

Genomic selection to increase EBI

Donagh Berry¹, Dawn Howard², Sinead Waters², Francis Kearney³

¹*Teagasc, Moorepark Dairy Production Research Centre, Fermoy, Co. Cork.*

²*Teagasc, Animal Bioscience Center, Grange, Dunsany, Co. Meath.*

³*Irish Cattle Breeding Federation, Highfield House, Bandon, Co. Cork.*

Summary

- Genomic selection is a new source of data for genetic evaluation of animals which involves translating the DNA signature of an animal into its EBI and genetic merit for all other evaluated traits including milk production, fertility, health, beef and type
- The greater the amount of data on an animal or its relatives, the greater the reliability of the animal's EBI. Therefore genomic selection is expected to increase the reliability of EBI, the extent to which it increases the reliability will depend on the on-going research but the improvement will no doubt improve over the coming years also
- The DNA of an individual is present in every cell and remains the same throughout the animal's life. Therefore, we can get DNA from an animal at one day of age and allowing for laboratory work and statistical analysis it should be possible to have a more accurate EBI of an animal at a young age
- Genomic selection will alter breeding programs. It will be relatively cheap to screen a large number of young bulls in Ireland as well as foreign bulls thereby increasing the genetic pool and thus genetic gain. Genetic gain using genomic selection can increase by 50% over current methods.

Introduction

Traditionally most breeding programs in Holstein-Friesian populations selected aggressively for milk production, with some breeding programs also favouring angularity. It is now well accepted internationally that such a strategy has led to deterioration in genetic merit for health and fertility in dairy cattle. Although some of the negative genetic consequences can be amortised through management such as hormonal intervention or allowing slippage in calving patterns, these approaches are not sustainable long term.

To address this decline in genetic merit for health and fertility, while continuing to increase genetic merit for milk solids, Ireland has been placing an ever increasing emphasis on health and fertility within the breeding goal. The Scandinavian countries have applied selection pressure on health and fertility for many decades and the benefits

can easily be seen in the superior health and fertility of the Scandinavian breeds. Other countries are now placing more and more emphasis on fertility in their breeding goal and the gap between the emphasis on fertility in Ireland and most other countries is narrowing as other realise that fertility is a serious issue and can be reversed through genetics.

Current approaches to genetic evaluation

Fundamental for genetic gain in profitability is a pertinent breeding goal such as the Economic Breeding Index (EBI) as well as the tools for accurately differentiating between animals for genetic merit of a given trait or index, and a structure to disseminate the genes of genetically superior animals into the commercial population. Essential for accurately estimating the genetic merit of a group of animals is high quality data, often referred to as phenotypes. Great strides have been made by Irish farmers and other organisations, under the guide of the ICBF, to input and collate high quality data on Irish dairy cattle. These data are invaluable in the estimation of accurate genetic proofs for bulls and cows. GEN€ IRELAND[®] is the vehicle used to disseminate the genes of genetically superior animals into the commercial population.

National genetic evaluations in Ireland and all other countries are currently based on the theory of “contemporary comparison” which compares groups of animals producing under similar conditions (e.g., same herd and year and season of calving, same age etc.). However, the estimated breeding value, or EBI, of a new born calf or heifer is derived mainly from the genetic merit of its parents. We know that only one quarter of the variation in the performance of paternal half-sibs is attributable to their sire with the same being true for the contribution of a dam to maternal half-sibs. Half the variation is due to which combination of genes they get from the sire and dam. Therefore, it’s no wonder that we can’t fully predict how good an animal will perform based on the genetic merit of its parents, even if both parents are accurately evaluated. Because of this, the reliability of the EBI of a young animal is low.

The reliability of a young test sire or heifer as well as most stock bulls is approximately 30% for EBI compared to a reliability of 95 to 99% for proven sires. With a reliability of 30% the actual EBI of an animal can vary by $\pm\text{€}98$ from its published EBI; in a small proportion of cases it can vary more than that. This variation is reduced to $\pm\text{€}12$ in a well proven bull. This is an important point to remember when considering purchasing stock bulls. However, low reliability can be overcome by using a “team of bulls” which as the name suggests, means using more than one bull across your herd. Using a team of bulls means that the reliability of the average EBI of the team is considerably higher than the individual reliability of each bull. The analogy would be purchasing shares in more than one commodity thereby minimising the risk of some shares (or bulls) possibly decreasing since others are likely to increase. The calculation of the reliability of a team of unrelated bulls is simple and is given by the formula:

$$\textit{Team reliability} = 1 - \left(\frac{1 - \textit{average reliability of the individual bulls in the team}}{\textit{number of bulls in the team}} \right)$$

So the EBI of a team of 5 young bulls, each with a reliability of 30% (i.e., 0.30 for the equation above), will have a reliability of 86% associated with it.

Nonetheless, it is obvious that a more accurate estimate, or higher reliability, of the EBI of a young bull calf would be very beneficial both in giving farmers more confidence in their choice of young bulls as well as facilitating breeding organisations to more accurately identify better bull calves for entry in to GEN€ IR€LAND[®] be they sourced in Ireland or abroad. DNA technologies may offer a solution.

Previous research in genomics for animal breeding

DNA is the building blocks of genes, which in turn determine the attributes of an animal such as if a cow will milk more or be more fertile. DNA is present in all cells and, within an animal, is identical in all cells. Furthermore, DNA does not change over the lifetime of an animal (same as in humans). For this reason a great deal of research resources has been expended in recent decades on attempting to use DNA to identify genes that affect animal performance.

The theory was that genes would be found and by identifying cohorts of animals with the different variants of the gene it would be possible to determine which was the “good” variant and which was the “bad” variant. The objective of subsequent breeding programs would be to increase the frequency of the good variant in the population. This technique was termed gene assisted selection. The difficulty in this approach was that it was proving very difficult to find the important genetic variants. These difficulties arose because genes are very large and the genome, which contains all the genes in the cells is thus very large. Because of this it was difficult to determine where on the genome the important genetic variant was.

For this reason, another approach called “marker assisted selection” was proposed whereby genetic variants, now commonly referred to as genetic markers, that could be easily measured in the DNA of an animal were associated with performance. These genetic markers associated with good performance were to be bred into the population in the hope that this approach also bred in the genetic mutation causing the effect. This approach although suffering from the fact that each genetic marker may only have a small effect on performance in their own right did not require as large resources to identify potential useful genetic markers. However, marker assisted selection had other faults. The link between the “good” marker variant and the “good” genetic mutation actually causing the effect did not always hold across breeds, populations within breeds and even across sire lines within breeds. This meant that resources would have to be expended on a continuous basis to ensure marker assisted selection was successful across the entire population. Very few countries have implemented marker assisted selection in their dairy

cattle breeding programs and where applied, the benefits of this approach are thought to be very small.

Genomic selection – The Theory

Genomic selection is more or less a large scale version of the previously described marker assisted selection. Rather than selection based on a few markers where the effect of each marker has been previously calculated, genomic selection involves simultaneous estimation of the effects of many thousands of genetic markers. Animals are then screened for all markers and their individual marker effects summed to give an EBI. These genetic markers are called single nucleotide polymorphisms abbreviated to SNPs (pronounced “snips”). The idea of genomic selection began back in 2001 based on a scientific publication by Meuwissen et al. (2001) and genomic selection became possible when US companies like Illumina and Affymetrix developed platforms that could cheaply screen an animal for tens of thousands of genetic markers at the one time. Technologies are currently available to quantify what variants each animal has across 54,001 different SNPs; this platform is commonly referred to as a "SNPchip" (pronounced snip-chip) and the number of SNPs on these chips is likely to increase considerably in the coming years. Small scale laboratories can generate information on genetic markers on hundreds of animals a week; what genetic markers an animal has is commonly called the animal's genotype. This is essentially a “DNA signature”.

Advantages of genomic selection are a) resources do not have to be expended on identification and validation of individual genes that affect traits of importance, b) the reliability of EBIs is considerably higher for dams of potential young test sires and also for the young sires themselves, and c) generation intervals (i.e., average age of the parents when their progeny are born) of all potential selection pathways may be reduced thereby increasing annual genetic gain. Schaeffer (2006) in his evaluation of genomic selection in Canadian Holstein dairy cattle showed that genetic gain for a moderately heritable trait could be twice that of their current progeny testing scheme with a 92% savings on costs; he continued by stating that “the potential advantages of a genome-wide selection scheme are too great to ignore”.

Fundamental to the success of a genomic selection program is high quality phenotypes (e.g., accurate EBIs) and genotypes (i.e., data from the SNPchip) on a large number of high reliability animals. These resources will allow the disentangling of the individual effects of the thousands of genetic markers on the trait of interest, which in Ireland is mainly the EBI. Estimating the effects of the each of the thousands of genetic markers or SNPs is commonly referred to as "training the SNPchip". The majority of the costs in a genomic selection program is in the training of the SNPchip and the more accurate this procedure is undertaken the greater than reliability of the young test sires when applied in practice.

Genomic selection – Current status in Ireland

Research on genomic selection in Ireland started back in 2006 when semen samples of commonly used dairy and beef sires were stored by Teagasc through funding from the Department of Agriculture's Research Stimulus Fund, Teagasc, The National Cattle Breeding Center and the ICBF. Semen straws were collected from Irish and foreign AI organisations as well as Irish farmers. Genotyping of Holstein-Friesian sires began in January 2008. DNA for the animals used to train the SNPchip, commonly referred to as the training population or reference population, was extracted from semen at Teagasc, Athenry. Genotyping was carried out by a commercial company in Denmark, since this facility is currently not available in Ireland. To date over 1,000 Holstein-Friesian sires have been genotyped. A collaboration agreement is in place with New Zealand to attempt to utilise their sire genotypes in our analyses ensuring that more accurate EBIs will be made available to Irish farmers.

Teagasc Moorepark and the ICBF are currently in strong collaboration with other countries and international breeding organisations/genetic evaluation centers to refine the statistical methodology to simultaneously estimate the effect of each of the SNPs on performance in Ireland; this methodology is well advanced although will need fine-tuning over the coming years as experience gathers in analysing such data. The end product of this research is simply a list of the thousands of SNPs and their effect on the different traits measured in Ireland such as EBI, its subindexes and component traits such as protein yield and calving interval as well as the traits not directly included in the EBI such as protein percentage, the individual type traits, and overall type. This table of effects is often referred to as a "key" and in this case it will be the "Irish key". It is referred to as a key because it is unique to Ireland and is of limited use in another country. Similarly a key developed in another country will be of limited use in Ireland. This is why it is crucially important that this research is undertaken in Ireland using Irish data. However, because the DNA of an animal never changes, a genotype generated in Ireland, Australia, New Zealand, Holland or any other country will be the same. Therefore, these genotypes may be passed across country borders, the SNP key applied and the most suitable bulls for Ireland identified. The same is true of screening Irish bull calves for their suitability in other countries.

It is anticipated that genomically calculated EBIs, generated by the ICBF, will be available to Irish dairy farmers in Spring 2009. If successful, Ireland could be one of the first countries in the world to produce publicly available genomic values of animals from a national genetic evaluation center and have the facility also available to generate genomic values on cows. Research contracted by the ICBF on the potential of genomic selection has shown that genetic gain in Ireland can be increased by 50% compared to current progeny testing schemes. How the AI organisations decide to exploit genomic selection is under consideration by the individual AI organisations. The cost of the research program to-date in Ireland (including labour) is estimated to be €450,000.

Soon after Spring 2009 research will focus on the possibility of calculating genomic EBIs for calves and cows. If successful, farmers will be able to obtain genomic EBIs (i.e., EBIs derived from the current method of genetic evaluation "blended" with

genetic evaluation based on DNA) on their own animals. ICBF and Teagasc, Moorepark are currently evaluating the potential to extract DNA from an "ear punch" of an animal. This will remove the necessity of taking a blood sample since the early indications suggest that DNA extracted from hair samples is not consistently of good quality although research on this is still on-going. If adequate quality and quantity of DNA can be extracted from an ear punch then a farmer can tag a new-born calf at 1 day of age, send the sample to a laboratory where DNA extraction and genotyping will be undertaken. The DNA signature of the animal will be sent from the laboratory to the ICBF who in turn will undertake a genetic evaluation and post, e-mail or put up on the farmer's HERDPLUS internet site the genomic EBI of the calf which will be more accurate than the currently used parent average. Of fundamental importance is that the genotype of the animal remains in the ICBF database so it can be used to improve the estimates of the effects of each SNP thereby improving the reliability of the genetic evaluations.

Future research in genomic selection

Internationally, genomic selection is very much in its infancy and although considerable research is underway across the world a lot more questions need answering. However, arguably two of the most important areas of research for Ireland in the short term is the evaluation of alternative breeding programs to maximise the full exploitation of genomics for Irish farmers and the development of a "low cost chip" for screening animals. The best genetics for Ireland are in Ireland but a national breeding program must also introduce germplasm from other countries. Genomic selection, if used correctly, will aid greatly in identifying suitable germplasm in other countries.

The cost of genotyping an animal with the large SNPchip is approximately €250-300. This is prohibitively expensive for individual farmers to screen all potential replacement heifers or breeding females. Although not known for certain, we expect only a few hundred genes to affect individual traits such as milk yield. Therefore it seems plausible, in theory at least, to develop a SNPchip with hundreds rather than thousands of relevant SNPs thereby reducing the overall cost to screen animals. Although the cost will depend on the number of SNPs eventually included in the reduced chip as well as the number of animals screened, an estimate would be that the cost of results from the smaller would be approximately one third of that of the larger chip.

Conclusions

Key to maximising genetic gain in profitability are a well defined breeding objective, high quality data to distinguish between animals of good and poor genetic merit, and a structured breeding program to efficiently disseminate superior genes into the commercial population. All three factors are well advanced in Ireland. To-date estimated EBIs of sires/animals have been derived from parental information as well as

phenotypic records of the daughter/cow for traits such as milk yield and calving dates. Genomic selection, which measures differences in the DNA signature of animals, offers another source of data. However, these data are available once a biological sample can be taken from the animal. EBI reliability of a young animal with genomic information (i.e., information on its DNA) will be greater than based solely on parent average. Genomic selection has the potential to increase genetic gain in EBI by 50%, or in other words from approximately €23/year to €35/year. This equates to over €2.5m annually for the Irish dairy industry which is cumulative and permanent.

Acknowledgements

Authors gratefully acknowledge the financial contribution of the Research Stimulus Fund (06-0328 and 06-0353), Teagasc, The National Cattle Breeding Center, the Irish Cattle Breeding Federation and the Department of Agriculture Animal Genetic Conservation.

References

- Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157:1819-1829.
- Schaeffer, L.R., 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218-223.