characterizes the evidence of association between loci and traits, SNP heritability reflects the proportion of phenotypic variance explained at the population level under a given model and marker set, and polygenic risk scores (PRS/PGS) further integrate these effects into predictive functions at the individual level (Fang and Wu, 2026; Fang, 2026a; 2026b). Nevertheless, this inferential chain remains incomplete at the level of causal interpretation. In particular, when LD structure is complex and multiple causal variants may coexist within a region, the central question is no longer merely whether a genomic region is associated with a trait, but rather which variants are more likely to be truly causal and how much uncertainty is attached to such judgments. Fine-mapping emerged precisely in response to this problem. Its objective is not to repeat the screening of significant loci, but to provide a probabilistic characterization of the plausibility of causal variants under given data and modeling assumptions, thereby introducing causal probability as a more informative inferential target in statistical genetics.

To address this objective, fine-mapping generally takes posterior inclusion probability (PIP) and credible sets as its core analytical quantities. Unlike conventional GWAS, which mainly relies on significance thresholds to make binary decisions, fine-mapping emphasizes continuous probabilistic estimation of causal configurations conditional on the observed data and model assumptions. As a result, the focus of inference shifts from whether an association exists to which candidate variants should be retained and how likely each of them is to be causal. Within this framework, a credible set is no longer treated as a subsidiary result centered on a single candidate locus, but rather as the smallest set of candidate variants constructed under a predefined coverage probability. This approach directly incorporates uncertainty into the inferential process and provides a more operational basis for subsequent functional validation and cross-omics integration (Kichaev et al., 2016; Hutchinson et al., 2020).

However, moving from association signals to causal probability distributions is far from straightforward. First, complex LD structure induces strong statistical correlations among multiple SNPs, giving rise to substantial non-identifiability and making it difficult to isolate the true causal variant from a group of correlated variants based on any single statistic alone. Second, the presence of multiple causal variants is not uncommon in complex traits, which means that traditional stepwise regression or conditional analysis cannot always identify independent effects in a stable manner and may instead lead to incorrect inference when the model is inadequately specified (Hutchinson et al., 2020). Therefore, the development of fine-mapping should not be viewed merely as a further refinement of analytical procedures. More fundamentally, it requires probabilistic models that explicitly characterize LD structure and the space of causal configurations while also allowing uncertainty to be properly represented and propagated.

In recent years, Bayesian probabilistic mapping approaches have gradually become the dominant framework in this field. These methods typically perform joint modeling of GWAS summary statistics and LD structure, while incorporating prior distributions to estimate the posterior probability that each candidate variant is causal, thereby providing a probabilistic representation of the causal architecture within local genomic regions. Although these methods all pursue the same general objective, they differ in their modeling emphasis. fastPAINTOR constructs annotation-informed priors by integrating functional annotations and improves computational efficiency through approximate Bayesian inference, making it more suitable for large-scale and multi-trait analyses. By contrast, CAVIAR and its extension MsCAVIAR do not rely on functional annotation, but instead model LD structure and multiple causal configurations directly, while improving the consistency and resolution of causal localization through the integration of cross-study or multi-ancestry data (Lapierre et al., 2020). Compared with traditional single-variant analysis, these approaches show clearer advantages in statistical power, the compactness of credible sets, and the consistency of results across cohorts (Hutchinson et al., 2020).

On this basis, colocalization analysis further establishes the connection between fine-mapping and functional genomics. When GWAS and molecular QTL signals overlap within the same genomic region, the key question is not simply whether they co-occur, but whether that overlap is driven by the same causal variant. Colocalization analysis is designed precisely to address this issue. Under a unified probabilistic framework, it evaluates the likelihood that GWAS signals and molecular QTL signals, such as eQTLs, sQTLs, pQTLs, or TWAS signals,