

THE EFFECT OF QUANTITY AND BREED COMPOSITION OF GENOTYPES FOR GENOMIC PREDICTION IN PUREBRED OR CROSSBRED CATTLE

Heather Bradford¹

¹Kansas State University, Manhattan

Introduction

The implementation of genomics enabled producers to more accurately select young animals for breeding resulting in a decrease in generation interval. Beef cattle typically have long generation intervals compared with species like poultry and swine, and genomic selection can increase response to selection. Genomic selection should have the most benefit for traits that are hard to measure, measured late in life, sex-limited, and measured after harvest (Hayes and Goddard, 2010). Traits like female fertility are sex-limited and difficult to measure while being economically relevant to producers. Selection for female fertility would benefit greatly from the inclusion of genomic data to increase accuracy. There are many economically relevant traits that beef producers could better select for by using genomics.

Review of Literature

Linkage

A quantitative trait loci (QTL) is a gene that affects a quantitative trait. Phenotype results from the total of the effects of all QTL including dominance and any gene interactions, environment, and the interaction of genetics and environment. A single nucleotide polymorphism (SNP) is a single base difference in a DNA sequence that may or may not be located within a gene. Linkage disequilibrium (LD) results when a SNP allele and QTL allele are linked and inherited together more often than expected (Hayes and Goddard, 2010). Meusissen et al. (2001) first proposed genomic selection using all SNP markers simultaneously. This method relies on dense SNP panels with

the intent that all QTL affecting a trait are in LD with at least 1 marker (Hayes and Goddard, 2010).

Because of the LD between SNP and QTL, the association between SNP and QTL affecting a trait of interest can be used to create genomic predictions. These genomic predictions are the sum of the effect of each SNP on the trait of interest. Across-breed LD is much more restricted than within-breed LD due to differential selection since the divergence of individual breeds (Hayes and Goddard, 2010). Because of the difference in LD across breeds, genomic predictions historically needed to be breed-specific. The accuracy of genomic predictions was largely the result of LD, and the loss of LD resulted in less accuracy in subsequent generations (Habier et al., 2007). Thus, SNP effects have to be periodically re-estimated because of the erosion of LD.

A population of genotyped animals with phenotypes or very accurate breeding values is typically used for estimating SNP effects. This group is referred to as a training or reference population while a separate group of animals with genotypes and phenotypes or breeding values is the validation population. The SNP effects estimated from the training population are used to predict genetic merit in the validation population. The accuracy of the genomic prediction, referred to as a genomic breeding value (GBV), direct genomic value (DGV), or molecular breeding value (MBV), can then be assessed in the validation population, because the validation population was independent of the training population that was used to develop the predictions. Genetic correlations between the genomic prediction and phenotypic trait data can be estimated with a two trait animal model using REML (Kachman, 2008). The square of this genetic correlation is the

percent of the additive genetic variance that was explained by the genomic test (Thallman et al., 2009). Genomic results are then incorporated into traditional genetic evaluations resulting in genomic-enhanced expected progeny differences (GE-EPD).

Number of Genotyped Animals

As more animals are genotyped, researchers can better estimate SNP effects resulting in more accurate genomic predictions. Simulations with various training population sizes and relationships to the validation population showed accuracy increases when the size of the training population increases, even if those animals are many generations removed from the validation population (Saatchi et al., 2010). In a study of 9 breeds for feed efficiency and carcass traits, breeds with larger training populations had greater accuracies than average (Bolormaa et al., 2013). The number of animals in training had a greater impact on lowly heritable traits, and the relationship to the training population became less important for those traits (Saatchi et al., 2010). As beef cattle training populations increase, the greatest impact on accuracy should be for lowly heritable traits.

As the adoption of genomic technology has increased, there has been the opportunity to evaluate the realized increase in accuracy resulting from an increase in the size of the reference population. Direct genomic values for Uruguayan Herefords were more accurate when using predictions for American ($n = 1,081$) instead of Uruguayan Herefords ($n = 395$; Saatchi et al., 2013). The difference in accuracy likely resulted from the larger training population for the American Hereford prediction and not the relationship to the population used in training. There was a linear increase in accuracy exceeding 0.10 as the size of the reference population increased from 500 to 2,500 head of crossbred sheep (Daetwyler et al., 2012b). However, the increase in accuracy would not be expected to continue to increase linearly as the size of the reference population continues to increase. Predictions from small reference populations with fewer than 1,000 individuals become considerably more accurate as more animals are included in the reference population.

Accuracies continue to improve as the reference population grows from 1,000 to 1,000's of individuals. When combining 3 Nordic Red populations with individual reference populations of 1,562 animals or fewer, reliabilities increased by a magnitude of 3 to 8% on average with a total reference population of 3,735 animals (Brøndum et al., 2011). Increasing the training population from 1,300 to 5,250 animals while using the same methodology resulted in predictions that on average explained 18% more genetic variation and increased accuracy by 0.40 (Bodhareddy et al., 2014). Reliabilities were 5 to 32% greater when using a combined Chinese and Nordic Holstein population of 7,387 instead of only 2,171 Chinese Holsteins (Zhou et al., 2013). Predictions were more accurate even when combining populations of the same breed from different countries. There has been a consistent increase in accuracy as more animals were added to small to moderate sized reference populations.

Research in dairy cattle has evaluated the impact of larger training populations when there are many genotyped animals. Reliabilities increased 10% on average when combining European Holstein populations to create reference populations with more than 9,000 bulls compared with individual country reference populations with 3,000 to 4,000 bulls (Lund et al., 2011). Improvements can still be made when reference populations contain several thousand head. Adding 3,593 foreign bulls to the U.S. Holstein evaluation with over 10,000 genotypes increased reliability by 2 to 3% (VanRaden et al., 2012). As the number of genotypes increases, the improvement in accuracy from a larger reference population isn't as substantial and would be expected to continue to decline.

Because of the accuracy increase that results from larger training populations, combining genotyped populations to develop genomic predictions is of interest. The main focus in the beef industry has been genotyping purebred populations to develop predictions that are then used within that breed. Due to the cost of genomic testing and the number of proven animals needed for training, combining these populations could improve prediction accuracy. However, a very diverse training population could result in less accurate predictions because the training population is now less related to the individual breeds that are being

FRANK BAKER MEMORIAL SCHOLARSHIP

predicted. Many simulations have been performed in addition to research in beef cattle and other species to evaluate the impact of the relationship between the training and validation populations.

Relationship between Training and Validation

The relationship between the reference animals and the populations that the genomic predictions will be used in affects accuracy. Daetwyler et al. (2012a) demonstrated that a large proportion of the accuracy of predictions results from the strong relationship between the reference and validation populations. When the training population consisted of generations that were more similar to the validation population, prediction accuracy was greater than if distant generations were used (Saatchi et al., 2010; Pszczola et al 2012). The animals in the more recent generations tend to be more related to the young animals in which the genomic tests are being used. The importance of the relationship of the training and validation populations likely resulted from recombination that took place between generations and reduced the LD between the markers and QTL (Saatchi et al., 2010). As LD erodes, accuracy decreases when a different SNP is associated with the QTL than in the reference population. When animals had a greater average squared relationship to the reference population, those animals had greater reliabilities (Pszczola et al., 2012). If an animal has relatives in the reference population, more confidence can be placed on the resulting genomic predictions because relatives with similar LD between SNP and QTL were used to estimate SNP effects. Because of the importance of the relationship between the reference and validation populations on accuracy, there has been much research on the breed specificity of genomic predictions.

Within-breed Prediction

In the beef industry, much emphasis has been placed on developing predictions for use within-breed with the results incorporated into national cattle evaluations. A simulation study by Kizilkaya et al. (2010) found slightly greater correlations between true and estimated breeding values when training and validating in purebreds compared with training and validating in a multibreed population. Estimates for

genetic correlations between GBV and the respective traits were 0.14 to 0.81 in Angus (Saatchi et al, 2011; Northcutt, 2013; Boddhireddy et al., 2014), 0.18 to 0.52 in Hereford (Saatchi et al., 2013), 0.39 to 0.76 in Limousin (Saatchi et al., 2012), and 0.29 to 0.65 in Simmental (Saatchi et al., 2012). Generally, there was sufficient LD between SNP and genes for traits included in national cattle evaluations to achieve strong genetic correlations. Because of these results, several beef breed associations currently publish GE-EPD.

Connectedness within-breed can also affect the accuracy of predictions for animals that are distantly related to the training population. Genomic predictions developed for American Herefords were less accurate when used in Argentinian, Canadian, or Uruguayan Herefords, possibly resulting from lesser relationships to the training population or genetic by environment interactions (Saatchi et al., 2013). When comparing reliabilities for Red dairy cattle from 3 European countries, the within country predictions were always more reliable than if predictions were developed in 1 country and used in the others (Brøndum et al., 2011). Because the reference populations were of similar size, the loss in accuracy again resulted from the lack of connectedness between countries. Further analysis revealed that the correlation of LD phase between countries ranged from 0.46 to 0.86 (Brøndum et al., 2011). This correlation suggests there has been some divergence in the Red breed in these countries which impacts the ability to develop across-country genomic predictions. When using the American Hereford prediction in Argentinian Herefords, animals with American Herefords in their pedigree had on average greater correlations between DEBV and GBV than those without American genetics (Saatchi et al., 2013). Similarly, DGV accuracies were less when using Angus predictions in an Angus herd that was closed for many generations and was less related to the training population (Saatchi et al., 2011). These studies demonstrated the importance of the training population being a representative sample of a breed to obtain accurate estimates of genomic merit across the population.

Across-breed Prediction

It would be convenient if predictions could be developed in 1 breed and used for other breeds, but

this approach has produced very poor accuracies. Simulations trained in one breed and predicted in another resulted in significantly less accuracy than training in the breed of interest (Toosi et al., 2010). When Angus trained predictions were used in other breeds, simulations accounted for less than one-third of the genetic variation that was achieved in Angus (Kizilkaya et al., 2010). These simulations did not demonstrate favorable results for using breed-specific predictions across breeds. When breed-specific predictions for Angus, Hereford, or Limousin were used across breeds, in most cases the genetic correlation was not significant and in a few instances was slightly negative despite a moderate, positive genetic correlation when validating within-breed (Kachman et al., 2013). On average, the genetic correlation from using predictions developed in Herefords on Angus sires was not different from 0 while predictions developed specifically for Angus had the greatest accuracies (0.24 to 0.61; Weber et al., 2012b). Similar results were observed in Holstein and Jersey dairy cattle (Hayes et al., 2009). Using 50,000 SNP has not been sufficient density to use predictions across breeds. Breeds have different LD which erodes the accuracy of predictions developed in one breed when the prediction was used in another breed. The use of genomic predictions across breeds has not been feasible due to the limited prediction accuracy.

Multi-breed for Purebred Prediction

Another approach would be to combine data for many purebreds to develop predictions that were then used for individual breeds. Combining populations resulted in greater accuracy, especially for lowly heritable traits, than training on each population individually (de Roos et al., 2009). Predictions were more accurate when trained in a multi-breed population instead of a purebred population and validated in the same purebred (Bolormaa et al., 2013). This possibly results from capturing more of the variants within the breed of interest. Genetic correlations averaged 0.47 (0.10 to 0.73) when trained on only Simmentals and 0.55 (0.18 to 0.91) when trained on Simmental, Angus, Red Angus, Gelbvieh, Brangus, Hereford, and Charolais (Saatchi and Garrick, 2013). This is of interest because animals registered with the American Simmental Association do not have to be

purebred Simmental. Only calving ease maternal and weaning weight maternal were not more accurate with the multi-breed training population (Saatchi and Garrick, 2013). Stayability was unchanged because there was no information for this trait in the other breeds. With the additional breeds, the size of the reference population more than doubled compared with only Simmental animals (Saatchi and Garrick, 2013). These studies suggest a benefit from combining single breed reference populations. Not only is the size of the reference population greater, but the reference population can capture more of the variation within the breed of interest.

Predictions that were developed without including the breed of interest were less accurate than if that breed had been included in training. The accuracy of multi-breed prediction in Australian sheep was always less if the breed for prediction was excluded from the training population (Daetwyler et al., 2012a). However, if the breed to be predicted was included in the reference population, multi-breed predictions were no more accurate than single-breed predictions (Pryce et al., 2011). If the within-breed predictions are based on a large enough reference population, the potential benefit from multi-breed predictions might be very minimal. If multi-breed predictions someday achieved equivalent or greater accuracies than within-breed predictions, all breeds of interest would need to be included in the training population.

There was very little difference in the accuracy of GBV when using a Holstein or Holstein and Jersey reference population to validate in Holsteins, but there was an increase in accuracy when using the combined instead of the Jersey reference population to validate in Jerseys (Hayes et al., 2009). This could result from the very small number of Jersey bulls with genotypes. The addition of more genotypes, despite the breed, allowed for more accurate genomic predictions. Using a Holstein, Jersey, and Brown Swiss reference population, resulted in an increase in accuracy for some traits in Jersey and Brown Swiss above that of the single-breed prediction (Olson et al., 2012). Again, there was no benefit for Holsteins to use multi-breed predictions but some benefit for smaller breeds with fewer genotypes. Incorporating 2 breeds in the reference population to predict a third breed increased prediction

FRANK BAKER MEMORIAL SCHOLARSHIP

accuracy compared with using 1 of the breeds to predict a different breed (Pryce et al., 2011). Again, the improvement in accuracy could result from the larger reference population used to predict marker effects in a different breed. Given a reference population of sufficient size, there has been no consistent benefit to using a multi-breed population to develop predictions for use in purebreds. Yet, while training populations of sufficient size are being collected, multi-breed predictions could help improve accuracy until enough animals were genotyped to produce reliable predictions.

Crossbred for Purebred Prediction

Another scenario is collecting data on crossbred animals for use in purebreds although this situation is unlikely in the beef industry with a lack of pedigree and performance recording in crossbred cattle. Simulations demonstrated, as the number of breeds represented in the crossbred population increased, the accuracy of predicting one of the purebreds decreased (Toosi et al., 2010). This decrease in accuracy could result from a decrease in the prevalence of haplotypes from the breed of interest in the training population as population size was held constant. That breed would then have a lesser contribution to the estimation of marker effects. Using the crossbred U.S. Meat Animal Research Center Germplasm Evaluation Program (GPE) population for training and validating in a purebred population resulted in MBV accuracies generally ranging from 0.20 to 0.40 with less accurate predictions in Charolais for most traits, likely a result of limited Charolais influence in the training population (Weber et al., 2012a). Validation in the 2,000 Bull Project animals, consisting of influential bulls representing 16 beef breeds, resulted in genetic correlations ranging from 0.19 to 0.37, which were similar to validation in purebreds (Weber et al., 2012a). Greater accuracies are being achieved in the beef industry by using within-breed predictions (Saatchi et al., 2011; Saatchi et al., 2012; Northcutt, 2013; Saatchi et al., 2013; Boddhireddy et al., 2014). An analysis of crossbred sheep of primarily Merino decent resulted in greater accuracies for Merinos than for terminal breeds (Daetwyler et al., 2010). Thus, the breed makeup for the crossbred population was important, and the breed of interest needed to be well represented in the crossbred genetics. There are many challenges associated

with using crossbred genotypes in the beef industry, mainly the lack of complete pedigree and performance recording outside of research herds. The use of crossbred predictions for many breeds appears less feasible.

Another approach to using crossbred genotypes has been to model breed-specific SNP effects. Modeling with breed-specific compared with across-breed SNP effects resulted in similar prediction accuracies for a variety of simulation scenarios (Ibán z-Escriche et al., 2009). As marker density increased up to 2,000 markers on one chromosome, there was less value in using breed-specific SNP effects (Ibán z-Escriche et al., 2009). The use of breed-specific SNP effects required large breed differences to justify the additional effects in the model, and this model had an advantage when large training populations were used (Ibán z-Escriche et al., 2009). Developing reference populations of sufficient size to justify the use of breed-specific SNP effects will be challenging in the beef industry. Very few of the 2,500 SNP with the largest effect were common to the GPE and 2,000 Bulls populations (Weber et al., 2012a). These results suggest a potential need for breed-specific effects to better account for both differences in LD and the magnitude of the SNP effect across breeds.

Crossbred Prediction

Although genetic evaluation of crossbred beef cattle is not common, a cheap genomic test for economically relevant traits would be a valuable genetic selection tool for commercial producers. In addition, many beef breed associations include hybrid animals in their genetic evaluations, and breed-specific predictions might not be as accurate in those composite animals. Genomic predictions based on 3,000 SNP for feed efficiency in Angus-Brahman crosses had accuracies ranging from -0.13 to 0.36 (Elzo et al., 2012). The small number of SNP could have contributed to the limited accuracy that was achieved from those genomic predictions. The accuracy of crossbred predictions was numerically less in most cases than within-breed predictions; however, those estimates had large standard errors (Mujibi et al., 2011). Larger reference populations incorporating a broader sample of the possible breed crosses might improve accuracy as more of the population of interest would be used to

develop predictions.

Purebred or Multi-breed for Crossbred Prediction

Because most phenotypes in the beef industry are collected on purebreds, creating predictions based on the purebred data for use in selecting crossbreds could be beneficial. Training on Angus, Angus and Red Angus, or Hereford resulted in weak MBV accuracy (0.01 to 0.43) for growth and carcass traits when validating in the crossbred GPE population (Weber et al., 2012a). MBV accuracy tended to be less than that achieved with a multi-breed training population consisting of sires in the 2,000 Bull Project (Weber et al., 2012a). Using a multi-breed instead of purebred training population should produce better predictions for crossbred animals because breeds differ in the LD between SNP and QTL. Training on 2,000 Bull Project and validating on GPE yielded moderate genetic correlations (0.13 to 0.42) with little or no improvement from including breed effects in the DEBV for the 2,000 Bull Project (Weber et al., 2012a). Training on a multi-breed population instead of a purebred population increased accuracy more for composite breeds than purebreds (Bolormaa et al., 2013). Genotypes and phenotypes on purebreds can be useful to develop predictions for crossbreds. The American Angus Association and Zoetis currently market a genomic test for commercial Angus-influence cattle. This test provides predictions for a couple economically relevant traits, and commercial producers are using this test to add value to feeder cattle and to select replacement females. If the beef industry were to move toward larger scale crossbred genetic evaluation, establishing genomic predictions from existing purebred databases appears to be the most feasible method. Including many breeds in the reference population would help make these predictions more relevant for a wider array of commercial producers.

Conclusions and Implications to Genetic Improvement of Beef Cattle

As the use of genomic testing in the beef industry grows, reference populations of greater size are being established by individual breed associations. Breed-specific genomic predictions are becoming

more accurate as a result of the increase in genomic testing. Research on using genomic predictions developed in 1 breed for use in another has not been favorable. Yet, pooling genotypes from multiple breeds to develop predictions for a purebred was promising for increasing accuracy past that achieved with only the purebred genotypes. The use of multi-breed predictions could be of interest to smaller breeds with fewer genotyped animals and breeds that register percentage animals. Smaller breeds could benefit from the larger reference population that could be assembled from combining genotypes. Breeds with hybrid animals could benefit from the inclusion of the LD from other breeds to develop more accurate predictions.

There is potential benefit from genetic evaluation at the commercial level. Because there is no infrastructure for performance recording in commercial cattle, genomic testing is a more feasible option to identify the genetically superior crossbred cattle. Preliminary research has demonstrated the feasibility of developing accurate genomic predictions from purebred or multi-breed populations to use in crossbred individuals. As genomic predictions become more refined in the seedstock industry, there is the potential to develop cheaper genomic tests for economically relevant traits in crossbred cattle.

Literature Cited

- Boddhireddy, P., M. J. Kelly, S. Northcutt, K. C. Prayaga, J. Rumph, and S. DeNise. 2014. Genomic predictions in Angus cattle: Comparisons of sample size, response variables, and clustering methods for cross-validation. *J. Anim. Sci.* 92:485-497.
- Bolormaa, S., J. E. Pryce, K. Kemper, K. Savin, B. J. Hayes, W. Barendse, Y. Zhang, C. M. Reich, B. A. Mason, R. J. Bunch, B. E. Harrison, A. Reverter, R. M. Herd, B. Tier, H.-U. Graser, and M. E. Goddard. 2013. Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in *Bos taurus*, *Bos indicus*, and composite beef cattle. *J. Anim. Sci.* 91:3088-3104.

FRANK BAKER MEMORIAL SCHOLARSHIP

- Brøndum, R. F., E. Rius-Vilarrasa, I. Strandén, G. Su, B. Guldbandsen, W. F. Fikse, and M. S. Lund. 2011. Reliabilities of genomic predictions using combined reference data of the Nordic Red dairy cattle populations. *J. Dairy Sci.* 94:4700-4707.
- Daetwyler, H. D., J. M. Hickey, J. M. Henshall, S. Dominik, B. Gredler, J. H. J. van der Werf, and B. J. Hayes. 2010. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. *Anim. Prod. Sci.* 50:1004-1010.
- Daetwyler, H. D., K. E. Kemper, J. H. J. van der Werf, and B. J. Hayes. 2012. Components of the accuracy of genomic prediction in a multi-breed sheep population. *J. Anim. Sci.* 90:3375-3384.
- Daetwyler, H. D., A. A. Swan, J. H. J. van der Werf, and B. J. Hayes. 2012b. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. *Genet. Sel. Evol.* 44:33.
- de Roos, A. P. W., B. J. Hayes, and M. E. Goddard. 2009. Reliability of genomic predictions across multiple populations. *Genetics.* 183:1545-1553.
- Elzo, M. A., G. C. Lamb, D. D. Johnson, M. G. Thomas, I. Misztal, D. O. Rae, C. A. Martinez, J. G. Wasdin, and J. D. Driver. 2012. Genomic-polygenic evaluation of Angus-Brahman multi-breed cattle for feed efficiency and postweaning growth using the Illumina 3K chip. 2012. *J. Anim. Sci.* 90:2488-2497.
- Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics.* 177:2389-2397.
- Hayes, B. and M. Goddard. 2010. Genome-wide association and genomic selection in animal breeding. *Genome.* 53:876-883.
- Hayes, B. J., P. J. Bowman, A. C. Chamberlain, K. Verbyla, and M. E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:51.
- Ibán z-Escriche, N., R. L. Fernando, A. Toosi, and J. C. M. Dekkers. 2009. Genomic selection of purebreds for crossbred performance. *Genet. Sel. Evol.* 41:12.
- Kachman, S. D. 2008. Incorporation of marker scores into national genetic evaluations. In: Proc. Beef Improv. Fed. 9th Genet. Pred. Workshop., Kansas City, MO. p. 92-98.
- Kachman, S. D., M. L. Spangler, G. L. Bennett, K. J. Hanford, L. A. Kuehn, W. M. Snelling, R. M. Thallman, M. Saatchi, D. J. Garrick, R. D. Schnabel, J. F. Taylor, and E. J. Pollak. 2013. Comparison of molecular breeding values based on within- and across-breed training in beef cattle. *Genet. Sel. Evol.* 45:30.
- Kilzikaya, K., R. L. Fernando, and D. J. Garrick. 2010. Genomic prediction of simulated multi-breed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. *J. Anim. Sci.* 88:544-551.
- Lund, M. S., A. P. W. de Roos, A. G. de Vries, T. Ducrocq, S. Fritz, F. Guillaume, B. Guldbandsen, Z. Liu, R. Reents, C. Schrooten, F. Seefried, and G. Su. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. *Genet. Sel. Evol.* 43:43.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* 157:1819-1829.
- Mujibi, F. D. N., J. D. Nkrumah, O. N. Durunna, P. Stothard, J. Mah, Z. Wang, J. Basarab, G. Plastow, D. H. Crews, Jr., and S. S. Moore. 2011. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *J. Anim. Sci.* 89:3353-3361.
- Northcutt, S. L. 2013. AAA experience with incorporating genomics into genetic evaluation. In: Proc. Beef Improv. Fed. 10th Genet. Pred. Workshop., Kansas City, MO. p. 5-7.

- Olson, K. M., P. M. VanRaden, and M. E. Tooker. 2012. Multi-breed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 95:5378-5383.
- Pryce, J. E., B. Gredler, S. Bolormaa, P. J. Bowman, C. Egger-Danner, C. Fuerst, R. Emmerling, J. Sillkner, M. E. Goddard, and B. J. Hayes. 2011. Short communication: Genomic selection using a multi-breed, across-country reference population. *J. Dairy Sci.* 94:2625-2630.
- Pszczola, M., T. Strabel, H. A. Mulder, and M. P. L. Calus. 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389-400.
- Saatchi, M., S. R. Miraei-Ashtiani, A. Nejati-Javaremi, M. Moradi-Shahrehabak, and H. Mehrebani-Yeghaneh. 2010. The impact of information quantity and strength of relationship between training set and validation set on accuracy of genomic estimated breeding values. *Afr. J. Biotechnol.* 9:438-442.
- Saatchi, M., M. C. McClure, S. D. McKay, M. M. Rolf, J. W. Kim, J. E. Decker, T. M. Taxis, R. H. Chapple, H. R. Ramey, S. L. Northcutt, S. Bauck, B. Woodward, J. C. M. Dekkers, R. L. Fernando, R. D. Schnabel, D. J. Garrick, and J. F. Taylor. 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. *Genet. Sel. Evol.* 43:40.
- Saatchi, M., R. D. Schnabel, M. M. Rolf, J. F. Taylor, and D. J. Garrick. 2012. Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle. *Genet. Sel. Evol.* 44:38.
- Saatchi, M., and D. J. Garrick. 2013. Improving genomic prediction in Simmental beef cattle using a multi-breed reference population. In: Proc. West. Sec. Am. Soc. Anim. Sci., Bozeman, MT. p. 174-177.
- Saatchi, M., J. Ward, and D. J. Garrick. 2013. Accuracies of direct genomic breeding values in Hereford beef cattle using national or international training populations. *J. Anim. Sci.* 91:1538-1551.
- Thallman, R. M., K. J. Hanford, R. L. Quaas, S. D. Kachman, R. J. Templeman, R. L. Fernando, L. A. Kuehn, and E. J. Pollak. 2009. Estimation of the proportion of genetic variation accounted for by DNA tests. In: Proc. Beef Improv. Fed., Sacramento, CA. p. 184-209.
- Toosi, A., R. L. Fernando, and J. C. M. Dekkers. 2010. Genomic selection in admixed and crossbred populations. *J. Anim. Sci.* 88:32-46.
- VanRaden, P. M., K. M. Olson, D. J. Null, M. Sargolzaei, M. Winters, and J. B. C. H. M. van Kaam. 2012. Reliability increases from combining 50,000- and 777,000- marker genotypes from four countries. *Interbull Bull.* 46:75-79.
- Weber, K. L., R. M. Thallman, J. W. Keele, W. M. Snelling, G. L. Bennett, T. P. L. Smith, T. G. McDanel, M. F. Allan, A. L. Van Eenennaam, and L. A. Kuehn. 2012a. Accuracy of genomic breeding values in multi-breed beef cattle populations derived from deregressed breeding values and phenotypes. *J. Anim. Sci.* 90:4177-4190.
- Weber, K. L., D. J. Drake, J. F. Taylor, D. J. Garrick, L. A. Kuehn, R. M. Thallman, R. D. Schnabel, W. M. Snelling, E. J. Pollak, and A. L. Van Eenennaam. 2012b. The accuracies of DNA-based estimates of genetic merit derived from Angus or multi-breed beef cattle training populations. *J. Anim. Sci.* 90:4191-4202.
- Zhou, L., X. Ding, Q. Zhang, Y. Wang, M. S. Lund, and G. Su. 2013. Consistency of linkage disequilibrium between Chinese and Nordic Holsteins and genomic prediction for Chinese Holsteins using a joint reference population. *Genet. Sel. Evol.* 45:7.