Proc. Assoc. Advmt. Anim. Breed. Genet. Vol 14

REVIEW OF QTL MAPPING IN THE NEW ZEALAND AND DUTCH DAIRY CATTLE POPULATIONS

R.J. Spelman[1], W. Coppieters[2], B. Grisart[2], S. Blott[2] and M. Georges[2]

[1]Livestock Improvement Corporation, Private Bag 3016 Hamilton, New Zealand
[2]Department of Genetics, Faculty of Veterinary Medicine, University of Liege, 4000-Liege, Belgium

SUMMARY

A joint venture between Livestock Improvement and Holland Genetics where monetary, biological and data resources were pooled, has proven successful in identifying QTL through work undertaken at the University of Liege. The majority of the work has been undertaken with a granddaughter design where 31 families have been genotyped with over 300 microsatellites and single nucleotide polymorphisms. Four QTL for milk production are reported in this paper; on chromosomes 6, 14, 16 and 20. Fine mapping has been undertaken for these QTL, with the QTL identified on chromosome 14 being fine mapped to a very defined region. QTL for non-production traits have been identified but differ between the New Zealand and Dutch dairy populations. The milk production QTL have been utilised in both of the companies breeding schemes via marker assisted selection.
Keywords: Dairy cattle, QTL.

INTRODUCTION

In 1994 Livestock Improvement Corporation (LIC) and Holland Genetics (HG) entered into a joint venture to detect and utilise quantitative trait loci (QTL) in dairy cattle. The majority of the molecular work has been undertaken at University of Liege, Belgium, by the group headed by Michel Georges, who had experience in this area from his work at Genmark (Georges *et al.* 1995).

The primary objective of the project was to detect QTL for milk production, and utilise the QTL within funding partners respective breeding schemes. This paper reviews the experimental strategy used, statistical techniques, QTL identified and the implementation of those QTL within the respective breeding schemes.

MATERIALS AND METHOD

Granddaughter design. The granddaughter design (Weller *et al.* 1990) was the central experimental design. In total there were 31 families that had an average of 39 sons each, with a range of 11-153. There were 22 grandsire families (all Holstein-Friesian) from the Netherlands (NL) and 9 from New Zealand (NZ) (7 Holstein-Friesian and 2 Jersey). All of the NZ families had DNA from all sons progeny tested, whereas some of the NL families only had selected sons. This was due to semen only being retained from higher genetic merit sons for a number of years. Initially selected sons were removed from the dataset (Spelman *et al.* 1996) based on the work of MacKinnon and Georges (1992). However, the work of Coppieters *et al.* (1998b) showed that statistical power increased when these animals were included in the design, and thus all animals have been retained in the dataset for subsequent analysis.

Daughter design. In addition, two NZ daughter families (Holstein-Friesian) were used. Both of the sires for the daughter design were also represented in the granddaughter design. The two sires had 914 and 1018 randomly selected daughters. Selective genotyping (Darvasi and Soller 1992) was applied to the daughter design. The daughters were sampled as lactating cows 3 years before genotyping commenced, resulting in the number of lactations ranging from 1 to 4 for the daughters.

Protein was used as the trait to select the animals to genotype. Approximately 17% of daughters (168) from each tail of the trait distribution were selected. This selected fraction has 85-90% of the statistical power of genotyping all of the animals. Daughters that failed paternity testing were excluded resulting in 307 for sire one and 299 for sire two.

Grand^2-daughter design. The pedigree structure of the NL and NZ granddaughter design had many links between the 31 half-sib family structures. This was exploited by the grand^2-daughter design where the contrast between two homologs of a given founder sire is measured in its maternal or paternal grandsons (Coppieters *et al.* 1998c). In essence, the transmission of a grandsires alleles are followed through ungenotyped daughters to genotyped sires, who are already in the dataset (ie sons of other grandsires). In the dataset there were 731 maternal grandsons from 18 foundation sires and 277 paternal grandsons from three founder sires. The number of grandsons per maternal grand-sire ranged from 1 to 222 and paternal grandsons 22 to 226. This design was used to confirm QTL that been detected in the granddaughter and daughter designs without having to undertake further genotyping.

Genotyping. Approximately 300 microsatellites and single nucleotide polymorphisms (SNPs) have been used in the whole genome scan. Microsatellite genotypes were produced by autoradiography after incorporation of dCTP^{32} as described (Georges *et al.* 1995) or using Applied Biosystems "four dye/one lane" technology and instrumentation (ABI 373 or 377). SNPs were genotyped using a PCR/oligation nucleotide assay (Landegren *et al.* 1988) with electrophoretic separation of the ligation products using an automatic ABI373 sequencer.

Map construction. Two-point lodscores were generated with ANIMAP (Georges *et al.* 1995) and subjected to Morton's test for heterogeneity in order to identify potential genotyping errors (Ott 1992). Multipoint linkage analyses were performed across families using CRIMAP (Lander and Green 1987) to determine the most likely order, followed by ANIMAP to determine the most likely recombination rates between adjacent markers.

Phenotypes. Milk production traits; fat yield, protein yield, milk volume and fat and protein percent were analysed for all three experimental designs. Daughter yield deviations (van Raden and Wiggans 1991) were used for the granddaughter design and the grand^2-daughter design for the fore-mentioned traits. In the daughter design yield deviations (YD) were used that were adjusted for the number of lactations for each cow to ensure variation of YD was constant over all lactations. Non-production traits were analysed separately for both the NZ (Spelman *et al.* 1999) and NL (Schrooten *et al.* 2000) populations. In both instances breeding values were used as the trait measurement. In the NZ population, 3 grandsires were analysed for milk colour (Davis *et al.* 1998) using breeding values.

Statistical analysis. A number of statistical techniques have been used to analyse the dataset. Two regression techniques were used extensively in the genome scan for the granddaughter design; least square regression (Knott *et al.* 1996) and non-parametric regression (Coppieters *et al.* 1998a). Both methods were used in an across family setting, where statistical values calculated for each individual half sib family were summed over all families. For fine mapping of areas where QTL had been detected, identity by descent mapping (Riquet *et al.* 1999) and linkage disequilibrium mapping (Farnir *et al.* 2001) were used.

In the daughter design two techniques were used; least squares regression with adjustment for selective genotyping (Bovenhuis and Spelman 2000) and a Monte Carlo Expectation Maximisation (MCEM) method (Johnson *et al.* 1998). Both techniques were used in a within family setting. In the grand^2-daughter design a rank-based non-parametric statistic was used (Coppieters *et al.* 1998c) in an across family setting.

Threshold values were calculated for most of the above statistical methods using permutation (Churchill and Doerge 1994). Stringent threshold values, termed experimentwise, were used where repeat hypothesis testing for independent traits and chromosomes were accounted for. Nominal threshold values were used to identify segregating families at identified QTL. Confidence intervals for QTL were calculated using bootstrapping (Visscher *et al.* 1996).

Confirmation of QTL on an independent sample was undertaken before being utilised through MAS or before commencing fine-mapping. In this study the grand^{2}daughter design and to a lesser extent the daughter design was used for this role.

RESULTS

Milk production. Six QTL have been identified that are significant at experimentwise threshold values. Four of these QTL are reported in this paper, with the other two QTL being classed as commercially sensitive. The QTL on chromosome six was first reported in 1996 based on the NL part of the granddaughter design (Spelman *et al.* 1996). The QTL was detected for protein percent (Figure 1). This trait is a ratio of the protein yield divided by milk volume. Therefore the QTL must be acting on one or both of the underlying traits.. There are 3 families segregating on chromosome six. The QTL effects in these 3 families are not consistent, with two families having an effect only on milk volume, and the other family having an increase in protein and fat yield, and no change in milk volume. Two families are significant near the position reported in Spelman *et al.* (1996) and the other family 40cM towards the casein cluster. However, in the across family analysis only the former QTL position is significant at the experimentwise threshold level. The confidence interval covers some 80 cM over the chromosome. The additive effects are some 150*l* for milk volume for two of the families and 5 kg for protein and 8 kg for fat yield for the other family. Although the 2 QTL model was not significant in this study there have been reports in the literature that would indicate that there is more than one milk production QTL segregating on chromosome six (Mosig *et al.* 2001).

Chromosome 14 harbours the most significant QTL and best defined QTL in this study (Coppieters *et al.* 1998c and Riquet *et al.* 1999). The most significant trait is fat percent where the test statistic is greater than the highest value generated from 50,000 permutations. Of the 31 grandsire families, 7

are segregating, and another 2 are very close to significance. The effects for this QTL are more consistent than those for chromosome six. In all of the significant families, the fat and protein effects are in opposite directions ie. when the QTL effect is positive for fat, the effect on protein and milk volume is negative. The difference in direction of effects for protein and milk is contrary to the genetic correlation between the two traits which is approximately 0.8. The average sizes of the additive effects are approximately 175 *l* for milk volume, 3 kg for protein yield and 4 kg for fat yield. The QTL is segregating in both populations in the granddaughter design and was identified in both families in the daughter design and verified in the grand^2-daughter design.

The confidence interval for this QTL is some 10cM based on the regression techniques applied to the granddaughter design. This has been further refined by applying identity by descent methods (Riquet *et al.* 1999). Riquet *et al.* (1999) genotyped the 7 segregating sires for a high density marker map. The seven chromosomes that increased fat percent were identified to carry a common piece of chromosome that spans some 5cM. The chromosomal position has been further refined through the application of linkage disequilibrium methods (Farnir *et al.* 2001). This QTL region is being investigated with further genetic studies and sequencing of candidate genes in the region.

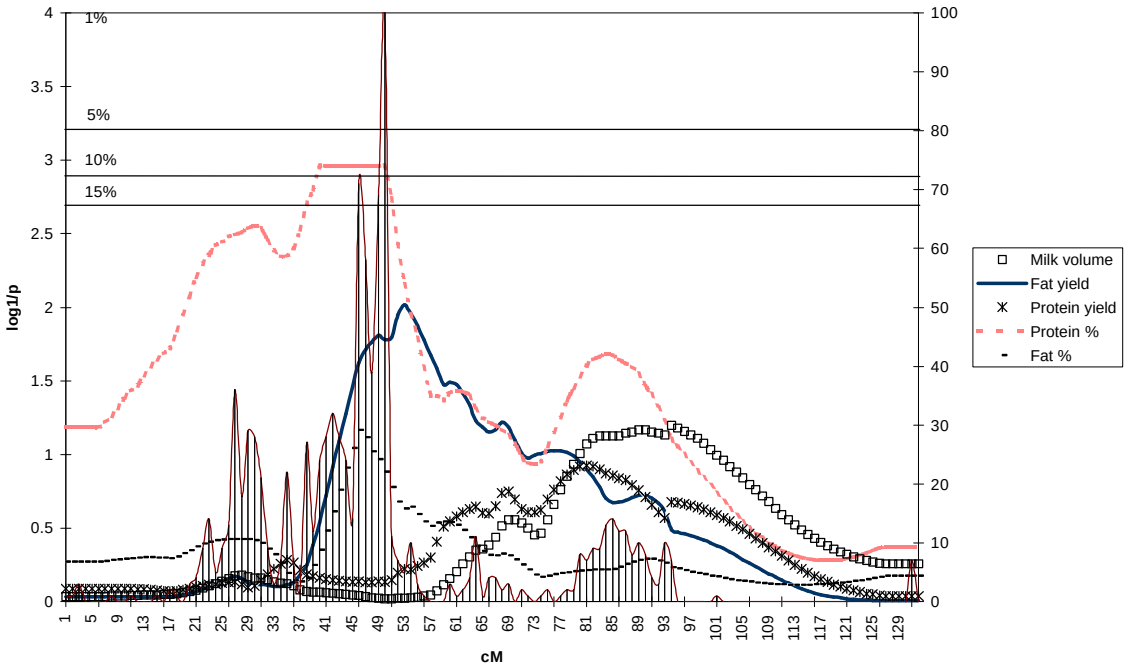

Figure 1. Test statistics for different positions on chromosome six and confidence interval based on distribution of bootstrap values. (The heavy vertical lines show the distribution of the bootstrap values)

The daughter and granddaughter design has identified a QTL segregating for fat percent on chromosome 16. Fat percent is significant at the 5% experimentwise level, but the other traits do not approach significance. When investigating each individual family there does not appear to be any families that are significant, except for one family in the daughter design. The QTL effect in this family is primarily for milk volume. The most likely position for this QTL is near marker BMS1207.

Chromosome 20 has been identified to segregate for a QTL for protein percentage in the granddaughter design and was verified in the grand^2-daughter design (Arranz *et al.* 1998). Three families are near significance for the chromosome. Two families appear to be segregating near BM5004 and the other family segregating near marker TGLA153. The confidence interval for this chromosome covers some 50% of the chromosome. The average effect for the three families is some 200 *l* for milk volume and near zero effects for fat and protein yield. IBD and linkage disequilibrium mapping techniques are being applied to fine map the QTL on this chromosome. Candidate genes are also being investigated on this chromosome.

Non-production. QTL have been detected in both the NZ and NL populations, although not at the same level of significance as that of the milk production traits. In the NZ dataset the most significant QTL was for stature on chromosome 14 (Spelman *et al.* 1999). The most likely position of this QTL is approximately 20-30 cM away from the most likely position for the milk production QTL. In the NL dataset there were 10 QTL that were significant at the genomewise level (less stringent than experimentwise as number of independent traits are not taken into account) (Schrooten *et al.* 2000). The most significant detected was for front udder attachment on chromosome 13. Interestingly all QTL identified in one population have not been confirmed in the other population. Davis *et al.* (1998) identified a QTL for milk colour, that was significant at the 25% experimentwise level on chromosome 23.

Marker assisted selection. Both LIC and HG have implemented the milk production QTL in their respective breeding schemes through marker assisted selection (MAS). The MAS schemes that the companies have implemented are quite different. HG use an animal model setting where they estimate both polygenic and QTL effects, for each animal in their nucleus herd and bulls entering progeny testing. A limited number of bulls have been selected based on this procedure. LIC uses marker information on a within family setting to identify which full sib sons have received the favourable QTL allele from their sire (Spelman and Garrick 1998). This MAS scheme has been used by LIC for bulls progeny tested in 2000 and 2001.

DISCUSSION

The QTL project that involves LIC and HG has been very successful in identifying QTL for milk production and to a lesser extent for non-production traits. There are other QTL that have not been reported. Some of these QTL are commercially sensitive and are being investigated through fine-mapping and the others are not as statistically significant as those reported here. The application of the QTL within the respective companies breeding schemes has not been as successful as one would have hoped. For LIC this has been due to poor reproductive performance in generating multiple male full-sibs that are required in the "bottom up" MAS scheme that is being implemented (Spelman 1998). However, with some potentially successful results from fine mapping this will make MAS

easier to apply as the markers can be used across families instead of the current within family restriction.

REFERENCES

Arranz, J-J., Coppieters, W., Berzi, P., Cambisano, N., Grisart B., Karim, L., Marcq, F., Riquet, J., Simon, P., Vanmanshoven, P., Wagenaar, D. and Georges, M. (1998) *Anim. Genet.***29(2):** 107.

Bovenhuis, H. and Spelman, R.J. (2000) *J. Dairy Sci.* **83:** 173.

Churchill, G.A. and Doerge, R.W. (1994) *Genetics* **138:** 963.

Coppieters, W., Kvas, A., Farnir, F., Arranz, J-J., Grisart, B., MacKinnon, M. and Georges, M. (1998a) *Genetics* **149:** 1547.

Coppieters, W., MacKinnon, M. and Georges, M. (1998b) *J. Herit.* **89(2):** 193.

Coppieters, W., Riquet, J. Arranz, J-J, Berzi, P., Cambisano, N., Grisart, B., Karim, L., Marcq, F., Moreau, L., Nezer, C., Simon, P., Vanmanshoven, P., Wagenaar, D. and Georges M. (1998c) *Mam. Genome*. **9:** 540.

Darvasi, A. and Soller, M. (1992) *Theor. Appl. Genet.* **85:** 353.

Davis, K.L., Spelman, R.J., Coppieters, W., Winkelman, A.M., Garrick. D.J. and Georges, M. (1998) *An. Genetics* **29(Suppl. 1):** 62.

Farnir, F., Coppieters, W. and Georges, M. (2001) *submitted*.

Georges, M., Nielsen, D., MacKinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., Sargeant, L.S., Sorenson, A., Steele, M.R., Zhao, X. Womack, J.E. and Hoeschele I. (1995) *Genetics* **139:** 907.

Johnson, D.L., Jansen, R.C and van Arendonk, J.A.M. (1998) *Genet. Research* **73:** 75.

Knott, S.A., Elsen, J.M. and Haley, C.S. (1996) *Theor. Appl. Genet.* **93:** 71.

Landegren, U., Kaiser, R., Sanders, J. and Hood, L. (1988) *Science* **241:** 1077.

Lander, E. and Green, P. (1987) Proc. Natl. Acad. Sci. **84:** 2363.

MacKinnon, M. and Georges, M. (1992) *Genetics* **132:** 1177.

Mosig, M.O., Lipkin, E., Khutoreskaya, G., Tchourzyna, E., Soller, M. and Friedmann, A. (2001) *Genetics* **157:** 1683.

Ott, J. (1992) Analysis of human genetic linkage. The John Hopkins University Press Ltd., London.

Riquet, J., Coppieters, W., Cambisano, N., Arranz, J-J., Berzi, P., Davis, S.K., Grisart, B., Farnir, F., Karim, L., Mni, M., Simon, P., Taylor, J., Vanmanshoven, P., Wagenaar, D., Womack, J.E. and Georges, M. (1999) *Proc. Nat. Acad. Sci.* **96:** 9252.

Schrooten, C., Bovenhuis, H., Coppieters, W. and J.A.M. van Arendonk (2000) *J. Dairy Sci.* **83:** 795.

Spelman, R.J. (2000) PhD thesis, Wageningen Agricultural University.

Spelman, R.J., Coppieters, W., Karim, L., van Arendonk, J.A.M. and Bovenhuis, H. (1996) *Genetics* **144:** 1799.

Spelman, R.J. and Garrick, D.J. (1998) *J. Dairy Sci.* **81:** 2942.

Spelman, R.J., Huisman, A.E., Singireddy, S.R., Coppieters, W., Arranz, J., Georges, M. and Garrick, D.J. (1999) *J. Dairy Sci.* **82:** 2514.

Van Raden, P.M. and Wiggans, G.R. (1991) *J. Dairy Sci.* **74:** 2737.

Visscher, P.M., Thompson, R. and Haley, C.S. (1996) *Genetics* **143:** 1013.

Weller, J.L., Kashi, Y. and Soller, M. 1990 *J. Dairy Sci.* **73:** 2525.