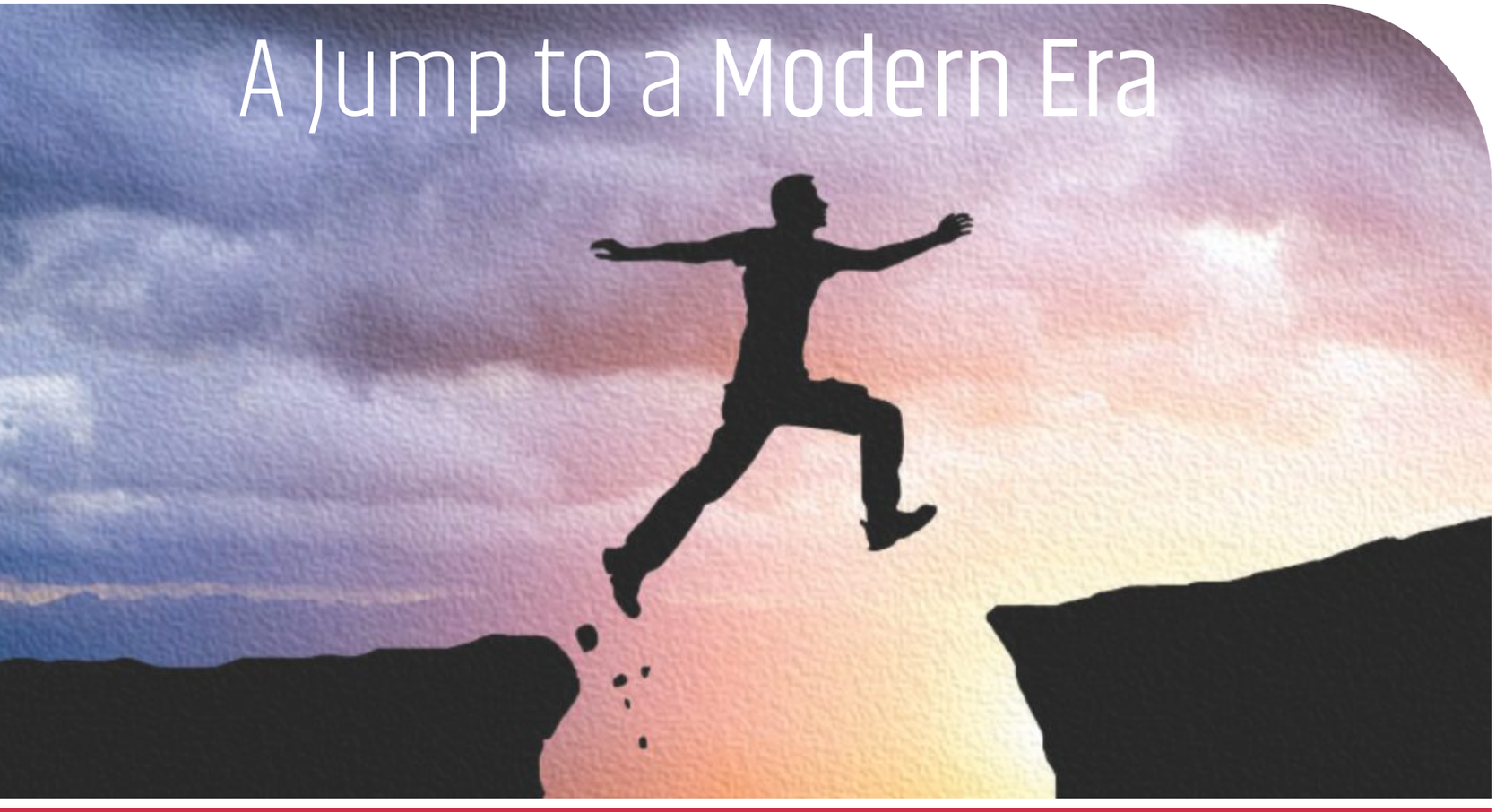


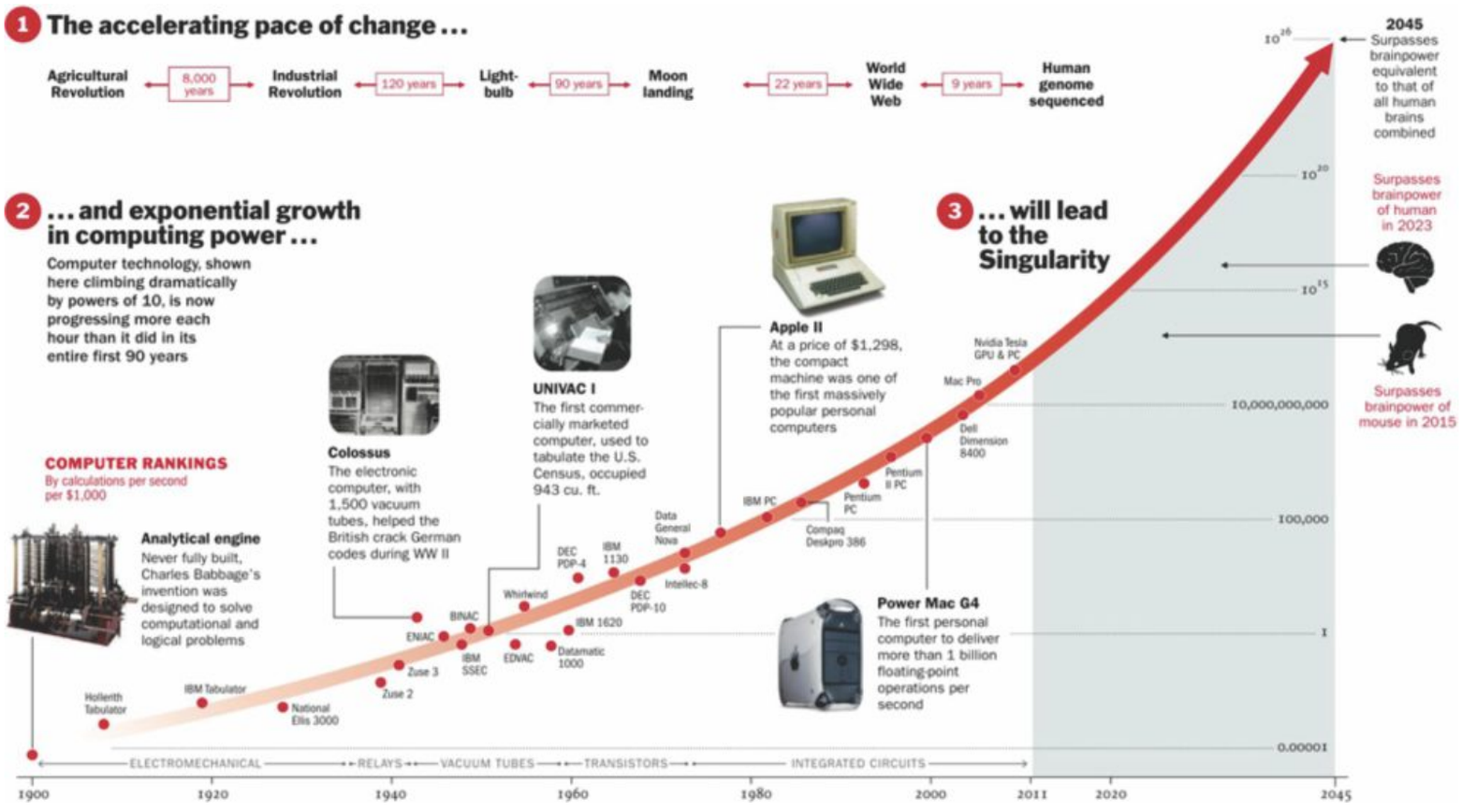
SYNERGY IN DATA RECORDING: FROM VISUAL TO SELECTION

by **Japie van der Westhuizen:** SA Stud Book and Animal Improvement Association

PART 2: Application in the present

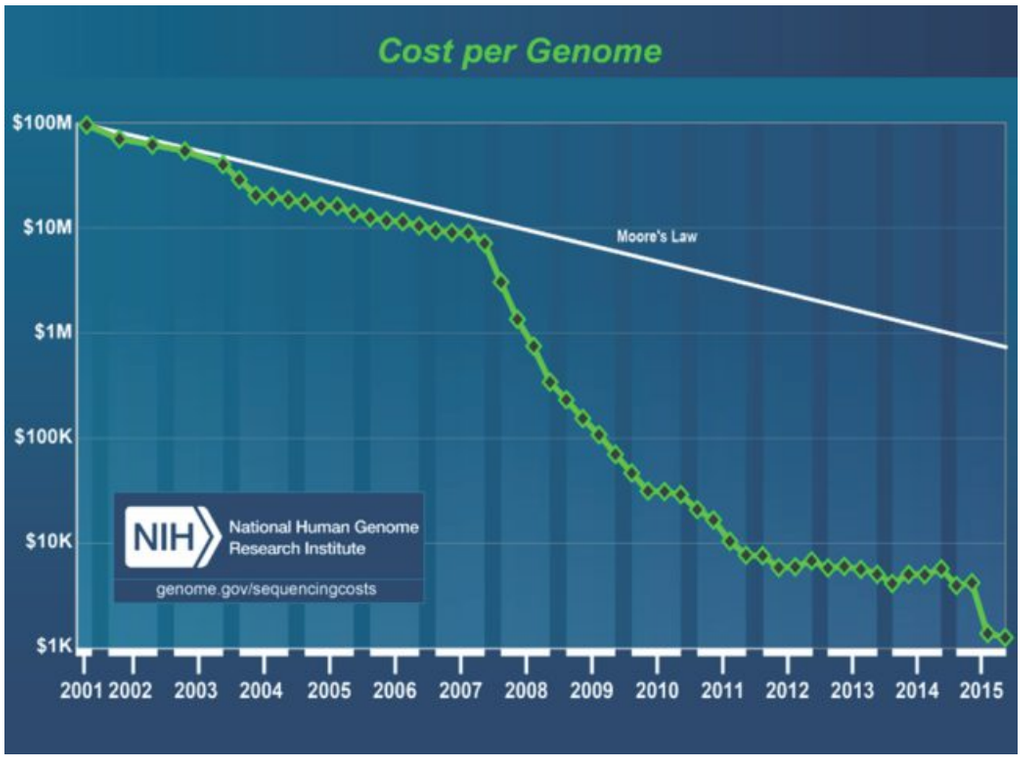
In 1965, Gordon Moore, co-founder of Intel Corporation (the computer processor micro chip manufacturer) stated "that the number of transistors per square inch on integrated circuits have doubled every two years since their invention. Since this statement was made, the debate, especially in the computer industry, has been if this pace can be maintained. Furthermore, the same principle has been claimed in other industries, namely a doubling of efficiency or speed annually. More than 50 years later, Intel still uses Moore's statement on their Web landing page!

Some would even warn against the fast pace of technological advance, including the possibility of "artificial intelligence". The current predictions are that computing power will surpass the human brain capacity ("singularity") by the year 2023 (and human brains on earth by 2045!). The following figure illustrates the advances in computer capabilities, keeping pace with Moore's expectations, since the early 1900s, and with the future predictions (from "2045: The Year Man Becomes Immortal"; Time 2011; content.time.com).



Animal breeding and genetics technologies are underpinned by a host of sciences and seemingly unrelated technologies. The initial curiosity in how properties, characteristics and traits are inherited was closely linked to the biological processes and physiology. As the realisation grew that, in many cases, differences in performance or expression of properties is the result of, in many cases, a multitude of genes (usually each with a relative small, so called "additive" effect) basically not detectible as single units, a different approach to predict genetic merit developed. This has led to the 'marriage' of breeding and genetics with the statistics associated with probabilities. **It is now relatively common knowledge that Professor Charles Henderson opened a new world of possibilities when he laid down the foundation of mixed model equations to predict the genetic merit of farm animals.**

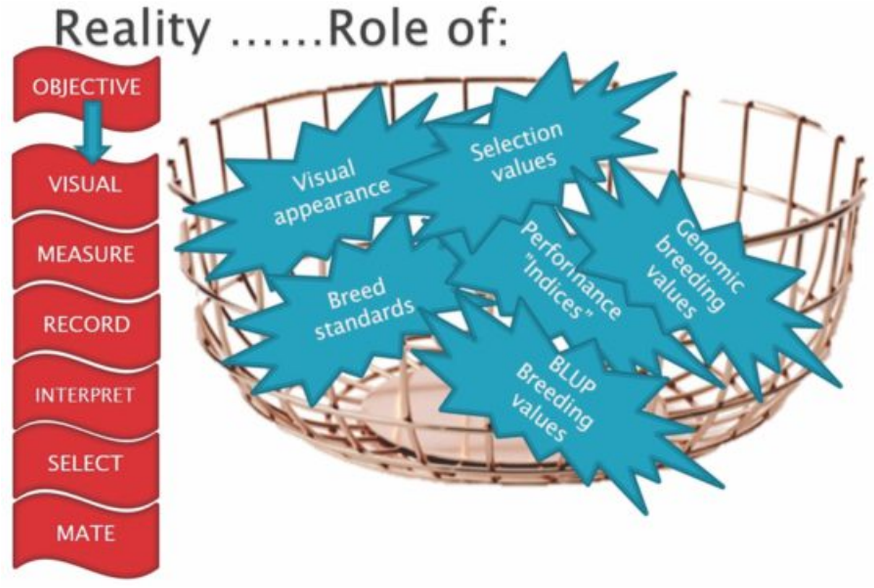
One of the major events in the unravelling the way inheritance work, was the Human Genome Project, where the genome (nucleotide base pairs) sequence of making up human DNA was determined. The project ran over a period of 13 years, involved numerous research groups from all over the globe and the costs ran over $3 billion (US Dollars). It is now possible to obtain a complete genome sequence of a person within a few days at a cost of about $1 000. Truly Moore's law was beaten by the Geneticists and the technologies they use. The next figure, presented by the National Human Genome Research Institute illustrates this.



Meanwhile, the fastest development in the animal breeding domain has been innovate means of integrating genomic information (usually in the form of partial information of the base pairs, or SNPs, on chromosomes of selection candidates) with BLUP breeding values. These developments are still based on the initial, age old theories where it is known that a magnitude of additive genes contribute to the differences in genetic merit for the traits of economic value, among these farm animals.

BACK TO REALITY...

Livestock breeders are sometimes confused with the many sources of information, methodology, institutional rules and role of common sense to apply in animal breeding.

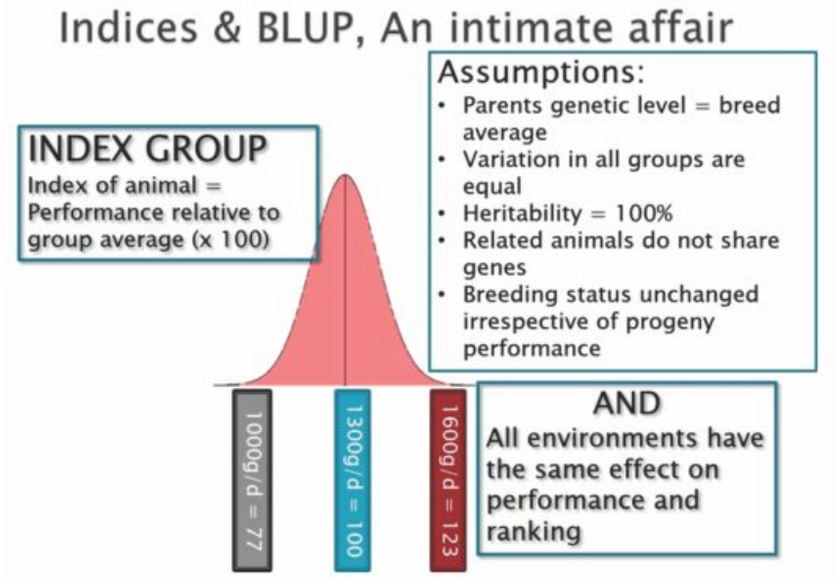


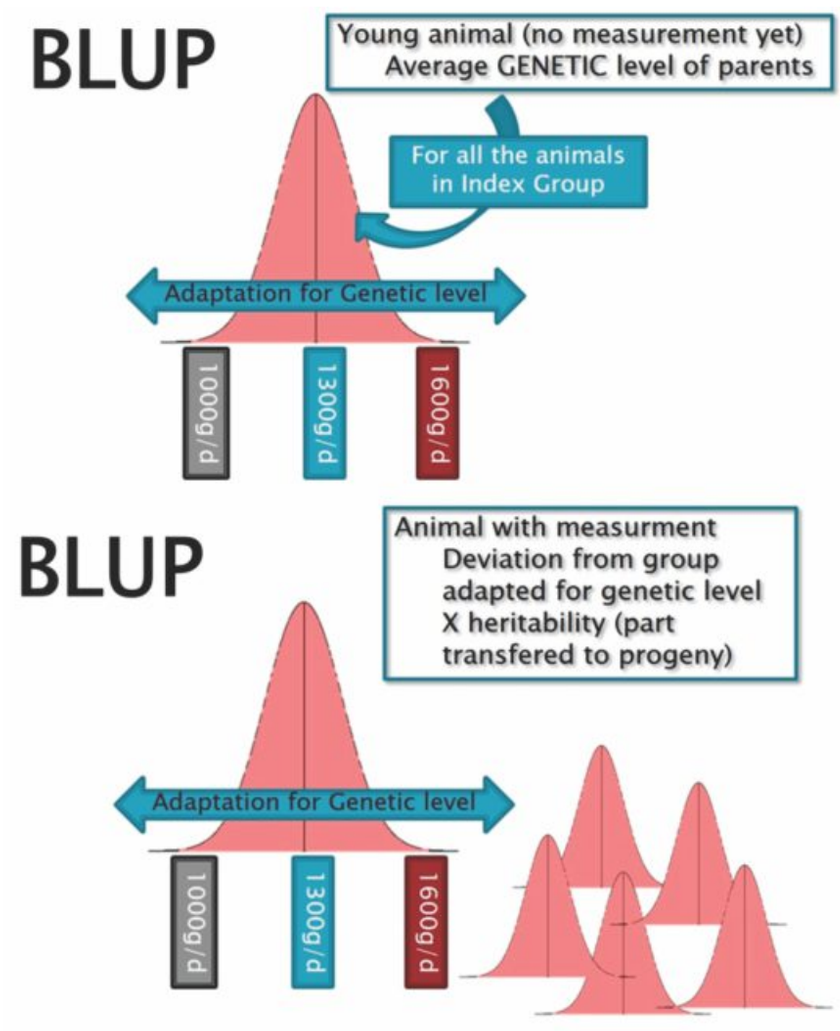
Visual appraisal or assessment still plays a very important role in livestock breeding. Not only can it act as first line (gatekeeper) in determining suitability of the right animal for the right environment or conditions, but it can also serve as source for more refined genetic assessment. **The challenge is therefore to link visual assessment to objective recording of these observations.** The dairy industry has been very successful applying these principles where functional traits are linked to linear classification and subsequent genetic merit predictions. Visual assessment also serves as the identification of elimination of unwanted genotypes. Many Breeders' Societies will also support these practices as part of their unique product identification and 'branding'. Nothing wrong with that!

Pedigree information (with visual assessment) used to be the cornerstone of seed stock (Stud) breeding. Registers of many breeds go back for decades and many stud farmers can quote linage of their valued animals. Sometimes a very neat and structured system is also applied in the naming of animals, based on their linage. At one stage, when the recording of performance traits became popular, there were voices denouncing the recording of pedigree information. The assumption was that the differences in performance among animals within contemporary or treatment groups will suffice to identify superior animals suited as selection candidates. Luckily stud breeders did not abandon pedigree recording. **Pedigree information serves many (scientific) purposes, among others, the following:**

- Incorporation into genetic merit prediction, making use of the principle that related animals share common genes. This is one of the cornerstones of BLUP breeding value predictions.
- Optimising breeding strategies and incorporation in plans to curb excessive increases in inbreeding and the loss of genetic diversity, especially in smaller populations (breeds with fewer numbers). Modern day mating computer software programs use relationships among animals to very good effect.
- Picking up the mode of inheritance and frequency of single genes, either deleterious or those desired in a cattle population.
- Gaining other very important information and patterns, such as retention and culling rates and reasons why the progeny of certain sires are culled more frequently.

Since the nineteen seventies, proper recording of objective measurement of traits with economic importance (performance recording) became the norm for most livestock breeders and producers. Although the obvious advantages relate to the genetic improvement of herds and breeds by identifying superior individuals as selection candidates, it has additional advantages, such as using the information as a benchmark, comparing the suitability of animals in specific production systems and environments, making timely management adjustments to ensure **sustainable profitability** and building reliable databases for reference and statistics.





The principle of comparing animals for selection purposes is to make these comparisons fair. This means that the breeder needs to define the test (contemporary) group well and then compare how each animal performs relative to their contemporaries or peers. In South Africa, we call this an **"INDEX"** which is a simple calculation where an adjusted weight (usually adjusted for the age of the animal and the age of the dam) or weight gain of each animal is compared to the average of the group, usually expressed as a deviation from 100. **This system served the breeding industry well, but has some flaws:**

- It is assumed that the superiority (or inferiority) as reflected in the difference in performance from the group can be transferred directly to the progeny. **In other words, the assumption is that the heritability is 100%. This is never the case.**
- Correction applied for factors such as age and age of dam is constantly applied across contemporary groups, but in real life they should be different. This can lead to re-ranking in groups.
- **Selection of top animals in groups tells you something about performance differences within the group and nothing about possibilities of superiority on other farms, seasons and birth years.** These facts could be the most important considerations for any buyer of a bull.
- **The performance label allocated to an animal because of its performance in its group will stick to it for life, irrespective of its success as a parent (breeding animal).** It is known that this is not true as a more accurate prediction of an animal's breeding value will be the result of progeny performance (relative to the performance of progeny of other parents).
- No clear separation can be made in the case of maternally influenced traits, such as birth and weaning weights.

The successful combination of pedigree information with performance measurements and the consideration of the heritability (the performance differences that can be transferred to the progeny) and genetic correlations (how selection for one trait will genetically alter the population with regards to a different trait), lead to a much more accurate prediction of genetic merit. Basically, all the flaws of the "Indexes" are addressed by **BLUP breeding values.** There are, however, still some very important aspects that need addressing to raise the reliability of genetic merit an extra bar. Low genetic merit prediction accuracies can be expected in the following cases (where the only source of information might be a, so called 'Parent Average' based breeding value prediction:

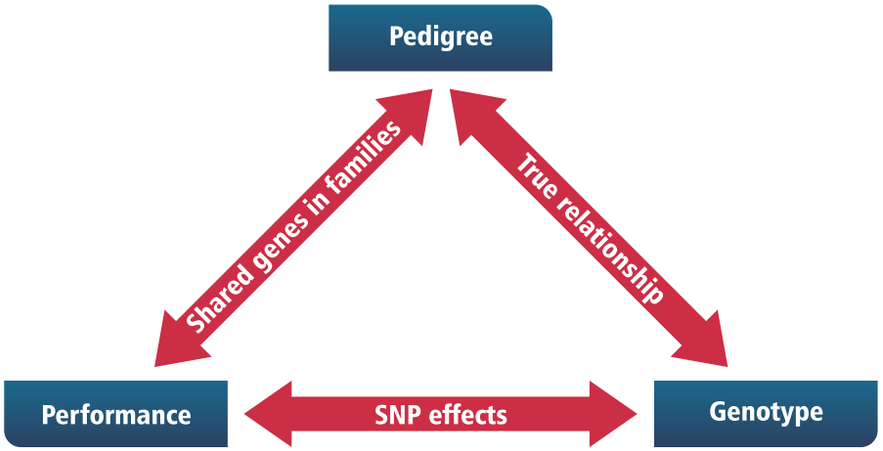
- The animal was still too young to be measured for the economically important trait. For example, if a breeder wants to select for herd life (longevity), he needs to wait until an animal is past its breeding age.
- Some traits can only be measured or recorded on one of the sexes. These are, ironically, some of the most economically important ones, such as the dam's effect on birth and weaning weights and female fertility (age at first calving and calving interval, as measurements within contemporary groups).
- Carcass characteristics can only be measures on slaughtered animals. BLUP breeding values for these properties will therefore rely on the recordings of sibs (mostly progeny groups of collateral related sibs).

Adding an additional source of information, namely **genomic information,** to pedigrees and performance records enhances the prediction accuracy to decide what animals will be of superior genetic merit. Adding this information to the BLUP breeding value predictions addresses all the shortcomings of BLUP but also goes beyond that, namely taking the true relationships of the genotyped animals into consideration.

It is known that collateral sibs (such as brothers and sister, cousins, etc.) might deviate from the expected mean relatedness. The same could be true for 'new' animals allowed in a breed, where pedigree information is not known. These GEBVs gives a better reflection of the expectation that an animal received 'good' or 'bad' parts of the chromosomes carrying the genes from its parents. The advent of **BLUP breeding values has increased the need for superior computing power and novel mathematical algorithms to resolve very complex calculations.**

Moore's Law has therefore been very helpful. **The big jump, however, is the incorporation of genomic information in the equations.** Even when using one of the less dense SNP chips, at least 7 000 records are added to each of the genotyped animals' data (it is more common to use SNP chips with 20 000 or more SNPs).

Modern day breeders therefore have access to state of the art technology, although most of the grinding happens behind the scenes. The sources of information used for predicting the value of animals as possible parents (genetic merit) is illustrated in the following figure:



Without ANY of the three building blocks, genomic EBVs are of very little value.

A LAST THOUGHT...

Embracing appropriate technology and identifying the value of different sources of information to make correct decisions, separates the exceptional breeders from the rest. The challenge is there to make use of the right tools and to put effort into recording that matters most to clients and buyers of seed stock.

"IN THE ERA OF GENOMICS, RECORDING IS KING"

(Adapted from Mike Coffey, EGENES, Edinburgh, Scotland).

Simply because the matter of selection with proper breeding objectives in mind and ways to include different traits in these objectives were not addressed in the article, does not mean it is of lesser importance. That is a topic of another day. ■