inducible effects (Lu et al., 2024). Importantly, results that do not support a given hypothesis should also be retained, as they contribute to refining candidate prioritization and avoiding overinterpretation.

7.4 Integration with emerging multi-omics and dynamic systems

The expansion of multi-omics data has made it possible to characterize genetic effects across multiple biological layers. Epigenomic and three-dimensional genome data provide direct evidence for regulatory mechanisms, allowing the physical relationships between variants, regulatory elements, and target genes to be examined. For example, different classes of QTL and chromatin interaction data can jointly describe the connections between enhancers and promoters (Hu et al., 2018; Bhattacharya et al., 2021).

Incorporating these data into integrative frameworks allows for more detailed interpretation of regulatory pathways. When multiple molecular layers-such as expression, splicing, and protein abundance-are considered simultaneously, multivariable approaches can help distinguish their relative contributions.

Single-cell and multimodal data further extend this framework by enabling analyses at the level of cell types and cellular states. These approaches have already revealed highly specific regulatory patterns in several systems, offering new perspectives on the mechanisms underlying complex traits.

Future directions are likely to include the incorporation of temporal and perturbation data, allowing dynamic processes to be modeled more explicitly, as well as network-based approaches that consider groups of genes or regulatory modules. Coupling statistical analyses with high-throughput experimental validation may ultimately lead to integrated systems in which data-driven inference and experimental testing inform one another, advancing the transition from statistical associations to mechanistic understanding and, eventually, to targeted intervention (Colomé-Tatché and Theis, 2018; Bhattacharya et al., 2021).

8 Conclusion

A central challenge in complex trait genetics is not the application of individual methods, but the establishment of a coherent analytical path that connects statistical associations to biological interpretation. In this context, colocalization analysis occupies a critical intermediate position, serving to evaluate the consistency of signals across data sources and to guide the transition from molecular association to downstream inference.

When GWAS and molecular QTL signals show stable correspondence within a genomic region, this provides a basis for prioritizing candidate genes and regulatory elements. However, such evidence reflects compatibility at the level of shared signal rather than direct insight into underlying mechanisms. In other words, colocalization supports entry into further analysis but does not, on its own, establish how a genetic variant influences a trait. Interpreting it as evidence of causality without additional support risks conflating shared genetic architecture with specific biological pathways.

Building on this foundation, Mendelian randomization offers a means to evaluate potential relationships in terms of direction and magnitude. By leveraging genetic variants as instruments, MR extends the analysis from signal alignment toward effect estimation. At the same time, its conclusions remain contingent on a set of assumptions, including the strength and validity of the instruments and the absence of alternative pathways. As a result, MR findings should be interpreted alongside diagnostic measures and complementary methods, and discrepancies should prompt re-examination of earlier analytical steps rather than isolated interpretation.

Taken together, the framework outlined here is best understood not as a fixed pipeline, but as a process of progressive refinement. Starting from GWAS signals, fine-mapping reduces the candidate variant space; molecular QTL data and colocalization analyses help identify signals that are consistent across datasets; expression-based models further narrow the focus to gene-level candidates; and, where appropriate, MR is used to evaluate potential relationships. The outputs of this process extend beyond lists of candidate genes, incorporating levels of supporting evidence and consistency that can guide experimental prioritization.