Incorporating DNA Information Into EPDs for Angus Cattle and Potential for Other Breeds

Matthew L. Spangler

INTRODUCTION

Genomic information, in the form of Single Nucleotide Polymorphisms (SNP), has always held the promise to increase the accuracy of Expected Progeny Differences (EPD). This promise has finally been realized for those breeds that incorporate this information into their EPD calculations. For those breeds that have not, genomic information for complex traits (those controlled by many genes) is available to producers in a disjoined context and is published separately from EPD.

One key advantage to genomic predictors (i.e. Molecular Breeding Values (MBV)) is that this information can be garnered early in the life of the animal thus enabling an increase in the accuracy of EPD particularly on young animals which have not yet produced progeny. However, the benefit of the inclusion of genomic predictions into EPD estimates is proportional to the amount of genetic variation explained by the genomic predictor (Thallman et al., 2009). In beef cattle to date, only the American Angus Association has produced marker-assisted EPD (MacNeil et al., 2010) although several other breeds are moving towards this goal (Saatchi et al., 2011a and 2011b).

BACKGROUND

The US Beef Industry has witnessed considerable evolution in terms of the genomic tests available in the market place. The tests that are currently being included in EPD are comprised of either 384 SNP or 50,000 (50K) SNP, although the research community is commonly using 50K or 770K genomic tests for discovery of "novel" traits (i.e. feed efficiency, disease susceptibility). The American Angus Association (AAA) began including genomic predictions into EPD calculations to producer Marker-Assisted EPDs (MA-EPD) in 2009.

Marker-Assisted EPD were first estimated for carcass traits and then evolved to other production traits for which EPD already existed. This is due to the need for phenotypes to train (process of developing prediction equations using significant SNP above some threshold) the genomic predictions. Consequently, genomic tests for "novel" traits such as different measures of efficiency or disease susceptibility require a significant effort in order to build large resource populations of animals with both phenotypes and genotypes. These two particular suite of traits (feed efficiency and Bovine Respiratory Disease) are currently the focus of two integrated USDA projects.

IMPLEMENTATION

The underlying question commonly asked by producers is "does it work?". It is critical to understand that this is not a valid question, as the true answer is not binary (i.e. yes or no). The important question to ask is "how well does it work?", and the answer to that question is related to how much of the genetic variation the marker test explains. The magnitude of the benefits will depend on the proportion of genetic variation (%GV) explained by a given marker panel, where the %GV is equal to the square of the genetic correlation multiplied by 100. Table 1 shows the relationship between the genetic correlation (true accuracy), %GV and Beef Improvement Federation (BIF) accuracy. BIF accuracy is the standard for all U.S. beef breeds. Table 2 summarizes the genetic correlations for the two tests that AAA currently utilizes.

MacNeil et al., (2010) utilized Angus field data to look at the potential benefits of including both ultrasound records and MBV for carcass traits in genetic evaluations.

Table 1. The relationship between true accuracy (r), proportion of genetic variation explained (%GV), and Beef Improvement Federation (BIF) accuracy.

r	%GV	BIF
0.1	1	0.005
0.2	4	0.020
0.3	9	0.046
0.4	16	0.083
0.5	25	0.132
0.6	36	0.200
0.7	49	0.286

Matthew L. Spangler, Ph.D.
Assistant Professor / Beef Genetics Extension Specialist
University of Nebraska
C204f Animal Science, Lincoln NE 68583-0908
mspangler2@unl.edu

Table 2. Genetic correlations (rg) between traits and their genomic indicators used by the American Angus Association by company.

Trait	Igenity rg (384 SNP)	Pfizer rg (50K SNP)
Calving Ease Direct	0.47	0.33
Birth Weight	0.57	0.51
Weaning Weight	0.45	0.52
Yearling Weight	0.34	0.64
Dry Matter Intake	0.45	0.65
Yearling Height	0.38	0.63
Yearling Scrotal	0.35	0.65
Docility	0.29	0.60
Milk	0.24	0.32
Mature Weight	0.53	0.56
Mature Height	0.56	0.56
Carcass Marbling	0.65	0.57
Carcass Ribeye Area	0.58	0.60
Carcass Fat	0.50	0.56
Carcass Weight	0.54	0.48

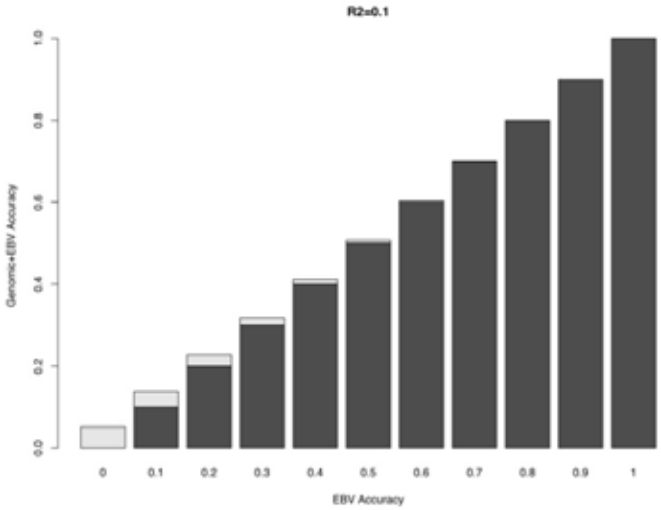

Figure 1. Increase in accuracy from integrating genomic information that explains 10% of the genetic variation into Estimated Breeding Values (EBV).

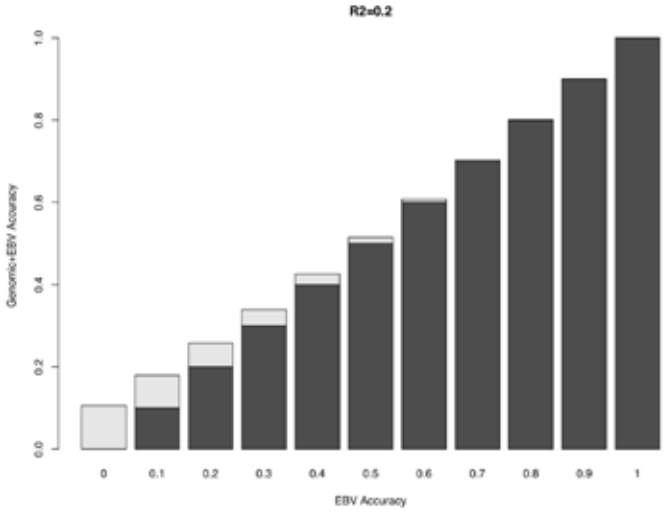

Figure 2. Increase in accuracy from integrating genomic information that explains 20% of the genetic variation into Estimated Breeding Values (EBV).

The MBV evaluated were produced specifically for Angus cattle and provided to AAA by Igenity. The MBV were developed using genotypes and EPD from 1,710 Angus bulls. The genetic correlations between the MBV and carcass traits are reflected in table 2 above. Although the genetic correlations between the MBV and the Economically Relevant carcass traits are moderate, they are not perfect predictors.

In contrast to the thought process of DNA marker panel results being a separate and disjoined piece of information, these test results should be thought of as a potentially useful indicator that is correlated to the trait of interest. As such, the MBV can be included in National Cattle Evaluations (NCE) as a correlated trait following methods of Kachman (2008). This is the approach that AAA is currently using. Other methods have been proposed including "blending" the EPD and MBV which is the equivalent to forming an index of the two where the index weights reflect the accuracy of the two components. Yet another approach is to use the actual SNP genotypes to form a genomic relationship matrix that would allow for known relationships between animals based on genotypes across SNP loci (Hayes et al., 2009; Legarra et al., 2009). The latter approach requires access to the genotypes, not just the MBV. Combining these sources of information, molecular tools and traditional EPD, has the potential to allow for the benefits of increased accuracy and increased rate of genetic change. Increased rate of genetic change can occur by increasing the accuracy of EPD, and thus the accuracy of selection, and by decreasing the generation interval. This decrease in the mean generation interval could occur particularly for sires if they are used more frequently at younger ages given the increased confidence in their genetic superiority due to added genomic information.

Figures 1-4 illustrate the benefits of including a MBV into EPD (or Estimated Breeding Value (EBV) which is twice the value of an EPD) accuracy (on the BIF scale) when the MBV explains 10, 20, 30, or 40% of the genetic variation (GV), which is synonymous with R^2 values of 0.1, 0.2, 0.3, and 0.4. The darker portion of the bars shows the EPD accuracy before the inclusion of genomic information and the lighter colored portion shows the increase in accuracy after the inclusion of the MBV into the EPD calculation. As the %GV increases, the increase in EPD accuracy becomes larger. Additionally, lower accuracy animals benefit more from the inclusion of genomic information and the benefits decline as the EPD accuracy increases. Regardless of the %GV assumed here, the benefits of including genomic information into EPD dissipate when EPD accuracy is between 0.6 and 0.7. On the other hand, when %GV is 40, an animal with 0

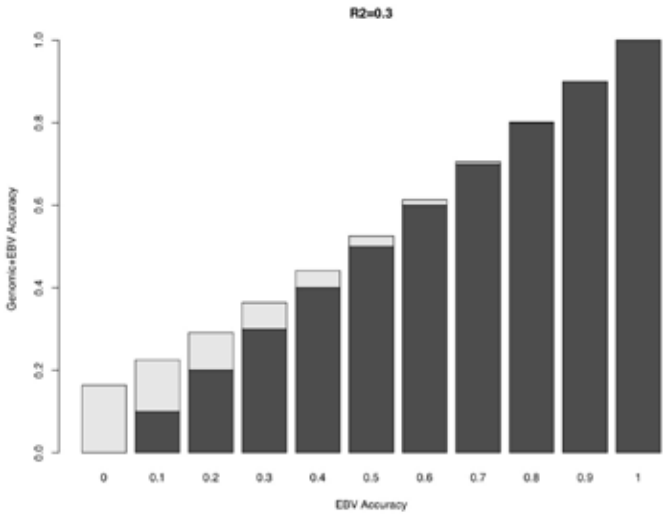

Figure 3. Increase in accuracy from integrating genomic information that explains 30% of the genetic variation into Estimated Breeding Values (EBV).

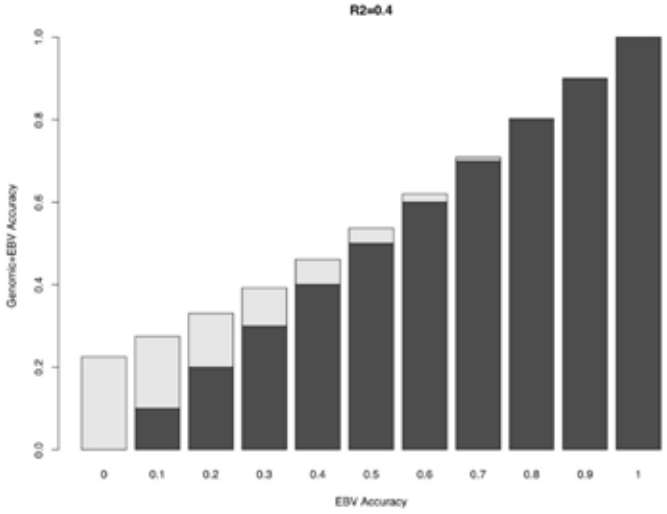

Figure 4. Increase in accuracy from integrating genomic information that explains 40% of the genetic variation into Estimated Breeding Values (EBV).

accuracy could exceed 0.2 accuracy with genomic information alone. This would be comparable to having approximately 4 progeny for a highly heritable trait or 7 progeny for a moderately heritable trait (Table 3).

CURRENT WORK IN OTHER BREEDS

Although AAA was the first to augment their EPD with genomic information, several other breeds have shown interest in taking advantage of this technology. Saatchi et al., (2011a and 2012) has shown moderate to high genetic correlations between several traits of interest and MBV for Hereford and Limousin (carcass traits only). Kachman et al., (2012) used growth traits (weaning weight and yearling weight) to illustrate the efficacy of BovineSNP50 (50,000 SNP assay) based MBV when the MBV was evaluated in the same breed as training and when it was evaluated in a different breed than training. Three single-breed MBV were created for each growth trait: Angus specific, Hereford specific and Limousin specific. The authors showed that when the MBV is used in the same breed that it was trained in, typical genetic correlations were between 0.28 and 0.42. However, the same authors found that when a breed-specific MBV was used in a different breed, the genetic correlations clustered around zero. This shows the unfortunate breed specificity issues surrounding these tools. This is consistent with other results that show the predictive power of MBV begin to erode as the genetic distance between the training and target (or evaluation) populations increase (Ibanez-Escriche et al., 2009; Toosi et al., 2010).

Table 3. Approximate number of progeny needed to reach accuracy levels (true (r) and the BIF standard) for three heritabilities (h^2).

	Accuracy	Heritability Levels		
r	**BIF**	**h^2 (0.1)**	**h^2 (0.3)**	**h^2 (0.5)**
0.1	0.01	1	1	1
0.2	0.02	2	1	1
0.3	0.05	4	2	1
0.4	0.08	8	3	2
0.5	0.13	13	5	3
0.6	0.2	22	7	4
0.7	0.29	38	12	7
0.8	0.4	70	22	13
0.9	0.56	167	53	30
0.999	0.99	3800	1225	700

Some breeds do not have the luxery of immediately having thousands of genotyped aniamls for use in developing a breed-specific genomic test. Consequently, the use of a robust across-breed set of genomic prediction equations would be beneficial. There are two primary methods of constructing an across-breed training data set: Pool purebred animals from multiple breeds or use crossbred animals. The first option requires the use of de-regressed EPD (Garrick et al., 2009) as "phenotypes" for training similar to the within breed scenario with the exception of correcting for breed effects in the model. The second option requires the use of adjusted phenotypes to train the genomic predictors. Weber et al., (2012) and Kachman et al., (2012) both evaluated the efficacy of across breed genomic predictors derived from two training data sets: the USMARC Germ Plasm Evaluation Project (GPE), and the USMARC 2,000 Bull Project. Both authors showed moderate genetic correlations between MBV and growth traits using the 2,000 Bull MBV in multiple purebred beef breeds. Both authors also showed lower genetic correlations when using the GPE derived MBV for growth traits across multiple purebred populations. The difference between the two across-breed MBV is that the 2,000 Bull training population leverages more information, since

the phenotypes are really de-regressed EPD that include several progeny records, while the GPE MBV relies on adjusted phenotypes. So while more genotyped animals were used to train the GPE MBV, the amount of phenotypic information used in training was less. Kachman et al., (2012) concluded that developing MBV using a training population of a pooled group of purebred animals can produce reliable MBV if the breed in which the MBV is to be used is also contained in the training population (i.e. if the MBV is to be used in Charolais, Charolais animals must be represented in the training data).

CONCLUSIONS

Genomics and the corresponding Marker-Assisted or Genomic-Enhanced EPD, have become a reality. Within-breed genomic predictions based on 50K genotypes have proven to add accuracy, particularly to young bulls, for several traits. The push going forward will be the adoption of this technology by other breed associations. Furthermore, methodology related to the use of this technology in crossbred or composite cattle is critically needed. The crux of adoption will be getting commercial bull buyers to see the value in, and thus pay, for increased EPD accuracy. There is still a need to collect and routinely record phenotypic information by seedstock producers. Commercial producers need to realize that EPD, and economic index values, are the currency of the realm for beef cattle selection. Genomic technology only makes these tools stronger, it does not replace them.

REFERENCES

Garrick D J; Taylor J F; Fernando R L. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics Selection Evolultion 41: 55.

Hayes B; Visscher P; Goddard M. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91:47-60.

Ibanez-Escriche N; Fernando R L; Toosi A; Dekkers J C. 2009. Genomic selection of purebreds for crossbred performance. Genetics Selection Evolution 41: 12.

Kachman S. 2008. Incorporation of marker scores into national cattle evaluations. Proc. 9th Genetic Prediction Workshop, Kansas City, MO, pp. 88-91.

Kachman S D; Spangler M L; Bennett G L; Hanford K J; Kuehn L A;
Snelling W A; Thallman R M; Saatchi; Garrick D J; Schnabel R D; Taylor J F; Pollak E J. 2012. Comparison of within and across breed trained molecular breeding values in seven breeds of beef cattle. Genetic Selection Evolutions (submitted).

Legarra A; Aguilar I; Misztal I. 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science 92**:**4656-4663.

MacNeil M D; Nkrumah J D; Woodward B W;Northcutt S L. 2010. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. Journal of Animal Science 88: 517.

Northcutt S L. 2011. Genomic Choices. http://www.angus.org/AGI/GenomicChoice070811.pdf Accessed 09/29/11.

Saatchi M; Garrick D J; Ward J. 2011a. Accuracy of genomic breeding values in Hereford beef cattle using k-means clustering method for cross-validation. Journal of Animal Science 89 E-Suppl. 2: 59.

Saatchi M; Garrick D J; Taylor J F. 2012. Genetic correlations between carcass traits and genomic breeding values in Limousin cattle. Journal of Animal Science 90 E-Suppl. 2: 66.

Saatchi M; McClure M C; Rolf M M; Kim J W; Decker J E; Taxis T M;
Chapple R H; Ramey H R; Northcutt S L; Bauck S;Woodward B; Dekkers J C M; Fernando R L;Schnabel R D; Garrick D J; Taylor J F. 2011b. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genetics Selection Evolultion 43: 40.

Thallman R M; Hanford K J; Quaas R L; Kachman S D; Tempelman R J; Fernando R L; Pollak E J. 2009. Estimation of the Proportion of Genetic Variation Accounted for by DNA Tests. Proceedings of the 41st Annual research symposium and annual meeting, Beef Improvement Federation: 184-209.

Toosi A;Fernando R L; Dekkers J C. 2010. Genomic selection in admixed and crossbred populations. Journal of Animal Science 88: 32-46.

Weber K L; Thallman R M; Keele J W; Snelling W M; Bennett G L; Smith T P L; McDaneld T G; Allan M F; VanEenennaam A L; Kuehn L A. 2011. Accuracy of genomic breeding values in multi-breed beef cattle populations derived from deregressed breeding values and phenotypes. Journal of Animal Science (Submitted).