Oat Newsletter oatnews.org volume 56, no.7, 2019

Theoretical unification and practical integration of conventional methods and genomic selection in plant breeding

Yan, W., Tinker, N.A., Bekele, W.A., Mitchell Fetch, J., and Frégeau-Reid, J. (2019) Crop Breed. Genet. Genom. **1**:e190003

DOI: https://doi.org/10.20900/cbgg20190003

Contact:
Weikai Yan, Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, 960 Carling Ave., Ottawa, ON K1A 0C6, Canada.

Email: weikai.yan@canada.ca

Summary:

Following successful application in dairy cow breeding, genomic selection (GS) has become a hot topic among plant geneticists and breeders. GS and conventional breeding methods have the same goal of identifying the best genotypes for a given crop and region, and should follow the same principles, particularly in dealing with genotype-by-environment interactions (GE).

Dealing with GE includes dividing a target region into meaningful mega-environments (MEs) based on repeatable GE patterns, and then selecting for each ME. This requires that different GS models be developed for different MEs. To account for the unrepeatable GE in the ME and to estimate the genetic main effect (G_P), selection for each ME requires testing at multiple locations for multiple years. For a ME with large GE, multiple and diverse cultivars should be selected and recommended. The number of locations, years, and replications within trials required to achieve a certain level of heritability (i.e., selection reliability, or prediction accuracy) for a trait (e.g., yield) can be estimated from existing multi-location, multiyear variety trial data.

Instead of direct selection for G_P, conventional selection has to resort to using indirect selection, negative selection (culling), and a lengthy breeding cycle, with low selection reliability. GS offers the possibility for direct selection and positive selection for G_P, thereby overcoming random GE and shortening the breeding cycle under conditions where high selection reliability can be achieved. To this end, GS models must be able to predict G_P for a breeding population, and its prediction accuracy should be measured by $r(G_M,G_P)$, with G_M being GS-predicted breeding values.

This definition of prediction accuracy is the essential connection between GS and phenotype-based selection. It implies that GS models must be developed and evaluated using phenotypic data from multiple locations and multiple years representing the target ME. Using a single training population for model development and a different population for model evaluation will allow estimation of $r(G_M,G_P)$, but it can be costly and may have limited relevance to a breeding program.

A pragmatic GS framework was proposed in this paper, with the following steps:

1) for model development, use data from yearly preliminary yield trials, in which a large number of new breeding lines are tested at several locations,
2) develop a GS model for each training population-by-trial (location-year) combination,
3) use all available GS models to make predictions, leading to a genotype-by-model two-way table of predictions for the current-year's breeding population,
4) to make selection decisions and to evaluate and select models as a breeder would do for their test environments, analyze this two-way table the way a breeder would do with a genotype-by-environment two-way table of a trait, and
5) to measure the prediction accuracy of GS and to identify a new cultivar, use the "rate of success", defined as the inverse of the number of genotypes that have to go through yield trials, instead of $r(G_M,G_P)$.

The prediction accuracy or rate of success determines the best pathway to integrate GS into a practical breeding program. Preliminary results from applying GS in the Ottawa oat breeding program showed promise; the GS prediction accuracy is expected to improve with time as more GS models are developed and as the breeding populations and the target ME are better represented in the GS models.