Appendix B. Comparative Summary of Statistical Objectives, Assumptions, and Applicability of GREML Extensions and Variants

Supplementary Table 2 summarizes the differences among commonly used extensions of the GREML framework in terms of their statistical objectives, core assumptions, applicable data structures, and key diagnostic considerations. This comparison table is intended to provide guidance for selecting appropriate extended models under different research scenarios.

Table S2 Statistical Objectives, Assumptions, and Applicability of GREML Extensions and Variants

Method extension	Primary statistical issue addressed	Core assumptions	Typical applicable data structures	Recommended diagnostics and cautions
LOCO (Leave-One-Chromosome-Out)	Proximal contamination, leading to inflation of local effects and overestimation of variance components	Genetic contributions from different chromosomes are approximately independent; excluding the target chromosome does not substantially compromise estimation of genome-wide background effects	Genomes with a limited number of chromosomes and large LD blocks; scenarios with high-effect loci; commonly applied in crop genomic datasets	Compare results from standard GRM and LOCO GRM; LOCO is designed primarily to address proximal contamination and should not be used as a general correction for population structure or LD heterogeneity
Functional partitioning of heritability (multiple GRMs)	Heterogeneous distribution of genetic variance across genomic functional regions; limited biological interpretability of aggregate heritability	SNP effect-size distributions and LD patterns differ systematically across functional annotations and can be identified through partitioned modeling	Large sample sizes; reliable functional annotations; sufficient SNP density within each category; suitable for enrichment and variance partitioning analyses	Highly sensitive to annotation correlation and multicollinearity; perform sensitivity analyses; avoid over-interpreting enrichment signals from individual categories
Bivariate / multi-trait GREML	Identifiability of genetic covariance and genetic correlation between traits; integration of information across traits	Genetic effects for multiple traits are captured by a common GRM; covariance structure is identifiable in the sample	Adequate sample size; traits measured in the same or highly overlapping samples; applications in multi-trait breeding and trade-off analyses	Sensitive to phenotypic measurement error and sample overlap; carefully inspect standard errors and confidence intervals of genetic correlations; avoid over-interpretation under limited sample sizes