



# Value of Collecting Phenotypes

Matthew Spangler  
University of Nebraska, Lincoln  
mspangler2@unl.edu

Dr. Darrh Bullock  
Dr. Jared Decker  
Dr. Megan Rolf  
Dr. Matthew Spangler  
Dr. Alison Van Eenennaam  
Dr. Robert Weaver



This factsheet was developed as part of USDA NIFA grants # 2013-68004-20364 #2011-68004-30367 #2011-68004-30214



United States Department of Agriculture  
National Institute of Food and Agriculture

## Introduction

From a historical point of view, there have been considerable changes made to National Cattle Evaluations (NCE) over time. More recently, many beef breed associations have augmented EPD with genomic information. This step alone has included many rapid evolutions both in terms of methods of incorporation and the source of genomic information. Changes include new genotyping platforms, the usefulness of genomic information in predicting genetic merit, and our understanding of how best to utilize it.

Prior to genomic information being first integrated into NCE by the American Angus Association in 2009, genomic information (Molecular Breeding Values; MBV) were viewed by some producers as competing sources of information to traditional EPD. This created confusion as to which piece of information to utilize. Even after the incorporation of genomic predictors into NCE, new implementation issues were evident in the beef seedstock industry. Retraining, or recalibration (the process of re-estimating SNP effects and refining the resulting genomic prediction equation), became a necessity and the beef industry understood that the efficacy of genomic predictors were not robust (persistent) over several generations. The lack of predictive ability across breeds was also very clear, and the use of genomic predictors trained in Angus

could not be used with any beneficial degree of accuracy in a closely related breed like Red Angus. Consequently, for breeds to capitalize on the benefits of augmenting traditional EPD with genomic information, they must first make an initial investment in developing a “training” population of genotyped and phenotyped animals, upon which to train the genomic prediction equations. Generally speaking, breed associations were advised to genotype a minimum of 1,000 animals that preferably had moderate to high accuracy EPD. To date several breed associations have met this mark and are currently computing EPD incorporating genomic information.

## Accuracy

How well a particular genomic test improves the accuracy of an EPD in the context of selection is related to how much of the genetic variation the marker test explains. The magnitude of the benefits depends on the proportion of genetic variation (%GV) explained by a given marker panel, where the %GV is equal by the square of the genetic correlation multiplied by 100. Table 1 shows the relationship between the genetic correlation (correlation between predicted and true genetic merit; true accuracy), %GV, and the Beef Improvement Federation (BIF) accuracy. BIF accuracy is the standard for all U.S. beef breeds.

**Table 1.** The relationship between true accuracy (r), proportion of genetic variation explained (%GV), and Beef Improvement Federation (BIF) accuracy.

<b>R</b>	<b>%GV</b>	<b>BIF</b>
0.1	1	0.005
0.2	4	0.02
0.3	9	0.046
0.4	16	0.083
0.5	25	0.132
0.6	36	0.2
0.7	49	0.286

From Table 1 it is clear that even when the %GV is exceptionally large, the corresponding BIF accuracy is relatively low. This suggests that although genomics has the potential to add additional information, by itself it is far from a perfect predictor of an animal's genetic merit.

Figures 1 and 2 illustrate the benefits of including genomic information into EPD (or Estimated Breeding Value (EBV) which is twice the value of an EPD) accuracy (on the BIF scale) when the genomic information explains 10 or 40% of the genetic variation (GV), which is synonymous with R<sup>2</sup> values of 0.1 and 0.4. The darker portion of the bars shows the EPD accuracy before the inclusion of genomic information and the lighter colored portion shows the increase in accuracy after the inclusion of the genomic information into the EPD calculation. As the %GV increases, the increase in EPD accuracy becomes larger. Additionally, lower accuracy animals benefit more from the inclusion of genomic information, and the benefits decline as the EPD accuracy increases.

Regardless of the %GV assumed here, the benefits of including genomic information into EPD dissipate when EPD accuracy is between 0.6 and 0.7.

On the other hand, when %GV is 40, an animal with 0 accuracy could exceed an accuracy of 0.2 with genomic information alone. This would be comparable to having approximately 4 progeny for a highly heritable trait, or 7 progeny for a moderately heritable trait (Table 2).

### Phenotypes in the Genomic Era

While these gains in accuracy are impressive, particularly for non-parent animals, it is clear that genomic information alone cannot “prove” a sire. Or in other words, additional information is required before an animal can achieve very high levels of BIF accuracy. To reach high levels of accuracy it is necessary to collect and submit phenotypic information on the animal's progeny.

There is still a need, and tremendous benefit from, the continued collection of phenotypes in the context of genomic selection. The benefits fall into two broad classifications:

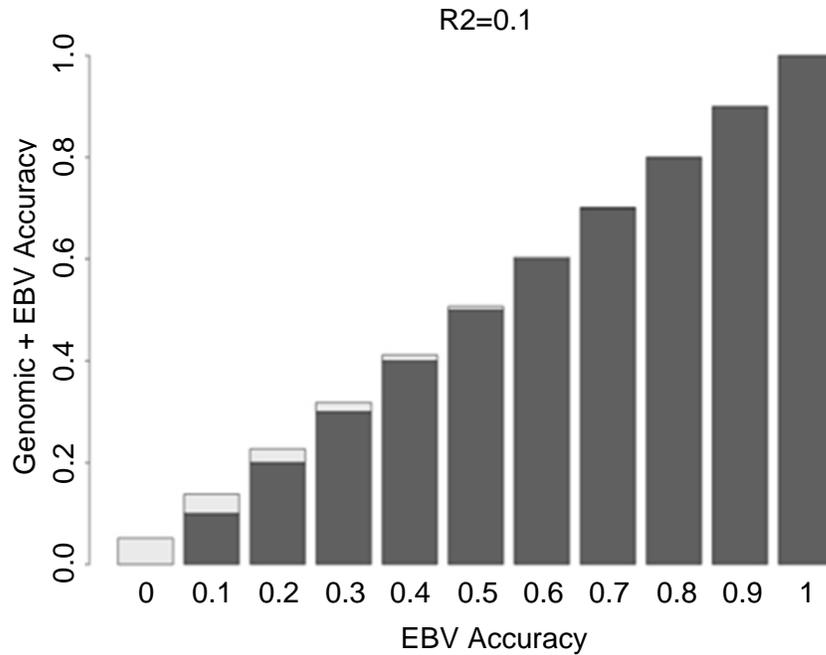
#### 1) *Training Population*

Animals with phenotypes are needed in order to develop the initial training population. Ideally these animals have moderate to high accuracy EPD, which

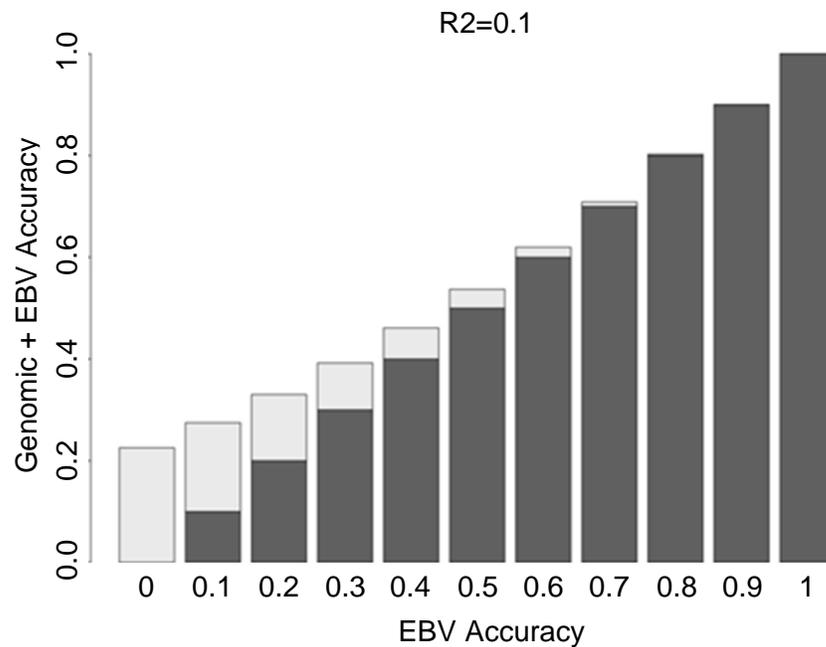


This factsheet was developed as part of USDA NIFA grants  
 # 2013-68004-20364  
 #2011-68004-30367  
 #2011-68004-30214





**Figure 1.** Increase in accuracy from integrating genomic information that explains 10% of the genetic variation into Estimated Breeding Values (EBV).



**Figure 2.** Increase in accuracy from integrating genomic information that explains 40% of the genetic variation into Estimated Breeding Values (EBV).

This factsheet was developed as part of USDA NIFA grants # 2013-68004-20364 #2011-68004-30367 #2011-68004-30214



United States Department of Agriculture  
National Institute of Food and Agriculture

**Table 2.** Approximate number of progeny with phenotypic information needed to reach accuracy levels (true  $r$ ) and the BIF standard) for three heritabilities ( $h^2$ ).

Accuracy		Heritability Levels		
R	BIF	$h^2$ (0.1)	$h^2$ (0.3)	$h^2$ (0.5)
0.1	0.01	1	1	1
0.2	0.02	2	1	1
0.3	0.05	4	2	1
0.4	0.08	8	3	2
0.5	0.13	13	5	3
0.6	0.2	22	7	4
0.7	0.29	38	12	7
0.8	0.4	70	22	13
0.9	0.56	167	53	30
0.999	0.99	3800	1225	700

would require that they have several progeny (refer to Table 2) that have the phenotype recorded. If routine phenotype collection does not occur, building the initial training set will be problematic.

Genomic predictions need to be “retrained” overtime, and to do so requires additional animals to be included in the training population. Similar to building the initial training population, this requires that newly selected animals are routinely measured for the trait of interest thus building EPD accuracy and providing additional information from which more reliable genomic predictors can be derived.

### 2) Added Accuracy

Although genomic predictors have been shown to increase EPD accuracy (refer to Figures 1 and 2), in isolation

they do not have the ability to increase BIF accuracy to high levels alone. To continue to build the accuracy of an animal’s EPD, the animal must have progeny recorded that have been measured for the trait of interest.

### Conclusions

The inclusion of genomic predictors into NCE offers an exciting and powerful tool to increase the rate of genetic gain by increasing accuracy of EPD, particularly of young animals, and by reducing the generation interval if younger sires are used more heavily. However, genotyping animals does not replace the need for phenotyping. Doing so inherently limits the upper bound of accuracy far below what is possible if additional phenotypes are collected. Genomic predictors should be viewed as an additional source of information for EPD calculations, not the complete picture.



This factsheet was developed as part of USDA NIFA grants # 2013-68004-20364 #2011-68004-30367 #2011-68004-30214



United States Department of Agriculture  
National Institute of Food and Agriculture