

# Improving Precision and Accuracy of Genetic Mapping with Genotyping-by-Sequencing Data in Outcrossing Species

## Background/Objective

- Genotyping-by-sequencing (GBS) is a widely used strategy for obtaining large numbers of genetic markers in model and non-model organisms. In crop plants, GBS-derived marker datasets are frequently used to perform quantitative trait locus (QTL) mapping.
- To identify strategies for increasing precision and accuracy of QTL mapping using GBS data in outbred crops, we conducted two complementary studies: (1) a comparison of SNP-calling and genetic map-building pipeline parameters in an empirical dataset of 18 biomass yield traits from a biparental *Miscanthus sinensis* F1 population using RAD-Seq SNPs, and (2) a simulation study to estimate the relationship between genome-wide error rate, read depth, and marker number.

## Approach

For the study of empirical *M. sinensis* data, we compared two SNP-calling methods (one non-reference-based and one reference-based), a series of depth filters (12×, 20×, 30×, and 40×) and two map-construction methods (linkage-only and order-corrected based on a reference genome).

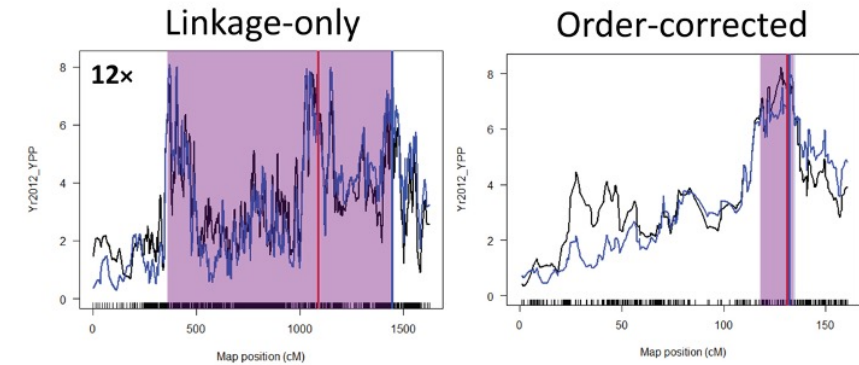
## Results

We found that correcting the order of markers on a linkage map by using a high-quality reference genome improved QTL precision (shorter confidence intervals). For typical GBS datasets of between 1,000 and 5,000 markers to build a genetic map for biparental populations, a depth filter set at 30× to 40× applied to outbred populations provided a genome-wide genotype-calling error rate of less than 1%, improved accuracy of QTL point estimates, and minimized type I errors for identifying QTL.

## Significance/Impacts

The use of a high-quality reference genome to correct the order of GBS markers on a genetic map is highly advantageous for improving the precision of QTL mapping in outbred populations and should be standard practice when possible. We recommend using a reference genome to correct the marker order of genetic maps and a robust genotype depth filter to improve QTL mapping for outbred crops. These improved methods will increase the efficiency and speed of breeding outbred bioenergy crops.

LaBonte, et al. 2024. “Improving Precision and Accuracy of Genetic Mapping with Genotyping-by-Sequencing Data in Outcrossing Species.” *GCB Bioenergy*. DOI: 10.1111/gcbb.13167.



**LOD curves for biomass yield (black = 2012, blue = 2013) on LG7 in *Miscanthus sinensis* population ‘Strictus’ × ‘Kaskade’, with linkage-based maps on the left and equivalent maps with order correction from the reference genome on the right (markers called with TASSEL-GBS).**