Negative short-range genomic autocorrelation of causal effects on human complex traits

Armin P. Schoech, Omer Weissbrod, Luke J. O'Connor, Nick Patterson,
Huwenbo Shi, Yakir Reshef, Alkes L. Price

Abstract

Most models of complex trait genetic architecture assume that signed causal effect sizes of each SNP (defined with respect to the minor allele) are uncorrelated with those of nearby SNPs, but it is currently unknown whether this is the case. We develop a new method, autocorrelation LD regression (ACLR), for estimating the genome-wide autocorrelation of causal minor allele effect sizes as a function of genomic distance. Our method estimates these autocorrelations by regressing the products of summary statistics on distance-dependent LD scores. We determined that ACLR robustly assesses the presence or absence of nonzero autocorrelation, producing unbiased estimates with well-calibrated standard errors in null simulations regardless of genetic architecture; if true autocorrelation is nonzero, ACLR correctly detects its sign, although estimates of the autocorrelation magnitude are susceptible to bias in cases of certain genetic architectures. We applied ACLR to 31 diseases and complex traits from the UK Biobank (average N=331K), meta-analyzing results across traits. We determined that autocorrelations were significantly negative at distances of 1-50bp ($P = 8 \times 10^{-6}$, point estimate -0.35 ± 0.08) and 50-100bp ($P = 2 \times 10^{-3}$, point estimate -0.33 ± 0.11). We show that the autocorrelation is primarily driven by pairs of SNPs in positive LD, which is consistent with the expectation that linked SNPs with opposite effects are less impacted by natural selection. Our findings suggest that this mechanism broadly affects complex trait genetic architectures, and we discuss implications for association mapping, heritability estimation, and genetic risk prediction.

Introduction

Analyses of complex trait genetic architecture, i.e. the genome-wide distribution of causal genetic effects, have substantially improved our understanding of complex trait biology, informing genetic association studies and genetic risk prediction[1-16]. One aspect of genetic architecture that has not been well-studied is genomic autocorrelation, i.e. autocorrelation of signed causal effect sizes (here defined with respect to the minor allele) as a function of genomic distance. On the one hand, genomic autocorrelation could potentially be positive, if the minor alleles of two nearby SNPs disrupt the same functional element[17;18]. On the other hand, genomic autocorrelation could potentially be negative, if linked SNPs with opposite trait effects escape negative selection, a phenomenon known as linkage masking[19]. Several previous studies have provided specific examples of concordant[20] and opposite[19;21] trait effects of nearby SNPs in humans, as well as concordant[22] and opposite[22;23] trait effects in model organisms. However, genome-wide genomic autocorrelation has not yet been investigated. It is therefore unclear if genome-wide autocorrelation exists, and if so, what its sign, magnitude, and genomic distance range is.

Here, we introduce a new method, autocorrelation LD regression (ACLR), for estimating the genome-wide autocorrelation of causal genetic effects as a function of genomic distance. Specifically, we define genomic autocorrelation as the genome-wide Pearson correlation coefficient between causal minor allele effects of all pairs of genetic variants that are a specific base pair distance apart. ACLR regresses products of genome-wide association study (GWAS) summary statistics on distance-dependent linkage disequilibrium (LD) scores. We evaluate ACLR by performing simulations under a range of genetic architectures,

and apply ACLR to 31 disease and complex traits from the UK Biobank[24] (average N=331K). We further investigate how genomic autocorrelations of nearby SNPs depend on their pairwise LD. We discuss the implications of genomic autocorrelation for analyses of heritability and genome-wide association studies.

Results

Overview of methods

ACLR estimates the genome-wide autocovariance and autocorrelation of causal minor allele effects as a function of base pair distance by regressing products of GWAS summary statistics on distance-dependent LD scores. Products of GWAS summary statistic of two SNPs will tend to be positive if their causal effects are correlated, but also if both SNPs are in LD with (i.e. tag) the same causal SNP. Intuitively, ACLR accounts for the effect of tagging to specifically infer the correlation of causal effects of SNP pairs depending on the base pair distance between them.

In detail, ACLR partitions SNP pairs into several base pair distance ranges (e.g. 1-50bp, 50-100bp, etc.) and assumes a fixed causal effect covariance of per-allele effect sizes for each distance range. We show that the product of GWAS summary statistics of two SNPs i and j has the following expectation (see Methods):

$$\mathrm{E}(\alpha_i\alpha_j) = R_{\cdot i}^T R_{\cdot j}\sigma_0^2 + \sum_{m=1}^n R_{\cdot i}^T S^{(m)} R_{\cdot j}\sigma_m + R_{ij}\sigma_\varepsilon^2/N, \tag{1}$$

where R is the LD matrix (specifically the minor allele count covariance matrix), σ_0^2 is the genome-wide causal per-allele effect size variance, $S^{(m)}$ is the design matrix of distance range m (with $S_{kl}^{(m)}$ taking values of 0 or 1 indicating whether the distance between SNP k and SNP l falls into the mth distance range), σ_m is the genome-wide causal minor allele effect size autocovariance for distance range m, σ_ε^2 is the variance of environmental (non-genetic) effects on the trait, and N is the sample size. σ_m can hence be inferred by regressing $\alpha_i\alpha_j$ on distance-dependent LD scores $R_{\cdot i}^T S^{(m)} R_{\cdot j}$. Unlike previous methods that use external reference LD panels to estimate LD between SNPs[5;25;26], we use in-sample LD estimates from imputed genotypes of the target samples (see Methods), which has been preferred in some recent analyses[27–29]. Distance-dependent causal minor allele effect autocorrelation estimates $\hat{\rho}_m$ are computed by dividing the corresponding autocovariance estimates ($\hat{\sigma}_m$) by estimates of the genome-wide causal per-allele effect size variance ($\hat{\sigma}_0^2$). Standard errors are estimated via genomic block-jackknife (see Methods). We emphasize that autocovariances and autocorrelations are reported with respect to per-allele effects, and not per-standardized allele effects.

Systematic differences in causal per-allele effect size variances between SNPs with different minor allele frequencies (MAF) and levels of LD can lead to substantial bias in related approaches that estimate different quantities[4;8;30–32]. Thus, ACLR allows for "MAF- and LD-corrected" estimates that explicitly allow for these differences (see Methods). Specifically, we model different causal per-allele effect size variances using 10 common SNP MAF bins, 10 low-frequency SNP MAF bins, and 6 LD-related annotations[8;33]. We have publicly released open-source software implementing the ACLR method (see URLs).

Simulations

We evaluated ACLR in simulations under a range of genetic architectures using UK Biobank genotype data (see Methods). We used the same set of 337,502 unrelated British-ancestry individuals as in our analyses of UK Biobank traits. Due to the large number of simulations, we restricted most simulations to chromosome 22 (232,046 SNPs with MAF $\geq$ 0.1%); however, we also performed a selected set of simulations using genome-wide data (16,525,661 SNPs with MAF $\geq$ 0.1%). We performed simulations with uncorrelated causal minor allele effects (null simulations), as well as simulation with nonzero genome-wide autocorrelations (causal simulations). ACLR assumes that causal effect autocovariances are constant for SNP pairs within each distance range, but we included causal simulations in which this assumption

is violated. We set genome-wide heritability to either 0.5 or 0.2. In most simulations we specified standardized causal effect sizes using the baseline-LF model of Gazal et al.[33], which models MAF- and LD-dependent effects; however, we also performed simulations in which causal per-allele effect sizes are independent of MAF and LD. We evaluated the accuracy of both autocovariance and autocorrelation estimates, as autocovariance estimates were more accurate (see below) but autocorrelation estimates are more intuitively interpretable.

We first performed null simulations to assess bias and calibration when true causal effects are uncorrelated. Results are reported in Table 1. We determined that MAF- and LD-corrected autocovariance and autocorrelation estimates were unbiased, with well-calibrated standard errors, for all genetic architectures simulated. However, as expected, estimates without MAF- and LD-correction were biased in the case of MAF- and LD-dependent genetic architectures. Although these simulations used data from chromosome 22 only, a subset of simulations using genome-wide data yielded similar conclusions (see Supplementary Table 1).

We then performed causal simulations, to assess how well ACLR can detect and quantify nonzero autocovariance and autocorrelation. We first discuss autocovariance estimates in causal simulations (Table 2). We determined that MAF- and LD-corrected autocovariance estimates were unbiased, with well-calibrated standard errors, for all genetic architectures in which true causal effect autocovariances were constant for SNP pairs within each distance range (as assumed by the inference method). On the other hand, autocovariance estimates were biased when true causal effect autocovariances were not constant for SNP pairs within each distance range - although the sign of the nonzero autocovariance was still correctly inferred. This bias due to unmodeled variation in autocovariance signal across pairs of SNPs is analogous to biases that can arise in heritability estimation due to unmodeled variation in per-SNP heritability[4;8;30–32].

We next discuss autocorrelation estimates in causal simulations (Table 2). In simulations without MAF- and LD-dependent genetic architectures, autocorrelation estimates were unbiased, with well-calibrated standard errors, analogous to null simulations. However, in simulations with MAF- and LD-dependent genetic architectures, autocorrelation estimates were biased - even when true causal effect autocovariances were constant for SNP pairs within each distance range, such that autocovariance estimates were unbiased; in each case, the sign of the nonzero autocorrelation was still correctly inferred. Since autocorrelation estimates are computed by dividing autocovariance estimates by SNP causal effect variance estimates, bias in autocorrelation estimates in the absence of bias in autocovariance estimates can be explained by bias as well as noise in SNP causal effect variance estimates, although we did not observe the latter playing a significant role in our simulation results.

In summary, ACLR yields unbiased autocovariance and autocorrelation estimates with well-calibrated standard errors in all null simulations. ACLR detects nonzero autocovariance and autocorrelation and correctly infers their sign in all causal simulations, although the magnitude of the autocovariance and autocorrelation estimates can be biased depending on the genetic architecture simulated. Thus, any significantly nonzero autocovariance/autocorrelation reported by ACLR is a robust result, in the sense that the true autocovariance/autocorrelation will be nonzero with the reported sign.

Negative short-range genomic autocorrelation across 31 UK Biobank traits

We estimated the genome-wide autocovariance and autocorrelation of causal minor allele effect sizes, as a function of genomic distance, for 31 strongly heritable diseases and complex traits from the UK Biobank (Supplementary Table 2; average N=330,618). We analyzed $16,525,661$ genome-wide SNPs with MAF $\geq 0.1\%$ and applied ACLR to summary association statistics computed using BOLT-LMM[7;34] together with in-sample LD from imputed genotypes (see Methods). We meta-analyzed results across 25 independent traits with pairwise genetic correlation less than 0.5 (see Methods).

Autocorrelation estimates are reported in Figure 1 and Supplementary Table 3, and the corresponding autocovariance estimates are reported in Supplementary Table 4 (autocovariance estimates are less intuitively interpretable, with absolute values ranging from the order of 10^{-8} to 10^{-10} across traits for variants less than 1000bp apart). We detected significantly negative autocorrelations in the meta-analysis at distances of 1-50bp ($P = 8 \times 10^{-6}$, point estimate -0.35 ± 0.08) and 50-100bp ($P = 2 \times 10^{-3}$, point esti-

mate -0.33 ± 0.11). Results were fairly consistent across traits (e.g. 1-50bp: negative estimates for 15/16 traits with standard errors < 0.5), and we did not detect statistically significant heterogeneity across traits for any distance range (see Supplementary Table 5). We observed no significant deviation from zero for 100-200bp and 200-500bp, and a slightly positive autocorrelation at 500-1000bp ($P = 3 \times 10^{-3}$, point estimate 0.07 ± 0.02). No single-trait autocorrelation estimate was significantly different from zero for any distance range after Bonferroni correction, though several were nominally significant at a significance threshold of 0.05. We also estimated the effect autocorrelation for SNPs more than 1000bp apart, but estimates were very small (absolute value $< 10^{-3}$ for all traits) and not significantly different from zero (Supplementary Table 6).

In a more stringent meta-analysis across 13 strictly independent traits with pairwise genetic correlation less than 0.1 (Supplementary Table 7), the negative autocorrelation at 1-50p remained highly significant ($P = 10^{-4}$) but the remaining autocorrelations were not significant after correcting for hypotheses tested for six distance ranges. We thus consider the negative short-range autocorrelation to be a robust result. As noted above, our simulations showed that conclusions about deviations from zero are robust, as ACLR was unbiased and well-calibrated in null simulations regardless of genetic architecture. Our simulations also showed that, in the case of true nonzero autocorrelations, autocorrelation estimates can have biased magnitudes. Specifically, autocorrelation estimates can be biased if true causal effect autocovariances are not constant for SNP pairs within each distance range (in which case autocovariance estimates are biased); it is possible that this phenomenon impacts our results on real traits. Autocorrelation estimates can also be biased when true causal effect autocovariances are constant for SNP pairs within each distance range (in which case autocovariance estimates are unbiased, but autocorrelation estimates can still be biased due to noise (or bias) in per-allele effect variance estimates); it is unlikely that this phenomenon significantly impacts our results on real traits, because heritability estimates produced by ACLR based on effect variance estimates are similar to previously published estimates (Supplementary Table 2) and because we only included traits that are strongly heritable (z-score ≥ 6 for nonzero heritability; see Methods).

Previous studies have pointed out that linked SNPs with opposite effects escape the action of negative selection, a phenomenon known as linkage masking[19;23]; specifically, a haplotype harboring two SNPs with opposite effects will be less impacted by negative selection because the effects will (partially) cancel out in any individual carrying that haplotype. Linkage masking could potentially explain the systematically negative short-range genomic autocorrelation that we inferred in this study. The very short genomic distance range (on the order of 100bp) of the autocorrelation that we detected does not rule out this explanation: although substantial LD is common at much longer ranges, it is weaker at longer ranges, which could potentially reduce the effects of linkage masking[35]. In addition, linkage masking effects at very short genomic distance range could be amplified by functional similarity at this range, e.g. due to compensatory mutations in binding sites or coding regions; SNP with opposite effects may be much less likely to arise at longer ranges.

We hypothesized that, under the linkage masking hypothesis, the negative short-range autocorrelation would be stronger for pairs of SNPs with positive LD (between minor alleles). We repeated our analyses while stratifying pairs of SNPs by their pairwise LD, in addition to genomic distance (see Methods); we confirmed that this approach produces robust results in simulations (Supplementary Table 8). Consistent with our hypothesis, short-range (1-100bp) autocorrelations were significantly negative for pairs of SNPs in positive LD, but significantly positive for pairs of SNPs in negative LD, in a meta-analysis across the 25 independent traits (Supplementary Table 9). We note that positive autocorrelation of trait effects of alleles in negative LD is also consistent with linkage masking.

Heritability estimation methods commonly assume that SNP-heritability is equal to the genome-wide sum of standardized per-allele effect variances[2–6;8;10;12–15]. While this assumption is correct in commonly used complex trait models that assume uncorrelated per-allele effects, it is incorrect if per-allele effects are significantly correlated, as we have demonstrated. Specifically, if linked SNPs have negatively correlated effects, the true SNP-heritability will be lower than the sum of standardized per-allele effect variances. We introduce a measure that we call "heritability shrinkage" that quantifies the relative decrease in heritability due to negatively correlated SNP effects (see Methods). Based on our autocorrelation estimates stratified by pairwise LD between pairs of SNPs (Supplementary Table

9), we derived a hypothetical heritability shrinkage estimate of $(13 \pm 2)\%$ (see Methods), indicating that genomic autocorrelation may substantially impact the genetic architecture of a complex trait. We caution that our heritability shrinkage estimate should be viewed only as a hypothetical estimate, as it relies on specific autocorrelation estimates, which can be biased according to our simulations. Although negative genomic autocorrelation leads to reduced heritability compared to this commonly used definition, negative autocorrelation can also lead underestimation of genetic effects due to linkage masking. As these biases act in opposite directions it has been suggested that they might cancel each other out [14], though the overall impact of genomic autocorrelation on these heritability estimation methods has not been explicitly analyzed.

Discussion

We developed a new method, ACLR, to estimate genomic autocovariance and autocorrelation of causal effects on human complex traits as a function of base pair distance. We determined that ACLR is unbiased with well-calibrated standard errors in null simulations with zero autocorrelation; ACLR produced biased autocorrelation estimates in causal simulations with nonzero autocorrelation under some genetic architectures, but always detected the correct sign of the autocorrelation. Thus, any significantly nonzero autocorrelation reported by ACLR is a robust result, in the sense that the true autocorrelation must be nonzero with the reported sign. We applied ACLR to 31 diseases and complex traits from the UK Biobank [24], detecting significantly negative autocorrelation at very short genomic distance ranges (up to 100bp) when meta-analyzing results across traits. The negative autocorrelation could be explained by opposite trait effects of linked SNPs cancelling each other out and escaping the action of negative selection, a phenomenon known as linkage masking [19]. The very short genomic distance range of the negative autocorrelation is unsurprising, as linkage masking effects at very short genomic distance range could be amplified by functional similarity at this range, e.g. due to compensatory mutations in binding sites or coding regions; SNP with opposite effects may be much less likely to arise at longer ranges. We further determined that the negative autocorrelation is primarily driven by pairs of SNPs in positive LD, consistent with the linkage masking hypothesis. We note that the negative autocorrelation has significant implications for association mapping, heritability estimation, and genetic risk prediction (see below).

Several previous studies have provided specific examples of opposite (and concordant) trait effects in humans and model organisms (although none of these studies estimated genome-wide autocorrelation of causal trait effects). In humans, Brown et al. [19] used a multi-SNP association approach to identify linked pairs of SNPs with opposite effects, on the other hand results by Zhou et al. [20] based on deep learning-predicted genetic effects showed that mutational effects on gene expression often tend to be concordant within genes. In model organisms, Bernstein et al. [23] used high throughput phenotyping in *Caenorhabditis elegans* to identify several instances in which neighboring genetic variants have opposite effects, and She and Jarosz [22] estimated single nucleotide effects of several quantitative traits in a yeast cross experiment, finding an excess of nearby mutations that are functionally coupled, with examples of both aligned and opposite signed effects. Brown et al. [19] attributed their findings to linkage masking caused by negative selection; however, our results imply that linkage masking is far more pervasive than previously known. (We note that Brown et al. [19] introduced the term linkage masking to emphasize linked genetic variants with opposite signed effects escaping detection in GWAS rather than escaping natural selection). Negative selection is known to broadly impact complex trait architectures, resulting in MAF-dependent [4;9;15;33], LD-dependent [8] and extremely polygenic [14] architectures. Recently, Garcia and Lohmueller [21] showed that negative autocorrelation of fitness effects of linked SNPs can arise under additive negative selection in evolutionary forward simulations, and detected signatures of this effect when comparing synonymous and non-synonymous SNP pairs in 1000 Genomes Project data [36]; however, Garcia and Lohmueller [21] did not analyze complex trait data. In addition, Zhou et al. [37] developed a method to estimate covariances between complex trait effects across fixed sets of genetic variants; however, unlike the method presented here, their method cannot be used to estimate autocovariance or autocorrelation.

Our results have several implications for future work. First, they indicate that linked SNPs with opposite effects are pervasive throughout the genome, likely leading to many disease risk variants that

are undetected (or detected with underestimated effect) in GWAS. This motivates the use of improved association methods that allow for joint testing of multiple proximal variants[19]. Second, our findings violate the assumption of uncorrelated SNP effects commonly used in fine mapping methods[38]. Correcting this assumption might substantially improve the accuracy of these methods. Third, our results have important implications for defining and estimating SNP-heritability[2–6;8;10;12–15]. We show that negative causal minor allele effect autocorrelation between linked SNPs may cause true SNP-heritability to be substantially smaller than the heritability defined under these models, although further work is needed to conclusively determine the impact of negative autocorrelation on SNP-heritability estimates. Fourth, negative autocorrelation may contribute to trans-ethnic disparities in polygenic risk prediction[39–41], as SNPs with opposite causal effects that are linked in the training population but not the target population are a source of heritability in the target population that cannot be predicted using the training population.

We note several limitations in our work. First, ACLR is unbiased and well-calibrated when the true autocorrelation is zero, but can produce biased autocorrelation estimates when the true autocorrelation is nonzero, though it detects the correct sign. Second, ACLR's correction for MAF- and LD-dependent genetic architectures is essential. Our correction is based on a MAF- and LD-dependent genetic architecture model that is supported by empirical data[8;33;42;43], but we cannot guarantee that this model is perfectly accurate. Third, we have not partitioned genomic autocorrelations by functional annotation or allele frequency (of one or both SNPs). As functional annotations and allele frequency ranges are differentially impacted by negative selection[8], we hypothesize that genomic autocorrelations are likely to differ, making this an interesting direction for future analyses. Fourth, we recommend that ACLR should be applied using in-sample LD (analogous to recent studies in other settings[27–29]), which may be unavailable for some data sets. Fifth, our analysis focuses on minor allele effects but does not consider derived allele effects. While minor and derived alleles mostly match for SNPs with lower MAF, this is not the case for very common minor alleles. We note that our heritability shrinkage analysis is not affected by this definition, since switching the reference allele changes the sign of both LD and effect concordance. Sixth, our method does not model nonzero directional signed minor allele effects genome-wide, which might be expected for traits under directional selection. However, under our model, such effects would be expected to lead to positive long-range autocorrelation estimates, which we did not observe. Seventh, we did not include very rare SNPs (MAF < 0.1%) in our analysis. Shared tagging of unmodeled, very rare causal SNPs could produce spurious positive genomic autocorrelations, but would not be expected to lead to the negative genomic autocorrelations that we report. Finally, analogous to other studies that employ linear complex trait models, we have not investigated the potential impact of epistatic interactions on our estimates; however, the impact of epistatic interaction on these models is hypothesized to be small[44;45]. We note that epistatic interactions have also been suggested to alter selection effects on nearby genetic variants[23;46] and might hence modulate the effect of linkage masking. Despite these limitations, our work provides an improved understanding of genomic autocorrelations of causal minor allele effect sizes.

Acknowledgements

Methods

Basic approach of the ACLR method

We assume a complex trait model, where genetic variants in the genome have a linear effect on the target trait and the trait is also affected by non-genetic environmental noise:

$$y = X\beta + \epsilon, \epsilon \sim (0, I \times \sigma_\epsilon^2), \beta \sim (0, \Sigma) \tag{2}$$

Here, y is the mean-centered, variance standardized trait vector, X is the column-mean-centered genotype matrix (minor allele counts with column means set to zero), each row representing one study individual and each column a genetic variant locus. ϵ is the environmental effect vector, which we assume to be uncorrelated between individuals. Furthermore, we define β as the vector of causal minor allele effects of genetic variants across the genome. In our model, these effects are random, drawn from an unspecified distribution with mean zero and covariance matrix Σ. Note that we do not constrain off-diagonal elements in Σ to be zero, which allows correlated minor allele effects. Since certain elements in β might be zero or negligibly small, this model does not necessarily assume that all variant loci in the genome have significant effects on the target trait.

We define summary statistics of variant i as $\alpha_i = (X_i^T y)/N$ and in-sample LD matrix $R = (X^T X)/N$, with N representing the number of individuals in the data set. From Equation 2 follows

$$\mathrm{E}(\alpha_i \alpha_j) = R_{.i}^T \Sigma R_{.j} + R_{ij} \sigma_\epsilon^2 / N \tag{3}$$

with $R_{.i}$ being the ith column of R. Intuitively, this equation states that the correlation of summary statistics is a function of both the LD matrix as well as the correlation structure of the causal genetic effects, plus a term related to the environmental noise.

A critical assumption in our model is that the covariance between a pair of genetic effects is a function of the base pair distance between them, i.e. Σ_{ij} is a function of the distance between locus i and locus j. Specifically, we define n distance ranges with constant covariances, with the covariance between effects of variants i and j being σ_k if the distance between the two loci falls into the kth distance range. In this case we can decompose Σ into a sum of distance dependent components:

$$\Sigma = I\sigma_0^2 + S^{(1)}\sigma_1 + ... + S^{(n)}\sigma_n \tag{4}$$

Here, $S^{(k)}$ is the design matrix of distance range k, where $S_{ij}^{(k)}$ is 1 if the distance between variant i and j falls into the kth distance range, and 0 otherwise. Combining Equations 3 and 4, we get

$$\mathrm{E}(\alpha_i \alpha_j) = R_{.i}^T R_{.j} \sigma_0^2 + R_{.i}^T S^{(1)} R_{.j} \sigma_1 + ... + R_{.i}^T S^{(n)} R_{.j} \sigma_n + R_{ij} \sigma_\epsilon^2 / N \tag{5}$$

Distance dependent genetic effect covariances σ_k can then be estimated by regressing the product of summary statistics $\alpha_i \alpha_j$ jointly on distance dependent LD scores $R_{.i}^T S^{(k)} R_{.j}$. If model assumptions hold, this method will produce unbiased estimates of σ_k.

Details of ACLR and implementation

We developed ACLR, a software package implementing the above inference procedure. UK Biobank genotype data is read in from the compressed binary file format BGEN (see URLs) using a fast costumery C++ subroutine. For computational feasibility, in-sample LD matrices are calculated with a banded matrix approach, specifically we assume that LD between variant sites more than 1Mb apart are zero. We developed a Python implementation based on sparse matrix methods to efficiently calculate distance-dependent LD scores as defined above. Summary statistics were either calculated following the above definition of α_i or using the corresponding mixed model association summary statistic calculated using BOLT-LMM[7] to increase statistical power.

We used a weighted regression when estimating effect covariances to further improve power. Following the derivation of Bulik-Sullivan et al.[26], regression weights were based on the variance of the regressand under a normal model, specifically $\mathrm{Var}(\alpha_i\alpha_j) = (R_{.i}^T\Sigma R_{.i} + R_{ii}\sigma_\epsilon^2/N) \times (R_{.j}^T\Sigma R_{.j} + R_{jj}\sigma_\epsilon^2/N) + (R_{.i}^T\Sigma R_{.j} + R_{ij}\sigma_\epsilon^2/N)^2$. Since Σ and σ_ϵ^2 are unknown prior to inference, we assumed a diagonal covariance matrix with diagonal entries and σ_ϵ^2 corresponding to a 50% heritable trait. Note that approximations of the variance of the regressand used to calculate regression weights cannot bias estimates but might only lead to suboptimal statistical power. Correlation results reported in the analysis are defined as $\hat{\rho}_k = \hat{\sigma}_k/\hat{\sigma}_0^2$ (see Equation 5). Uncertainty in $\hat{\sigma}_0^2$ and $\hat{\sigma}_k$ was inferred using a block jackknife approach, dividing the

genome into 200 equally sized blocks of variants (10 jackknife blocks were used in simulations based on chromosome 22 only). The reported uncertainty in $\hat{\rho}_k$ is based on the assumption of $\hat{\sigma}_0^2$ and $\hat{\sigma}_k$ being independent and normally distributed. If model assumptions hold, $\hat{\sigma}_k$ are unbiased estimates of σ_k independent of the set of SNP pair summary statistics used, however they do affect statistical power. Maximum base pair distances between summary statistics pairs were considered in simulations, showing comparable power when only using $\alpha_i\alpha_j$ terms with $i = j$. We therefore only used these terms in the analyses shown in this work after consideration these simulation results and computational constraints, even though ACLR allows using $i \neq j$ cross terms too.

Equation 4 implicitly assumes that all SNPs have the same effect size variance σ_0^2. However, if SNPs with certain MAF and LD patterns have consistently larger or smaller effects, this assumption is broken and our method can yield biased results, consistent with previous work from related methods[8]. We devised a MAF- and LD-corrected version of ACLR that allows for effect size variance differences depending on the MAF and LD of SNPs and can remove this bias. Allowing for these differences assures that the model stays well specified in the presence of MAF- and LD-dependent architectures and hence mitigates possible biases in SNP effect autocovariance estimates. Following a model by Gazal et al.[33], we divided all SNPs into 10 MAF bins for common SNPs (MAF $\geq 5\%$) and 10 MAF bins for low-frequency SNPs (MAF $< 5\%$), for each choosing the MAF thresholds such that an equal number SNPs falls in each one of them. We then allow SNPs in each bin to have a distinct SNP effect variance. Furthermore, we used 6 LD-related functional annotations from Gazal et al.[33], specifically predicted allelic age, recombination rate, level of LD in African populations, nucleotide diversity, background selection statistic, and CpG content. For all these LD-related annotations, apart from allelic age, we allow for different effect coefficients for low-frequency and common SNPs. These annotations linearly affected the SNP effect variances, following previous models[5;8]. Specifically, Equation 4 is replaced by

$$\Sigma = \qquad\qquad V + S^{(1)}\sigma_1 + ... + S^{(n)}\sigma_n \qquad\qquad (6)$$
$$V = \quad D^{(1)}\sigma_{0,1}^2 + ... + D^{(20)}\sigma_{0,20}^2 + F^{(1)}\tau_1 + ... + F^{(11)}\tau_{11} \qquad\qquad (7)$$

where V is the diagonal matrix of SNP effect variances, $\sigma_{0,k}^2$ is the effect variance of the kth MAF bin and $D^{(k)}$ is a diagonal matrix with $D_{ii}^{(k)}$ being 1 if the ith SNP falls into the kth MAF bin and 0 otherwise. $F^{(k)}$ is a diagonal matrix with $F_{ii}^{(k)}$ being the value of the kth LD annotation of SNP i and τ_k the linear effect of the annotation on the SNP effect variance. Equation 5 is then replaced by

$$\mathrm{E}(\alpha_i\alpha_j) = \quad R_{.i}^T D^{(1)} R_{.j}\sigma_{0,1}^2 + ... + R_{.i}^T D^{(20)} R_{.j}\sigma_{0,20}^2 + R_{.i}^T F^{(1)} R_{.j}\tau_1 + ... + R_{.i}^T F^{(11)} R_{.j}\tau_{11} \qquad (8)$$
$$+ R_{.i}^T S^{(1)} R_{.j}\sigma_1 + ... + R_{.i}^T S^{(n)} R_{.j}\sigma_n + R_{ij}\sigma_\epsilon^2/N \qquad (9)$$

Covariances σ_k are calculated by jointly solving this linear regression model for all variance and covariance parameters $\sigma_{0,1}^2, ..., \sigma_{0,20}^2, \tau_1, ..., \tau_{11}, \sigma_1, ..., \sigma_n$ and σ_ϵ^2. Autocorrelation estimates $\hat{\rho}_k = \hat{\sigma}_k/\hat{\sigma}_0^2$ were calculated using the covariance estimates $\hat{\sigma}_k$ from the MAF-LD corrected method, but $\hat{\sigma}_0^2$ from the uncorrected method, since the former does not directly yield an estimate of the overall SNP effect variance and using an appropriate combination $\sigma_{0,i}^2$ and τ_i did not improve the accuracy of the autocorrelation estimate.

UK Biobank genotype data

In our analysis we used genotype data from UK Biobank data set (see URLs). To avoid problems with population stratification, we only used data from $337,502$ individuals with self-reported and confirmed British ancestry that are unrelated (pairwise genetic relatedness $< 5\%$ after LD-pruning). We used genotype information of autosomal variant loci that were imputed using sequencing data from the UK10K project[47] and 1000 Genomes Phase 3 (ref.[36]). These variant sites include both SNPs and small insertions and deletions, but for simplicity we are referring to all of them as SNPs. We only used SNPs with MAF $> 0.1\%$ to avoid using poorly imputed rare variants, resulting in $16,525,661$ variant loci genome

wide. Genotype data was read in from BGEN format files (see URLs), which contains the posterior probabilities after imputation of a given locus being homozygous reference, homozygous alternative and heterozygous, for each individual in the data set. In the analysis performed here, we used minor allele dosage genotypes (i.e. the expected number of minor alleles given these probabilities) as genotype values. Summary statistics were only calculated for a subset of SNPs, specifically SNPs from the HapMap Project Phase 3 (ref.[48]). However, we did not use summary statistics from SNPs close to the human leukocyte antigen genes, specifically SNPs between 25.5 and 33.5Mb into chromosome 6.

Simulations

All simulations used genotypes from 337,502 unrelated UK Biobank individuals with self-reported and confirmed British ancestry (see above section). Due to the large number of simulations, we restricted most simulations to chromosome 22 ($232,046$ UK Biobank SNPs with MAF $\geq 0.1\%$); however, we also performed a selected set of simulations using genome-wide data ($16,525,661$ UK Biobank SNPs with MAF $\geq 0.1\%$; see above section). For null simulations, i.e. simulations with uncorrelated SNP effects, SNP effects β were drawn from independent mean-zero normal distributions. Depending on the specific simulation analysis, the effect variances were either identical for all SNPs, or varied according the inferred variances by Gazal et al.[33], using their MAF and LD related annotations mentioned above. Since the model by Gazal et al.[33] infers effect variances for variance-standardized genotypes, in contrast to the non-standardized genotypes in this analysis, we rescaled MAF bin variances and LD-related variance effects by $2p(1-p)$, where p is the median MAF of the respective MAF range. SNPs that had negative effect variances under this model, were assigned zero trait effects. Although we did not explicitly fix trait polygenicity in simulations, polygenicity is implicit in simulations that use MAF and LD related annotations, as they result in 27% of SNPs having negative effect variances and hence zero trait effects.

For causal simulations, correlated SNP effects were created by dividing all SNPs into blocks of $n = 3$ adjacent SNPs. In each block, SNP effects were drawn from a multivariate normal distribution with pairwise effect correlation being set to a fixed correlation parameter $\rho_{\text{block}} = \pm 0.4$. Effects of SNPs from different blocks were independent. Note that by construction, ρ_{block} does not correspond to the true correlation for any base-pair distance range used in the inference method. Instead, we used the sample correlation of all SNP pairs in a given distance bin as the ground truth to compare our inference results to. When using a MAF and LD-dependent architecture, effect variances were set as described for the null simulations above.

By construction, the above described simulations do not have constant effect covariances per distance range. We also performed causal simulations where we kept the effect covariance constant within each distance range. For simulations without MAF- and LD-dependent (and hence constant) effect variances, SNP base pair positions where altered in simulations such that blocks of $n = 3$ SNPs were within 100bp of each other, while other SNPs were further away, guaranteeing a fixed effect covariance within these blocks and zero covariance between blocks. In simulations with MAF- and LD-dependent variances and constant covariances per distance range, we chose a simpler architecture model than the one by Gazal et al.[33], which makes it possible to have constant covariances per distance range, without effect covariance matrices being not positive semi-definite. Specifically, effect variances were set to 1.0 for low-frequency SNPs ($< 5\%$ MAF) and 0.6 for common SNPs ($> 5\%$ MAF). Also, we chose the effect of the LLD African functional annotation[33] to be $\tau = -0.1$ and zero for all other annotations.

Given a set of genome-wide SNP effects that were simulated, trait values were then simulated using Equation 2. σ_ϵ^2 was chosen such that for a given heritability value, $h^2 = \frac{\text{var}(X\beta)}{\text{var}(X\beta)+\sigma_\epsilon^2}$.

UK Biobank phenotype data

We analyzed 31 heritable traits and diseases from the UK Biobank data set (see URLs). Specifically, for each trait included in the final analysis we required heritability estimates from our method to be at least 6 standard errors away from zero. Furthermore, we removed all traits with missing trait values for more than 10% of the $337,502$ individuals used in this study for the following reason: the distance-dependent LD scores were only calculated once for all traits due to the large computational cost of this

step; since our method assumes that summary statistics and LD scores are computed on the same set of individuals, we only included traits where this is approximately correct. The resulting set of 31 traits and diseases includes 9 quantitative traits, 13 case-control traits, as well as 9 quantitative blood cell traits (see Supplementary Table 2). When calculating summary statistics using the BOLT-LMM software[7], we included sex, age, age squared, the assessment center, the genotyping platform, and 20 top principal components of the genotype matrix as fixed effects to mitigate possible confounding.

Stratifying pairs of SNPs by their pairwise LD

A basic assumption of our method is that the SNP effect covariance only depends on the base pair distance between the respective SNP pair. However, we also performed an analysis that additionally stratified SNP pairs by pairwise LD between them. Specifically, instead of Equation 4, we now assume

$$\Sigma = I\sigma_0^2 + S^{(1,1)}\sigma_{1,1} + S^{(1,2)}\sigma_{1,2} + ... + S^{(n,m)}\sigma_{n,m} \tag{10}$$

where $\sigma_{i,j}$ is the covariance of effects between SNP pairs that fall into the ith distance range and the jth pairwise LD bin, with $S^{(i,j)}$ being defined accordingly. Here, pairwise LD is defined as the Pearson correlation coefficient r_{LD} between the minor allele counts of two SNPs in the study population. We define four pairwise LD bins: negative LD ($-1 \leq r_{\text{LD}} < -0.01$), no LD ($-0.01 \leq r_{\text{LD}} < 0.01$), moderate positive LD ($0.01 \leq r_{\text{LD}} < 0.3$), strong positive LD ($0.3 \leq r_{\text{LD}} \leq 1$). For computational feasibility we reduced the distance bin number to three in these analyses: 1-100bp, 100-1000bp, >1000bp. Hence, $\sigma_{2,3}$ here defines the effect covariance of SNP pairs that are 100-1000bp apart and the correlation between their respective minor alleles is between 0.01 and 0.3.

Estimation of $\sigma_{i,j}$ parameters was performed identically to purely distance dependent covariance parameters and included MAF and LD bias correction, i.e. allowing for MAF and LD dependent SNP effect variances. For computational feasibility, LD between SNP pairs defining $S^{(i,j)}$ were calculated only using a random subset of 1000 individuals.

Heritability shrinkage estimation

In the special case of all SNP effects being uncorrelated and the trait variance standardized to 1, the linear complex trait model described in Equation 2 implies that the trait heritability is equal to the genome-wide sum of variances of standardized genotype effects, a fact that has been used in several previous heritability estimation methods[2–6;8;10;12–15]. We call this quantity h_{uncorr}^2, since it is only equal to the heritability if SNP effects are uncorrelated, but this is not true in the case of significant effect autocorrelation. If linked SNPs have systematic negatively correlated allele effects, these effects will cancel each other out and the heritability will be smaller than the sum of squared standardized SNP effects. Given a general linear complex trait model that allows correlated allele effects, we derive the following extended definition of heritability:

$$h_{\text{full}}^2 := \text{Var}(\beta^T x) = \beta^T \text{Cov}(x,x)\beta = \beta^T R\beta \approx \text{E}(\beta^T R\beta) = \text{E}(\text{tr}(\beta\beta^T R)) = \text{tr}(\Sigma R) \tag{11}$$

Similar to Equation 2, R is the genome-wide population LD matrix which we define as the minor allele count covariance matrix, x is a random genotype (i.e. minor allele count; not variance standardized but mean centered) vector, β is the causal minor allele effect vector, and Σ is the covariance matrix of β. The approximation is accurate as long as the trait is sufficiently polygenic. The equality $\text{tr}(\Sigma R) = \sum_{i,j}(\Sigma \circ R)_{ij}$, the grand sum of the element-wise product, directly shows how positive LD anticorrelated effect SNPs as well as negative LD correlated effect SNPs reduce heritability. This is in contrast to the aforementioned definition used in previous analyses which assume uncorrelated effects, in which case Σ is a diagonal matrix and hence $h_{\text{uncorr}}^2 \approx \sum_i \sigma_i^2 R_{ii} = \sum_i \tilde{\sigma}_i^2$, where σ_i^2 is the variance of the ith per-allele effect and $\tilde{\sigma}_i^2 = \sigma_i^2 R_{ii}$ the variance of the ith allele effect after standardizing the genotype variance.

In order to assess the genome-wide magnitude of the cancelling of linked anticorrelated effects, we defined a new measure called "heritability shrinkage" equal to $s_h = 1 - h_{\text{full}}^2/h_{\text{uncorr}}^2$, the fraction the heritability is reduced due to correlated SNP effects. We use our meta-analyzed UK Biobank trait

10

autocorrelation estimates to calculate a rough estimate of s_h, to get an approximate quantification of the impact of effect autocorrelation genome-wide. To do that, we ignore the MAF- and LD-dependent genetic architecture and assume $\Sigma_{ii} = \sigma_0^2$ for all SNPs in the genome. Following our pairwise LD and distance stratification model of Σ in Equation 10, it follows that

$$h_{\text{full}}^2 = \text{tr}(\Sigma R) = h_{\text{uncorr}}^2 + \text{tr}\left(S^{(1,1)}R\sigma_{1,1}\right) + ... + \text{tr}\left(S^{(n,m)}R\sigma_{n,m}\right) \tag{12}$$

and hence,

$$s_h = -\frac{1}{\text{tr}(R)}\left[\text{tr}\left(S^{(1,1)}R\rho_{1,1}\right) + ... + \text{tr}\left(S^{(n,m)}R\rho_{n,m}\right)\right] \tag{13}$$

In practice, we calculated s_h following the above equation using our pairwise LD and distance stratified autocorrelation estimates in Supplementary Table 9 and a banded LD matrix estimated from chromosome 10. Standard errors are directly based on the uncertainties in $\hat{\rho}_{i,j}$ in Supplementary Table 9.

URLs

Open-source software package implementing ACLR, `https://github.com/arminschoech/ACLR`; UK Biobank website, `http://www.ukbiobank.ac.uk/`; BGEN file format, `http://www.well.ox.ac.uk/~gav/bgen_format/`; UK Biobank genotype imputation manual, `http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/imputation_documentation_May2015.pdf`;

References

[1] International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder. *Nature*, 460(7256):748, 2009.

[2] Jian Yang, Beben Benyamin, Brian P McEvoy, Scott Gordon, Anjali K Henders, Dale R Nyholt, Pamela A Madden, Andrew C Heath, Nicholas G Martin, Grant W Montgomery, et al. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42:565–569, 2010.

[3] S Hong Lee, Teresa R DeCandia, Stephan Ripke, Jian Yang, Patrick F Sullivan, Michael E Goddard, Matthew C Keller, Peter M Visscher, Naomi R Wray, Schizophrenia Psychiatric Genome-Wide Association Study Consortium, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics*, 44:247–250, 2012.

[4] Jian Yang, Andrew Bakshi, Zhihong Zhu, Gibran Hemani, Anna AE Vinkhuyzen, Sang Hong Lee, Matthew R Robinson, John RB Perry, Ilja M Nolte, Jana V van Vliet-Ostaptchouk, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature genetics*, 47(10):1114, 2015.

[5] Hilary K Finucane, Brendan Bulik-Sullivan, Alexander Gusev, Gosia Trynka, Yakir Reshef, Po-Ru Loh, Verneri Anttila, Han Xu, Chongzhi Zang, Kyle Farh, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics*, 47(11):1228, 2015.

[6] Po-Ru Loh, Gaurav Bhatia, Alexander Gusev, Hilary K Finucane, Brendan K Bulik-Sullivan, Samuela J Pollack, Teresa R de Candia, Sang Hong Lee, Naomi R Wray, Kenneth S Kendler, et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance components analysis. *Nature Genetics*, 47:1385–1392, 2015.

[7] Po-Ru Loh, George Tucker, Brendan K Bulik-Sullivan, Bjarni J Vilhjalmsson, Hilary K Finucane, Rany M Salem, Daniel I Chasman, Paul M Ridker, Benjamin M Neale, Bonnie Berger, et al. Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics*, 47(3):284, 2015.

[8] Steven Gazal, Hilary K Finucane, Nicholas A Furlotte, Po-Ru Loh, Pier Francesco Palamara, Xuanyao Liu, Armin Schoech, Brendan Bulik-Sullivan, Benjamin M Neale, Alexander Gusev, et al. Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. *Nature genetics*, 49(10):1421, 2017.

[9] Jian Zeng, Ronald Vlaming, Yang Wu, Matthew R Robinson, Luke R Lloyd-Jones, Loic Yengo, Chloe X Yap, Angli Xue, Julia Sidorenko, Allan F McRae, et al. Signatures of negative selection in the genetic architecture of human complex traits. *Nature genetics*, page 1, 2018.

[10] Yan Zhang, Guanghao Qi, Ju-Hyun Park, and Nilanjan Chatterjee. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nature genetics*, 50(9):1318–1326, 2018.

[11] Xiang Zhu and Matthew Stephens. Large-scale genome-wide enrichment analyses identify new trait-associated genes and pathways across 31 human phenotypes. *Nature communications*, 9(1):1–14, 2018.

[12] Kangcheng Hou, Kathryn S Burch, Arunabha Majumdar, Huwenbo Shi, Nicholas Mancuso, Yue Wu, Sriram Sankararaman, and Bogdan Pasaniuc. Accurate estimation of snp-heritability from biobank-scale data irrespective of genetic architecture. *Nature genetics*, page 1, 2019.

[13] Gleb Kichaev, Gaurav Bhatia, Po-Ru Loh, Steven Gazal, Kathryn Burch, Malika K Freund, Armin Schoech, Bogdan Pasaniuc, and Alkes L Price. Leveraging polygenic functional enrichment to improve gwas power. *The American Journal of Human Genetics*, 104(1):65–75, 2019.

[14] Luke J O'Connor, Armin P Schoech, Farhad Hormozdiari, Steven Gazal, Nick Patterson, and Alkes L Price. Extreme polygenicity of complex traits is explained by negative selection. *The American Journal of Human Genetics*, 105(3):456–476, 2019.

[15] Armin P Schoech, Daniel M Jordan, Po-Ru Loh, Steven Gazal, Luke J O'Connor, Daniel J Balick, Pier F Palamara, Hilary K Finucane, Shamil R Sunyaev, and Alkes L Price. Quantification of frequency-dependent genetic architectures in 25 uk biobank traits reveals action of negative selection. *Nature communications*, 10(1):1–10, 2019.

[16] Luke R Lloyd-Jones, Jian Zeng, Julia Sidorenko, Loïc Yengo, Gerhard Moser, Kathryn E Kemper, Huanwei Wang, Zhili Zheng, Reedik Magi, Tonu Esko, et al. Improved polygenic prediction by bayesian multiple regression on summary statistics. *Nature communications*, 10(1):1–11, 2019.

[17] ENCODE Project Consortium et al. An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57–74, 2012.

[18] Anshul Kundaje, Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Alireza Heravi-Moussavi, Pouya Kheradpour, Zhizhuo Zhang, Jianrong Wang, Michael J Ziller, et al. Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539):317–330, 2015.

[19] Brielin C Brown, Alkes L Price, Nikolaos A Patsopoulos, and Noah Zaitlen. Local joint testing improves power and identifies hidden heritability in association studies. *Genetics*, 203(3):1105–1116, 2016.

[20] Jian Zhou, Chandra L Theesfeld, Kevin Yao, Kathleen M Chen, Aaron K Wong, and Olga G Troyanskaya. Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. *Nature genetics*, 50(8):1171–1179, 2018.

[21] Jesse A Garcia and Kirk E Lohmueller. Negative linkage disequilibrium between amino acid changing variants reveals interference among deleterious mutations in the human genome. *bioRxiv*, 2020.

[22] Richard She and Daniel F Jarosz. Mapping causal variants with single-nucleotide resolution reveals biochemical drivers of phenotypic change. *Cell*, 172(3):478–490, 2018.

[23] Max R Bernstein, Stefan Zdraljevic, Erik C Andersen, and Matthew V Rockman. Tightly linked antagonistic-effect loci underlie polygenic phenotypic variation in c. elegans. *Evolution letters*, 3(5): 462–473, 2019.

[24] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, 2018.

[25] Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J Daly, Alkes L Price, Benjamin M Neale, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, 47(3):291, 2015.

[26] Brendan Bulik-Sullivan, Hilary K Finucane, Verneri Anttila, Alexander Gusev, Felix R Day, Po-Ru Loh, Laramie Duncan, John RB Perry, Nick Patterson, Elise B Robinson, et al. An atlas of genetic correlations across human diseases and traits. *Nature genetics*, 47(11):1236, 2015.

[27] Christian Benner, Aki S Havulinna, Marjo-Riitta Järvelin, Veikko Salomaa, Samuli Ripatti, and Matti Pirinen. Prospects of fine-mapping trait-associated genomic regions by using summary statistics from genome-wide association studies. *The American Journal of Human Genetics*, 101(4): 539–551, 2017.

[28] Jacob C Ulirsch, Caleb A Lareau, Erik L Bao, Leif S Ludwig, Michael H Guo, Christian Benner, Ansuman T Satpathy, Vinay K Kartha, Rany M Salem, Joel N Hirschhorn, et al. Interrogation of human hematopoiesis at single-cell and single-variant resolution. *Nature genetics*, 51(4):683–693, 2019.

[29] Omer Weissbrod, Farhad Hormozdiari, Christian Benner, Ran Cui, Jacob Ulirsch, Steven Gazal, Armin P. Schoech, Bryce van de Geijn, Yakir Reshef, Carla Márquez-Luna, Luke O'Connor, Matti Pirinen, Hilary K. Finucane, and Alkes L. Price. Functionally-informed fine-mapping and polygenic localization of complex trait heritability. *bioRxiv*, 2020. doi: 10.1101/807792. URL https://www.biorxiv.org/content/early/2020/04/02/807792.

[30] Doug Speed, Gibran Hemani, Michael R Johnson, and David J Balding. Improved heritability estimation from genome-wide SNPs. *The American Journal of Human Genetics*, 91:1011–1021, 2012.

[31] Alexander Gusev, Gaurav Bhatia, Noah Zaitlen, Bjarni J. Vilhjalmsson, Dorothée Diogo, Eli A. Stahl, Peter K. Gregersen, Jane Worthington, Lars Klareskog, Soumya Raychaudhuri, Robert M. Plenge, Bogdan Pasaniuc, and Alkes L. Price. Quantifying missing heritability at known gwas loci. *PLOS Genetics*, 9:1–19, 12 2013. doi: 10.1371/journal.pgen.1003993. URL https://doi.org/10.1371/journal.pgen.1003993.

[32] Doug Speed, Na Cai, Michael R Johnson, Sergey Nejentsev, David J Balding, UCLEB Consortium, et al. Reevaluation of SNP heritability in complex human traits. *Nature Genetics*, 49:986–992, 2017.

[33] Steven Gazal, Po-Ru Loh, Hilary K Finucane, Andrea Ganna, Armin Schoech, Shamil Sunyaev, and Alkes L Price. Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations. *Nature genetics*, 50(11):1600–1607, 2018.

[34] Po-Ru Loh, Gleb Kichaev, Steven Gazal, Armin P Schoech, and Alkes L Price. Mixed-model association for biobank-scale datasets. *Nature genetics*, 50(7):906–908, 2018.

[35] Montgomery Slatkin. Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9(6):477–485, 2008.

[36] 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature*, 526:68–74, 2015.

[37] Xuan Zhou, Hae Kyung Im, and S Hong Lee. Core greml for estimating covariance between random effects in linear mixed models for complex trait analyses. *Nature Communications*, 11(1):1–11, 2020.

[38] Daniel J Schaid, Wenan Chen, and Nicholas B Larson. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, 19(8):491–504, 2018.

[39] Carla Márquez-Luna, Po-Ru Loh, South Asian Type 2 Diabetes (SAT2D) Consortium, SIGMA Type 2 Diabetes Consortium, and Alkes L Price. Multiethnic polygenic risk scores improve risk prediction in diverse populations. *Genetic epidemiology*, 41(8):811–823, 2017.

[40] Alicia R Martin, Masahiro Kanai, Yoichiro Kamatani, Yukinori Okada, Benjamin M Neale, and Mark J Daly. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature genetics*, 51(4):584–591, 2019.

[41] L Duncan, H Shen, B Gelaye, J Meijsen, K Ressler, M Feldman, R Peterson, and B Domingue. Analysis of polygenic risk score usage and performance in diverse human populations. *Nature communications*, 10(1):1–9, 2019.

[42] Steven Gazal, Carla Marquez-Luna, Hilary K Finucane, and Alkes L Price. Reconciling s-ldsc and ldak functional enrichment estimates. *Nature genetics*, 51(8):1202–1204, 2019.

[43] Doug Speed, John Holmes, and David J Balding. Evaluating and improving heritability models using summary statistics. *Nature Genetics*, 52(4):458–462, 2020.

[44] William G Hill, Michael E Goddard, and Peter M Visscher. Data and theory point to mainly additive genetic variance for complex traits. *PLOS Genetics*, 4:e1000008, 2008.

[45] Asko Mäki-Tanila and William G Hill. Influence of gene interaction on complex trait variation with multilocus models. *Genetics*, 198(1):355–367, 2014.

[46] Tuuli Lappalainen, Stephen B Montgomery, Alexandra C Nica, and Emmanouil T Dermitzakis. Epistatic selection between coding and regulatory variation in human evolution and disease. *The American Journal of Human Genetics*, 89(3):459–463, 2011.

[47] UK10K Consortium et al. The uk10k project identifies rare variants in health and disease. *Nature*, 526:82–90, 2015.

[48] International HapMap 3 Consortium et al. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311):52, 2010.

Tables

Table 1: ACLR results in null simulations with zero autocorrelation

MAF-LD corrected	MAF-LD architect	h^2	$\hat{\sigma}_{1-100bp}$	$\hat{\rho}_{1-100bp}$	$\hat{\sigma}_{100-1000bp}$	$\hat{\rho}_{100-1000bp}$
yes	yes	0.5	-0.008 ± 0.008 (0.008)	-0.004 ± 0.010 (0.011)	0.002 ± 0.001 (0.001)	0.002 ± 0.002 (0.002)
yes	no	0.5	-0.002 ± 0.011 (0.011)	0.010 ± 0.011 (0.011)	-0.001 ± 0.002 (0.002)	-0.001 ± 0.002 (0.002)
yes	yes	0.2	-0.007 ± 0.008 (0.008)	-0.003 ± 0.010 (0.010)	0.002 ± 0.001 (0.001)	0.002 ± 0.002 (0.002)
no	no	0.5	-0.003 ± 0.011 (0.011)	0.013 ± 0.011 (0.011)	-0.005 ± 0.002 (0.002)	-0.004 ± 0.002 (0.002)
no	yes	0.5	-0.182 ± 0.008 (0.009)	-0.212 ± 0.010 (0.011)	-0.052 ± 0.001 (0.001)	-0.063 ± 0.002 (0.002)

For each genetic architecture simulated, we report autocovariance estimates ($\hat{\sigma}$) and autocorrelation estimates ($\hat{\rho}$) for 1-100bp and 100-1000bp genomic distance ranges, and their empirical standard errors. We also report mean jackknife standard error estimates in parentheses. Autocovariance estimates ($\hat{\sigma}$) are scaled based on a true variance of 1 for ease of comparison; true values of σ and ρ are equal to 0. "MAF-LD corrected" denotes MAF- and LD-corrected estimates. "MAF-LD architect" denotes MAF- and LD-dependent generative architectures. h^2 denotes the heritability simulated. Results are averaged across 1000 simulations in each case.

Table 2: ACLR results in causal simulations with nonzero autocorrelation

MAF-LD architect	h^2	Const covar	$\sigma_{1-100bp}$	$\hat{\sigma}_{1-100bp}$	$\hat{\rho}_{1-100bp}$	$\sigma_{100-1000bp}$	$\hat{\sigma}_{100-1000bp}$	$\hat{\rho}_{100-1000bp}$
no	0.5	yes	0.200	0.203 ± 0.016 (0.017)	0.215 ± 0.016 (0.017)	0.000	-0.000 ± 0.003 (0.003)	0.003 ± 0.003 (0.003)
no	0.5	yes	-0.200	-0.206 ± 0.014 (0.015)	-0.197 ± 0.014 (0.014)	0.000	-0.002 ± 0.003 (0.003)	0.000 ± 0.003 (0.003)
yes	0.5	yes	0.230	0.220 ± 0.014 (0.014)	0.297 ± 0.018 (0.018)	0.000	-0.005 ± 0.003 (0.003)	-0.002 ± 0.003 (0.003)
yes	0.5	yes	-0.250	-0.256 ± 0.009 (0.009)	-0.356 ± 0.013 (0.014)	0.000	-0.002 ± 0.002 (0.002)	-0.000 ± 0.003 (0.003)
yes	0.2	yes	0.230	0.219 ± 0.014 (0.015)	0.296 ± 0.018 (0.018)	0.000	-0.005 ± 0.003 (0.003)	-0.002 ± 0.003 (0.003)
yes	0.2	yes	-0.250	-0.255 ± 0.009 (0.009)	-0.354 ± 0.013 (0.014)	0.000	-0.002 ± 0.002 (0.002)	-0.000 ± 0.003 (0.003)
yes	0.5	no	0.149	0.034 ± 0.009 (0.009)	0.051 ± 0.010 (0.010))	0.033	0.009 ± 0.001 (0.001)	0.010 ± 0.002 (0.002)
yes	0.5	no	-0.149	-0.048 ± 0.007 (0.007)	-0.056 ± 0.009 (0.009)	-0.033	-0.007 ± 0.001 (0.001)	-0.009 ± 0.001 (0.001)

For each genetic architecture simulated, we report the true autocovariances (σ), autocovariance estimates ($\hat{\sigma}$) and autocorrelation estimates ($\hat{\rho}$) for 1-100bp and 100-1000bp genomic distance ranges, and their empirical standard errors. We also report mean jackknife standard error estimates in parentheses. True autocovariances (σ) and their estimates ($\hat{\sigma}$) are scaled based on a true variance of 1 for ease of comparison; thus, the true ρ is equal to the true σ. All estimates were MAF- and LD-corrected. "MAF-LD architect" denotes MAF- and LD-dependent generative architectures. h^2 denotes the heritability simulated. "Const covar" denotes genetic architectures in which true autocovariances (σ) are constant for SNP pairs within each distance range. Results are averaged across 1000 simulations in each case.
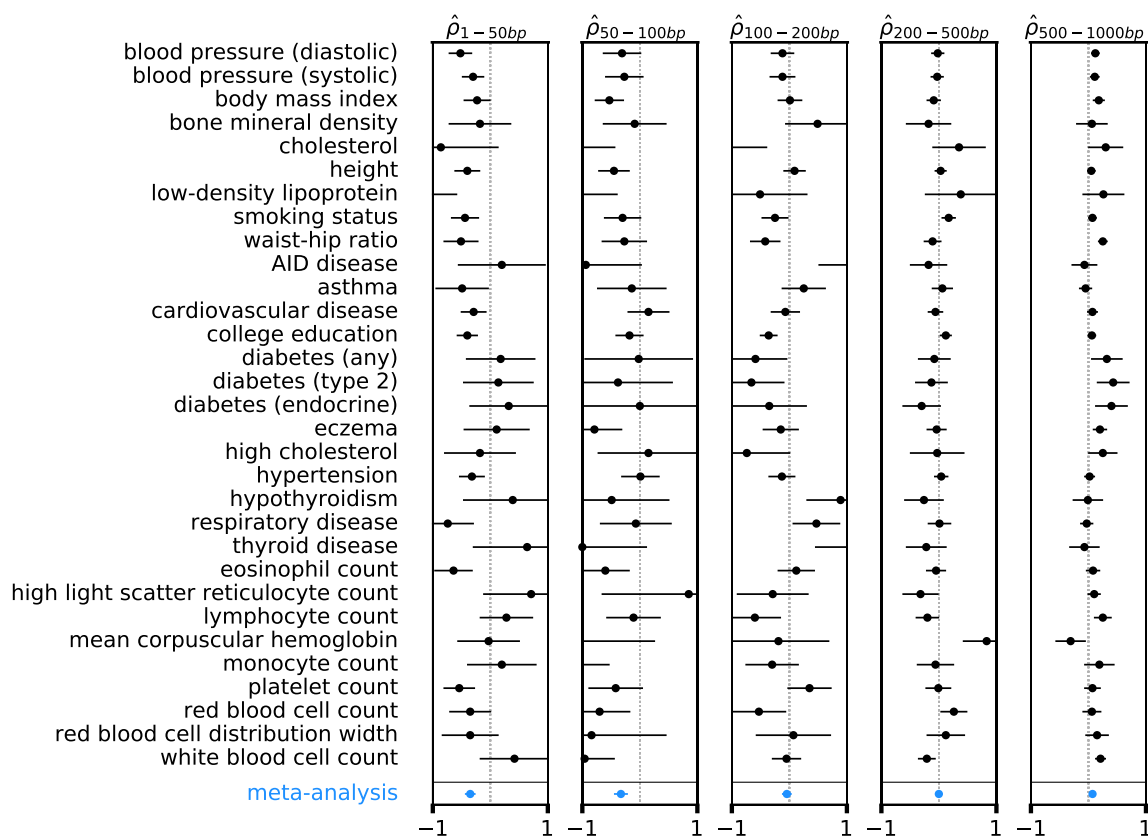
Figures

Figure 1: ACLR results for 31 UK Biobank traits. We report autocorrelation estimates ($\hat{\rho}$) across 5 different genomic distance ranges for each of 31 UK Biobank traits, including 9 quantitative traits, 13 case-control traits, as well as 9 quantitative blood count traits. We also show meta-analysis results across a subset of 25 independent traits in the last row. Error bars denote block-jackknife standard errors. Numerical results and autocovariance estimates are reported in Supplementary Tables 3 and 4 respectively.