Empirical studies have shown that in cross-ancestry GWAS analyses, such as those of lipid traits and type 2 diabetes, MsCAVIAR can significantly reduce the size of the credible set relative to single-study methods, often by about 20% or more, while still maintaining high causal coverage. This indicates that cross-study integration improves not only statistical power, but also the robustness of causal inference.

4.3 Methodological trade-offs: likelihood vs prior, robustness vs resolution

From a methodological standpoint, (Ms)CAVIAR and fastPAINTOR form a natural contrast within the unified framework:

(1) fastPAINTOR: improves identifiability by optimizing the prior

(2) CAVIAR / MsCAVIAR: improves identifiability by refining the likelihood

For CAVIAR and MsCAVIAR, one prominent advantage is that they do not rely on functional annotations and are therefore less sensitive to bias in external information. This property is particularly important when annotation resources are limited, when annotation quality is unstable, or when the study organism is not a model species with well-developed functional information. In comparison, these methods depend more directly on the LD structure itself and the statistical constraints it provides, and thus often show greater robustness when functional information is lacking.

A further advantage of MsCAVIAR lies in its use of cross-population differences in LD structure. LD patterns are not identical across populations, and this provides additional leverage for separating causal signals from non-causal correlated signals. For this reason, MsCAVIAR can improve resolution through cross-study integration in situations where single-study analysis is insufficient. This gain in information derived from heterogeneity is essentially a cross-LD decoupling mechanism that is not available within a single-study framework.

However, the main limitation of these methods is also clear, namely their relatively high computational cost. Because they require evaluation of high-dimensional causal configuration spaces, computational complexity increases rapidly as the number of candidate SNPs grows and the number of potential causal variants rises. In MsCAVIAR, this problem is further amplified by joint modeling across multiple studies (Lapierre et al., 2020). As a result, in practical applications, different methods tend to serve distinct roles. fastPAINTOR is better suited for genome-wide scanning and high-throughput analysis, whereas (Ms)CAVIAR is more appropriate for fine-scale dissection of candidate regions and validation studies under cross-population settings.

With continued algorithmic optimization and ongoing improvements in computational resources, cross-study fine-mapping is expected to play an increasingly important role in complex trait research and may gradually become part of the standard analytical workflow.

5 Colocalization and eQTL/TWAS Interfaces: A Multi-Trait Causal Inference Layer

5.1 From signal overlap to causal sharing: reformulating the problem

Although GWAS has identified a large number of loci associated with complex traits, these statistical signals do not directly reveal their underlying molecular mechanisms. Molecular quantitative trait loci (QTLs), including expression QTLs (eQTLs), splicing QTLs (sQTLs), and protein QTLs (pQTLs), provide critical insights into how genetic variants influence phenotypes through regulatory processes at the transcriptional and translational levels (Okamoto et al., 2023). In parallel, transcriptome-wide association studies (TWAS) aggregate genetic effects at the gene level by predicting gene expression, thereby prioritizing candidate genes.

It should be noted, however, that the overlap of GWAS signals with molecular QTL or TWAS signals within the same genomic region does not necessarily imply that they are driven by the same causal variant. Because linkage disequilibrium (LD) and allelic heterogeneity may coexist, multiple highly correlated variants within a region can affect different phenotypes separately, thereby producing overlapping signals at the statistical level without reflecting true causal consistency. This phenomenon is commonly referred to as false colocalization. Colocalization analysis is designed precisely to address this question by modeling cross-trait causal sharing within a probabilistic framework (Giambartolomei et al., 2013; Hormozdiari et al., 2016).