

**ICAR – 28/05/2012**

**Genomics – what does the future hold for dairy farmers?**

*J.E. Pryce<sup>1,2</sup>, B.J. Hayes<sup>1,2,3</sup> and M.E. Goddard<sup>1,2,4</sup>*

<sup>1</sup>*Biosciences Research Division, Department of Primary Industries, Bundoora, Victoria 3083, Australia*

<sup>2</sup>*Dairy Futures Cooperative Research Centre, Victoria 3083, Australia*

<sup>3</sup>*La Trobe University, Bundoora, Victoria 3086, Australia*

<sup>4</sup>*Faculty of Land and Food Resources, University of Melbourne, Parkville 3010, Australia*

**Abstract**

Genomics is a new, exciting technology that has rapidly been adopted by the dairy industry. Young bulls are now being offered with their proofs calculated using information from DNA markers. As the reliabilities of bull proofs increases, we will see a shift from traditional progeny-testing schemes, to hybrid schemes where smaller teams are selected on the basis of their genomic breeding values, to schemes where young sires are marketed on their genomic breeding values and are replaced after only a few years by the next, genetically superior crop of bulls. These bulls will probably still get conventional breeding values, although perhaps only for those traits that data is required for other purposes. For example, yield, health and fertility data will still be needed for management reasons and will continue to be recorded. However, for other traits such as conformation (type) traits, that breeding companies have traditionally made an effort to record on young progeny tested bulls, it is not clear how the mature genomic-evaluation model will look. One option is that breeding companies or farmer co-operatives will work intensively with a representative sample of herds, on which they will record a raft of traits of importance. Increasing the reliability of genomic proofs is one of the key challenges right now. Ways in which this can be achieved are: 1) increasing the size of the reference population, which is likely to be more cow-focused in the future as the length of time which individual bulls are used for decreases 2) increasing the density of genetic markers or using the entire genome sequence 3) using better statistical methods. As progeny-testing reduces, the number of young bulls used a year is also expected to reduce, which will present challenges for maintaining the reference population – for which genotyping herd recorded cows is an obvious solution. Additionally, it is likely that young sires will be turned over

more rapidly, so bulls that are “breed icons” may become a thing of the past and replaced with team-products. This could mean that breeding decisions evolve into exclusively computer-based decisions as remembering extensive cow and bull pedigrees so that inbreeding is avoided becomes more difficult – even for the fanatic. Finally, as genotyping becomes cheaper, we might see genotyping of even commercial cows become normal practice. This information can be used to 1) identify parentage with certainty; 2) generate sophisticated mating plans that help to optimise genetic gain and avoid inbreeding (at the genomic rather than pedigree level); 3) select replacements or candidates to sell; 4) avoid genetic defects and 5) select for difficult to measure traits such as feed conversion efficiency or methane emissions.

## **Introduction**

Genomic selection refers to selection decisions based on genomic breeding values (GEBV). The GEBV are calculated as the sum of the effects of dense genetic markers that are approximately equally spaced across the entire genome, thereby potentially capturing most of the genes that cause differences between animals in the traits we are interested in.

There are very few examples where the price of a commodity has dropped as dramatically as genotyping. In 2001, the first human being was sequenced at a price of US\$3 billion (Venter et al., 2001). That is deciphering every single piece of the genetic make-up of a person, or every base pair in a very long string of DNA (which in reality was >3 billion base pairs which they had to read over 20 times over to ensure the sequence was accurate). Ten years later, Illumina lowered its full genome sequencing price to \$5,000 per human genome, or \$4000 for orders of more than 50. One of the biggest successes in agricultural science in recent years has been to leverage off investment made in medical genetics and applying it to make smart new ways to select the best individuals for the next generation. Sequencing of bulls (key ancestors of dairy and beef breeds) is now happening around the world, with an ambitious aim to have 1,000 bulls fully sequenced in the next couple of years. In this paper, we will describe how this fits in with the future of genomics in dairying.

## **Will progeny testing come to an end?**

One of the features of genomic selection is that it allows the genetic merit of young candidate bulls calves to be assessed with much greater accuracy than at the same age under progeny-testing. A day old bull calf could have a reliability based on pedigree information of around 30% (before being selected for progeny-testing). Through genomics, a day old bull calf can have a reliability of over 60%, which is roughly the same as 30 daughters for a trait with a heritability of 30%, such as protein yield. For comparison, progeny-tested bulls have first proof EBVs with reliabilities of around 85% (>60 daughters).

The increase in reliability at birth through genomics compared to parent average, means that fewer young bulls are being selected for progeny testing by breeding companies. In fact, a reliability of 60% is high enough to be confident enough about a bull's genetic merit to use him widely. This means that conventional progeny testing is already coming to an end in many countries, with young genomically tested sires having large daughter groups of 100s or 1000s of daughters in their first "progeny test". One of the advantages of progeny-testing is that it allows a pre-screen for genetic diseases to be made. Some breeding companies are already implementing a two-stage release of young bulls marketed on their genomic evaluations, with the first stage being a limited use of young bulls to

### **What will happen to data recording?**

There is a very good reason why we actually need to continue collecting data on cows. That is to avoid a reduction in reliability because the reference population (used to estimate marker effects) becomes more distantly related to the predicted population as we rapidly turn-over generations. At current marker densities (around 50k Single Nucleotide Polymorphisms or SNPs), the prediction of GEBVs using associations between SNP data and phenotypes in a reference population of bulls needs to be continuously updated, otherwise the reliability of the GEBV erodes rapidly over generations due to the decay of genetic relationships between the reference and validation populations (e.g. Habier et al. 2007). Therefore very high quality phenotypes (measurements on all traits of importance) are vitally important to make our genomic predictions as accurate as possible now and into the future. *Information*

*herds*, that are meticulous record keepers could provide the information needed to make sure prediction equations are accurate.

Most EBVs are already a by-product of data recorded for management purposes, such as for milk recording and health and fertility management. So, for as long as data is being collected it will be used to estimate breeding values. As more farmers move to inline monitoring the task here is to link farm computers to central data repositories for genetic evaluations. Data that is collected by breeding companies specifically for progeny-testing, such as conformation or type data might be more challenging to obtain in the genomic era. One option is that breeding companies, or even genetic evaluation units will start working intensively with a representative sample of herds, on which they will record a raft of traits of importance. All cows in these herds would probably be genotyped and they would become *information herds* for genetic evaluations. Under a breeding company scenario, this would probably mean that genomic evaluations would be done in-house by the breeding companies, which could lead to a pyramid of genetic dissemination, not dissimilar to the pig and poultry industries. So, where does this leave national genetic evaluation centres (usually paid for by farmer levies)? The questions around this sort of arrangement are who pays and how should the farmers of the herds be rewarded for their efforts?

Another advantage of having information herds is that novel, expensive or difficult to measure traits can potentially be recorded. Examples include traits related to feed intake, such as energy balance and residual feed intake and some measures of fertility, such as commencement of luteal activity. Residual feed intake is the difference between actual (measured) feed intake and predicted feed requirements, based on liveweight, growth rate and level of milk production (for lactating cows). One of the strengths of genomic selection is the opportunity it presents to select for these traits. Provided a reference population with genotypes and phenotypes is established i.e. the *information herds*, then genomic predictions can be calculated for animals with genotypes, but without phenotypes.

Herd-recording of production, health and fertility traits will continue for reasons other than genomic selection, such as herd management. Genetic and genomic evaluations

will continue to make good use of these data in determining genetic merit of cows and bulls to improve selection decisions.

### **Will the reliabilities of genomic selection reach progeny-testing?**

There is a lot of research effort going into achieving this important objective. In fact this is one of the main aims of the Animals Program of the Australian Dairy Futures Cooperative Research Centre (CRC). Here, our strategy has been 1) to extend the reference population to include females as well as males; 2) use the sequence of 1000 bulls to increase the density of SNPs in the reference population 3) improve the statistical methodology used.

Recently, the Australian Dairy Futures Cooperative Research Centre's 10,000 Holstein Cow Genomes project and Jer-nomics project embarked on collecting DNA samples and genotyping 10,000 Holstein and 4,000 Jersey cows (from commercial herds). This information has recently become part of the reference population. This has led to an approximate 8% improvement in the reliability of most breeding values (From 50% to 62% depending on the trait). Careful analysis confirmed addition of the cows was not biasing the GEBV.

Having the entire sequences of bulls may help to increase the accuracy further. The idea behind sequencing key ancestors of cattle breeds, is that we will have the causative mutations in the data set, i.e. we will be able to capture more of the genetic variation in a trait.

By sequencing the key ancestors of say, the Holstein and Jersey breed we will attempt to impute (fill in the gaps), of commercially available SNP chips, such as the popular 50k SNP panel up to the entire sequence. One of the advantages with moving to sequence data is that it is thought that the reference population will not need to be refreshed so often, as the associations will be between genetic markers that are closer to the genetic variants actually affecting traits like fertility and production. .

### **Will there be new traits?**

One of the most powerful potential uses of genomic selection is to select for economically important, yet difficult or expensive to measure traits. Provided the accuracy of genomic breeding values is high enough, these traits lend themselves well to genomic selection. This is because you need a dedicated reference population, such as a group of herds on which the expensive measurements are made. Associations between the genotypes and phenotypes estimated in this population can be extended to animals with genotypes but no phenotypes.

Examples include some measures of fertility, such as commencement of luteal activity and traits related to feed intake, such as energy balance and residual feed intake (RFI). RFI is the difference between actual (measured) feed intake and predicted feed requirements, based on liveweight, growth rate and level of milk production (for lactating cows). It is important because feed is a major component of on-farm costs. Recently, Pryce et al. (2012) reported a cross-validation study of 2,000 6-month-old heifers from Australia and New Zealand to predict the accuracy residual feed intake using genomic selection and assessed using cross-validation. All heifers were genotyped with the Illumina High Density Bovine SNP chip ([www.illumina.com/agriculture](http://www.illumina.com/agriculture)) and had detailed phenotypes on feed intake and liveweight enabling RFI to be calculated. The average accuracy was 0.37 in Australian heifers and 0.31 in New Zealand heifers. Hayes et al. (2011) demonstrated that at this accuracy, including RFI in the Australian Profit Ranking (APR), and selection for the index, with DNA marker derived breeding values for RFI, would improve the rate of annual gain for profitability by 3.8%.

Selection to improve “difficult” traits is often hampered by the cost of measurement and as a consequence they are generally only available on small numbers of animals. Optimal strategies to collect the phenotypes and genotypes for these traits are required if they are to be incorporated in future breeding objectives. One way in which this can be done is to set up dedicated resource populations in which the phenotypes are collected. Alternatively, there may be opportunities for research organisations to combine data on genotypes and comparable phenotypes.

## **How cheap does genotyping need to be to make it worthwhile genotyping my whole herd?**

As genotyping costs reduce for low density SNP chips, the prospect of using this technology on commercial dairy farms becomes increasingly attractive. Especially if genotyping is used for several strategies, such as mating plans to control inbreeding, select the best replacements and parentage verification. The situation is different for pedigree breeders who may realise better sale prices from genotyped high genetic merit heifers (Pryce and Hayes. 2012).

For selecting replacement heifers using EBI, assuming a standard deviation of EBI of €62 (Donagh Berry personal communication) and genotyping costs are €29/cow, the net profit of genotyping 40 heifers to select the top 20 as replacements (per 100 cows) would be worth €46/cow per cow discounted over four lactation. This is after spreading the cost of genotyping over the replacements kept. However, using parent average estimated breeding value information is free and can already be used to select replacement heifers. If genotyping is used solely as a tool to select replacements, the cost would need to be €15/cow before it becomes economically worthwhile. This calculation excludes rearing and sale costs, as it is assumed that there is no deliberate rearing of extra heifers specifically to sell. Currently, there is a lively export market of heifers from Australia to China. The net profit (i.e. sale minus rearing costs) for in-calf heifers is in the region of \$500 (€390), making it attractive to rear and sell extra heifers for the export market. If extra heifers are reared specifically for selling, genotyping may make it easier to decide which animals to sell or keep.

Extra value from can also be captured by using other strategies to profit from genotyping. For example, mating plans that use genomic relationships rather than pedigree relationships to capture inbreeding are superior in terms of reducing progeny inbreeding at a desired level of genetic gain, although pedigree does an adequate job. A 1% reduction of inbreeding, valued at AU\$5 per annum or AU\$14.20 discounted over four lactations, or (€11.09), can be achieved. Ascertainment of sire (and dam if genotyped) is an additional use of genotyping (Hayes, 2011) and is potentially worth around AU\$36 per cow (i.e. that is the cost of a parentage test obtained using a DNA sample; Holstein Australia) and assumed to be similar in Europe (€28).

The net value of selecting replacements, avoiding inbreeding and ascertainment of pedigree is €8.44/cow when genotyping costs are €9/cow, this value is profit, i.e. already includes the genotyping costs. If genotyping falls to €5/cow then the net benefit is €6/cow. Avoidance of genetic diseases and selling of heifers reared specifically to sell have extra value. There could also be extra benefits from selecting for difficult to measure traits such as feed conversion efficiency or methane emissions.

## **Conclusions**

In the near future, farmers can look forward to higher reliabilities for genomic breeding values, opportunities to select the best heifers as herd replacements, greater use of sophisticated mating plans to optimise bull usage and avoid inbreeding and new traits to select for (e.g. feed efficiency). Milk-recording and recording health and fertility events will still be very important in providing the data required for robust genomic evaluations.

## **Acknowledgements**

Funding is gratefully acknowledged from the Dairy Futures CRC, Melbourne, Australia

## **List of References**

- Habier D., R.L. Fernando, J.C. Dekkers, 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177: 2389–2397.
- Hayes B.J. 2011. Technical note: Efficient parentage assignment and pedigree reconstruction with dense single nucleotide polymorphism data. *Journal of Dairy Science* 94:2114-2117.
- Hayes, B.J., J.H.J. van der Werf, J.E. Pryce, 2011. Economic benefit of genomic selection for residual feed intake (as a measure of feed conversion efficiency) in Australian dairy cattle. *Recent Advances in Animal Nutrition* 18: 31-36.



- Pryce, J.E. and Hayes, B.J. 2011. A review of how dairy farmers can use and profit from genomic technologies. *Animal Production Science* (<http://dx.doi.org/10.1071/AN11172>).
- Pryce, J.E., J. Arias, P.J. Bowman, S.R. Davis, K.A. Macdonald, G.C. Waghorn, W.J. Wales, Y.J. Williams, R.J. Spelman and B.J. Hayes 2012. Accuracy of genomic predictions of residual feed intake and 250 day bodyweight in growing heifers using 625,000 SNP markers. *J. Dairy Sci.* 95: 2108-2119
- Venter, J.C., et al. 2001. The sequence of the human genome. *Science* 291: 1304-1351.