Proceedings, Applied Reproductive Strategies in Beef Cattle - Northwest
September 30 – October 1, 2011; Boise, ID

IMPROVING EPD ACCURACY BY COMBINING EPD INFORMATION WITH DNA TEST RESULTS[1]

Alison L. Van Eenennaam

Department of Animal Science, University of California, Davis, CA

Introduction

Molecular or DNA sources of information represent a rapidly developing technology with regards to beef cattle selection. Given the rapid commercialization of DNA marker panels, producers have been able to see changes first hand in reporting styles, panel sizes, and traits for which panels are available. Ultimately, the power of this technology will only be fully exploited when it is seamlessly integrated into National Cattle Evaluations (NCE).

Utilization of Molecular Information to Date

DNA information has been effectively utilized to identify animals that are carriers of recessive alleles. This has been of particular interest for genetic defects, color, and horned/polled status. Prior to the advent of this technology the only way to test if a sire was a carrier of a particular genetic defect was to mate him to a given number of known carriers of the defect or an even greater number of his own daughters of unknown genotype. Even then definitive conclusions could only be drawn if he sired an afflicted calf. If all corresponding offspring were free of the defect, then it would be possible to assign a certain probability to the sire being a non-carrier with the probability being dependent on the number of calves born from a particular mating. Today carrier status can be determined with a simple DNA test. DNA-marker technology has also proven very beneficial in determining parentage. More recently, SNP panels have been developed to test for a portion of the genetic merit of an animal for a variety of traits ranging from fertility and longevity, to growth and carcass merit.

Methods of Reporting (past and present)

Many of the early recording systems to relay marker panel results were categorical in nature. For instance, systems existed that provided one star for each favorable allele regardless of the proportion of variation explained by the marker. Others provided a 1-10 scale where genotypes were categorized by the impact they had on the trait of interest. Neither of these systems allowed for the inclusion of DNA results in NCE. More recently marker panel results have been reported as Molecular Breeding Values (MBVs). Although MBV is the term that is being used by the scientific community, DNA testing companies have created unique names to identify their respective products in the market place (e.g. Molecular Value Prediction [MVP] and Genetic Prediction Difference [GPD]).

[1] Research summarized in this manuscript was supported by National Research Initiative Competitive Grant no. 2009-55205-05057 ("*Integrating DNA information into beef cattle production systems*") from the USDA National Institute of Food and Agriculture. Some of the material contained in this manuscript is excerpted from Spangler, M., and Van Eenennaam, A. L. 2010. DNA-based Biotechnologies. Pages 79-84 in Beef Sire Selection Manual. 2nd Edition. National Beef Cattle Evaluation Consortium http://www.nbcec.org/producers/sire.html (with permission).

Differences Between MBVs and EPDs

EPDs provide an estimate of the genetic potential of an animal as a parent based upon ancestral information; his/her own records, and the records of his/her progeny. With this in mind, an EPD accounts for all the genes that affect a particular trait, regardless of the magnitude of their affect. While an EPD accounts for all of the additive genetic variation, the specific sources of the variation (genes) are unknown. Conversely, DNA marker tests reveal the genotype of an animal for specific DNA markers for a particular trait but, to date, do not account for all of the genetic variation. This is simply due to the fact that the markers or genes with the largest effects are the easiest to identify and become the logical candidates for inclusion in marker panels. The potentially vast numbers of markers or genes with much smaller effects are more difficult to identify, and consequently have not been included in the development of marker panels.

There is also an inherent difference between an EPD and an estimated breeding value (EBV). An EBV is the genetic merit of an animal whereas an EPD is the genetic merit of an animal as a parent. Given that an animal can only pass on a sample half of its alleles to the next generation, the relationship between the two is as follows: EPD = ½(EBV).

Although some DNA companies report results in a form that looks similar to an EPD in that it is reported in units of the trait, the values are EBVs based on molecular information. To determine how much better one animal is versus another as a parent, EBVs must be divided by two.

It is critical to understand that a desirable genetic test result with current commercially available panels is not always associated with a desirable EPD. For instance, it would be possible for an animal to be homozygous for the favorable allele for a DNA marker for marbling but still have a marbling EPD that is below breed average. This could occur because, although the animal has the favorable form of both alleles of one of the genes affecting marbling, it may have unfavorable alleles for numerous other unknown genes that affect marbling as well.

Accuracy

When DNA from an animal is submitted for a DNA-test, there is an accuracy associated with that result. Accuracy has multiple meanings. One definition is “degree or extent of freedom from mistake or error”. Genotyping generally has a high level of technical accuracy or precision, meaning that there are rarely errors in the actual results of the DNA test. However, in animal breeding, accuracy refers to how well an estimate of the genetic merit (e.g. EPD, or DNA-test result) predicts the true genetic merit of an animal. One measure of this genetic prediction accuracy is the correlation (r) between a genetic merit estimate and the true genetic merit of that animal. Accuracy values can range from 0 (in which case the estimate has no relationship to an animal’s true genetic merit) to 1 (in the theoretical situation where the estimated breeding value is equal to the true breeding value). In practice accuracy values never reach the theoretical limit of 1, although very high accuracy of extensively used AI sires can reach 0.99.

In some countries, the accuracy of a genetic prediction (EPD in the U.S.) is reported as the correlation between the estimated value and the “true” value. With progeny test information, this accuracy measure quickly attains a high value as progeny numbers exceed 20, especially for

traits with moderate to high heritability. Traits that are lowly heritable, such as reproductive traits, require more progeny records to attain the same level of accuracy as a trait that is moderately to highly heritable.

The U.S. beef industry reports accuracy using standards suggested by the Beef Improvement Federation (BIF). The BIF accuracy (ACC_{BIF}) scale is based on minimizing Prediction Error Variance (a measure of the magnitude of errors in predicting breeding values), rather than using the correlation between the estimated and true breeding value. The two are interchangeable using the following equations.

$$Accuracy(r) = \sqrt{1-(1-ACC_{BIF})^2}$$

$$ACC_{BIF} = 1-\sqrt{1-r^2}$$

BIF accuracies are more conservative than the simple correlation, in that they require more data (e.g. progeny records in the case of a bull evaluated from a progeny test) to achieve high accuracy values (see Table 1).

Table 1. Accuracies of estimated breeding values based on (A) the correlation with true breeding values (r), and (B) the BIF accuracy (ACC_{BIF}), and the number of progeny test records required to obtain these accuracy values for traits of low (0.1) and moderate (0.3) heritability.

		Number of progeny records required	
A. Correlation (r)	**B. BIF accuracy (ACC_{BIF})**	**Low heritability (0.1)**	**Moderate heritability (0.3)**
.1	.01	1	1
.2	.02	2	1
.3	.05	4	2
.4	.08	8	3
.5	.13	13	5
.6	.20	22	7
.7	.29	38	12
.8	.40	70	22
.9	.56	167	53
0.99	.93	1921	608
0.995	.99	3800	1225

The accuracy associated with EPDs increases as more information becomes available. Initially EPDs are derived from the average of an animal's parents (called a pedigree estimate). Once an animal has its own measurement or performance record, the accuracy of the EPD increases and continues to do so as the animal has recorded progeny. Unfortunately this takes time and for some economically relevant traits (ERTs) it is not possible for animals to have a record themselves (e.g. milk production in bulls), or the record may occur very late in life (e.g. stayability). New metrics for estimating the "accuracy" of DNA tests have been developed based on the relationship between MBVs and the trait of interest, some of which are published by DNA testing companies to accompany marker panel results. It is critical to understand that at present, these values are not directly comparable to the BIF accuracy values associated with EPDs.

Example:
Assume that a DNA test has a genetic correlation (r) of 0.8 with the trait of interest. This would equate to a BIF accuracy of 0.40 (see Table 1). For traits that are hard to measure or measured late in life this would be very beneficial. Seedstock producers could identify superior animals earlier in life and commercial producers who purchase unproven sires could reduce the risk associated with low accuracy values. However, if the genetic correlation between the molecular test and the trait of interest is low ($r = 0.2$) then the value of using only the genetic test score for the purposes of selection is dramatically decreased, especially if an animal has EPDs for the trait of interest. The greatest benefit in accuracy should come from the integration of DNA tests scores along with phenotypic records in the calculation of EPDs.

The reason that DNA tests are able to increase the accuracy of EPDs is that they have the ability to account for a phenomenon called "Mendelian sampling". This term is used to describe the sampling that occurs when parents pass on a random sample of half of their DNA to their offspring. Every allele, good or bad, has an equal likelihood of being inherited. One could envision a scenario where an animal could receive only the most desirable alleles from both parents resulting in a large favorable Mendelian sampling effect or the exact opposite which could result in a large unfavorable sampling effect. Perhaps the best example of this is a set of flush mates. Although all of them have the same pedigree estimate, they may differ considerably in terms of their performance and ultimately their EPDs due to Mendelian sampling. This effect can be quantified using contemporary group deviations and is a measure of how much better or worse an animal is compared to the average of its parents. Mendelian sampling is the reason that performance records on the individual and its progeny are required to obtain accurate genetic predictions. Individual records provide some information on the sampling of alleles inherited by an animal, and progeny information provides even greater insight as to the sum of the additive effects that the animal is passing to the next generation. DNA tests have the potential to view into the black box of Mendelian sampling at birth and reveal what alleles an animal inherited.

The accuracy of a DNA test at predicting the true genetic merit of an animal is primarily driven by the amount of additive genetic variation accounted for by the DNA test. Thallman et al. (2009) found that the best predictor of this proportion was the square of the genetic correlation between the MBV and the trait of interest. The first generation of DNA tests for complex traits in beef cattle did not have high accuracies because the small number of markers included in these tests were associated with only a small proportion of the additive genetic variation for the trait of interest (Allan and Smith, 2008). As the number of informative markers in a DNA test increases so will the proportion of additive genetic variation explained by the test.

Since the first marker tests were developed, a large number of SNP markers have been identified in the bovine genome. As marker panels grow in size they will be able to track the inheritance of an increased number of genes, and if these genes are associated with genetic variability in the trait under selection then these tests will explain a larger proportion of the overall genetic variation for that trait. What will be the benefit of higher accuracy values on young sires? For the seedstock producer, it will enable the selection of truly superior animals earlier in life. It will also allow seedstock producers to supply clientele with a product that has less risk of change associated with it. The benefit to commercial producers lies in the ability to buy yearling bulls with more certainty surrounding their EPDs.

Example:
Assume a commercial producer wants to purchase a calving ease bull for use on heifers. If a bull does not have a record of calving ease himself, the BIF accuracy might be 0.20. Assume that the possible change[2] value associated with this accuracy level is 6 and that his published EPD is +5 (breed average in this case). In this situation, we would be 68% confident that this bull's "true" EPD for calving ease is between -1 and +11 realizing that for calving ease a larger number is more desirable since it is interpreted as the percentage of unassisted births. However, if the accuracy was higher (0.5) this would mean a small possible change value (4) so we would then be 68% confident that his true EPD would be between +1 and +9.

Increased accuracy values can aid in the selection of truly superior animals. For instance, if calving ease is a concern for a commercial producer who buys yearling bulls then there is an inherent risk that the bull's true genetic merit and his predicted genetic merit are not close. It would be advantageous to have more information from which to predict the genetic merit of yearling animals so that the predicted value was a closer estimate of the true value.

Example:
Assume that two yearling bulls both have a calving ease direct EPD of +5 and that the possible change values associated with them are +6. In this scenario both bulls would be equally likely to be candidates for selection. However, assume that we were able to garner more information, in the form of a marker panel test, and thus increase the accuracy values of both bulls by joining the results of the marker panel and the information included in the EPD. Perhaps we would find that one bull is actually a -1 and the other bull's is a +11. In this case the two bulls seemed equally valuable based on their low accuracy EPDs but as the accuracy values increased and we were able to get a clearer picture of their true genetic potential as parents we found one bull is actually superior over the other. In this example, the difference between the two bulls is actually 12 or one bull is likely to have 12% fewer assisted births than the other. If multiple bulls were purchased with the same low accuracy EPDs (in this case +5) it might be argued that the average of the "true" values would still be close to +5 even though some are likely to be higher and some lower. However, for a trait like calving ease, it is advantageous to eliminate bulls that may create calving difficulty even if the average of an entire bull battery is acceptable.

Shorter Generation Interval

Combining phenotypic and molecular data, particularly for traits that cannot be measured early in life, can lead to faster genetic change. The factors that impact the rate of genetic change are the accuracy of selection, the genetic standard deviation, the selection intensity, and the generation interval. Generation interval is defined as the average parental age when the offspring are born. Typically this is six years of age in beef cattle. Genetic change (ΔG) per year can be derived by the breeder's equation shown on the following page.

[2] Possible change values are standard deviations and are a measure of risk associated with different accuracy values. Possible change values differ between breeds and between traits. Updated possible change values can be found on breed association websites.

THE BREEDER'S EQUATION

$$\Delta G = \frac{[(\text{Accuracy of Selection})*(\text{Selection Intensity})*(\text{Genetic Standard Deviation})]}{\text{Generation Interval}}$$

It is clear that if the generation interval were to decrease then the rate of genetic change would increase. For seedstock producers, the ability to use a yearling sire heavily due to increased confidence in his EPD could reduce generation interval and thus lead to faster genetic progress. The benefits of including molecular information in the calculation of EPDs for yearling bulls will depend on the marker panel itself. The more genetic variation that is explained by the panel (i.e. the higher the genetic correlation (*r*)) the larger the increase in accuracy. Marker panel results should be thought of as another phenotype, correlated to the trait of interest, which can be included in the genetic prediction. In other words, the addition of the DNA panel phenotype adds to the amount of information and consequently provides an increase in accuracy proportional to the amount of variation explained by the panel.

Paradigm of Disjointed Pieces of Information

Differences in reporting styles, between EPDs and molecular test results and even between DNA companies, have led to a plethora of confusion. There are seemingly two distinct pieces of information, marker panel results and EPDs, which due to the sources of information included in them can potentially be in disagreement. This has often begged the question of which to use. Sometimes it has led to the belief that one must be incorrect.

Benefits of Combining Molecular and Phenotypic Data

An obvious benefit of combining traditional phenotypic based EPDs and the results from marker panel results is less confusion. No longer would there be a question as to which one to use. However there are other, more quantitative benefits such as the potential to increase the accuracy of EPDs and therefore selection in young animals thereby potentially enabling a decrease in the generation interval leading to more rapid rates of genetic change (see breeder's equation above).

Methods of Combining MBVs and EPDs

Rather than thinking of DNA-marker panel results as being separate and disjointed pieces of information, test results should be thought of as an indicator trait that is correlated to the trait of interest. As such, the MBVs can be included in national cattle evaluations (NCE) as a correlated trait. In this scenario it will be important to estimate the heritability of the marker score and the genetic correlation between it and other production traits as well as the phenotypic variation of the marker score. Kachman (2008) suggested that marker scores (MBVs) have a number of advantages over using the marker panel data (genotypes) directly.

Three primary advantages of incorporating MBVs into national cattle genetic evaluations are:
1) it reduces the amount of data that must be processed when conducting a genetic evaluation,
2) markers used in the test (panel) do not have to be identified, and
3) it allows for the rapid adoption and incorporation of advances in DNA tests and statistical methodology.

To incorporate DNA test information into genetic evaluations, it is necessary to estimate the accuracy (r) of DNA tests in a validation or calibration population outside the discovery set of animals. This process provides the parameters required for their incorporation into NCE. In the US this is being done by genomics companies in collaboration with some breed associations (e.g. American Angus Association (Northcutt, 2011; MacNeil et al. 2010)).

In the US there is an agreement between Angus Genetics Inc. (AGI) and both IGENITY® (Duluth, GA) and Pfizer Animal Genetics (Kalamazoo, MI) to calculate genomic-enhanced EPDs for multiple traits using DNA and traditional (performance records, pedigree) information sources and the DNA test genetic correlations shown in bold in Table 2. The bold numbers were estimated based by the American Angus Association based on records in their breed database.

Table 2. Genetic correlation (r) between commercial DNA tests and target traits for Angus cattle. Bold traits are being incorporated into American Angus Association national genetic evaluations on a weekly basis. Estimates for some traits are not publicly-available.

Trait	h^2	IGENITY® Angus Profile		Pfizer HD 50K for Angus	
		Included	AGI Genetic correlation (r)[3]	Included	AGI Genetic correlation (r)[3,4]
Average Daily Gain	0.28	X		X	.55[4]
Net/residual Feed Intake	0.50	X		X	.35[4]
Dry matter intake	**0.31**	**X**	**.45**	**X**	**.65**
Tenderness	0.37	X		X	.51[4]
Calving Ease (Direct)	**0.20**	X	**.47**	X	**.33**
Birth weight	**0.42**	**X**	**.57**	**X**	**.51**
Weaning Weight	**0.20**	**X**	**.45**	**X**	**.52**
Yearling Weight	**0.20**	**X**	**.34**	**X**	**.64**
Yearling Height	0.45	X			
Calving ease (maternal)	0.12	X		X	.63[4]
Milking Ability	**0.14**	**X**	**.24**	**X**	**.32**
Heifer Pregnancy	0.20	X			
Docility	**0.37**	**X**	**.47**		
Mature Height	0.82	X			
Mature Weight	0.55	X			
Scrotal Circumference	0.43	♂			
Stayability	**0.10**	♀			
Carcass weight	**0.31**	**X**	**.54**	**X**	**.48**
Backfat thickness	**0.26**	**X**	**.50**	**X**	**.56**
Ribeye area	**0.32**	**X**	**.58**	**X**	**.60**
Marbling score	**0.26**	**X**	**.65**	**X**	**.57**
Percent choice	--	X			

[3] Northcutt, S.L. (2011) Genomic Choices. American Angus Association®/Angus Genetics Inc. release. http://www.angus.org/AGI/GenomicChoice070811.pdf (Updated July 8, 2011)
[4] Pfizer Animal Genetics. 2010. Technical Summary. https://animalhealth.pfizer.com/sites/pahweb/US/EN/PublishingImages/Genetics%20Assets/HD50K/50K%20Tech%20Summary%204-13-10.pdf

The results of genomic predictions from both IGENITY® Profile for Angus and Pfizer HD 50K for Angus are being incorporated weekly into 10 genetic evaluations (Table 3). Docility genetic evaluations include only IGENITY MBVs as this trait is not included in the Pfizer product (Northcutt 2011). Incorporation of other listed traits (SC, YH, MW, MH) is likely forthcoming.

Table 3. Traits that are currently being incorporated into weekly American Angus Association genetic evaluations (http://www.angus.org/Nce/WeeklyEvalGenomicData.aspx).

Trait	Igenity	Pfizer
Calving Ease (CED)	✓	✓
Growth (BW, WW, YW, Milk)	✓	✓
Residual Average Daily Gain (RADG)	✓	✓
Docility (DOC)	✓	----
Carcass (CW, Marb, RE, FAT)	✓	✓
Scrotal/Yearling Height (SC, YH)	-----	----
Mature Weight/Height (MW, MH)	-----	----

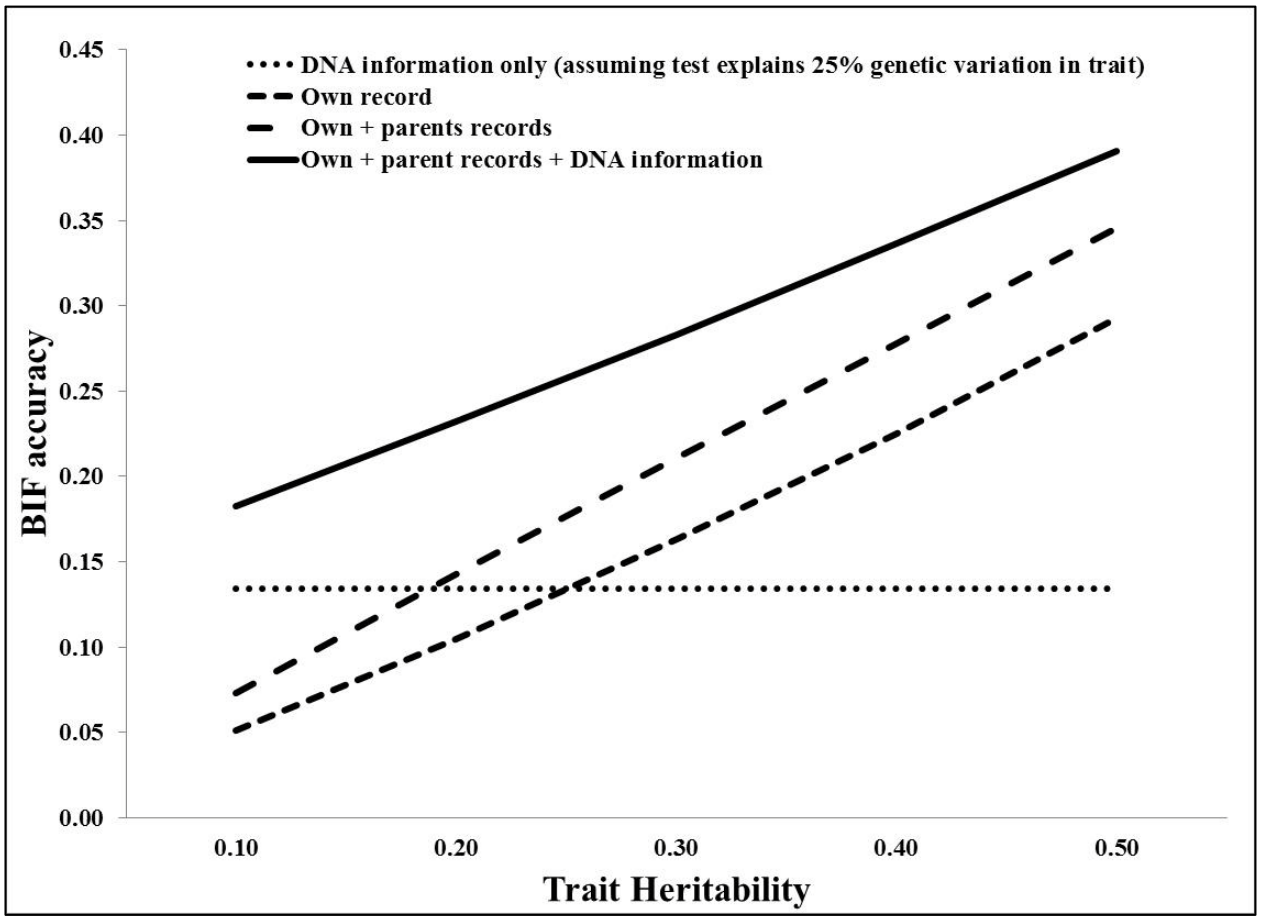
How much does molecular (DNA) information improve the accuracy of genetic evaluation?

Genomic-enhanced EPDs are produced using DNA and traditional (performance records, pedigree) information sources. The inclusion of DNA information should improve the accuracy of EPDs, especially for young animals with little performance data. Producers often ask how much a DNA test improves accuracy. The answer depends on the amount of genetic variation accounted for by the test, and the accuracy of the EPD for that trait in the absence of DNA information.

Figure 1 shows the BIF accuracy of a genetic test that is associated with 25% of the genetic variation in a trait ($r = 0.5$). If no other information existed on that trait, then the BIF accuracy of the EPD based on the DNA data alone would be 0.13 (dotted line, $ACC_{BIF} = 1\text{-}\sqrt{1 - r^2}$ where r = genetic correlation or "true" accuracy). If an animal and its parents have records for the trait (large dashed line), then the BIF accuracy would range between 0.07 and 0.35 depending upon the heritability of the trait. In that situation, the addition of DNA test information would boost the BIF accuracy most (0.11) for the low heritability trait, and least (0.05) for the most highly heritable trait. In practice this means that DNA testing is most useful for traits for which no other information exists (e.g. meat tenderness), and/or low heritability traits (e.g. calving ease).

As DNA testing becomes more comprehensive and encompasses a larger number of traits, it will provide a selection tool for traits where no other information or selection criteria exist. There are many economically-relevant traits in this category including cow and feedlot feed efficiency, and disease resistance (Pollak, 2005). This will enable the development of more comprehensive selection indexes that include all of the economically-relevant traits (ERTs) of relevance to U.S. beef production systems. One of the most important of these is likely to be feedlot health.

Figure 1. Effect of DNA information (for a test with r = 0.5) on BIF accuracy of EPDs given different sources of information and trait heritability



The accuracy associated with EPDs increases as more information becomes available. DNA information will therefore have very little impact on high accuracy AI sires. The phenotypic records of their progeny will already result in high accuracy genetic evaluations. However, EPDs are initially derived from the average of an individual's parents (called a pedigree estimate; ACC_{BIF} = 0.05). It is on these young animals with no phenotypes that DNA testing offers the most opportunity to improve the accuracy of genetic merit estimates (i.e. EPDs). This is especially true for low heritability traits where phenotypic records are collected later in life (e.g. female reproductive traits). DNA information is more valuable for low heritability traits as the boost it provides to accuracy is equivalent to a larger number of progeny records (i.e. as trait heritability decreases, DNA tests provide data equivalent to an increasing number of "progeny equivalents").

Example:

Table 4 illustrates the effect of incorporating the Pfizer HD 50K for Angus MVPs into Angus evaluations on young bulls with no phenotypic records (i.e. EPDs are based on pedigree estimate). These numbers do not reflect the fact that many young bulls will have their records (e.g. birth, weaning and yearling weight, ultrasound), and hence the ACC_{BIF} improvement will be less for traits where animals with their own phenotypic records. DNA information is therefore

particularly valuable for traits that are not available on young animals (e.g. carcass traits, fertility traits, female reproduction and longevity). Research suggest that large numbers of records will be required to obtain accurate DNA tests for hard-to-measure and low h^2 traits. DNA testing provides an attractive approach to obtain previously-absent criteria on these traits, although costs involved with obtaining sufficient records for these traits should not be underestimated, especially given the number of different beef cattle breeds.

Table 4[5]. Change in BIF accuracy (BIF_{ACC})resulting from the incorporation of Pfizer Animal Genetics molecular breeding values into AAA EPDs for young sires with no other phenotypic information (i.e. parent-average EPDs; $BIF_{ACC} = 0.05$), and average number of progeny records (Approx. progeny equivalents) it would require to obtain a similar ACC_{BIF} increase based on progeny test records.

Trait	AGI[6] Trait heritability (h^2)	AGI HD 50K Correlation (r)	Change in ACC_{BIF} – (from .05[7]) from DNA testing	Approx. Progeny Equivalents
Birth Weight (BW)	0.42	0.51	0.25	8
Weaning Weight (WW)	0.20	0.52	0.23	16
Yearling Weight (YW)[8]	0.20	0.64	0.27	20
Residual Average Daily Gain[9]	0.31	0.65	0.27	13
Milk	0.14	0.32	0.15	12
CW	0.31	0.48	0.17	7
Marb[10]	0.26	0.57	0.24	12
RE[9]	0.32	0.60	0.23	9
FAT[9]	0.26	0.56	0.23	11

Beyond Angus

To date, data suggest that tests developed for one breed are unlikely to work very well in another. Even within Angus, tests trained in North American Angus were associated with less genetic variation when used in the Australian/New Zealand Angus population, and required regional recalibration for that population and production system (Johnston *et al*., 2010). One possible reason for this is that a portion of the accuracy (*r*) associated with genetic testing is the result of relationships between the animals in the training population (the population used to develop the test) and the animals in the population where the test is being used. Tests are likely to be less accurate (explain less of the genetic variation) when used to evaluate more-distantly related animals, such as those in another country where animals may be several generations removed from U.S. genetics.

[5] Table courtesy of Dr. Kent Anderson, Pfizer Animal Genetics, June 2011.
[6] Angus Genetics Inc. Heritabilities from American Angus http://www.angus.org/Nce/Heritabilities.aspx
[7] These changes are less for higher initial accuracy values.
[8] Post-weaning average daily gain (ADG).
[9] Dry matter intake.
[10] Carcass progeny, not scanned progeny.

Bigger Panels

It's becoming increasingly clear that to obtain accurate DNA tests, it's necessary to train on large numbers of records. The dairy industry is currently training on records from tens of thousands of animals including well-proven Holstein sires to obtain their genomic-enhanced genetic merit estimates. Obtaining similarly large numbers of records from beef breeds, with the possible exception of Angus, poses a challenge for the beef industry. The recent development of very high-density (~700K SNP) bovine marker panels (e.g. Affymetrix Bovine 650K, Illumina Bovine 770K HD SNP Array) provides a potential solution.

With the widely used Illumina Bovine SNP50 BeadChip assay (50K) marker panel, a marker associated with a trait in one breed may not have the same association in another breed (Figure 2). The reason is that the marker is often located a "long" distance from the gene and so the marker is not always found to be associated with the variant of the gene that is causing the effect in all breeds. By increasing the number of SNP markers to 700 K, markers are more closely spaced and so there is a greater likelihood of finding SNPs that are close to the gene (linked), and hence the marker will "work" in both breeds.

Figure 2. Marker location relative to the gene of interest in two breeds when using (A) the 50K SNP marker panel (markers spaced at 70,000 bp (70 kb) intervals), or (B) the high density 700K SNP marker panels (markers spaced at 5,000 base pair (5 kb) intervals).

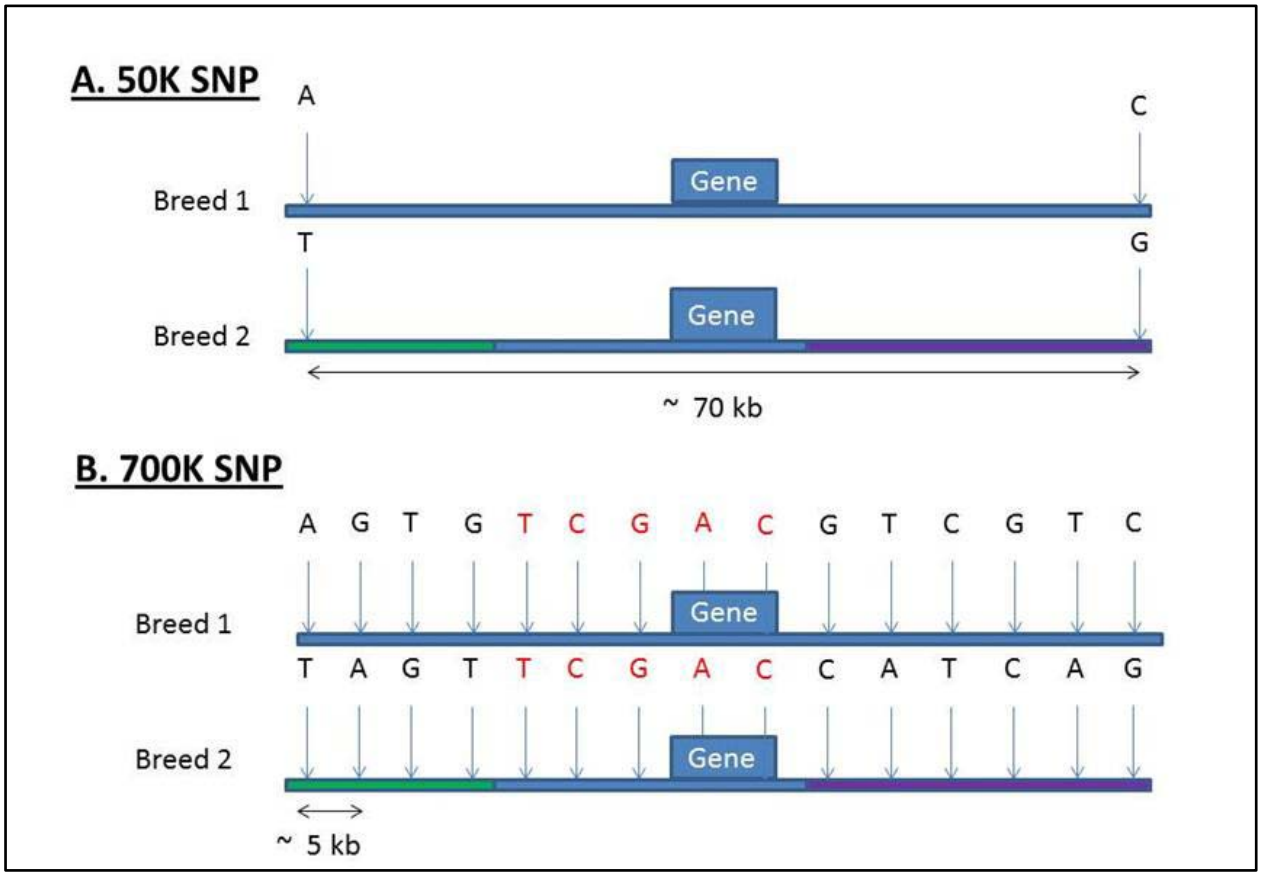
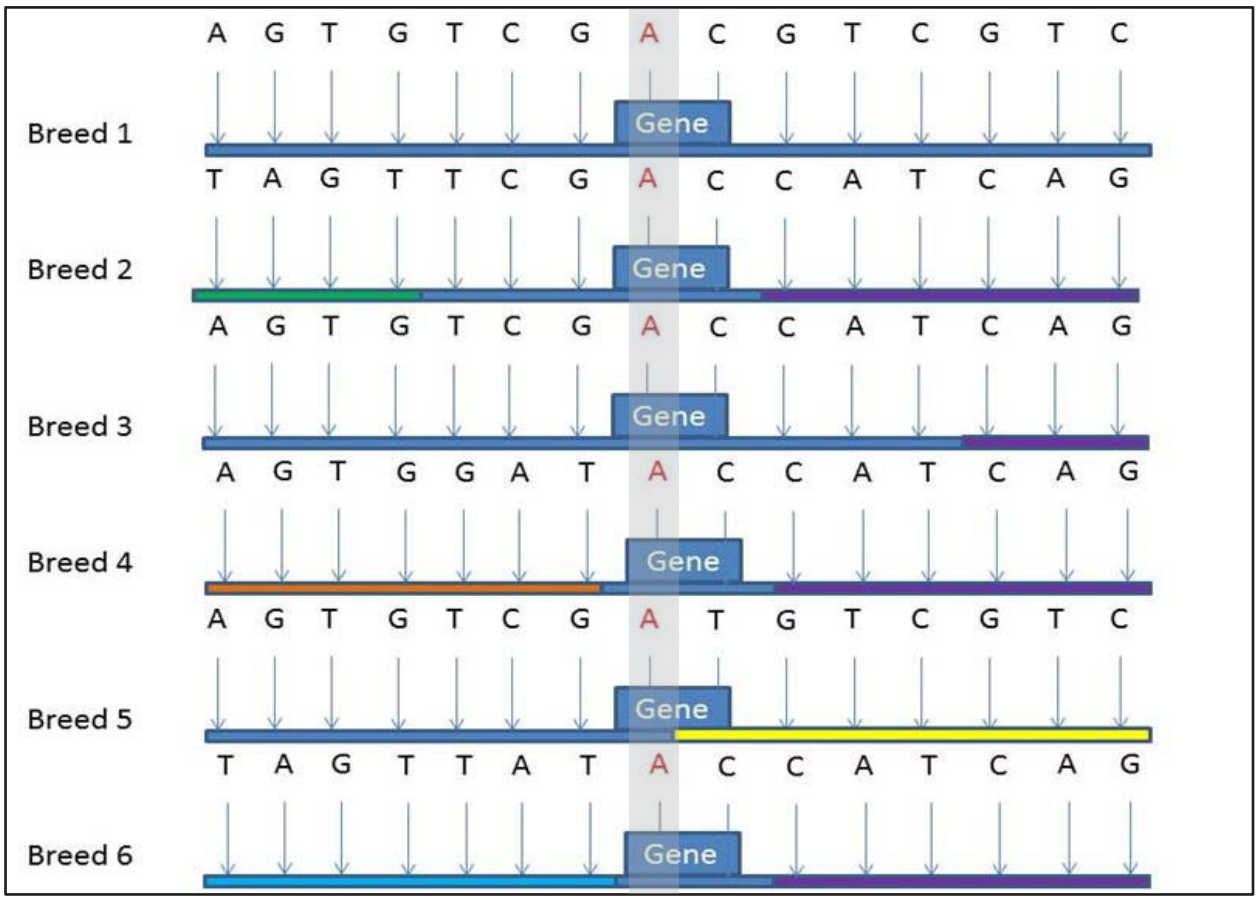


Figure 3. High density SNP panels may help to identify markers that have the same association with the trait of interest and hence the marker will "work" in multiple breeds (e.g. shaded "A").



These very high-density bovine marker panels may also provide an opportunity for breeds to pool information. It is possible that high density panels will have markers with consistent associations across breeds (Figure 3; shaded "A" marker). Developing large multi-breed training data sets to develop genetic tests may collectively improve the accuracy (*r*) of tests for all breeds, more than any single breed can do on its own due to the larger number of combined records.

These high-density panels come with increased cost (~$200 each), but this price will likely drop in the future. Ultimately the availability of high-density arrays may enable the use of DNA test information to improve EPD accuracy for many breeds of cattle, including those with a relatively small number of performance-recorded animals. Although this field is evolving quickly, the ultimate goal for using genomics remains the same, and that is to cost-effectively improve accuracy of EPD produced in National Cattle Evaluation (NCE). As outlined by BIF, "*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*" (Tess, 2008).

Economics

DNA testing presents a marketing opportunity for bull sellers. Early adopters, those who have panel information sent to breed associations for inclusion in genetic evaluations, may have a competitive advantage over other seedstock producers who do not. This assumes that bull buyers are willing to pay more for yearling bulls with higher accuracy values. The process of collecting DNA samples and then paying for a diagnostic test for a particular trait represents an additional cost to the breeder. Some seedstock producers are currently DNA-testing their bulls to provide potential buyers with DNA information. The value of that information to the buyer is will be determined by the market. If the value is more than the cost of testing and is reflected in the bull purchase price, then the seedstock producer will have improved their bottom line.

DNA testing also presents an opportunity to improve the accuracy of EPDs and accelerate the rate of genetic progress. The response to selection and therefore the value associated with DNA testing is dependent upon how much the DNA information improves the accuracy of genetic evaluations at the time of selection, and the value of a unit of genetic improvement. One way to determine the value of a unit of genetic improvement is to develop an economic "selection index" that weights all traits on their relative economic importance. Indexes consider both the "input" or expense side and the income side of selection decisions and enable cattle producers to make balanced selection decisions, taking into account the economically-relevant growth, carcass and fertility attributes of each animal to identify which animals are the most profitable for their particular commercial enterprise. Melton (1995) suggested that US cow-calf producers should have a relative economic emphasis of 47% on reproduction, 24% on production, and 30% on carcass traits whereas producers in an integrated system should have a relative economic emphasis of 31% on reproduction, 29% on production and 40% on carcass traits. This relative emphasis will depend on how much the value derived from genetic gain in carcass traits is shared with the producer in the integrated system.

Van Eenennaam et al. (2010) undertook a simulation study to determine "*What is the value of DNA tests to increase the accuracy of beef bull selection in the seedstock sector?*" A seedstock operation consisting of a closed nucleus of 600 breeding females was modeled. It was assumed that all of bulls in the calf crop were DNA tested and that in the absence of DNA test information, EPDs on young, untested bulls were informed by their own performance records along with those of their sire, dam and 20 paternal-half sibs. Each year the top 8 bulls were selected to be stud sires, and 125 (remaining bulls from the top half of the calf crop) were made available for sale to commercial producers. Commercial sires were then used to sire four calf crops at a mating ratio of 25 females: 1 male (i.e. they were exposed to a total of 100 cows). DNA test information from a hypothetical DNA test panel that accurately predicted ALL of the traits in the selection index was combined with performance records to increase the accuracy of EPDs. DNA test information increased $Index selection response 20-41% over that obtained with performance recording alone, depending upon the traits in the breeding objective. It was assumed that the seedstock breeder paid for the DNA tests, and so results were presented as the value derived per DNA test purchased. It was reasoned that two times this value (to DNA test both the bull that was sold and the one that was not because he was in the bottom half of the calf crop) would be the amount of additional revenue the seedstock operator would need to recover from the sale of genetically-superior commercial bulls to breakeven on the cost of DNA testing.

Commercial Bull Selection

The value derived from using DNA-test information to enable more accurate selection of genetically-superior commercial bulls ranged from AU$61-135 per bull. Assuming that the entire bull calf crop (n = 267) was tested and that the top half of the bulls (n=125) were sold as commercial sires, the breakeven value of the genetic gain derived from DNA testing ranged from **$30-67** per DNA test depending upon the market being targeted. These values assumed commercial producers were willing to pay a price premium for genetically-superior bulls, and some form of industry vertical integration or profit sharing between sectors such that the rewards for improvement in processor traits (e.g. dressing %, marbling score, etc.) were transferred along to commercial producers and breeders. The value of DNA tests to improve traits of direct value to commercial cattle enterprises (e.g. fertility) would be less than this value. For example, 69% of the returns from including DNA data in commercial sire selection for an integrated system index were derived from traits that generate a direct return to processors.

Commercial Replacement Heifer Selection

DNA testing could also be used to select replacement commercial females, many of which have no EPD information. The beef industry would benefit greatly from improvement in traits directly affecting maternal performance (Roughsedge et al. 2005). The value of testing heifers will depend upon the information available at the time of selection, the power of the genetic test, and the selection intensity (i.e. what proportion of the available heifers you require for replacements). The latter is dependent upon the calving and replacement rates. The breakeven cost of testing all of the potential replacement heifers in a commercial herd with a replacement rate of 20% and 45 potential replacement heifers born per 100 cows per year using a DNA test with an index accuracy (r) of 0.25 was calculated to be less than $20 using the assumptions outlined in Van Eenennaam et al. (2010), and that the commercial producer was not recording heifer performance such that there were no other data upon which to base heifer replacement decisions (Van Eenennaam and Drake, 2011). In practice, selection for replacement heifers is frequently driven by birth date and size as heifers that are born later in the calving season are often too young to be cycling in time for the first potential breeding season.

The value derived DNA test is less for replacement heifers than for bulls as commercial females produce fewer descendants in their lifetime. Unless DNA tests have been shown to have high accuracies (r) for maternal traits, DNA test information should be used in conjunction with available phenotypic data to select replacements. And here is the quandary when developing accurate tests for replacement female selection. Traits that are of the most economic value to self-replacing herds are low heritability reproductive traits including age at first calving, reproductive success and replacement rate (Roughsedge et al. 2005). Research results suggest that very large numbers of records will be required in training populations to obtain accurate DNA tests for low heritability traits (Goddard, 2009). However as commercial producers often have little information upon which to base their replacement heifer selection decisions, DNA testing provides an attractive approach to obtain previously-absent selection criteria on reproductive traits prior to replacement selection, although such tests will likely need to be inexpensive to be commercially viable.

The Future

Marker panels are likely to continue to grow in size and in the future it may even become cost-effective to obtain whole genome sequence on individual animals, i.e. sequence all 3 billion nucleotide base pairs of the bovine genome! Some significant hurdles remain for the successful implementation of DNA-based selection of genetically-superior animals in the beef industry. These include the need for large discovery populations (i.e. thousands) of genotyped and phenotyped cattle to enable DNA tests to accurately predict genetic merit in unrelated animals (Goddard, 2009). It also seems probable that additional phenotyped, genotyped populations of young animals will be needed in the future to continually update the associations between DNA markers and traits of interest (Van Eenennaam, 2011).

The value proposition of DNA testing to improve the accuracy of EPDs may shift if the value of genetic gain changes appreciably. This might happen if genomic or other technologies result in the introduction of novel traits into breeding objectives. This could be driven by new production system requirements, health concerns, or through emerging technologies which enable selection for traits previously omitted from breeding objectives due to a lack of selection tools or criteria. A DNA test accounting for 50% of the genetic variation in a difficult-to-measure, economically-important trait like feed efficiency is likely be more valuable than a DNA test accounting for 50% of the variation in an easily measured trait like weaning weight. Likewise, there would be great value in the development of accurate genetic tests for the identification of animals resistant to bovine respiratory disease (BRD) or pneumonia. BRD, also known as "shipping fever", is the leading natural cause of death in U.S. beef and dairy cattle, resulting in annual losses of more than one million animals and ~$700 million. It has been estimated that to maximize profit in a commercial production system retaining ownership through the feedlot, a multi-trait selection index for Angus terminal sires would place **six** times more emphasis on EPD for BRD resistance (if such an EPD existed) than on EPD for weaning weight (Van Eenennaam and MacNeil, 2011). As a result of the potential value associated with selection approaches to improve hard-to-measure, economically important traits; two 5-year, multimillion dollar, multistate projects to discover genetic markers associated with both feed efficiency and BRD resistance were funded by the USDA National Institute of Food and Agriculture (NIFA) in 2011. The successful outcome of these projects would be of considerable value to the beef industry.

Summary

The advent of molecular information in the form of both tests for simply inherited traits and complex traits has created both excitement and some confusion. The lag between discovery and application has been decreased, allowing for technology to be rapidly delivered to industry. In certain cases this has created uncertainty as to how DNA information should be used in making selection decisions. DNA marker tests results should not be used to replace traditional selection based on EPDs and economic index values, but rather should be seen as providing an additional source of information from which to predict genetic merit. When included in EPDs, DNA information improves the accuracy of genetic predictions, especially on young animals. DNA testing holds the greatest promise for economically-relevant traits which are too expensive to measure, and for which no good selection criteria exist (e.g. feed efficiency). Meaningful incorporation of these traits into national cattle evaluations will be required to make the best use of information and ensure that appropriate economic weighting is given to DNA test information.

Literature Cited

1. Allan, M. F. and T. P. L. Smith. 2008. Present and future applications of DNA technologies to improve beef production. Meat Science 80:79-85.
2. Goddard, M.E. 2009. Genomic selection: prediction of accuracy and maximization of long term response. Genetica 136: 245-257.
3. Johnston, D.J., Jeyaruban G.J. and Graser H.-U. 2010. Evaluation of Pfizer Animal Genetics HD 50K MVP Calibration. http://agbu.une.edu.au/pdf/Pfizer_50K_September%202010.pdf.
4. Kachman, S. 2008. Incorporation of marker scores into national cattle evaluations. Proc. 9th Genetic Prediction Workshop, Kansas City, MO, pp. 92-98. http://www.beefimprovement.org/PDFs/Kansas%20City%20Missouri%202008.pdf
5. 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: 517-522.
6. Melton, B.E. 1995. Conception to consumption: The economics of genetic improvement. In Proceedings of the Beef Improvement Federation 27th Annual Meeting and Research Symposium. pp. 40-87.
7. Northcutt, S.L. (2011) Genomic Choices. American Angus Association/Angus Angus Genetics Inc. release. http://www.angus.org/AGI/GenomicChoice070811.pdf (Updated July 8, 2011).
8. Pollak, E.J. 2005. Application and impact of new genetic technologies on beef cattle breeding: a 'real world' perspective. Australian Journal of Experimental Agriculture 45:739-748.
9. Roughsedge T., P.R. Amer, R. Thompson, and G. Simm. 2005. Development of a maternal breeding goal and tools to select for this goal in UK beef production. Animal Science 81, 221-232.
10. Tess, M.W. 2008. Guidelines for combining molecular and quantitative approaches in genetic evaluation. Proc. 9th Genetic Prediction Workshop, Kansas City, MO. pp. 76-82. http://www.beefimprovement.org/PDFs/Kansas%20City%20Missouri%202008.pdf
11. Thallman, R.M., K.J. Hanford, R.L. Quaas, S.D. Kachman, R.J. Templeman, R.L. Fernando, L.A. Keuhn, and E.J. Pollak. 2009. Estimation of the proportion of genetic variation accounted for by DNA tests. Proc. 41st Beef Improvement Federation, Sacramento, CA.
12. Van Eenennaam, A. L., J.H. van der Werf, and M.E. Goddard. 2010. Value of DNA information for beef bull selection. 9th World Congress of Genetics Applied to Livestock Production. August 1-6, 2010, Leipzig, Germany http://www.kongressband.de/wcgalp2010/assets/pdf/0094.pdf
13. Van Eenennaam, A. L. 2011. Beef translational genomics: Lessons from the literature. Proceedings of the Association for the Advancement of Animal Breeding and Genetics 19: 271-278.
14. Van Eenennaam, A. L., and D. J. Drake. 2011. Where in the beef cattle supply chain do DNA tests generate value? Animal Production Science. *In press.*
15. Van Eenennaam, A. L. and M. D. MacNeil. 2011. What weighting should be given bovine respiratory disease (BRD) resistance in selection decisions? Proceedings of the Beef Improvement Federation 43rd Annual Meeting and Research Symposium. Pages 61-68. http://www.beefimprovement.org/PDFs/2011-BIF%20Proceedings%202011%20final.pdf.