An Introduction to Genomics

With several breed societies already incorporating genomic information into their BREEDPLAN analyses, and several more aiming to move to genomic evaluations in the not-too-distant future, this Technical Note has been developed to give beef producers an introduction to genomic technology.

This Technical Note outlines what genomics is, how genomics works, and how the inclusion of genomic information into the BREEDPLAN analysis can benefit beef producers.

WHAT IS GENOMICS?

The genome is the genetic material of an organism; that is, all of an organism's DNA. Genomics is the study of the genome, while genomic selection refers to the inclusion of DNA information into a genetic evaluation and selection program (such as BREEDPLAN). The DNA information (or genotypes) used for genomic selection is in the form of thousands of genetic markers, known as Single Nucleotide Polymorphisms (SNPs). Genotypes containing these thousands of SNP markers are generated by analysing a DNA sample from the animal on a SNP chip. SNP chips are available in different densities; for example the SNP chip might have 20,000 SNP markers (20K), 50,000 SNP markers (50K) or 800,000 SNP markers (800K).

To generate genomic information on their animals, breeders take a hair sample from their animals and request the breed society or a laboratory to do a genomic test on the sample. The DNA gets extracted from the hair sample, and the DNA is run on a SNP chip (density specified by the breeder, with the denser chips being more expensive). The resulting SNP genotypes can then be used, along with pedigree information and performance information, in the calculation of BREEDPLAN EBVs.

HOW DOES GENOMIC SELECTION WORK?

When genomic information is not included in the BREEDPLAN analysis (the current situation for most Australian breeds), the BREEDPLAN analysis uses pedigree information and performance data (both on the individual and the related animals) to generate EBVs. When genomic information is implemented for a breed specific GROUP BREEDPLAN analysis, breeders will be able to take a hair sample on an individual animal, send the sample to the lab, and have the sample genotyped on one of the available SNP chips. The genotype information needs to be supplied to the breed to be included in the BREEDPLAN analysis and used, in conjunction with pedigree and performance information, to generate EBVs.

For genomic selection to work, a reference population is required. The reference population consists of thousands of animals which have both phenotypes (performance data) and genotypes (Figure 1). Setting up a reference population has been one of the challenges in implementing genomics in many breeds of beef cattle; for most breeds to date, there simply have not been enough animals with both phenotypes and genotypes available to form an effective reference population. The ideal reference population has phenotypes collected on all traits of economic importance and interest to the breed.

The second group of animals in Figure 1 are those which have genotypes, but do not yet have phenotypes collected. Typically these are young animals which have not yet reached an age where they can be performance recorded. However; any animal with a genotype but no performance data fits into this second group.

Genomic selection uses the known relationships between the phenotypes and genotypes of the animals in the reference population to calculate genomic EBVs for young animals (Figure 1).

There are several factors that will influence how well genomic selection works. Firstly, the size of the reference population is critical. For genomic selection to work successfully in Australian beef cattle, a reference population with a minimum of three to five thousand
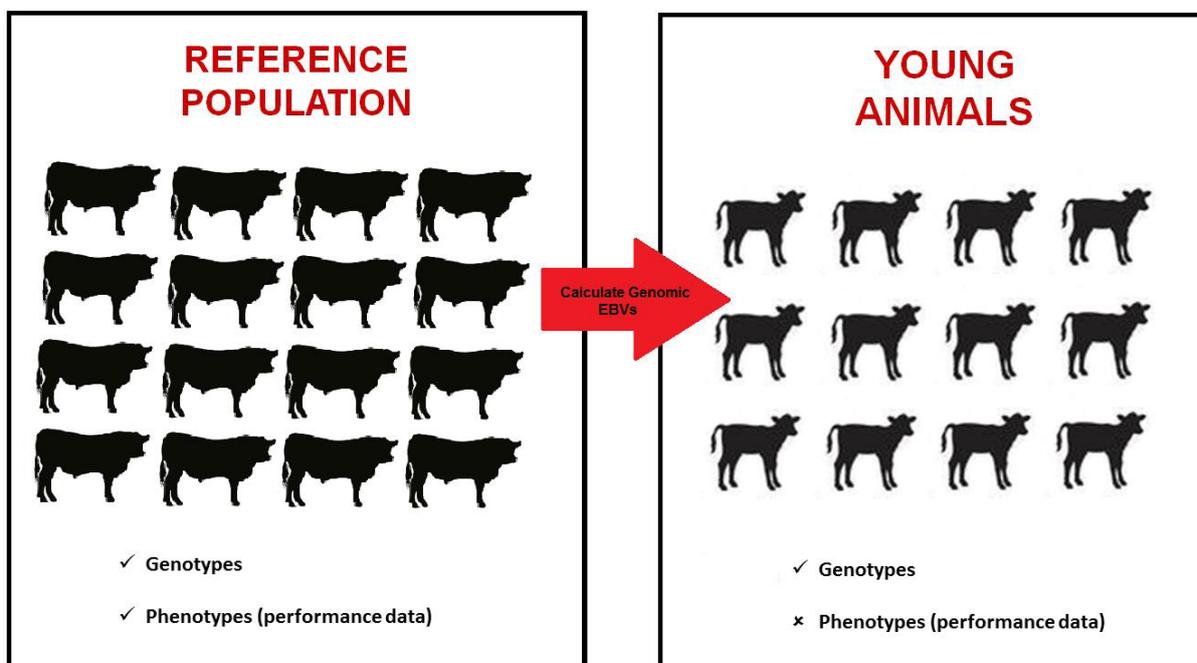
Figure 1. A reference population is critical to the implementation and ongoing success of genomic evaluations. The reference population consists of animals which have phenotypes (performance data for economically important traits) and genotypes available. The relationship between the genotypes and phenotypes of the reference population can be used to calculate genomic EBVs for young animals which have genotypes but do not have phenotypes (performance data).

animals with both phenotypes (performance data) and genotypes will be needed. Secondly, genomic selection works best when the reference population is closely related to the young animal population for which genomic EBVs are being calculated. For this reason, the reference population should be designed to represent the whole genetic pool of a breed, rather than just a subset of genetics within a breed. This is also one of the reasons why genomics will not replace performance recording – there is a requirement that animals from the next generation have both genotypes and phenotypes coming into the reference population over time. It is important that seedstock producers understand genomics will not replace performance recording; the work that you as seedstock producers do to performance record your animals will be critical for the success of genomics in the future.

WHAT BENEFITS CAN BEEF PRODUCERS EXPECT FROM GENOMICS?

When genomics is implemented, a breeder will be able to take a hair sample on an individual animal, send the sample away for genotyping, and the genotype information will be included in the BREEDPLAN analysis and used to generate EBVs. This will have two main applications for seedstock producers:

1. *EBVs can be generated for animals which do not have performance data*

    Within any breed, there will be a number of animals which do not have performance information as they are from herds which do not record performance data. In the future, with the inclusion of genomic information into the BREEDPLAN analysis, these animals could be genotyped and get BREEDPLAN EBVs. These animals could be from both seedstock and commercial herds.

    There are also a number of animals which are in BREEDPLAN herds, but do not have performance information for some traits. This may be because:

■ The animal is too young to have been measured for that trait. For example, a 200 day old calf will not have been ultrasound scanned, so is unlikely to have EMA, Rib Fat, Rump Fat or IMF EBVs. Where the 200 day old calf does have carcase EBVs, these are likely to be mid-parent EBVs of fairly low accuracy.

■ The trait is hard and/or expensive to measure. For example, Net Feed Intake (NFI) is measured in feedlot trials where the animals are on ad libitum feed for nearly 100 days (including the pre-trial adjustment period). This makes NFI very expensive to measure,

and thus NFI measurements are usually only collected on animals in progeny test programs. Retail Beef Yield is another good example; measuring Retail Beef Yield is very expensive because the carcase has to be completely boned out and the individual retail cuts trimmed and weighed.

- The trait is only able to be measured in one sex. For example, Mature Cow Weight is only recorded for females.

- The trait can only be measured once the animal is dead. For example, abattoir carcase information, including Retail Beef Yield and Marbling, is only measured on carcases, and not from live animals. The beef industry utilises live animal ultrasound scanning measurements as a way around this problem, but the actual carcase measurements can only be done on dead animals.

- Even when an animal does have performance information, this information may not be able to be used effectively by the BREEDPLAN analysis. For example, when an animal is placed in a single animal contemporary group, its performance information cannot be used by the BREEDPLAN analysis to calculate EBVs. As a result, most animals in single animal contemporary groups have mid-parent EBVs until performance information can be collected on their own progeny or other relatives.

Whatever the reason an animal does not have performance information; the animal will either not have EBVs for the trait in question, or have mid-parent EBVs of fairly low accuracy.

With genomics, such animals could be genotyped and get EBVS for a range of traits (provided that the inclusion of the genomic information into the BREEDPLAN analysis means that the EBVs reach the minimum accuracy threshold required to report). Animals which are too young to be performance recorded for a trait could be genotyped at a young age (e.g. at birth) and get EBVs that normally they would not receive until they were much older (e.g. rising 2 year olds with scan data).

Similarly, where a seedstock producer wanted EBVs on stud animals for hard to measure traits, genomics would mean that relevant animals could be genotyped and EBVs generated using the genomic information.

2. *More accurate EBVs can be generated for animals with limited performance information*

Currently, a young animal with no performance data (either of its own or its progeny) will have mid-parent EBVs with quite low accuracy. With the addition of its own performance data and performance data of its progeny, the EBVs will change (depending on how well the individual and its progeny perform within their contemporary groups) and the accuracy associated with the EBVs will increase.

Genomics will "boost" the accuracy of BREEDPLAN EBVs; this benefit is most pronounced when the animal has EBVs with low accuracies. For example, a young animal may have an accuracy of 30% for one EBV; with the inclusion of a genomic test, the accuracy for that EBV might become 40%. However, an older animal, which might have an accuracy of 90% for the same EBV, would likely only have an increase to 92% accuracy for that EBV with the inclusion of genomic information.

In this way genomics can be considered similar to the addition of progeny performance data into the BREEDPLAN analysis; when the accuracy of an EBV is low additional data has a large effect, and when the accuracy of an EBV is high, additional data has a small effect. Of course, the improvement in the accuracy of an EBV due to the inclusion of genomic information will vary for each trait (depending on the size of the reference population and the heritability of the trait) and for each animal (depending on how closely the animal is related to the reference population).

The applications of genomics as discussed above will allow seedstock producers to identify elite bulls and heifers at younger ages than is currently possible. For example, consider an ET program where 10 full sibling bulls have been born (Figure 2). As full siblings, these ET calves will share on average 50% of their DNA (e.g. full sibling ET calves are not genetically identical). These young bull calves, at marking time, will have identical mid-parent EBVs with low accuracy. This is illustrated in Figure 2A, where all 10 bull calves have a mid-parent Selection Index value of $100. Which one(s) should be kept as bull(s) for breeding purposes, and which ones should be steered? This is a difficult decision, because the breeder has limited knowledge of the genetic potential of these
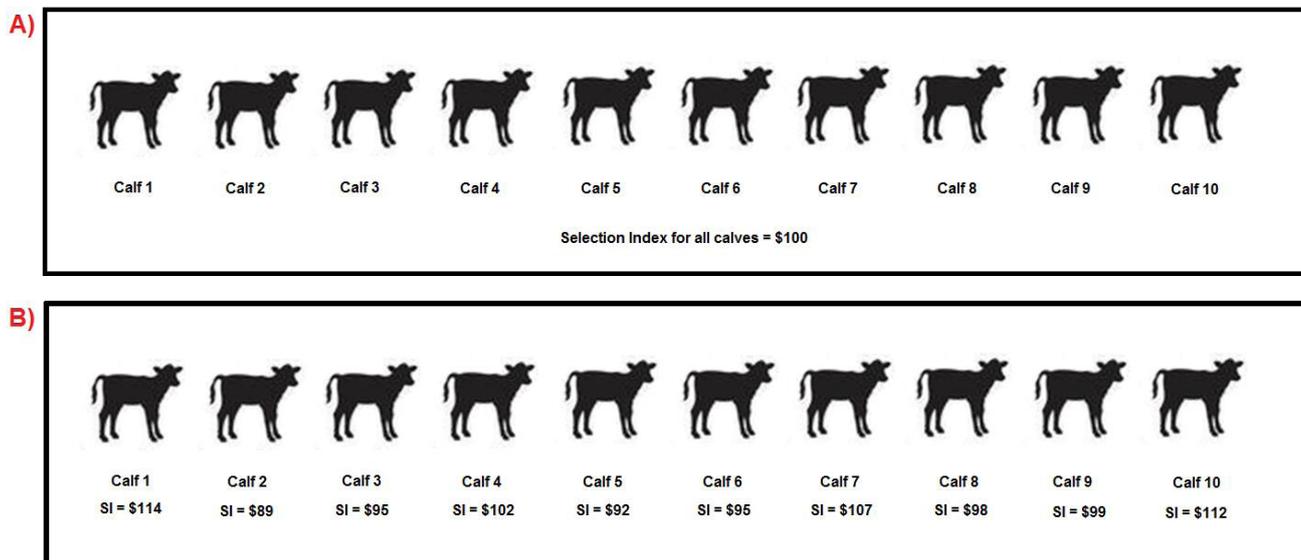
Figure 2. An ET breeding program results in 10 full sibling bull calves. In A) these 10 calves are too young to have performance information recorded, and so have identical mid-parent EBVs and Selection Index values. In this instance, all 10 calves have a Selection Index value of $100. In B) each of the 10 calves has had a genomic test done, and the genomic information has been included in the BREEDPLAN analysis. Here, due to the inclusion of genomic information, the Selection Index values range from $89 to $114.

young bull calves, and thus has no way of differentiating these calves based on their genetic potential.

Figure 2B shows the Selection Indexes generated after the inclusion of genomic information into the BREEDPLAN analysis. As illustrated in Figure 2B, following the inclusion of genomic data into the BREEDPLAN analysis, the Selection Index values for these 10 full-sibling bull calves now range from $89 to $114. As the breeder can now differentiate between calves on their genetic potential, the breeder can identify which of these young full-sibling bull calves has the desired genetics for their breeding program. The breeder can now use the best young bulls for yearling mating. Of course, the use of one Selection Index is for illustrative purposes only; in reality the breeder would need to consider both Selection Indexes and individual EBVs of importance, and do a visual assessment for structure, when making selection decisions.

As the Figure 2 example shows, the power of genomics is the identification of elite bulls and heifers at a young age. This allows cattle breeders to make selection decisions at younger ages than are currently possible, and thus shorten the generation interval. In turn, shortening the generation interval should increase the rate of genetic improvement, both for individual breeders and the breed as a whole.

CONCLUSION

Several breed societies are already incorporating genomic information into their BREEDPLAN analyses, and several more are aiming to move to genomic evaluations in the foreseeable future. The introduction of genomic evaluations will have several benefits to Australian cattle producers, including the potential to calculate EBVs on animals that cannot or have not been measured for particular traits. In addition, the inclusion of genomic data in the BREEDPLAN analysis is likely to increase the accuracy of EBVs for young animals with limited performance data. Therefore, genomics looks set to be an important tool for Australian beef producers who wish to continue making genetic improvement into the future.

For further information on genomics, please contact staff at Southern Beef Technology Services (SBTS) or Tropical Beef Technology Services (TBTS).

Southern Beef Technology Services
Telephone: (02) 6773 3555
Email: office@sbts.une.edu.au

Tropical Beef Technology Services
Telephone: (07) 4927 6066
Email: office@tbts.une.edu.au