DETECTION AND ANALYSIS OF A LOCUS AFFECTING MILK CONCENTRATION IN THE US AND ISRAELI DAIRY CATTLE POPULATIONS

M. Ron[1], D. W. Heyen[2], J. I. Weller[1], M. Band[1], E. Feldmesser[1], H. Pasternak[1], Y. Da[3], G. R. Wiggans[3], P. M. Vanraden[3], E. Ezra[4] and H. A. Lewin[2]

[1]Institute of Animal Sciences, A. R. O,. The Volcani Center, Bet Dagan 50250, Israel
[2]Department of Animal Sciences, University of Illinois, Urbana, Illinois 61801, USA
[3]Agricultural Research Service, USDA, Beltsville, MD, USA
[4]Israel Cattle Breeders Association, Pardes Hana, Israel

SUMMARY

DNA microsatellites were used to detect individual loci affecting economically important quantitative traits in dairy cattle via the granddaughter design. Eighteen US Holstein grandsires and 1555 of their sons were genotyped for 30 genetic markers located on 19 of the 29 bovine autosomes. From 16 to 205 sons were genotyped per family. Of 14,650 son genotypes determined, 77% were informative. The genotype data was matched to the bulls' daughter yield deviations for milk, fat, and protein production, fat and protein percentage, productive life, and somatic cell score. The within-family allele effect was significant at $p<0.01$ for TGLA263 on chromosome 3, CSRM60 on chromosome 10, and CSSM66 on chromosome 14. TGLA263 and CSSM66 had significant effects on more than one trait, but the effects on fat percent were greatest for both loci. The effect of CSSM66 on fat percent was significant at $p<10^{-7}$. CSSM66 also had a significant effect on fat percent in the Israel Holstein population ($p<0.001$), which was analyzed by a daughter design of seven families. By maximum likelihood, it was determined that the QTL has a substitution effect of about 0.28% fat, and is probably located 10 to 20 cM from CSSM66 proximal to the centromere.
Keywords: Quantitative trait loci, genetic markers, dairy cattle, granddaughter design.

INTRODUCTION

Many studies have shown that individual quantitative trait loci (QTL) affecting economic traits can be detected and mapped via linkage to genetic markers (reviewed by Weller, 1996). Weller *et al.* (1990) showed that, assuming an appropriate population structure, power per individual genotyped can be increased if sons of sires heterozygous for the genetic marker are genotyped, and the records of their daughters, the granddaughters of the original heterozygous sires, are analyzed. It is still necessary to genotype hundreds of sons for power to detect a QTL accounting for less than 10% of the additive genetic variance.

With the advent of DNA-level genetic markers, especially microsatellites, the number of genetic markers that can be generated in any species of interest is virtually unlimited (Lander and Kruglayk, 1995). If many markers are analyzed, then the single comparison significance level is no longer appropriate. This problem is more acute if multiple traits and families are

analyzed. Lander and Kruglayk (1995) therefore proposed that QTL detection should only be considered "confirmed" if statistically significant results are found in two independent samples. Even in the largest dairy cattle populations, a second independent sample of sons is not available. Ron *et al.* (1994) proposed that a daughter design analysis could provide confirmation of significant effects found by a granddaughter design analysis.

We present results from a new analysis of the US Holstein population. Maximum likelihood estimates (MLE) were derived for QTL substitution effect, location, allele frequency, and grandsire genotype, under the assumption that a single QTL with two alleles linked to the marker is segregating in the population. Results for one highly significant effect were confirmed by a daughter design analysis of the Israeli Holstein population.

MATERIALS AND METHODS

Genotyping of US Holstein bulls and Israeli Holstein cows: Nine North American AI organizations contributed semen to the Dairy Bull DNA Repository. The 18 grandsires with the most sons were selected for genotyping. DNA was extracted from semen samples, and grandsires and their sons were genotyped using methods described previously (Heyen *et al.*, 1997). Multiplex PCR and multiple loadings for gel electrophoresis were used to facilitate the large number of assays. Thirty markers were selected for genotyping the sons, based on informativeness, ease of genotyping, and distribution throughout the genome. Of these, 27 were poly-AC microsatellites, and the remaining three were poly-AGC. The number of sons genotyped ranged from 16 to 205. Nineteen autosomes had at least one marker; chromosomes 1, 4, 7, 11, 18, 21, and 23 had two markers; and chromosomes 9 and 10 had three markers. Of 14,650 son genotypes that were determined, 11,210 were informative (77%). Sons were considered "informative" if their genotype was different from their sires' (Ron *et al.*, 1994). The number of families genotyped for each marker ranged from 3 to 13, with a mean of 7.4. The number of sons genotyped per marker ranged from 235 to 723, with a mean of 488. The number of informative sons ranged from 154 to 641, with a mean of 374. Three additional markers on chromosome 14; BM4513, BM302, and ILSTS11; were also genotyped for families 4 and 5, which had significant contrasts ($p<0.01$) for the effect of CSSM66 on fat percentage. The map locations of these markers is given in Figure 1.

To confirm the QTL linked to locus CSSM66, 1498 daughters of seven Israeli Holstein sires were also genotyped. One microliter milk samples were used for PCR templates, as described by Ron *et al.* (1994). Out of 1295 informative genotypes, 1279 had genetic evaluations for milk production traits. The number of informative daughters with evaluations per family ranged from 130 to 208.

The traits analyzed and statistical methods for QTL detection: The genotype data were matched to the bulls' daughter yield deviations (DYD) (Georges *et al.*, 1995) from the July, 1996, US National Holstein evaluation for seven traits; milk, fat, and protein production; fat and protein percent; somatic cell score; and productive life. Preliminary QTL detection was

based on a linear model analysis of DYD including the effect of grandsire and grandsire allele nested within grandsire. A significant allele effect indicated a segregating QTL linked to that marker. For the daughter design analysis, the dependent variable was the estimated cow breeding value from an animal model evaluation of the milk-recorded population, the grandsire effect was replaced with the sire effect, and marker effect was nested within sire family.

An expectation maximization algorithm for analysis of DYD and a single marker linked to a QTL, based on Jansen (1992) was developed to derive MLE of QTL substitution effect, a; frequency of the "low" allele, p; recombination frequency between the QTL and the genetic marker, r; grandsire QTL genotype, and grandsire polygenic effect. The method was applied to estimate the parameters for the effect of locus CSSM66 on fat percent. The model assumed only two QTL alleles segregating in the population.

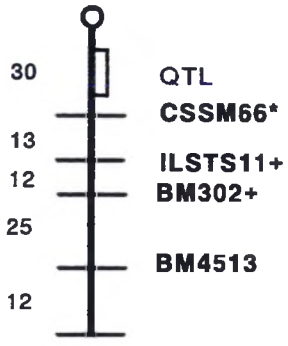

Figure 1. Map of chromosome 14 (Barendse *et al.*, 1997; Eggen *et al.*, 1994) showing markers that were genotyped and recombination frequencies between them. CSSM66 is marked with a star. Other markers with significant effects on fat percentage in either families 4 or 5 are marked with a plus. The centromere is denoted with a circle. The putative location of the QTL is also marked.

RESULTS AND DISCUSSION

The linear model analyses for QTL effects: Three loci; TGLA263, CSSM66, and CSRM60; located on chromosomes 3, 10, and 14; significantly affected at least one trait at $p<0.01$. The F values and probabilities are given in Table 1. TGLA263 had significant effects on both fat and protein percent, which have a correlation of 0.55. The effect of CSSM66 on fat percentage was significant at $p<10^{-7}$. This locus also had a highly significant effect on fat yield, which has a correlation of 0.54 with fat percent. Less than half of the heterozygous families had significant contrasts at $p<0.05$. Between-allele contrasts for fat yield and percent fat for CSSM66 were both significant for families 4 and 5. The contrasts for fa[illegible] 4 and 5 were 15.2 and 18.2 lbs. fat, and 0.130 and 0.065% fat, respectively.

Mapping the QTL linked to CSSM66: The MLE computed over all heterozygous families were a = 0.141% fat, p = 0.714, and r = 0.184. The probabilities that families 2, 7, 9, and 14 were homozygous for the low QTL allele were all > 0.999. The effect computed over all families was slightly larger than the CSSM66 contrast estimated for family 4. The QTL was located adjacent to CSSM66 when the interval ILSTS11-CSSM66 was analyzed. Combining data from families 4 and 5, the estimate of recombination frequency between the QTL and CSSM66 from the analysis of interval BM302-CSSM66 was 0.06. However, the standard error of recombination between the QTL and CSSM66 was about 0.05 for both analyses.

Since the hypothesis of complete linkage between the QTL and CSSM66 was rejected by the single locus analysis, it can be deduced that the QTL is most likely located centromeric to CSSM66, as indicated in Figure 1.

Table 1. Significant within grandsire allele effects ($p<0.01$) by the linear model analysis.

Locus	TGLA263			CSRM60	CSSM66	
Chromosome	3			10	14	
No. families	6			9	8	
No. Sons	452			423	457	
Trait	Milk lbs.	Fat %	Protein %	Protein %	Fat lbs.	Fat %
F-value	2.61	4.29	3.35	2.90	4.71	6.55
Probability	.0168	.0003	.0031	.0025	.0001	10^{-7}

Analysis of CSSM66 in the Israeli Holstein population. The effect of CSSM66 was significant for fat percent at $p<0.0003$, but not for any of the other traits. The within-family contrast was significant only for Israeli family 2 at $p<0.0001$, and the magnitude of the contrast was 0.102% fat, as estimated at the level of the cows' breeding values. Although over all families the effect of CSSM66 on fat yield was not significant, the effect in family 2, 13.2 lbs fat, was significant at $p<.0026$. This is slightly less than the effects found in US families 4 and 5. Using single locus ML, convergence was obtained with $r=0.02$. However, the 95% confidence interval for r spanned zero to 0.34. Thus, all the QTL mapping results are consistent over both populations. It can therefore be deduced that the same quantitative trait locus is segregating in both populations. Development of a comparative map between BTA14 and its human homologue, HSA8, will allow the isolation of candidate genes for the QTL.

Acknowledgments: This research was supported by the United States - Israel Binational Agricultural Research and Development Fund (BARD)

REFERENCES

Barendse, W., *et al.*, (1997) Mammalian Genome **8**:21-28.

Eggen A., Steinbach, D., Fries, R. and Leveziel, H. (1994) *Proc. XXIV Internat. Conf. Anim. Genet. Prague*, Czech Republic. p 111.

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

Heyen, D. *et al.*(1997) *Anim. Genet.* **28**:21-27.

Jansen, R.C., (1992) *Theor. Appl. Genet.* **85:** 252-260.

Lander, E. and Kruglyak, L. (1995) *Nature Genet.* **11**:241-247.

Ron, M., Band, M., Yanai, A. and Weller, J.I. (1994) *Anim. Genet.* **25**:259-264.

Weller, J.I. (1996) In: "Beltsville Symp. Agric. Res. XX: Biotechnology's Role in the Genetic Improvement of Farm Animals", p. 259, editors R.H. Miller, V.G. Pursel, and H. D. Norman, American Society of Animal Science, Savoy, IL.

Weller, J. I., Kashi, Y. and Soller, M. (1990) *J. Dairy Sci.* **73**:2525-2537.