CABBI
CENTER FOR ADVANCED BIOENERGY AND BIOPRODUCTS INNOVATION

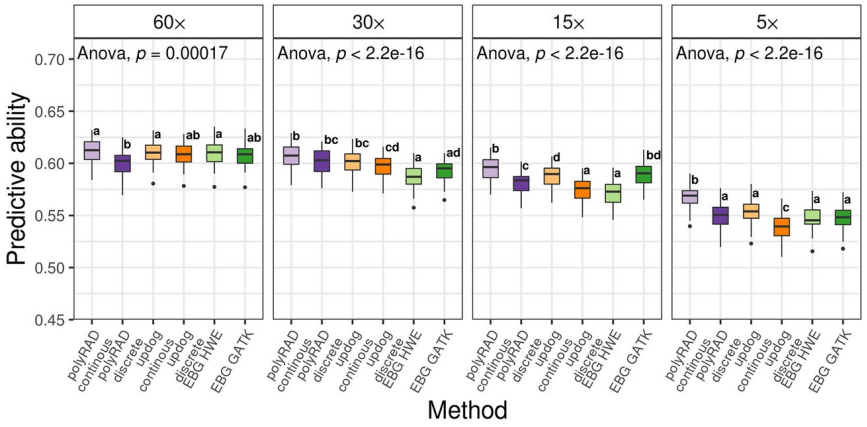
Impact of Genotype-Calling Methodologies on Genome-Wide Association and Genomic Prediction in Polyploids

Background/Objective

Discovery and analysis of genetic variants underlying agriculturally important traits are key to molecular breeding of crops. However, accurate genotype calling from cost-efficient, next-generation sequencing data is challenging, particularly in polyploid species due to their genome complexity. This study used empirical and simulated data to evaluate three Bayesian algorithms and demonstrate their impact on the power of genome-wide association study (GWAS) analysis and the accuracy of genomic prediction.

Approach

Recently developed Bayesian statistical methods, polyRAD, EBG, and updog, were used. We further incorporated uncertainty in allelic dosage estimation by testing continuous genotype calls and comparing their performance to discrete genotypes in GWAS and genomic prediction. We tested the genotype-calling methods using data from two autotetraploid species, *Miscanthus sacchariflorus* and *Vaccinium corymbosum*, and performed GWAS and genomic prediction.



Boxplots of predictive ability (*y*-axis) in the simulated study at sequencing depths 60×, 30×, 15×, and 5×, comparing six genotype-calling methods (*x*-axis).

Results

In the empirical study, the tested Bayesian genotype-calling algorithms differed in their downstream effects on GWAS and genomic prediction, with some showing advantages over others. Through subsequent simulation studies, we observed that at low read depth, polyRAD was advantageous in its effect on GWAS power and limit of false positives. Additionally, we found that continuous genotypes increased the accuracy of genomic prediction, by reducing genotyping error, particularly at low sequencing depth.

Significance/Impacts

Our results indicate that by using the Bayesian algorithm implemented in polyRAD and continuous genotypes, we can accurately and cost-efficiently implement GWAS and genomic prediction in polyploid crops. Future studies that integrate uncertainty in dosage estimation may be able to improve the accuracy of genomic prediction, especially when low read depths are used to keep costs low and efficiency high.

Njuguna et al. 2023. "Impact of Genotype-Calling Methodologies on Genome-Wide Association and Genomic Prediction in Polyploids." *The Plant Genome*. DOI: 10.1002/tpg2.20401.