

# Balancing Phenotype and Genomics

by Dr. Ryan Boldt, IGS Lead Geneticist

*With advancements in genomic technology, do I still need to turn in birth weight, weaning weight, and yearling weight data, and other phenotypes on my herd?*

In the age of genomic testing and genomically enhanced EPD, do we really need to collect phenotypes? This is a question that is asked of me from time to time, and it is a good one. For those who are wanting a summary, the answer is yes. **The collection of phenotypic information is still necessary and arguably more necessary than prior to the incorporation of genomics into single-step genetic evaluations.**

To show an example in terms of EPD prediction and impact of genomics, some statistics were pulled from the IGS database to compare the average accuracy based on the amount of information included in the genetic evaluation. For this analysis weaning weight (WW) will be used as the example trait. The following table displays the effect on the WW EPD accuracy of non-parent animals with various combinations of data available to the genetic evaluation:

Data Available on Non-Parent Animals	Avg. WW EPD Accuracy
Pedigree information only	0.14
Pedigree and individual phenotype information only	0.25
Pedigree information and a genomic test (no phenotypes)	0.41
Pedigree information, genomic test, and phenotype information	0.44

This illustrates that as data is added to the prediction of non-parent animals, the accuracy of the EPD increases. However, the question becomes that without a phenotype, an animal with genomics has a higher accuracy than an animal with only a phenotype recorded. This tradeoff is largely the reason for the misunderstanding that the genetic evaluation is at a point where phenotypes are not required. However, previous studies have shown that if we do not continue to supply the evaluation with both genomic and phenotypic information, the accuracy of the genomic portion of the EPD calculation quickly diminishes. The ability of genomic data to predict and improve EPD accuracy is dependent on the amount of the variation that can be explained by the markers included in the model.

In terms of genetic prediction, the types of variation that are most important are phenotypic and genetic variation. Phenotypic variation is the amount of variation that is seen in the measurements that are collected on animals or the total amount of variation. This variation is influenced by not only the genetics of the animal, but also the environment that the animal experiences.

When it comes to genetic prediction and genetic improvement, the amount of variation that can be explained due to differences in genetics that can be inherited is known as additive genetic variation. The important distinction here is that we want to focus on the genetics that can be inherited from one generation to the next. In terms of genomics being the only source of data used for genetic prediction, then the genotype would be required to account for 100% of the genetic variance of the trait. While genotypes can account for a certain portion of the variation in each trait, the ability of genotype information to account for all genetic variation has not been achieved in any trait. Also, the amount of variation that the genotype can account for does improve over time as more animals have both genotypes and phenotypes included in the evaluation.

There are situations where the ability to get genetic predictions for a trait is solely through genomic data. These situations are when collection of phenotypes is extremely cost prohibitive. Therefore, the only tangible way to be able to create predictions is through collection of the phenotypes on a subset of animals with genomic data and identify which genomic markers are most influential and their effect. Then this can be applied to a larger group of animals.

Inclusion of genomic information into EPD calculations does improve EPD accuracy, especially on young, non-parent animals. **The inclusion of both phenotype and genotype information on animals leads to the maximum accuracy at a young age to make selection decisions.** On a larger scale, the more animals that have the maximum amount of information, the more accurate predictions will be for all animals included in a genetic evaluation. **S**