Short communication

Cross-validation of genetic and genomic predictions of temperament in Nellore–Angus crossbreds

Lauren L. Hulsman Hanna [a,*], Dorian J. Garrick [b], Clare A. Gill [a], Andy D. Herring [a], James O. Sanders [a], David G. Riley [a]

[a] Department of Animal Science, Texas A&M University, College Station, TX 77843, United States
[b] Department of Animal Science, Iowa State University, Ames, IA 50011, United States

ABSTRACT

The objective of this study was to cross-validate predictions of genomic merit for overall temperament at weaning (TEMP) by alternative partitioning of training and validation sets in a Nellore–Angus crossbred population. Nellore–Angus F_2 calves (embryo transfer and natural service reciprocal crosses), paternal half-siblings to the embryo transfer calves, and F_3 calves had records for TEMP (subjectively scored on a 1 to 9 scale). Calves with DNA available were genotyped on the Illumina Bovine SNP50 version 1 and 2 assays, and, after quality control filtering, there were 34,913 SNP markers available for use. Calves with TEMP records and genotypes ($n=769$; mean $\pm$ std of all TEMP records is 3.975 ± 2.062) were used in this study. BayesB procedures with $\tilde{\pi}=0.997$ were used to obtain direct genomic values (DGV) with alternative partitioning of data into training and validation sets utilizing the family structure of this population. Training was conducted (scenarios 1–4) using the progeny of all but one of the four sire or grandsires, (5) using only F_2 progeny, (6) by random assignment, and (7) using only embryo transfer F_2 and natural service half-sibling cattle. In random assignment, the number of animals included in the validation population was the average of scenarios 1–5. The DGV generated in these training and validation sets were compared to traditional, pedigree-based breeding values from an animal model with equivalent training and validation sets. The training model included fixed effects of sex, birth year-season combination, and temperament scoring pen nested in birth year-season combination. Standardized accuracies were higher for DGV using BayesB procedures (0.226 on average) compared to EBV using a traditional pedigree-based animal model (0.122 average). Random allocation of individuals into training and validation groups resulted in the highest accuracies for DGV (0.503) and EBV (0.354) of the validation groups. Overall, accuracies of animals in validation groups were low (i.e., less than 0.35) using genomic predictions, which could be due to inadequate sample size, insufficient marker density, or limited relationship between individuals in the training and validation populations. The genomic method generated unique DGV among full- and half-siblings within families whereas pedigree-based prediction results in identical EBV among full-siblings. This is of particular interest for the cattle industry as genomic methods can provide more plausible estimates of genetic merit for crossbred or purebred cattle without records, even if the initial accuracy is low.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Direct genomic values (DGV), for the purpose of this study, are predictions of genetic merit using only genomic information. Accuracy of such predictions (i.e., how well the estimates predict the true breeding value) is reliant on the training population used, where studies have shown that the size and relationship of the training population to the validation population are key factors (e.g., Habier et al., 2010; Luan et al., 2009; Saatchi et al., 2010). This is especially relevant when the training and validation populations consist of more than one breed or include crossbred or composite animals. As relatedness between animals in training and validation decreases, so does prediction accuracy (e.g., de Roos et al., 2009; Ibáñez-Escriche et al., 2009) often due to varying degrees of linkage disequilibrium (LD) and patterns of LD among cattle breeds (Bovine HapMap Consortium, 2009; Gautier et al., 2007).

Generating DGV on animals in the training population should result in fairly accurate predictions (≥ 0.70; e.g., Hulsman Hanna et al., 2014b). Cross-validation techniques can be used to

determine how well the prediction method will work for younger animals in the population and, have been adopted to assess accuracy of genomic predictions (e.g., Boddhireddy et al., 2014; Erbe et al., 2010; Luan et al., 2009; Saatchi et al., 2010, 2011, 2012, 2013). Methods of partitioning populations have included random assignment (e.g., Saatchi et al., 2011, 2012, 2013), using year cohorts (Luan et al., 2009), generations (Saatchi et al., 2010), and K-means clustering (Boddhireddy et al., 2014; Saatchi et al., 2011, 2012, 2013).

Temperament, defined as the reaction of an animal to human handling (Burrow and Dillon, 1997; Fordyce et al., 1982), is a complex trait and has received little attention in genetic evaluation procedures to date. The primary focus has been the use of docility score (Beef Improvement Federation, 2010) due to its adoption by a handful of breed associations that have incorporated it into genetic evaluations. New methodology in genetic evaluations are focusing on incorporating genomic data, however, a limited number of studies have investigated whether previous findings concerning the relationship of training and validation populations remain the same with a measurement of temperament (e.g., Boddhireddy et al., 2014), although it is expected it will. Phenotype can directly impact genetic evaluation results due to its ability to adequately capture genetic variation of interest, therefore, evaluation of phenotype scoring, such as temperament scoring systems, is relevant. This evaluation of phenotype is also important to understand the potential impact on genetic predictions due to family structure, population size, and trait architecture.

A Nellore–Angus crossbred population at the Texas A&M AgriLife McGregor Research Center provides an opportunity to conduct a unique cross-validation approach in testing a prediction method when using a temperament scoring system. The objective of this study was to compare alternative partitioning based on family structure in this Nellore–Angus crossbred population and to evaluate the impact of those data partitions on predictions of genetic and genomic merit for overall temperament at weaning.

2. Materials and methods

2.1. Cattle population and traits

The cattle, traits, and genotypes used in this study were previously described by Hulsman Hanna et al. (2014a). In short, cattle were Nellore–Angus F_2 calves (embryo transfer (ET NA) and reciprocal natural service crosses), natural service paternal half-siblings to ET NA F_2 animals, and F_3 calves from ET NA F_2 cattle. These cattle were characterized using 3 distinct cycles: Cycle 1 (ET NA F_2 and natural service half-sibling calves), Cycle 2 (reciprocal natural service F_2 calves unrelated to ET NA F_2), and Cycle 3 (natural service F_3 calves from ET NA F_2). All procedures involving animals were approved by the Texas A&M Institutional Animal Care and Use Committee. Animals used in this study were genotyped on the Bovine SNP50 version 1 and 2 assays (Illumina Inc., San Diego, CA), which, after quality edits, provided 34,913 SNP (excluding the Y chromosome) for prediction analyses. Quality edits included checking call rates, minor allele frequency, and Hardy–Weinberg equilibrium of genotype frequencies. Overall temperament at weaning (TEMP) was utilized for this study, which was a subjective measurement of temperament on a 1 to 9 scale, where 1 is docile and 9 is wild or unruly ($n=772$). Calves were evaluated for TEMP by 4 evaluators, where the collection process has been previously described in Hulsman Hanna et al. (2014a). The average score across all evaluators was used for analyses in this study.

2.2. Statistical analysis

Utilizing the same set of animals as this study, Hulsman Hanna et al. (2014a) used a final statistical model for TEMP that included fixed effects of sex ($n=2$), family nested within sire ($n=31$), birth year-season combinations ($n=10$), and evaluation pen nested within birth year-season combinations ($n=42$). The population cycle ($n=3$), which is based on breeding type, was evaluated to account for crossbreeding effects that may appear due to family structure. It was considered separate from the effect of family nested within sire, but was not found to be a significant effect ($P=0.802$). Further investigation of the final model effects when using them for genomic predictions indicated that inclusion of fixed effects accounting for family structure (e.g., family nested within sire) detrimentally impacted the Bayesian estimation procedures (Hulsman Hanna et al., 2014b). Therefore, the final model for TEMP included all previously listed effects except family nested within sire for Bayesian procedures and the pedigree-based animal model. Only those animals with phenotypes, genotypes and records associated with the fixed effects were used for further analyses ($n=769$). This group of animals consisted of 441 ET NA F_2, 259 natural service paternal half-siblings to ET NA F_2, 20 reciprocal natural service F_2 calves, and 50 natural service F_3 calves from ET NA F_2.

2.3. Bayesian methods employed

Hulsman Hanna et al. (2014b) showed that the use of BayesB (Meuwissen et al., 2001) and BayesC (Habier et al., 2011; Kizilkaya et al., 2010) procedures did not differ significantly in their accuracy or bias, however BayesB did have numerically lower bias in the estimates. As there is more scientific literature utilizing BayesB procedures, this study will utilize those procedures with $\bar{\pi}=0.997$ and seed values as described by Hulsman Hanna et al. (2014a) to investigate alternative partitioning of training and validation data. This choice of $\bar{\pi}$ would result in approximately 105 markers being fit in a given model per iteration, which was determined to be appropriate for TEMP in Hulsman Hanna et al. (2014a). Analyses were performed using GenSel software (Fernando and Garrick, 2009), which also generate DGV for animals included in the training phase. To predict breeding values in GenSel for validation animals, the "Predict" procedures (Fernando and Garrick, 2009) were run. In both cases, DGV were calculated for each individual by summing across loci the posterior mean of each random marker effect times the centered number of copies (i.e., -10, 0, 10) of the minor allele. In addition, DGV generated using BayesB methods were compared to traditional EBV generated using ASReml software (Gilmour et al., 2009). ASReml incorporates the same fixed effects as used in genomic analyses, random animal effects with variance-covariance determined from pedigree information, and random residuals. After reformatting, the same phenotype files used for training procedures in GenSel were used to generate the EBV in ASReml. This would allow animals in the validation population to have pedigree-based EBV predicted without their own phenotype records as a control for direct comparison to genomic methods.

Cross-validation techniques incorporated family structure to determine the efficacy of predicting breeding values in the validation population and resulted in partitioning animals into 7 sets of training and validation groups. Training was conducted (scenarios 1-4) using the progeny of all but one sire or grandsire of the 4 sires/grandsires in the population, (5) using only F_2 progeny (ET NA and natural service), (6) by random assignment, and (7) using only Cycle 1 progeny (Table 1). In the case of random assignment, the number of animals included in the validation population was the average of scenarios 1–5. Summary statistics of additive