DNA PROVEN BULLS: APPLICATION OF GENOMIC TECHNOLOGY

R J Spelman, P T Gatley and B L Harris
LIC, Private Bag 3016, Hamilton, New Zealand

Summary

Marker assisted selection has been applied in dairy cattle breeding schemes with minor to moderate improvements in genetic gain. The cost effectiveness of it has, at best, been neutral in most cases. The completion of the sequencing of the bovine genome in 2006 has generated a large number of single nucleotide polymorphisms that are now commercially available. In addition, the cost of genotyping single nucleotide polymorphisms is markedly reduced compared to microsatellite genotyping. LIC has undertaken extensive genotyping of the sires that have been progeny tested in the last 20-30 years. Analysis of this data has allowed LIC to identify combinations of markers that allow the genetic evaluation of bulls from their DNA. By utilising this genomic information, semen from two-year-old sires is now being sold to New Zealand dairy farmers. It is expected that selection based on genomic information (genomic selection) will increase the rate of genetic gain in New Zealand by 30-50%.

Introduction

LIC has been investing in DNA technology since the early 1990s. The first application of DNA information being used in breeding schemes was for parentage testing when it transferred from blood protein polymorphisms to microsatellite markers in the mid 1990s. At the same time, the detection of quantitative trait loci (QTL) for dairy cattle had just commenced (Georges et al. 1995) and LIC was at the fore-front. A flood of QTL have been identified for dairy cattle traits but there are only three instances where the underlying causative mutation for the QTL has been identified; chromosome 14 – DGAT1 (Grisart et al., 2002), chromosome 20 – GHR (Blott et al. 2003) and chromosome 6 (Ron et al 2006). LIC were the researchers involved in the chromosome 14 and 20 work and these genes have been utilised and marketed under the brand names of Quantum and Optimum respectively.

The utilisation of the QTL in dairy cattle breeding schemes via marker-assisted selection (MAS) commenced at LIC in 1998 using six QTL that had been identified for protein, fat and milk volume. The MAS breeding scheme relied on the generation of multiple full-sib sons from

Notes:

a given sire and dam using MOET or IVP reproductive programmes. The resulting male offspring were genotyped, with the sons that received the desirable alleles for the QTL selected to enter progeny testing.

The reproductive performance of the donor cows was poor and very few of the families had enough sons to allow within-family marker-assisted selection. After two years of poor reproductive performance the within-family MAS was abandoned. LIC recommenced MAS in 2003 after the identification of both DGAT1 (Quantum) and GHR (Optimum) genes. All of the bull dams and bulls entering the progeny testing scheme were genotyped for the two genes for the next four years.

In 2006 the sequencing of the bovine genome by the international consortium was completed (Kappes et al. 2006). The sequencing generated a large number (million plus) of single nucleotide polymorphisms (SNPs) compared to the thousands of the previous type of marker (microsatellites). In addition to the greater number of SNPs, the cost of genotyping moved from NZ$2.50 per genotype for a microsatellite to less than one NZ cent when tens of thousands of SNP are typed in parallel. The technology shift to large scale SNP genotyping has had, and will have, a major effect on the utilisation of markers in dairy cattle breeding schemes.

Genomic selection

Meuwissen et al. (2001) first proposed the use of dense marker maps in a genomic selection (GS) setting. Through simulation they found that that the reliabilities of the estimating breeding values from genomic information were 0.75-0.85. At the time of the paper being published, cost estimates to undertake the work on 2000 sires with 4000 microsatellites was approximately NZ$20 million. With the sequencing of the bovine genome and the commercial SNP panels being developed, the cost of this experiment is a now in the region of $1m-$2m.

LIC has completed the genotyping of approximately 4500 sires that have been progeny tested over the preceding 30 years. LIC had the foresight to store DNA from every sire that was progeny tested since 1980. This has enabled LIC to genotype sires that were the best and the worst of their progeny test cohort and thus evaluate markers across the genetic range. Genotyping has been undertaken on the Illumina 50K SNP panel.

The dataset was split into two parts: i) Research dataset and ii) Validation dataset. The Research dataset was all bulls that were progeny tested prior to 2002. There were 2450 sires in the Research dataset. The genotypes for the 2450 sires were analysed to identify which SNPs affected the estimated breeding values for the sires for 25 different traits. The objective of the Research dataset was to find a subset of SNPs that predict (or explained) the genetic merit of the sires. Once the effects for the SNPs had been estimated they were then tested on the Validation dataset. There were 850 sires in the Validation dataset and they were the bulls that had been progeny tested between 2002 to 2004. The Validation dataset was used to “test or validate” the

SNP effects that had been estimated from the Research dataset. The validation was undertaken by comparing the progeny test BVs with the genomic BVs estimated for the relevant sires. The degree of accuracy of the DNA-based BVs was measured by their correlation with the progeny test BVs. The correlations varied from 0.45 to 0.60 for the production and non-production traits for the Holstein-Friesian breed (Table 1).

Table 1. Correlations between DNA-based and progeny test BVs in the validation population for the Holstein-Friesian breed

Trait	Correlation
Protein yield	0.57
Milk fat yield	0.47
Milk volume	0.60
Live weight	0.51
Fertility	0.59
Somatic cell score	0.54
Total longevity	0.60
Shed temperament	0.47
Farmer opinion	0.45
Udder overall	0.45

Currently there are three sources of information that can be used in the estimation of breeding values; parental information, own performance and progeny performance. DNA is now the fourth source of information. A dairy bull of one year of age has only information from his parents for dairy related traits. The parental information is combined with the DNA information to estimate a genomic breeding value for the bull. Approximately 50-60% of the genomic breeding value is parental information with the balance being DNA information. Through the research it has been identified that there is an overestimation in the parental information. Reasons for this are potentially epistatic genetic effects (positive interactions between genes that are not transmitted to the next generation and unintentional preferential

Notes:

treatment of bull dams entering IVP or ET reproductive programmes. To account for this bias the bull's evaluations were reduced by approximately 18 BW.

This was undertaken in New Zealand in 2008 for the first time with the release of the genomically selected teams marketed as "DNA Proven". These teams included bulls that were 2 or 3 years of age and thus did not have lactating daughters. The reliabilities for the genomic breeding values for the DNA Proven bulls are approximately 50-55%. This is in contrast to a reliability of 75%-85% for a progeny tested ("Daughter Proven") bull and 30-35% for a bull entering progeny test.

During the 2008 mating season the HF DNA Proven team was withdrawn from the market. This was driven by further information from the 2005 SPS where the DNA proof for the HF team did not predict the daughter proof as well as had been seen in the two previous SPS years. Given that the HF DNA Proven team had an advantage of only approximately 10 BW it appeared that there was a significant probability that the DNA Proven team could have a lower BW than the Daughter Proven team. Over the 2008 season the three DNA Proven teams (Holstein Friesian, Jersey, KiwiCross^{TM}) undertook approximately 15% of the total inseminations. More accurate genetic evaluations for the individual bulls will occur in the 2009-10 and 2010-11 seasons when their daughters begin their first lactations. The latest genomic evaluations indicate that the bulk of the DNA Proven inseminations (60% of which were KiwiCross^{TM}) will be found to have delivered an advantage over the Daughter Proven option.

The BW reliability of the DNA Proven sires is approximately 55%. For a sire that has an estimated BW of 250 with a reliability of 55%, it is expected that 95 times out of 100 his true genetic merit will be between 160 and 340. For a Daughter Proven sire with an 85% reliability the range is between 200 and 300. Utilising the bulls in teams of 15-18 reduces the effect of the lower reliability of the DNA Proven sires. Team reliabilities of 97% are achieved. This is reflected in that 95 out of 100 times the true mean of the team will be ± 23 BW points of that estimated from the DNA proofs. In contrast for the Daughter Proven team the true mean of the team will be ± 13 BW points from that estimated from progeny testing. Therefore farmers using teams of DNA Proven bulls can expect to see greater volatility in the proofs of the DNA team compared to the Daughter Proven team. LIC advocates if the farmers are uncomfortable with a change of up to 25 BW points in the team average then the Daughter Proven team used solely, or in conjunction with the DNA team, may be better suited to the farmer's risk profile. In June 2009 the national (NZ Animal Evaluation Ltd) genetic evaluation will include genomic information for the first time. Prior to this LIC and Ambreed have been calculating in-house breeding values.

The LIC breeding scheme, which has been based on progeny testing 300 bulls per annum, has been altered to incorporate the DNA information. In 2009, LIC screened over 1000 bulls

based on their DNA profile and then selected the best 150-180. These bulls enter progeny test as one year olds and then their semen will be sold commercially as two-year-old bulls. The bulls will be retained until they receive their progeny test evaluation and semen will be sold to farmers who prefer to purchase genetics based on progeny test information. With the shorter generation interval due to bulls being used as two year olds instead of five year olds, it is estimated that the rate of genetic gain in the New Zealand population could increase by 30-50%.

The current SNP technology is based around genotyping of 50,000 SNP markers. There are expectations that there will be a 500,000 marker panel available for commercial genotyping by the end of the year, and there are also indications that there could be a million marker panel in the next 12 months. This progression follows what occurred in the human genomics area where they have gone from 100,000 SNPs to one million over a 12-18 month period. In addition to the technology change with genotyping, the major shift is occurring with sequencing. When the bovine genome was sequenced in 2006-08, it cost approximately $US 50 million. If repeated this year it is expected that the sequencing could be completed for less than $US 100,000 and by 2012 there are indications that a genome could be sequenced for $US 1,000. It can be expected that sequencing will become part of the selection process for dairy breeding schemes in the next 10 years.

The utility of DNA in dairy cattle breeding schemes has now reached the level of accuracy that enables dramatic changes and improvements to breeding schemes. With denser marker maps becoming available in the coming years, more sophisticated statistical tools, and in the longer term, with sequencing, it is expected that the level of accuracy from genomic information will continue to improve.

REFERENCES

Blott S, Kim, J J, Moisio S, Schmidt-Kuntzel A, Cornet A, Berzi P, Cambisano N, Ford C, Grisart B, Johnson D, Karim L, Simon P, Snell R, Spelman R, Wong J, Vilkki J, Georges M, Farnir F and Coppieters, W. 2003. Molecular Dissection of a Quantitative Trait Locus: A Phenylalanine-to-tyrosine Substitution in the Transmembrane Domain of the Bovine Growth Hormone Receptor is Associated with a Major Effect on Milk Yield and Composition. *Genetics* 163:253.

Notes:

Georges M, Nielsen D, Mackinnon M, Mishra A and Okimoto, R et al. 1995. Mapping Quantitative Trait Loci Controlling Milk Production in Dairy Cattle by Exploiting Progeny Testing. *Genetics* 139:907.

Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Cambisano N, Mni M, Reid S, Spelman R, Georges M and Snell R. 2002 Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine DGAT1 Gene with Major Effect on Milk Yield and Composition. *Genome Res*. 12(2):222.

Kappes S M, Green, R D and Van Tassell C P. 2006. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production.

Meuwissen T H E, Hayes B J and Goddard M E. 2001. Prediction of Total Genetic Value Using Genome-wide Dense Marker Maps. *Genetics* 157:1819.

Ron M, Cohen-Zinder M, Peter C, Weller J I and Erhardt G. 2006. Short Communication: A Polymorphism in ABCG2 in Bos Indicus and Bos Taurus Cattle Breeds. *J Dairy Sci.* 89(12):4921.