SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL – SBMA
IX Simpósio Brasileiro de Melhoramento Animal
20-22 de junho de 2012, João Pessoa, PB, Brasil

9 th Biennial Symposium of the Brazilian Society of Animal Breeding
June 20-22, 2012, João Pessoa, Paraíba, Brazil

CURRENT USE AND EXPECTATIONS ABOUT THE REALISTIC APPLICATION OF GENOMICS IN ANIMAL BREEDING PROGRAMS

Matthew L. Spangler

University of Nebraska

Introduction

Genomic information has long held the promise to increase the accuracy of prediction, particularly for traits that are hard or expensive to measure. The arena of genomics has proven to be far from stagnant and the intermediate steps of commercialization of prediction equations and reduced assays has often created confusion among livestock producers, particularly in less integrated industries (i.e. beef). However, undeniable progress has been made, both in terms of basic science and application, across several species.

Paternity (Parentage) Testing

A critical factor in estimating reliable EBV is the proper determination of parentage. Errors in pedigree may have significant negative impacts on the reliability of genetic evaluations and potential genetic gains (Geldermann et al., 1986; Israel and Weller, 2000). Increases in misidentification of an animal's parentage results in progressively more biased estimates of genetic parameters and this bias severely compromises potential genetic gains from selection (Van Vleck, 1970; Senneke et al., 2004). By utilizing genomic technology to determine parentage these inaccuracies can be greatly diminished (Dodds et al., 2005). There are currently two methods being utilized to ascertain parentage; microsatellites and SNP. Both methods provide a probability of parentage, which is influenced by the sensitivity of the test and the relationship of the potential parents.

Utilizing genomics for parentage allows producers to manage multiple sire breeding pastures (i.e. sheep and beef production) and settle AI/natural sire discrepancies when birth dates alone are insufficient to make this determination inconclusiveThe use of genomics parentage testing to resolve the paternity of offspring produced by multi-sire breeding systems with subsequent use of their pedigree and phenotypes in a progeny test genetic evaluation has been proposed (DeNise, 1999; Goddard and Goddard, 1997). More recently, genomic-based pedigree structures coupled with strategically collected performance records have been used to compute EBV for both seedstock and commercial producers (Weaber, 2005; Van Eenennaam et al., 2007).

Qualitative Traits

Markers for many qualitative traits such as coat color, horned/polled and a variety of genetic defects have been identified and are commercially available. This technology can now be used to identify animals that are carriers of recessive alleles facilitating selection against the carriers, if desired; or more informed mating decisions.

Historically, when lethal recessives were identified the common method of eliminating them from the population was an aggressive culling campaign, often eliminating entire lines of seedstock

animals. On November 15, 2008 the American Angus Association (AAA) recognized Arthrogryposis Multiplex (AM) as a genetic defect that had emerged in the breed and on June 12, 2009 the same determination was made for Hydrocephalus (NH). Within a relatively short time after each defect was identified, markers were identified to clearly distinguish heterozygous (carrier) animals from those that were homozygous normal. Animals within the affected lines could then be eliminated, without the complete elimination of a prominent line of Angus cattle. Other examples such as osteopetrosisin Red Angus cattle (Meyers et al., 2010) orovine progressive pneumonia in sheep (Heaton et al., 2012) illustrate the decreased time from discovery to the availability of a diagnostic test, all enabled by genomics. The initial benefits to the livestock industry of sequencing will be the identification of deleterious mutations. Instead of dealing with these one at a time, livestock organizations will likely be faced with a plethora of "defects" to manage. This will require a level of sophisticated mating strategies that some species may not be prepared for as purging all carrier animals will not be an option, rather optimization strategies contemplating the probability of abnormalities, genetic gain in the overall breeding objective, and inbreeding will be required.

Quantitative Traits

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

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 EBV particularly on young animals, which have not yet produced progeny. However, the benefit of the inclusion of genomic predictors into EBV estimates is proportional to the amount of genetic variation explained by the genomic predictor. Although AAA was the first US beef breed association to augment their EBV with genomic information, several other breeds have shown interest in taking advantage of this technology. Saatchi et al., (2011 and 2012) has shown moderate to high genetic correlations between several traits of interest and MBV for Hereford and Limousin (carcass traits only).

Various livestock industries have witnessed considerable evolution in terms of the genomic tests available in the market place. The tests that are currently being included in EBVin the US Beef Industry are comprised of either 384 SNP or 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). In the US Beef Industry, marker-Assisted EBV were first estimated for carcass traits and then evolved to other production traits for which EBV already existed. This is due to the need for phenotypes (deregressed EBV or adjusted phenotypes) for training. 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 (Thallman et al., 2009). Table 1 summarizes the genetic correlations for the two tests that AAA currently utilizes.

Table 1. Genetic correlations (rg) between traits and their genomic indicators used by the American Angus Association by company.

Trait	Igenityrg (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

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

There are four basic ways of combining genomic and phenotypic information into a single selection tool (I. Marker or Genomic Enhanced EBV). The first method is to compute independent values, both EBV and MBV, and to then include both pieces of information in a selection index whereby each "trait" is weighted proportionally to the respective amount of genetic variation that they account for. A second approach is through genomic relationships whereby marker information is used to fit a genomic relationship matrix (relationship among animals at each SNP locus; Hayes et al., 2009; Legarra et al., 2009) that is used to augment estimated relationships based on pedigree

information. For this method it is necessary to know the actual SNP genotypes rather than having a marker score or MBV. This method is currently being used in dairy genetic evaluations. The first method deployed by the beef industry, and which is currently used by the AAA, is the correlated trait approach. MBV information is included in National Cattle Evaluation (NCE) as a correlated trait (Kachman, 2008), similarly to the way ultrasound information is utilized in a multiple trait model in the estimation of EBV for carcass traits. As the genetic correlation between the indicator trait, MBV in this case, and the trait of interest increases so does the EBV accuracy, particularly for younger (lower accuracy) animals. The final method is to treat MBV as if they were external EBV (EBV from an animal that is external to the population or breed; Quaas and Zhang, 2000). This method is currently being used by the American Simmental Association and allows for MBV to influence the accuracy of EBV differently, thus making use of the variation around the MBV estimates. This individual animal MBV prediction error variance (PEV) can vary depending on the relationship between the animal with the MBV and the training population and the contribution of an animal to the training set (i.e. higher accuracy animals contribute more to training than do lower accuracy animals).

Combining these sources of information, molecular tools and traditional EBV, 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 EBV, 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.

As the %GV increases, the increase in EBV accuracy becomes larger. Additionally, lower accuracy animals benefit more from the inclusion of genomic information and the benefits decline as the EBV accuracy increases. The benefits of including genomic information into EBV dissipate when EBV accuracy is between 0.6 and 0.7. However, for an animal that has an accuracy of 0, as might be the case for traits that are new or not densely recorded, including MBV into its EBV that account for 40% GV would increase said animals accuracy to a level similar to having had approximately 4 progeny for a highly heritable trait or 7 progeny for a moderately heritable trait (Table 2).

Table 2. Approximate number of progeny needed to reach accuracy levels (r) for three heritabilities (h^2).

Accuracy	Heritability Levels		
r	h^2 (0.1)	h^2 (0.3)	h^2 (0.5)
0.1	1	1	1
0.2	2	1	1
0.3	4	2	1
0.4	8	3	2
0.5	13	5	3
0.6	22	7	4
0.7	38	12	7
0.8	70	22	13
0.9	167	53	30
0.999	3800	1225	700

The choice of which animals to genotype, and with what assay, varies by specie. The dairy industry has made use of the BovineSNP50 assay and, in particular, the Holstein breed now has a densely genotyped population. Consequently, the use of lower-density (LD) assays (i.e 7K; Illumina, 2011) can be used effectively to genotype cows and heifers and then can be imputed up to 50K. The beef industry is not as lucky. Although the same LD chip exists for use in beef cattle, industry or breed-wide imputation is not as simple given the general lack of genotypes in many breeds and the fact that strategic genotyping is not employed, rather animals are chosen for genotyping in a rather ad-hoc manor by individual producers, although here are some efforts underway in US beef breeds to strategically target influential sires for genotyping in order to build a training set. There is no doubt that efficiencies could be gained if animals were targeted for genotyping in a more sophisticated manor, from a breed perspective instead of at the individual producer level. Several methods have been proposed to do this, including targeting animals (males only or both males and females) that contribute the greatest to the population in terms of average relationship and number of offspring (Spangler et al., 2008) or employing the use of machine learning algorithms such as Ant Colony Optimization (Spangler et al., 2009). Other industries can more efficiently make use of genomic information by genotyping nucleus animals with higher-density assays and developing reduced SNP sets for use in specific lines for a targeted suite of traits (i.e. swine).

Robust Across Breed Predictions—An Industry Relevant Crux

Kachman et al., (2012) used growth traits in beef cattle (weaning weight and yearling weight) to illustrate the efficacy of BovineSNP50 based MBV when the MBV were evaluated in the same breed as training and when they were 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 was 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).

Some organizations (i.e. smaller 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 EBV (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 in beef cattle 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 EBV 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).

Understanding and Adoption—The Human Crux

An ongoing challenge relative to technology adoption is a general understanding and familiarity with genomics. To increase the knowledge base and aid in the adoption of genomics, an integrated project referred to as the Weight Trait Project (WTP) was initiated in 2009 as a means of educating the US beef industry about the utility of genomic tools and to build a resource population for development and evaluation of methodology for incorporating molecular information into NCE. Twenty-four seedstock producers from the Northern Plains region of the US were nominated by their respective breed associations to participate in the WTP. These seedstock producers represent Angus, Red Angus, Charolais, Gelbvieh, Hereford, Limousin, and Simmental. As part of the WTP, they collected hair samples on the natural service (NS) sires and other animals used in their herds as a source of DNA for genotyping and a wide array of phenotypic data are collected on the progeny. The population has evolved into a valuable resource for the demonstrating the efficacy of genomically enhanced EBV on traits of economic importance. Through this ongoing integrated effort, key technology adopters are able to learn by doing, using their own animals as a demonstration of genomic predictors and methodology.

Although genomic information has the potential to generate value for multiple livestock industries, adoption must be economically driven. Less integrated industries (i.e. beef) suffer the most from a general lack of economic signal due to the disparity and lack of communication between differing segments. Other, more integrated industries such as dairy, poultry, and swine, have the opportunity to realize and capture more value immediately given the vertical nature of the structure of these industries. Using genetic tests to increase the accuracy of selection in the nucleus sector has the potential to generate large returns throughout all sectors. Improving the accuracy of EBV on elite young seedstock animals will accelerate the rate of genetic gain and impact the genetic merit of many descendants thereby amplifying the value of each unit of genetic improvement (Van Eenennaam et al, 2011). The economic value resulting from increases in productivity via improvements in net genetic merit may be captured through a variety of methods. Some of the improved economic value will be captured through improvements in sector specific economically-relevant traits (ERT). Genetic improvements that result in improved production efficiencies through more successful reproduction, growth and end product merit should provide improved revenue streams throughout and entire industry.However, industries with a higher degree of segmentation suffer to capture this increased value throughout the value chain.

In less integrated systems, genomic enabled selection strategies at the seedstock level may support higher testing costs due to the seedstock sector's ability tocapture the value of improved

genetic merit. In contrast, genetic tests for selection of animals in the commercial sector will need to be much less expensive due to the lower per unit returns in the commercial sector. It should also be noted that the traits of greatest importance to commercial producers will likely represent the suite of traits for which genomic technology is challenged the most (i.e. fertility, disease susceptibility).The beef feedlot sector could potentially use genomic information for marker-assisted management (MAM). However, slim profit margins in this sector requires a substantial return on investment before adoption of new technologies takes place

Conclusions

Genomics and the corresponding Marker-Assisted or Genomic-Enhanced EBV, have become a reality. Within-breed genomic predictions based on medium density (i.e. 50K) genotypes have proven to add accuracy, particularly to young animals, for several traits. Cost undoubtedly impedes deployment of this technology, however recent advances show promise in dramatically decreasing the cost of genotyping (Thallman, 2012). Methodology related to the use of this technology in crossbred or composite animals is critically needed to fully benefit the commercial livestock sector. The crux of adoption will be getting commercial producers to see the value in, and thus pay for, increased EBV accuracy. There is a still a need to collect and routinely record phenotypic information by seedstock producers and commercial producers need to realize that EBV, and economic index values, are the currency of the realm for selection. Genomic technology only makes these tools stronger, it does not replace them.

References

DeNise, S. K. 1999. Using parentage analysis in commercial beef operations. Pages 183-190 in Proc. Beef Improvement Federation 31st Annual Research Symposium and Annual Meeting, Roanoke, VA.

Dodds, K.G., M. L. Tate and J. A. Sise. 2005. Genetic evaluation using parentage information from genetic markers. J. Anim. Sci. 83: 2271–2279.

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

Geldermann, H., U. Pieper, and W. E. Weber. 1986. Effect of misidentification on the estimation of breeding value and heritability in cattle. J. Anim. Sci. 63:1759-1768.

Goddard, M. E. and D. E. Goddard. 1997. Using DNA fingerprinting of extreme offspring to progeny test sires. Proceedings of the Association for the Advancement of Animal Breeding and Genetics 12:438-441.

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

Heaton, M., M. Clawson, C. C. Mckown, K. Leymaster, T. Smith, G. Harhay, s. White, L. Hoesing, M. Mousel, G. Lewis, T. Kalbfleisch, J. Keen, and W. Laegreid. 2012. Reduced lentivirus susceptibility in sheep with TMEM154 mutations. PLoS Genetics. 8(1) e1002467.

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

Illumina. 2011. BovineLD Genotyping Beadchip. Accessed Nov. 18, 2011. http://www.illumina.com/Documents/products/datasheets/datasheet_bovineLD.pdf

Israel, C. and J. I. Weller. 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy cattle populations. J. Dairy Sci. 83:181-187.

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

A relationship matrix including full pedigree and genomic information. J. Dairy Sci.92:4656-4663.
MacNeil, M. D.,J. D. Nkrumah, B.W. Woodward, and S.L. Northcutt. 2010. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. J. Anim. Sci. 88: 517.

Meyers, S.N., T.G. McDaneld, S.L. Swist, B.M. Marron, D. J. Steffen, D. O'Toole, J. R. O'Connell, J. E. Beever, T. S. Sonstegard, and T. P. L. Smith. 2012. A deletion mutation in bovine SLC4A2 is associated with osteopetrosis in Red Angus Cattle. BMC Genomics11:337.

Quaas, R.L., and Z. Zhang. 2000. Incorporating external infomration in multi-breed genetic evaluation. J. Anim. Sci. 79: Suppl.(1).

Saatchi, M., D.J.Garrick, and J.Ward. 2011. Accuracy of genomic breeding values in Hereford beef cattle using k-means clustering method for cross-validation. J. Anim. Sci. 89 E-Suppl. 2: 59.

Saatchi, M., D.J. Garrick, and J. F. Taylor. 2012. Genetic correlations between carcass traits and genomic breeding values in Limousin cattle. J. Anim. Sci. 90 E-Suppl. 2: 66.

Senneke, S. L., M. D. MacNeil, L. D., Van Vleck.2004. Effects of sire misidentification on estimates of genetic parameters for birth and weaning weights in Hereford cattle. J. Anim. Sci. 82:2307-2312.

Spangler, M. L., R. L. Sapp, M. D. MacNeil, J. K. Bertrand, and R. Rekaya. 2008. Different methods of selecting animals for genotyping to maximize the amount of genetic information known in the population. J. Anim. Sci. 86:2471-2479.

Spangler, M. L., K. R. Robbins, M. D. MacNeil, J. K. Bertrand, and R. Rekaya. 2009. Ant colony optimization as an alternative method for genotype sampling. Animal Genetics DOI: 10.1111/j.1365-2052.2008.01835.x.

Thallman,R. M., K.J. Hanford, R.L.Quaas, S.D. Kachman, R.J.Tempelman, R.L.Fernando, and E.J.Pollak. 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., R.L. Fernando, and J.C. Dekkers. 2010. Genomic selection in admixed and crossbred populations.J. Anim. Sci. 88: 32-46.

Van Eenennaam, A.L., van der Werf, J.H.J., and M. E. Goddard. 2011. The value of using DNA markers for beef bull selection in the seedstock sector. J AnimSci 89: 307-320.

Van Eenennaam, A. L., R. L. Weaber, D. J. Drake, M. C. T. Penedo, R. L. Quaas, D. J. Garrick, and E. J. Pollak. 2007. DNA-based paternity analysis and genetic evaluation in a large, commercial cattle ranch setting. J. Anim. Sci. 85:3159-3169.

Van Vleck, L. D. 1970. Misidentification and sire evaluation. J. Dairy Sci. 53:1697-1702. Weaber, R. L. 2005.A simulation study of replacement sire selection and genetic evaluation strategies for large commercial ranches.PhD Diss. Cornell Univ., Ithaca, NY.

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. 2012. Accuracy of genomic breeding values in multi-breed beef cattle populations derived from deregressed breeding values and phenotypes. Journal of Animal Science (Submitted).