are more sensitive to LD structure, such as SumHer, helps assess the degree to which the estimates depend on assumptions embedded in GRM construction (Speed and Balding, 2018; Speed et al., 2022).

However, even when model specification is appropriate, the statistical stability of the estimates still needs to be evaluated separately. The convergence of the REML algorithm, the magnitude of standard errors, and the width of confidence intervals are all key indicators for judging result reliability. In particular, when boundary solutions occur, such as genetic variance estimates approaching zero or reaching the upper bound of the parameter space, statistical explanations such as insufficient sample size or limited model information should be considered first, rather than assigning direct biological meaning to such results. For complex traits or studies with limited sample sizes, resampling methods such as jackknife or bootstrap can be used to evaluate estimation variability, and increasing sample size through multi-cohort joint analysis has also been shown to be an effective way to improve estimation precision (Evans et al., 2017; Wainschtein et al., 2022).

Given these multidimensional constraints, interpretation based on a single method is clearly limited. Cross-validating GREML estimates with other methods is therefore a key strategy for improving the robustness of conclusions in current research. Because different methods differ in how they capture genetic variance, their estimates for the same trait often show systematic deviations. Comparing individual-level GREML results with summary-statistics-based LDSC or SumHer estimates can help identify biases introduced by differences in data structure or model assumptions (Speed et al., 2016; Speed and Balding, 2018). Especially when SNP heritability is substantially lower than family-based heritability, the result should be interpreted comprehensively in terms of marker coverage, LD structure, non-additive genetic effects, and gene-environment interactions, rather than being simply attributed to methodological limitations or missing genetic information (Yang et al., 2017; Wainschtein et al., 2022).

In essence, SNP heritability estimated under the GREML framework is a quantitative expression of the "genetic variance identifiable under given data and model conditions." The interpretation of SNP heritability results should follow the standardized checklist shown in Supplementary Table S1. Only when data quality, model specification, statistical stability, and methodological consistency have all been adequately verified can the estimate serve as an important basis for understanding the genetic architecture of complex traits. Integrating statistical inference with the biological background of the trait and developing an interpretive pathway based on multiple lines of evidence has become a mainstream paradigm in contemporary statistical genetics.

## 7 Discussion

### 7.1 Implications of SNP-based heritability estimation for the "missing heritability" debate

The issue of "missing heritability" has long been a central debate in quantitative genetics and population genomics. Pedigree-based studies often report relatively high heritability estimates, whereas SNP-based approaches-such as GREML, which estimates genetic variance via the genome-wide relationship matrix-typically yield lower values (Evans et al., 2017; Yang et al., 2017). This discrepancy arises from multiple factors, including the limited capture of low-frequency and rare variants, incomplete linkage disequilibrium (LD) between genotyped markers and causal mutations, the omission of epistatic interactions, and the cumulative effects of numerous small-effect alleles underlying highly polygenic traits (Hou et al., 2019; Holland et al., 2020). In recent years, with the increasing use of whole-genome sequencing and the development of more refined LD-aware modeling approaches, this gap has narrowed to some extent. However, for highly polygenic traits, a portion of heritability remains unexplained (Evans et al., 2017; Hou et al., 2019).

Importantly, SNP-based heritability should not be interpreted simply as an underestimate of the true heritability, but rather as a quantitative characterization of the variance explained by the observed set of markers (Yang et al., 2017). This perspective has prompted a conceptual shift in how heritability is defined: the issue is not whether heritability is truly "missing," but whether the association between genotyped markers and causal variants is incomplete. Consequently, SNP-based heritability serves as an important indicator of the capture efficiency of genotyping platforms and provides a theoretical basis for designing higher-density genotyping strategies and improving the dissection of complex traits.