

# Benchmarking local genetic correlation estimation methods using summary statistics from genome-wide association studies

Chi Zhang<sup>1</sup>, Yiliang Zhang<sup>1</sup>, Yunxuan Zhang<sup>1</sup> and Hongyu  
Zhao<sup>1,2\*</sup>

<sup>1</sup>Department of Biostatistics, Yale School of Public Health, New  
Haven, CT, United States.

<sup>2</sup>Program of Computational Biology and Bioinformatics, Yale  
School of Public Health, New Haven, CT, United States.

\*Corresponding author(s). E-mail(s): [hongyu.zhao@yale.e.com](mailto:hongyu.zhao@yale.e.com);  
Contributing authors: [chi.zhang.cz354@yale.edu](mailto:chi.zhang.cz354@yale.edu);  
[yiliang.zhang@yale.edu](mailto:yiliang.zhang@yale.edu); [yunxuan.zhang@yale.edu](mailto:yunxuan.zhang@yale.edu);

## Abstract

Local genetic correlation evaluates the correlation of genetic effects between different traits across genetic variants in a local region. It has been proven informative for understanding the genetic similarities of complex traits beyond that captured by global genetic correlation calculated across the whole genome. Several summary-statistics-based approaches have been developed for estimating local genetic correlation, including  $\rho$ -hess, SUPERGNOVA, and LAVA. However, there has not been a comprehensive evaluation of these methods to offer practical guidelines on the choices of these methods. In this study, we conduct benchmark comparisons of the performance of these three methods through extensive simulation and real data analyses. We focus on two technical difficulties in estimating local genetic correlation: sample overlaps across traits and local linkage disequilibrium (LD) estimates when only the external reference panels are available. Our simulations suggest that the type-I error and estimation accuracy are highly dependent on the estimation of the local LD matrix. These observations

are corroborated by real data analyses of 31 complex traits. Overall, our results offer insights into post-GWAS local correlation studies and highlight issues that demand future methodology developments.

**Keywords:** benchmark, complex traits, GWAS, local genetic correlation, linkage disequilibrium, sample overlap

## 1 Introduction

In recent years, genome-wide association analyses (GWAS) have identified tens of thousands of genetic variants associated with numerous complex traits and diseases[1–4]. Various post-GWAS approaches[5], such as fine mapping, genetic correlation, functional enrichment, and polygenic risk score (PRS), are routinely conducted to gain a further understanding of the genetic variants and biological mechanisms behind the observed statistical associations. Estimating genetic correlation using GWAS data is an essential part of the post-GWAS analysis that can quantify genetic similarities and uncover shared genetic basis of complex traits and disorders. Additionally, genetic correlation results can help increase statistical power in genetic association studies [6, 7], and improve polygenic risk score prediction accuracy [8–12]. Genetic correlations can be characterized globally, summarizing the average correlation of genetic effects across the genome, or locally highlighting specific regions having correlated effects on different traits.

While global genetic correlations have been extensively studied [13] with many methods developed for their estimations using GWAS data [14–17], they may not capture local genetic correlations that may be distinct across different regions [18–22]. These include the existence of opposing correlations in different regions, which can lead to a non-significant global genetic correlation. Furthermore, global genetic correlations provide limited insight into the

shared biological mechanisms, when different genomic regions have different correlation levels. For example, when investigating shared genetic architecture between COVID-19 severity and idiopathic pulmonary fibrosis (IPF), the global genetic correlation between these two diseases was 0.35 ( $p = 0.001$ ), however, the effect of MUC5B and ATP11A revealed opposing effects for these two diseases[23]. In order to capture local correlation patterns, several methods have been developed for estimating or detecting local genetic correlation including  $\rho$ -HESS[24], SUPERGNOVA[20], and LAVA[21].

$\rho$ -hess[19] and SUPERGNOVA[20] focus on evaluating bivariate local genetic correlations, whereas LAVA[21] used partial correlation and multiple regression to estimate bivariate and multivariate genetic correlations. These methods also differ in model assumptions relating genetic variants to their effects on traits. Whereas  $\rho$ -hess[19] and LAVA[21] are based on fixed effects models, SUPERGNOVA[20] is based on a random effects model. Although there was earlier research evaluating the performance of these approaches[20, 21], there has not been a comprehensive study of their performance through extensive simulation and real data analyses. Given the importance of inferring local genetic correlations, there is a need for objectively benchmarking the performance of these three methods with user-defined genome partitions in realistic settings. As all three methods both infer shared genetic effects and estimate local heritability in a local region, we evaluate their performances for both tasks, i.e. local genetic covariance/correlation estimation between two traits and local heritability estimation.

We conducted simulations using the observed genotype data from the UK Biobank (UKB)[25] and compared different methods using both in-sample and external reference panels to estimate local linkage disequilibrium (LD) structure. We used genotype data from 1KG Phase 3[26] as the external reference

panel. We considered binary and continuous traits with varying sample overlaps and region sizes. We assessed the robustness of each method against both infinitesimal and non-infinitesimal models, and whether the effect sizes follow the underlying assumption of the random effect method, SUPERGNOVA. Additionally, we investigated the stability of  $\rho$ -hess and LAVA with different reference panels. After simulations, we applied these methods to analyze 31 complex traits with publicly available GWAS summary statistics. To validate the accuracy of these methods in real data, we applied LDSC[14] to estimate global genetic covariances and heritability and compared these estimates with the sum of local heritability and local genetic covariance. For these real data, we also assessed the stability of the point estimates and inferences using different reference panels and conducted polygenic risk score analyses using markers from regions with significant positive and negative correlations. The observations from our simulation and real data analyses offer valuable insights into the statistical properties, advantages, and limitations of each method.

## 2 Methods

### 2.1 Study population and quality control of genotype data

The UKB is a large, prospective study that aims to examine complex traits and diseases in middle-aged adults. We performed simulations using imputed genotype data from UKB and selected samples from genetically unrelated participants of White British ancestry ( $n=276,731$ ). For real data analysis, we used phenotype and genotype data from UKB to perform polygenic risk score (PRS) analysis for four traits: coronary artery disease (CAD), type 2 diabetes (T2D), low-density lipoprotein (LDL), and body mass index (BMI). Of the participants included in the analysis, 4,765 individuals were diagnosed with

CAD, and 40,361 individuals were diagnosed with T2D. The mean LDL level was 3.37 (*mmol/L*) with a standard deviation of 1 and the mean BMI was 31.23 (*kg/m<sup>2</sup>*) with a standard deviation of 5.8.

For real data analysis, in addition to the UKB dataset, we also used data from SPARK (Simons Foundation Powering Autism Research)[27] for autism spectrum disorder (ASD) patients. We accessed the first release of the combined multi-batch SPARK WES dataset, which includes phenotype data for the SPARK Collection Version 7. The details of these samples are available on the SFARI website, <https://www.sfari.org/resource/spark/>. This dataset includes 69,592 samples processed on the Illumina Global Screening Array and is provided in PLINK[28] format. After removing samples with estimated ancestry other than European (EUR) and missing genotype data, 51,658 samples remained for further analysis. We applied pre-imputation quality control using PLINK[28], including the removal of SNPs with low genotype call rates ( $<0.95$ ), minor allele frequencies ( $<0.01$ ), or deviations from Hardy-Weinberg equilibrium ( $<1e-06$ ), as well as samples with high missing genotype rates ( $>0.05$ ). This left us with 455,444 SNPs and 43,891 samples. The genotype data were then phased and imputed to the HRC reference panel using the Michigan Imputation Server[29]. After imputation, we applied additional quality control, including the removal of SNPs with low imputation quality ( $<0.8$ ) or minor allele frequency ( $<0.01$ ). Finally, the SPARK study data contained 7,194,844 SNPs on the GRCh37/hg19 build, of which 5948,083 SNPs were also included in the EUR 1KG Phase 3 data. We then retained 12,264 individuals who were ASD probands and also had intelligence quotient (IQ) scores to assess the association between PRSs and IQ scores in ASD probands. In our analysis, we quantified cognitive performance using full-scale IQ, verbal IQ, and

non-verbal IQ. There were 1,026, 785, and 830 ASD probands in SPARK, who had both these IQ scores and qualify-controlled genotype data, respectively.

## 2.2 Genome partition

Both  $\rho$ -hess and SUPERGNOVA use LDetect[30] to partition the genome into non-overlapping blocks with an average width of 1.6 cM per block. However, LDetect is a heuristic method that may not always produce optimal results. In contrast, LAVA divides the genome by recursively splitting the largest block into two smaller blocks, selecting a new breakpoint that minimizes local LD between the resulting blocks. To compare the performance of different partitions fairly, we used `snp_ldsplit`[31] to partition the genome, which uses dynamic programming to minimize the sum of squared correlations between variants in different blocks. We compared the performance of different genome partitions and found that the partitions generated from `snp_ldsplit` led to smaller sum of squared correlations between SNPs in different blocks (Supplementary Material, Appendix A).

We used the 1KG Phase 3[26] reference panels for our analysis. We selected the European samples by the SuperPopulation information provided by 1KG and then excluded all duplicated and ambiguous SNPs. We applied quality control to the 1KG data for EUR ancestry using PLINK[28] (`-geno 0.05 -hwe 1e-10 -mind 0.05 -maf 0.05`) and generated a genetic map using the [https://plink.readthedocs.io/en/latest/plink\\_mani/](https://plink.readthedocs.io/en/latest/plink_mani/) website. In addition, we conducted quality control on the UKB data, creating two UKB reference panels with 503 randomly selected non-overlapping samples from the unrelated White British individuals (the same sample size as the EUR 1KG Phase 3 reference panel).

To partition the genome, we excluded the MHC block on chromosome 6 (30-31Mb) and applied `snp_ldsplit` to each chromosome in parallel. To avoid

LD leakage and biased estimates, we set the minimum size of each block to be at least 0.5 cM. We adaptively searched for the optimal values of `max_r2` (the maximum squared correlation allowed for one pair of variants in two different blocks) and `max_size` (the maximum number of variants in each block) for each chromosome to make the LD blocks as independent as possible. It is important to note that a larger block may result in increased computational and memory requirements, and obscure local signals. To find the minimal combination of `max_r2` and `max_size` that can generate partitions with a mean block size smaller than 1.6cM, we searched for values of `max_r2` from 0.3 to 0.72 and `max_size` from 5cM to 13cM. From the partitions found, we selected the partitions that resulted in the minimal cost (the sum of squared correlations between SNPs at different blocks). The final `max_r2` and `max_size` values for each chromosome are shown in Supplementary Table 1 and the partitions used in this analysis are in Supplementary Table 2. (Note that genomic coordinates for this paper are in reference to the human genome build 37.)

## 2.3 Methods for genetic correlation estimation

We compared the performance of three local genetic correlation estimation methods:  $\rho$ -hess, SUPERGNOVA, and LAVA. These three approaches are based on the analysis of summary statistics, with  $\rho$ -hess and LAVA using fixed effects models and SUPERGNOVA adopting a random effects model. In the following, we first briefly introduce the concept of local genetic covariance and then describe the underlying statistical framework for these methods.

Let  $X_i$  denote the standardized genotype vector of size  $m_i$  in block  $i$ , where  $m_i$  is the number of markers in this block,  $i = 1, \dots, I$ , and  $\beta_i$  and  $\gamma_i$  are the effect size vectors of the  $m_i$  markers within block  $i$  for two traits, then the local genetic contributions for the two traits in block  $i$  are  $g_{1i} = X_i^T \beta_i$

and  $g_{2i} = X_i^T \gamma_i$ , respectively. Local genetic covariance is defined as  $\rho_i = Cov(g_{1i}, g_{2i})$ . The local genetic correlation  $r_i$  can be estimated as  $r_i = \frac{\rho_i}{\sqrt{h_{1i}^2 h_{2i}^2}}$ , where  $h_{1i}^2$  and  $h_{2i}^2$  are the local heritability in block  $i$ . Furthermore, the global genetic covariance matches the sum of local genetic covariance when the genetic components in different partitions are independent:

$$\begin{aligned}
Cov(g_1, g_2) &= Cov\left[\sum_{i=1}^I X_i^T \beta_i, \sum_{i=1}^I X_i^T \gamma_i\right] \\
&= \sum_{i=1}^I \sum_{j=1}^I Cov[X_i^T \beta_i, X_j^T \gamma_j] \\
&= \sum_{i=1}^I Cov[X_i^T \beta_i, X_i^T \gamma_i] \\
&= \sum_{i=1}^I \rho_i.
\end{aligned} \tag{1}$$

where  $g_1$  and  $g_2$  are the genetic contributions for traits 1 and 2, respectively.

In the outputs of these three methods,  $\rho$ -hess and SUPERGNOVA give estimates of local genetic covariance, whereas LAVA provides estimates of local genetic correlation. All three methods yield estimates of local heritability with p-values provided by  $\rho$ -hess and LAVA, so we can obtain both local genetic correlation and covariance for all three methods.

### 2.3.1 $\rho$ -hess

Based on the definition of local genetic covariance introduced above,  $\rho$ -hess defines local genetic covariance in block  $i$  as  $\rho_i = Cov(g_{1i}, g_{2i}) = Cov(X_i^T \beta_i, X_i^T \gamma_i) = \beta_i^T V_i \gamma_i$ , where  $V_i$  is the local LD matrix in this block, and  $\beta_i$  and  $\gamma_i$  are the fixed effect size vectors for two traits in block  $i$ .

When there are two GWASs, with  $n_1$  samples for trait 1 ( $\phi_1$ ) and  $n_2$  samples for trait 2 ( $\phi_2$ ),  $\rho$ -hess assumes that  $\phi_1 = \sum_{i=1}^I Y_i \beta_i + \epsilon$  and  $\phi_2 =$



$\sum_{i=1}^I Z_i \gamma_i + \delta$ , where  $\phi_1$  is the vector of trait 1 for  $n_1$  samples and  $\phi_2$  is the vector of trait 2 for  $n_2$  samples,  $Y_i$  and  $Z_i$  are the standardized genotypes in block  $i$  for  $n_1$  and  $n_2$  individuals, respectively, and  $\epsilon$  and  $\delta$  are the vectors of noises with  $\text{var}(\epsilon) = \sigma_\epsilon^2$  and  $\text{var}(\delta) = \sigma_\delta^2$ . Assume the first  $n_s$  samples overlap, then

$$\text{cov}(\epsilon_{j_1}, \delta_{j_2}) = \begin{cases} \rho_\epsilon, 1 \leq j_1 = j_2 \leq n_s \\ 0, \text{otherwise} \end{cases} \quad (2)$$

and  $\epsilon_{j_1}$  and  $\delta_{j_2}$  are the noises for individual  $j_1$  in trait 1 and for individual  $j_2$  in trait 2, respectively.

The marginal effect size estimates of SNPs in block  $i$  from GWAS,  $\hat{\beta}_i$  and  $\hat{\gamma}_i$ , follow the normal distribution  $\hat{\beta}_i \sim N(V_i \beta_i, \frac{\sigma_\epsilon^2}{n_1} V_i)$ ,  $\hat{\gamma}_i \sim N(V_i \gamma_i, \frac{\sigma_\delta^2}{n_2} V_i)$ , so in the absence of sample overlap,  $\rho$ -hess estimates the local genetic covariance in block  $i$  by

$$\rho_i = \hat{\beta}_i^T V_i^{-1} \hat{\gamma}_i. \quad (3)$$

However, due to sample overlap, the estimation based on (3) using  $\hat{\beta}_i$  and  $\hat{\gamma}_i$  will have a bias term. Besides, in practice  $\rho$ -hess uses truncated SVD to address rank-deficiency of LD matrix  $V_i$  to improve stability, especially when only the external reference panel is available. By defining  $\hat{s}_{ij} = \frac{1}{\hat{w}_{ij}} (\hat{\beta}_{ij}^T \hat{u}_{ij}) (\hat{\gamma}_{ij}^T \hat{u}_{ij})$ , where  $\hat{w}_{ij}$  and  $\hat{u}_{ij}$  are the  $j$ th top eigenvalue and its corresponding eigenvector of local LD matrix in block  $i$  from the external reference panel, and  $g(\hat{\beta}_i, \hat{\gamma}_i, k) = \sum_{j=1}^k \hat{s}_{ij}$ ,  $\rho$ -hess estimates local genetic covariance after correcting for the bias by

$$\hat{\rho}_i = \frac{n_1 n_2 g(\hat{\beta}_i, \hat{\gamma}_i, k) - n_s \rho_\epsilon k}{n_1 n_2 - n_s k}, \quad (4)$$

where  $k$  is the number of top eigenvalues and their corresponding eigenvectors used which can be input by the user and is the same for different blocks.

For testing significance,  $\rho$ -hess assumes that the sampling distributions of the local genetic correlation and covariance are normal, and uses a parametric bootstrap approach to estimating the standard errors.

In our real data analysis, we followed  $\rho$ -hess’s suggestion [24] to estimate  $\lambda_{GC}$  from GWAS data. We then used the estimated  $\sqrt{\lambda_{GC}}$  to re-inflate the effect sizes before estimating the local SNP heritability and genetic correlation. To improve the accuracy of  $\rho$ -hess, we used the global heritability and its standard error from LDSC as extra inputs. The number of shared samples used in our analysis was based on the consortium from which each GWAS was generated. As indicated in Supplementary Table 4, when two traits have samples from the same consortium, we fixed the shared sample size to the minimum sample size of the common consortium. We set the shared sample size to zero when two traits came from completely different consortia. All other parameters were kept at their default values.

### 2.3.2 SUPERGNOVA

SUPERGNOVA also assumes traits follow the same linear models shown in  $\rho$ -hess. However, SUPERGNOVA models genetic effects  $\beta_i$  and  $\gamma_i$  as random rather than fixed. More specifically,  $\beta_i$  and  $\gamma_i$  in block  $i$  follow a multivariate normal distribution:

$$\begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{h_{1i}^2}{m_i} I_{m_i} & \frac{\rho_i}{m_i} I_{m_i} \\ \frac{\rho_i}{m_i} I_{m_i} & \frac{h_{2i}^2}{m_i} I_{m_i} \end{bmatrix} \right) \quad (5)$$

where  $h_{1i}^2$  and  $h_{2i}^2$  are the local heritability of traits 1 and 2 in block  $i$ ;  $\rho_i$  is the local genetic covariance between traits 1 and 2;  $I_{m_i}$  is the identity matrix of size  $m_i$ ; and  $m_i$  is the number of SNPs in block  $i$  as defined before.

The estimator used in SUPERGNOVA is defined in terms of the marginal z-statistics of a single SNP  $j$  in block  $i$ , which is given by  $z_{1ij} = \frac{\hat{\beta}_{ij}}{se(\hat{\beta}_{ij})}$  and  $z_{2ij} = \frac{\hat{\gamma}_{ij}}{se(\hat{\gamma}_{ij})}$ , where  $\hat{\beta}_{ij}$  and  $\hat{\gamma}_{ij}$  are the marginal effect sizes from GWAS. SUPERGNOVA performs eigen decomposition of the local LD matrix ( $V_i = U_i \Sigma_i U_i^T$ ) and chooses the first  $K_i$  eigenvectors to transform and decorrelate association statistics in a given block  $i$ , where  $K_i$  is determined adaptively. After decorrelation, local genetic covariance  $\rho_i$  is estimated by modeling the expected value of the products of the projected z-statistics,

$$E[\tilde{z}_{1ij} \tilde{z}_{2ij}] = \frac{\sqrt{n_1 n_2}}{m_i} \rho_i w_{ij}^2 + \frac{n_s \rho_t}{\sqrt{n_1 n_2}} w_{ij}, \quad (6)$$

where  $w_{ij}$  is the  $j$ th eigenvalue, where  $1 \leq j \leq m_i$ ,  $\rho_t$  is the sum of genetic covariances and non-genetic covariance, i.e.,  $\rho_t = \sum_{i=1}^I \rho_i + \rho_e$ .  $n_1$ ,  $n_2$ , and  $n_s$  are the sample sizes for each trait and the sample size shared by two GWASs, respectively, as defined above. Besides,  $\frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}}$  is the estimation of  $\frac{n_s \rho_t}{\sqrt{n_1 n_2}}$  using the intercept of cross-trait LDSC[32]. Then a weighted least squares regression is used to regress  $\eta_{ij} = \tilde{z}_{1ij} \tilde{z}_{2ij} - \frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}} w_{ij}$  on predictor  $w_{ij}^2$  with the weight as the reciprocal of  $q_{ij}^2 = (\frac{n_1 h_1^2}{m_i} w_{ij}^2 + w_{ij})(\frac{n_2 h_2^2}{m_i} w_{ij}^2 + w_{ij})$ , where  $h_1^2$  and  $h_2^2$  are the global heritability for traits 1 and 2, respectively.

SUPERGNOVA adopts an adaptive procedure to determine the number of eigenvalues/eigenvectors for each block. This is accomplished by choosing  $K_i$  which minimizes the maximum between the theoretical variance and the empirical variance of local genetic covariance:

$$Var[\hat{\rho}_i \parallel \frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}}] = (\frac{m_i^2}{n_1 n_2}) / \sum_{j=1}^{K_i} \frac{w_{ij}^4}{q_{ij}^2}, \quad (7)$$

$$\begin{aligned} \text{Var}[\hat{\rho}_i \parallel \frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}}] = & [(\frac{m_i^2}{n_1 n_2}) / \sum_{j=1}^{K_i} \frac{w_{ij}^4}{q_{ij}^2}] \\ & [\sum_{j=1}^{K_i} \frac{\eta_{ij}^2}{q_{ij}^2} - \frac{(\sum_{j=1}^{K_i} \eta_{ij} w_{ij}^2 / q_{ij}^2)^2}{\sum_{j=1}^{K_i} w_{ij}^4 / q_{ij}^2}] / (K_i - 1). \end{aligned} \quad (8)$$

Finally, the variance of local genetic covariance has the following form:

$$\text{Var}(\hat{\rho}_i) = \text{Var}(E[\hat{\rho}_i \parallel \frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}}]) + E[\text{Var}(\hat{\rho}_i \parallel \frac{n_s \hat{\rho}_t}{\sqrt{n_1 n_2}})]. \quad (9)$$

### 2.3.3 LAVA

Same as  $\rho$ -hess, LAVA also assumes that the genetic effect sizes are fixed and denotes local genetic covariance in block  $i$  as  $\rho_i = \beta_i^T V_i \gamma_i$ . LAVA first applies singular value decomposition to the local LD matrix in block  $i$  which is  $V_i = U_i \Lambda_i \Lambda_i^T$ , and then defines  $U_{i*}$  as the  $m_i$  by  $k_i$  pruned eigenvector matrix and  $\Lambda_{i*}$  as the corresponding  $k_i$  by  $k_i$  diagonal singular value matrix, where  $m_i$  is the number of SNPs in block  $i$  and  $k_i$  is the number of top eigenvalues that could explain 99% variance of the local LD matrix in block  $i$ . Thus, the inverse of  $V_i$  can be approximated as  $U_{i*} (\Lambda_{i*} \Lambda_{i*})^{-1} U_{i*}^T$ .

Furthermore, LAVA defines the scaled principal component (PC) matrix  $W_i = X_i^T U_{i*} \Lambda_{i*}^{-1}$  and the corresponding PC effects  $\zeta_{1i} = \Lambda_{i*} U_{i*}^T \beta_i$  and  $\zeta_{2i} = \Lambda_{i*} U_{i*}^T \gamma_i$ , such that  $W_i \zeta_{1i} \approx X_i^T \beta_i$  and  $W_i \zeta_{2i} \approx X_i^T \gamma_i$ . Thus, the local genetic covariance can be represented by the covariance of the PC effects:

$$\rho_i = \beta_i^T V_i \gamma_i \approx \beta_i^T U_{i*} \Lambda_{i*} \Lambda_{i*} U_{i*}^T \gamma_i = \zeta_{1i}^T \zeta_{2i}. \quad (10)$$

Assume that  $\hat{\beta}_i$  and  $\hat{\gamma}_i$  are the vectors of the marginal effects of two traits in block  $i$ , then based on the distribution of marginal effect sizes, PC effects can be estimated as  $\hat{\zeta}_{1i} = \Lambda_{i*}^{-1} U_{i*}^T \hat{\beta}_i$  and  $\hat{\zeta}_{2i} = \Lambda_{i*}^{-1} U_{i*}^T \hat{\gamma}_i$ , and  $\hat{\zeta}_i = [\hat{\zeta}_{1i}, \hat{\zeta}_{2i}]$  follows the distribution  $MVN(\zeta_i, \hat{\Sigma})$ , where  $\zeta_i = [\zeta_{1i}, \zeta_{2i}]$  and  $\hat{\Sigma}$  represents the sampling covariance matrix.

The method of moments can be used to estimate:

$$E[\hat{\zeta}_i^T \hat{\zeta}_i] = \zeta_i^T \zeta_i + k_i \hat{\Sigma}. \quad (11)$$

In the absence of sample overlap,  $\hat{\Sigma}$  is defined as the diagonal matrix with diagonal elements as the sampling variances of trait 1 and trait 2. When accounting for sample overlap, LAVA first applies LDSC to create a covariance matrix with the intercepts for the global genetic covariance for the off-diagonal elements and each trait’s univariate LDSC intercept as the diagonal elements. Then LAVA converts this covariance matrix to a correlation matrix,  $C$ , and computes the sampling correlation matrix as  $\hat{\Sigma} = \text{diag}(\hat{\sigma}) \times C \times \text{diag}(\hat{\sigma})$ , where  $\hat{\sigma}^2$  is a vector with the sampling variances of the traits.

Once estimated, the significance of  $\rho_i$  is evaluated using simulation-based P values. Based on the definition of non-central Wishart distribution and  $\hat{\zeta}_i$  follows  $MVN(\zeta_i, \hat{\Sigma}_i)$ , the statistic  $\hat{\zeta}_i^T \hat{\zeta}_i$  has a non-central Wishart distribution with  $k_i$  degrees of freedom, scale matrix  $\hat{\Sigma}_i$  and non-centrality matrix  $\hat{\zeta}_i^T \hat{\zeta}_i - k_i \hat{\Sigma}_i$ .

## 2.4 Simulation settings

The genotype data in all the simulation settings are generated by white British individuals from UKB[25] while the reference panel is either the in-sample reference panel from UKB or the external reference panel from EUR 1KG Phase

3 data across the settings. We first conducted extensive simulations on blocks with different sizes to evaluate the performance of  $\rho$ -Hess, SUPERGNOVA, and LAVA under 1) varying sample overlaps between two GWASs, 2) both continuous and binary traits, 3) both infinitesimal and non-infinitesimal models, 4) the presence or absence of correlations between effect sizes and LD, and 5) different reference panels. Then based on the above simulation results, we further investigated the impact of the number of eigenvalues and eigenvectors on the stability and inference of  $\rho$ -hess and LAVA. We repeated each simulation setting 100 times and summarize the simulation settings in Table 1 and describe the details below.

We selected overlapping SNPs from chromosome 1 in UKB, EUR 1KG Phase3, and HapMap3[33] datasets for efficient simulations and to ensure sufficient SNP coverage. We then selected SNPs with MAF  $> 5\%$ , genotype missing rate  $< 5\%$ , and Hardy-Weinberg equilibrium P-value  $> 1e-10$ . After removing SNPs with ambiguous alleles, 71,609 SNPs remained for our simulation. We randomly selected 20,000 unrelated white British individuals from UKB and divided them into two subgroups of 10,000 individuals each, labeled as set1 and set2, respectively. We formed another set3 with 5,000 individuals from set1 and 5,000 individuals from set2. We randomly selected four blocks on chromosome 1 having 525 SNPs (POS: 60197393-61754126), 743 SNPs (POS: 3264297-5311384), 1033 SNPs (POS: 245966297-249239303), and 2315 SNPs (POS: 113753415-146215362), respectively. We treated one block as the local region of interest in each simulation and the other SNPs as the background SNPs. We simulated two traits whose SNP effects followed the multivariate normal distribution, with correlation only for SNPs within the chosen region of interest. The correlation of the local genetic effects was set to be 0, 0.3,

0.6, and 0.9, respectively. The remaining SNPs on chromosome 1 were considered background SNPs without genetic correlation. We set the heritability of two traits to be 0.5 which was evenly distributed to all SNPs (71,609 SNPs), so the local heritability of the above four blocks was 0.0037, 0.0052, 0.0072, and 0.0162, respectively. Genome-wide Complex Trait Analysis (GCTA)[34] was applied to simulate continuous traits  $\phi_1$  and  $\phi_2$ . We used PLINK[28] to analyze the simulated traits and generate GWAS summary statistics.

We considered no sample overlap, partial sample overlap, and complete sample overlap. When there was no sample overlap, two continuous traits,  $\phi_1$  and  $\phi_2$ , were simulated on set1 and set2, respectively. For  $\phi_1$  and  $\phi_2$  with partial sample overlap, set1 and set3 were used and the covariance of non-genetic effects was set to 0.2. As for the case where the samples were completely overlapping, we used set1 to simulate both  $\phi_1$  and  $\phi_2$  and the covariance of non-genetic effects was still set to 0.2. Additionally, we considered the situation that 20% of the SNPs were causal SNPs, where we randomly chose 20% of the SNPs in the regions of interest and 20% of SNPs in the background to be causal. For this case, we considered the no-sample overlap case where the two traits were continuous. We also conducted a simulation where the two traits were binary with no sample overlap. We considered the same local regions of interest, heritability, and genetic correlation for continuous traits. We used a liability model to simulate the binary traits. We first simulated continuous traits  $\phi_1$  and  $\phi_2$  and then the binary traits were set to be  $I[\phi_1 > \gamma]$  and  $I[\phi_2 > \gamma]$ , where  $\gamma$  was the quantile of standard normal distribution. Since we considered two simulation settings with  $\gamma$  to be 80% or 50%, the prevalence of the binary traits in the two simulations was 0.2 or 0.5. Besides, we also considered the situation where the effect sizes were correlated with local LD, which was the baseline in the LAVA simulation which was also mentioned in

SUPERGNOVA. In the simulation setting, LAVA first decomposed the local LD matrix of the reference panel as  $V_i = Q_i \Lambda_i Q_i^T$  for block  $i$  and obtained the subset of eigenvalues  $\Lambda_i^*$  and eigenvectors  $Q_i^*$  that explained 99% of the variance. We denote the number of eigenvalues thus selected as  $q$ . LAVA defined the projected genotype matrix in its simulation setting for block  $i$  as  $W_i = X_i Q_i^* \Lambda_i^{*-0.5}$  where  $X_i$  is the standardized genotype of the reference panel in block  $i$ . It then generated  $\delta_*$ , a  $q \times 2$  matrix with 0 means and identity variance, and decomposed the variance-covariance matrix of the genetic components as  $\Omega = Q' \Lambda' Q'^T$  and set  $\delta = Q' \Lambda'^{0.5} \delta_*$ . It simulated the genotype component for two traits as  $G_i = W_i \delta$ . Thus, the effects in the simulation settings of LAVA were  $Q_i^* \Lambda_i^{*-0.5} \delta$  where the effect sizes were correlated with local LD. We note that SUPERGNOVA also conducted simulations when the effect sizes were associated with ldsc. Thus, we also considered the simulation setting where the effect sizes were generated similarly from the simulation setting in LAVA[21] so that the effect sizes were related to local LD. For each simulation setting described above, we used both the in-sample reference panel from the UKB set1 samples and one external reference panel from the 1KG Phase 3 data.

Across the above simulation settings, the performance of LAVA was sensitive to the choice of reference panels. To further investigate the stability of LAVA, we applied LAVA using six different reference panels, 1) EUR 1KG Phase3 reference panel, 2) UKB reference panel with 500 randomly selected individuals from set1, 3) UKB reference panel with 5,000 randomly selected individuals from set1, 4) UKB reference panel with all 20,000 individuals from set1 and set2, 5) UKB reference panel with 20,000 individuals randomly selected from unrelated white British populations in UKB which do not



overlap with set1 and set2, and 6) 20,000 CEU individuals simulated using HAPGEN2[35](CEU refers to Northern Europeans from Utah).

Even though the  $\rho$ -hess-based estimates were more stable when using different reference panels, there was a substantial difference in statistical inference. As  $\rho$ -hess allows users to change the number of eigenvalues, we considered different reference panels and varied the number of eigenvalues to further investigate the performances of  $\rho$ -hess (the EUR 1KG Phase 3 reference panel; the UKB reference panel with samples from set1 and set2; and the CEU reference panel using HAPGEN2 with 20,000 individuals). For each reference panel, we varied the number of eigenvalues to explain 99%, 95%, 90%, 85%, 80%, and 70% variance in the above-selected blocks.

## 2.5 GWAS summary statistics

We analyzed 31 complex traits whose GWAS summary statistics are publicly available. These GWASs were primarily generated using individuals of European ancestry. The sources, sample sizes, and global heritability for these traits are listed in Table 2. To prepare data for analysis, we employed the `munge_sumstats.py` script from LDSC[32] to reformat and conducted quality control on the datasets, including the elimination of strand-ambiguous SNPs and the intersection of the remaining SNPs with those from the 1KG Project. In our analysis, we considered only autosomal SNPs with  $MAF > 5\%$  and excluded the MHC block on chromosome 6 (30-31Mb).

## 2.6 Polygenic risk score (PRS) analysis

We used the positively correlated and negatively correlated blocks from  $\rho$ -hess, SUPERGENOVA, and LAVA (with  $FDR < 0.1$ ) between ASD and CP, CAD and LDL, and T2D and BMI to construct PRS+ and PRS- for ASD, CAD, and T2D, respectively. These SNPs were clumped using PLINK, with a

significance threshold of 1 for index SNPs, an LD threshold of 0.1 for clumping, and a physical distance threshold of 250kb. PRSs were generated for ASD probands in the SPARK cohort and CAD and BMI cases in the UKB dataset. In addition, we compared CP (measured by IQ), LDL, and BMI between patients with high PRS+ and those with high PRS- for relevant disorders.

## 3 Results

### 3.1 Simulation Results

#### 3.1.1 Basic Simulation analysis

We compared the performance of  $\rho$ -hess, SUPERGNOVA, and LAVA by point estimation of local genetic correlation, local genetic covariance, local heritability, type I error, and statistical power. Since all three methods can use customized reference panels, we performed simulations on both the in-sample reference panel and the external reference panel with matched ancestry to investigate the robustness of these methods to the choice of LD reference panels.

We considered the simulation settings described in the Methods section to simulate traits based on the genotype data in the UKB and used the EUR 1KG Phase3 data and sample set1 from UKB as reference panels. For the continuous traits generated from non-overlapping samples (set1 and set2), SUPERGNOVA provided unbiased estimates for local genetic correlation, local genetic covariance, and local heritability (Figure 1A and Supplementary Figure 2 and 3) in all settings, and the results were robust to the choice of reference panel. However, SUPERGNOVA sometimes had inflated type-I error (Figure 1B and Supplementary Figure 4). When the EUR 1KG Phase 3 reference panel was used, which did not match the GWAS samples, LAVA overestimated local

genetic covariance and local heritability and underestimated local genetic correlation (Figure 1A and Supplementary Figure 2 and 3), and the higher the local genetic covariance, the less accurate the point estimates obtained from LAVA. LAVA had a higher inflated type-I error (about 20%) (Figure 1B and Supplementary Figure 4) than SUPERGNOVA and  $\rho$ -hess. On the other hand, if the in-sample UKB reference panel was used, LAVA yielded unbiased estimates for local genetic covariance and local heritability, and more accurate local genetic correlation estimates with well-controlled type-I error (Figure 1 and Supplementary Figure 2-4). In contrast, regardless of the reference panels used,  $\rho$ -hess always underestimated local genetic covariance and local heritability, particularly when local genetic covariance and local heritability were high (Supplementary Figure 2 and 3), but it provided unbiased local genetic correlation estimation, which may be due to compensation for both underestimated local genetic covariance and local heritability. The statistical test based on  $\rho$ -hess was overly conservative, leading to reduced statistical power (Figure 1B, Supplementary Figure 4).

Since the shared sample size between two traits needs to be provided to  $\rho$ -hess, we used both the correct shared sample size and incorrect overlapping sample size (1,000) to investigate the impact of this parameter on  $\rho$ -hess for the partial and complete overlapping scenarios. In this case, the performance for point estimate and inference by SUPERGNOVA was the same (Supplementary Figures 5-10). However, LAVA did not have well-controlled type-I error with overlapping samples, even when the in-sample reference panel was used (Supplementary Figures 5 and 8). With an incorrect overlapping sample size,  $\rho$ -hess had much reduced statistical power (Supplementary Figures 5 and 8).

When only some SNPs were causal, the performance of different methods was similar when all the SNPs were set to be causal, except for

LAVA having some inflated type-I error even using the in-sample reference panel. This suggests that the sparsity of causal SNPs does not have much impact on the performances of local genetic correlation/covariance estimation (Supplementary Figure 11-13).

Since all three methods can be applied to binary traits, we also considered binary traits in our simulation. Even though the genetic covariance is estimated on the observed scale, there is no distinction between observed- and liability-scale genetic correlation[14]. The estimates for the genetic correlation of binary traits had similar performances to that for continuous traits except with larger variations across the 100 repeats (Supplementary Figures 14A and 15A). However, the statistical power for binary traits was lower for all methods compared to continuous traits, especially  $\rho$ -hess which barely detected any significant blocks in our simulations (Supplementary Figures 14B and 15B).

### **3.1.2 LD-related effect sizes**

For simulations where the effect sizes were associated with the local LD structure, similar to the simulation setting in LAVA, there was a substantial underestimation of local genetic covariance (Supplementary Figure 16) with nearly no significant block detected by SUPERGNOVA and  $\rho$ -hess with the default settings. When we gave  $\rho$ -hess the number of eigenvalues that could explain 99% variance and used the UKB reference panel, the performance of  $\rho$ -hess improved but still had lower power than LAVA. In this setting, LAVA had the best performance in this simulation setting except for the inflated type-I error when using EUR 1KG Phase 3 data as the reference panel.

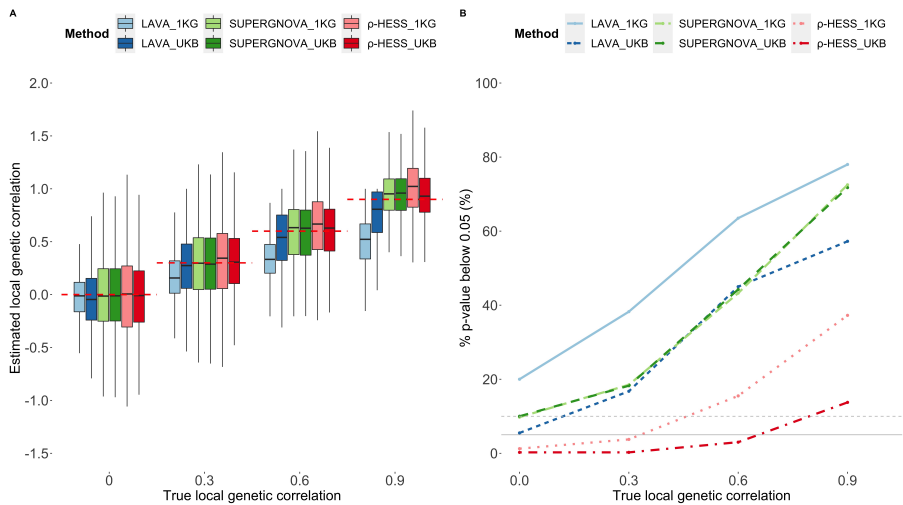
### 3.1.3 Robustness to reference panels

Our simulation results suggest that with different reference panels, the results from LAVA can be unstable, thus we further investigated the choice of the reference panels on LAVA. We considered the EUR 1KG Phase 3 reference panel, four different UKB reference panels, and one CEU reference panel (Method). As shown in Figure 2, when we used two UKB reference panels with 20,000 participants, LAVA had unbiased estimations and well-controlled type-I errors. With smaller UKB reference samples, LAVA could not provide reliable local genetic correlation estimates and well-controlled type-I errors. By comparing the performance between using the CEU reference panel and the UKB reference panel with the same sample size (20,000), the sample size of the reference panel was not a key factor for the performance of LAVA. Our results suggest that LAVA only performs well with enough individuals from the genotype data cohorts used as the reference panel.

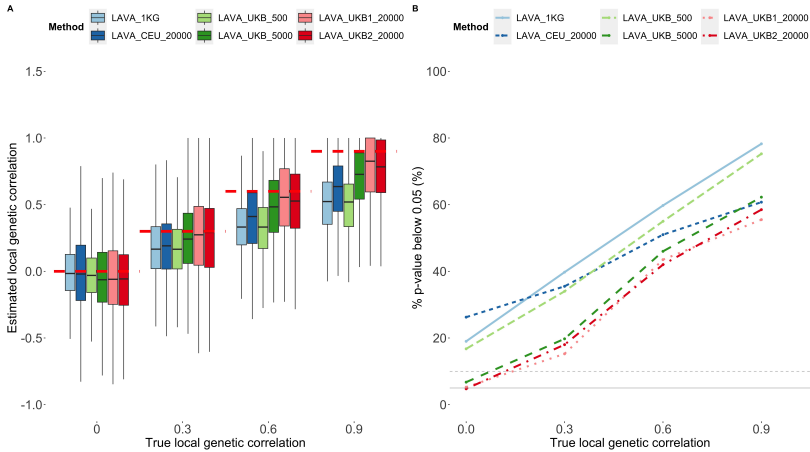
### 3.1.4 Number of eigenvalue input

By using the EUR 1KG Phase3, UKB samples from set1 and set2 and 20,000 CEU individuals(Method) as reference panels, we investigated whether the optimal number of eigenvalues used in  $\rho$ -hess stays more or less the same for different blocks and different reference panels. The optimal number of eigenvalues here is the one that could result in well-controlled type-I errors and higher powers, although the point estimates derived using these numbers could still be biased. When the reference panel was the same as that from the GWAS samples, the more eigenvalue used, the better the performance for  $\rho$ -hess, although it still had limited statistical power (Supplementary Figure 18 and Supplementary Figure 4). With an external reference panel that is different from the GWAS samples, there was no consistent pattern observed, and the optimal

number of eigenvalues varied for blocks and reference panels (Supplementary Figures 17, 19 and Supplementary Table 3). When using the EUR 1KG Phase 3 reference panel, the optimal number of eigenvalues was 93 (95%), 153 (90%), 146 (95%), and 85 (70%), respectively. For the CEU reference panel, the optimal number was 93 (95%), 310 (99%), 42 (70%), and 572 (99%), respectively. These observations could explain the poorer performance of  $\rho$ -hess in Figure 1 when using the in-sample UKB reference panel than the external 1KG Phase3 reference panel. This is because when the in-sample reference panels were used, the larger the number of eigenvalues used, the better the performance, while in Figure 1, the default number was 50.



**Fig. 1** Evaluation of local genetic correlation/covariance methods on continuous traits from non-overlapping datasets (set1 and set2) using EUR 1KG Phase3 and UKB reference panel. (A) local genetic correlation estimates. The red dashed lines represent the true value of local genetic correlation. (B) type-I error and statistical power. The solid grey line represents 5% p-values below 0.05 in 100 repeats, and the grey dashed line represents 10% p-values below 0.05 in 100 repeats.



**Fig. 2** Evaluation of local genetic correlation estimated by LAVA on continuous traits from non-overlapping datasets (set1 and set2) using different reference panels. (A) local genetic correlation estimates. The red dashed lines represent the true value of local genetic correlation. (B) type-I error and statistical power. The solid grey line represents 5%, and the grey dashed line represents 10%. LAVA\_1KG represents the EUR 1KG Phase3 reference panel; LAVA\_CEU\_20000 represents the CEU reference panel using HAPGEN2; LAVA\_UKB\_500 represents the UKB reference panel using 500 individuals randomly selected from set1; LAVA\_UKB\_5000 represents the UKB reference panel using 5,000 individuals randomly selected from set1; LAVA\_UKB1\_20000 represents the UKB reference panel using 20,000 individuals from set1 and set2; LAVA\_UKB2\_20000 represents the UKB reference panel using 20,000 individuals that do not overlap with set1 and set2.

## 3.2 Local genetic correlation/covariance of 31 complex traits

We considered 31 complex disorders or traits to compare the performance of the three methods. Table 2 summarizes these traits, abbreviations, sample sizes (the number of cases and the number of controls for binary traits), global heritability and its standard error derived from LDSC[14], and the original papers.

### 3.2.1 Stability based on different reference panels

When using only EUR 1KG Phase 3 genotype data as our reference panel to estimate the local genetic correlation for the above 31 complex traits,

there were substantial differences in point estimates and inferences by  $\rho$ -hess, SUPERGNOVA, and LAVA (Supplementary Material Appendix B). we focused on comparing the point estimates and detecting significant blocks using different reference panels for the same method in this section because the simulation results showed the importance of the reference panels for both  $\rho$ -hess and LAVA. Since the heritability of Height is the largest among all the traits, for clearer and more efficient comparison we decreased the number of trait pairs estimated and compared the results of the genetic correlation between Height and other traits using two other reference panels which were generated using different randomly selected white British UKB samples (Methods) and EUR 1KG Phase 3 reference panel. We compared the local heritabilities and local genetic correlations in the same block for the same trait or trait pairs using the same methods but different reference panels. As seen in Supplementary Figures 26-31, SUPERGNOVA displayed the most stable point estimates for local genetic correlation and local heritability using different reference panels and the estimates from  $\rho$ -hess were more stable than LAVA. Even with two different references from the same cohort, i.e. the two UKB reference panels, LAVA resulted in different estimates for the same block for the same pair of traits (Supplementary Figures 30 and 31).

Since the sum of local heritability should equal global heritability and the local genetic covariance should equal global genetic covariance (Methods), we further compared the sum of local heritability and local genetic covariance with global heritability and global genetic covariance with different reference panels. As shown in Supplementary Figure 32, LAVA tended to overestimate local heritability which is consistent with simulation results when the samples of the reference panel and the GWAS did not match. The sum of local heritability was highly concordant with the estimated global heritability for  $\rho$ -hess. For



SUPERGNOVA, except for three traits, Lupus, OCD, and T1D, the sum of local heritability was also highly concordant with the global heritability. The sum of local genetic covariance had a high correlation with the global genetic covariance for SUPERGNOVA and LAVA but was lower for  $\rho$ -hess (Figure 3). As shown in Figure 4 and Supplementary Figure 34, the significant blocks found by SUPERGNOVA and  $\rho$ -hess were consistently detected using different reference panels, while the results differed substantially for LAVA with different reference panels.

### 3.2.2 PRS analysis

Several studies[20, 36] including SUPERGNOVA have investigated the shared genetics among autism spectrum disorder (ASD), Attention-deficit/hyperactivity disorder (ADHD), and cognitive ability (CP) by utilizing local genetic information. To further compare the results of  $\rho$ -hess, SUPERGNOVA, and LAVA, we applied these methods to ASD, ADHD, and CP. By using a false discovery rate (FDR) cutoff of 0.1, we identified one block by  $\rho$ -hess, 55 blocks by SUPERGNOVA, and 126 blocks by LAVA with significant local genetic covariances between ADHD, ASD, and CP (Supplementary Table 8), respectively. The only block identified by all three methods was on chromosome 6 which was positively correlated between ASD and CP (POS: 97094444-98938023). Additionally, this is the same block that was significantly correlated between ASD and CP by both LAVA and SUPERGNOVA. The global genetic correlation between ASD and CP was 0.2 ( $p=1.8e-10$ ), between ASD and ADHD was 0.36 ( $p=1.14e-11$ ), and between ADHD and CP was -0.38 ( $p<1e-11$ ) revealing that the local correlations of CP with ASD and ADHD were bidirectional. As in Supplementary Figure 33, there was no significant block with a negative correlation between ASD and ADHD identified using LAVA and there were only two such blocks detected by SUPERGNOVA.

Besides, there was no block where ASD and ADHD showed opposite correlations with CP. SUPERGNOVA identified 12 blocks with positive correlations and four blocks with negative correlations between ASD and CP. LAVA identified 14 positively correlated blocks and 14 negatively correlated blocks between the two traits. We constructed positive and negative polygenic risk scores (Methods), referred to as PRS+ and PRS-, of ASD based on independent SNPs from blocks with significant positive or negative local correlations between ASD and CP detected by SUPERGNOVA or LAVA, respectively, for 1,026 ASD probands who had both genotypes and IQ scores in SPARK (Methods).

We observed probands with high PRS+ had higher IQ than probands with high PRS- only in PRSs generated utilizing SUPERGNOVA (Figure 5A-I). No negative blocks were detected by  $\rho$ -hess, resulting in only PRS+ constructed based on  $\rho$ -hess (Figure 5C, F, I). When using PRS+ and PRS- based on SUPERGNOVA, there was a sharp change in the right tails of the PRS distribution analysis of the average full-scale IQ, from 84.7 and 83.1 in the 75th percentile to 89.9 and 75.0 in the 99th percentile for PRS+ and PRS-, respectively (Figure 5A). Similarly, the average non-verbal IQ (Figure 5D) and verbal IQ (Figure 5G) also showed a sharp change in the right tail of the PRS distribution, with respective changes from 93.2 and 92.5 in the 75th percentile to 101.7 and 84.0 in the 99th percentile, and from 94.9 and 91.4 in the 75th percentile to 102.1 and 80.4 in the 99th percentile for PRS+ and PRS-, respectively.

The LAVA study[21] explored the relationship between LDL and CAD, and between BMI and T2D from the angle of multivariate correlation. Here we conducted a PRS analysis using the results of the bivariate local genetic correlation between LDL and CAD, and BMI and T2D. The global genetic correlation between LDL and CAD was 0.3 ( $p < 1 \times 10^{-15}$ ). SUPERGNOVA identified 36 positive blocks and 5 negative blocks with significant local genetic

correlations between LDL and CAD at an FDR level of 0.1, and LAVA identified 108 positive blocks and 30 negative blocks (Supplementary Table 9). No significant block was identified using  $\rho$ -hess. SUPERGNOVA and LAVA identified 22 common blocks that had consistent correlation directions, including 21 positive blocks and one negative block on chromosome 5. As displayed in Figure 5J-5K, CAD cases with high PRS+ had higher LDL than cases with high PRS- for both SUPERGNOVA and LAVA, with an average LDL changing from 3.41 and 3.32 for the 75 percentile to 3.54 and 3.10 for the 99 percentile when using SUPERGNOVA. However, the trend in LAVA was less apparent, with the average LDL moving from 3.39 and 3.38 for the 75 percentile to 3.38 and 3.21 for the 99 percentile.

When analyzing the local correlations between T2D and BMI whose global genetic correlation was 0.57 ( $p < 1 \times 10^{-15}$ ),  $\rho$ -hess identified 279 significant blocks, SUPERGNOVA identified 176 ones, and LAVA identified 589 blocks (Supplementary Table 10). A total of 93 blocks were found by all three methods with just one block on chromosome 3 showing a negative correlation between T2D and BMI, and all these 93 blocks had consistent correlation direction. Among the significant blocks,  $\rho$ -hess found 271 that were positively correlated and eight that were negatively correlated, SUPERGNOVA identified 170 that were positively correlated and 6 that were negatively correlated, and LAVA identified 66 that were positively correlated and 23 that were negatively correlated. As demonstrated in Figure 5L-5N, T2D cases with high PRS+ had a greater BMI than cases with high PRS- for all three methods. For SUPERGNOVA, the average LDL changed from 32.1 to 31.2 for the 75 percentile to 32.9 and 31.6 for the 99 percentile. For LAVA, the average LDL changes from 32.5 to 30.9 for the 75 percentile to 33.1 and 30.2 for the 99

percentile. For  $\rho$ -hess, the average LDL changes from 32.4 and 30.9 for the 75 percentile to 33.4 and 30.7 for the 99 percentile.

## 4 Discussion

In recent years, there has been an increasing interest in inferring local genetic correlation in post-GWAS analyses in addition to global genetic correlation. This trend can be attributed to advancements in methodologies for estimating local genetic correlation and detecting locally significant blocks, as well as a growing knowledge of the limitations of global genetic correlation for revealing the underlying genetic similarity between complex traits. Local genetic correlation has also been utilized to improve association studies and PRS prediction.

The first step for local genetic correlation is determining how to partition the whole genome into approximately independent blocks. The larger the blocks, the more independent the partitions, but larger blocks may mask local information in the same way that global genetic correlation does. On the other hand, smaller blocks may result in LD leakage and biased estimates. The three methods compared in this paper all provide their own partitions, but also allow users to use their own partitions. Another issue that needs to be addressed for local genetic correlation is also considered for global correlation, i.e. how to deal with pervasive sample overlap across GWASs. The common solution for these three methods is to utilize the cross-trait LDSC intercept to calculate the phenotypic correlation.  $\rho$ -hess is the only method that requires the shared sample size between two GWASs as input. However, as the number of GWASs generated by meta-analysis grows, the exact number of overlapping sample sizes is difficult to obtain. Our simulation results suggest that the power of  $\rho$ -hess will decrease if an incorrect number of shared sample size

is given. The other two methods have more stable performances in terms of the sample-overlapping level. The third and most crucial challenge with these three methods is estimating the local LD structure using external reference panels. Ideally, the external reference panels applied should have the same LD structure as the genotype data used to calculate summary statistics. In the real world, because access to individual-level data from the GWAS dataset is typically limited due to practical constraints, it is common to choose an external reference panel. Through extensive simulations and real data analysis, we have demonstrated that the choice of the local LD matrix is critical for both estimation and inference. SUPERGNOVA is the most robust method for the choice of reference panels because it has an adaptive procedure to choose the number of eigenvalues and eigenvectors used for different blocks and different reference panels. However, the type-I error of SUPERGNOVA is still inflated in some simulation settings which indicates a better adaptive procedure is still needed. LAVA recommends using the number of eigenvalues and eigenvectors that explain 99% of the variances and performed the best when the genotype data and the reference panel were perfectly matched. However, with different reference panels, LAVA could provide different estimations, and the significant blocks detected were also not consistent.  $\rho$ -hess needs to be given the number of eigenvalues as input and the default number is set to be 50. However, the optimal number of eigenvalues and eigenvectors depends on both the local LD structure of the reference panels and the LD structure of the blocks in the genotype data. In summary,  $\rho$ -hess can provide unbiased estimates if the proper number of eigenvalues is selected based on different reference panels, while SUPERGNOVA yields unbiased estimates assuming the underlying assumption holds true. Additionally, LAVA produces unbiased estimates when

an in-sample reference panel with sufficient sample sizes is utilized. While  $\rho$ -hess generally has well-controlled type-I error rates, it may have lower power. SUPERGENOVA is generally more stable across different reference panels, but may have slightly inflated type-I error at times. LAVA only produces well-controlled type-I error rates when an in-sample reference panel with sufficient sample sizes is used.

Despite extensive simulation settings and real data sets considered, there are limitations in our study. First, the methods compared in this study are those that can reveal correlated blocks between two traits within a single population (e.g. European). All these three methods can provide both estimates and references with user-defined partitions. However, there are other methods that could detect corrected blocks between different populations (e.g. European and African) for the same trait[37] or evaluate the concordance of two traits on the method-defined regions[22]. Thus, a more general comparison or review is needed. Secondly, there is no gold standard to compare these methods in real-world data applications since true local genetic correlations or significantly correlated blocks between phenotypic pairs are unknown. Even though we have conducted PRS analysis to help assess the performances of these methods, other downstream analyses can be done to compare the performance of different methods.

## 5 Acknowledgments

We conducted the research using the UKBB resource under approved data requests (access ref: 29900). This study makes use of summary statistics from many GWAS consortia. We thank the investigators in these GWAS consortia

for generously sharing their data. We are grateful to all the families participating in the Simons Foundation Powering Autism Research for Knowledge (SPARK) study.

## 6

## References

- [1] Visscher, P.M., Brown, M.A., McCarthy, M.I., Yang, J.: Five years of gwas discovery. *American journal of human genetics* **90** **1**, 7–24 (2012)
- [2] Loos, R.J.F.: 15 years of genome-wide association studies and no signs of slowing down. *Nature Communications* **11** (2020)
- [3] Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., Yang, J.: 10 years of gwas discovery: Biology, function, and translation. *American journal of human genetics* **101** **1**, 5–22 (2017)
- [4] Abdellaoui, A., Yengo, L., Verweij, K.J.H., Visscher, P.M.: 15 years of gwas discovery: Realizing the promise. *American journal of human genetics* (2023)
- [5] Gallagher, M.D., Chen-Plotkin, A.S.: The post-gwas era: From association to function. *American journal of human genetics* **102** **5**, 717–730 (2018)
- [6] Turley, P., Walters, R.K., Maghzian, O., Okbay, A., Lee, J.J., Fontana, M.A., Nguyen-Viet, T.A., Wedow, R., Zacher, M., Furlotte, N.A., Magnusson, P.K.E., Oskarsson, S., Johannesson, M., Visscher, P.M., Laibson, D.I., Cesarini, D., Neale, B.M., Benjamin, D.J.: Multi-trait analysis of genome-wide association summary statistics using mtg. *Nature genetics* **50**, 229–237 (2017)

- [7] Grotzinger, A.D., Rhemtulla, M., de Vlaming, R., Ritchie, S.J., Mal-lard, T.T., Hill, W.D., Ip, H.F., Marioni, R.E., McIntosh, A.M., Deary, I.J., Philipp, D., Koellinger, Harden, K.P., Nivard, M.G., Tucker-Drob, E.M.: Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. (2019)
- [8] Maier, R.M., Zhu, Z., Lee, S.H., Trzaskowski, M., Ruderfer, D.M., Stahl, E.A., Ripke, S., Wray, N.R., Yang, J., Visscher, P.M., Robinson, M.R.: Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nature Communications* **9** (2018)
- [9] Hu, Y., Lu, Q., Liu, W., Zhang, Y., Li, M., Zhao, H.: Joint modeling of genetically correlated diseases and functional annotations increases accuracy of polygenic risk prediction. *PLoS Genetics* **13** (2017)
- [10] Zhou, G., Zhao, H.: A fast and robust bayesian nonparametric method for prediction of complex traits using summary statistics. *PLoS genetics* **17** **7**, 1009697 (2021)
- [11] Pain, O., Lewis, C.M.: Using local genetic correlation improves polygenic score prediction across traits. *bioRxiv* (2022)
- [12] Miao, J., Guo, H., Song, G., Zhao, Z., Hou, L., Lu, Q.: Quantifying portable genetic effects and improving cross-ancestry genetic prediction with gwas summary statistics. *bioRxiv* (2022)
- [13] Zhang, Y., Cheng, Y., Jiang, W., Ye, Y., Lu, Q., Zhao, H.: Comparison of methods for estimating genetic correlation between complex traits using gwas summary statistics. *bioRxiv* (2020)
- [14] Bulik-Sullivan, B.K., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R.,



- Loh, P.-R., Duncan, L.E., Perry, J.R.B., Patterson, N.J., Robinson, E.B., Daly, M.J., Price, A.L., Neale, B.M.: An atlas of genetic correlations across human diseases and traits. *Nature genetics* **47**, 1236–1241 (2015)
- [15] Lu, Q., Li, B., Ou, D., Erlendsdottir, M., Powles, R., Jiang, T., Hu, Y., Chang, D., Jin, C., Dai, W., He, Q., Liu, Z., Mukherjee, S., Crane, P.K., Zhao, H.: A powerful approach to estimating annotation-stratified genetic covariance via gwas summary statistics. *American journal of human genetics* **101** **6**, 939–964 (2017)
- [16] Ning, Z., Pawitan, Y., Shen, X.: High-definition likelihood inference of genetic correlations across human complex traits. *Nature Genetics*, 1–6 (2020)
- [17] Zheng, J., Erzurumluoglu, M.A., Elsworth, B.L., Kemp, J.P., Howe, L.J., Haycock, P.C., Hemani, G., Tansey, K.E., Laurin, C., Genetics, E., Consortium, L.E.E., Pourcain, B.S., Warrington, N.M., Finucane, H.K., Price, A.L., Bulik-Sullivan, B.K., Anttila, V., Paternoster, L., Gaunt, T.R., Evans, D.M., Neale, B.M.: Ld hub: a centralized database and web interface to perform ld score regression that maximizes the potential of summary level gwas data for snp heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2016)
- [18] van Rheenen, W., Peyrot, W.J., Schork, A.J., Lee, S.H., Wray, N.R.: Genetic correlations of polygenic disease traits: from theory to practice. *Nature Reviews Genetics*, 1–15 (2019)
- [19] Shi, H., Mancuso, N., Spendlove, S., Pasaniuc, B.: Local genetic correlation gives insights into the shared genetic architecture of complex traits. *American journal of human genetics* **101** **5**, 737–751 (2017)

- [20] Zhang, Y., Lu, Q., Ye, Y., Huang, K., Liu, W., Wu, Y., Zhong, X., Li, B., Yu, Z., Travers, B.G., Werling, D.M., Li, J.J., Zhao, H.: Local genetic correlation analysis reveals heterogeneous etiologic sharing of complex traits. *bioRxiv* (2020)
- [21] Werme, J., van der Sluis, S., Posthuma, D., de Leeuw, C.A.: An integrated framework for local genetic correlation analysis. *Nature genetics* **54** **3**, 274–282 (2022)
- [22] Guo, H., Li, J.J., Lu, Q., Hou, L.: Detecting local genetic correlations with scan statistics. *Nature Communications* **12** (2019)
- [23] Partanen, J.J., Häppölä, P., Zhou, W., Lehisto, A., Ainola, M., Sutinen, E., Allen, R.J., Stockwell, A.D., Oldham, J.M., Guillen-Guio, B., Flores, C., Noth, I., Yaspan, B.L., Jenkins, R.G., Wain, L.V., Ripatti, S., Piriinen, M., Kaarteenaho, R., Myllärniemi, M., Daly, M.J., Koskela, J.T.: Leveraging global multi-ancestry meta-analysis in the study of idiopathic pulmonary fibrosis genetics. *Cell Genomics* (2021)
- [24] Shi, H., Kichaev, G., Pasaniuc, B.: Contrasting the genetic architecture of 30 complex traits from summary association data. *American journal of human genetics* **99** **1**, 139–53 (2016)
- [25] Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O’Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N.E., Donnelly, P., Marchini, J.: The uk biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018)
- [26] Clarke, L., Fairley, S., Bradley, X.Z., Streeter, I., Perry, E., Lowy-Gallego,

- E., Tassé, A.M., Flicek, P.: The international genome sample resource (igsr): A worldwide collection of genome variation incorporating the 1000 genomes project data. *Nucleic Acids Research* **45**, 854–859 (2017)
- [27] Feliciano, P., Daniels, A.M., Snyder, L.G., Beaumont, A.L., Camba, A., Esler, A.N., Gulsrud, A.G., Mason, A., Gutierrez, A., Nicholson, A.G., Paolicelli, A.M., McKenzie, A.P., Rachubinski, A.L., Stephens, A.N., Simon, A.R., Stedman, A., Shocklee, A.D., Swanson, A.R., Finucane, B.M., Hilscher, B.A., Hauf, B., j. O’Roak, B., McKenna, B.G., Robertson, B.E., Rodriguez, B., Vernioia, B.M., Metre, B.V., Bradley, C.C., Cohen, C., Erickson, C.A., Harkins, C.M., Hayes, C., Lord, C., Martin, C.L., Ortiz, C., Ochoa-Lubinoff, C., Peura, C., Rice, C., Rosenberg, C.R., Smith, C.J., Thomas, C.M., Taylor, C.M., White, L.C., Walston, C.H., Amaral, D.G., Coury, D.L., Sarver, D.E., Istephanous, D., Li, D.D., Nugyen, D.C., Fox, E.A., Butter, E.M., Berry-Kravis, E.M., Courchesne, E., Fombonne, E., Hofammann, E., Lamarche, E., Wodka, E.L., Matthews, E.T., O’Connor, E., Palen, E., Miller, F., Dichter, G.S., Marzano, G., Stein, G., Hutter, H., Kaplan, H.E., Li, H., Lechniak, H., Schneider, H.L., Zaydens, H., Arriaga, I., Gerdtts, J., Cubells, J.F., Cordova, J.M., Gunderson, J., Lillard, J., Manoharan, J., McCracken, J.T., Michaelson, J.J., Neely, J., Orobio, J., Pandey, J., Piven, J., Scherr, J.F., Sutcliffe, J.S., Tjernagel, J., Wallace, J., Callahan, K., Dent, K., Schweers, K.A., Hamer, K.E., Law, J.K., Lowe, K.R., O’Brien, K., Smith, K., Pawlowski, K.G., Pierce, K.L., Roeder, K., Abbeduto, L., Berry, L.N., Cartner, L.A., Coppola, L.A., Carpenter, L.A., Cordeiro, L., DeMarco, L., Grosvenor, L.P., Higgins, L., Huang-Storms, L.Y., Hosmer-Quint, L., Herbert, L.M., Kasparson, L., Prock, L.M., Pacheco, L.D., Raymond, L.W., Simon, L., Soorya, L.V., Wasserburg, L., Lazar, M., Alessandri,

- M., Brown, M., Currin, M.H., Gwynette, M.F., Heyman, M., Hale, M.N., Jones, M., Jordy, M., Morrier, M.J., Sahin, M.C., Siegel, M., Verdi, M.B., Parladé, M.V., Yinger, M., Bardett, N., Hanna, N., Harris, N., Pottschmidt, N.R., Russo-Ponsaran, N.M., Takahashi, N., Ousley, O.Y., Juárez, A.P., Manning, P., Annett, R.D., Bernier, R.A., Clark, R.D., Landa, R.J., Goin-Kochel, R.P., Remington, R., Schultz, R.T., Brewster, S.J., Booker, S., Carpenter, S., Eldred, S., Francis, S.M., Friedman, S.L., Horner, S., Hepburn, S., Jacob, S., Kanne, S.M., Lee, S.J., Mastel, S., Plate, S., Qiu, S., Sandhu, S., Thompson, S., White, S., Myers, V.J., Singh, V., Yang, W.S., Warren, Z., Amatya, A., Ace, A.J., Chatha, A.S., Lash, A.E., Negron, B., Rigby, C., Ridenour, C., Stock, C.M., Schmidt, D., Fisk, I., Acampado, J., Nestle, J.L., Nestle, J.L., Layman, K., Butler, M.E., Kent, M., Mallardi, M.D., Carriero, N., Lawson, N., Volfovsky, N., Edgar, R., Marini, R.A., Rana, R., Ganesan, S., Shah, S., Ramsey, T., Chin, W., Jensen, W., Krentz, A.D., Gruber, A.J., Sabo, A., Salomatov, A., Eng, C.M., Muzny, D.M., Astrovskaya, I., Gibbs, R.A., Han, X., Shen, Y., Reichardt, L.F., Chung, W.K.: Spark: A us cohort of 50,000 families to accelerate autism research. *Neuron* **97**, 488–493 (2018)
- [28] Purcell, S.M., Neale, B.M., Todd-Brown, K.E.O., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J.B., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C.: Plink: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81** **3**, 559–75 (2007)
- [29] Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A.M., Vrieze, S.I., Chew, E.Y., Levy, S.E., McGue, M., Schlessinger, D., Stambolian, D.E., Loh, P.-R., Iacono, W.G., Swaroop, A., Scott, L.J., Cucca,

- F., Kronenberg, F., Boehnke, M., Abecasis, G.R., Fuchsberger, C.: Next-generation genotype imputation service and methods. *Nature Genetics* **48**, 1284–1287 (2016)
- [30] Berisa, T., Pickrell, J.K.: Approximately independent linkage disequilibrium blocks in human populations. *bioRxiv* (2015)
- [31] Privé, F.: Optimal linkage disequilibrium splitting. *Bioinformatics* **38**, 255–256 (2021)
- [32] Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N.J., Daly, M.J., Price, A.L., Neale, B.M.: Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291–295 (2015)
- [33] Tanaka, T.: The international hapmap project. *Nature* **426**, 789–796 (2003)
- [34] Boomsma, D.I., Hottenga, J.-J., Walters, R.K., Laurin, C., de Geus, E.J.C., Willemsen, G., Smit, J.H., Middeldorp, C.M., Penninx, B.W.J.H., Vink, J.M., Lubke, G.H.: Genome-wide complex trait analysis (gcta) for complex traits including major depressive disorder and smoking. (2011)
- [35] Su, Z., Marchini, J., Donnelly, P.: Hapgen2: simulation of multiple disease snps. *Bioinformatics* **27** **16**, 2304–5 (2011)
- [36] Rao, S., Baranova, A.V., Yao, Y., Wang, J., Zhang, F.: Genetic relationships between attention-deficit/hyperactivity disorder, autism spectrum disorder, and intelligence. *Neuropsychobiology* **81**, 484–496 (2022)

- [37] Shi, H., Burch, K.S., Johnson, R., Freund, M.K., Kichaev, G., Mancuso, N., Manuel, A.M., Dong, N., Pasaniuc, B.: Localizing components of shared transethnic genetic architecture of complex traits from gwas summary data. *bioRxiv* (2019)
- [38] Zhang, H., Ahearn, T.U., Lecarpentier, J., Barnes, D., Beesley, J., Qi, G., Jiang, X., O'Mara, T.A., Zhao, N., Bolla, M.K., *et al.*: Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nature genetics* **52**(6), 572–581 (2020)
- [39] Mahajan, A., Spracklen, C.N., Zhang, W., Ng, M.C., Petty, L.E., Kitajima, H., Yu, G.Z., Rüeger, S., Speidel, L., Kim, Y.J., *et al.*: Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nature genetics* **54**(5), 560–572 (2022)
- [40] Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., *et al.*: Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature genetics* **51**(2), 237–244 (2019)
- [41] De Lange, K.M., Moutsianas, L., Lee, J.C., Lamb, C.A., Luo, Y., Kennedy, N.A., Jostins, L., Rice, D.L., Gutierrez-Achury, J., Ji, S.-G., *et al.*: Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature genetics* **49**(2), 256–261 (2017)
- [42] Wang, Y.-F., Zhang, Y., Lin, Z., Zhang, H., Wang, T.-Y., Cao, Y., Morris, D.L., Sheng, Y., Yin, X., Zhong, S.-L., *et al.*: Identification of 38 novel

- loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. *Nature communications* **12**(1), 772 (2021)
- [43] Meier, S.M., Trontti, K., Purves, K.L., Als, T.D., Grove, J., Laine, M., Pedersen, M.G., Bybjerg-Grauholm, J., Bækved-Hansen, M., Sokolowska, E., *et al.*: Genetic variants associated with anxiety and stress-related disorders: a genome-wide association study and mouse-model study. *JAMA psychiatry* **76**(9), 924–932 (2019)
- [44] Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R.K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O.A., Anney, R., *et al.*: Identification of common genetic risk variants for autism spectrum disorder. *Nature genetics* **51**(3), 431–444 (2019)
- [45] Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., Bækvad-Hansen, M., *et al.*: Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature genetics* **51**(1), 63–75 (2019)
- [46] Howard, D.M., Adams, M.J., Clarke, T.-K., Hafferty, J.D., Gibson, J., Shirali, M., Coleman, J.R., Hagenaars, S.P., Ward, J., Wigmore, E.M., *et al.*: Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience* **22**(3), 343–352 (2019)
- [47] Watson, H.J., Yilmaz, Z., Thornton, L.M., Hübel, C., Coleman, J.R., Gaspar, H.A., Bryois, J., Hinney, A., Leppä, V.M., Mattheisen, M., *et al.*: Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature genetics* **51**(8),

1207–1214 (2019)

- [48] Arnold, P.D., Askland, K.D., Barlassina, C., Bellodi, L., Bienvenu, O., Black, D., Bloch, M., Brentani, H., Burton, C.L., Camarena, B., *et al.*: Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular psychiatry* **23**(5), 1181–1181 (2018)
- [49] Trubetskoj, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.-Y., Dennison, C.A., Hall, L.S., *et al.*: Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**(7906), 502–508 (2022)
- [50] Mullins, N., Forstner, A., O’Connell, K.: Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* **53**(6), 817–829 (2021)
- [51] Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghziyan, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Linnér, R.K., Fontana, M.A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P.N., Walters, R.K., Willoughby, E.A., Yengo, L., Alver, M., Bao, Y., Clark, D.W., Day, F.R., Furlotte, N.A., Joshi, P.K., Kemper, K.E., Kleinman, A., Langenberg, C., Mägi, R., Trampush, J.W., Verma, S.S., Wu, Y., Lam, M., Zhao, J.H., Zheng, Z., Boardman, J.D., Campbell, H., Freese, J., Harris, K.M., Hayward, C., Herd, P., Kumari, M., Lencz, T., Luan, J., Malhotra, A.K., Metspalu, A., Milani, L., Ong, K.K., Perry, J.R.B., Porteous, D.J., Ritchie, M.D., Smart, M.C., Smith, B.H., Tung, J.Y., Wareham, N.J., Wilson, J.F., Beauchamp, J.P., Conley, D.C., Esko, T., Lehrer, S.F., Magnusson, P.K.E., Oskarsson, S., Pers, T.H., Robinson, M.R., Thom, K.,

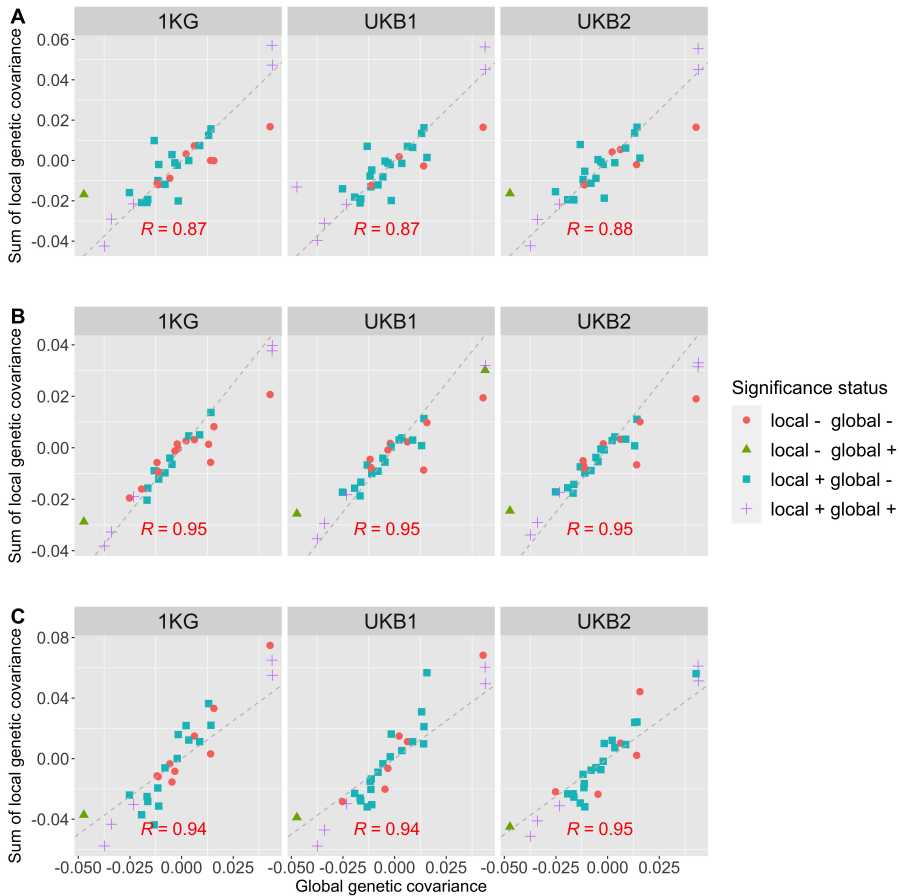


- Watson, C., Chabris, C.F., Meyer, M.N., Laibson, D.I., Yang, J., Johannesson, M., Koellinger, P.D., Turley, P., Visscher, P.M., Benjamin, D.J., Cesarini, D.: Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* **50**, 1112–1121 (2018)
- [52] McKay, J.D., Hung, R.J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D.C., Caporaso, N.E., Johansson, M., Xiao, X., Li, Y., *et al.*: Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nature genetics* **49**(7), 1126–1132 (2017)
- [53] Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T.J., van der Sluis, S., Andreassen, O.A., Neale, B.M., Posthuma, D.: A global overview of pleiotropy and genetic architecture in complex traits. *Nature genetics* **51**(9), 1339–1348 (2019)
- [54] Pult, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., *et al.*: Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of european ancestry. *Human molecular genetics* **28**(1), 166–174 (2019)
- [55] Dashti, H.S., Jones, S.E., Wood, A.R., Lane, J.M., Van Hees, V.T., Wang, H., Rhodes, J.A., Song, Y., Patel, K., Anderson, S.G., *et al.*: Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nature communications* **10**(1), 1100 (2019)

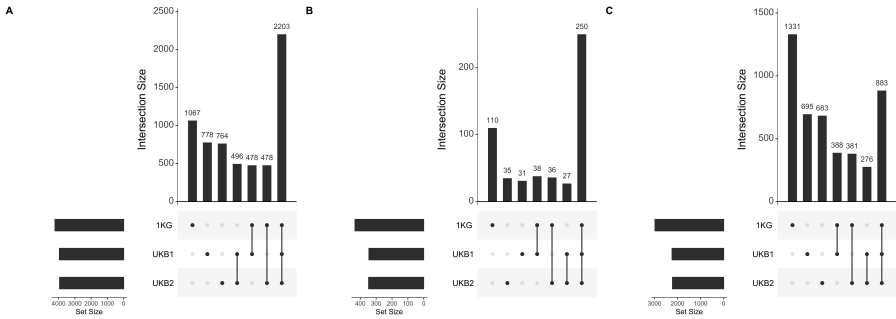
- [56] Luciano, M., Hagenaars, S.P., Davies, G., Hill, W.D., Clarke, T.-K., Shirali, M., Harris, S.E., Marioni, R.E., Liewald, D.C., Fawns-Ritchie, C., *et al.*: Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nature genetics* **50**(1), 6–11 (2018)
- [57] Klimentidis, Y.C., Arora, A., Newell, M., Zhou, J., Ordovas, J.M., Renuquist, B.J., Wood, A.C.: Phenotypic and genetic characterization of lower ldl cholesterol and increased type 2 diabetes risk in the uk biobank. *Diabetes* **69**(10), 2194–2205 (2020)
- [58] Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., *et al.*: Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature genetics* **54**(4), 437–449 (2022)
- [59] Van Der Harst, P., Verweij, N.: Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circulation research* **122**(3), 433–443 (2018)
- [60] Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., *et al.*: Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**(7488), 376–381 (2014)
- [61] Forgetta, V., Manousaki, D., Istomine, R., Ross, S., Tessier, M.-C., Marchand, L., Li, M., Qu, H.-Q., Bradfield, J.P., Grant, S.F., *et al.*: Rare genetic variants of large effect influence risk of type 1 diabetes. *Diabetes* **69**(4), 784–795 (2020)

Evaluation factors	Overlapping level	Traits	Reference panels	Genetic correlation	non-genetic effect	Prevalence	Figures
Sample overlapping level	No sample overlap (set1 and set2)	continuous	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0	/	Figure 1 Supplementary Figure 2-4
	Partial sample overlap (set1 and set3)	continuous	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0.2	/	Supplementary Figure 5-7
Infinitesimal model (20% causal SNPs)	Complete sample overlap (set1 and set1)	continuous	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0.2	/	Supplementary Figure 8-10
	No sample overlap (set1 and set2)	continuous	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0	/	Supplementary Figure 11-13
Binary traits	No sample overlap (set1 and set2)	binary	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0	0.2	Supplementary Figure 14
		binary	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0	0.5	Supplementary Figure 15
Effect sizes correlated with LD	No sample overlap (set1 and set2)	continuous	1. UKB from set1, 2. EUR IKG	0, 0.3, 0.6, 0.9	0	/	Supplementary Figure 16
			1. UKB 500 from set1 2. UKB 5,000 from set1 3. UKB 20,000 from set1 and set2				
Stability of LAVA on reference panels	No sample overlap (set1 and set2)	continuous	4. UKB 20,000 non-overlap with set1 and set2 5. EUR IKG 6. CEU 20,000 using HAPPEN2	0, 0.3, 0.6, 0.9	0	/	Figure 2
			1. UKB 20,000 non-overlap with set1 and set2 2. EUR IKG 3. CEU 20,000 using HAPPEN2				
# eigenvalues and eigenvectors in $\rho$ -hess	No sample overlap (set1 and set2)	continuous		0, 0.3, 0.6, 0.9	0	/	Supplementary Figure 17-19

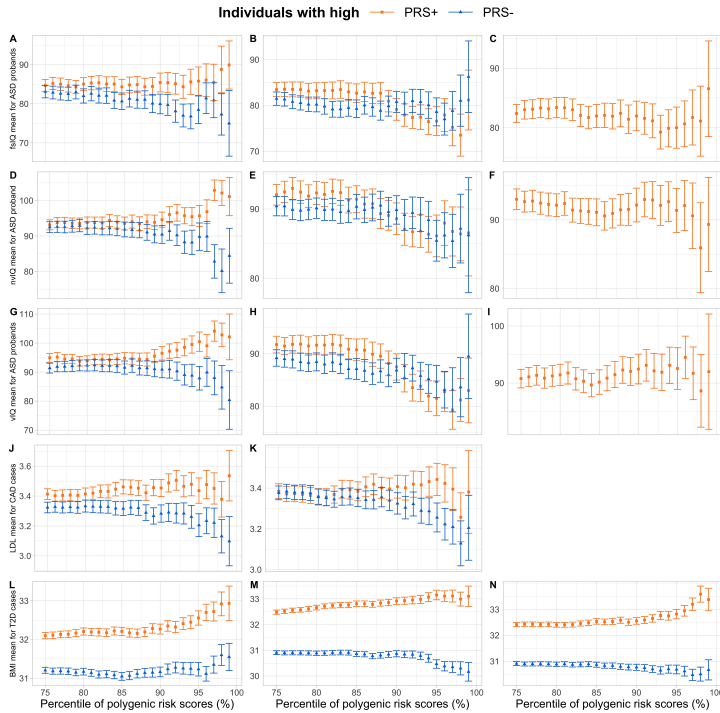
Table 1 Details of each simulation setting



**Fig. 3 Comparisons of the sum of local genetic covariance for 465 trait pairs with the global genetic covariance derived from LDSC.** A. Comparisons of the sum of local genetic covariances estimated from A  $\rho$ -hess, B SUPERGENOVA or C LAVA with the global genetic covariances using three different reference panels. Each point represents a trait pair. The color and shape of each data point denote the significance status in global and local correlation analyses. 'local +' denotes that there are significant blocks detected between that trait pair and 'local -' denotes that there are no corrected blocks detected. 'global +' denotes the global genetic correlation is significant, while 'global -' denotes that the global genetic correlation is not significant. The figures are divided into multiple panels, with each panel corresponding to different reference panels (EUR 1KG reference panel and two UKB reference panels with different samples). The dashed, grey reference line with a slope of 1 represents the line of perfect correlation in each panel. The strength of the relationship is indicated by Pearson correlation coefficients, which are displayed at the bottom of each panel.



**Fig. 4 Comparisons of blocks with significant local genetic correlations when using different reference panels.** These plots used bars to break down the Venn diagram of overlapped significant blocks using different reference panels using FDR at 0.1 level detected by A.  $\rho$ -hess, B. SUPERGENOVA, and C. LAVA.



**Fig. 5 Phenotype heterogeneity of ASD probands, CAD and T2D patients with high PRS+ and PRS.** Average full-scale IQ is computed for different groups defined by PRS based on the significant blocks found by A. SUPERGENOVA B. LAVA and C.  $\rho$ -hess. Average non-verbal IQ is computed for different groups defined by PRS based on the significant blocks found by D. SUPERGENOVA E. LAVA and F.  $\rho$ -hess. Average verbal IQ is computed for different groups defined by PRS based on the significant blocks found by G. SUPERGENOVA H. LAVA and I.  $\rho$ -hess. Average LDL is computed for different groups defined by PRS based on the significant blocks found by J. SUPERGENOVA K. LAVA. Average BMI is computed for different groups defined by PRS based on the significant blocks found by L. SUPERGENOVA M. LAVA and N.  $\rho$ -hess. Each interval indicated the standard error of the average values.

trait	Abbreviation	Sample Size	Number of Cases	Number of Controls	Global Heritability (s.e.)	Pubmed ID
Breast Cancer	BC	247,173	133,384	113,789	0.1253(0.0100)	32424353[38]
Type 2 Diabetes	T2D	933,970	80,154	853,816	0.0467(0.0022)	35551307[39]
Alcoholism	DrnkWk	537,349			0.0492(0.0022)	30643251[40]
Smoking Behavior	SmkInit	632,802			0.0678(0.0022)	30643251[40]
Ulcerative Colitis	UC	45,975	12,366	33,609	0.2487(0.0247)	28067908[41]
Inflammatory Bowel Disease	IBD	59,957	25,042	34,915	0.2981(0.0263)	28067908[41]
Crohn's Disease	Crohn	40,266	12,194	28,072	0.4215(0.0438)	28067908[41]
Lupus	Lupus	12,615	4,576	8,039	0.0954(0.0459)	33536424[42]
Anxiety Disorder	AXD	31,880	12,655	19,225	0.2073(0.0204)	31116379[43]
Autism Spectrum Disorder	ASD	46,351	18,382	27,969	0.2028(0.0164)	30804558[44]
Attention-Deficit/Hyperactivity Disorder	ADHD	53,293	19,099	34,194	0.2362(0.0158)	30478444[45]
Major Depressive Disorder	MDD	500,199	170,756	329,443	0.0611(0.0026)	30718901[46]
Anorexia Nervosa	AN	72,517	16,992	55,525	0.1802(0.0124)	31308545[47]
Obsessive Compulsive Disorder	OCD	9,725	2,688	7,037	0.2918(0.0516)	28761083[48]
Schizophrenia	SCZ	130,644	77,258	53,386	0.3637(0.0126)	35396580[49]
Bipolar Disorder	BD	413,466	41,917	371,549	0.0708(0.0028)	34002096[50]
Cognitive Performance	CP	257,828			0.1909(0.0069)	30038396[51]
Lung Cancer	LC	85,716	29,266	56,450	0.0795(0.0134)	28604730[52]
Asthma	Asthma	385,822	44,301	341,521	0.0494(0.0042)	31427789[53]
Hypertension	Hypertension	298,307	99,665	189,642	0.1318(0.0056)	31427789[53]
Heart Rate	HR	361,411			0.1337(0.0081)	31427789[53]
Body Mass Index	BMI	806,834			0.1880(0.0055)	30239722[54]
Height	Height	385,748			0.4376(0.0194)	31427789[53]
Sleep Duration	SD	446,118			0.0676(0.0031)	30846698[55]
Neuroticism	NSM	329,821			0.1073(0.0046)	29255261[56]
Low-density Lipoprotein Cholesterol	LDL	431,167			0.1990(0.0225)	32493714[57]
Educational Attainment	EA	765,283			0.1285(0.0032)	353361970[58]
Coronary Artery Disease(including angina)	CAD	547,261	122,733	424,528	0.0578(0.0035)	29212778[59]
Waist/hip Ratio	WHR	697,734			0.1325(0.0045)	30239722[54]
Rheumatoid Arthritis	RA	58,284	14,361	58,284	0.1378(0.0189)	24390342[60]
Type 1 Diabetes	T1D	25,063	9,358	15,705	0.1824(0.0364)	32005708[61]

Table 2 Overview of the traits included in this study