Application of Genomic Selection in the New Zealand Dairy Cattle Industry

R.J. Spelman**,** *M.D. Keehan, V. Obolonkin, A.M. Winkelman, D.L. Johnson and* **Bevin Harris***.*

Livestock Improvement Corporation (LIC), New Zealand
LIC, Private Bag 3016, Hamilton 3240, New Zealand.

Abstract

The New Zealand (NZ) dairy industry has used genomic information to enhance its genetic evaluations for the last 13 years. The first 10 years provided little benefit to farmers and breeding companies. Sequencing of the bovine genome in 2006 generated a pool of SNPs that have been commercialised and subsequently used by the bovine research community. The improvement of the accuracy obtained by including genomic information on thousands of animals in the national evaluation system is sufficient to predict that NZ's dairy breeding programme will be revolutionised by genomic technology. The genomically-enhanced breeding (GEBVs) of one-year-old bulls are more reliable than BVs based on parent average, thus allowing them to be reliably selected and used in the national herd. Traditionally, use of young bulls was limited and bulls were not used extensively until they were five years old when the more reliable progeny test results became available. Using young sires, as opposed to progeny-tested sires, in the breeding programme dramatically reduces the generation interval, thereby increasing the rate of genetic gain by 40%-50%. Young sires haven been marketed on their GEBVs in NZ for the last four years. Initial results have shown that the genomic estimates were over-estimated. Adjustments have been introduced into the national evaluation to reduce the bias. Statistical tool development, genotype exchanges and genotyping of cows have been undertaken to increase the accuracy and reduce the bias of the genomic predictions. In addition, sequencing of the dairy cattle population has just commenced in an effort to further improve the genomic predictions and also to detect causative mutations that underlie traits of economic performance.

Introduction

LIC has been investing in DNA technology since the early 1990s. The first application of DNA information in the LIC breeding scheme was for parentage testing in the mid 1990s. At the same time, the detection of quantitative trait loci (QTL) for dairy cattle had begun (Georges et al., 1995). QTL discovered in the LIC research programme were used via marker assisted selection (MAS) in 1998 and 1999 until it was determined that the cost of the programme was greater than its economic return (Spelman, 2002).

Meuwissen et al. (2001) first proposed the use of dense marker maps in a genomic selection (GS) setting. In 2006, the sequencing of the bovine genome by the international consortium was completed (Kappes et al. 2006). The sequencing generated a large number (one million plus) of single nucleotide polymorphisms (SNPs). The cost of genotyping dropped from NZ$2.50 per marker for a microsatellite to less than one NZ cent per marker when tens of thousands of SNP are genotyped in parallel. The technology shift to large-scale SNP genotyping allowed the application of the theory that Meuwissen et al. (2001) proposed. This paper describes the use of genomic information in the LIC dairy cattle breeding scheme.

Material and Methods

Genotyping. Genotyping was done using two different marker panels; Illumina 50K panel and the Illumina 777K panel (high density (HD) panel). In total, approximately 23,000 animals have been genotyped using the 50K panel and 3,000 animals using the HD panel. The genotyped population includes all of the sires that have been progeny tested since 1995, and half of the sires that were progeny tested in the previous 10 years. This group is made up of 8160 Holstein-Friesian (HF) sires, 2861 Jerseys (JE) sires and 723 KiwiCross™ (KX). KiwiCross™ is the name given to crossbred animals that are less than 87.5% of one of the main dairy breeds. The majority of the females that have been genotyped are daughters of sires that were progeny tested in the last five years. All of the 50K genotypes were imputed to HD using Beagle software (Browning and Browning 2009).

Genetic evaluation. Genetic evaluation of the animals is done using a two-step process that incorporates information from the genomic relationship matrix (GRM) (Harris and Johnson, 2010) into the traditional BVs. The first stage involves the prediction of genomic breeding values for genotyped individuals. Because not all ancestors of genotyped animals are genotyped, a selection index procedure is used to blend genomic predictions with traditional ancestral information that is lost between the process of deregression of the national breeding values and subsequent re-estimation using the genomic relationship matrix, thereby producing genomically enhanced BVs (GEBVs). Finally, the GEBVs are filtered through to non-genotyped descendants using a regression procedure.

Application of Genomic Selection

Commercial use of genomically selected bulls. Genomically-evaluated two- and three-year old bulls were first released commercially in New Zealand in 2008 with the creation of genomically-selected bull teams. Since 2010, yearling bulls have also been selected for commercial use based on their GEBVs. The genomically selected teams are marketed as "DNA Proven" whereas the teams with sires that are progeny tested are marketed as "Daughter Proven". The DNA proven teams have a larger number of bulls (25) compared to the daughter proven teams to reduce the risk associated with using bulls that have reliabilities that are lower than that of the progeny-tested bulls. The DNA proven teams (separate team for each breed) are sold to the market at a NZ$5 premium over that of the daughter proven teams to reflect the genetic superiority of the young bulls. The average cost of a daughter proven straw of semen is approximately NZ$18. The average superiority of the DNA proven team compared to the daughter proven teams is approximately 25 BW points, which is equivalent to 2½ years of genetic improvement. The use of DNA proven sires has increased over the last 5 years, increasing from15% of inseminations in 2008 to just over 40% in 2011.

Breeding scheme. Prior to the introduction of the GS scheme, LIC had been progeny testing 300 bulls per annum. In 2008, the number of bulls progeny tested was reduced to approximately 160 and has remained at this size since then. The reduction in size stems from the ability to screen a large number of young bulls based on their GEBVs. The 160 bulls are selected from approximately 2000 to 3000 bulls with genomic information. The bulls have been genotyped on a number of different panels; a custom-made 384 SNP panel for in the first year and a 50K panel for the next four years. The low-density (LD, approximately 12K) panel will likely be used in 2012.

Performance of genomics. New Zealand Animal Evaluation Limited (NZAEL), which oversees the national dairy genetic evaluation, first included genomic information in the national evaluation in 2009. The method used is described by (Harris and Johnson, 2010). Subsequent analysis by NZAEL in 2010 identified that the genomic estimates were biased upwards. At that time NZAEL advised removing the bias by subtracting 15-30 BW units from the GEBVs. In 2011, the genomic bias adjustment factor was further re-fined so that it accounted for the breed, ancestry and genetic merit of the animal. The majority of the adjustment factors now range from 20 to 40 BW units. Also in 2011, NZAEL decided to remove genomics from the main index (Breeding Worth) but continue to publish a secondary index that includes genomic information.

The first two crops of DNA-proven sires, used in 2008 and 2009, have now received their progeny test proofs. The initial GEBVs of these sires were found to be over-estimated and the use of genomically-proven sires did not result in the expected genetic advantage. Therefore, as a gesture of appreciation to the early adopters of genomic evaluation, LIC credited the $5 premium that the farmers paid. Since 2009 the genomic adjustments have been in place and thus the level of over-estimation observed in the 2008 and 2009 teams should not occur in the forthcoming years.

Improving genomic estimation. LIC is committed to the use of genomic information in its breeding programme. To increase confidence in the technology, there has been a concerted effort to improve the accuracy and reduce the bias of the genomic evaluations. Increasing the number of genotyped animals has been the main thrust to increase the accuracy of evaluation ~~and~~ statistical methodology developments to decrease the genomic bias.

All of the sires that have been progeny tested in the last 15 to 20 years have been genotyped and thus to increase the number of genotyped animals there are two options: genotype exchanges and genotyping females. LIC has undertaken both of these options. Genotypes have been exchanged with Ireland,

Australia and more recently with CRV Ambreed (another NZ AI company). Genotype-by-environment interactions have reduced the utility of the genotypes from both Australia and Ireland. The exchange of NZ progeny-tested bulls with CRV resulted in increases of approximately 3% in reliability for the milk production traits. To date, LIC has genotyped 25,000 females from the NZ dairy herd and is in the process of genotyping another 25,000 cows. Most of the females that have been genotyped are the daughters of progeny-test sires. Integration of the female genotypes into the genetic evaluation has been challenging. Recent statistical advancements have seen extra accuracy being generated from these animals.

Statistical tool development has focused on the single-step method (Misztal et al. 2009) for the NZ multi-breed population.. In a recent analysis, the mixed model equations of the single-step method included pedigree and genomic information for all bulls born in 2006 and earlier. Daughter records were included for sires born prior to 2005 but not for sires born in 2005 and 2006. Hence, GEBVs for the latter group were obtained using only on the genomic and pedigree information. The GEBVs from this group were compared to the BVs obtained from the traditional progeny test analysis. Comparisons of the progeny test evaluations to those obtained from to the single-step, two-step and parent average analyses are shown in Table 1 and 2. Two aspects of the prediction are considered; accuracy (correlation squared) between progeny test BVs and GEBVs and inflation – the slope of the regression of progeny test BVs on GEBVs. A slope of less than one indicates inflation of the GEBV.

Table 1 shows the accuracy of prediction of the parent average BVs and GEBVs obtained from the two-step and single-step methods. On average, a 10% increase in realized reliability is achieved between parent average and genomic proofs. The average increase in the reliability of the single-step over the two-step method is 5%.

Table 1: Accuracy of evaluation obtained from parent average, two-step and single-step genomic prediction.

Evaluation Method	Trait	KiwiCross	HF	Jersey
Parent average	Milk volume	0.38	0.26	0.23
	Fat	0.28	0.33	0.23
	Protein	0.46	0.35	0.36
Two-Step				
	Milk volume	0.51	0.34	0.39
	Fat	0.32	0.31	0.36
	Protein	0.63	0.46	0.40
Single-step				
	Milk volume	0.53	0.40	0.43
	Fat	0.35	0.43	0.39
	Protein	0.64	0.51	0.48

Table 2 shows the estimates of inflation of the parent average BVs and the GEBVs obtained from the two-step and single-step. The GEBVs obtained from the single-step method have lower bias than those obtained from the two-step method. The use of the single-step method removes most of the bias for the HF and KX sires, but not the KX sires.

Table 2: Average inflation estimates for the three milk production traits for each breed with the different methodology and datasets.

Evaluation method	KiwiCross	HF	Jersey
Parent average	0.95	0.85	0.80
Two-step	0.76	0.70	0.68
Single-step	0.99	0.96	0.86

Future direction. LIC will continue to invest into improving the accuracy of genomic evaluation. This currently focuses on increasing the number of genotyped animals and increasing the density of genetic markers. This has been through the use of HD marker panels and, more recently, through whole genome sequencing. Integrating biological information to identify markers that have causative effects rather than anonymous markers is also an area that will be pursued.

Conclusion

The introduction of genomic information into LIC dairy cattle breeding scheme has been a steep learning curve over the last five years. Dairy farmers have utilised the new technology and to date have not benefited to the degree that was expected, which is a common situation with new technology. Ongoing investment will be required to continue to improve and maintain the accuracy of genomic selection. With this investment it is expected that the accuracy of genomics will continue to improve and breeding schemes will utilize genomic information further at the expense of progeny testing.

References

Browning B.L. and Browning S.R. (2009) Am J Hum Genet 84:210-223.

Georges, M., Nielsen, D., Mackinnon, M., *et al.* (1995). *Genetics* 139:907-920.

Harris, B.L. and Johnson, D.L. (2010). *J. Dairy Sci.* 93:1243-1252.

Kappes, S.M.; Green, R.D. and van Tassell, C.P. 2006. In *Proc. 8th WCGALP*.

Meuwissen, T.H.E.; Hayes, B.J. and Goddard, M.E. (2001). *Genetics* 157:1819-1829.

Misztal, I., Legarra, A. and Aguilar, I. (2009) *J. Dairy Sci*. 92: 4648-4655.