

Research Report

Open Access

The Causal Inference Layer in Complex Trait Genetics: A Unified Statistical Framework from Fine-Mapping to Cross-Trait Integration

Xuanjun Fang 

Hainan Provincial Key Laboratory of Crop Molecular Breeding, Hainan Institute of Tropical Agricultural Resources (HITAR), Sanya, 572025, Hainan, China

 Corresponding email: xuanjunfang@hitar.orgBioscience Evidence, 2026, Vol.16, No.3 doi: [10.5376/be.2026.16.0015](https://doi.org/10.5376/be.2026.16.0015)

Received: 22 May, 2026

Accepted: 22 Jun., 2026

Published: 30 Jun., 2026

Copyright © 2026 Fang, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Fang X.J., 2026, The causal inference layer in complex trait genetics: a unified statistical framework from fine-mapping to cross-trait integration, Bioscience Evidence, 16(3): 186-201 (doi: [10.5376/be.2026.16.0015](https://doi.org/10.5376/be.2026.16.0015))

Abstract Genome-wide association studies (GWAS) have identified a large number of loci associated with complex traits and diseases. However, most of these signals arise from linkage disequilibrium (LD) rather than directly reflecting causal variants, thereby limiting their mechanistic interpretability. Probabilistic fine-mapping addresses this limitation by introducing posterior inclusion probabilities (PIPs) and credible sets, shifting the inferential target from a single significant locus to a distribution of causal probabilities and enabling a systematic characterization of genetic uncertainty. In recent years, with the growing availability of functional annotation data, multi-ancestry studies, and multi-omics resources, fine-mapping methods have continued to expand in both model architecture and application scope. Nevertheless, a unified theoretical perspective across these methods remains lacking. In this study, we develop a unified statistical framework centered on causal configurations by systematically integrating fine-mapping and colocalization analyses within a Bayesian inferential framework. Under this framework, fastPAINTOR constructs annotation-informed priors using functional annotations, whereas CAVIAR and its extension McCAVIAR strengthen likelihood-based constraints through LD structure and cross-study information. Colocalization analysis further extends the inferential target from a single-trait setting to a multi-trait space, enabling probabilistic modeling of cross-trait causal consistency. Accordingly, research on complex traits can be organized into a continuous inferential pipeline from GWAS to fine-mapping, and further to colocalization and transcriptome-wide association studies (TWAS), thereby progressively translating statistical associations into biological mechanistic interpretation. On this basis, we further propose a method-selection strategy based on inferential hierarchy, clarifying the trade-offs among computational complexity, causal resolution, and information sources across different methods, and summarizing a practical workflow of “hierarchical inference.” This framework is applicable not only to studies of human complex diseases, but also to applied contexts such as crop genetic improvement, where it can be used to assess causal consistency across environments and populations. By unifying fine-mapping and colocalization within the same causal inference layer, this study provides statistical genetics with a consistent conceptual language and analytical paradigm, thereby facilitating the systematic transition of complex trait research from association discovery to mechanistic interpretation and causal inference.

Keywords Complex traits; Fine-mapping; Causal inference; Colocalization analysis; Credible sets; Multi-omics integration

1 Introduction

Genome-wide association studies (GWAS) have substantially advanced the dissection of the genetic architecture of complex traits and diseases over the past decade, leading to the identification of numerous genomic regions associated with phenotypic variation in both human populations and crop systems (Lapierre et al., 2020). However, the evidence provided by GWAS is primarily statistical association rather than direct causal interpretation. Because linkage disequilibrium (LD) is widespread across the genome, significantly associated loci often function only as marker SNPs correlated with the true functional variants, rather than representing the actual causal variants with biological effects. At the same time, multiple independent causal signals may coexist within a single associated region, making the assumption that the most significant SNP directly corresponds to the causal variant statistically untenable in many cases (Hutchinson et al., 2020). Accordingly, how to distinguish causal signals from correlated signals within associated regions has become an important issue linking population genetic analysis with functional genomic interpretation.

From the perspective of the development of statistical genetics, genomic analysis of complex traits has not been built upon a single statistic, but has instead progressed around a series of different inferential targets. GWAS

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,

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:

$$\text{PIP}_j = P(\gamma_j = 1 | \text{Data}, \text{LD})$$

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).

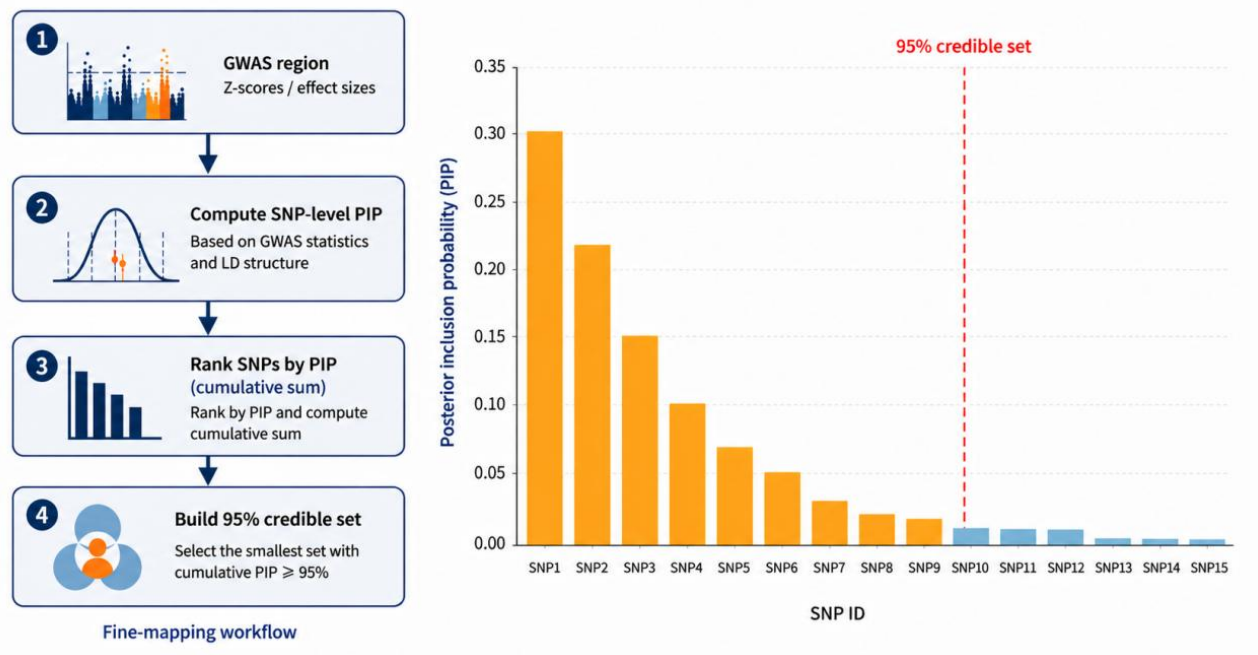


Figure 1 Workflow and illustrative example of credible set construction based on posterior inclusion probability (PIP)

Note: The left panel illustrates the core statistical workflow of fine-mapping: starting from GWAS summary statistics (e.g., Z-scores or effect sizes) within a genomic region, SNP-level posterior inclusion probabilities (PIPs) are computed, ranked, and cumulatively summed to construct a credible set under a predefined coverage threshold (e.g., 95%). The right panel presents an example PIP distribution, where each bar represents the PIP of a candidate SNP, and the dashed line indicates the cutoff at which the cumulative probability reaches 95%

In practical applications, the size of a credible set is not fixed, but is jointly determined by the complexity of local LD structure and the number of underlying causal variants. When LD relationships are relatively simple and only a single causal variant exists within a region, the credible set is usually compact. By contrast, under high-LD backgrounds or in the presence of multiple causal variants, the candidate set often expands substantially, which essentially reflects increased statistical non-identifiability (Hutchinson et al., 2019). In recent years, calibrated credible sets have further improved coverage estimation, bringing it closer to the underlying causal architecture (Shrestha et al., 2024). Therefore, this probability-set-based analytical framework not only enhances the interpretability of fine-mapping results, but also provides a relatively unified input form for functional annotation integration and cross-omics analysis.

2.2 Paradigm shift from single-variant testing to probabilistic inference

From the perspective of statistical inference, the difference between fine-mapping and traditional single-variant GWAS testing is not merely a refinement of analytical steps, but a fundamental shift in both the inferential target and the object of inference. Traditional GWAS mainly relies on single-SNP tests and determines whether a locus is associated with a trait through a significance threshold (Schaid et al., 2018). This approach is efficient for genome-wide scanning, but it rests on a strong simplifying assumption, namely, that the most significant SNP can serve as a proxy for the causal variant. This assumption often fails in the presence of complex LD structure or multiple causal signals, thereby leading to incorrect attribution or bias in the prioritization of candidate loci.

In contrast, Bayesian fine-mapping does not attempt to select a single “best” locus directly from significance results. Rather, it reformulates the problem as the estimation of the posterior distribution over all possible causal configurations, conditional on the observed data and the LD structure. Within this framework, PIP provides a continuous probabilistic characterization of the causal plausibility of an individual locus, whereas the credible set defines the candidate space under a prespecified probability constraint. For example, if three SNPs within a region have PIPs of 0.60, 0.25, and 0.10, respectively, they jointly constitute a 95% credible set, rather than only the

most significant locus being retained (Hutchinson et al., 2020). This treatment more adequately reflects the uncertainty of local signals and substantially reduces the risk of prioritizing false positives driven by LD structure.

More importantly, once statistical inference is advanced from significance testing to probabilistic characterization, the output of fine-mapping is no longer limited to the interpretation of local association signals, but can be further integrated with other statistical modules. PIP and credible sets can be used for functional annotation integration, support cross-population comparisons, and facilitate colocalization analysis, thereby maintaining a relatively consistent measurement basis across different data domains (Kichaev et al., 2014; Gerber et al., 2023). For this reason, fine-mapping should not be understood merely as a downstream filtering step following GWAS, but rather as an important inferential layer linking association analysis to mechanistic interpretation.

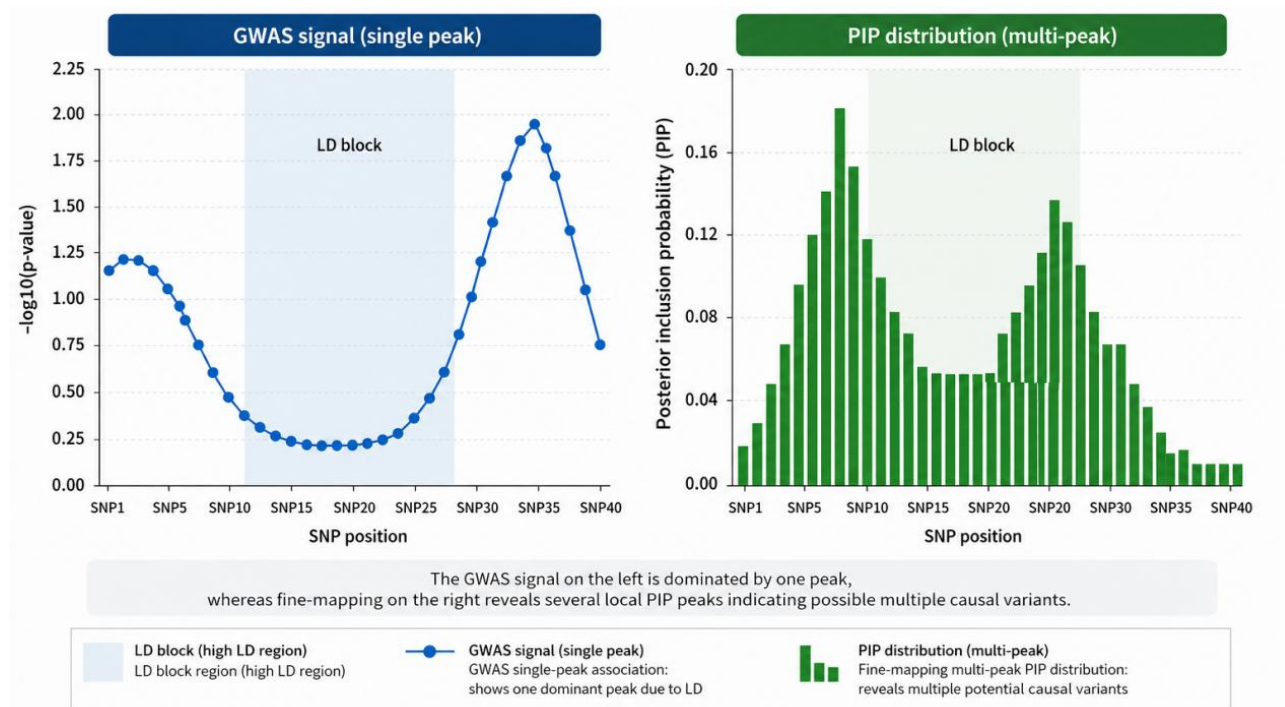


Figure 2 p-value vs PIP under an LD block: contrast between single-peak association signals and multi-peak causal probability distributions

Note: The left panel shows GWAS association signals represented by $-\log_{10}(p)$, typically characterized by a single dominant peak within an LD block. The right panel shows the corresponding posterior inclusion probability (PIP) distribution obtained from fine-mapping, revealing multiple local peaks that reflect potential causal variants. The shaded region denotes the LD block

2.3 Theoretical implications: from point estimates to causal distributions

From a more general statistical perspective, the central significance of fine-mapping lies in its transformation of causal inference from the identification of a single locus into the estimation of a posterior distribution. A credible set can be understood as the smallest set estimator constructed under a coverage probability constraint, and this logic is highly consistent with that of interval estimation and confidence set inference (Hutchinson et al., 2019; 2020). Within this framework, the effects of LD structure and the presence of multiple causal variants are no longer treated as nuisances in the analysis, but are instead directly incorporated into the model, allowing causal inference to shift from deterministic judgment to probabilistic characterization. Compared with approaches that rely on a single significant locus to draw conclusions, this strategy more faithfully captures the uncertainty inherent in the genetic architecture of complex traits and avoids excessive simplification of local association signals.

At the same time, once causal probabilities are expressed in the form of PIP and credible sets, they can be propagated into downstream analyses, thereby providing a common probabilistic basis for integrative studies at different levels. For example, in joint analyses of GWAS and eQTL data, researchers may compare the overlap of

credible sets or examine joint posterior probabilities to assess whether different signals are likely to share the same causal variant. In this sense, fine-mapping is not merely a tool for improving localization accuracy within a genomic region; it also serves to transmit uncertainty into subsequent analyses, enabling functional annotation integration, colocalization analysis, and cross-omics research to proceed on a unified probabilistic scale.

In addition, because PIP is fundamentally defined in probability space and does not depend on the form of a specific statistical measure, such as a p-value or effect size, it is more readily comparable across different study designs, populations, and omics data types. This property allows it to serve as a common language for expressing causal information across data domains and provides a methodological basis for cross-study integration (Shrestha et al., 2024). This is particularly valuable in multi-ancestry studies, where differences in LD structure across populations can be regarded as a naturally occurring additional constraint. Different LD patterns alter the correlations between non-causal variants and true causal loci, thereby helping to further shrink credible sets and improve causal resolution (Spain and Barrett, 2015). This feature also gives fine-mapping greater practical value in multi-ancestry research and naturally extends to the development of subsequent methods such as MsCAVIAR.

3 fastPAINTOR: An Annotation-Informed Model for Causal Probability

3.1 Model framework: joint inference of statistical evidence and functional annotation

Within the unified Bayesian framework of fine-mapping, the key distinction among methods lies not in whether probabilistic inference is performed, but in how the prior distribution over causal configurations is specified. From this perspective, fastPAINTOR can be understood as a prototypical *annotation-informed causal model*, whose central objective is to establish a probabilistic mapping between GWAS statistical evidence and functional annotations.

In general form, the fine-mapping problem can be expressed as estimating the posterior distribution over the causal configuration vector γ :

$$p(\gamma|D,R,A) \propto p(D|\gamma,R) p(\gamma|A)$$

From the perspective of statistical genetics, a basic fine-mapping model characterizes the distribution of causal probabilities under LD constraints, whereas fastPAINTOR further describes an annotation-informed causal probability distribution conditional on functional annotations, thereby giving its inferential target stronger biological interpretability. To estimate the posterior distribution within this framework, fastPAINTOR adopts approximate Bayesian methods, such as variational inference or importance sampling, in order to avoid the exponential enumeration of the full causal configuration space or the high computational cost associated with traditional Markov chain Monte Carlo (MCMC) approaches (Talukdar et al., 2023). Its optimization objective can be expressed as the evidence lower bound (ELBO):

$$\Pr_{\tilde{q}}(\gamma_j=1|a_j) = \text{logit}^{-1}(\alpha_0 + \alpha^T a_j)$$

where a_j denotes the annotation vector for SNP j (e.g., eQTL status, chromatin accessibility, transcription factor binding sites). This formulation effectively transforms biological functional potential into statistical prior probability, enabling a principled integration of functional and association evidence (Kichaev et al., 2014).

From the perspective of statistical genetics, a basic fine-mapping model characterizes the distribution of causal probabilities under LD constraints, whereas fastPAINTOR further describes an annotation-informed causal probability distribution conditional on functional annotations, thereby giving its inferential target stronger biological interpretability. To estimate the posterior distribution within this framework, fastPAINTOR adopts approximate Bayesian methods, such as variational inference or importance sampling, in order to avoid the exponential enumeration of the full causal configuration space or the high computational cost associated with traditional Markov chain Monte Carlo (MCMC) approaches (Talukdar et al., 2023). Its optimization objective can be expressed as the evidence lower bound (ELBO):

$$\mathcal{A}(q) = \mathbb{E}_{q(\gamma)}[\log p(D|\gamma,R) + \log p(\gamma|A)] - \mathbb{E}_{q(\gamma)}[\log q(\gamma)]$$

This approximation achieves a favorable balance between computational efficiency and inferential accuracy, enabling scalable application to genome-wide datasets and multi-trait analyses.

3.2 Methodological features: information integration, scalability, and resolution

From a methodological perspective, the strengths of fastPAINTOR are first reflected in its ability to directly incorporate external functional information into the modeling of causal probabilities. Variants located in enhancer or promoter regions, or supported by eQTL evidence, are assigned higher prior weights by the model. As a result, the posterior inclusion probability (PIP) reflects not only the strength of statistical association, but also signals of functional relevance, thereby giving candidate causal variants greater biological interpretability (Kichaev et al., 2016).

At the same time, fastPAINTOR also shows strong scalability at the computational level. By adopting approximate Bayesian inference, the method avoids the computational bottleneck of MCMC in high-dimensional causal configuration spaces, allowing it to maintain relatively high efficiency under large-scale GWAS data and multiple annotation settings. This property is particularly important for current large-sample studies represented by biobank-scale datasets.

On this basis, the incorporation of functional annotations can further improve the resolution of causal localization. In particular, under multi-trait or multi-annotation settings, fastPAINTOR can often effectively reduce the size of the credible set. Previous studies have shown that, while maintaining relatively high coverage, the inclusion of functional annotations can reduce the candidate set by approximately 40%~60% (Kichaev et al., 2016). For studies in which downstream experimental validation resources are limited, such compression of the candidate space has direct practical value.

When these features are considered within a unified framework, the main advantage of fastPAINTOR can be understood as arising from the optimization of prior structure. In other words, when the likelihood model remains unchanged, the introduction of functional annotations improves the identifiability of the causal probability distribution, thereby enhancing the model's ability to distinguish true causal variants.

3.3 Model limitations and sources of bias

Although fastPAINTOR has clear advantages in information integration and computational efficiency, its inferential performance also depends heavily on the validity of the prior model itself. As a result, while this method improves interpretability, it also introduces new methodological risks.

3.3.1 Prior misspecification

When functional annotations are noisy, incomplete, or biased, the prior distribution may distort posterior inference, leading to over-prioritization of non-causal variants or underestimation of true causal variants. This issue can be viewed as a prior-induced shift in the causal estimand and may require mitigation through multi-annotation integration or sensitivity analyses.

3.3.2 Limited transferability across populations

Functional annotations are often tissue-specific or population-specific. In cross-ancestry studies or non-model organisms, annotation information may not generalize, potentially reducing model performance. Consequently, interpretation of results in such contexts should be approached with caution.

3.3.3 Approximation error

While variational inference improves scalability, it may underestimate posterior uncertainty in regions with complex LD structure or multiple causal variants. This can result in overly compact credible sets or the omission of true causal variants.

4 (Ms)CAVIAR: Likelihood-Driven Modeling of Cross-Study Causal Consistency

4.1 CAVIAR: an LD-driven model of causal configurations

Within the unified Bayesian framework of fine-mapping, CAVIAR (CAusal Variants Identification in Associated Regions) represents a class of models in which inference is driven primarily by the likelihood structure. In contrast to fastPAINTOR, which incorporates functional annotations through an explicit prior, CAVIAR relies solely on GWAS summary statistics and LD structure to infer the posterior distribution over causal configurations, without invoking external biological information (Lapierre et al., 2020).

Formally, CAVIAR models the joint distribution of GWAS Z-scores. For a genomic region containing m SNPs, let $Z=(Z_1, \dots, Z_m)$ denote the vector of marginal association statistics. Conditional on a causal configuration γ , the model assumes:

$$Z \sim N(0, \Sigma + \sigma^2 I)$$

where Σ is the covariance matrix determined by LD structure, and σ^2 captures residual variance. The key feature of this model lies in its explicit incorporation of LD structure into the likelihood function, thereby allowing correlated signals and causal signals to be distinguished within the same statistical framework.

Under this framework, CAVIAR evaluates (either exactly or approximately) the space of possible causal configurations and computes posterior inclusion probabilities (PIPs) for each SNP, from which credible sets are constructed under a specified coverage constraint (e.g., 95%). Importantly, the model allows for multiple causal variants ($K \geq 1$) within a locus, thereby addressing the “signal ambiguity” that arises in LD-rich regions.

From a unified perspective, CAVIAR characterizes a causal probability distribution that is primarily determined jointly by LD structure and association statistics, namely an LD-driven causal estimand. Compared with models that rely on annotation information, it is not directly affected by the quality of external functional data. As a result, it often shows greater robustness when functional annotations are unavailable, when the annotations themselves may be biased, or when the study is conducted in a cross-species context.

4.2 MsCAVIAR: extending the likelihood to cross-study causal consistency

Building on the single-study formulation, MsCAVIAR (Multiple Study CAVIAR) extends the inferential target from identifying causal variants within a single study to assessing causal sharing across studies or populations.

This extension fundamentally alters the causal estimand: rather than estimating the posterior probability that a variant is causal within a single dataset, MsCAVIAR estimates the posterior probability that a variant is jointly causal across multiple studies.

To achieve this, MsCAVIAR introduces a random-effects model to capture heterogeneity in effect sizes across studies. Let β_{js} denote the effect of SNP j in study s , then:

$$\beta_{js} \sim N(\mu_j, \tau_j^2)$$

Where μ_j represents the shared (mean) effect across studies and τ_j^2 captures cross-study heterogeneity. By jointly modeling summary statistics from multiple studies together with their respective LD structures, MsCAVIAR can estimate not only study-specific PIPs, that is, within-study causal probabilities, but also the shared causal probability across studies.

The importance of this modeling strategy lies in the fact that differences in LD structure across populations are no longer treated merely as background conditions of the analysis, but rather as a source of information that can be actively exploited. If the correlation structure of a given SNP differs across populations, then the association strength generated by non-causal variants will usually fluctuate with changes in LD patterns, whereas true causal variants are more likely to exhibit relatively consistent signals across studies. MsCAVIAR takes advantage of this contrast to eliminate spurious signals and further reduce the size of the credible set (Lapierre et al., 2020).

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).

From the perspective of this study, if fine-mapping characterizes the within-trait causal probability distribution (via PIP), then colocalization extends this to: the joint posterior distribution over shared causal configurations across traits, thereby forming a critical bridge between locus-level causal inference and mechanistic interpretation.

5.2 Bayesian foundations: from single-causal to multi-causal models

The statistical foundation of colocalization lies in evaluating competing causal hypotheses under a Bayesian framework. A canonical example is COLOC, which defines five mutually exclusive hypotheses within a genomic region:

- H_0 : neither trait is associated
- H_1 : only the GWAS trait is associated
- H_2 : only the QTL trait is associated
- H_3 : both traits are associated, but with distinct causal variants
- H_4 : both traits share the same causal variant

Posterior probabilities $P(H_i|\text{data})$ are computed for each hypothesis, with $P(H_4|\text{data})$ (commonly denoted as PP4) serving as the primary measure of colocalization evidence (Giambartolomei et al., 2013).

The strength of this framework lies in the fact that it does not rely on a single lead SNP for inference. Instead, it makes use of the overall information contained in the credible set at the regional level and explicitly incorporates uncertainty into the inferential process. Therefore, compared with approaches that rely solely on peak overlap, COLOC provides a more rigorous statistical interpretation. However, an important assumption of COLOC is that each trait has at most one causal variant within the region.

In studies of complex traits, multiple causal signals are not uncommon, and this single-causal-variant assumption is often overly restrictive. For this reason, a series of extended methods have been developed to accommodate more complex genetic architectures. eCAVIAR allows multiple causal variants within a region and combines LD structure to calculate the colocalization posterior probability (CLPP); moloc extends the colocalization framework to joint modeling of multiple traits; and HyPrColoc is further designed for high-dimensional multi-trait colocalization analysis. Although these methods differ in their specific implementations, they share a common conceptual basis: all of them extend the causal configuration space in fine-mapping into a joint space across multiple phenotypes. From a unified perspective, colocalization can therefore be understood as a further extension of fine-mapping along the phenotypic dimension.

5.3 Interface with fine-mapping: from variant-level to cross-trait inference

Colocalization analysis is naturally coupled with fine-mapping methods such as fastPAINTOR and (Ms)CAVIAR. Their relationship can be summarized as follows:

fine-mapping provides variant-level posterior probabilities (PIPs) and credible sets for each trait
 colocalization evaluates whether these probabilities imply shared causality across traits

Formally, cross-trait causal sharing can be conceptualized as a function of trait-specific PIPs and LD structure:

$$P(\text{shared causal}) \approx f(\text{PIP}_{\text{GWAS}}, \text{PIP}_{\text{QTL}}, \text{LD})$$

In practical research, this relationship is usually implemented as a two-stage inferential workflow. The first stage operates at the locus level, where researchers apply fastPAINTOR or (Ms)CAVIAR to perform fine-mapping for a single phenotype, thereby obtaining high-resolution PIP distributions and credible sets. The second stage then moves to the cross-phenotype level, where methods such as COLOC, eCAVIAR, or moloc are used, based on the results of the first stage, to evaluate the probability of colocalization between GWAS signals and QTL or TWAS signals.

The practical value of this analytical workflow lies in the clear complementarity among these methods. fastPAINTOR primarily reduces the candidate space by integrating annotation information, whereas (Ms)CAVIAR improves inferential robustness through LD structure and cross-population differences.

Colocalization analysis further examines causal consistency through joint inference across phenotypes. When combined with TWAS-based integration at the gene level, this workflow ultimately forms a complete inferential chain from GWAS to fine-mapping, then to colocalization and TWAS, and finally to candidate gene identification:

GWAS→fine-mapping→colocalization→TWAS→candidate gene identification

This continuous analytical path indicates that the focus of statistical genetics has moved beyond the identification of association signals and is gradually shifting toward the systematic interpretation of molecular mechanisms (Hormozdiari et al., 2016; Okamoto et al., 2023).

6 Practical Guidelines for Method Selection: A Layer-Based Decision Framework

6.1 A unified view of trade-offs: computation, identifiability, and information sources

In fine-mapping and cross-omics analysis, method selection is not simply a matter of tool preference, but rather the result of a systematic trade-off among different inferential strategies. At its core, this trade-off usually involves three mutually constraining aspects: computational complexity, causal resolution or identifiability, and the structure of information sources, including LD, functional annotations, and cross-study data. Within a unified Bayesian framework, the differences among methods can essentially be understood as different ways of imposing constraints on the same causal configuration space by drawing on different types of information (Schaid et al., 2018; Hutchinson et al., 2020).

(1) FastPAINTOR: improving resolution through prior information

The core of fastPAINTOR lies in its use of functional annotations to construct informed priors, thereby increasing the identifiability of causal variants while keeping the likelihood structure, that is, the LD structure, unchanged (Kichaev et al., 2014, 2016). The strengths of this method are mainly reflected in two respects. On the one hand, it can operate efficiently at the genome-wide scale and is therefore suitable for large-scale analyses. On the other hand, when functional annotation information is sufficiently informative, it can substantially shrink the credible set and thus improve the ranking resolution of candidate causal variants. Previous studies have shown that integrating functional annotations not only reduces the size of the candidate set but also improves the prioritization of causal variants (Kichaev et al., 2016; Zou et al., 2021).

However, this very advantage also defines its limitation. The inferential results of fastPAINTOR are highly sensitive to the quality of annotations. Once functional annotations contain noise, missingness, or systematic bias, these defects may be propagated into the posterior inference and introduce systematic error. From a statistical perspective, fastPAINTOR is therefore essentially an annotation-informed prior model that improves causal resolution by relying on prior information.

(2) CAVIAR: robustness through likelihood modeling

Unlike fastPAINTOR, CAVIAR does not rely on functional annotations, but instead distinguishes correlated signals from causal signals by explicitly modeling LD structure at the likelihood level (Schaid et al., 2018). This characteristic makes it less dependent on external information and therefore generally more robust when functional annotations are sparse or of uncertain reliability. At the same time, CAVIAR allows multiple causal variants to exist within a region, giving it considerable flexibility in handling complex genetic architectures.

This robustness, however, comes at a cost. Because CAVIAR must evaluate high-dimensional causal configuration spaces, its computational complexity increases rapidly as the number of candidate SNPs and the number of potential causal variants rise. CAVIAR is therefore best understood as a likelihood-driven model that emphasizes robustness in causal inference.

(3) MsCAVIAR: enhancing identifiability through cross-study information

MsCAVIAR extends CAVIAR into cross-study or cross-population settings. By jointly modeling summary signals from multiple studies together with their respective LD structures, it further improves the identification of causal variants (Lapierre et al., 2020). Its most important advantage lies in its ability to exploit differences in LD patterns

across studies, thereby removing non-causal signals more effectively and improving the assessment of causal consistency across studies.

At the same time, MsCAVIAR places higher demands on the data. It depends on the completeness and comparability of multi-study datasets, and because it combines statistical information across several studies, its computational burden is also substantially increased. MsCAVIAR can therefore be regarded as a cross-study model that extends the scope of causal inference and improves identifiability through cross-study information.

(4) Colocalization methods: inference of cross-trait causal consistency

Colocalization methods, such as COLOC and eCAVIAR, extend the inferential target from a single phenotype to multiple phenotypes, focusing on whether different phenotypes share the same causal variant and thereby enabling probabilistic modeling of cross-trait causal consistency (Giambartolomei et al., 2013; Hormozdiari et al., 2016). Whereas fine-mapping primarily concerns locus-level causal probabilities, colocalization methods further expand the inferential target to shared causal probability, allowing GWAS signals to be connected more directly to molecular mechanisms.

The performance of these methods, however, is typically influenced by the quality of downstream molecular QTL data and the degree of matching across tissues and cell types. For this reason, colocalization methods are best understood as a framework for causal inference at the cross-trait level.

6.2 A layer-based decision strategy

In practical research, method selection is better organized according to inferential layers rather than based on empirical preference alone. Different stages of analysis address different questions, and the methods used should accordingly change with the inferential goal.

(1) Genome-wide scanning stage

At the genome-wide scanning stage, the main objective is usually to narrow the candidate space as quickly as possible in order to complete an initial screen in large-scale datasets. In this context, fastPAINTOR is often an appropriate choice. Its use presupposes access to high-quality functional annotation information. By leveraging annotation-informed prioritization, the method can balance efficiency with candidate-space reduction in high-throughput analysis settings (Kichaev et al., 2016; Zou et al., 2021).

(2) Regional fine-resolution stage

Once the analysis enters the stage of fine-scale dissection of candidate regions, the focus shifts from broad screening to the robust identification of sets of causal variants. At this stage, CAVIAR is often more advantageous. Especially when functional annotations are insufficient or local LD structure is complex, its independence from prior information allows its robustness to become more evident (Schaid et al., 2018).

(3) Cross-study integration stage

When the research goal further shifts to evaluating causal consistency across studies or ancestries, MsCAVIAR becomes more suitable. It can exploit LD differences across studies to enhance identification, thereby improving both causal consistency assessment and localization resolution in multi-cohort or multi-ancestry studies (Lapierre et al., 2020).

(4) Mechanistic interpretation stage

At the stage of mechanistic interpretation, the focus is no longer limited to locating causal variants, but rather to linking genetic variants with molecular functional processes. At this point, colocalization analysis combined with TWAS can more effectively evaluate cross-trait causal consistency and connect GWAS findings with potential molecular mechanisms (Giambartolomei et al., 2013; Okamoto et al., 2023).

6.3 A unified strategy: hierarchical inference pipeline

Integrating the above considerations, method selection in complex trait genetics can be summarized as a unified inferential pipeline:

GWAS→fine-mapping→colocalization→TWAS/functional validation

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).

From a unified perspective, these two paths correspond to improving the identifiability of causal inference either by enhancing prior information or by expanding observed data. In practice, the combination of high-quality functional annotations and multi-ancestry data integration often provides a favorable balance between efficiency and robustness.

7.3 From single-trait to multi-trait inference: extending the causal layer

Fine-mapping mainly addresses the distribution of causal probabilities under a single-trait setting, whereas colocalization analysis further extends this inferential framework into the multi-trait space by evaluating whether different traits share causal variants, thereby enabling probabilistic modeling of cross-trait causal consistency (Giambartolomei et al., 2013; Hormozdiari et al., 2016). In statistical terms, this extension corresponds to a shift from univariate posterior distributions to joint posterior distributions.

In practical workflows, this extension forms a continuous inferential chain:

GWAS→fine-mapping→colocalization→TWAS

This pipeline embeds statistical association signals into a causal pathway linking variants, genes, and phenotypes, thereby facilitating the transition from association discovery to mechanistic interpretation (Okamoto et al., 2023).

7.4 Implications for crop genetic improvement

Although probabilistic fine-mapping methods were primarily developed in human genetics, they also hold considerable promise for crop genetic improvement. Plant traits are often shaped by strong environmental dependence and complex population structure, which means that GWAS results obtained under a single environment or within a single population often have limited stability (Schaid et al., 2018).

In this context, the probabilistic framework centered on PIP and credible sets can be used to evaluate the consistency of causal signals across multiple environments and populations, thereby identifying more stable genetic factors. More specifically, MsCAVIAR can be applied to validate causal consistency across populations (Lapierre et al., 2020), colocalization analysis can help connect genetic variants with molecular mechanisms (Giambartolomei et al., 2013), and functional annotation integration can be used to prioritize candidate loci (Kichaev et al., 2016). The combination of these approaches provides a more reliable statistical basis for marker-assisted selection (MAS) and the prioritization of targets for gene editing.

7.5 Integration with downstream causal inference methods

Fine-mapping and colocalization analysis together provide an important foundation for higher-level causal inference methods. In Mendelian randomization (MR) analysis, selecting instrumental variables on the basis of credible sets can reduce false positives driven by LD and thereby improve the validity of the instruments (Broekema et al., 2020). At the same time, by combining colocalization and TWAS results, it is possible to construct multilayer causal networks linking variants, genes, pathways, and phenotypes, thus extending the analytical perspective from single-gene interpretation to systems-level biological interpretation (Okamoto et al., 2023).

Therefore, within a unified framework, the role of fine-mapping is not merely to refine association signals locally, but to transform statistical association results into candidate sets that can be used for downstream causal inference, thereby serving as a key intermediate layer in the broader causal analysis chain.

7.6 Future directions: toward an integrated causal inference framework

Future developments in statistical genetics are likely to continue advancing along the direction of a unified causal inference framework. Important directions include the deeper integration of multi-ancestry and multi-omics data (Lapierre et al., 2020), the development of multi-trait and network-level models such as Genomic SEM, and the further application of machine learning methods to high-dimensional causal inference. However, despite the expansion of methodological forms, the central challenge remains unchanged: how to maintain interpretability and reproducibility in causal inference under complex data structures.

Against this background, the probabilistic framework centered on PIP and credible sets will continue to serve as an important connecting foundation across different methods, functioning as a common language of representation and supporting the ongoing transition of statistical genetics from association analysis toward causal inference (Schaid et al., 2018).

8 Conclusion

The development of probabilistic fine-mapping marks an important shift in the research paradigm of statistical genetics. The focus of complex trait research is no longer confined to the detection of association signals, but is gradually moving toward an inferential framework centered on causal probability distributions. Within this framework, genetic variants are no longer ranked solely according to significance levels, but are instead characterized probabilistically through posterior inclusion probabilities (PIPs) and credible sets, thereby enabling a more systematic representation of uncertainty in causal inference. This shift means that research on complex traits is moving from the localization of association signals toward the analysis of causal mechanisms.

The main contribution of this study lies in its attempt to bring different fine-mapping methods into a common causal configuration framework from a unified Bayesian perspective. Although these methods differ in their implementation, they all essentially serve to constrain and characterize the same causal space. Methods that rely on functional annotations primarily improve resolution through prior information, whereas methods that rely on LD structure enhance robustness through likelihood-based modeling, and cross-study integrative methods further improve the identifiability of causal signals by expanding the data basis. Thus, the differences among these methods are better understood as arising from the way information is utilized, rather than from differences in the inferential target itself.

On this basis, the inferential framework can also be naturally extended to the multi-trait level. By modeling the probability that different phenotypes share causal variants, colocalization analysis extends causal inference from the single-trait setting to the evaluation of cross-trait causal consistency, thereby establishing a statistical connection between genetic variation and molecular mechanisms. In this way, research on complex traits gradually forms a continuous inferential path, beginning with GWAS, proceeding through fine-mapping and colocalization analysis, and ultimately moving toward functional interpretation and candidate gene identification.

From a practical perspective, method selection should be organized systematically according to inferential layers and data structure. Different methods are not simple substitutes for one another, but rather different steps designed to address different questions within the same causal inference process. Only by combining these methods appropriately in specific research contexts is it possible to achieve a balance among computational efficiency, inferential robustness, and biological interpretability, and thereby improve the overall reliability of complex trait analysis.

Author Contributions

Xuanjun Fang conducted the study, including literature review, data analysis, and drafting and revising the manuscript. The author has read and approved the final version of the manuscript.

Acknowledgments

This work was supported by the Major Program of the National Natural Science Foundation of China (Grant No. 30490254).

References

- Broekema R.V., Bakker O.B., and Jonkers I.H., 2020, A practical view of fine-mapping and gene prioritization in the post-genome-wide association era, *Open Biology*, 10(1): 190221.
<https://doi.org/10.1098/rsob.190221>
- Fang X.J., 2026a, Genome-wide relationship matrix-based heritability estimation: statistical interpretation, comparability, and practical diagnostics in the GCTA-GREML framework, *Computational Molecular Biology*, 16(3): 159-180.
- Fang X.J., 2026b, Polygenic Risk Scores (PRS/PGS) across Multi-ancestry and Cross-domain Settings: Statistical Framework, Methodological Advances, and Robustness Evaluation, *Genomics and Applied Biology*, 17(3): 138-153.
- Fang X.J., and Wu W.R., 2026, Evolution of statistical genetic paradigms: from linkage analysis and candidate gene strategies to GWAS, *Molecular Plant Breeding*, 24(9): 2817-2829.

- Gerber Z., Fisun M., Aschard H., and Djebali S., 2023, PainterPipe: A pipeline for genetic variant fine-mapping using functional annotations, *Bioinformatics Advances*, 4: vbad188.
<https://doi.org/10.1093/bioadv/vbad188>
- Giambartolomei C., Vukcevic D., Schadt E.E., Franke L., Hingorani A. D., Wallace C., and Plagnol V., 2013, Bayesian test for colocalisation between pairs of genetic association studies using summary statistics, *PLoS Genetics*, 10(5): e1004383.
<https://doi.org/10.1371/journal.pgen.1004383>
- Hormozdiari F., Kostem E., Kang E. Y., Pasaniuc B., and Eskin E., 2016, Identifying causal variants at loci with multiple signals of association, *Genetics*, 198(2): 497-508.
<https://doi.org/10.1534/genetics.114.167908>
- Hutchinson A., Watson H., and Wallace C., 2019, Correcting the coverage of credible sets in Bayesian genetic fine-mapping, *bioRxiv*, 781062.
<https://doi.org/10.32614/CRAN.package.corrcoverage>
- Hutchinson A., Watson H., and Wallace C., 2020, Improving the coverage of credible sets in Bayesian genetic fine-mapping, *PLoS Computational Biology*, 16(4): e1007829.
<https://doi.org/10.1371/journal.pcbi.1007829>
- Kichaev G., Yang W., Lindstrom S., Hormozdiari F., Eskin E., Price A. L., Kraft P., and Pasaniuc B., 2014, Integrating functional data to prioritize causal variants in statistical fine-mapping studies, *PLoS Genetics*, 10(10): e1004722.
<https://doi.org/10.1371/journal.pgen.1004722>
- Lapierre N., Taraszka K., Huang H., He R., Hormozdiari F., and Eskin E., 2020, Identifying causal variants by fine mapping across multiple studies, *PLoS Genetics*, 17(4): e1009733.
<https://doi.org/10.1371/journal.pgen.1009733>
- Okamoto K., Schmitz R. J., and Shinozaki K., 2023, Multi-omic integration for functional genomics in plant and human GWAS, *Nature Reviews Genetics*, 24(1): 30-44.
- Schaid D. J., Chen W., and Larson N. B., 2018, From genome-wide associations to candidate causal variants by statistical fine-mapping, *Nature Reviews Genetics*, 19(8): 491-504.
<https://doi.org/10.1038/s41576-018-0016-z>
- Shrestha M., Bai Z., Gholipourshahraki T., Hjelholt A., Hu S., Kjolby M., Rohde P., and Sørensen P., 2024, Evaluation of Bayesian linear regression models as a fine mapping tool, *bioRxiv*.
<https://doi.org/10.1101/2023.09.01.555889>
- Spain S. L., and Barrett J. C., 2015, Strategies for fine-mapping complex traits, *Human Molecular Genetics*, 24(R1): R111-R119.
<https://doi.org/10.1093/hmg/ddv260>
- Talukdar G., Orr H.T., and Lei Z.X., 2023, The PERK pathway: beneficial or detrimental for neurodegenerative diseases and tumor growth and cancer, *Human Molecular Genetics*, 32(16): 2545-2557.
<https://doi.org/10.1093/hmg/ddad103>
- Zou Y., Carbonetto P., Wang G., and Stephens M., 2021, Fine-mapping from summary data with the "Sum of Single Effects" model, *PLoS Genetics*, 18(7): e1010299.
<https://doi.org/10.1371/journal.pgen.1010299>



Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
