02
<i>Enhancer (Andersson)</i>	5.56(2.06)	1.39(1.12)	2	0
<i>Promoter Flanking (Hoffman)</i>	5.46(1.50)	0.98(0.54)	2	0.01
<i>DGF (ENCODE)</i>	5.27(0.42)	4.28(0.61)	18	0.15
<i>Enhancer (Hoffman)</i>	4.94(0.34)	3.22(0.30)	19	0.04
<i>UTR 3 (UCSC)</i>	4.89(0.41)	2.87(0.51)	11	0.01
<i>Promoter (UCSC)</i>	4.71(0.48)	2.57(0.50)	12	0.05
<i>CTCF (Hoffman)</i>	4.59(1.25)	0.61(0.27)	1	0.02
<i>TFBS (ENCODE)</i>	4.56(0.35)	3.23(0.30)	18	0.14
<i>H3K9ac (Trynka)</i>	4.02(0.15)	4.44(0.42)	21	0.14

<i>H3K4me3 (Trynka)</i>	3.62(0.16)	4.29(0.47)	23	0.14
<i>Fetal DHS (Trynka)</i>	3.42(0.24)	3.27(0.44)	14	0.09
<i>DHS (Trynka)</i>	3.31(0.22)	2.36(0.29)	15	0.18
<i>Super Enhancer (Hnisz)</i>	2.79(0.08)	3.47(0.26)	27	0.17
<i>H3K27ac (PGC2)</i>	2.53(0.09)	3.57(0.33)	26	0.28
<i>H3K27ac (Hnisz)</i>	2.06(0.04)	2.58(0.26)	28	0.4
<i>Transcribed (Hoffman)</i>	1.33(0.05)	1.06(0.14)	23	0.36
<i>Intron (UCSC)</i>	1.12(0.03)	1.53(0.22)	23	0.4
<i>Low-frequency (UK10K)</i>	0.40(0.02)	0.44(0.06)	15	0.3

Supplementary Table 6 Polygenicity and heritability enrichment for functional annotations, meta-analyzed across well-powered traits. The number of traits used in the meta-analysis is indicated; traits were excluded if the heritability estimate for a trait-annotation pair was not significantly different from zero, or if the standard error on the M_a estimate was greater than 4 times the median point estimate for that annotation across traits. Annotations were excluded if the number of remaining traits was less than 10 or if the meta-analyzed heritability enrichment estimate was less than 1 (except for the low-frequency category). Standard errors are also reported.

Supplementary Table 7 (see Excel file) Numerical results from Figure 6.

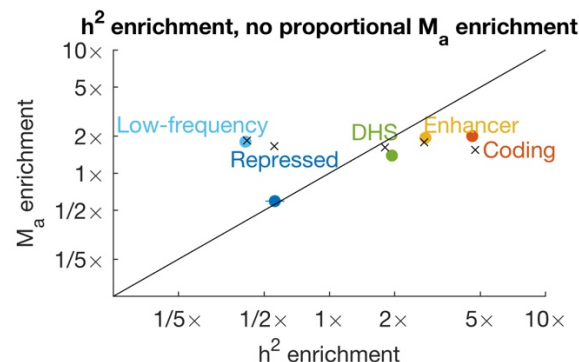
	<i>Enrichment</i>	<i>log₁₀Enrichment</i>
<i>Polygenicity</i>	1.82	0.27 (0.02)
<i>Heritability</i>	1.35	0.13(0.007)
<i>Average effect size</i>	0.71	-0.15 (0.02)

Supplementary Table 8 Polygenicity and heritability enrichment of ExAC genes.

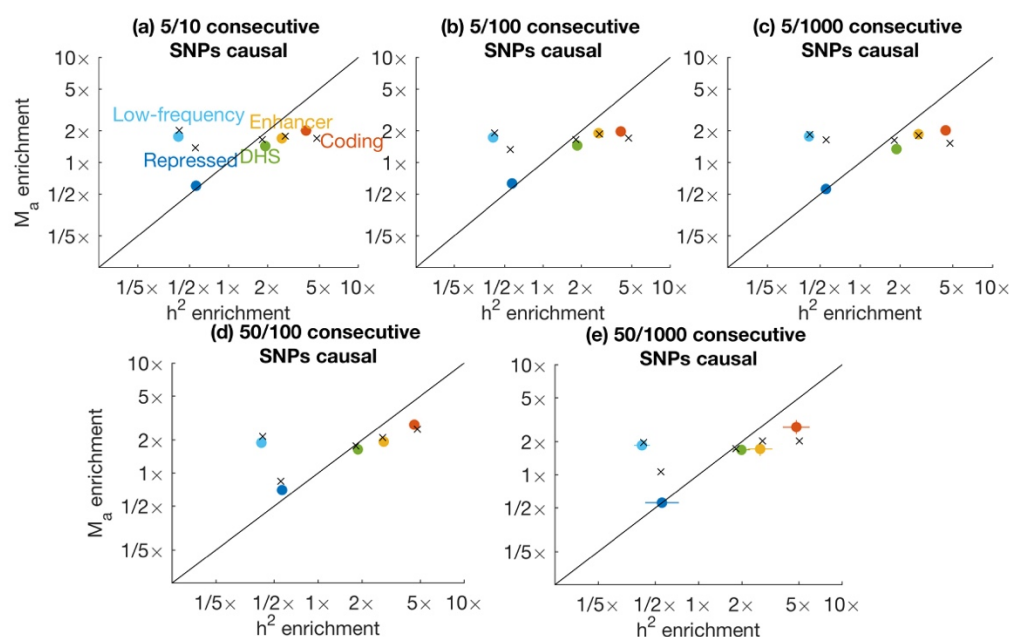
Enrichments are reported for SNPs in and near ExAC LoF-intolerant genes³² compared with SNPs near any gene (in and near is defined as the gene body plus or minus 50kb). The average effect size is equal to the heritability divided by the polygenicity (Figure 2; Methods). Standard errors are also reported.

Supplementary Table 9 (see Excel file) Fine-mapped IBD genes from ref.³⁵ harboring coding and noncoding variants. Coding and noncoding SNPs with > 50% posterior probability and 1-2 nearby genes are listed. pLI values are averaged for SNPs with 2 genes. Genes with both coding and noncoding variants are included in Supplementary Figure 7a and not in panel b.

Supplementary Figures

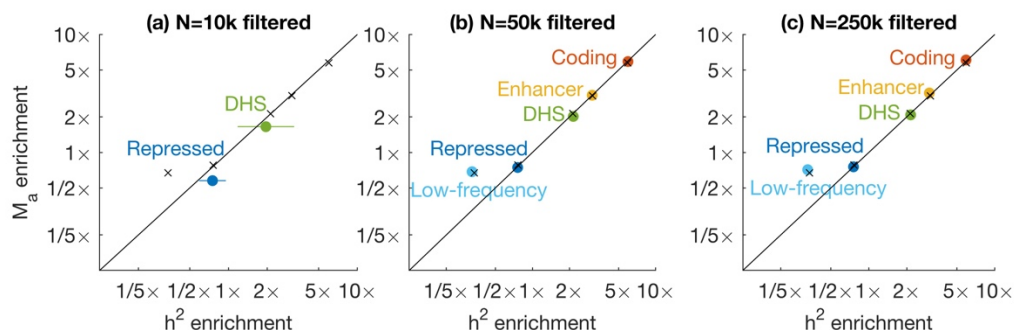


Supplementary Figure 1 Simulations with heritability enrichment and no proportional polygenicity enrichment. Due to some sparsity being driven by differences in heritability enrichment across categories, there is lower sparsity (polygenicity enrichment) within each individual category than within their union. We observed strong downward bias for the repressed category, which is depleted for heritability; we hypothesize that this bias is the result of imperfect resolution to distinguish LD to this category from LD to nearby SNPs, which leads to inflated estimated fourth moments because the nearby SNPs have larger causal effect sizes. This bias has little effect on our estimates for categories depleted for heritability because the nearby SNPs have smaller causal effect sizes than the SNPs in the category, and therefore very little effect on fourth moments; it also has little effect on our estimates for low-frequency SNPs because these SNPs are never in strong LD with common SNPs. In analyses of real traits, we do not report polygenicity enrichment estimates for categories that are depleted for heritability. Based on 1,000 simulations. Error bars indicate 95% confidence intervals. Numerical results are contained in Supplementary Table 1.

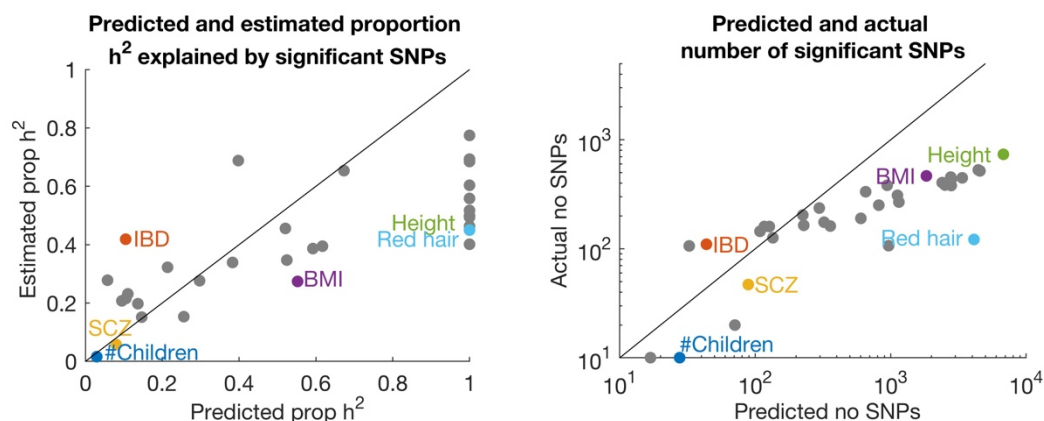


Supplementary Figure 2 Simulations with clustering of causal SNPs. The probability for a

SNP to be causal was zero for most of the genome, and nonzero for contiguous blocks of SNPs of different sizes. (a) Blocks of 10 SNPs with a 50% chance of being causal. (b) Blocks of 100 SNPs with a 5% chance of being causal. (c) Blocks of 1,000 SNPs with a 0.5% chance of being causal. (d) Blocks of 100 SNPs with a 25% chance of being causal. Based on 1,000 simulations. Error bars indicate 95% confidence intervals. Numerical results are contained in Supplementary Table 1.

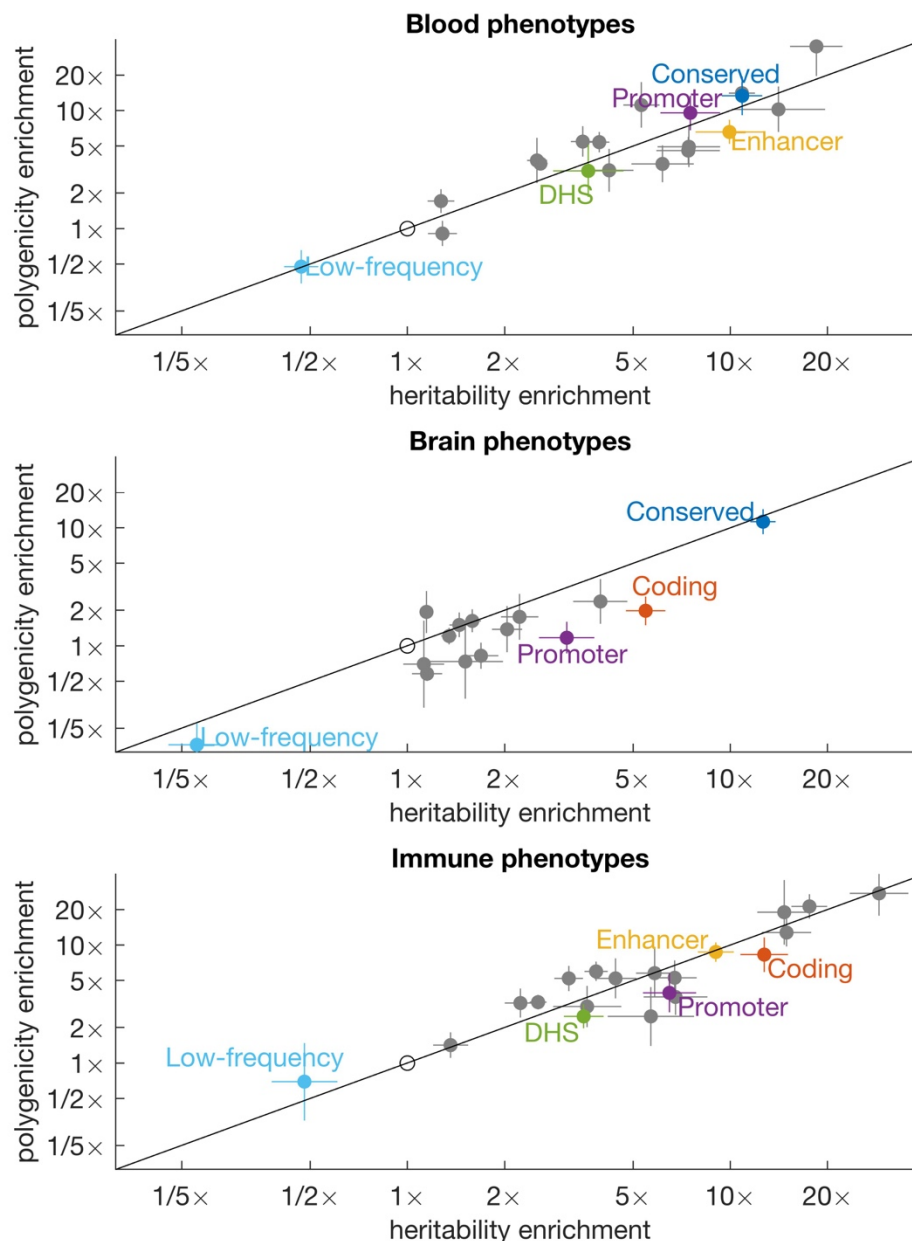


Supplementary Figure 3 Simulations with ascertainment based on power. For each annotation, simulation runs were discarded if heritability for that annotation was not significantly different from zero ($Z > 2$) or if the standard error of the M_a estimate for the annotation was larger than two times the median M_a point estimate. Results are not shown for annotations where the fraction of retained simulations was less than 1%, and confidence intervals are large for annotations where the fraction of retained simulations is low. See Supplementary Table 2 for the fraction of simulations that were retained in each case. Based on 1,000 simulations (before filtering). Error bars indicate 95% confidence intervals. Numerical results are contained in Supplementary Table 1.



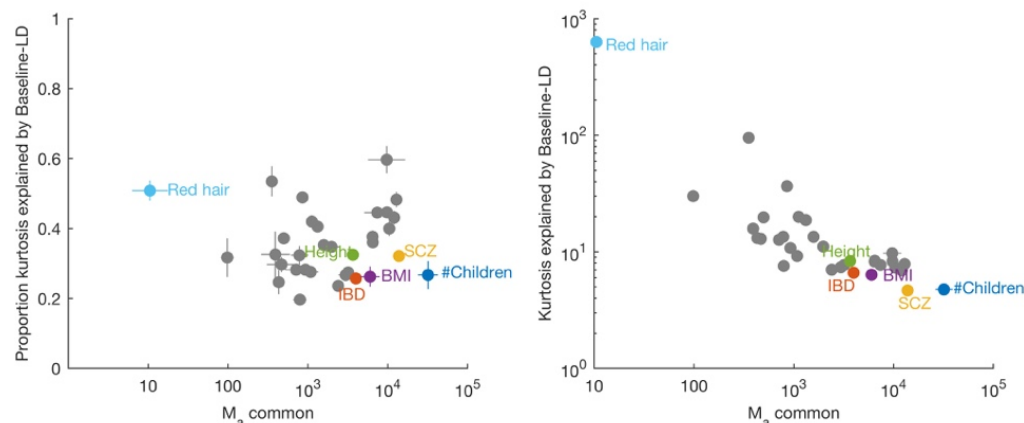
Supplementary Figure 4 Heritability explained by genome-wide significant SNPs. M_a gives a predicted upper bound on the proportion of heritability explained by significant SNPs; similarly, it also gives an upper bound on the number of significant SNPs (Methods). Significant SNPs ($\chi^2 > 30$) were chosen using a greedy pruning procedure, with a gap of at least 0.5cM

between them. We estimated the proportion of heritability explained by these SNPs as the sum of their estimated marginal effect size magnitudes. We caution that this estimate may be upwardly biased due to winner's curse and due to subtle LD between these SNPs.



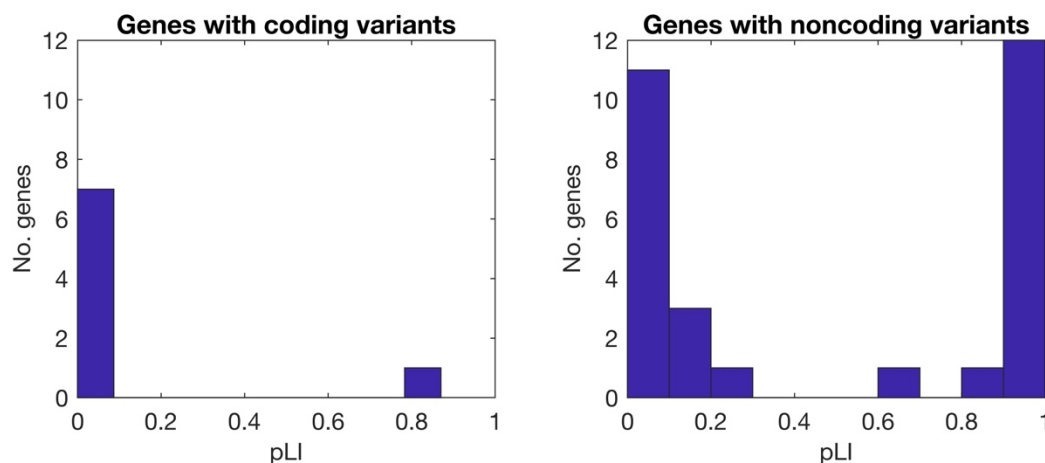
Supplementary Figure 5 Heritability and polygenicity enrichment meta-analyzed across related traits. There were 6 blood-related phenotypes (eosinophil count, platelet count, platelet distribution width, RBC count, RBS distribution width, and WBC count); 8 brain-related phenotypes (Alzheimer's, BMI, college, morning person, neuroticism, smoking status, number of

children and schizophrenia); and 5 immune-related phenotypes (all autoimmune, asthma, eczema, IBD, and RA). Results for each annotation are meta-analyzed across well-powered traits within each group, and each annotation is plotted if at least three traits had a well-powered polygenicity estimates for that annotation.



Supplementary Figure 6 Common-variant sparsity explained by baseline-LD model.

Sparsity is defined as M_e/M_a . Some sparsity is expected due to differences in per-SNP heritability across categories; S-LDSC was used to estimate the causal effect size of each regression SNP, and these estimates were multiplied by the observed χ^2 statistics to estimate the sparsity explained by the model. (a) The proportion of kurtosis explained was defined as the ratio of logs (log-sparsity explained divided by log-sparsity). It ranged between 20% and 50% for most traits, and there was no clear tendency for traits with greater polygenicity to have greater or smaller percent sparsity explained. (b) The amount of sparsity explained was strongly negatively correlated with polygenicity (and positively correlated with total sparsity).



Supplementary Figure 7 Distribution of pLI values for IBD genes with fine-mapped coding and noncoding variants. Causal variants for Crohn's disease and ulcerative colitis were fine mapped in ref. ³⁴. 47 fine-mapped variants with >50% posterior probability and 1-2 annotated protein-coding genes were selected. There were 7 SNPs (all noncoding) with 2 annotated genes, and pLI values were similar (within 0.02) for 6/7 genes; pLI values were averaged for these loci.

There were 3 genes with multiple causal variants (ranging from 2-6), including both coding and noncoding variants; these were counted only once, as coding genes. pLI values were significantly smaller for genes with coding variants than for genes with noncoding variants (single-tailed rank-sum test $p = 0.006$). We caution that pLI values are computed based on allele frequencies of loss of function variants, although high pLI genes may have common missense variants. SNP-by-SNP data is contained in Supplementary Table 7.