

frequency, the resulting matrix provides a comparable measure of genetic similarity across loci with different allele frequency scales, thereby establishing a foundation for subsequent decomposition of genetic variance.

However, in real data, genetic effects are typically not uniformly distributed across all variant sites; instead, they are jointly influenced by allele frequency and linkage disequilibrium (LD) structure. Based on this understanding, a stratified GRM construction strategy can be further introduced, in which SNPs are grouped according to minor allele frequency (MAF) intervals or LD levels, and multiple sub-GRMs are constructed accordingly (i.e., the GREML-LDMS framework). This approach allows different classes of variants to contribute heterogeneously to genetic variance. By incorporating a more refined structural representation at the model level, this strategy helps mitigate fitting biases of the standard GRM under complex genetic architectures, thereby improving the interpretability and stability of genetic parameter estimates.

### **S1.3 Estimation procedures**

After the genomic relationship matrix is constructed, the estimation of SNP heritability mainly relies on two methodological pathways: individual-level data and summary statistics. When complete individual-level data are available, linear mixed models (LMMs) are typically used to decompose phenotypic variance, with the basic form given as:

$$y = X\beta + g + \varepsilon$$

Here,  $g$  represents the additive genetic effects captured by genome-wide SNPs and is typically assumed to follow a normal distribution with mean zero and a covariance structure defined by the genomic relationship matrix (GRM), i.e.,  $g \sim N(0, \sigma_g^2 G)$ . The residual term  $\varepsilon$  reflects the random error not explained by the model and satisfies  $\varepsilon \sim N(0, \sigma_e^2 I)$ . Within this framework, the genetic variance and environmental variance can be estimated using the restricted maximum likelihood (REML) method, from which SNP heritability can be further derived:

$$h_{\text{SNP}}^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

In practical applications, GCTA-GREML, BOLT-REML, and several closed-form estimation methods can all implement the above estimation process. While they differ in computational efficiency and model approximation, they are theoretically grounded in the same variance decomposition framework.

When individual-level data are unavailable, methods based on GWAS summary statistics provide an alternative approach. Among these, LD score regression (LDSC) estimates heritability by regressing the  $\chi^2$  statistics on LD scores, and its expected form can be expressed as:

$$E[\chi^2] = 1 + \frac{N h^2 l_j}{M}$$

This method utilizes the LD structure to weight summary statistics, thereby enabling heritability inference without requiring individual-level data. Furthermore, the SumHer method extends this framework by introducing weighting schemes that account for dependencies on minor allele frequency (MAF) and LD, allowing genetic effects to be non-uniformly distributed across different frequency ranges and LD regions. As a result, it generally exhibits greater flexibility under complex genetic architectures. It should be noted that the choice of LD reference structure has a substantial impact on estimation results; in practice, one may use external references (such as the 1000 Genomes Project) or, preferably, within-sample LD to improve matching and estimation accuracy.

### **S1.4 Model diagnostics and robustness analysis**

To ensure the reliability of heritability estimates, it is necessary to conduct systematic model diagnostics from multiple perspectives. First, from the standpoint of matrix properties, analyzing the eigenvalue spectrum of the genetic relationship matrix (GRM) can help identify potential numerical instabilities, such as near-singularity or anomalous structures. These issues often indicate the presence of inadequately controlled population structure or sample dependencies within the data. Second, by varying the composition of the SNP set—such as comparing the