Genomic Selection of Purebred Animals for Crossbred Performance under Dominance

J. Zeng[1], A. Toosi[1], R. L. Fernando[1], J.C.M. Dekkers[1], and D.J. Garrick[1]
[1]Department of Animal Science, Iowa State University, Ames, Iowa, 50011, US.
jdekkers@iastate.edu

Crossbreeding is extensively used to take advantage of heterosis and breed complementarity. In livestock, purebreds (PB) that are crossed to produce the commercial crossbreds (CC) are typically raised under superior management conditions. Thus, due to genetic and environmental differences, performance of PB can be a poor predictor of CC performance. Genomic selection (GS) with training on CC performance may overcome this problem, but in most GS studies additive gene action has been assumed yet dominance is the likely basis for heterosis.

In this study, crossbreeding was simulated under a dominance model, and performance of GS under a BayesCπ framework with training on CC data using an additive model that assumed a common substitution effect (α_c) of SNP alleles across breeds was compared to training using a dominance model that included dominance (d) in addition to additive effects (a) of SNP without accounting for breed origin. The comparison was based on response to selection in crossbreds and predictive accuracies of PB breeding values for CC improvement in a variety of training populations. The simulated genome consisted of 1,000 SNP and 100 QTL on a single chromosome of 100 cM. Breeds were separated 50 generations prior to training, where broad and narrow sense heritability was 1/2 and 1/6. Mean crossbred response to selection over 20 generations was calculated for each model over 100 replications. The true breeding value of a PB individual for CC performance is a linear function of breed-specific allele substitution effects defined as $\alpha_s = a + (1-2p)d$, where p is the allele frequency in the breed of mates. Estimates of α_s in the dominance analysis and estimates of α_c in the additive analysis were used for genomic prediction of breeding values. Mean accuracy of genomic prediction summarized over 24 replications was the correlation between predicted and true breeding values for CC performance.

Results showed that without retraining, the dominance model steadily gave greater crossbred response during selection, accumulating to an advantage of over 35% by generation 20 compared to the additive model. Extra response was the result of an increase in heterosis but at some cost to improvement of PB performance. The rate of decline in accuracy during selection was less with the dominance versus additive model because SNP allele substitution effects were recomputed with updated allele frequencies every generation under the dominance model. For various training populations, the dominance model resulted in higher accuracies than the additive model (by 2 to 13%), particularly when the genomic background of the training population had less overlap with the target crossbreds. When training was on an admixed population, including breed composition in either model did not improve accuracy of GS.

In conclusion, when dominance is present, GS with a dominance model enables more effective PB selection for CC performance than an additive model and allows lower frequency of retraining, which is appealing for traits that are hard or expensive to measure. As the training population becomes less related to the target crossbreds, PB breeding values

inaccurately predicted target CC performance, which is the situation that favors the dominance model for training.