(annotation-informed Bayesian methods versus LD-aware baseline approaches). Third, model selection should be driven by external validation, comparing relative R^2/AUC, calibration slope, and decision-curve net benefit in held-out or independent datasets, with the final model applied only after being frozen (Kachuri et al., 2024).

At the data and modeling levels, future improvements in PRS performance require both expanded data coverage and advances in statistical methodology. On the one hand, efforts should focus on increasing the representation of multi-ancestry GWAS and LD reference panels, improving the capture of low-frequency and structural variants, and implementing standardized quality control across centers and platforms in human studies, as well as covering the target population of environments (TPE) through multi-environment trials (METs) in crop systems (Wang et al., 2018). On the other hand, modeling strategies should incorporate ancestry-aware LD structures, hierarchical random effects, and transfer learning approaches to decompose shared and population-specific effects. In target populations, recalibration (e.g., slope and intercept adjustment) and reweighting (stacking) can reduce predictive bias, while the integration of local ancestry and functional annotations can further improve cross-population robustness (Cai et al., 2021; Zhang et al., 2023).

At the evaluation level, a shift from single performance metrics to a systematic assessment framework is needed. In addition to reporting R^2 and AUC, studies should report the transferability ratio ($T = R^2_{target}/R^2_{source}$), population-stratified calibration metrics (slope and intercept), and decision-curve net benefit. Sensitivity analyses with respect to LD reference panels, functional annotation strength, and model parameters should also be conducted, forming a closed-loop evaluation framework of "training-extrapolation-recalibration-re-evaluation."

At the application level, PRS/PGS are evolving from predictive tools toward decision-support systems. In medicine, combining PGS with age, family history, biomarkers, and lifestyle factors enables stratified screening and personalized interventions for high-burden diseases such as cardiovascular, metabolic, and certain cancers. In crop breeding, PGS is methodologically aligned with genomic selection (GS), and can be used for early-stage preselection and multi-environment reaction norm modeling to achieve coordinated optimization of "general adaptability and environment-specific selection," thereby substantially improving genetic gain per unit time in complex stress-related traits (Wang et al., 2018; Alemu et al., 2024).

Despite these advances, several key bottlenecks remain. First, imbalances in multi-ancestry data and functional annotation resources lead to training bias and uncertainty in evaluation. Second, low-frequency and rare variants, as well as structural variants (e.g., SVs and CNVs), are not fully captured under current "tag SNP" frameworks, requiring advances in pangenome references, long-read sequencing, and multi-omics data integration for improved causal inference. Third, cross-platform batch effects and residual population structure may further amplify extrapolation errors (Du et al., 2025). From an ethical and governance perspective, medical applications must address risks of genetic determinism and discrimination, and establish frameworks for dynamic consent and data sovereignty; in breeding, considerations of biodiversity conservation, equitable benefit sharing, and support for smallholder systems are essential to avoid structural bias driven by performance optimization alone (Gorjanc et al., 2017; Broesch et al., 2020).

Looking forward, the development of PRS/PGS can be summarized within an integrated paradigm: multi-ancestry data expansion + causal and functional annotation integration + ancestry-aware modeling + environment coupling + recalibration and fairness evaluation

This paradigm will facilitate the transition of PRS/PGS from standalone predictive tools to comprehensive platforms that are transferable, interpretable, and governable.

6 Conclusion

Polygenic risk scores (PRS/PGS) systematically integrate GWAS-derived effect estimates to transform locus-trait associations into actionable individual-level predictive measures. In both human medicine and crop breeding, they support risk stratification, early screening, and selection decisions, serving as a key bridge between fundamental genetics and translational applications.