Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations

Steven Gazal\textsuperscript{1,2*}, Po-Ru Loh\textsuperscript{2,3}, Hilary K. Finucane\textsuperscript{2,4}, Andrea Ganna\textsuperscript{2,5,6}, Armin Schoech\textsuperscript{1,2,7}, Shamil Sunyaev\textsuperscript{2,3,8} and Alkes L. Price\textsuperscript{1,2,7*}

Common variant heritability has been widely reported to be concentrated in variants within cell-type-specific non-coding functional annotations, but little is known about low-frequency variant functional architectures. We partitioned the heritability of both low-frequency (0.5%≤ minor allele frequency <5%) and common (minor allele frequency ≥5%) variants in 40 UK Biobank traits across a broad set of functional annotations. We determined that non-synonymous coding variants explain 17 ± 1% of low-frequency variant heritability \((h^2_{lf})\) versus 2.1 ± 0.2% of common variant heritability \((h^2_c)\). Cell-type-specific non-coding annotations that were significantly enriched for \(h^2_{lf}\) of corresponding traits were similarly enriched for \(h^2_c\) for most traits, but more enriched for brain-related annotations and traits. For example, H3K4me3 marks in brain dorsolateral prefrontal cortex explain 57 ± 12% of \(h^2_{lf}\) versus 12 ± 2% of \(h^2_c\) for neuroticism. Forward simulations confirmed that low-frequency variant enrichment depends on the mean selection coefficient of causal variants in the annotation, and can be used to predict effect size variance of causal rare variants (minor allele frequency <0.5%).

To investigate functional enrichments of low-frequency variants (defined here as 0.5%≤ MAF<5%), we extended stratified linkage disequilibrium (LD) score regression\textsuperscript{22–23} (S-LDSC) to partition the heritability of both low-frequency and common variants. Our method produces robust (unbiased or slightly conservative) results in simulations. We applied our method to partition the heritability of low-frequency and common variants in 40 heritable traits from the UK Biobank\textsuperscript{24–26} (average \(n=363\) K UK-ancestry samples) across a broad set of coding and non-coding functional annotations\textsuperscript{28,29,30,31}. We performed forward simulations to connect estimated low-frequency and common variant functional enrichments to the action of negative selection, and to predict the effect size variance of causal rare variants (MAF<0.5%) within each functional annotation.

**Results**

**Overview of methods.** S-LDSC\textsuperscript{22–23} is a method for partitioning the heritability causally explained by common variants across overlapping discrete or continuous annotations using GWAS summary statistics for accurately imputed variants and an LD reference panel. Here, we extended S-LDSC to partition the heritability causally explained by low-frequency variants using GWAS summary statistics for accurately imputed and poorly imputed variants. We included separate annotations for low-frequency and common variants, and used whole-genome sequencing (WGS) data from 3,567 UK10K samples\textsuperscript{18} as an LD reference panel to ensure accurate LD information for low-frequency variants in the UK-ancestry target samples analyzed in this study (see Methods).

We jointly analyzed 163 annotations (referred as the “baseline-LF model”), including 33 main binary annotations, MAF bins, and LD-related annotations (Supplementary Table 1 and Supplementary Table 2; see Methods). We note that the inclusion of MAF- and LD-related annotations implies that the expected causal heritability

---

\textsuperscript{1}Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.  \textsuperscript{2}Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA.  \textsuperscript{3}Division of Genetics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA.  \textsuperscript{4}Schmidt Fellows Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA.  \textsuperscript{5}Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA.  \textsuperscript{6}Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA.  \textsuperscript{7}Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.  \textsuperscript{8}Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA.  \textsuperscript{*}e-mail: sgazal@hsph.harvard.edu; aprice@hsph.harvard.edu
of an SNP is a function of MAF and LD. We first estimated the heritability causally explained by all low-frequency variants \( (h^2_{LF}) \) and the heritability causally explained by all common variants \( (h^2_c) \). For the 33 main binary annotations, we computed their low-frequency variant enrichment (LFVE), defined as the proportion of \( h^2_{LF} \) causally explained by variants in the annotation divided by the proportion of low-frequency variants that lie in the annotation, and common variant enrichment (CVE), defined analogically. Further details of the method are provided in the Methods section. We have released open-source software implementing the method, and have made our annotations publicly available (see URLs).

Simulations of extending S-LDSC to low-frequency variants. Although S-LDSC has previously been shown to produce robust results for partitioning common variant heritability using overlapping binary and continuous annotations\cite{1,2,3}, we performed additional simulations to assess our extension to low-frequency variants. We first confirmed that S-LDSC with the UK10K LD reference panel produced unbiased heritability estimates for variants with MAF \( \geq 0.5\% \) in simulations using UK10K target samples (see Supplementary Fig. 1, Supplementary Table 3, and Supplementary Note). We subsequently performed more realistic simulations using target samples from the UK Biobank interim release\cite{4}, so that LD (and MAF) in the target samples and UK10K LD reference panel do not perfectly match (see Methods and Supplementary Fig. 2). S-LDSC was run either by restricting regression variants to accurately imputed variants (that is, INFO score \( \geq 0.99 \)), as we recommended previously\cite{5}, or by including all variants (regardless of INFO score). We focused our simulations on two representative annotations spanning roughly 1% of the genome: coding and enhancer. We considered various MAF-dependent architectures\cite{6,7,8}, and conservatively specified our generative model to be different from the additive model assumed by S-LDSC (see Methods). For each of the two annotations, we simulated scenarios with no functional enrichment (“No Enrichment”) and scenarios with CVE roughly equal to \( 7\% \) and lower LFVE (“Lower LFVE”), similar LFVE (“Same LFVE”), or higher LFVE (“Higher LFVE”), respectively. For both annotations, we observed that including all variants in the regression produced slightly conservative LFVE estimates and unbiased LFVE/CVE ratio estimates, while restricting to accurately imputed variants produced upward biases (Fig. 1, Supplementary Table 4). The slightly conservative \( h^2_c \) and LFVE estimates are due to LD-dependent architectures (coding and enhancer variants have lower than average levels of LD, as do other enriched functional annotations\cite{9}) because we observed nearly unbiased estimates when creating shifted annotations with average levels of LD (see Methods and Supplementary Fig. 3). We thus recommend including all variants in the regression when running S-LDSC using the baseline-LF model. Our simulations indicate that this method is robust (unbiased or slightly conservative) in estimating low-frequency and common variant functional enrichments and LFVE/CVE ratios across a wide range of genetic architectures, even in the presence of poorly imputed variants, a target sample that does not exactly match the UK10K LD reference panel, and a MAF-dependent architecture that does not match the additive model assumed by S-LDSC.

Low-frequency functional architecture of UK Biobank traits. We applied S-LDSC with the baseline-LF model to 40 polygenic, heritable complex traits and diseases from the full UK Biobank release\cite{10} (average \( n = 363 K \); Supplementary Table 5). Analyses were restricted to the set of 409 K individuals with UK ancestry\cite{10} to ensure a close ancestry match with the UK10K LD reference panel. Summary statistics were computed by running BOLT-LMM v2.3 (ref. \cite{11}) on imputed dosages, which were made publicly available (see URLs). S-LDSC results were meta-analyzed across 27 independent traits (average \( n = 355 K \); see Supplementary Note). We observed a roughly linear relationship between estimates of \( h^2_c \) and \( h^2_{LF} \) (Fig. 2 and Supplementary Table 5), with low-frequency variants explaining \( 6.3 \pm 0.2\% \) less heritability and having \( 4.0 \pm 0.1\% \) lower per-variant heritability than common variants on average. These ratios are consistent with a model in which the variance of per-normalized genotype effect sizes is proportional to \( (2p(1-p))^\alpha \) (where \( p \) is the MAF; ref. \cite{12,13}) with \( \alpha = -0.37 \) (95% confidence interval \( CI = [-0.40; -0.34] \); similar to previous \( \alpha \) estimates from raw genotype–phenotype data\cite{14,15,16}), and consistent with a model in which low-frequency variants have smaller heritability but larger per-allele effect sizes\cite{17,18,19,20} (Supplementary Fig. 4).

We compared the LFVE and CVE of the 33 main binary functional annotations of the baseline-LF model, meta-analyzed across traits (Fig. 3, Supplementary Table 6). LFVE were highly correlated to CVE (\( r = 0.79 \) and larger than CVE on average (regression slope = 1.85). We identified nine main functional annotations with significantly different LFVE and CVE (Fig. 3, Supplementary Table 6). Non-synonymous variants had the largest LFVE and largest difference versus CVE (5.0x ratio; \( LFVE = 38.2 \pm 2.3\% \) versus CVE = 7.7 \pm 0.9\% ; \( P = 3 \times 10^{-10} \) for difference). As non-synonymous variants comprise 0.45% of low-frequency variants versus only 0.27% of common variants due to strong negative selection on non-synonymous mutations\cite{21,22,23,24,25,26} (later), this difference is even larger when comparing the proportion of heritability they explain (8.2x ratio; 17.3 \pm 1.0\% of \( h^2_c \) versus 2.1 \pm 0.2\% of \( h^2_{LF} \); \( P = 5 \times 10^{-10} \)). Non-synonymous variants predicted to be deleterious by PolyPhen-2 (ref. \cite{27}) had larger LFVE and LFVE/CVE ratio than non-synonymous variants predicted to be benign (Supplementary Fig. 5).

We also observed LFVE significantly larger than CVE for coding variants (2.5x ratio; \( PF = 1 \times 10^{-10} \)), 5’ UTR (2.5x ratio; \( PF = 1 \times 10^{-7} \)) and the five main conserved annotations\cite{28,29,30} (ratios 1.5x \( \sim 2.2\% \); each \( P < 5 \times 10^{-7} \); Fig. 3, Supplementary Table 6). Surprisingly, phastCons regions conserved in primates\cite{31} were more enriched than phastCons regions conserved in vertebrates or conserved in mammals\cite{32} (even though regions conserved in more distant species may be viewed as more biologically critical). We observed that the significantly larger LFVE (compared to CVE) for all five conserved annotations was mainly due to conserved regions that were coding, and that coding enrichments were similar for regions conserved across different species (Supplementary Fig. 6). Finally, we observed significantly smaller LFVE than CVE for intronic variants (0.85x ratio; \( PF = 8 \times 10^{-4} \)). These results were generally consistent across the 40 UK Biobank traits analyzed (Supplementary Fig. 7).

We also observed significantly larger enrichment/depletion for LFVE than for CVE in the first and/or last quintile of LD-related continuous annotations related to negative selection\cite{33} (Supplementary Fig. 8 and Supplementary Table 7). Our forward simulations from ref. \cite{34} confirmed larger effects of low-frequency variants in these LD-related annotations (Supplementary Table 8). Overall, our results suggest that LFVE is substantially larger than CVE only for annotations that are strongly constrained by negative selection because the strongest differences were observed for coding and non-synonymous variants, which are known to be under strong negative selection\cite{35,36}. A more detailed interpretation of the LFVE/CVE ratio is provided later (see Forward simulations).

CTSS enrichments of low-frequency variants. We sought to investigate the contribution to low-frequency variant architectures of CTS annotations\cite{37,38,39} (that is, reflecting regulatory activity in a given cell type) with excess contributions to common variant architectures. For each of the 40 UK Biobank traits, we selected the subset of 396 CTS Roadmap annotations\cite{40} with statistically significant common variant enrichment after conditioning on (non-CTS annotations in) the baseline-LD model\cite{41} (see Methods). We selected a total of 637 trait–annotation pairs, with at least one CTS annotation for 36 of 40 traits (25 of 27 independent traits) (Supplementary Table 9); the
Fig. 1 | Simulations to assess LFVE estimates. We report estimates of LFVE and LFVE/CVE ratio in simulations under a coding-enriched architecture (first row) or enhancer-enriched architecture (second row). We considered four different simulation scenarios (see main text). S-LDSC was run either by restricting regression variants to accurately imputed variants (S-LDSC - INFO ≥0.99) or by including all variants (S-LDSC - All variants). We do not report LFVE/CVE ratio for the No Enrichment simulation (CVE = LFVE = 1) due to unstable estimates; however, all analyses of real traits in this paper focus on annotations with significant CVE. Results are averaged across 1,000 simulations. Error bars represent 95% CI. Numerical results are reported in Supplementary Table 4.

637 CTS annotations contained 2.7% of common variants and 3.0% of low-frequency variants on average (Supplementary Table 10). We analyzed each of these trait–annotation pairs using the baseline-LF model (Fig. 4a and Supplementary Table 10). For the 25 trait–annotation pairs with the most statistically significant CVE for each of the 25 independent traits (critical CTS annotations), LFVE and CVE were similar, with LFVE 1.12 ± 0.13 larger than CVE on average (other definitions of critical CTS annotations produced similar conclusions; see Supplementary Fig. 9).

We observed Bonferroni-significant differences (after correcting each trait for 1–53 annotations tested) for two traits. The most significant trait–annotation pairs were neuroticism and H3K4me3 in brain dorsolateral prefrontal cortex (4.4× ratio; LFVE = 30.8 ± 6.4× versus CVE = 6.9 ± 1.2×; \(P = 2 \times 10^{-4}\) for difference; 56.9 ± 11.7% of \(h_c^2\) versus 11.7 ± 2.0% of \(h_c^2\), and age of first birth in females and H3K4me3 in brain germinal matrix (5.1× ratio; LFVE = 42.1 ± 10.2× versus CVE = 8.3 ± 1.5×; \(P = 0.001\); 63.2 ± 15.4% of \(h_c^2\) versus 11.1 ± 2.0% of \(h_c^2\)). We note that these results are not driven by the fact that H3K4me3 marks are often located in 5′ UTR and exons38 (Supplementary Table 10). Interestingly, these two annotations (and 55 of all 62 CTS annotations with LFVE/CVE > 2) are brain-specific, implicating stronger selection against variants impacting gene regulation in brain tissues (see Forward simulations and Discussion).

Although CTS annotations generally have only moderately large LFVE (for example, smaller than non-synonymous variants; Fig. 4a), they often explain a large proportion of \(h_c^2\) (for example, larger than non-synonymous variants; Fig. 4b) due to large annotation size, as with CVE. In particular, H3K4me1 in regulatory T-cells (3.7% of low-frequency variants) explains 86.2 ± 20.8% of \(h_c^2\) for all autoimmune diseases (versus 3.4% of common variants explaining 48.9 ± 9.1% of \(h_c^2\)), and H3K4me1 in primary monocytes (4.8% of low-frequency variants) explains 79.3 ± 18.1% of \(h_c^2\) for monocyte count (versus 4.6% of common variants explaining 70.8 ± 8.6%...
would be informative for the action of negative selection, which constrains strongly selected variants to lower frequency\cite{14,17}.

To investigate this, we performed forward simulations\cite{8} using a genetic architecture involving annotations mimicking non-synonymous variants (1% of the simulated genome), functional non-coding variants (1%), and ordinary non-coding variants (98%), with different respective distributions of selection coefficients \( s \) (Supplementary Fig. 11). For each of these three annotations we specified the probability for a de novo variant to be deleterious (\( \pi_{dl} \)), the mean selection coefficient for de novo deleterious variants (\( s_{dl} \)) and the probability for a deleterious variant to be causal for the trait (\( s_{del,causal} \)); the probability for a de novo variant to be causal for the trait is \( \pi = \pi_{dl} s_{del,causal} \). Per-allele trait effect sizes were specified to be proportional to \( \kappa \) (see Methods), implying that only deleterious variants have non-zero effects (see Methods). We investigated how the LFVE and CVE of the functional non-coding annotation varied as a function of the values of \( s_{dl} \) and \( \pi \) for that annotation. To achieve a realistic simulation framework, we fixed the remaining values: \( \pi_{dl} = 0.6 \) for the functional non-coding annotation (similar results for \( \pi_{dl} = 0.4 \); see Methods); \( \pi_{dl} = 80\% \) (ref. 13), \( s_{dl} = -0.003 \) (ref. 13), and \( \pi = 8\% \) for the non-synonymous annotation; \( \pi_{dl} = 40\% \), \( s_{dl} = -0.0001 \), and \( \pi = 4\% \) for the ordinary non-coding annotation; and \( \tau_{EW} = 0.75 \). We note that our fitted value of \( \tau_{EW} \) is larger than previous estimates\cite{14,15,16,17} (see Discussion).

We determined that the CVE of the functional non-coding annotation in our simulations depends on both \( \pi_{dl} \) and \( \pi \) (Fig. 6a), while the LFVE/CVE ratio depends primarily on \( s_{dl} \) (Fig. 6b). When de novo deleterious variants are under strong selection (\( s_{dl} \leq -0.0003 \), corresponding to LFVE/CVE ratio \( \geq 1.2 \times \); Fig. 6b), the CVE depends primarily on \( \pi \) (Fig. 6a) because the mean selection coefficient of deleterious common variants varies only weakly with \( s_{dl} \) (because most deleterious common variants have \( s < s_{del} \); Fig. 6c). Finally, we observed that functional non-coding annotations with similar CVE and LFVE tend to have causal variants with slightly stronger selection coefficients (that is \( s_{del} \approx 0.0002 \)) than ordinary non-coding causal variants (\( s_{del} \approx -0.0001 \)), for which LFVE is lower than CVE (Fig. 6b). We note that the LFVE/CVE ratio can be used to infer the mean selection coefficient of deleterious causal variants as a function of MAF (see Fig. 6c) because this ratio depends primarily on \( s_{dl} \) and because the selection coefficients of de novo deleterious causal variants are drawn from a distribution with mean \( s_{dl} \).

Our forward simulations provide an interpretation of the LFVE/CVE ratios of different functional annotations that we estimated for UK Biobank traits and annotations. First, they confirm that non-synonymous variants (which are strongly deleterious\cite{41}; large \( \pi_{dl} \) and \( s_{del} \)) can have a limited contribution to common variant architectures (2.1% of \( h^2 \)) but a large contribution to low-frequency variant architectures (17.3% of \( h^2 \)) (Fig. 3a). Second, they indicate that the proportion of causal variants (\( \pi \)) is larger for critical CTS annotations than for non-synonymous variants (based on their CVE; Fig. 4a), but that the causal variants in critical CTS annotations have only slightly larger selection coefficients than ordinary non-coding variants, except for some brain annotations that are under much stronger selection (much larger \( s_{del} \), based on their LFVE/CVE ratios; Fig. 4a). Third, they explain the extremely large CVE for non-synonymous variants inside genes predicted to be under strong negative selection\cite{14} (large \( s_{del} \), Fig. 5), which are expected to correspond to genes with an extremely large proportion of deleterious non-synonymous variants (large \( s_{del} \), implying large \( \pi = \pi_{dl} \cdot s_{del,causal} \)).

### Analysis

**Fig. 3** Functional low-frequency and common variant architectures across 27 independent UK Biobank traits. We plot LFVE versus CVE (log scale) for the 33 main functional annotations of the baseline-LF model (meta-analyzed across the 27 independent traits), highlighting annotations for which LFVE is significantly different from CVE. Numbers in the legend represent the proportion of common/low-frequency variants inside the annotation, respectively. The first three conserved annotations are based on phastCons elements\cite{15}; Conserved in mammals\* is based on GERp RS scores\cite{16} (2×), and Conserved in mammals\*\* is based on ref. 20. The promoter flanking annotation has (non-significantly) negative LFVE and is not displayed for visualization purposes. The solid line represents LFVE = CVE; dashed lines represent LFVE = constant multiples of CVE. Error bars represent 95% CI. Numerical results are reported in Supplementary Table 6. of \( h^2 \); Fig. 4b and Supplementary Table 10). Thus, CTS annotations often dominate low-frequency architectures, analogous to common variant architectures\cite{14}.

**Larger non-synonymous enrichments in genes under selection.** Recent studies have identified gene sets that are depleted for non-synonymous variants\cite{17,19}. To further investigate the connection between functional enrichment and negative selection, we stratified the CVE and LFVE of non-synonymous variants (Fig. 3a) based on the strength of selection on the underlying genes. We considered five bins of estimated values of selection coefficients for heterozygous protein-truncating variants\cite{11} (\( s_{dn} \)), with 3,073 protein-coding genes per bin, and added annotations based on non-synonymous variants within each bin to the baseline-LF model (see Methods). We determined that both the LFVE and CVE of non-synonymous variants correlated strongly with the predicted strength of selection on the underlying genes (Fig. 5 and Supplementary Table 11). In particular, we observed extremely strong enrichments for non-synonymous variants in genes under the strongest selection (bin 1: LFVE = 102.0 ± 7.9× and CVE = 41.5 ± 4.8×). However, the LFVE/CVE ratio was smaller for non-synonymous variants in genes under the strongest selection (bin 1: 2.5×) than in genes under the weakest selection (bins 4 + 5: 5.8×). We discuss this surprising result later (see Forward simulations). We obtained similar results when stratifying non-synonymous variants in genes under varying levels of selective constraint based on other related criteria (Supplementary Fig. 10).

**Forward simulations confirm role of negative selection.** We hypothesized that the LFVE and CVE of different functional annotations...
Fig. 4 | Low-frequency and common variant architectures of CTS annotations. a, b, For 637 trait-annotation pairs with conditionally statistically significant common variant enrichment, we report LFVE versus CVE (log scale) (a) and proportion of $h^2_c$ versus proportion of $h^2_f$ explained (b). The dashed black line in a represents the regression slope for 25 critical CTS annotations for independent traits (see main text). Brain-specific annotations are denoted in blue. Two trait-H3K4me3 annotation pairs with LFVE significantly larger than CVE are denoted in dark blue (see main text); error bars represent 95% CI. The two arrows in b denote All autoimmune diseases (H3K4me1 in regulatory T-cells; left arrow) and Monocyte count (H3K4me1 in primary monocytes; right arrow) (see main text). Results for coding and non-synonymous annotations (meta-analysis across 27 independent traits) are denoted in red; error bars represent 95% CI. Numerical results are reported in Supplementary Table 10.

Fig. 5 | Low-frequency and common variant enrichments for non-synonymous variants vary with the strength of selection on the underlying genes. We report LFVE versus CVE (log scale) for non-synonymous variants in five bins of $s_n$ (see main text), meta-analyzed across 27 independent UK Biobank traits; bins 4+5 are merged for visualization purposes. Numbers in the legend represent the proportion of common/low-frequency variants inside the annotation, respectively. The solid line represents LFVE = CVE; dashed lines represent LFVE = constant multiples of CVE. Error bars represent 95% CI. Numerical results for each bin are reported in Supplementary Table 11.

However, despite extremely large CVE and LFVE, this class of variants had a smaller LFVE/CVE ratio than that of non-synonymous variants inside genes predicted to be under weak selection (Fig. 5), a surprising result that appears to suggest a smaller $\beta_{lf}$ (Fig. 6b) despite the extremely large value of $\beta_{dn}$. We performed additional forward simulations to show that a larger $|\beta_{lf}|$ does not produce larger LFVE/CVE ratios for annotations with extremely large values of $\beta_{dn}$, for which the ratio between the proportion of low-frequency variants that are deleterious and the proportion of common variants that are deleterious is reduced to 1 (Supplementary Fig. 12).

Although our focus is primarily on low-frequency variants (0.5% ≤ MAF ≤ 5%), we also used our forward simulation framework to draw inferences about rare variant (MAF < 0.5%) architectures of non-coding functional annotations, based on LFVE and CVE estimates from UK Biobank (Fig. 4a). Specifically, we compared the mean squared per-allele effect size of rare causal variants in annotations mimicking functional non-coding variants and non-synonymous variants, respectively. We inferred disproportionate causal effects of rare variants in annotations under very strong selection ($\beta_{dn} = -0.0003$, similar to non-synonymous variants13), with mean squared causal effect sizes 11×, 26×, and 60× larger than annotations with $\beta_{dn} = -0.0006$, $\beta_{dn} = -0.0003$, and $\beta_{dn} = -0.0002$, respectively (Fig. 6d and Supplementary Table 12; similar results for different choices of $\pi$, Supplementary Fig. 13). These results indicate that an annotation with large CVE needs to have even larger LFVE (for example, LFVE/CVE ratio ≥ 2×, corresponding to $\beta_{dn} ≤ -0.0006$; Fig. 6b) in order to harbor rare causal variants with substantial mean squared effect sizes (for example, only an order of magnitude smaller than rare causal non-synonymous variants; Fig. 6d). Unfortunately, most of the non-brain CTS annotations that we analyzed do not achieve this ratio (Fig. 4a), motivating further work on more precise non-coding annotations (see Discussion).

Discussion

In this study, we partitioned the heritability of both low-frequency and common variants in 40 UK Biobank traits across numerous functional annotations, employing an extension of stratified LD score regression5–7 to low-frequency and common variants, which produces robust (unbiased or slightly conservative) results. Meta-analyzing functional enrichments across 27 independent traits, we highlighted the critical impact of low-frequency non-synonymous variants (17.3% of $h^2_f$, LFVE = 38.2×) compared to common non-synonymous variants (2.1% of $h^2_c$, CVE = 7.7×). Other annotations previously linked to negative selection, including non-synonymous variants with high PolyPhen-2 scores26, non-synonymous variants in genes under strong selection11, and LD-related annotations23, were also significantly more enriched for $h^2_f$ compared to $h^2_c$. Finally, at the trait level, we observed that CTS annotations6,24 also dominate the low-frequency architecture and that significant CVE tended
to have similar LFVE or larger LFVE for brain-related annotations and traits. This last observation implicates the action of negative selection on low-frequency variants affecting gene regulation in the brain, and is consistent with the interaction between brain enhancers and genes under stronger purifying selection, with the excess of rare de novo mutations in regulatory elements active in fetal brain in patients with neurodevelopmental disorders. Using forward simulations, we demonstrated that the CVE of an annotation depends primarily on its proportion of causal variants (π), while its LFVE/CVE ratio depends primarily on the mean selection coefficient for de novo deleterious variants (s\textsubscript{dn}), and thus to the mean selection coefficient of causal variants (Fig. 6). These conclusions are consistent with previous studies of the role of selection, including pleiotropic selection, in maintaining variants with large effects on complex traits at low frequencies. Overall, our work quantifies the relationship between the strength of selection in specific functional annotations (both coding and non-coding) and low-frequency and CVE for human diseases and complex traits, providing an interpretation of the enrichments estimated for UK Biobank traits and annotations.

Our results on low-frequency variant functional architectures have several implications for downstream analyses. First, our results provide guidance for the design of association studies targeting low-frequency variants. Non-synonymous variants should be strongly prioritized at the low-frequency variant level because they explain a large proportion of h\textsuperscript{2} and directly implicate causal genes (and specifically implicate core disease genes rather than peripheral genes), avoiding the challenge of mapping non-coding variants to genes. However, we observed that all coding and UTR variants jointly explained only 26.8 ± 1.9% of h\textsuperscript{2} (Supplementary Table 6), providing an upper bound of the proportion of low-frequency signal captured by whole-exome sequencing (WES) studies. This underscores the advantages of large GWAS (with imputed genotypes obtained using large reference panels), compared to WES or exome chip data, for querying low-frequency variation. Furthermore, using functionally informed association tests that assign higher weight to low-frequency non-synonymous variants or CTS annotations should significantly improve power in these analyses.

Second, our results provide guidance for the design of association studies targeting rare (MAF < 0.5%) variants, which require large sequencing datasets. Although WES datasets have been successfully used to detect new coding variants, genes, and gene sets associated with human diseases and complex traits, there is an increasing focus on WGS that can capture rare non-coding variants. However, our LFVE and CVE results for critical CTS annotations (Fig. 4), coupled with our predictions of causal rare variant effect size variance (Fig. 6d), suggest that in most instances these annotations do not harbor causal variants with large mean squared effect sizes (with brain-related annotations and traits as a notable exception; also see ref. 21), highlighting the need for more precise non-coding annotations for prioritization in WGS. As a first step toward this goal, we estimated the LFVE and CVE of annotations constructed using a...
wide range of recently developed non-coding variant prioritization scores. We identified only one annotation, defined using the top 0.5% of Eigen scores, with an LFVE/CVE ratio significantly larger than 1 (1.7× ratio; LFVE = 22.0 ± 2.2× versus CVE = 13.0 ± 1.4×; P = 7 × 10^{-4} for difference; Supplementary Fig. 14). However, even for this annotation, the LFVE/CVE ratio < 2 again implies that this annotation does not harbor causal variants with substantial mean squared effect sizes (only an order of magnitude smaller than rare causal non-synonymous variants; Fig. 6d). Third, our results were consistent with strong coupling between selection coefficient and trait effect size (Eyre–Walker coupling parameter r_{EW} = 0.75; robust to error bars in LFVE and CVE estimates, see Supplementary Fig. 15), implicating a larger impact of negative selection on complex traits than previously reported \cite{13,14} and much larger effect sizes for rare variants in functional annotations with strong selection coefficients. This can be explained by the fact that our inference procedure explicitly allows different distributions of selection coefficients for non-synonymous and non-coding variants s_{α}^{NN} = -0.003 and s_{α}^{NC} = -0.0001, respectively; Supplementary Fig. 16). Finally, the different LFVE/CVE ratios that we inferred for different functional annotations suggest it may be appropriate to allow annotation-specific α values when using the α model (per-normalized genomic type effect size proportional to (2p(1−p))^{1/2}; ref \cite{1,2,3}). In the extreme case of non-synonymous variants, we explored different choices of α values for non-synonymous and other variants, and determined that a value of α = −1.10 for non-synonymous variants and α = −0.30 for other variants provided the best fit for our UK Biobank heritability and enrichment results (Supplementary Table 13).

Although our work has provided insights on low-frequency variant architectures of human diseases and complex traits, it has several limitations (see Supplementary Note). Despite these limitations, our low-frequency and CVE results convincingly demonstrate and quantify the action of negative selection across coding and non-coding functional annotations.

URLs. ldsc software, http://www.github.com/bulik/ldsc. Baseline-LF annotations, https://data.broadinstitute.org/alkesgroup/LDSCORE/baselinel.F.tar.gz. BOLT-LMM association statistics computed in this study are available at https://data.broadinstitute.org/alkesgroup/UKBB/UKBB_409K.phastCons elements, http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/. Flanking bivalent TSS/enhancers, http://egg2.wustl.edu/roadmap/data/byFileType/chromHmmSegmentations/ChmmModels/coreMarks/jointModel/final/. BOLT-LMM software, https://data.broadinstitute.org/alkesgroup/BOLT-LMM. SLiM2 software, https://messerlab.org/slim/.

Online content Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41588-018-0231-8.

Received: 6 April 2018; Accepted: 7 August 2018; Published online: 8 October 2018

References

1. Maurano, M. T. et al. Systematic localization of common disease-associated variation in regulatory DNA. Science 337, 1190–1195 (2012).
2. Trynka, G. et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat. Genet. 45, 124–130 (2013).
3. Gusev, A. et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. Am. J. Hum. Genet. 95, 535–552 (2014).
4. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. Am. J. Hum. Genet. 94, 559–573 (2014).
5. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 47, 1228–1233 (2015).
6. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015).
7. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omogenomic. Cell 169, 1177–1186 (2017).
8. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat. Genet. 50, 621–629 (2018).
9. 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 (2015).
10. Zeng, J. et al. Signatures of negative selection in the genetic architecture of human complex traits. Nat. Genet. 50, 746–753 (2018).
11. Mancuso, N. et al. Quantification of frequency-dependent genetic architectures and action of negative selection in 25 UK Biobank traits. Preprint at https://www.biorxiv.org/content/early/2017/09/13/188806 (2017).
12. Eyre-Walker, A. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. Proc. Natl. Acad. Sci. USA 107, 1752–1756 (2010).
13. Agarwala, V., Flannick, J., Sunyaev, S., GoT2D Consortium & Altshuler, D. Evaluating empirical bounds on complex disease genetic architecture. Nat. Genet. 45, 1418–1427 (2013).
14. Zuk, O. et al. Searching for missing heritability: designing rare variant association studies. Proc. Natl. Acad. Sci. USA 111, E455–E464 (2014).
15. Fuchsberger, C. et al. The genetic architecture of type 2 diabetes. Nature 536, 41–47 (2016).
16. Simons, Y. B., Bullaughey, K., Hudson, R. R. & Sella, G. A population genetic interpretation of GWAS findings for human quantitative traits. PLoS Biol. 16, e2002985 (2018).
17. The UK10K Consortium. The UK10K project identifies rare variants in health and disease. Nature 526, 82–90 (2015).
18. Acevedo, G. et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. Nat. Genet. 48, 314–317 (2016).
19. Marouli, E. et al. Rare and low-frequency coding variants alter adult human height. Nature 542, 186–190 (2017).
20. Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: study designs and statistical tests. Am. J. Hum. Genet. 95, 5–23 (2014).
21. Gazal, S. et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. Nat. Genet. 49, 1421–1427 (2017).
22. Lupin, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of common diseases of middle and old age. PLoS Med. 12, e1001779 (2019).
23. Bycroft, C. et al. Genome-wide genetic data on ~500,000 UK Biobank participants. Preprint at https://www.biorxiv.org/content/early/2017/07/20/166298 (2017).
24. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. Nat. Genet. 50, 906–908 (2018).
25. Stiegel, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genomes Res. 15, 1034–1050 (2005).
26. Davydov, E. V. et al. Identifying a high fraction of the human genome to be under selective constraint using GERP+++. PLoS Comput. Biol. 6, e1001025 (2010).
27. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248 (2010).
28. Lindblad-Toh, K. et al. The high-resolution map of human evolutionary constraint using 29 mammals. Nature 478, 476–482 (2011).
29. Cassa, C. A. et al. Estimating the selective effects of heterozygous protein-truncating variants from human exome data. Nat. Genet. 49, 806–810 (2017).
30. Gazal, S., Finucane, H. K. & Price, A. L. Reconciling S-LDSC and LDAK functional enrichment estimates. Preprint at https://www.biorxiv.org/content/early/2018/01/30/256342 (2018).
31. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. 11, 499–511 (2010).
32. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. Am. J. Hum. Genet. 91, 1011–1021 (2012).
33. Lee, S. H. et al. Estimation of SNP heritability from dense genotype data. Am. J. Hum. Genet. 93, 1151–1155 (2013).
34. Gazal, S. et al. Recessequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. Nat. Genet. 42, 969 (2010).
35. Tennessen, J. A. et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337, 64–69 (2012).
36. Shiyueva, D., Stemple, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. 15, 272–286 (2014).
An Analysis

39. Ganna, A. et al. Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. Am. J. Hum. Genet. 102, 1204–1211 (2018).

40. Haller, B. C. & Messer, P. W. SLiM 2: flexible, interactive forward genetic simulations. Mol. Biol. Evol. 34, 230–240 (2017).

41. Kryukov, G. V., Pennacchio, L. A. & Sunyaev, S. R. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. Am. J. Hum. Genet. 80, 727–739 (2007).

42. Short, P. J. et al. De novo mutations in regulatory elements in neurodevelopmental disorders. Nature 555, 611–616 (2018).

43. Won, H. et al. Chromosome conformation elucidates regulatory relationships in developing human brain. Nature 538, 523–527 (2016).

44. Claussnitzer, M. et al. FTO obesity variant circuitry and adipocyte browning in humans. N. Engl. J. Med. 373, 895–907 (2015).

45. Kichaev, G. et al. Leveraging polygenic functional enrichment to improve GWAS power. Preprint at https://www.biorxiv.org/content/early/2017/11/20/222265 (2017).

46. Ritchie, G. R. S., Dunham, I., Zeggini, E. & Flicek, P. Functional annotation of non-coding sequence variants. Nat. Methods 11, 294–296 (2014).

47. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46, 310–315 (2014).

48. Ionita-Laza, I., McCallum, K., Xu, B. & Buxbaum, J. D. A spectral approach integrating functional genomic annotations for coding and noncoding variants. Nat. Genet. 48, 214–220 (2016).

49. Huang, Y.-F., Gulko, B. & Siepel, A. Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. Nat. Genet. 49, 618–624 (2017).

50. di Iulio, J. et al. The human noncoding genome defined by genetic diversity. Nat. Genet. 50, 333–337 (2018).

Acknowledgements
We thank A. Gusev, C. Marquez-Luna, M. Hujoel, Y. Reshef, F. Hormozdiari, O. Weissbrod, B. Neale, A. Siepel, and S. M. Gazal for helpful discussions. This research has been conducted using the UK Biobank Resource (application number 16549). This research was funded by NIH grants U01 HG009379, R01 MH101244, R01 MH107649, R01 MH109978 and U01 HG009088. P.R.L. was supported by a Burroughs Wellcome Fund Career Award at the Scientific Interfaces and the Next Generation Fund at the Broad Institute of MIT and Harvard.

Author contributions
S.G. and A.L.P. designed experiments. S.G. performed experiments. S.G., P.R.L., H.K.F., A.G., and A.S. analyzed data. S.G. and A.L.P. wrote the manuscript with assistance from P.R.L., H.K.F., A.G., A.S., and S.S.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41588-018-0231-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.G. or A.L.P.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Extension of S-LDSC to low-frequency variants. S-LDSC\textsuperscript{12} is a method for partitioning heritability explained by common variants across overlapping annotations (both binary and continuous\textsuperscript{12}) using GWAS summary statistics. More precisely, S-LDSC models the vector of per-normalized genotype effect size $\beta$ as a mean-0 vector whose variance depends on $D$ continuous-valued annotations $a_1, a_2, \ldots$.

\[
\text{Var}(\beta) = \sum_{d=1}^{D} a_d \tau_d \tag{1}
\]

where $a_d(j)$ is the value of annotation $a_d$ at variant $j$, and $\tau_d$ represents the per-variant contribution of one unit of the annotation $a_d$ to heritability. We can thus perform a regression to infer the values of $\tau$ using the following relationship with the expected $\gamma^2$ statistic of variant $j$:

\[
E[\gamma_j^2] = N \sum_{d=1}^{D} a_d(j) \tau_d + NB + 1 \tag{2}
\]

where $(j, d) = \sum_{d=1}^{D} a_d(k) j_d^2$ is the LD score of variant $j$ with respect to continuous values $a_d(k)$ of annotation $a_d$, $r_{jk}$ is the correlation between variant $j$ and $k$ in an LD reference panel, $N$ is the sample size of the GWAS study, and $b$ is a term that measures the contribution of confounding biases\textsuperscript{5,6}. Thus, the heritability caused by a subset of variants $S$ can be estimated as $h_{S}^2 = \sum_{S \subseteq D} \sum_{d \in S} a_d \tau_d$. We note that this definition, used here to define and estimate the expected $\gamma^2$ statistic of variants in the annotation (see Supplementary Note for a justification of the expected $\gamma^2$ statistic of variant $j$), which refers to the heritability tagged by a set of genotyped and/or imputed variants.

To allow different effects for low-frequency and common variants inside a functional annotation $a_d$, we modeled the variance of the per-normalized genotype effect sizes using different $\tau_d$ for these two categories of variants. Where we consider $D$ functional annotations, we write:

\[
\text{Var}(\beta) = \sum_{d=1}^{D} a_d(j) (1_{\text{coding}}(j) \gamma_d^{(0)} + 1_{\text{non-coding}}(j) \gamma_d^{(1)}) \tag{3}
\]

where $1_{\text{coding}}(j)$ (respectively, $1_{\text{non-coding}}(j)$) is an indicator function with value 1 if variant $j$ is a low-frequency (respectively, common) variant and 0 otherwise, $\gamma_d^{(0)}$ (respectively, $\gamma_d^{(1)}$) represents the per-variant contribution of one unit of the annotation $a_d$ to the heritability explained by low-frequency (respectively, common) variants. These parameters can be estimated using S-LDSC by writing equation (3) in the form:

\[
\text{Var}(\beta) = \sum_{d=1}^{D} a_d(j) (1_{\text{coding}}(j) \gamma_d^{(0)} + 1_{\text{non-coding}}(j) \gamma_d^{(1)}) \tag{4}
\]

where $a_d^{(0)}(j)$ (respectively, $a_d^{(1)}(j)$) is an annotation equal to $a_d(j)$ if variant $j$ is a low-frequency (respectively, common) variant and 0 otherwise. In all analyses we added one annotation containing all variants, five MAF bins for low-frequency variants, and 10 MAF bins for common variants to take into account MAF-dependent effects\textsuperscript{5,6,23}. For each functional binary annotation of interest $a_d$, we compared its LFVE and CVE, defined as the proportion of $h_{S}^2$ (respectively, $h_{S}^2$) explained by the annotation, divided by the proportion of low-frequency (respectively, common) variants in the annotation (see Supplementary Note for a justification of the denominator). Standard errors were computed using a block jackknife procedure\textsuperscript{12,23}.

Baseline-LF model and functional annotations. We considered 34 main functional annotations from the baseline-LD model v1.1 (27 binary and seven continuous annotations, including LD-related annotations; ref\textsuperscript{3,12,23,25}, including coding, UTR, promoter and intronic regions, the histone marks monomethylation (H3K4me1) and trimethylation (H3K4me3) of histone H3 at lysine 3, acetylation of histone H3 at lysine 9 (H3K9ac) and two versions of acetylation of histone H3 at lysine 27 (H3K27ac), open chromatin as reflected by DNase I hypersensitivity sites, combined chromHMM and Segway predictions (which make use of many Encyclopedia of DNA Elements (ENCODE) annotations to produce a single partition of the genome into seven underlying chromatin states), three different conserved annotations, two versions of super-enhancers, FANTOM5 enhancers, typical enhancers, and six LD-related continuous annotations (see Supplementary Table 1).

To further dissect the set of coding variants, which was a major focus of this study, we included one coding annotation using ANNOVAR\textsuperscript{26} and added one synonymous and one non-synonymous annotation to our model. We also added three new annotations based on phastCons\textsuperscript{27} conserved elements (46 way) in vertebrates, mammas and primates, and one annotation based on flanking bivalent transcription starting sites (TSS)/enhancers from Roadmap data (see URLs). These six new annotations led to a total of 33 main binary annotations (see Supplementary Table 1).

We included 500 bp windows around each binary annotation and 100 bp windows around four of the main annotations, leading to a total of 74 main functional annotations. All annotations were then duplicated for low-frequency and common variants, as described in equation (4), except for the predicted allele age annotation\textsuperscript{28} (which had too many missing values for low-frequency variants). Finally, we included one annotation containing all variants, 10 high-MAF MAF bins (as in the baseline-LD model\textsuperscript{5}) and five low-frequency variant five MAF bins. We thus obtained a set of 163 total annotations. We refer to this set of annotations as the "baseline-LF model" (see Supplementary Table 2), which we used for all our S-LDSC analyses. More details on the baseline-LF model are provided in the Supplementary Note.

We note that the inclusion of MAF and LD-related annotations in this model implies that the expected causal heritability of a SNP is a function of MAF and LD. More details on LD-related heritability models are provided in the Supplementary Note.

Simulations using UK Biobank target samples to assess extension of S-LDSC to low-frequency variants. To assess possible biases in heritability and enrichment estimates under a more realistic scenario, we simulated quantitative phenotypes from chromosome 1 of UK Biobank interim release dataset with imputed variants from 1,000 genomes\textsuperscript{12} and UK10K\textsuperscript{11} (113,851 unrelated individuals, 1,023,659 variants with allele counts greater or equal to five in the UK10K cohort) and randomly sampled integer-valued genotypes from UK Biobank imputation dosage data. Second, we set trait heritability to $h^2 = 0.5$, selected $M = 100,000$ causal variants, and performed simulations under a coding-enriched architecture by simulating the variance of per-normalized genotype effect sizes proportional to $1_{\text{coding}}(2p(1-p)) + c_{\text{coding}}(2p(1-p))$, where $1_{\text{coding}}(j)$ is an indicator function tagging the value 1 if variant $j$ belongs (respectively, does not belong) to the coding annotation, $p$ is the frequency of the causal variant in the simulated UK Biobank genotypes dataset, $\alpha_s$ was set to $-0.25$, and $c$ and $\alpha_{\text{coding}}$ were chosen to produce four different genetic architectures (see Supplementary Table 4). We note that this generative model is different and more complex than the additive inference model implemented in S-LDSC, but may be more realistic because the effect size of coding variants depends now directly on their allele frequency (and not on their low-frequency/common status). We also performed simulations under an enhancer-enhanced architecture by simulating the baseline ChromHMM/Segway weak-enhancer annotation, which has similar properties to the coding annotation (22.8% of reference low-frequency variants versus 1.83% for coding, and elements with a mean length size of 249 bp versus 315 bp for coding). To investigate the impact of the LD-dependent architecture created by the enrichment of these two annotations (coding and weak-enhancer variants tend to have low levels of LD\textsuperscript{28}), we randomly created 100 shifted coding (respectively, weak-enhancer) annotations and selected the annotation with an average level of LD (that is, the LD-related annotation with the 50th percentile of the level of LD computed on low-frequency variants; see ref\textsuperscript{28} for a definition of level of LD). Third, we used v2.3 of BOLT-LMM software\textsuperscript{27} (see URLs) to compute association statistics on UK Biobank DNA datasets to mimic the fact that we computed summary statistics on imputed data. Finally, we used S-LDSC with our baseline-LF model (except that the six new functional annotations were not included in the simulation analyses) to estimate $h_{S}^2$ and coding/enhancer CVE and LFVE. S-LDSC was run by restricting regression variants to accurately imputed variants (that is, INFO score $\geq 0.99$, as we suggested previously), or to all variants (irrespective of INFO score). We reported results when using an INFO score threshold of 0.5 or 0.9, when we did not improve the results (see Supplementary Table 4). We also considered including INFO score explicitly in the regression to downweight poorly imputed variants (that is, replacing equation (4) with $E[\gamma_j^2] = IN_{\theta_{\text{LD}}} + \sigma^2 f_{\text{LD}}(j) + NB + 1$). Where, $I$ is the INFO score of variant $j$ and $\theta = \sum_{j} I_n$ this approximation assumes that genotype uncertainty decreases the association test statistics), but this did not improve the results, which is consistent with the summary statistics computed from dosage data to downweight poorly imputed variants (Supplementary Table 5).

We performed 1,000 simulations for each simulation scenario. In each case, we removed 0–3 outlier simulations in which the estimate of $h_{S}^2$ was below 0.0001. We did not observe any such outlier results in analyses of real traits (minimum $h_{S}^2 = 0.006$; Supplementary Table 5).

S-LDSC analyses of UK Biobank data. We applied S-LDSC with the baseline-LF model to 40 UK Biobank traits, estimated $h_{S}^2$ and $h_{S}^2$, and the $h_{S}^2/h_{S}^2$ ratio using the 15 MAF bin annotations, and computed their standard errors using a jackknife procedure. We meta-analyzed the $h_{S}^2/h_{S}^2$ ratio and multiplied it by the ratio of the number of low-frequency and common variants in the LD reference sample.
Unlike our previous forward simulation framework, where we designed these simulations to have a realistic DFE for annotations mimicking both non-synonymous and non-coding variants. Briefly, we created 30 non-synonymous elements (with a realistic length 200bp (10kb in total, 1% of the 1Mb simulated genome) separated by non-coding elements of size 14.9kb (99% of the simulated genome; Supplementary Fig. 11a). To mimic non-synonymous elements, we used $r_{del}=0.80$, $r_{syn}=−3.16×10^3$, and $\theta=0.32$, as previously estimated. Then, we estimated that fixing $\pi_{c}=40\%$, $\pi_{d}=−1.00×10^4$, $\delta=0.32$ for non-coding variants and $\pi_{d}=0.75$ provide a good fit for our UK Biobank heritability and non-synonymous enrichment results (see Supplementary Note).

In most subsequent simulations, we fixed the probability of a deleterious variant to be causal ($\pi_{del\rightarrow\text{causal}}$) at 10% so that the proportion of de novo non-synonymous variants that are causal ($\pi_{\text{causal}}$, defined as $\pi_{\text{causal}}=\pi_{\text{del}\rightarrow\text{causal}}$) is 8% (respectively, 4% for non-coding variants). This allows non-synonymous variants to have higher LFVE and CVE on the order of magnitude as the LFVE and CVE observed for the non-synonymous variants inside genes predicted to be under strong negative selection (102.0x and 41.4x, respectively; Fig. 5). We note that we replicated our main results when using $\pi_{\text{del}\rightarrow\text{causal}}=5\%$ (Supplementary Fig. 18).

Next, we investigated the impact of $\pi_{\text{d}}$ and $\pi_{\text{c}}$ on a “functional non-coding” annotation. To do so, we alternatively considered 200kb functional elements as non-synonymous elements (1% of the simulated genome) or as functional non-coding elements (1% of the simulated genome) separated by “ordinary non-coding” elements of size 9.8kb (98% of the simulated genome; Supplementary Fig. 11b). For each functional non-coding element, we fixed $\pi_{\text{d}}=60\%$ and $\theta=0.32$ (equal to the value of $\theta$ for non-synonymous and overall non-coding elements). We chose a value $\pi_{\text{c}}=40\%$ for both functional non-coding annotations (Supplementary Fig. 19). We varied $\pi_{\text{d}}$ and $\pi_{\text{c}}$ (and thus $\pi_{\text{d}}$) of the functional non-coding annotation while retaining $\pi_{\text{d}}=10\%$ for the variants in the non-synonymous and ordinary non-coding elements. We varied $\pi_{\text{c}}$, on the log scale and reported truncated values in the manuscript for simplicity; for example, $\pi_{\text{c}}=−0.003$ stands for $\pi_{\text{c}}=0.32\%$; see Supplementary Table 12 for exact $\pi_{\text{c}}$ values. For each scenario, we simulated 1,000 regions of 1 Mb for each scenario, merged the outputted variants, and considered 10 randomly chosen sets of causal variants.

When drawing inferences about rare variant (MAF<0.5%) architectures of non-coding functional annotations, we focused on simulations with $\pi_{\text{c}}=48\%$ for the functional non-coding annotations because the CVE and LFVE/CVE ratios for the functional non-coding annotations were lower than for non-coding functional annotations in the human genome. However, we note that we obtained similar results when choosing $\pi_{\text{c}}=40\%$ for the functional non-coding annotation (Supplementary Fig. 19). We varied $\pi_{\text{d}}$ and $\pi_{\text{c}}$ (and thus $\pi_{\text{d}}$) of the functional non-coding annotation while retaining $\pi_{\text{d}}=10\%$ for the variants in the non-synonymous and ordinary non-coding elements. We varied $\pi_{\text{c}}$, on the log scale and reported truncated values in the manuscript for simplicity; for example, $\pi_{\text{c}}=−0.003$ stands for $\pi_{\text{c}}=0.32\%$; see Supplementary Table 12 for exact $\pi_{\text{c}}$ values. For each scenario, we simulated 1,000 regions of 1 Mb for each scenario, merged the outputted variants, and considered 10 randomly chosen sets of causal variants.

When drawing inferences about rare variant (MAF<0.5%) architectures of non-coding functional annotations, we focused on simulations with $\pi_{\text{c}}=48\%$ for the functional non-coding annotations because the CVE and LFVE/CVE ratios for the functional non-coding annotations were lower than for non-coding functional annotations in the human genome. However, we note that we obtained similar results when choosing $\pi_{\text{c}}=40\%$ for the functional non-coding annotation (Supplementary Fig. 19). We varied $\pi_{\text{d}}$ and $\pi_{\text{c}}$ (and thus $\pi_{\text{d}}$) of the functional non-coding annotation while retaining $\pi_{\text{d}}=10\%$ for the variants in the non-synonymous and ordinary non-coding elements. We varied $\pi_{\text{c}}$, on the log scale and reported truncated values in the manuscript for simplicity; for example, $\pi_{\text{c}}=−0.003$ stands for $\pi_{\text{c}}=0.32\%$; see Supplementary Table 12 for exact $\pi_{\text{c}}$ values. For each scenario, we simulated 1,000 regions of 1 Mb for each scenario, merged the outputted variants, and considered 10 randomly chosen sets of causal variants.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. ldsc software is available at http://www.github.com/bulik/ldsc. A tutorial for running our extension of stratified LD score regression is available in https://data.broadinstitute.org/alkesgroup/LDSCORE/baselineF.tar.gz.

Data availability. Baseline-LF annotations are available at https://data.broadinstitute.org/alkesgroup/LDSCORE/baselineF.tar.gz. BOLT-LMM association statistics computed in this study are available at https://data.broadinstitute.org/alkesgroup/UKBB/UKBB_409K.

References.

51. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. 47, 291–295 (2015).
52. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 42, 565–569 (2010).
53. Lee, S. H. et al. Estimating the proportion of variation in susceptibility to human height. Nat. Genet. 42, 565–569 (2010).
54. Yzsm, J. et al. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 42, 565–569 (2010).
55. Moore, C. B. et al. Low frequency variants, collapsed based on biological knowledge, uncover complexity of population stratification in 1000 genomes project data. PLoS Genet. 9, e1003959 (2013).
56. Leslie, S. et al. The fine-scale genetic structure of the British population. Nature 519, 309–314 (2015).
57. Liu, X. et al. Functional architectures of local and distal regulation of gene expression in multiple human tissues. Am. J. Hum. Genet. 100, 605–616 (2017).
58. Hormozdian, F. et al. Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. Nat. Genet. 50, 1041–1047 (2018).
59. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164 (2010).
60. Rasmussen, M. D., Hubisz, M. J., Gronau, I. & Siepel, A. Genome-wide inference of ancestral recombination graphs. *PLoS Genet.* **10**, e1004342 (2014).
61. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
62. Hoffman, M. M. et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* **41**, 827–841 (2013).
63. Loh, P.-R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
64. Vahedi, G. et al. Super-enhancers delineate disease-associated regulatory nodes in T-cells. *Nature* **520**, 558–562 (2015).
65. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
66. Gravel, S. et al. Demographic history and rare allele sharing among human populations. *Proc. Natl. Acad. Sci.* **108**, 11983–11988 (2011).
67. Nordborg, M. & Krone, S. M. Separation of time scales and convergence to the coalescent in structured populations. In *Modern Developments in Theoretical Population Genetics: The Legacy of Gustave Malécot* (eds. Slatkin, M. & Veuille, M.) Ch. 12 (Oxford Univ. Press, New York, 2002).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- n/a  Confirmed
- □  The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- □  An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- □  The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- □  A description of all covariates tested
- □  A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- □  A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- □  For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- □  For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- □  For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- □  Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
- □  Clearly defined error bars
  - State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection  We did not collect data for this study. We made all summary statistics analyzed available at https://data.broadinstitute.org/alkesgroup/UKBB/UKBB_409K.

Data analysis  Our ldsc software is available at http://www.github.com/bulik/ldsc. Annotations and a tutorial is available at https://data.broadinstitute.org/alkesgroup/LDSCORE/baselineLF.tar.gz.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The URL section of the paper contains the link of all the publicly available datasets that we analyzed.
Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | We used the larger sample size that was available. We restricted our analyses to UK Biobank traits for which the z-score for nonzero h2c computed using S-LDSC with the baseline-LF model was at least 10, to maximize robustness of h2lf. |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | Our analyses were restricted to the set of 409K individuals with UK ancestry to ensure a close ancestry match with the UK10K LD reference panel. We excluded the HLA from all analyses and analyzed only autosomes. |
| Replication | No replication dataset was analyzed as our analyses are already based on a meta-analyses of several independent traits. |
| Randomization | We performed no randomization and analyzed all individuals from UK Biobank. |
| Blinding | We did not collect data for this study, but analyzed summary statistics from UK Biobank. |

Reporting for specific materials, systems and methods

| Materials & experimental systems | n/a |
|---------------------------------|-----|
| Involved in the study | - Unique biological materials
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants |

| Methods | n/a |
|---------|-----|
| Involved in the study | - ChiP-seq
- Flow cytometry
- MRI-based neuroimaging |