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.