Optimizing Traditional and Marker Assisted Evaluation in Beef Cattle

D. H. "Denny" Crews, Jr.,[1,2,3] Stephen S. Moore,[2] and R. Mark Enns[3]

1   *Agriculture and Agri-Food Canada Research Centre, Lethbridge, Alberta;*
2   *Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta;*
3   *Department of Animal Sciences, Colorado State University, Fort Collins, Colorado*

Introduction

Genetic evaluation systems wherein phenotypic data are used to predict breeding values are well accepted by the beef industry. Beef breed associations conduct regular national cattle evaluations (NCE) and genetic trend analyses have shown that, over the past two decades, the mean genetic merit of populations (breeds) reflect the widespread use of breeding values, or expected progeny differences (EPD) as selection tools. More recently, mapping of the bovine genome and the development of related genomic tools has prompted an interest in augmentation of traditional genetic evaluation systems with marker information. Marker assisted evaluation systems would optimally combine genomic with phenotypic and pedigree data to predict EPD with higher accuracy than would be expected from evaluations based solely on either data source alone. Such marker assisted EPD (MEPD) would be particularly useful for increasing the accuracy of evaluating young animals which have yet to make their own phenotypic record or produce progeny with records for economically relevant traits. Considering polygenic and marker breeding values as separate but correlated indicators of genetic merit allows for the application of selection index methods to optimally combine phenotypic and genomic information in MEPD. Decisions must be made with regard to the cost effectiveness of marker panel development and genotyping to increase producer uptake of the technology. Criteria for identifying likely candidate traits for genome assisted evaluation generally depend on accuracy of polygenic breeding value and the genetic variance attributable to the marker set. The potential benefits to industry from marker assisted evaluation and selection remain large, and will be realized as the efficacy of the marker breeding value component of this approach increases.

Traditional NCE

For more than 25 years, breeding values or EPD predicted using mixed model methods have been used by beef breed associations and their members as selection tools. The most commonly evaluated traits such as weight and growth rate have shown significantly positive genetic trend over the same period. Along with accuracy values, EPD are the most common tool used by producers for effective, additive, and permanent genetic improvement of cattle. Current NCE systems are robust to the common problems associated with bias in field data, and with a sufficiency of information and selection intensity, considerable genetic improvement can be made for heritable traits. In most beef breed associations that regularly conduct NCE, EPD and accuracy values are predicted for millions of animals in their pedigree and performance databases. Validation studies have conclusively shown the effectiveness of EPD-based selection, and the correspondence between expected and realized progeny differences (Thrift and Thrift, 2006). Recent advances in NCE programs include the addition of systems for input traits such as feed intake and efficiency, threshold traits related to pregnancy, stayability and (or) longevity, and a wide array of multiple trait indexes.

The recent shift in focus on new trait development is characterized by less emphasis on component traits (e.g., ribeye area is a component of retail yield; birth weight is a component of growth) and indicators (e.g., scrotal circumference is an indicator of fertility; birth weight is an indicator of calving difficulty) and more emphasis on economically relevant traits (ERT) which directly impact either cost or revenue (Golden et al. 2000). As traditional NCE continues to mature, ERT evaluation systems include traits that are more costly, time consuming, and difficult to measure. Additionally, some traits have been of interest for some time but rarely included in NCE because of low data density at the population level. For ERT, whether new or not, the effectiveness of NCE depends on information density (animals with data) which supports evaluation accuracy. In the interest of maximizing genetic progress, young sires with high accuracy evaluations are very desirable. However, these young animals are generally not evaluated accurately except with respect to the most basic traits which are easily measured and have at least moderate heritability, leaving the large proportion of young bulls with EPD predicted on the basis of older relatives with records and (or) progeny with records. With these factors considered, traits such as tenderness, feed intake and efficiency, stayability and longevity, health status, and reproductive ability are examples that would benefit from NCE that are augmented to increase accuracy.

The historical approach to increasing accuracy on complex traits is to measure indicators and include these in NCE as correlates of the ERT. In the case of carcass traits, for example, ultrasound measures on live animals are commonly included in carcass NCE as separate but correlated traits (e.g., Crews et al., 2003, 2004). This is an especially useful approach to evaluating carcass merit on young sires because of the ability to readily collect ultrasound measurements near yearling age. For a

trait such as stayability (e.g., Snelling et al., 1995), which evaluates genetic merit for production to a threshold related to profitable numbers of parities, heritability is low (i.e., requiring more data to attain high accuracy EPD compared to traits with higher heritability), and animals are 5 or more years of age before they make their own phenotypic record. In any case, the limitation of evaluations based largely on indicator traits is that accuracy is restricted by the genetic correlation between the ERT and the indicator. Indicators with higher genetic correlations with ERT are therefore more desirable, along with being easier and cheaper to measure on young animals.

In a growing number of novel ERT cases, few if any indicator traits have been identified. In the case of efficiency expressed as residual feed intake, indicator traits are purposefully eliminated in order to reduce genetic antagonisms with growth rate, mature size, and body composition. In these cases, the use of molecular tools has higher potential to usefully increase accuracy. For growth rate and weight, carcass merit to some extent, and other traits with at least moderate heritability and that are easily measured, traditional NCE approaches should remain the standard for genetic evaluation on the population level.

Marker Technologies and Tools

With the development of the bovine genome map, the density of marker information available for use in genetic evaluation is growing. Regions of the genome which show associations with phenotypes, often called quantitative trait loci (QTL) or QTL regions, have been identified in numerous studies. Also, point mutations referred to as single nucleotide polymorphisms (SNP) have been reported to associate with variation in quantitative traits. High throughput genotyping platforms are available to genotype animals for very large

numbers of SNP markers, producing a high volume of data on individual animals. The latest genotyping technology platforms can be used to generate 50,000 SNP genotypes on individual animals, and the capacity of these SNP marker chips is likely to increase in the near future. Genotyping costs remain a limitation to the numbers of animals that have, or will have, marker information. Also, methods to estimate molecular breeding value from the large numbers of genotypes now possible have yet to mature.

Over the last decade, quantitative trait loci (QTL) for a number of traits have been described in cattle. For example, QTL for fat depth have been reported on a number of chromosomes (Stone et al., 1999; Casas et al., 1999; Casas et al., 2000, 2003; Moore et al., 2003; Li et al., 2004). Some of these QTL, however, have been localized to larger chromosomal segments and are only weakly supported. Most QTL studies have focused on traits that are more easily measured in the live animal or on the carcass. So, despite its importance, very few attempts at identifying QTL for beef cattle feed intake and efficiency have been made in the past (Nkrumah et al, 2005; Moore et al., 2006). Recently, Barendse et al. (2007) reported a whole genome association study for feed efficiency traits in beef cattle.

Functional genomics approaches combine information from various marker tools to establish associations between sets of markers and ERT. This process typically involves SNP discovery using animal populations which are divergent for traits of interest. Polymorphisms within genes of known function can also be used as candidates in marker discovery, and to understand basic gene function. Associations of markers with ERT in discovery populations must be validated across a diverse array of non-discovery populations in order to establish how robust and effective sets of markers are likely to be across populations. Once the magnitude and direction of marker effects has been estimated from discovery populations, marker scores can be assigned to any animals with marker genotypes. The typical approach to validation is to regress phenotypes simultaneously on marker value and polygenic effects using a mixed inheritance model. This allows for estimation of marker effects in the presence of polygenes in the validation population. The partial regression of marker score on phenotype is evaluated and a significant association between the marker breeding value and the phenotype is a confirmation or validation of the marker set. However, the effects of markers and marker sets or "panels" can be breed- or population-specific (e.g., Angus versus Charolais, or Bos taurus versus Bos indicus), which affects the validation process. Factors such as phase (the linkage patterns among a group of markers) impact the extent to which markers can be validated and probably contribute to the variable effects of markers across groups. Requirements of the validation process emphasize the need for established discovery and separate validation populations of animals with known pedigree structure and an extensive range of recorded phenotypes. Both research herds and experimental herds in industry serve this need, but should be maintained at or above present levels.

Marker and Phenotypic Information

The advent of marker technology has the potential to increase the effectiveness of NCE. This depends, however, on the optimal combination of marker information with traditional phenotypic and pedigree information in evaluations. This topic has been of interest for more than a decade in the scientific literature although no clearly optimal approach has been proposed.

Early methodology was presented which simultaneously estimated marker and polygenic

effects in a mixed model. These early methods for marker assisted evaluation depended on highly compute-intensive algorithms and are, in general, not applicable to the large databases common to industry at the population or breed level. Selection index methodology has been more recently applied to the problem of optimally combining genomic and polygenic breeding values. Polygenic and marker derived breeding values can be combined in a linear index with weighting factors that depend solely on accuracy of the polygenic breeding value and the proportion of genetic variance attributable to the marker set. Simulation has confirmed that polygenic EPD accuracy and the gain in accuracy of evaluation due to inclusion of marker information are inversely related. This fact clearly illustrates that marker information contributes less and less to added evaluation accuracy as polygenic accuracy increases (e.g., Mrode, 2005). Therefore, for traits with high heritability, or that accumulate accuracy quickly for young animals are less viable candidates for genomic selection. On that same note, traits related to some aspects of carcass merit, feed intake, and longevity are examples of likely candidates for gene assisted evaluation because of the cost and time required for traditional polygenic evaluations to be accurate.

Once a validated marker set is available, animal genotypes for that set are used to produce marker breeding values. In effect, MEPD are a combination of marker and polygenic breeding value. The traditional EPD, with the addition of marker information, is partitioned into a component due to identifiable genetic markers, and a polygenic component that is due to genetic effects that are not linked to markers. In effect, a portion of additive genetic merit which was previously attributed to the small effects of many genes, is attributed in the MEPD to a set of markers that can be identified.

Predictive Power and Cost of Markers

Functional genomics studies have begun to report the association of SNP markers and marker panels with ERT in beef, but optimal marker assisted evaluation systems will require robust estimates of combined marker effects, and these have yet to be commonly reported. It is likely that marker panels will need to account for at least 10 to 15% of the genetic variance in the trait of interest before the gain in evaluation accuracy due to adding marker information will be worthwhile, or even cost effective considering genotyping cost.

As the availability of marker tools increases, producers will be faced with the decision of where genetic improvement investment should be directed. For young animals, this decision will essentially compare the options of genotyping versus phenotypic testing. The value of genotyping will reflect the predictive power of the marker panel. In the case of older animals, which may be considered population or breed founders, genotyping will not be a likely option. For these animals, then, no marker information will be available. In fact, the largest proportion of animals in any breed may never have marker breeding values, unless these are derived indirectly rather than from actual marker panel genotypes. The index approach and related methodologies would predict MEPD based largely on phenotypic data for animals without marker breeding values. For the younger animals with marker information, MEPD predictions would be from a combination of data sources, appropriately balanced to account for EPD accuracy and predictive power of the marker set. With respect to ERT where essentially no phenotypic data exists, MEPD would largely depend on marker breeding value, with small contributions from phenotypic indicators and perhaps records on relatives.

As the predictive power of marker sets increases, the proportion of genetic variance explained by that set also increases, and the estimate of genetic merit from markers becomes increasingly accurate. With whole genome selection methods, it may be possible to predict breeding value from markers alone with 80% or more accuracy. The development of validated and robust marker sets to this level of accuracy has yet to mature. Current genotyping costs for individual animals using high-density chips is > $400.

Information Packaging

During the period when genetic evaluation and improvement has become predominantly based on EPD, beef producers have gained experience and confidence with using NCE results in their program. One consideration in developing MEPD systems is to build upon that familiarity.

It can be shown that independent selection on both the traditional EPD and separate marker information can be antagonistic. Therefore, publishing EPD and marker scores separately will lead to sub-optimal and in some cases incorrect selection decisions. This becomes particularly relevant when the marker panel and resulting marker breeding values predict a small proportion of the genetic variance in the breeding objective.

Index approaches and similar methods combine polygenic and marker information into a package that would be familiar, and have accuracy that reflects the genetic information derived from phenotypic and marker data.

Summary

The potential benefits to industry from marker assisted evaluation and selection remains very large. Traditional polygenic evaluation systems such as national cattle evaluation have proven effective. For traits that generally have lower accuracy polygenic evaluation, the use of genomic tools could improve evaluation accuracy. In general, traits for which highly informative marker sets can be validated will be more viable candidates for marker assisted evaluation and selection compared to those for which traditional evaluation is effective. The combination of phenotypic and marker data into an evaluation depends solely on accuracy of traditional evaluation and the proportion of genetic variance attributable to the marker set. Uptake of marker assisted evaluations in industry is likely to depend on the effectiveness of marker panels and genotyping costs, as well as effective reporting of evaluation results.

References

Barendse, W., A. Reverter, R. J. Bunch, B. E. Harrison, D. J. Johnston, V. H. Oddy, W. Barris, and M. B. Thomas. 2007. A validated whole genome association study of efficient food conversion in cattle. Genetics 176:1893-1905.

Casas, E., S. D. Shackelford, J. W. Keele, R. T. Stone, S. M. Kappes, and M. Koohmaraie. 2000. Quantitative trait loci affecting growth and carcass composition of cattle segregating for alternate forms of myostatin. J. Anim. Sci. 78:560-569.

Casas, E., S. D. Shackelford, J. W. Keele, M. Koohmaraie, T. P. L. Smith, and R. T. Stone. 2003. Detection of quantitative trait loci for growth and carcass composition in cattle. J. Anim. Sci. 81:2976-2983.

Crews, D. H., Jr., E. J. Pollak, R. L. Weaber, R. L. Quaas, and R. J. Lipsey. 2003. Genetic parameters for carcass traits and their live animal indicators in Simmental cattle. J. Anim. Sci. 81:1427-1433.

Crews, D. H., Jr., E. J. Pollak, and R. L. Quaas. 2004. Evaluation of Simmental carcass EPD estimated using live and carcass data. J. Anim. Sci. 82:661-667.

Golden, B. L., D. J. Garrick, S. Newman, and R. M. Enns. 2000. A framework for the next generation of EPD. Proc. 32nd Beef Improvement Federation Annual Research Symposium and Meeting, Wichita, Kansas, pp. 2-13.

Li, C., J. A. Basarab, W. M. Snelling, B. Benkel, J. Kneeland, B. Murdoch, C. Hansen, and S. S. Moore. 2004. Identification and fine mapping of quantitative trait loci for backfat on bovine chromosomes 2, 5, 6, 19, 21, and 23 in a commercial line of Bos taurus. J. Anim. Sci. 82:967-972.

Moore, S. S., C. Li, J. A. Basarab, W. M. Snelling, J. Kneeland, B. Murdoch, C. Hansen, and B. Benkel. Fine mapping of quantitative trait loci and assessment of positional candidate genes for backfat in a commercial line of Bos taurus. J. Anim. Sci. 81:1919-1925.

Moore, S. S., D. H. Crews, Jr., and J. D. Nkrumah. 2006. Multiple and candidate gene approaches to genetic evaluation of feed efficiency in beef cattle. Proc. 8th World Congress on Genetics Applied to Livestock Production 14-01.

Mrode, R. A. 2005. Linear Models for the Prediction of Animal Breeding Values, 2nd Edition. CABI Publishing.

Nkrumah, J. D., C. Li, Z. Wang, R. Bartusiak, B. Murdoch, J. Basarab, D. H. Crews, Jr., and S. S. Moore. 2005. Full genome scan for quantitative trait loci (QTL) for net feed efficiency in beef cattle. J. Anim. Sci. 83 (Suppl. 1):13 (Abstr.).

Snelling, W. M., B. L. Golden, and R. M. Bourdon. 1995. Within-herd genetic analysis of stayability of beef females. J. Anim. Sci. 73:993-1001.

Stone, R. T., J. W. Keele, S. D. Shackelford, S. M. Kappes, and M. Koohmaraie. 1999. A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. J. Anim. Sci. 77:1379-1384.

Thrift, F. A., and T. A. Thrift. 2006. Review: Expected versus realized progeny differences for various beef cattle traits. Prof. Anim. Sci. 22:413-423.