Emerging Technologies in Genetic Improvement—Convergence of Quantitative and Molecular Tools

Mike Tess
Montana State University

Introduction

Research into the molecular basis of inheritance is progressing at a rapid pace. Technologies that permit the identification of molecular genetic differences (i.e., differences in deoxyribonucleic acid (DNA) sequence among animals) are also evolving very rapidly. Several DNA-based tools are being marketed in the beef industry; and, some as selection tools. These tools are known by a variety of names in the academic community and within the beef industry (e.g., genomic tests, DNA markers, molecular markers). For simplicity I will refer to them as "DNA tests."

DNA tests present opportunities and challenges to the U.S. beef industry. Accurate DNA-based selection tools will give beef cattle breeders opportunity to identify animals with superior breeding value (BV) as soon as a tissue sample can be collected, potentially leading to significant savings in time and money associated with performance testing and genetic evaluation. However, the current state of national cattle evaluation (NCE) in the beef industry provides no clear direction to breeders regarding how best to use these new DNA tests in their selection programs.

In this paper I will attempt to: 1) briefly describe the different types of DNA tests currently marketed, 2) discuss the potential as well as the limitations of current DNA tests for selection, 3) and present a model for NCE that provides better information to beef cattle breeders.

I would like to express my sincere appreciation to the current and former members of the BIF Commission on DNA Markers for their work during the past year: Bill Bowman, Ronnie Green (former member), Steve Kappes, Ronnie Silcox, and Darrell Wilkes.[1] The ideas and concepts presented here reflect their work. I would also like to thank John Pollak[2] and the National Beef Cattle Evaluation Consortium for their cooperation and support.

Current DNA Tests

There are a variety of DNA tests available to the beef industry today. The number of DNA tests marketed will likely increase rapidly over time. Following is a list of the broad types available based on their applications. All are based on identifying differences (or in some cases similarities) in DNA base-pair sequence among animals. The number of base-pairs involved, and the lab techniques employed vary.

Parentage Identification/Validation tests are used to identify or validate the parents of calves. They involve testing the calves and at least one parent.

Identification/Traceability tests are used to track animals and their tissues through the food production chain as animals and their products change ownership and move from location to location. Variation in DNA is used to identify individual animals. Each animal being tracked must be tested.

Management tests are used to predict the future phenotypes of the animals tested in specific

1 American Angus Association, USDA-ARS, USDA-ARS, University of Georgia, and ABS Global, respectively.
2 Cornell University.

production-marketing systems. They are based on identifying differences in total genetic merit among animals (i.e., additive and non-additive genetic merit).

Selection tests are used to estimate breeding value (i.e., distinguish among animals on the basis of their progeny performance). Traits may be qualitative or quantitative in nature. Qualitative traits are controlled by one or a few loci, and phenotypes generally fall into distinct classes (e.g., presence of horns, coat color, and certain genetic defects). Quantitative traits are controlled by many loci. Quantitative phenotypes may be measured on a continuous scale (e.g., weights) or in classes (e.g., pregnant or open).

The focus of what follows is on DNA tests for quantitative traits used for selection. Some DNA tests may be marketed for both management and selection purposes. These tests may measure non-additive as well as additive genetic variation. When used as selection tools their value depends on their ability to measure additive genetic variation.

Accuracy

In simple terms, DNA tests measure differences among animals in DNA sequence (sometimes called genotypes). DNA tests for selection are developed by genotyping large numbers of animals within a population and computing statistical associations among genotypes and phenotypes of the target trait. The phenotypes may be measured on the animals that were genotyped or on their progeny. Alternatively, associations may be computed between sire genotypes and sire EPD. The genotypes may be DNA sequences from within a single gene, or a "panel" of several genes. The process of developing a DNA test is called "discovery." The discovery process attempts to find genotypes that are highly predictive of desirable progeny phenotypes.

Hence, DNA tests measure genotypes for the purpose of predicting phenotypes. To be useful in genetic improvement, a DNA test must accurately predict the phenotypes of progeny. In other words, the DNA test must accurately predict breeding value.

Accuracy is a genetic evaluation term that is familiar to most breeders. BIF reports accuracy as a number between 0.0 and 1.0. Accuracy measures the degree or strength of the association between a predictor of breeding value (like an EPD) and the true breeding value. The stronger the association between the predictor and the true breeding value, the higher the accuracy. Another way of describing accuracy is that accuracy measures the amount of variation in true breeding value accounted for by the predictor.

Quantitative traits are controlled by hundreds if not thousands of genes. Quantitative traits are also affected by non-genetic factors that geneticists simply refer to as environmental effects. If we could genotype all the genes that affect a certain trait, and if we could predict how every allele of each gene affects the trait, in theory we could explain all the additive genetic variation for the trait and predict the breeding value of an animal for that trait with an accuracy of 1.0.

Current DNA tests are based on from a few to over 100 genes. The number of genes measured is likely to increase rapidly over the next few years. Although we know that all genes do not have equal effects, for any given trait the accuracy of the DNA test (i.e., its ability to accurately predict breeding value) is expected to increase as the number genes accounted for increases.

Validation and Assessment

DNA tests are developed based on associations between variations in DNA sequence with variations in phenotypes. The animal populations used to develop a test may or may not be representative of industry populations. The concept of validation generally involves the confirmation of rejection of these associations in populations different from those in which the tests were developed. Validation studies are considered to be more reliable if they are conducted by scientists who have no vested interest in the tests (e.g., development, commercialization, or marketing).

The concept of assessment involves determining how specific DNA tests are associated with each other and with non-target phenotypes. In other words, assessment seeks to determine how competing DNA tests overlap and how non-target traits will be influenced by selection based on these tests.

The Breeder's Dilemma

The status of current DNA tests and NCE presents a dilemma for beef cattle breeders. The components of this dilemma are related to the concepts of accuracy, validation, assessment, and the need for a common currency for genetic evaluation.

First, the accuracy of current DNA tests for predicting breeding value is essentially unknown. In other words, the fraction of additive genetic variation explained by any specific DNA test is unknown. Since most current DNA tests are based on genotyping only a fraction of the genes that affect the target trait, it's safe to assume the accuracies of these tests are low. Similar to an EPD with a low accuracy, the breeding value estimated by a DNA test with a low accuracy is likely to be quite different from the true breeding value.

Second, to date components of commercially available DNA tests have been validated via publication in peer-reviewed scientific journals or by the National Beef Cattle Evaluation Consortium (NBCEC) serving as an independent third party. However, not all DNA tests have been scientifically validated. Validation serves to reduce risk for breeders using the DNA tests as selection tools.

Third, currently there are three genomic companies marketing DNA tests in the United States. For some traits (e.g., marbling, tenderness), each of the companies markets competing tests. The results of these competing tests are reported in different units. It's safe to assume that competing tests are not identical (i.e., they don't measure the same genes), and are not equally accurate. Yet, it's also safe to assume that competing tests are positively correlated; in other words, the information on breeding value provided by the tests overlaps. Breeders need some way to compare and combine these competing tests as they seek to compare animals with test results from different companies.

Fourth, a critical part of designing a breeding plan is consideration for how selection based on a set of target traits might produce genetic changes in important non-target traits. Geneticists call these changes "correlated responses" which are due to genetic correlations among traits. Currently, we no very little about expected correlated responses to selection using DNA tests. It is unknown if selection on a subset of genes the affect a trait have the same correlated responses as traditional selection on the trait.

Fifth, DNA tests seek to estimate breeding value just as EPD estimate breeding value. Although based on different sources of information (DNA sequences versus phenotypes and pedigrees), the information provided by DNA tests and EPD for the same trait overlap. There is no valid scientific way for breeders to compare or combine DNA tests and EPD to estimate breeding value. It's safe to assume that DNA tests provide valid information on breeding value—information that can be collected at birth; however, we have no way of comparing these estimates to an EPD, or using this information to improve an EPD.

In North America, EPD have become the currency of genetic evaluation of beef cattle. EPD are the genetic currency of the beef industry. The beef cattle industry needs a NCE system of that utilizes all sources of information on economically important traits to estimate a single estimate of breeding value with an accompanying accuracy value. The beef industry needs EPD that are computed using all the information available—pedigrees, phenotypes, and DNA tests.

A Proposed Model

Performance testing and genetic evaluation are being conducted on an increasing number of economically relevant traits. The types of information available (i.e., available from a practical and economical view) varies among traits. Types of information include pedigree relationships, performance measurements (i.e., phenotypes), and DNA test results. Phenotypes may include direct and indirect measurements on the same traits. For example, carcass backfat may be measured indirectly via ultrasound or directly after slaughter. Table 1 illustrates the various combinations possible. Because most animals marketed in the U.S. as seedstock have known parentage the table assumes that pedigree relationships are known.

Some economically relevant traits are difficult to measure and there are no DNA tests for these traits available. These traits will likely be the focus of future research. In a second category are traits for which phenotypes are regularly measured in the field, systematically data-based, and for which EPD are computed. The emergence of DNA tests now permits the estimation of breeding values on animals for which little or no phenotypic information is available (a third category). A current example would be tenderness. Tenderness phenotypes are difficult and expensive to measure, but DNA tests are available. In a fourth category are traits where both phenotypes and DNA tests are available. A current example would be carcass marbling.

Over the past year the BIF Commission on DNA Markers and the National Beef Cattle Evaluation Consortium (NBCEC) worked to address these issues presented above and that accommodates traits with different types and amounts of information collected. Our guiding philosophy is summarized by the following statement drafted by the BIF Commission:

> *The BIF Commission believes that information from DNA tests only has value in selection when incorporated with all other available forms of performance information for economically important traits in NCE, and when communicated in the form of an EPD with a corresponding BIF accuracy. For some economically important traits information other than DNA tests may not be available. Selection tools based on these tests should still be expressed as EPD within the normal parameters of NCE.*

Figures 1 and 2 schematically present a proposed model for NCE that incorporates pedigree relationships, performance phenotypes, and DNA test information in the computation of EPD and accuracies. The model will accommodate traits with different amounts and types of information (i.e., pedigree relationships, indirect and direct measures of phenotype, and DNA tests from multiple companies). As envisioned this model would accommodate within-breed NCE as well as multi-breed NCE. The proposed model assumes that breed associations will continue to bear major responsibility for the delivery of EPD to the beef industry.

Statistical procedures for incorporating DNA test information into NCE and the computation of EPD and associated accuracies will be described in other media. Briefly, the method utilizes DNA test results in a manner analogous to using correlated traits in more traditional NCE. The method permits incorporation of several competing DNA tests (e.g., tests for the same trait) as well as pedigree and performance information. The method is applicable to any trait for which some information on breeding value is available (i.e., phenotypes and (or) DNA tests; 3 of the 4 cells in Table 1).

Evaluation of a DNA test as a Selection Tool
As represented in Figure 1, I have assumed that the NBCEC will coordinate validation and assessment efforts. At present, the future of NBCEC is uncertain. Nevertheless, the participation of an independent third party in the model is critical.

Under the proposed model, evaluation of a DNA test as a selection tool includes the concepts of validation and assessment; but also provides information on the accuracy of selection based on the DNA test. Evaluation of a DNA test as a selection tool includes: 1) estimation of the genetic relationship between the DNA test and the target trait (phenotype), 2) estimation of the genetic relationships among competing DNA tests for the target trait, 3) estimation of genetic relationships between the DNA test and non-target traits, and 4) computation of EPD and their associated accuracies. Completion of these tasks will require the genotyping (i.e., running the DNA tests) of reference populations, and statistical analyses of datasets that include DNA test scores, pedigree relationships, and phenotypes for the traits of interest. Results of these evaluations should be reported to the public in an efficient and timely manner.

Inclusion of DNA Test Information in NCE Programs
Results of the evaluation phase (outlined above) will provide all the needed statistical parameters needed for NCE. The decision to include a DNA test in a NCE system should be made by the breed association or the organization responsible for publishing the EPD. Consideration should be given to the heritability of the trait, the availability of producer-collected phenotypes, and the increase in accuracy provided by the addition of the DNA test information.

Reference Populations
As used here (Figure 1), reference populations are: 1) pedigreed herds representative of and genetically linked to commercial populations in the beef industry, 2) managed in production/marketing systems representative of the beef industry, and 3) measured for economically relevant traits. Herds useful for the purposes described here include: 1) cataloged data from research studies, and 2) existing herds. Ownership may be public or private; however, as envisioned here the most useful on-going reference populations are likely to be federally owned and managed.

Access to quality reference populations is absolutely critical to the success of a NCE system that incorporates DNA test information. The participation of USDA-ARS and Agriculture Canada will be vital to the successful implementation of DNA-based selection tools in the beef industry.

Databases
After a DNA test has been approved for inclusion in NCE, DNA test scores will need to be stored and accessed in an efficient manner. Figure 2 presents

a schematic of how NCE would incorporate DNA test information on an ongoing basis. The proposed model will require the storage and use of potentially large databases of DNA information. Important considerations include: 1) the marketed DNA tests are expected to change frequently over time, 2) multiple companies are likely to market DNA tests for the same target trait, and 3) access to the raw data may need to be restricted. Hence, it will be important that the database(s) accommodate these aspects.

The quality of any EPD is dependent on the quality of the data used to compute the EPD. Much like selective reporting of phenotypic measurements may bias EPD computed from pedigree and phenotypes, selective reporting of DNA tests may bias EPD computed from DNA tests. It will be important for breed associations to implement policies that encourage complete reporting of all DNA tests.

A Dynamic Future

The pace of change in the beef industry continues to be rapid. The evolution of genomic tools for the beef industry may be even more rapid. Participants in this year's BIF convention will get a glimpse of several new emerging technologies. As NCE systems evolve to accommodate new technologies, it will be important to do our best to implement systems that embrace a dynamic future.

One of the most significant changes on the horizon is that future DNA tests will undoubtedly be based on many more genes. Developments in molecular technologies now permit the genotyping of tens of thousands of gene segments at the same time. This number will no doubt continue to increase. Although these technologies are not without significant challenges, they should lead to new DNA tests based on hundreds, if not thousands of genes, leading to important improvements in the accuracy of breeding value estimates based on DNA tests.

Just as agriculture is a global industry, genomics is a global industry. The future will see greater collaboration among federal scientists across nations, as they seek to discover and evaluate DNA tests in different production/marketing environments.

Conclusion

The growing genomics industry represents a great future for selection tools and the genetic tailoring of cattle for specific production/marketing environments. Like most new technologies, genomics brings new challenges. National cattle evaluation must adapt to include DNA test information so that beef breeders can make efficient use of the information provided by these tests. As the beef industry moves forward national cattle evaluation will depend on the partnership of any increasing number of industry segments, including seedstock producers, breed associations, federal and university research institutions, seedstock marketers, and genomic companies.

Table 1. Traits categorized according to information available.

DNA Tests	Industry-collected Phenotypes	
	No	Yes
No	---	EPD
Yes	EPD	EPD

Figure 1

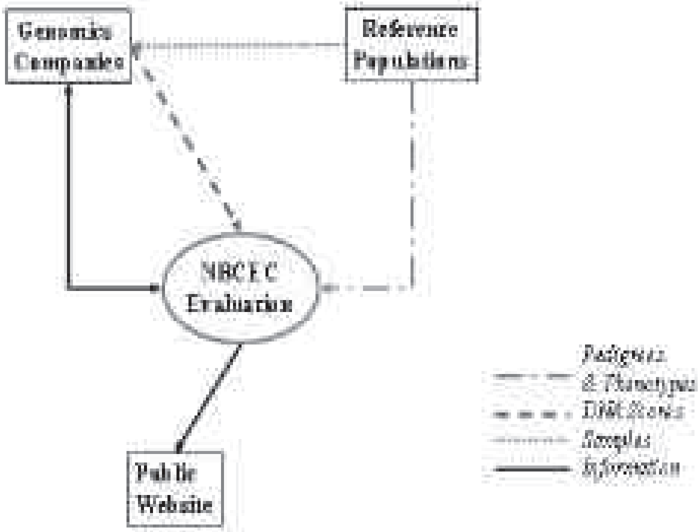

Figure 2