ACCURACIES OF GENOMIC PREDICTIONS IN US BEEF CATTLE

M. Saatchi[1] and D. J. Garrick[1,2]

[1]Department of Animal Science, Iowa State University, Ames, IA
[2]Institute of Veterinary, Animal and Biomedical Sciences, Massey University,
Palmerston North, New Zealand

SUMMARY

Molecular breeding values (MBV) derived from genomic information holds promise to increase the accuracy of genetic evaluation in young animals. The objective of this study was to derive and evaluate the accuracies of MBV for economically relevant traits for routine implementation in several beef cattle breeds. Within-breed application of genomic predictions improved the accuracies of genetic evaluations compared to traditional methods. Accuracies of MBV ranged from 0.37 to 0.68 in American Hereford, 0.37 to 0.85 in American Red Angus, and from 0.19 to 0.73 in American Simmental using within-breed genomic predictions. Within-breed genomic predictions had less utility when applied to other breeds. Within-breed genomic predictions were less accurate for animals that were less closely related to the training population, such as those bred in other countries. The accuracies of MBV improved slightly for some breeds when predictions were derived using multi-breed reference populations obtained by simply pooling the genotypic and phenotypic data from different breeds. Genomic information has now been implemented into routine genetic evaluation for breeders of American Angus, Hereford, Limousin, Red Angus and Simmental beef cattle and will soon be extended to other US breeds.

INRODUCTION

Now it is possible to genotype beef cattle for more than 50,000 single nucleotide polymorphisms (SNP) at relatively low cost. The resulting SNP genotypes can be used to produce molecular breeding values (MBV) for selection candidates that do not necessarily have phenotypes (Meuwissen *et al*. 2001). Selection of young animals using MBV could reduce generation intervals and increase genetic progress (Schaeffer 2006). The accuracies of resultant MBV are key to successful application of this new technology as genetic gain is directly proportional to the accuracy achieved. The objective of this study was to compare accuracies of genomic predictions using within- or multi-breed reference populations for American Hereford, Red Angus and Simmental beef cattle.

MATERIALS AND METHODS

Genotype and phenotype data. A total of 9,931 animals (3,550 Hereford, 3,178 Simmental, 1,766 Black Angus, 1,274 Red Angus, 124 Gelbvieh, 37 Brangus and 2 Charolais) were genotyped, mainly at GeneSeek (Lincoln, NE). Most animals were genotyped with the BovineSNP50 BeadChip (Illumina, San Diego, CA) but some animals (less than 5%) were genotyped with the BovineHD BeadChip (Illumina, San Diego, CA). For those animals, genotypes for markers present on the BovineSNP50 BeadChip were extracted. Deregressed estimated breeding values (DEBV) free of parent average effects, derived following Garrick *et al*. (2009), were used as response variables in weighted analyses to estimate SNP effects.

In total, 12 traits were analyzed (some traits were not recorded in all breeds, Table 1). The number of genotyped animals with DEBV varied among traits (2395 for scrotal circumference to 9443 for birth weight) because some animals had no individual or offspring information contributing to their expected progeny difference (EPD) and therefore had no information in their DEBV.

Statistical model. Habier *et al.* (2011) showed that the method "BayesC" (Kizilkaya *et al.* 2010) is less sensitive to prior assumptions than the method "BayesB" (Meuwissen *et al.* 2001). So, method BayesC was used to estimate marker effects for Red Angus and Simmental animals. However, the method BayesB was used for Hereford animals (all traits except fat thickness and marbling, which used BayesC) as the higher accuracies reported by Saatchi *et al.* (2013) for a subset of the data using BayesB method. Both methods assume that some known fraction of markers (π) have zero effect. For each trait, the following model was fit to the DEBV data for training: $$y_i = \mu + \sum_{j=1}^{k} z_{ij} u_j + e_i \,,$$
where y_i is the DEBV for animal *i*, μ is the population mean, *k* is the number of marker loci in the panel, z_{ij} is allelic state (i.e., number of B alleles from the Illumina A/B calling system) at marker *j* in individual *i*, u_j is the allele substitution effect for marker *j*, with $u_j \sim N(0, \sigma_u^2)$ (with probability *1 - π*) or $u_j = 0$ (with probability π) as described by Habier *et al.* (2011) for BayesB and BayesC methods, and e_i is a residual with heterogeneous variance, depending on the reliability of the DEBV information for animal *i* (Garrick *et al.* 2009). Parameter π was assumed to be 0.95 for all analyses. Markov chain Monte Carlo (MCMC) methods with 41,000 iterations were used to provide posterior mean estimates of marker effects and variances, after discarding the first 1,000 samples for burn-in.

The MBV for individual *i* within a validation set was derived as the sum over all *k* markers of posterior means of predicted SNP effects, as estimated in the training set, multiplied by the number of copies of the B allele. All analyses were performed using GenSel software (Fernando and Garrick 2010).

Within breed genomic predictions. The accuracies of MBV were evaluated by pooling estimates from either 5-fold or 6-fold cross-validation strategies for 2,980 American Hereford, 1,274 American Red Angus or 2,703 American Simmental animals. The K-means clustering method (Saatchi *et al.* 2011) was used for partitioning animals with the aim of increasing within-group and decreasing between-group relationships. Within-breed training analyses were performed by excluding one group when estimating marker effects, which were then used to predict MBV of individuals from the omitted group (validation set). Bivariate animal models were used for each trait to estimate the genetic correlation between DEBV and MBV as a reflection of the accuracy of genomic prediction (Saatchi *et al.* 2012).

Accuracies of genomic predictions were also evaluated for Red Angus animals using a multi-breed reference population (consisted of 3,178 Simmental, 1,766 Black Angus, 124 Gelbvieh, 37 Brangus, 31 Hereford and 2 Charolais plus 1,274 Red Angus animals). In multi-breed cross-validation, the same four Red Angus groups for each of the five training runs were used, except that animals from all the other breeds were always included in the training analyses. In multi-breed analyses only the accuracies of Red Angus predictions were of interest.

Across countries and across breed genomic predictions. The accuracies of genomic predictions were evaluated for 100 Argentine, 75 Canadian and 395 Uruguayan Hereford, 3,178 American Simmental and 1,274 American Red Angus animals using marker estimates from training on American Hereford animals. Simple correlations between MBV and DEBV were used as estimates of the accuracies of MBV in non American Hereford animals because pedigree information was not available to estimate genetic correlations between DEBV and MBV.

RESULTS

Accuracies of MBV ranged from 0.37 to 0.68 (average 0.53) in American Hereford, 0.37 to 0.85 (average 0.64) in American Red Angus, and from 0.19 to 0.73 (average 0.50) in American Simmental using within-breed genomic predictions (Table 1). Genomic predictions were more accurate in Red Angus using multi-breed rather than the single-breed reference population for all

traits except calving ease and weaning weight maternal (ranged from 0.32 to 0.90 with the average of 0.69, Table 1).

Table 1. Accuracies[1] of genomic predictions (±SE) in American Hereford, Simmental and Red Angus beef cattle using single- and multi-breed reference populations.

Trait	Single-breed			Multi-breed
	Hereford	Simmental	Red Angus	Red Angus
Birth weight	0.68±0.03	0.67±0.03	0.66±0.04	0.75±0.04
Calving ease direct	0.68±0.04	0.46±0.02	0.59±0.03	0.60±0.04
Calving ease maternal	0.51±0.04	0.31±0.02	0.37±0.03	0.32±0.04
Carcass weight	-	0.61±0.04	0.62±0.04	0.75±0.04
Fat thickness	0.48±0.04	0.19±0.02	0.85±0.16	0.90±0.15
Marbling	0.43±0.04	0.60±0.04	0.77±0.10	0.85±0.09
Rib eye muscle area	0.49±0.03	0.55±0.05	0.71±0.07	0.75±0.06
Scrotal circumference	0.43±0.04	-	-	-
Weaning weight direct	0.52±0.03	0.56±0.04	0.55±0.04	0.67±0.04
Weaning weight maternal	0.37±0.03	0.32±0.03	0.54±0.04	0.51±0.04
Yield grade	-	0.73±0.09	0.81±0.18	0.83±0.12
Yearling weight	0.60±0.03	0.45±0.02	0.57±0.04	0.69±0.03

[1]Measured as the genetic correlation between tderegressed estimated breeding values and molecular breeding values estimated from a bivariate animal model.

Table 2. Accuracies[1] of genomic predictions in non American Hereford, American Simmental and American Red Angus beef cattle using marker estimates from training on American Hereford.

Trait	Argentine Hereford	Canadian Hereford	Uruguayan Hereford	American Simmental	American Red Angus
Birth weight	0.15	0.48	0.24	0.29	0.28
Calving ease direct	-	0.58	0.28	0.30	0.27
Calving ease maternal	-	0.46	0.19	0.17	-0.11[2]
Fat thickness	-0.19[2]	0.30	0.12	0.09	0.15
Marbling	0.35	0.58	0.23	0.16	0.05
Rib eye muscle area	0.17	0.43	0.20	0.12	0.15
Scrotal circumference	-0.16	0.26	0.17	-	-
Weaning weight direct	0.10	0.31	0.28	0.22	0.21
Weaning weight maternal	0.00	0.24	0.20	0.05	0.07
Yearling weight	0.04	0.14	0.33	0.23	0.24

[1]Simple correlations between deregressed estimated breeding values (DEBV) and molecular breeding values (MBV).
[2]By definition, the accuracy cannot be a negative value. However, the negative value obtained here as the simple correlation between DEBV and MBV used as a measure of accuracy.

Within-breed genomic predictions were less accurate for animals that were less closely related to the training population, such as those bred in other countries (Table 2). Genomic predictions for Argentine, Canadian and Uruguayan Hereford were less accurate than those obtained for American Hereford animals (simple correlations between DEBV and MBV have not shown for American Hereford). Among non American Hereford animals, genomic predictions were most

accurate for Canadian Hereford (Table 2). Across-breed genomic predictions were less accurate than those obtained from within breed genomic predictions (Table 2). Among all traits, across-breed genomic predictions were higher for birth, weaning and yearling weights; and calving ease direct, than for other traits (Table 2).

DISCUSSION

This study applied genomic prediction to American and non American Hereford, American Red Angus and American Simmental beef cattle using single-breed or multi-breed reference populations. The accuracies of genomic predictions were more accurate than those reported by Saatchi *et al*. (2013) for American Hereford animals (the average of accuracies increased from 0.37 to 0.53). This is due to the larger training populations (2,980 vs. 1,081 animals) used in this study. Goddard and Hayes (2009) showed that the size of training population is a crucial factor influencing the accuracies of genomic predictions. Among non American Hereford animals, genomic predictions were more accurate for Canadian Hereford, which reflects the higher degree of genetic relationship between American and Canadian Hereford population in comparison with Argentine and Uruguayan Hereford populations (Saatchi *et al*. 2013).

Genomic predictions were more accurate using multi-breed than single-breed reference populations for Red Angus animals for all traits except calving ease and weaning weight maternal (the average of accuracies increased from 0.64 to 0.69). This may reflect the fact that some registered Simmental animals have a heterogeneous genetic background being admixed with other beef cattle breeds including Red Angus, as American Simmental Associations allow registration of crossbred animals with other beef cattle breeds. This demonstrates the benefit of using multi-breed reference population for American Red Angus beef cattle.

Across breed genomic predictions had less utility when applied to other breeds for most traits. This may reflect differences in linkage phase between markers and quantitative trait loci (QTL) or differences in causative mutations at the same QTL for these traits across different breeds. However, across-breed genomic predictions had some utility for birth, weaning and yearling weights; and calving ease direct traits due to the segregation of common QTL with large effects among these breeds (unpublished data).

Genomic information has now been implemented into routine genetic evaluation for breeders of American Angus, Hereford, Limousin, Red Angus and Simmental beef cattle and will soon be extended to other US breeds.

REFERENCES

Fernando R.L. and Garrick D.J. (2010) http://taurus.ansci.iastate.edu/ Accessed Sep. 1, 2010.
Garrick D.J., Taylor J.F. and Fernando R.L. (2009) *Genet. Sel. Evol.* **41**: 55.
Goddard M.E. and Hayes B.J. (2009) *Nat. Rev. Genet.* **10**: 381.
Habier D., Fernando R.L., Kizilkaya K. and Garrick D.J. (2011) *BMC Bioinformatics* **12**: 186.
Kizilkaya K., Fernando R.L. and Garrick D.J. (2010) *J. Anim. Sci.* **88**: 544.
Meuwissen T.H., Hayes B.J. and Goddard M.E. (2001) *Genetics* **157**:1819.
Saatchi M., Ward J. and Garrick D.J. (2013) *J. Anim. Sci.* **91**: 1538.
Saatchi M., McClure M.C., McKay S.D., Rolf M.M., Kim J., Decker J.E., Taxis T.M., Chapple R.H., Ramey H.R., Northcutt S.L., Bauck S., Boodward B., Dekkers J.C.M., Fernando R.L., Schnabel R.D., Garrick D.J. and Taylor J.F. (2011) *Genet. Sel. Evol.* **43**: 40.
Saatchi M., Schnabel R.D., Rolf M.M., Taylor J.F. and Garrick D.J. (2012) *Genet. Sel. Evol.* **44**: 38.
Schaeffer L.R. (2006) *J. Anim. Breed. Genet.* **123**: 218.