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.