Implementation of genomic breeding values for novel traits such as feed efficiency through female nucleus reference populations

J.E. Pryce[1] *and D.P. Berry*[2]

[1]*Department of Environment and Primary Industries and La Trobe University, Agribio, 5 Ring Road, Bundoora, Victoria 3083, Australia*

[2]*Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark, Co. Cork, Ireland*

Abstract

Genomic selection requires a large reference population of animals with both phenotypes and genotypes and until recently, most countries that have implemented genomic selection have used reference populations of sires. Female reference populations are now gaining in popularity, especially for novel traits, such as feed efficiency, methane emissions, and detailed reproductive measures which are expensive to measure and are therefore available for only some animals in the population. An example is individual feed intake data, which is mainly available from research herds. This is because the equipment required to measure these traits are too expensive for commercial farmers to invest in. So far two methods have been proposed to predict feed efficiency: 1) genomic selection and 2) mid infra-red technology applied to milk samples. Although using mid-infrared technology to predict variation in feed efficiency appears to be promising and has the advantage that it can be calculated for ungenotyped animals, these predictions (like genomic selection) still require feed intake data to form the predictions and they also need to be validated in independent populations. So, regardless of the method, a large reference population of animals with feed intake phenotypes are required to estimate the prediction equations and these data are generally challenging to obtain in substantial quantities. A large multi-national collaboration (called gDMI) has established a reference population with phenotypes and genotypes for genomic prediction. Challenges remain to 1) maintain the reference population into the future, and 2) implement breeding values within member countries.

Keywords: genomic selection; feed efficiency; novel traits.

Introduction

Genomic breeding values are available in many countries for traits that already have traditional pedigree based breeding values (Pryce and Daetwyler, 2012) and now there is a growing appetite for a broader range of genomic breeding values, provided they can be estimated with acceptable accuracy. Farmers value traits with a direct impact on profitability, such as feed efficiency and disease resistance, and these are likely to be among the first traits that will be available as "genomic-only" breeding values.

One of the main differences between "genomic-only" breeding values and traditional breeding values is likely to be that they are likely to be derived from phenotypes obtained from dedicated reference populations of females that are genotyped and also have phenotypes.

Female genomic reference populations – considerations

The three main considerations in establishing a genomic reference population of females is: 1) the size of the population required; 2) the phenotypes that are currently available or that need to be measured and 3) selection of herds/individuals to participate in the reference population.

The reliability of genomic prediction is a function of the size of the reference population, the length of the genome, the number of independent genomic segments, the effective population size, and the heritability of the trait (Daetwyler et al., 2008, Goddard, 2009). When only the heritability and number of females in the reference population are varied the reliability can be calculated deterministically (Figure 1). In order to achieve a reliability of 50%, the number of females that would need to be included in the genomic reference population varies between 12,000 (heritability of 0.5) and 120,000 (heritability 0.05) (Gonzalez-Recio, 2014). They showed that for traits that are expensive to measure or in the early stages of phenotype collection, female reference populations are preferable to male reference populations. As soon as many farmers collect phenotypes for a given trait, then switching to male reference populations (i.e. genotyped

males with ungenotyped progeny groups) may be preferable, especially when large numbers of individuals are required.

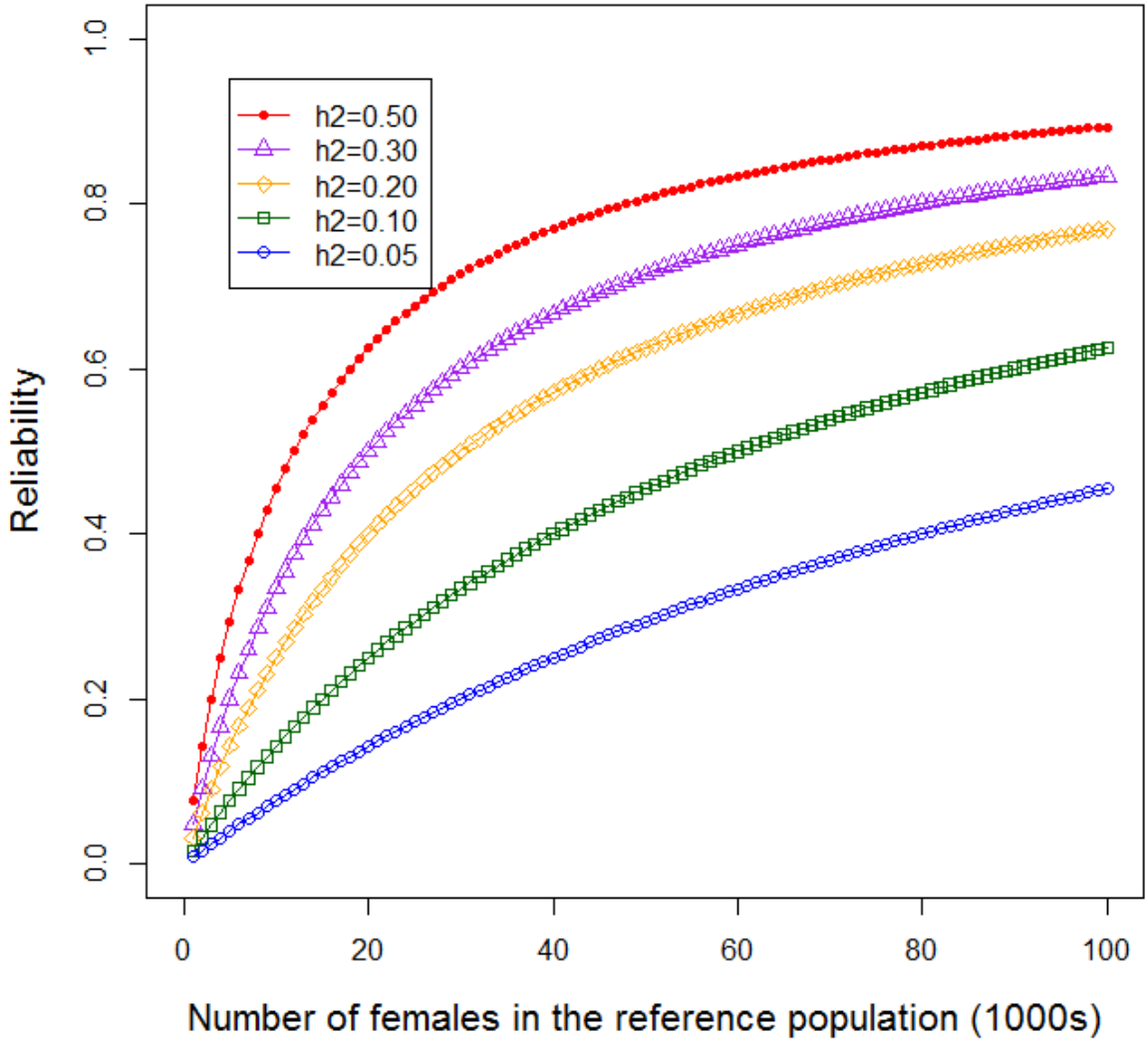

Figure 1. Expected reliability for reference populations of 0 to 100,000 (from (Gonzalez-Recio, 2014)

Building genomic reference populations of 120,000 females will be challenging especially for difficult or expensive to measure traits. For this reason, we believe, there will be a lot of emphasis in trying to achieve higher heritabilities for many traits that have traditionally been viewed to have low heritabilities, for instance health and fertility traits. More detailed

reproductive phenotypes may help to increase the heritability of fertility traits in a similar way to using progesterone profiles to obtain more heritable fertility phenotypes (Veerkamp et al., 2001).

In many countries, there are minimum acceptable reliability thresholds for publishing breeding values. However, until there is substantial advances in estimation methods used for genomic prediction, lowering of the reliability threshold for some traits that are only available via genomic selection may be necessary. Education and risk management are likely to become more important in ensuring that the use of this information is appropriate. For example, knowing a bull has a poor breeding value for feed efficiency (with low to moderate reliability), is likely to be better than the absence of any information. Already, in New Zealand, Livestock Improvement (LIC) publish bull genomic breeding values for residual feed intake (RFI) with reliabilities of around 10%; this prediction is based on a genomic reference population of around 2000 growing heifers genotyped at high density (Pryce et al., 2012).

The type of reference population may also depend on the phenotypes collected. For example, traits that require specialist and expensive equipment, such as feed intake are likely to be measured in research herds. Provided enough cows have phenotypes and genotypes, these individuals can be used to form a reference population.

For other traits, such as metabolic disorders, mastitis, lameness, methane emissions etc, there are already high-tech farm systems that assimilate and provide management reports using data from: pedometers, daily body weight, automatic oestrus detection, milk conductivity, breath analysers etc. There are already systems commercially available that record data on at least some of these phenotypes. While there is some concern about the accuracy of prediction of some of these systems (Rutten et al., 2013), as the technology improves, the accuracy of prediction and therefore the value of these data will also increase. Moreover, there is an opportunity to link into these systems by genotyping cows that have valuable phenotypes. For instance, daily changes in liveweight and distance walked could be used to enhance predictions of feed efficiency. While, electrical conductivity of milk (Norberg et al., 2004), lactate dehydrogenase (Chagunda et al., 2006), udder temperature (Berry et al., 2003), udder conformation (Rupp and Boichard, 1999)are promising measures to add to SCC, which has been the traditional breeding value used to select for mastitis resistance. To predict the risk of ketosis in dairy cows, beta hydroxybutyrate in

addition to acceleration in milk yield and body fatness at calving could be useful (Nielsen et al., 2005)as well as mid-infrared spectroscopy anlaysis of milk.

Genomic reference populations in Australia and Ireland (cattle)

Australia

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 4000 Jersey cows (from commercial herds). The cows were selected using a scoring system that quantified their contribution to breeding values. So, cows with multiple lactations of complete data (production, fertility, conformation) scored highly. In April 2012, the females were added to the male reference population (at the time the reference populations were around 4,000 Holstein males and 1,000 Jersey males) yielding a 4-8% improvement in the reliability of breeding values depending on trait. In fact, adding females to the Australian Jersey reference population led to the publication of genomic breeding values for Jerseys for the first time. Building on the success of these projects, a new genomic information nucleus (known as Ginfo) has been established. This time 100 herds that contribute the most data to breeding values have been invited to join the project. The herds were identified using almost exactly the same scoring system that we derived for the 10,000 Holstein cow genomes project and Jernomics; except that the herds with the highest mean score were selected instead of high scoring individuals.

Ireland

The genomic reference population in Ireland consists exclusively of 4,962 Holstein-Friesian AI sires. This has results in almost a doubling in the reliability of genetic evaluations compared to the traditional genetic evaluations. Retrospective analysis of 135 sires now with daughter proofs that were genomically tested in 2009 to 2011 shows that the genomic evaluations are up to 29% more accurate than the parental average. Research is underway into a one-step method which incorporates cow phenotypic (and genotypic) information into the genomic predictions. Research is also underway to implement an across-breed genomic evaluation.

Research into genomic evaluations in beef cattle is underway in Ireland. Approximately 120,000 beef cows will be genotyped in 2014. This is part of a national scheme where 15% of the cows in a herd that signs up to the initiative will be genotyped. Like in Australia, cows will be selected

for genotyping on the basis of information content (i.e., number of lactations within contemporary groups but also divergence in phenotypes after adjustment for systematic environmental effects), breed composition and genetic diversity (i.e., restriction on the number of progeny per sire). Genotyping will be undertaken using the custom built international dairy and beef genotyping platform.

International reference populations

Gold standard measurement of phenotypes such as feed intake are likely to only be measured in research populations. Using within country reference populations, the accuracy of genomic prediction of feed intake and residual feed intake (RFI) in Australia, UK and the Netherlands range between 0.35 and 0.4 (de Haas et al., 2012, Pryce et al., 2012). However, by using a multi-trait model that assumes dry matter intake is a different trait in each country (Australia, the Netherlands and UK), the accuracy increased by 5.5% compared to the univariate analysis (single country analyses) (de Haas et al., 2012). The success of this collaboration has led to the establishment of the global dry matter initiative (gDMI), where dry matter intake data of almost 9000 cows and heifers from 9 countries has been combined and genotypes from around 6000 animals with phenotypes have been imputed to high density (around 620k SNP) (Pryce et al., 2014). The next stage of this project is to estimate the accuracy of genomic prediction for each member country using the combined reference population.

Khansefid et al.(2014)examined the prospect of using a combined beef and dairy genomic reference populations (from research farms) to increase the accuracy of genomic prediction of RFI. They found that small increases in the accuracy of genomic prediction of RFI in dairy can be achieved by including beef cattle in the reference population. They also found that the age of measurement of phenotypes and diet fed could impact on accuracy of genomic prediction. The largest improvement of accuracy was observed when Angus cattle and Holstein heifers were combined; the Angus and Holstein animals used in the experiment were measured at about the same age and were fed a similar diet.

Conclusions

It is likely, that for traits that require expensive or specialist equipment, genomic reference populations that comprise animals from multiple research farms will become more common. Evidence so far suggests that there is value in combining genomic reference populations from research populations of different countries to: 1) minimize the costs for individual participants and 2) increase the accuracy of genomic prediction. Until the cost of phenotyping becomes cheaper and less complex, for traits such as feed efficiency and methane emissions, these are likely to remain in the domain of research organizations. For other easier and cheaper to measure phenotypes, there is likely to be an emergence of groups of commercial herds that are invited to become part of dedicated female genomic reference populations organized nationally (or by breeding companies). These farmers could be provided with software to aid measurement of complex phenotypes, the incentive for the farmers is likely to be the gain they could make in using the data they collect for more precise management of their herds.

Acknowledgements

Jennie Pryce acknowledges funding from Dairy Futures Cooperative Research Council and Donagh Berry acknowledges funding from the Irish Department of Agriculture, Food and Marine, research Stimulus Fund project 13/S4/96, RAPIDFEED and the Marie Curie project International Research Staff Exchange Scheme SEQSEL .

References

Berry, R., A. Kennedy, S. Scott, B. Kyle, and A. Schaefer. 2003. Daily variation in the udder surface temperature of dairy cows measured by infrared thermography: Potential for mastitis detection. Canadian journal of animal science 83(4):687-693

.

Chagunda, M. G., T. Larsen, M. Bjerring, and K. L. Ingvartsen. 2006. L-lactate dehydrogenase and N-acetyl-β-D-glucosaminidase activities in bovine milk as indicators of non-specific mastitis. Journal of dairy research 73(04):431-440.

Daetwyler, H. D., B. Villanueva, and J. A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PloS one 3(10):e3395.

de Haas, Y., M. P. Calus, R. F. Veerkamp, E. Wall, M. P. Coffey, H. D. Daetwyler, B. J. Hayes, and J. E. Pryce. 2012. Improved accuracy of genomic prediction for dry matter intake of dairy cattle from combined European and Australian data sets. J Dairy Sci 95(10):6103-6112.

Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. Genetica 136(2):245-257.

Gonzalez-Recio, O. C., M.P.; Pryce, J.E. 2014. On the value of phenotypes in genomic selection. Journal of Dairy Science.

Khansefid, M., J. Pryce, S. Miller, and M. Goddard. 2014. ACCURACY OF GENOMIC PREDICTION FOR RESIDUAL FEED INTAKE IN A MULTI-BREED CATTLE POPULATION. Pages 298-302 in Proc. Proc. Assoc. Advmt. Anim. Breed. Genet.

Nielsen, N. I., N. C. Friggens, M. G. Chagunda, and K. L. Ingvartsen. 2005. Predicting Risk of Ketosis in Dairy Cows Using In-Line Measurements of< i> β</i>-Hydroxybutyrate: A Biological Model. Journal of dairy science 88(7):2441-2453.

Norberg, E., H. Hogeveen, I. R. Korsgaard, N. Friggens, K. Sloth, and P. Løvendahl. 2004. Electrical conductivity of milk: ability to predict mastitis status. Journal of dairy science 87(4):1099-1107.

Pryce, J. and H. Daetwyler. 2012. Designing dairy cattle breeding schemes under genomic selection: a review of international research. Animal Production Science 52(3):107-114.

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 body weight in growing heifers using 625,000 single nucleotide polymorphism markers. J Dairy Sci 95(4):2108-2119.

Pryce, J. E., J. Johnston, B. J. Hayes, G. Sahana, K. A. Weigel, S. McParland, D. Spurlock, N. Krattenmacher, R. J. Spelman, E. Wall, and M. P. Calus. 2014. Imputation of genotypes from low density (50,000 markers) to high density (700,000 markers) of cows from research herds in Europe, North America, and Australasia using 2 reference populations. J Dairy Sci 97(3):1799-1811.

Rupp, R. and D. Boichard. 1999. Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins. Journal of dairy science 82(10):2198-2204.

Rutten, C., A. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. < i> Invited review:</i> Sensors to support health management on dairy farms. Journal of dairy science 96(4):1928-1952.