share the same causal variant, thereby distinguishing true causal concordance from superficial overlap generated solely by LD. With the development of multi-causal models, methods such as SuSiE have further improved the precision of causal decomposition and colocalization analysis, enabling cross-omics integration to move beyond simple signal overlap toward a more rigorous probabilistic characterization of causal consistency.

Against this background, the present study approaches fine-mapping from the perspective of the overall inferential framework of statistical genetics and interprets it as a key layer of causal inference situated between association discovery, variance explanation, and individual-level prediction. Focusing on representative methods including fastPAINTOR, CAVIAR, and MsCAVIAR, this study compares their differences in modeling assumptions, applicable boundaries, and computational characteristics, further clarifies the methodological meaning of PIP and credible sets as inferential targets for causal probability, and discusses their relationship with colocalization analysis. On this basis, the study seeks to propose a method-selection framework that balances computational efficiency, localization accuracy, and data conditions, thereby providing a more operational basis for empirical analysis in different research settings. From this perspective, fine-mapping should no longer be viewed merely as an auxiliary technique applied after GWAS, but rather as a core inferential step through which statistical genetics moves from association evidence toward causal interpretation.

2 Statistical Foundations of Fine-Mapping

2.1 Causal probabilities and credible sets: from association statistics to causal inference targets

The primary objective of fine-mapping is not merely to identify statistically significant loci, but to characterize, under complex linkage disequilibrium (LD) structures and multi-causal architectures, which variants are causal and how uncertainty is distributed across them. From this perspective, the inferential target of fine-mapping departs fundamentally from GWAS association statistics and instead becomes a new statistical object: the posterior distribution over causal configurations (i.e., a causal estimand).

In the conventional GWAS framework, single-variant tests provide evidence that a genomic region harbors an association signal but cannot distinguish true causal variants from correlated LD proxies (Spain and Barrett, 2015; Schaid et al., 2018). To address this limitation, Bayesian fine-mapping introduces the posterior inclusion probability (PIP) as a core quantity, defined as the probability that a given variant is causal conditional on the observed data and LD structure:

$$\mathrm{PIP}_j = P(\gamma_j = 1 \mid \mathrm{Data}, \mathrm{LD})$$

where γ_j is an indicator of whether SNP j is causal. Within this framework, PIP is not a surrogate for statistical significance but a formally defined estimand that quantifies causal probability under explicit model assumptions (Hutchinson et al., 2020).

By ranking SNPs according to their posterior inclusion probabilities (PIPs) and cumulatively summing them until a predefined coverage threshold, such as 95%, is reached, one can construct the smallest candidate variant set, namely the credible set (Hutchinson et al., 2020; Shrestha et al., 2024). Under the assumption that the model is correctly specified, this set contains the true causal variant with a given probability. Unlike the traditional point-estimation logic that directly treats the most significant SNP as the causal variant, the credible set embodies a set-based inferential framework. Rather than forcing a single locus to bear the full burden of causal interpretation, it retains, under a probabilistic constraint, a candidate space capable of covering the true causal variant. In this way, uncertainty in causal inference is explicitly incorporated into the statistical analysis. From the perspective of the broader framework of statistical genetics, GWAS primarily provides evidence of association between loci and traits, SNP heritability corresponds to the explanation of variance at the population level (Fang, 2026a), and PRS serves individual-level prediction (Fang, 2026b), whereas PIP and the credible set describe the causal probability distribution at the locus level, thereby forming a key intermediate layer between association analysis and functional interpretation (Figure 1).