



Demystifying **DNA Technology**

A Livestock Breeders Guide
to **Genomic Selection**



Demystifying **DNA Technology**

A Livestock Breeders Guide
to **Genomic Selection**

Acknowledgements

This document has been developed in collaboration between Neogen Australasia, the Animal Genetics and Breeding Unit, Meat & Livestock Australia, Sheep Genetics Australia, and Angus Australia.

The contribution and review of the content included in this document by the following people is gratefully acknowledged.

Mr Andrew Byrne

Senior Technical Product Specialist, Neogen Australasia

Dr Steve Miller

Director, Animal Genetics & Breeding Unit

Dr Andrew Swan

Chief Scientist, Animal Genetics & Breeding Unit

Dr Natalie Connors

Senior Research Fellow, Animal Genetics & Breeding Unit

Ms Peta Bradley

Operations Manager, Sheep Genetics

Mr Christian Duff

General Manager - Genetic Improvement, Angus Australia



CONTENTS

Section 1 – What is DNA?	4
DNA : The Blue Print of Life	4
Common DNA Terminology.....	6
DNA : An Instruction Manual	7
Section 2 – What is Genomic Selection?	8
An Introduction to Genomic Selection	8
Breeding Values 101	11
Section 3 – Genomic Selection : What is the Value Proposition	12
More Accurate Breeding Values	12
Selection Decisions Early in Life	13
Selection for Additional Traits.....	13
Selection for Additional Animals	14
Other Benefits.....	14
What Genomics is Not	15
Section 4 – Genomic Selection : Key Concepts	16
Gene Inheritance – The Random Shuffle of Genes	16
The Polygenic Nature of Traits	19
The Role of a Genomic Reference Population.....	20
Section 5 – Genomic Selection in Seedstock Animals	22
Single Step Genetic Evaluation	22
Analytical Approaches for Calculating Genomically-Enhanced Breeding Values	23
Section 6 – Genomic Selection in Commercial Animals	27
Section 7 – Genomic Selection : Does it Really Work?	28



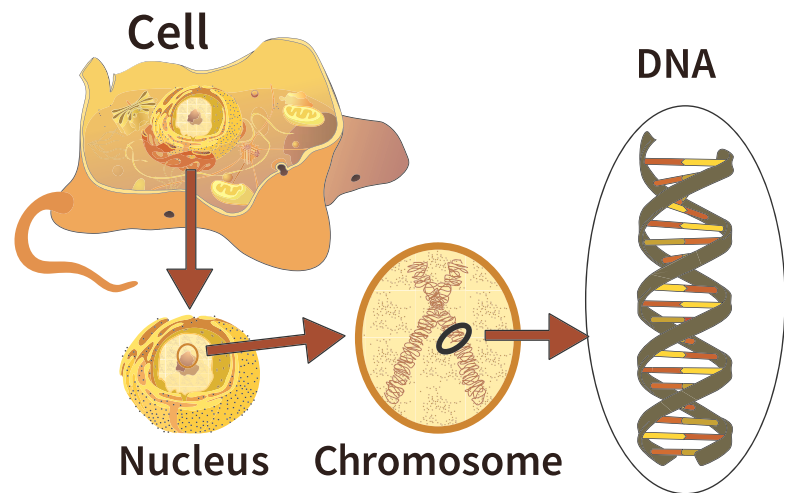
DNA: THE BLUEPRINT OF LIFE

In the world of livestock breeding, success hinges on the selection and mating of animals with superior genetics to produce progeny with desirable traits. But behind these desirable traits, lies a complex and fascinating molecule called 'DNA'.

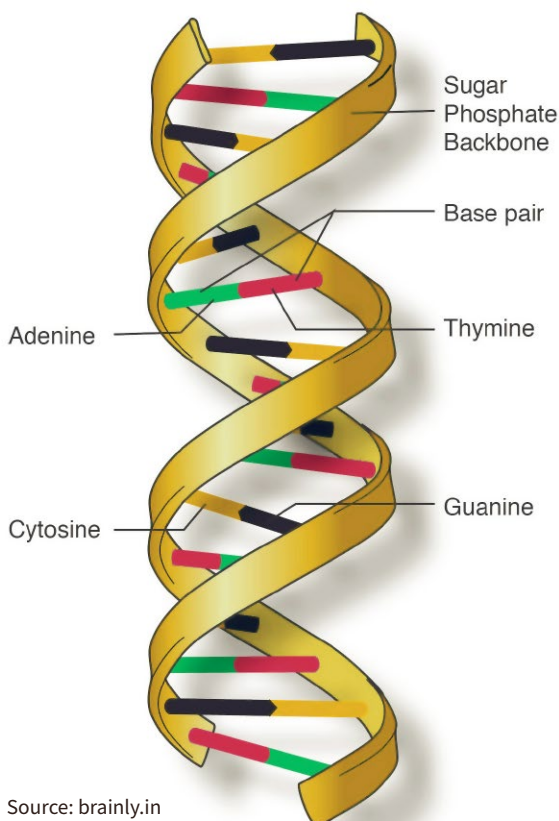
What is DNA?

DNA, or deoxyribonucleic acid, is a molecule present within the cells of all living organisms, including livestock. It serves as a repository of genetic information and is often described as the 'blueprint' of an organism, because it codes the instructions for how that organism will develop, function and appear.

In plants and animals, DNA is located within the cell nucleus, and is arranged, or packaged into bundles known as chromosomes.



Source: microscopemaster.com



Source: brainly.in

Structure of DNA

To understand the function of DNA, one must first understand the structure of the DNA molecule.

DNA is composed of smaller units called nucleotides, which are arranged in a sequence that looks like a long twisted ladder, or double helix.

The sides of the ladder are made up of alternating sugar and phosphate molecules. The rungs connect to the sugar molecules and are known as nitrogenous bases. There are four different nitrogenous bases, Adenine (A), Thymine (T), Guanine (G) and Cytosine (C).

Each rung of DNA is made up of two nitrogenous bases linked together, colloquially referred to as 'base pairs', with guanine (G) bases always linking to cytosine (C), and thymine (T) bases always linking to adenine (A).

The specific order, or sequence, of the base pairs provide the key functionality of DNA, and differences in the base pair sequence provides the basis for variation between individuals.

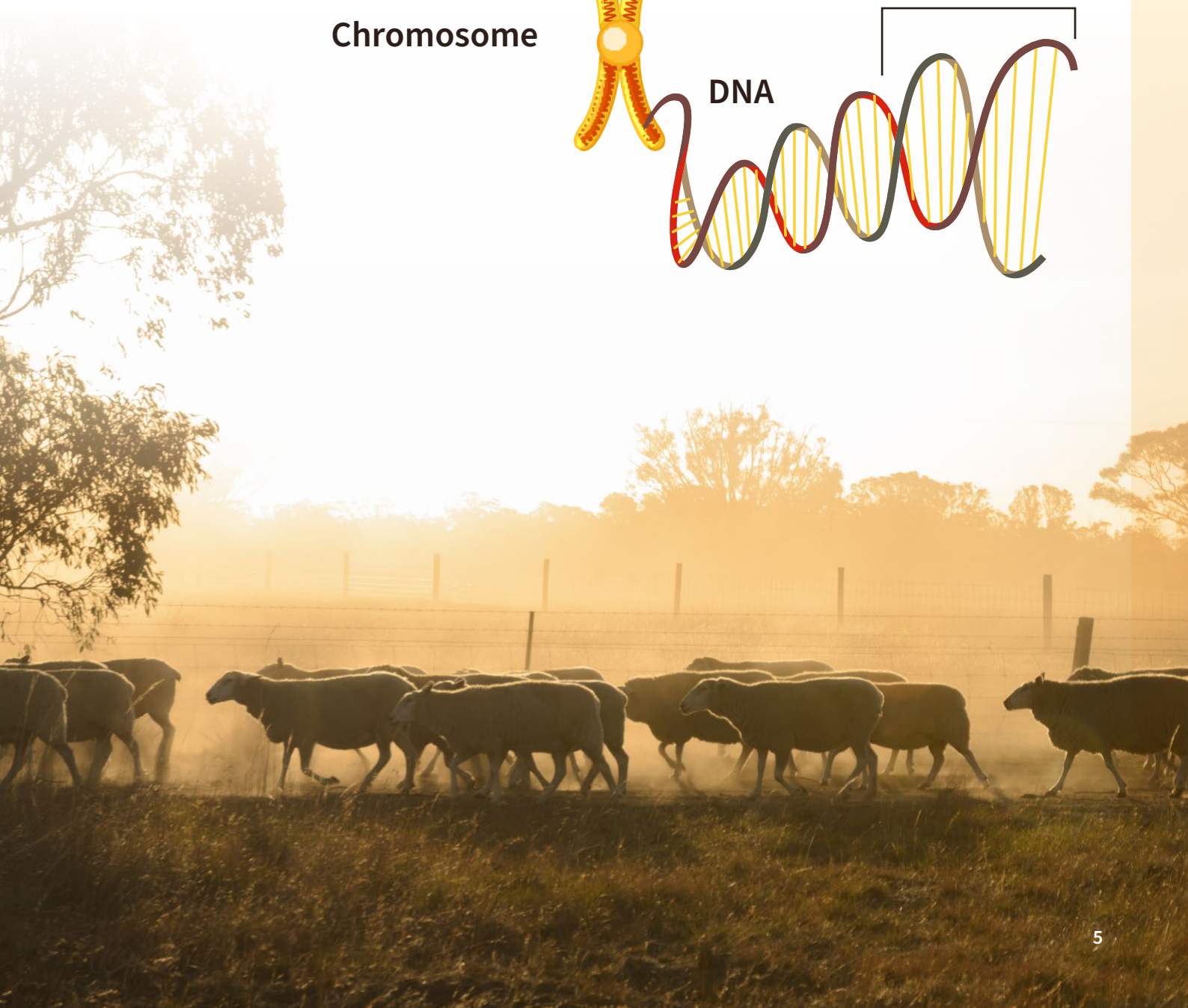
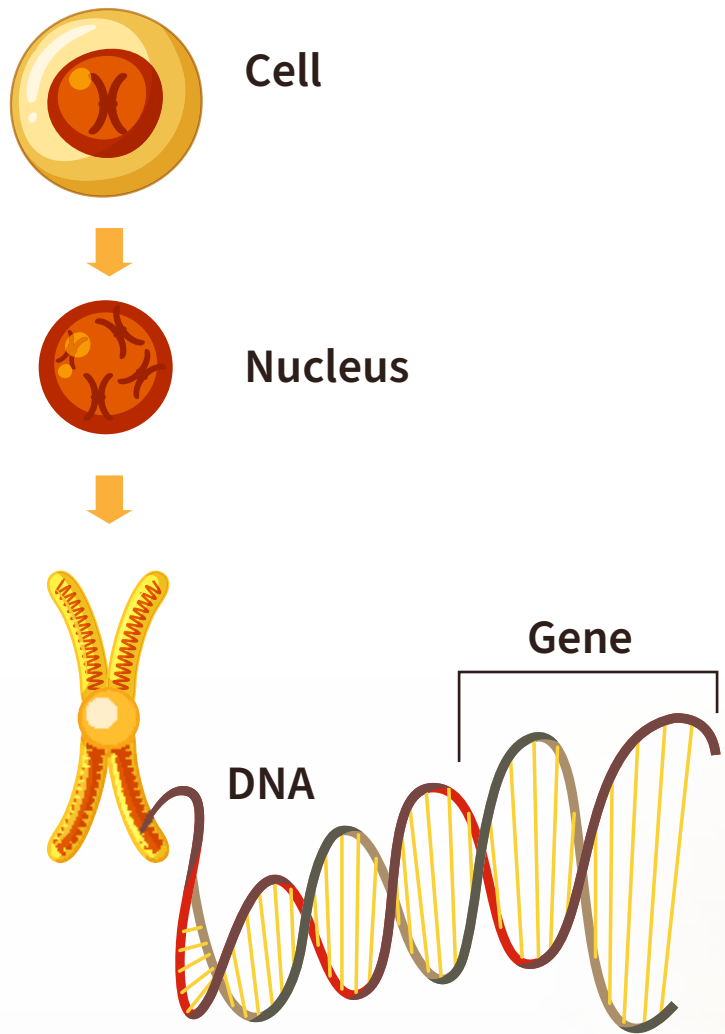
Determining the sequence of base pairs at specific locations within an animal's DNA, and understanding its function, also forms the foundation for the genomic testing of livestock.

What is a Gene?

The DNA of cattle and sheep is estimated to be approximately 2.7 billion base pairs in length, which equates to 2.7 gigabases (Gb), or 2,700 megabases (Mb).

Genes are segments of DNA, or a sequence of base pairs located at a specific location within an animal's DNA, that contain instructions for building and maintaining an organism. Each gene carries information about a specific trait, such as coat colour, polledness, growth rate, fertility and more.

It is estimated that there are between 20,000 to 25,000 genes present in the DNA of cattle and sheep.



COMMON DNA TERMINOLOGY

Term	Definition
Genomics	The study of the structure, function, evolution and mapping of genomes.
Genome	The complete set of genetic material present in an organism. Comprised of DNA.
DNA	The instructions for how an organism will develop, function and appear. Often referred to as an organism's genetic blueprint.
Chromosome	'Bundles' of DNA. That is, the DNA of an organism is arranged, or packaged into different 'bundles'. Each separate bundle is referred to as a chromosome.
Gene	Segments of DNA at a particular location in the genome that contain instructions for building or maintaining an organism. Each gene carries instructions for a different trait.
Locus	The specific location of a gene, or DNA sequence, in the genome.
Allele	Alternative forms of a gene, which can result in differences in phenotype or observable characteristics. There can be two or more alleles for any particular gene. For example, one of the genes influencing coat colour in cattle has three alleles, the 'e' allele for red coat colour, the 'ED' allele for black coat colour and a wild type 'E+' allele. Which alleles the animal has, and how those alleles interact with each other, determines the coat colour of the animal.
Homozygous	The two alleles within an animal's DNA for a gene are identical. . E.g. the animal has two 'ED' alleles.
Heterozygous	The two alleles within an animal's DNA for a gene are different. E.g. the animal has one 'ED' allele, and one 'e' allele.
Dominant	An allele is dominant over another allele, so that the phenotype of the dominant allele is always expressed over the recessive allele.
Recessive	An allele is recessive, so that the phenotype is only expressed when two copies of the allele are present (i.e. it is homozygous).
Carrier	An animal which is heterozygous and 'carries' a recessive allele which is not being expressed phenotypically
Qualitative Trait	A trait that is controlled by a single, or relatively small number of genes, with phenotypes falling into distinct categories. For example, polledness in cattle and sheep.
Quantitative Trait	A trait that is controlled by a large number of genes, with phenotypes being continuous, and often being distributed in a normal, bell curve distribution. Most production traits (e.g. live weight) are considered quantitative traits.
Gene Marker	A gene, or DNA, marker is a known, heritable base pair sequence which is associated with a particular gene. The DNA marker sequence does not necessarily have to be part of the gene that codes for the trait but only associated with it, therefore the value of the genetic marker will vary.
SNP	Single Nucleotide Polymorphism (SNP), pronounced 'snip', is a difference in a single base pair at a specific location within a chromosome. SNPs are often used as gene markers.
SNP Chip	A DNA chip, or 'microarray' that evaluates an animal's DNA at a fixed number of SNPs. E.g. a 50K SNP chip determines the animal's genetic makeup for 50,000 SNPs.

DNA : AN ANIMAL'S INSTRUCTION MANUAL

DNA can be thought of as an animal's instruction manual.

The whole instruction manual represents the animal's genome, being the complete set of genetic material present in the animal, or the complete set of instructions for how the animal will develop, function and appear.

The instruction manual is arranged into chapters, in a similar manner to how an animal's genome is arranged into chromosomes.

Within each chapter, there are paragraphs and words, in a similar manner to how each chromosome is broken into different segments, which form the animal's genes.

And lastly, the paragraphs and words in the instruction manual are comprised of letters, in a similar way to how genes are comprised of base pairs. The DNA 'alphabet' consists of only four letters, being A, C, T & G, with the sequence of the letters, or base pairs, determining the words, or 'instructions', for how the animal will develop, function and appear.

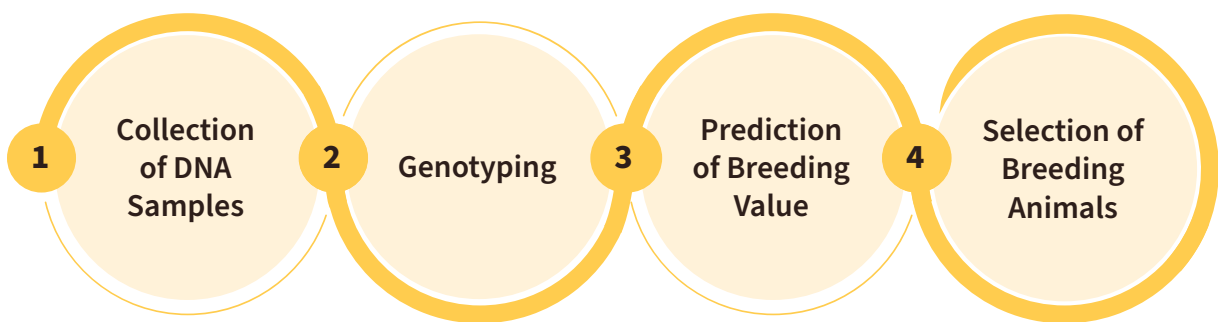


AN INTRODUCTION TO GENOMIC SELECTION

Genomic selection utilises information from an animal's DNA to predict its breeding value for specific traits.

Genomic selection has revolutionised breeding across all livestock species, with genomic selection now utilised in livestock breeding programs all around the globe.

Broadly, genomic selection consists of four key steps:



Step 1 – Collection of DNA Samples

Genomic testing of beef, sheep and dairy animals is most commonly conducted using Tissue Sampling Units, or TSUs.

TSUs are a fast, simple method of collecting DNA.

These units use an ear punch to take a small tissue sample, placing it directly into a sealed vial for processing. This replaces the more complex tasks of collecting hair, blood, semen or other sample types used for DNA testing.

A single-squeeze motion collects a sample in seconds with minimum need for restraint or distress to the animal.

The genetic material is sealed in a specially designed vial to prevent contamination. Tissue samples contain a large quantity of high-quality DNA for genetic analysis, yielding excellent lab results, maximising sample integrity and ensuring you have confidence in the results you receive for each animal.

The TSU is fully sealed and positively identified with both a 2D barcode and ID panel.

Other sample types, such as semen straws, blood cards or tail hair cards, can also be collected, however these sample types require additional steps when processing them within the DNA laboratory, and are not the preferred sample type in most circumstances.



Step 2 – Genotyping

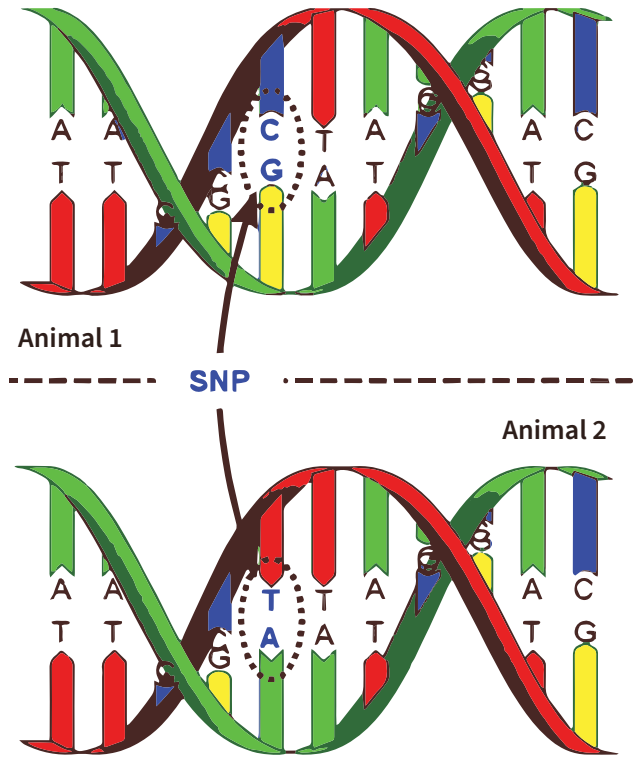
Once DNA samples have been collected, they are sent to the DNA laboratory for genotyping.

Genotyping involves extracting the animal’s DNA from the sample that has been collected, preparing the DNA for analysis, and ‘reading’ the DNA to identify the base pairs that exist at targeted locations within the animal’s DNA.

Most genotyping for genomic selection in livestock is based on single nucleotide polymorphisms, or SNPs (pronounced ‘snips’).

SNPs occur when there is a difference in a single base pair at a specific location with the genome of animals. For example, in the diagram, animal 1 has a C/G base pair at the specified location in its DNA, whereas animal 2 has a T/A base pair at the equivalent location in its DNA.

When genotyping animals for genomic selection, the animal’s DNA is ‘read’ at thousands of different SNPs, with the exact number of SNPs differing depending on the genotyping platform, or ‘SNP chip’, that is used. For example, a SNP chip might ‘read’ the animals DNA at 20,000 SNPs (20K), 50,000 SNPs (50K) or 100,000 SNPs (100K).



SNP Name	Allele 1	Allele 2
ARS-BFGL-BAC-19454	A	G
ARS-BFGL-BAC-27364	A	C
ARS-BFGL-BAC-29509	T	T
ARS-BFGL-BAC-29664	T	C
ARS-BFGL-BAC-31822	A	G
ARS-BFGL-BAC-35552	A	A
ARS-BFGL-NGS-10035	T	C
ARS-BFGL-NGS-100821	T	C
ARS-BFGL-NGS-101394	A	A
ARS-BFGL-NGS-101456	T	C
ARS-BFGL-NGS-102169	A	C
ARS-BFGL-NGS-102869	C	C
ARS-BFGL-NGS-102903	T	C
ARS-BFGL-NGS-103807	C	C
ARS-BFGL-NGS-104172	A	C
ARS-BFGL-NGS-106015	A	G
ARS-BFGL-NGS-108997	T	T
ARS-BFGL-NGS-110327	G	G
ARS-BFGL-NGS-110793	T	T
ARS-BFGL-NGS-111053	T	C
ARS-BFGL-NGS-111076	T	C
ARS-BFGL-NGS-111114	T	G
ARS-BFGL-NGS-112094	T	T
ARS-BFGL-NGS-112325	T	T
ARS-BFGL-NGS-112670	A	A

An example of the genomic profile that is created for an animal during genotyping, showing the animal’s genotype is made up of a series of A’s, C’s, T’s and G’s, across the different SNPs that have been read. In practice, the animal’s genomic profile will consist of many, many more rows than what has been depicted in the diagram. For example, the genomic profile for an animal genotyped with a SNP chip containing 100,000 SNPs will have 100,000 rows.

Step 3 – Prediction of Breeding Value

The third step in genomic selection involves analysis of the genomic profile that has been created for an animal to predict the animal's breeding value for individual traits.

In industry genetic evaluations, such as those conducted by BREEDPLAN, Sheep Genetics and DataGene, this step involves analysing the genomic profiles that are available, along with other sources of information, such as pedigree and phenotypic information, to simultaneously calculate a genomically enhanced breeding value.

A different, two-step analysis approach is used in tools such as Angus HeiferSELECT, Wagyu Feeder Check and Igenity Beef, where a genomic prediction equation is firstly developed, and then applied to an animal's genomic profile, to calculate a genomic breeding value.

A more detailed explanation of how the genomic profile for an animal is used to predict their breeding value is outlined in later sections of this document.

Step 4 – Selection of Breeding Animals

The last, and most important step in genomic selection is the use of the breeding values that have been calculated, in combination with other information that is available, to select the animals that will be retained for use in the breeding program.

Use of the breeding values that have been calculated to identify and select the animals that are most aligned with the breeding program's objective and goals will maximise the genetic gain that is achieved from genomic selection.



BREEDING VALUES 101

To take advantage of genomic selection, one must first understand the benefits of incorporating breeding values into the selection of breeding animals.

The key to successful animal breeding is the accurate description of the genetic differences that exist between animals, and the use of the genetic variation that exists to breed the next generation.

The problem is however that it can be very difficult to 'see' the genetic differences that exist between animals.

- Only a component of the differences in the phenotype, or the performance of animals are due to genetics. What you see is not necessarily what you get.
- Some traits are difficult to 'see' when assessing animals visually.
- Some traits are difficult to measure.
- Some traits are sex limited.
- Some traits are expressed later in life and can't be measured prior to selection.
- It can be difficult to compare the phenotype, or performance of animals that are being managed in different mobs, enterprises or environments.

This is where the incorporation of breeding values in the selection of animals for mating in a breeding program provides value.

Breeding values consider all available information about an animal, such as its performance, and the performance of animals to which it is related, to predict the genetic merit of the animal.

Importantly, breeding values 'belong to an animal's progeny', meaning the breeding value of an animal predicts the genetics that it will pass onto its progeny, rather than providing a description of the animal's phenotype or performance. This is important as the objective when making selection decisions is not to select the most superior animal, but to select the animal that will have the most superior progeny.

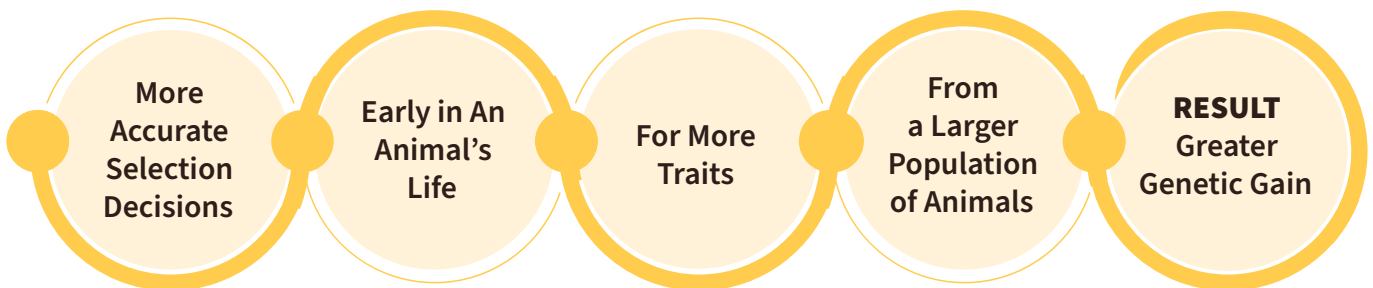
Breeding values in Australia are commonly published as Estimated Breeding Values (EBVs) for beef cattle, as Australian Sheep Breeding Values (ASBVs) for sheep, and as Australian Breeding Values (ABVs) for dairy cattle. In some beef cattle, breeding values may be published as Estimated Progeny Differences (EPDs).

The incorporation of breeding values in the selection of breeding animals, in association with other sources of information, facilitates more accurate selection decisions, and ultimately faster genetic improvement to be made in a livestock breeding program.

GENOMIC SELECTION : WHAT IS THE **VALUE** **PROPOSITION?**

Genomic selection offers wide ranging benefits in livestock breeding programs.

By comparison to traditional selection methods, genomic selection facilitates more accurate selection decisions, earlier in an animal's life, and for additional traits and animals.

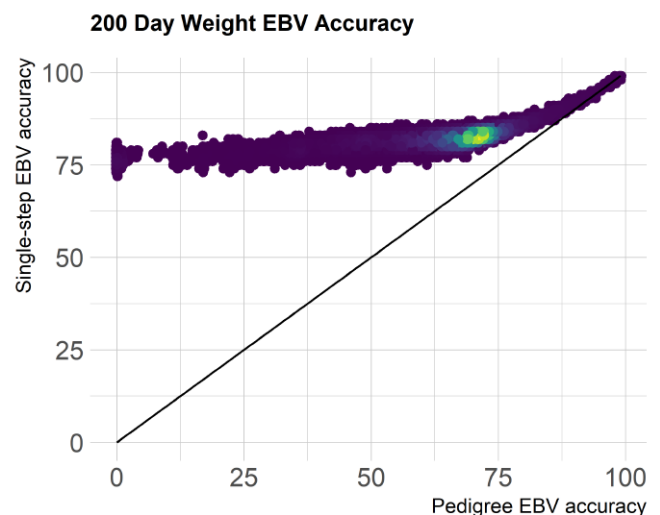


More Accurate Breeding Values

Genomic selection facilitates the prediction of breeding values with greater accuracy, which subsequently results in more accurate selection of breeding animals, and faster genetic improvement.

While considerable additional accuracy is observed across all traits, the most additional accuracy is realised for:

- for animals that have incomplete, or limited pedigree.
- for traits that are hard to measure, or traits that cannot be measured prior to an animal entering the breeding herd or flock.
- for traits that have a low heritability.
- in situations where collecting effective performance information is problematic, such as in small herds or flocks, or when an animal has been removed from its contemporary group.

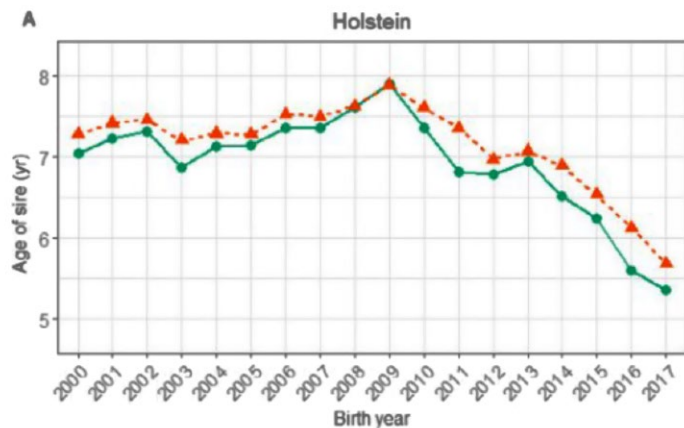


Accuracy of breeding values for 200 Day Weight in Australian Angus cattle with (blue) and without (diagonal) genomic information.

Selection Decisions Early in Life

Genomic selection facilitates the prediction of accurate breeding values early in an animal's life, which enables the mating of young animals with greater confidence, such as the use of ram lambs, yearling bulls or unproven sires, resulting in a shorter generation length, and faster genetic improvement.

The availability of accurate breeding values early in an animal's life also provides greater flexibility when making management decisions, such as the ability to make culling decisions well before animals reach breeding age.



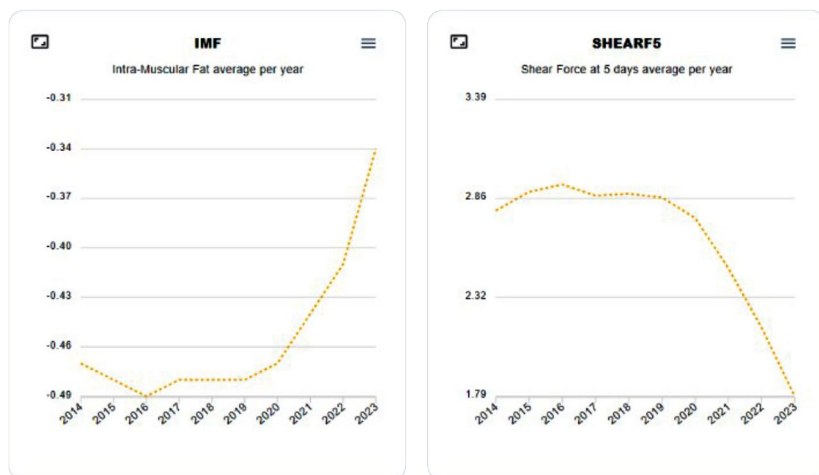
Generation interval in Holstein breeding programs in Australia has decreased dramatically with the introduction of genomic selection (Source: Scott et al, Journal of Dairy Science, 2021)

Selection for Additional Traits

Genomic selection facilitates a greater number of traits to be considered when making selection decisions by making selection possible for traits that can't be routinely, or easily measured on farm.

Examples include traits associated with the sustainability of a beef enterprise, such as animal resilience and methane emissions, or selection for intramuscular fat and meat tenderness in sheep breeding programs.

The consideration of additional traits makes selection possible for a greater proportion of a breeding program's overall breeding objective, resulting in faster genetic improvement.



Genomic selection has enabled selection for improved eating quality in the Australian sheep industry (Source: Sheep Genetics)

Selection for Additional Animals

By comparison to traditional genetic evaluation programs that required pedigree information for the calculation of breeding values, genomic selection enables the calculation of breeding values for animals in the absence of pedigree information.

This enables:

- additional populations of animals to be considered when making selection decisions.
- more accurate selection decisions to be made in commercial herds or flocks.
For example, when selecting female replacements, or males in multiplier breeding programs.
- drafting of commercial animals into production systems aligned with their genetics.
For example, the drafting of steers into short or long fed feedlot finishing programs.

Other Benefits

In addition to facilitating more accurate selection decisions, earlier in an animal's life, and for additional traits and animals, genomic selection also offers a number of other benefits in a livestock breeding program.

- **Genetic Benchmarking of Livestock Enterprises**

Genomic selection makes it possible to benchmark the genetics in a livestock enterprise across enterprises and over time, providing benefits such as more targeted ram and bull purchase decisions in commercial breeding programs, and the tracking of the genetic change that is occurring over time within a commercial breeding program.

- **Less Reliance on Performance Measurements**

Whilst in most situations the accuracy of breeding values will be maximised when both genomic information and effective performance information is available, subject to the availability of well recorded, genomic reference population, genomic selection makes it possible to achieve useful breeding value accuracy from genomic information alone.

This prevents the need for the on-farm collection of performance measurements in some situations, such as in smaller breeding enterprises, for difficult or expensive to measure traits, when animals are managed on agistment or lease blocks with limited facilities or labour, or where contemporary groups need to be dispersed for management purposes, such as in times of drought, flood or fire.

- **More Accurate Animal Pedigrees**

Genomic selection improves the accuracy of animal parentage information, reduces the effect of historic pedigree errors on the calculation of breeding values, facilitates better management of inbreeding, and increases the reliability of the information presented on sale animals.

- **Parentage Assignment**

Genomic selection makes it possible to assign parentage information from an animal's DNA profile, offering advantages such as the ability to utilise syndicate joining in seedstock operations, shortening the duration of calving and lambing periods in breeding operations utilising artificial insemination, and reducing the on-farm labour requirements associated with the collection of parentage information using traditional methods.

- **Marketing of Seedstock Animals**

Genomic selection provides a number of advantages to the marketing of sale animals in a seedstock operation. Utilisation of genomic selection not only demonstrates a seedstock operation is ‘progressive’, which may be appealing to some bull or ram buying clients, but it reduces the risk associated with post-sale changes in the breeding values of sale animals, and removes barriers to purchase by buyers who are reluctant to purchase animals without genomic information.

- **Compatibility with Add-On DNA Tests**

Genomic selection is compatible with other forms of DNA testing that can be conducted in association with genomic selection. Add-on DNA testing can be used for the management of important genetic conditions, to fast track the introgression of favourable qualitative traits into a breeding herd, such as polledness, and to meet the requirements for the registration of animals with breed associations.

What Genomics is Not?

While genomic selection offers many advantages, it is equally important to understand what genomic selection doesn’t provide to a livestock breeding program.

A common misconception is that genomic selection directly assesses the genes that control productivity traits in livestock breeding program, whereas in reality, the genes have largely not been identified. Rather than assessing the genes of an animal directly, genomic selection is predominantly based on gene markers, which are used to predict an animal’s genetic superiority or inferiority.

Likewise, genomic selection doesn’t replace the need for breeding values. Whilst some livestock breeders may have hoped that genomic technology would make breeding values redundant, the reality is the opposite. Genomic selection utilises the information from an animal’s DNA to predict its breeding value, rather than replacing the need for breeding values.

Another common misconception is that genomic selection replaces the need for the collection of performance measurements. The availability of performance measurements is a critical requirement of genomic selection, with performance measurements essential for the establishment and maintenance of a genomic reference population. Further, in most situations, the accuracy of breeding values will be maximised when both genomic information and effective performance measurements are included in the calculation of the breeding value.

In seedstock operations, genomic selection also doesn’t replace the need for the collection of important information that is required for the genetic evaluation of performance measurements, or for the registration of animals, such as birth date information.



GENE INHERITANCE : THE RANDOM SHUFFLE OF GENES

To understand the value of genomic selection, one must first understand how genes are inherited from one generation to the next.

An animal's genes are contained within its DNA, which is bundled into chromosomes. Livestock have pairs of chromosomes, with the number of chromosome pairs in an animal's DNA varying across the different species. For example, cattle have 30 pairs of chromosomes, or a total of 60 chromosomes. One chromosome in each chromosome pair is inherited from the animal's sire, and one from the animal's dam.

Species	Chromosome Pairs
Human	23 pairs
Cattle	30 pairs
Sheep	27 pairs
Goats	30 pairs
Pigs	19 pairs

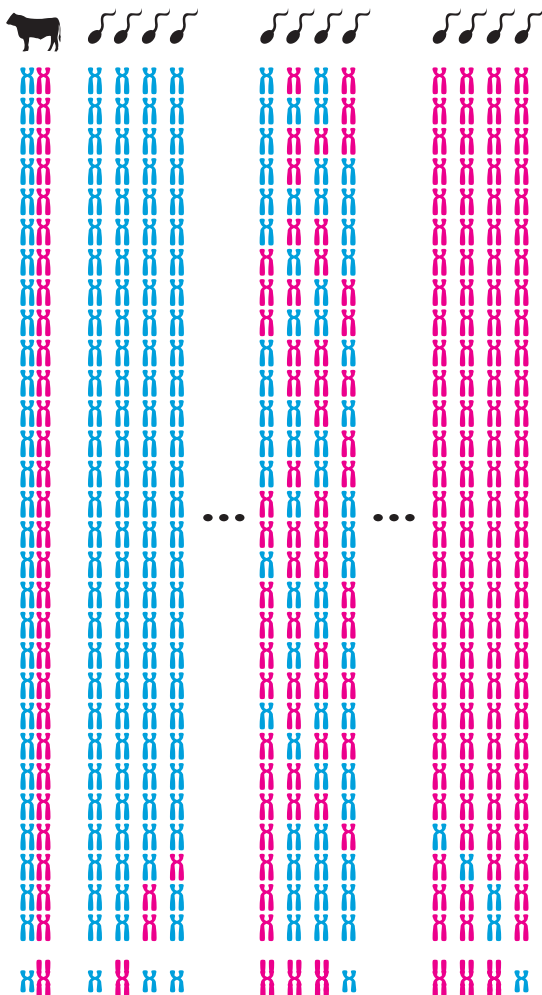


Figure 1. Illustration of the shuffling of chromosomes that occurs during sperm formation. The first column represents the bull's two sets of chromosomes. Chromosomes inherited from the bull's sire are in blue. Chromosomes from the bull's dam are in pink. The other columns depict possible combinations of paternal and maternal chromosomes in individual sperm cells. There are more than 1 billion possible combinations.

When the animal reproduces, for example, a sperm is produced by a sire, one of the chromosomes in each pair is selected at random and passed on to the progeny. So, when a sperm is produced, the selection of chromosomes that are passed on is similar to flipping a coin numerous times (i.e. 30 times in cattle, 27 times in sheep), with a different combination of heads and tails returned each time a new sperm is produced.

If we label the chromosomes the sire inherited from his father as blue (paternal) and the chromosomes inherited from his mother as pink (maternal), the number of possible combinations of the sire's paternal and maternal chromosomes that can be passed onto the progeny is illustrated in Figure 1.

Delving into this in more detail, the number of possible chromosome combinations that can be passed on in cattle is 1,073,741,824 (i.e. 2^{30}), while in sheep the number of possible chromosome combinations is 134,217,728 (i.e. 2^{27}). And this number ignores the exchange of segments between paternal and maternal chromosomes that occurs in a biological process called recombination.

The manner in which chromosomes are passed from one generation to the next is often described as the progeny receiving a random sample of the sire's genes, or mendelian sampling, and with the number of possible combinations, the genes passed on in no two sperm are exactly alike. The same number of possible combinations is true for each egg produced by a dam, meaning the number of possible chromosome combinations inherited by animals from their parents is literally in the billions.

Think for a moment about your favourite set of full siblings, being brothers or sisters with the same parents.

Perhaps this is your favourite set of embryo flush mates, your brothers and sisters, or your children. The dissimilarity between full siblings may be striking, for example, one may be short and the other tall, one may have light hair and the other dark hair, or one may be laid-back and the other excitable.

The similarities and dissimilarities between full siblings are due to differences in the chromosomes that they have inherited, and the environment they have experienced.

So what does this mean for genomic selection?

In their simplistic form, the breeding value calculated for an animal can be thought of as one half of the sire's breeding value, plus one half of the dam's breeding value, plus the mendelian sampling term. The mendelian sampling term represents the animal's difference from the average of its parents' breeding values and is due to the random sample of genes and chromosomes that the calf inherited.

$$\mathbf{EBV}_{\text{CALF}} = \frac{1}{2} \mathbf{EBV}_{\text{SIRE}} + \frac{1}{2} \mathbf{EBV}_{\text{DAM}} + \mathbf{mendelian\ sampling\ term}$$

In a traditional, pedigree based genetic evaluation, there is no data available when an animal is born, so the Mendelian sampling term can not be estimated, and the animal's EBV is reported as the parent average. This is often referred to as the 'mid-parent' EBV value.

As performance data is subsequently collected on the animal, and the animal's progeny, the mendelian sampling term can be estimated, and the animal's EBV either increases, decreases or remains unchanged, reflecting the random sample of genes that the animal has inherited from its parents.

But this approach has limitations and it can be problematic collecting effective performance information on the animal to enable the mendelian sampling term to be estimated. For example, traits that are difficult or expensive to measure, traits that can't be measured prior to the animal entering the breeding herd, traits that have a low heritability, or animals that are in small contemporary groups.

This is where genomic selection now provides considerable value.

In a genetic evaluation that incorporates genomic information, such as the industry single step genetic evaluations that are conducted in the beef, dairy and sheep industries in Australia and New Zealand (e.g. BREEDPLAN, DataGene, Sheep Genetics), genomic data provides an insight into the random sample of chromosomes that an animal has inherited, which enables an estimate to be made of the Mendelian sampling term early in an animal's life.

Knowledge of the chromosomes that an animal has inherited via genomic testing, and the resultant ability to better estimate the mendelian sampling term, results in the availability of more accurate and reliable breeding values for animals.





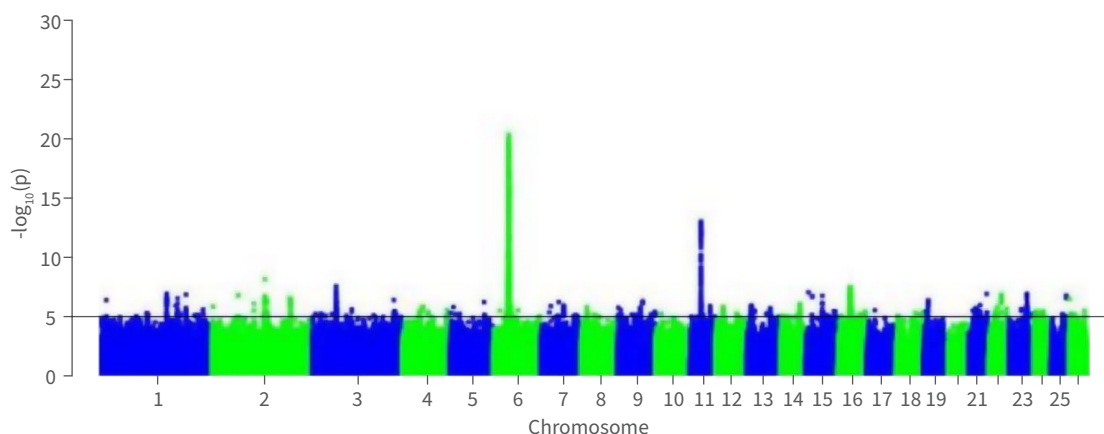
THE POLYGENIC NATURE OF PRODUCTION TRAITS

An important concept in understanding the value of genomic selection is that most traits that influence productivity in livestock enterprises are ‘polygenic’ traits.

Polygenic traits, for example growth rate in beef cattle and sheep, or milk production in dairy cattle, are controlled by many genes, with each individual gene having only a relatively small effect on the trait.

Polygenic traits differ to qualitative traits, like polledness or coat colour, which are controlled by a single, or small number of genes, which individually have a large effect on the trait.

The polygenic nature of productivity traits is illustrated in the plot below which demonstrates the level of significance between different locations across the sheep genome, and dressed carcass weight in Australian sheep. The plot shows that there are genes of significance across most, if not all chromosomes, with genetic differences in dressed carcass weight being the result of the cumulative actions of many genes.



A plot illustrating that there are genes across most, if not all chromosomes influencing dressed carcass weight in sheep. (Reference: Moghaddar et al, “Multi-trait genome wide association meta-analysis of body weight, carcass composition and eating quality traits in Australian sheep.”, AAABG, 2023.)

The polygenic nature of production traits means that the genes controlling productivity traits in livestock breeding programs have largely not been identified, which has consequences for genomic selection. Rather than assessing the genes of an animal directly, genomic selection is predominantly based on gene markers.

A gene, or DNA marker, is a base pair sequence that is associated with a difference in the phenotype or observable characteristics of animals. The gene marker may not necessarily be part of the gene that codes for a particular trait but only associated with it.

The most common form of gene marker used for genomic selection in livestock are single nucleotide polymorphisms, or SNPs (pronounced ‘snips’). Genomic selection involves analysing the SNPs in the genomic profile of an animal, alongside other sources of information (e.g. phenotypes), to predict an animal’s breeding value.

THE ROLE OF A GENOMIC REFERENCE POPULATION

Genomic selection, and the ability to utilise an animal's genomic profile to predict its breeding value, is dependent on the existence of a well recorded, relevant genomic reference population.

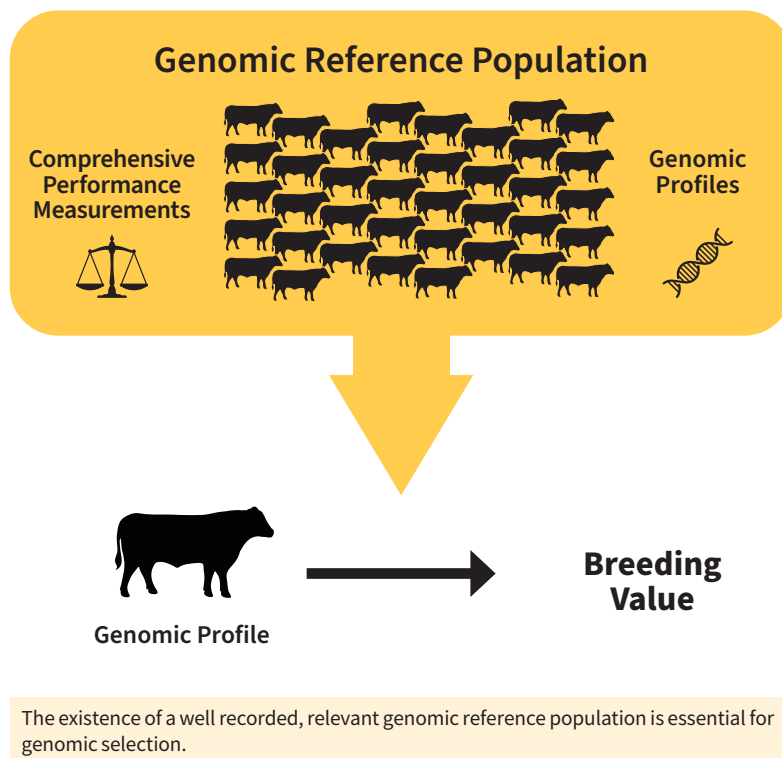
A genomic reference population refers to a large group of animals that have both genotypes (i.e. genomic profiles) and phenotypes (i.e. performance measurements) available.

In industry genetic evaluations, such as those conducted by BREEDPLAN, Sheep Genetics and DataGene, the genomic reference population includes animals, and their associated genotypes and phenotypes, from industry funded, structured reference population programs, and animals in individual breeder herds or flocks.

The accuracy provided by genomics to the prediction of an animal's breeding value is determined by the size and composition of the genomic reference population that is available.

Specifically, the accuracy provided by genomics is influenced by:

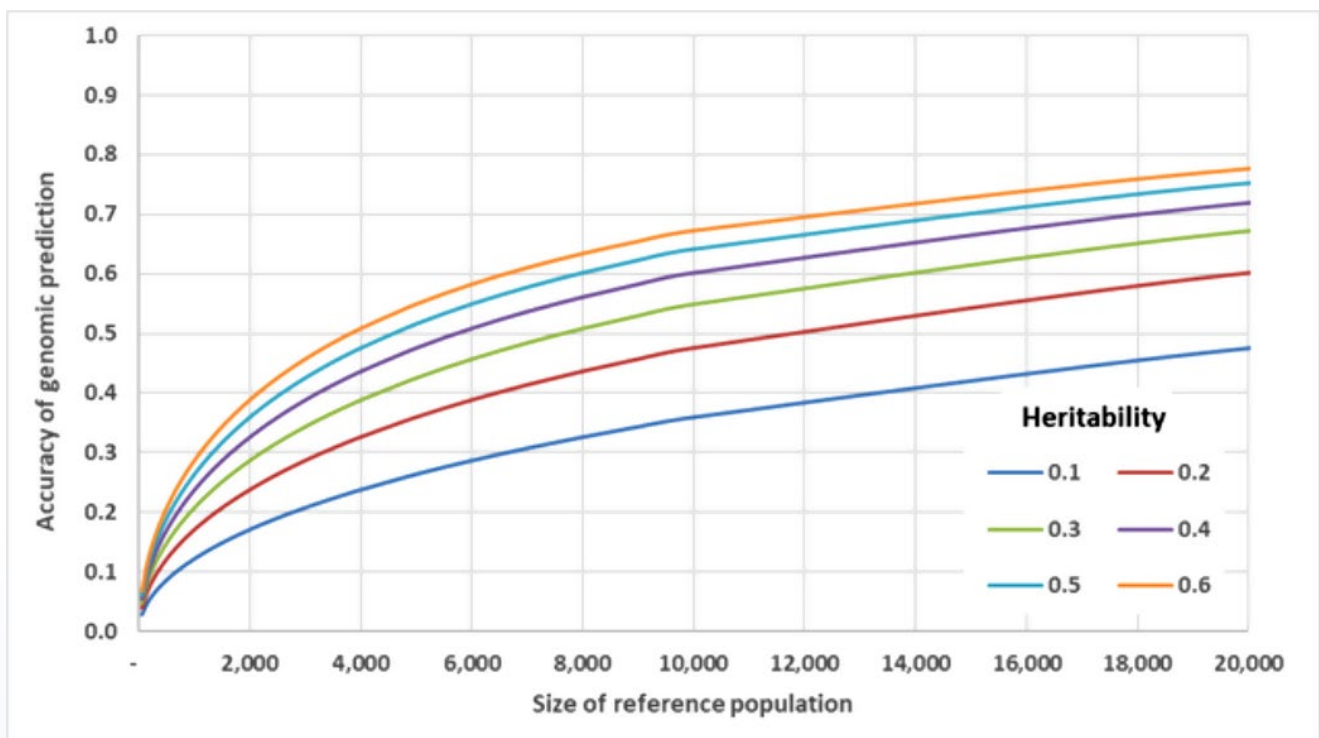
- the size of the genomic reference population. i.e. how many genotyped animals have phenotypes for a particular trait.
- the heritability of the trait. i.e. a larger number of genotyped animals with phenotypes is required for traits that have a lower heritability.
- the effective population size. i.e. the more genetically diverse a population is, the larger the genomic reference population that is required.
- the genetic relationship between the animal for which the breeding value is being calculated and the animals in the genomic reference population. i.e. the more closely related an animal is to the animals in the genomic reference population, the more value genomics provides. Conversely, if an animal is not related to animals in the genomic reference population, genomics provides minimal value, and can potentially lead to the calculation of biased breeding values.
- The relevance of the production system in which the phenotypes were collected.



In addition, the usefulness of a genomic reference population decreases over time with changes in the genetic makeup of the breeding population, so it is important that the reference population is routinely updated with new animals.

Importantly, genomics provides no value to the breeding value calculated for an animal in the absence of a well recorded, relevant genomic reference population.

For this reason, utilising genomic selection can be challenging, particularly in breeds with a smaller population of breeding animals, as there may not be enough animals with both phenotypes and genotypes available to form an effective genomic reference population.



The accuracy of genomic selection as the size of the genomic reference population increases for traits with varying heritability.



GENOMIC SELECTION IN SEEDSTOCK ANIMALS

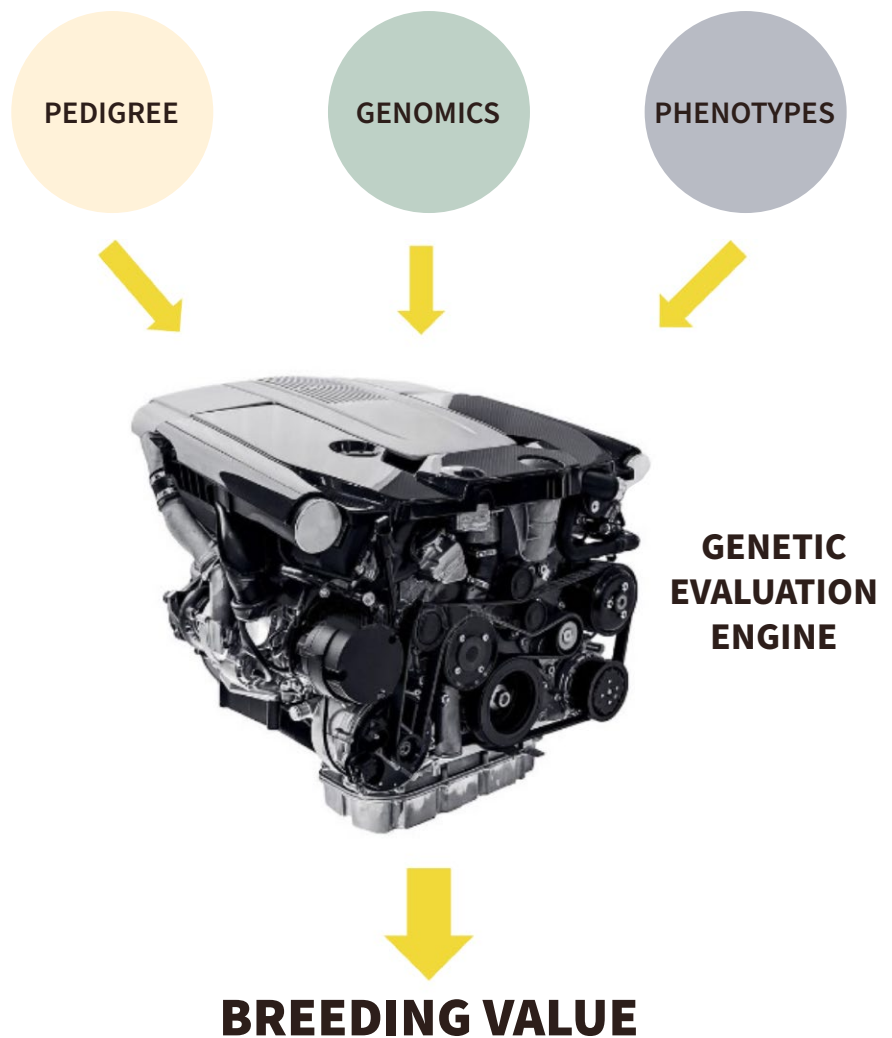
Genomic selection in seedstock animals most commonly involves the inclusion of genomic profiles in existing industry genetic evaluations, such as those conducted by BREEDPLAN, Sheep Genetics and DataGene, to calculate genomically enhanced breeding values.

Single Step Genetic Evaluation

The manner in which genomic information is included in industry genetic evaluations is often referred to as a 'single-step' genetic evaluation, reflecting that genomic profiles are analysed simultaneously with other sources of information (e.g. pedigree, phenotypes), placing appropriate emphasis on all available sources of information, to calculate a single, genomically enhanced breeding value.

A unique feature of single-step genetic evaluation is that breeding values can be calculated for both animals with and animals without genomic profiles. This differs to some earlier approaches that could only handle individuals with genomic profiles.

Likewise, in contrast to some earlier approaches, a 'genomic breeding value' is not calculated in a single step genetic evaluation, only the overall, genomically enhanced breeding value that is published.



In industry genetic evaluations, such as those conducted by BREEDPLAN, Sheep Genetics and DataGene, genomic profiles are analysed simultaneously with other sources of information (e.g. pedigree, phenotypes) to calculate a single, genomically enhanced breeding value.

Analytical Approaches for Incorporating Genomic Information

There are a number of different analytical approaches for incorporating genomic information in the calculation of breeding values within a single-step genetic evaluation.

Two commonly adopted approaches include the 'relationship' approach, and the 'marker effects' approach.

While the analytical method behind both analytical approaches differs, research conducted by the scientific community would indicate that when the amount of information included in the genetic evaluation grows in size (i.e. the number of animals, genomic profiles and phenotypes), the breeding values that are predicted using either approach are very similar.

The 'Relationship' Approach

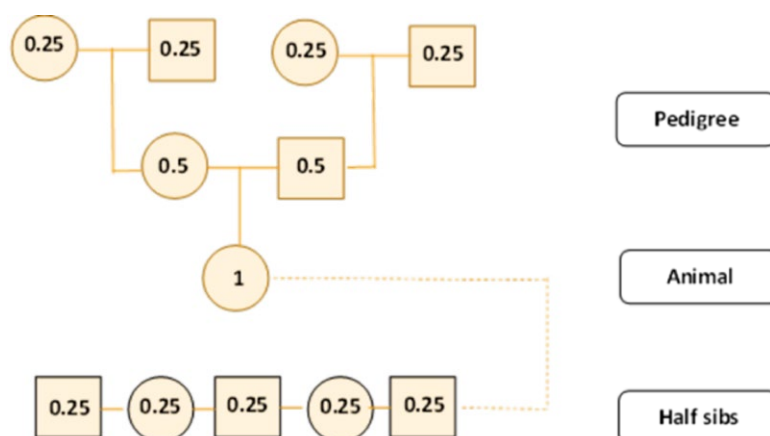
In the relationship approach, such as the TBLUP model that is used by Sheep Genetics & BREEDPLAN, or the APY model that is used in the World Angus Evaluation, the incorporation of genomics is about 'animal relationships'.

In a simplistic form, a breeding value is calculated by a) estimating the relationship between an animal and the other animals in the genetic evaluation, and b) using the performance information that has been collected for the animal and the animals to which it is related to predict an animal's breeding value.

The more closely an animal with a phenotype is related to the target animal whose breeding value is of interest, the more influence that its performance has on the breeding value being calculated.

In traditional, pedigree based genetic evaluation, the relationships between animals are determined by pedigree alone.

In a genetic evaluation that incorporates genomics using the 'relationship' approach, the relationships between animals are determined by combining pedigree and genomic information.

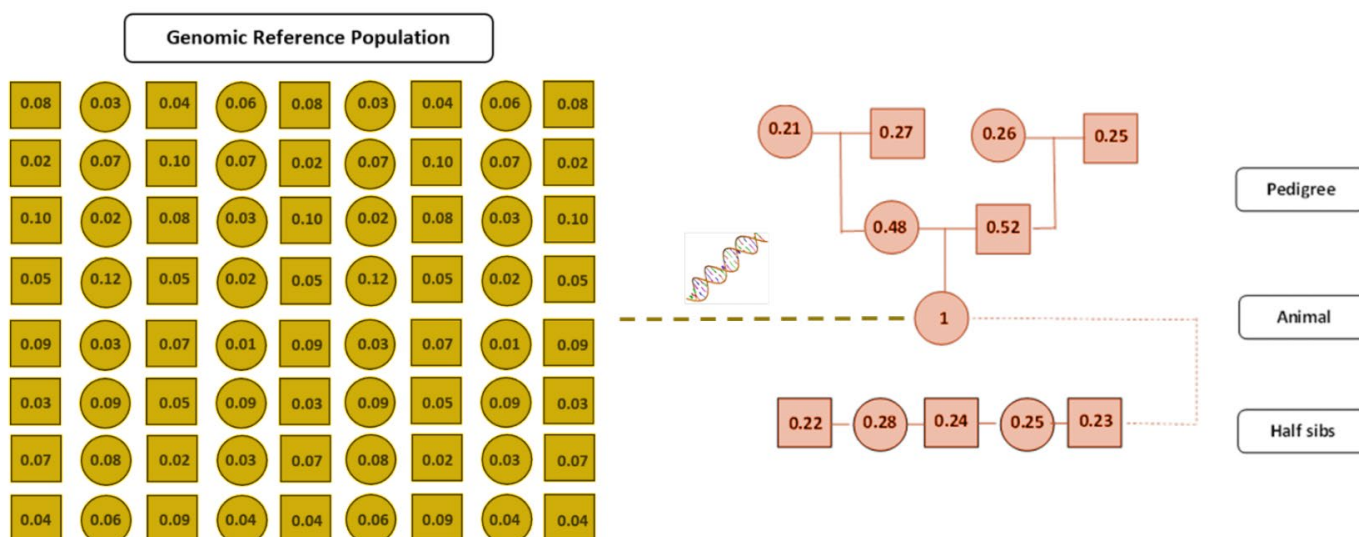


Example of relationships between animals in a traditional genetic evaluation. Based on pedigree relationships, the animal is estimated to have 50% of genes in common with each parent, 25% of genes in common with each grand-parent, and 25% of genes in common with each of its half siblings.

Utilisation of genomic relationships, being the genetic relatedness between two animals based on their genomic profiles, improves estimation of the relationship between animals in the genetic evaluation, and ultimately results in the calculation of more reliable breeding values.

Genomic relationships enable:

- the relationship between animals to be estimated when no, or incomplete pedigree information is available;
- a more precise estimate to be made of the genetic relationship between two animals who are related by pedigree. i.e. rather than assuming an animal has 25 percent of its genes in common with each grand parent (based on the law of averages), genomics assesses the random shuffle of chromosomes the animal has inherited to provide a more precise indication of the proportion of genes that an animal has in common with each relative;
- the relationship between animals unrelated by pedigree to be determined, and any performance information that is available on these animals to be used in the calculation of the animal's breeding value. e.g. pedigree may indicate that two animals are unrelated, but in reality, genomics indicates that they share a proportion of genes in common. This is particularly advantageous as it allows the performance information of animals in the genomic reference population to be utilised in the calculation of an animal's breeding value.



Example of relationships between animals in a single step genetic evaluation. Utilisation of genomic relationships enables a more precise estimate to be made of the relationship to each animal in the pedigree, while also enabling the relationship to be estimated to animals in the genomic reference population.

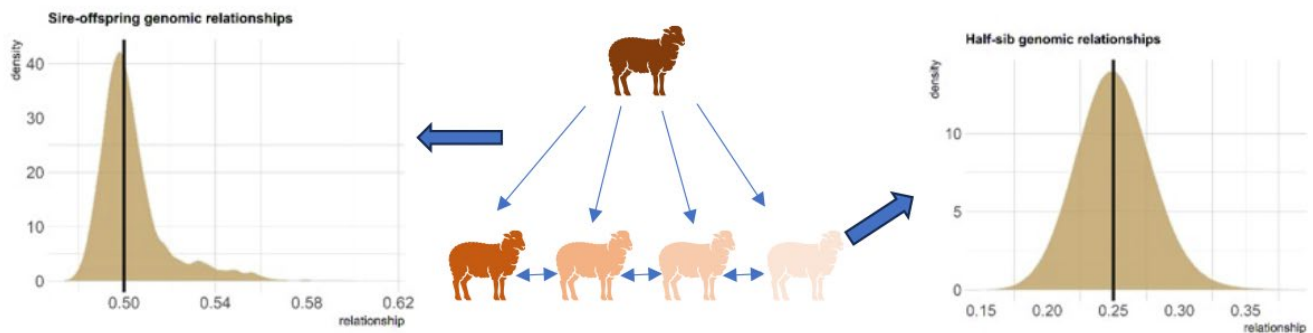


Genomic Relationships in Practice

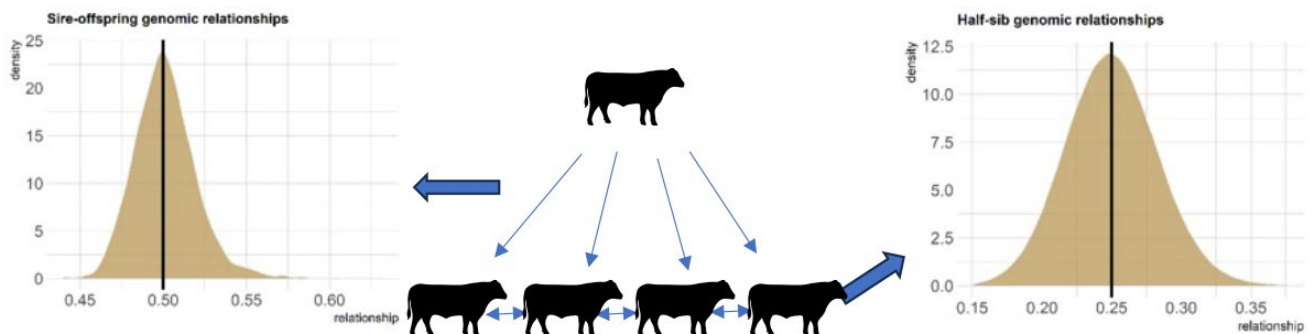
The real-world examples below demonstrate how utilisation of genomic relationships improves the estimation of the relationship between animals in a single step genetic evaluation.

Based on pedigree information, the sires would be estimated to have a relationship of 0.5 with each of their offspring. Similarly, the sire's offspring would each be expected to have a relationship of 0.25 with one another.

Genomic testing reveals that while, on average, these relationships are correct, the relationship between the sire and each individual offspring differs, and some offspring are more closely related to each other than they are with others.



Genomic relationships between 2,813 progeny of Merino sire, Anderson Poll, 160729 in the MerinoSELECT genetic evaluation conducted by Sheep Genetics. (Source: AGBU).



Genomic relationships between 2,259 progeny of Angus sire, Millah Murrh Paratrooper P15, in the TransTasman Angus Cattle Evaluation. (Source: AGBU).



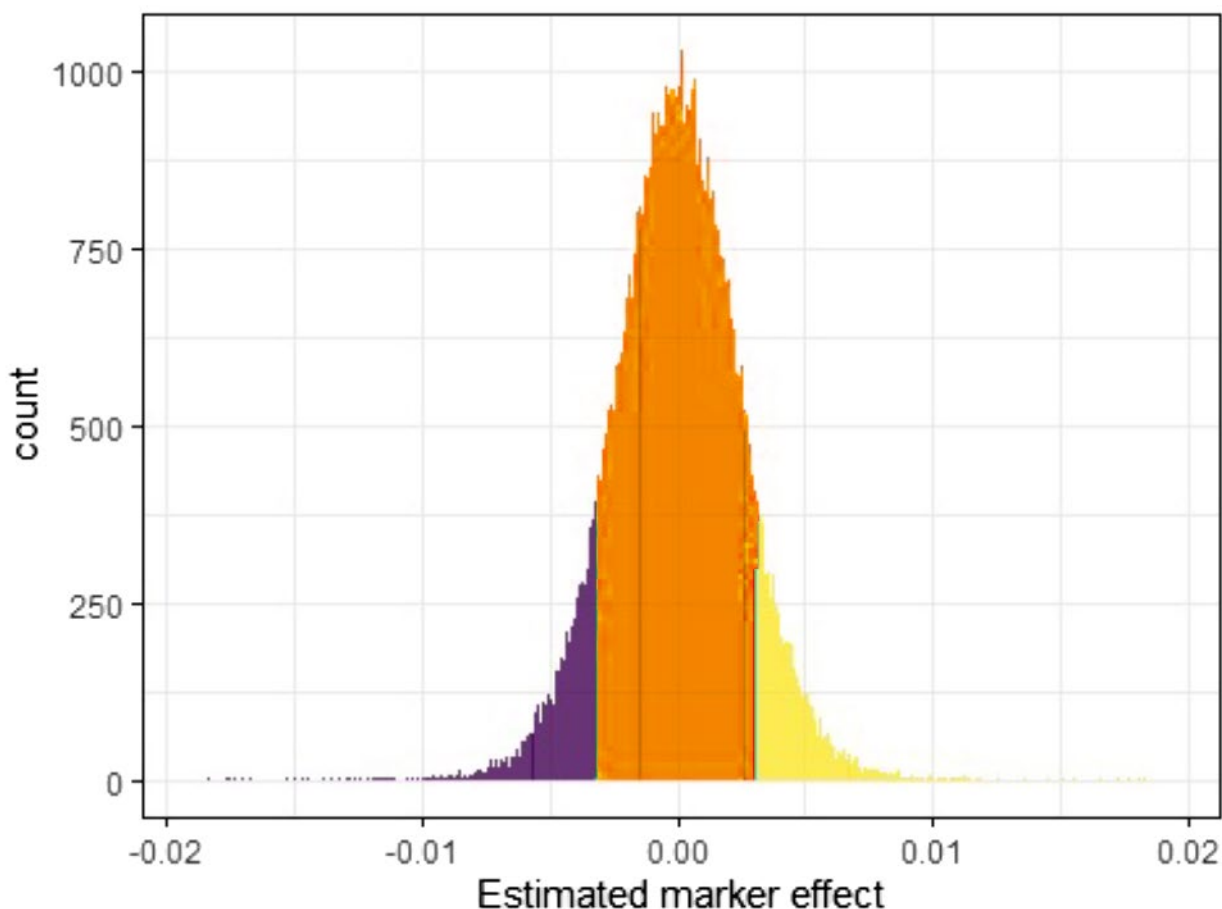
The ‘Marker Effects’ Approach

While the ‘relationship approach is about ‘animal relationships’, in a marker effects approach, such as the approach that is used by International Genetic Solutions (IGS), the incorporation of genomics is about the relationship between chromosome segments and the trait of interest.

In a marker effect approach, the influence, or size of effect that an individual SNP has on a particular trait is estimated from the genotypes and phenotypes of animals. For example, a particular SNP may be associated with decreasing birth weight by 10g, while another SNP may be associated with increasing birth weight by 15g. This is commonly referred to as the ‘marker’ or ‘SNP’ effect.

The genotype of the animal is assessed to determine the number of ‘doses’ of the SNP effect the animal receives. For example, an animal might have either 0, 1 or 2 copies of a particular SNP, which equates to the animal receiving either 0, 1 or 2 copies of the SNP effect.

When the breeding value is calculated by simultaneously analysing pedigree, performance and genomic information, the ‘dosage’ of marker effects across all SNPs contributes to the breeding value. For example, if an animal has a higher-than-average dosage of SNPs associated with increasing birth weight, the inclusion of genomic information will increase its birth weight breeding value. Conversely, if the animal has a higher-than-average dosage of SNPs that are associated with decreasing birth weight, its birth weight EBV will decrease with the inclusion of genomic information.



A plot illustrating the marker effect of different SNPs on birth weight in an Australian cattle breed (Source: AGBU)

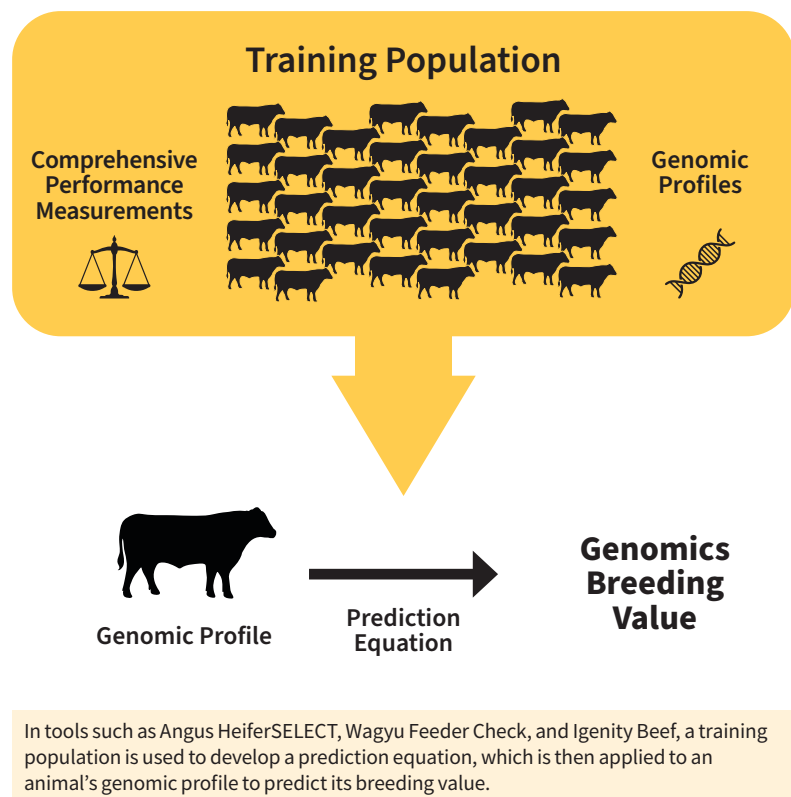
GENOMIC SELECTION IN COMMERCIAL ANIMALS

Genomic selection in commercial animals, such as tools Angus HeiferSELECT, Wagyu Feeder Check, and Igenity Beef, most commonly adopts a ‘two-step’ approach to predict an animal’s genomic breeding value.

Step 1 – a genomic reference population of animals with both phenotypes and genomic profiles is assembled. This population, often referred to as the ‘training population’, is used to develop a prediction equation. The prediction equation aggregates the influence, or size of effect that each individual SNP has on a particular trait within the training population.

Step 2 – each time that DNA sample is submitted for an animal, the DNA sample is analysed to create a genomic profile for the animal, following which the prediction equation is applied to the genomic profile to predict the animal’s breeding value.

One limitation of this approach is that the prediction equation needs to be routinely updated so that it remains relevant. One resolution to this limitation is to utilise the SNP effects from an industry single step genetic evaluation to construct the prediction equation, meaning the prediction equation is continually updated.



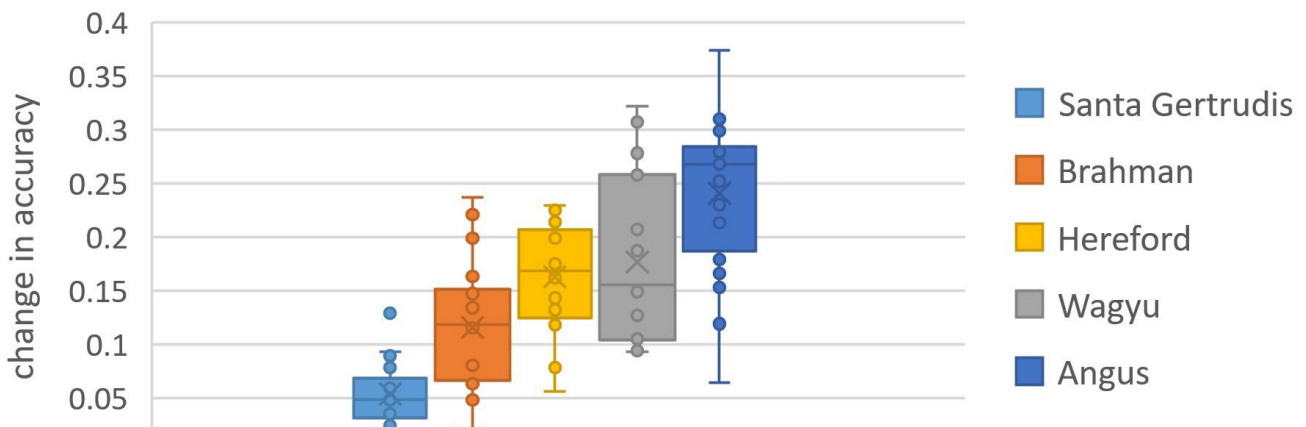
GENOMIC SELECTION : DOES IT REALLY WORK?

Genomic selection is a widely used, effective and proven technique in plant and animal breeding, and numerous research studies have been conducted demonstrating the value of genomic selection in livestock breeding programs.

One such study was undertaken by researchers at the Animal Genetics & Breeding Unit (AGBU), which demonstrated the benefit of genomic selection in the Australian beef industry.

The peer reviewed study, titled 'BREEDPLAN Single-Step Genomic Evaluation Delivers Increased Accuracies Across All Breeds and EBVs', compared the predictive ability of breeding values calculated in a traditional genetic evaluation with the predictive ability of genomically enhanced breeding values calculated in a single step genetic evaluation for a range of beef cattle breeds.

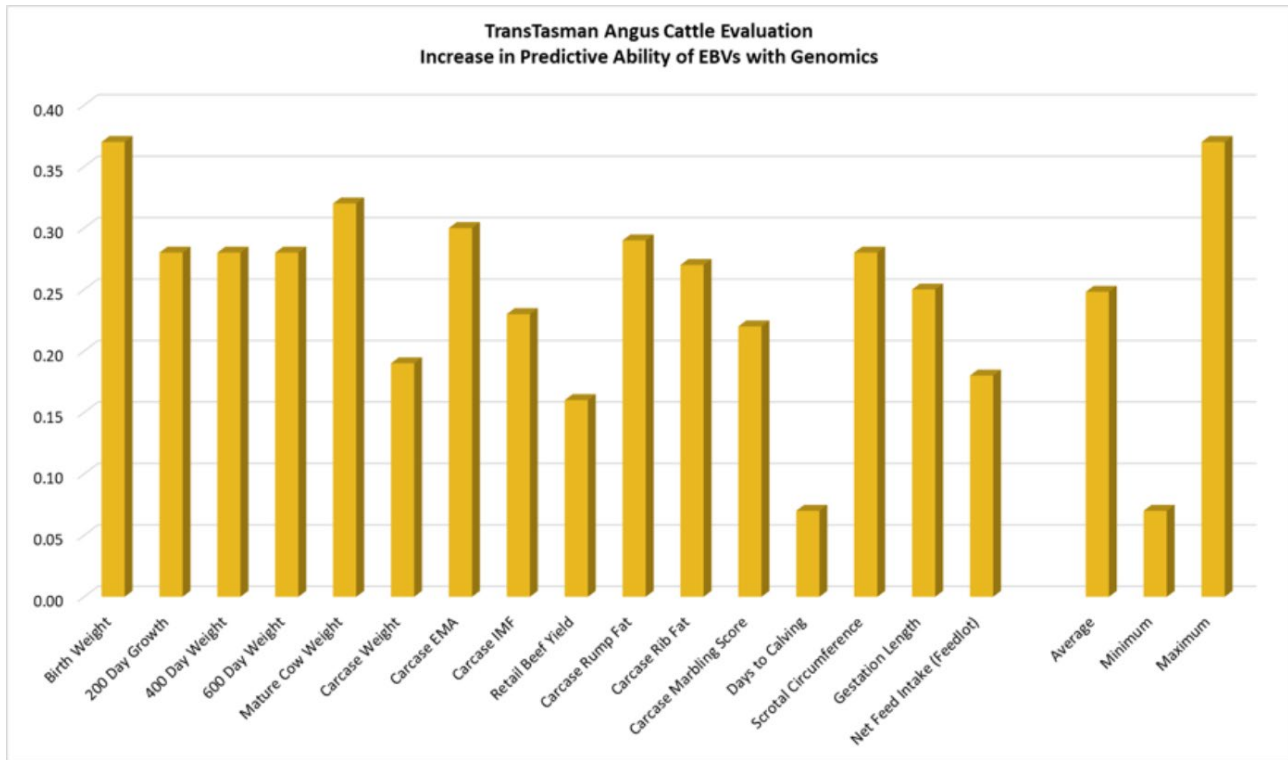
The study clearly demonstrated that the inclusion of genomic information increased the predictive ability of breeding values, with accuracy increases of between 5 – 30% observed.



Source: Johnston, D. J., Ferdosi, M. H., Connors, N. K., Cook, J., Girard, C. J., & Swan, A. A. (2023). BREEDPLAN single-step genomic evaluations deliver increased accuracies across all breeds and EBVs.

Further, the study showed that the increase in predictive ability that resulted from the inclusion of genomic information differed by breed and trait, with the magnitude of the increase largely determined by the size of the genomic reference population that was available for each breed:trait combination, the heritability of the trait, and the amount of genetic diversity present within the breed.

This is illustrated in the following image, showing that while the inclusion of genomic information increased the predictive ability of breeding values in the Angus breed on average by 25%, the increase observed varied from 7% to 37%, subject to the trait being analysed.



While this study clearly demonstrated the value of genomic selection, it also illustrated the critical role that the genomic reference population plays in genomic selection, and indicates:

- With the exception of some traits where the existing genomic reference population is very large, the increase in the predictive ability of breeding values is likely to be larger in the future as the size of the genomic reference population for each breed:trait combination increases.
- Organisations utilising genomic selection, via the provision of a single step genetic evaluation or otherwise, need clear strategies to maintain the genomic reference population if they are to continue to harness the benefit of genomic selection.
- Organisations who do not currently utilise genomic selection need clear strategies to generate a genomic reference population if they are to get any benefit from genomic selection in the future.

Demystifying DNA Technology

A Livestock Breeders Guide
to **Genomic Selection**



Neogen Australasia

Ph: +61 (7) 3736 2134
E: naa-lab@neogen.com
W: neogenaustralasia.com.au



Animal Genetics & Breeding Unit (AGBU)

Ph: +61 (2) 6773 2055
W: agbu.une.edu.au



Sheep Genetics

Ph: +61 (2) 8055 1818
E: info@sheepgenetics.org.au
W: sheepgenetics.org.au



Angus Australia

Ph: +61 (2) 6773 4600
E: office@angusaustralia.com.au
W: angusaustralia.com.au