

Within this framework, the various methods are not simple substitutes for one another, but rather different steps that progressively constrain and refine the same causal problem at different inferential layers (Schaid et al., 2018; Hutchinson et al., 2020). Among them, fastPAINTOR and CAVIAR mainly act on the causal space of a single phenotype, MsCAVIAR further addresses causal consistency across studies, and colocalization analysis extends the inference to cross-trait causal consistency.

The value of this layered inferential strategy lies in its ability to maximize the integration of information from different sources, including LD structure, functional annotations, cross-study data, and cross-omics information, while seeking a balance between computational efficiency and inferential precision and reducing, as much as possible, the systematic bias that may arise from any single method. In this sense, method selection should not be understood as a comparison among interchangeable techniques, but rather as a layered decision-making process centered on the same causal inferential objective.

## **7 Discussion: A Unified Framework from Association Signals to the Causal Inference Layer**

### **7.1 From method complementarity to a unified statistical framework**

Existing fine-mapping methods are often regarded as a collection of tools developed for different research scenarios. However, from the unified perspective of statistical genetics, these methods all essentially perform inference on the same causal configuration space under different informational constraints. The differences among fastPAINTOR, CAVIAR, and its extension MsCAVIAR can be understood as reflecting different emphases on distinct components of the Bayesian model: fastPAINTOR primarily strengthens the prior structure through functional annotations (Kichaev et al., 2014, 2016), CAVIAR reinforces likelihood constraints through explicit modeling of LD structure (Schaid et al., 2018), and MsCAVIAR further introduces additional data constraints through cross-study integration (Lapierre et al., 2020).

Accordingly, the complementarity among these methods should not be viewed as a simple empirical combination, but rather as the result of progressively strengthening constraints on the same causal estimand across different informational dimensions. Within this framework, method selection is no longer merely a matter of choosing among tools, but instead involves determining which path most effectively approximates the causal probability distribution under given data conditions. In this sense, fine-mapping methods should not be regarded as isolated techniques, but as different implementations within a unified statistical framework.

### **7.2 Dual pathways of information integration: prior enrichment and data expansion**

The key to improving fine-mapping resolution lies in increasing the identifiability of causal configurations. In the current methodological landscape, this objective is pursued mainly along two paths. One path relies on strengthening prior information through the incorporation of functional annotations. Methods represented by fastPAINTOR transform biological information, such as eQTLs, ATAC-seq, and ChIP-seq data, into informed priors, thereby increasing the posterior probability of candidate causal variants (Kichaev et al., 2016; Zou et al., 2021). When annotation quality is high, this strategy can not only shrink credible sets but also enhance the biological interpretability of the results. However, this path is highly dependent on annotation quality. Once functional information is biased or incomplete, it may lead to prior misspecification and consequently affect posterior inference (Kichaev et al., 2016).

The other path takes the form of expanding the observed data, particularly through the integration of multi-population or multi-study datasets to exploit differences in LD structure and thereby enhance the identification of causal variants. Methods such as MsCAVIAR follow this logic by jointly modeling summary signals from different studies, thus improving both the robustness and the resolution of causal inference (Lapierre et al., 2020). Compared with annotation-dependent methods, this approach does not directly rely on external functional information and can more effectively eliminate spurious signals driven by LD. Its cost, however, is that it places greater demands on data quality, comparability across studies, and computational resources (Schaid et al., 2018).