ASA EXPERIENCE WITH INCORPORATING GENOMICS INTO GENETIC EVALUATION

Lauren R. Hyde1,2, Wade R. Shafer1, Stephen C. McGuire1, Mahdi Saatchi3, and Dorian J. Garrick3

1American Simmental Association, Bozeman, MT
2Red Angus Association of America, Denton, TX
3Iowa State University, Ames

Background

At the annual meeting of the American Simmental Association (ASA) in January 2011, the Board of Trustees voted unanimously to fund and initiate the development of genomically enhanced expected progeny differences (GE-EPD). The project required a large number of DNA samples representing heavily used Simmental-influenced bulls. The association already had a repository of DNA samples on many of these bulls from donations made by Simmental breeders over several years. Bull studs (ABS, Accelerated Genetics, Genex and Select Sires) and cooperators in ASA’s Carcass Merit Program also made significant contributions to the DNA repository. The project to develop GE-EPD was the result of a multi-year collaborative effort between the ASA and several research entities, including the USDA, the University of Illinois, the University of Missouri, Montana State University, Iowa State University, the National Beef Cattle Evaluation Consortium (NBCEC), and GeneSeek.

Development of ASA’s 50K Training Panel

Genotyping of DNA samples was completed in the summer of 2011 either at the University of Missouri in Columbia or at GeneSeek in Lincoln, NE. Most of the 2,703 samples were genotyped with the BovineSNP50 BeadChip (Illumina, San Diego, CA), but 264 samples were genotyped with the BovineHD BeadChip (Illumina, San Diego, CA) (Saatchi et al., 2012).

Scientists at Iowa State University, led by Dorian Garrick, executive director of the NBCEC, computed deregressed EBV (DEBV) (Garrick et al., 2009) and corresponding weighting factors from EPD and accuracies supplied by ASA on the genotyped animals, their sires and their dams. Application of the BayesC method produced estimates of SNP marker effects, which were used to derive direct genomic breeding values (DGV). The researchers evaluated the accuracy of the DGV using K-means clustering and a 5-fold cross-validation strategy (Saatchi et al., 2012).

The DGV and DEBV from all 5 validation sets were then fitted together in a weighted bivariate animal model to estimate (co)variance components. The estimated covariance between DGV and DEBV was used to estimate genetic correlations that represent the accuracy of genomic predictions for each trait, and these ranged from .29 to .65 across 12 different traits (Table 1; Saatchi et al., 2012).
As shown in Table 1, genetic correlations for the ASA 50K panel were comparable to those of commercially available panels developed for the American Angus Association (Shafer, 2012). The results indicated that, for many traits, incorporation of DNA test results into ASA’s multibreed genetic evaluation system would add a significant amount of information, improving accuracy of predicted genetic merit on young animals, and in turn providing the opportunity to increase rate of genetic change within the population.

Table 1. Genetic correlations between DGV and traits developed for two breed associations by three different providers

<table>
<thead>
<tr>
<th>Trait</th>
<th>AAA (^1)</th>
<th>ASA (^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Igenity (Neogen)</td>
<td>Pfizer (Zoetis)</td>
</tr>
<tr>
<td>Direct calving ease</td>
<td>0.47</td>
<td>0.33</td>
</tr>
<tr>
<td>Birth weight</td>
<td>0.57</td>
<td>0.51</td>
</tr>
<tr>
<td>Weaning weight</td>
<td>0.45</td>
<td>0.52</td>
</tr>
<tr>
<td>Yearling weight</td>
<td>0.34</td>
<td>0.64</td>
</tr>
<tr>
<td>Milk</td>
<td>0.24</td>
<td>0.32</td>
</tr>
<tr>
<td>Maternal calving ease</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Stayability</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Carcass weight</td>
<td>0.54</td>
<td>0.48</td>
</tr>
<tr>
<td>Marbling</td>
<td>0.65</td>
<td>0.57</td>
</tr>
<tr>
<td>Ribeye area</td>
<td>0.58</td>
<td>0.60</td>
</tr>
<tr>
<td>Backfat</td>
<td>0.50</td>
<td>0.56</td>
</tr>
<tr>
<td>Shear force</td>
<td>na</td>
<td>na</td>
</tr>
</tbody>
</table>

\(^1\)American Angus Association.

\(^2\)American Simmental Association (including both Simmental and SimAngus).

**Business Model**

Following the successful development of ASA’s 50K test, we established a unique business model to serve the needs of our membership. Rather than relying on commercial entities to manage genotypes and provide molecular breeding values (MBV), ASA sends samples for genotyping directly to GeneSeek, a Neogen company that is collaborating with NBCEC. The resulting raw genotypes can then be directly accessed by ASA, along with the corresponding MBV that are computed at GeneSeek after GeneSeek imputes the genotypes to represent those used in the prediction equations previously developed from the ASA training population. The genotypes are routinely shared with NBCEC for ongoing research including derivation of improved prediction equations. This model gives the ASA considerable flexibility (Spangler, 2012), including the opportunity to compute MBV in-house or to use the genotypes directly in genetic evaluation.

**Incorporation of MBV into Genetic Evaluation**

In 2012 we developed the infrastructure required to incorporate DNA test results into ASA’s multibreed international cattle evaluation (MB-ICE) and herdbook interim systems. In the genetic evaluation system, MBV were treated like external EPD as described by Quaas and Zhang (2006). In the interim system, we applied Kachman’s (2012) blending method with breed
effect included in the base adjustment B. When ASA starts running more frequent evaluations, the herdbook interim system will not be necessary and will be eliminated.

External EPD Approach

The main advantage of the external EPD approach is that individual accuracies of MBV are taken into account, resulting in more accurate GE-EPD. Accuracies of MBV are largely dependent on how well an animal is connected through pedigree ties to the population of animals used to develop the test. The MBV on animals with several close relatives in the training population will be more informative (i.e., more accurate) than those on animals with few or no close relatives. Failure to account for differences in MBV accuracy will not properly weight the information provided by the MBV, resulting in less accurate or biased GE-EPD.

We first incorporated DNA in our fall 2012 MB-ICE, in which we used the external EPD approach to compute GE-EPD for calving ease, weight traits and carcass traits using MBV with reliabilities less than or equal to 0.50 on 1,263 animals. After release, we noted that some of these animals had higher accuracy values than expected given the amount of information on them. We attributed this to the method we use for approximating prediction error variances (PEV).

Approximation of Prediction Error Variance

Actual PEV can be computed from the inverse of the coefficient matrix of Henderson’s mixed model equations. However, for most beef cattle evaluations, there are millions of equations so it is computationally impossible to invert the coefficient matrix. Methods to approximate PEV have been developed, but none are perfect.

In the MB-ICE, PEV are approximated using a Taylor series of expansion of inverse matrices (Wang et al., 1995). During expansion some off-diagonal elements are ignored, resulting in artificially high accuracies on some animals (R. L. Weaber, Kansas State Univ., Manhattan, personal communication). Although ASA has developed and implemented a procedure that tempers inflated accuracies, it is computationally expensive and does not lend itself well to routine genetic evaluation. For this reason, we used only the blending method to compute GE-EPD and corresponding accuracies for the spring and fall 2013 MB-ICE.

Scaling of DGV

After the fall 2013 MB-ICE was released to the public, it became apparent that animals with better (worse) than average non-enhanced EPD had even better (worse) GE-EPD. In theory, however, DNA test results should have an equal chance of increasing or decreasing an animal’s EPD, and any changes should be independent of the previous prediction of genetic merit.

Mathematically, this means that the correlation between the original EPD and the change in EPD (i.e., the difference between the original EPD and the GE-EPD) should be 0. In addition, from Reverter et al. (1994), the expected value of the regression of the GE-EPD (more accurate EPD) on the original EPD (less accurate EPD) should be 1.
From analysis of birth weight EPD, within-fold correlations between original EPD and changes in EPD for the 5 cross-validation groups and a group representing animals genotyped since development of the initial set of prediction equations were -0.21, -0.20, 0.15, 0.17, 0.19 and 0.24. Corresponding regressions were 0.92, 0.91, 1.02, 1.04, 1.06 and 1.05. Although the regressions were generally close to 1, the correlations were not close to 0. Analysis of other traits revealed similar relationships.

These results can reflect inappropriate scaling of the DGV, or double counting of information in the DGV and EPD. Multiplicative rescaling factors were derived within-fold for all traits. Rescaling of the DGV moved the correlations closer to 0 and the regressions closer to 1. The ASA’s interim system was reprogrammed to accommodate the rescaling factors and the correct handling of breed effects, and the association re-released blended EPD on over 5,000 Simmental-influenced animals on November 1, 2013.

Although the ASA has had some negative experiences incorporating genomics to its MB-ICE, the association maintains that implementation of any new technology can lead to unforeseen difficulties and firmly believes in the potential of genomics to advance beef cattle breeding.

**Literature Cited**

ASA. 2012. 50K testing for SimGenetic DNA-enhanced EPDs.  


