Mammalian Genome 12, 27–31 (2001).
DOI: 10.1007/s003350010232

Mammalian Genome
Incorporating Mouse Genome

Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle

Jeyakumary Nadesalingam,[1,*] Yves Plante,[2,*] John P. Gibson[1,*]

[1]Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1
[2]Bova-Can Laboratories, Saskatchewan Research Council, 15 Innovation Blvd., Saskatoon, Saskatchewan, Canada S7N 2X8

Received: 19 November 1999 / Accepted: 31 August 2000

Abstract. Seventy to 75 sons of each of six Holstein sires were assayed for genotypes at a number of microsatellite loci spanning Chromosomes (Chrs) 1 and 6. The number of informative loci varied from three to eight on each chromosome in different sire families. Linkage order and map distance for microsatellite loci were estimated using CRI-MAP. Estimates of QTL effect and location were made by using a least squares interval mapping approach based on daughter yield deviations of sons for 305-d milk, fat, protein yield, and fat and protein percentage. Thresholds for statistical significance of QTL effects were determined from interval mapping of 10,000 random permutations of the data across the bull sire families and within each sire family separately. Across-sire analyses indicated a significant QTL for fat and protein yield, and fat percentage on Chr 1, and QTL effects on milk yield and protein percentage that might represent one or two QTL on Chr 6. Analyses within each sire family indicated significant QTL effects in five sire families, with one sire possibly being heterozygous for two QTLs. Statistically significant estimates of QTL effects on breeding value ranged from 340 to 640 kg of milk, from 15.6 to 28.4 kg of fat, and 14.4 to 17.6 kg of protein.

Introduction

It has long been recognized that genetic markers could be used to detect and track the inheritance of polymorphisms contributing to genetic variation, known as quantitative trait loci (QTLs). The discovery of highly polymorphic microsatellite markers (Litt and Luty 1989) and the subsequent development of reasonably dense microsatellite linkage maps for the bovine genome (Bishop et al. 1994; Barendse et al. 1994, 1997) have made marker mapping of QTLs a practical reality. Several studies recently have reported the presence of significant QTLs affecting milk production traits on several different chromosomes (e.g., Georges et al. 1995; Spelman et al. 1996; Lipkin et al. 1998; Zhang et al. 1998; Velmala et al. 1999). The sizes of QTL effects being discovered are more than sufficient to warrant their use in selection programs, particularly for the pre-selection of young bulls entering progeny testing (Kashi et al. 1990; Gomez-Raya and Gibson 1993).

We report here on the use of microsatellite markers to map QTLs on Chr 1 and 6 that contribute to variation in milk production traits in six Holstein sire families with a total of 434 sons progeny tested in Canada. The results confirm the existence of several QTLs detected in previous studies as well as detecting QTLs previously unreported.

* All authors contributed equally to this paper.
Correspondence to: J.P. Gibson; E-mail: j.gibson@cgiar.org

Materials and methods

Overview. In total, 434 sons of six prominent grandsires, with 71–75 sons per grandsire, were genotyped at eight and seven microsatellite loci on Chr 1 and 6. Least squares interval mapping was performed based on daughter yield deviations of the sons for 305-day milk, fat, and protein yield, and fat %, and protein %.

DNA preparation. Semen straws (250–500 μl) were thawed at room temperature and emptied into 1.5-ml Eppendorf tubes. The semen was washed three times in 1 × SSC, 2 mM EDTA to remove the cryoprotectant. Cells were resuspended in 1.0 ml TE buffer (pH 8.0) containing 100 μg Proteinase K and incubated at 65°C for 1 h to lyse potentially contaminating epithelial cells. After centrifugation (12,000 rpm for 2 min), the sperm cells were resuspended and lysed into 500 μl of 100 mM Tris-HCl (pH 8.0), 10 mM EDTA, 500 mM NaCl, 1% SDS, 2% β-mercaptoethanol, 400 μg proteinase K, and incubated for 16–24 h at 65°C. The lysates were extracted three times with phenol–chloroform–iso-amyl alcohol (25:24:1) and then three times with chloroform–iso-amyl alcohol (24:1). DNA was recovered by precipitation in the presence of 0.3 M sodium acetate and two volumes of 95% ethanol. DNA was spooled onto a small glass rod, washed in 70% ethanol, air dried, and resuspended in an appropriate volume of TE buffer (pH 8.0). DNA concentrations were estimated by spectrophotometry (O.D. 260 nm). DNA samples were stored at 4°C until needed.

Genotyping. Polymorphic bovine microsatellites were selected based on their map positions on Chrs 1 and 6. Each marker was assayed individually for annealing temperature and Mg^{++} concentration. Template DNA (50 ng) was initially denatured at 94°C for 5 min, followed by the addition of 20 pmol of each primer, 0.5 pmol kinase end-labeled (γ-^{32}P-dCTP) forward primer, 200 μM each of dNTP, 10 mM Tris-HCl pH 9.0, 1.5–3.0 mM $MgCl_2$ (depending on the primer pair), 50 mM KCl, 0.01% vol/wt gelatin, 0.1% Triton X-100, and 0.3 unit *Taq* DNA polymerase (Sigma) in a total volume of 12.5 μl. Samples were subjected to 30 cycles of amplification, each cycle consisting of 30 s of denaturation at 94°C, 30 s annealing at 55°C–68°C (depending on the primer pair), and 30 s extension at 72°C. A final 10-min extension step at 72°C was added at the end of the 30 cycles of amplification to insure complete extension of the PCR products.

Amplification reactions were stopped by the addition of 12.5 μl of sequencing stop buffer and denatured at 94°C for 5 min. Aliquots of 2.0 μl were loaded into 6% denaturing (7 M urea) polyacrylamide gels alongside an M13 sequencing ladder. Sequencing gels were electrophoresed at 90 W for 1.5–3.0 h depending on the expected size of the amplified microsatellite alleles. Gels were fixed (15% methanol, 10% acetic acid), air dried, and exposed to X-ray films overnight.

Three people independently scored allele sizes against an M13 sequencing ladder. Data were corrected based on disagreements among the three scorers, and samples having ambiguous genotypes were either re-amplified or the genotypes set as unknown.

Performance data. All QTL mapping analyses were based on daughter yield deviations (DYD) of sons based on progeny tests that generally involved from 50 to 100 daughters. DYD were supplied by G. Jansen and J. Jamrosik of the Canadian Dairy Network based on Canadian calculations

of estimated breeding values (EBV). DYD for fat % were derived from DYD for milk and fat yield by adding the population mean for milk and fat yield to each DYD, and then dividing the fat DYD by the milk DYD. The DYD for protein % was similarly derived.

Linkage mapping. Linkage orders and map distances among markers were estimated by using the CRI-MAP program (Green et al. 1990) with map distances based on Kosambi's mapping function. Results were compared with the Meat Animal Research Centre (MARC) linkage map as presented on their web site (http://sol.marc.usda.gov/genome/cattle/cattle.html). Our best map order had markers BP7 and BM2320 on Chr 6 reversed compared with the MARC map. The correct order was deemed to be the one on the MARC map, and so we forced the correct order and estimated recombination rates and distances for that order.

Statistical analyses. A weighted least squares interval mapping was performed by using a modified version of the program developed by S. Knott and C. Haley, with details as described by Knott et al. (1996). The combined-sire model for the analysis was,

$$DYD_{ij} = s_j + b_j p_{ij} + e_{ij},$$

where s_j is a fixed effect for sire j, b_j is the coefficient of regression of DYD on probability of QTL inheritance nested within sire j, p_{ij} is the probability of son i having inherited a particular QTL allele from sire j at a given chromosome position, and e_{ij} is a residual error, with variance approximately equal to σ^2_e/REL_{ij}, where REL_{ij} is the reliability of the DYD of son i within sire j. The weighting factor in the weighted least square analysis was $1/REL_{ij}$.

Analyses were performed at 1-cM intervals along the chromosome. For all analyses, the MARC map distances among markers were used. The justifications for doing this are a) the MARC map, with its much higher marker density and genotypes of both parents, should have a much lower genotyping error rate and therefore be more accurate; b) by doing so, all map positions of QTLs are expressed on an internationally recognized and easily accessible map; and c) we had previously demonstrated that it made essentially no difference to interpretation of results whether we used our map distances of those of MARC (Nadesalingam 1999).

Following the combined sire analyses, all analyses were repeated for each sire family separately.

Thresholds for testing significance of effects were obtained by 10,000 random permutations of the DYD data within sire families, with a full combined-sire and within-sire analysis repeated for each trait by chromosome combination for each permutation. The highest F ratio value for each analysis was stored, and the 1%, 5%, and 10% thresholds were found as the 100^{th}, 500^{th}, and 1000^{th} ranked F ratio values for each chromosome by sire (or combined-sire) by trait combination.

Results

The results of the linkage mapping are provided in Table 1. Map distances on Chrs 1 and 6 from our analysis are compared with those from MARC. The MARC map order for Chr 6 was the third best order on our map, but this order was a LOD score of only 0.2 worse than the best order. We therefore forced the order to the MARC order, and map distances are presented for this particular order. The map distances estimated for Chrs 1 and 6 in the study are inflated by approximately 22% and 19% compared with the respective MARC maps. This degree of inflation seems comparable to that of similar maps available on various web sites, and probably reflects the greater difficulty in detecting genotype errors in sparse maps, especially when only one parent is genotyped.

Figure 1 shows the F values for combined-sire analyses of Chrs 1 and 6. Figures 2 and 3 show F values for within-sire analyses of all traits on Chrs 1 and 6, respectively. Thresholds of F values for significance at the 5% level for combined-sire analyses were mostly between 2.8 and 3.0, while those for within-sire analyses ranged between 6.4 and 7.2. Exact probabilities do not change rapidly with F values, so it can safely be assumed that a 5% probability occurs at about F = 2.9 for across-sire analyses and at about F = 6.8 for within-sire analyses.

The across-sire analyses are generally consistent with the within-sire analyses, in that significant effects appear within sires at locations similar to peak values of the test statistic across sires that are close to or surpass significance thresholds. The within-sire analyses allow a more clear interpretation of the number, location, effect and sharing across families of the QTL, and further discussion focuses on these results. Taking the most likely location of significant QTL as being the peak of the F value for a given trait in a given sire family, estimates of statistically significant effects are presented in Table 2. Estimates of allele substitution effect on DYD for all five traits are presented at each significant trait location. There are 10 significant trait-by-location effects. Three additional trait locations are given in Table 2, representing effects that approached the 5% significance level and seemed clearly to be associated with significance at a similar location for the same or correlated trait in the same or another sire family.

Table 1. Comparison of map distances of present and MARC maps for Chromosomes 1 and 6.[a]

Chromosome 1				Chromosome 6			
		Sequential Distances				Sequential Distances	
Marker	MARC map	MARC	Here	Marker	MARC map	MARC	Here
TGLA49	1.9	1.9	0.0	BM1329	35.5	35.5	0
RM095	21.3	19.4	30.8	BM143	49.4	13.9	7.4
ILSTS004	32	10.7	11.0	BM4528	68.1	18.7	12.4
BM4307	35.2	3.2	12.4	BM415	76.3	8.2	6.7
INRA011	54.4	19.2	23.1	BM4311	89.1	12.8	23.7
MB6506	69.2	14.8	13.6	BP7	91.2	2.1	9.3
BM1824	108.6	39.4	42.6	BM2320	120.7	29.5	41.6
BM148	118.1	9.5	8.6				
Total Map[b]		116.2	142.1			85.2	101.1

[a] All distances in cM.
[b] Total distance measured from TGLA49 for Chr 1 and BM1329 for Chr 2.

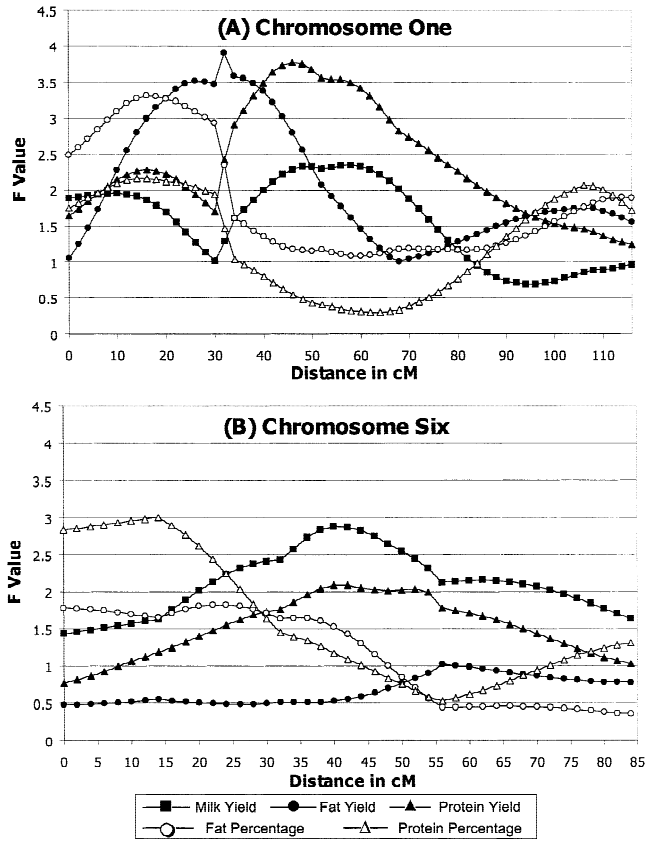

Fig. 1. Across-sire family F values for QTL effects on five milk production traits along Chr 1 (A) and 6 (B).

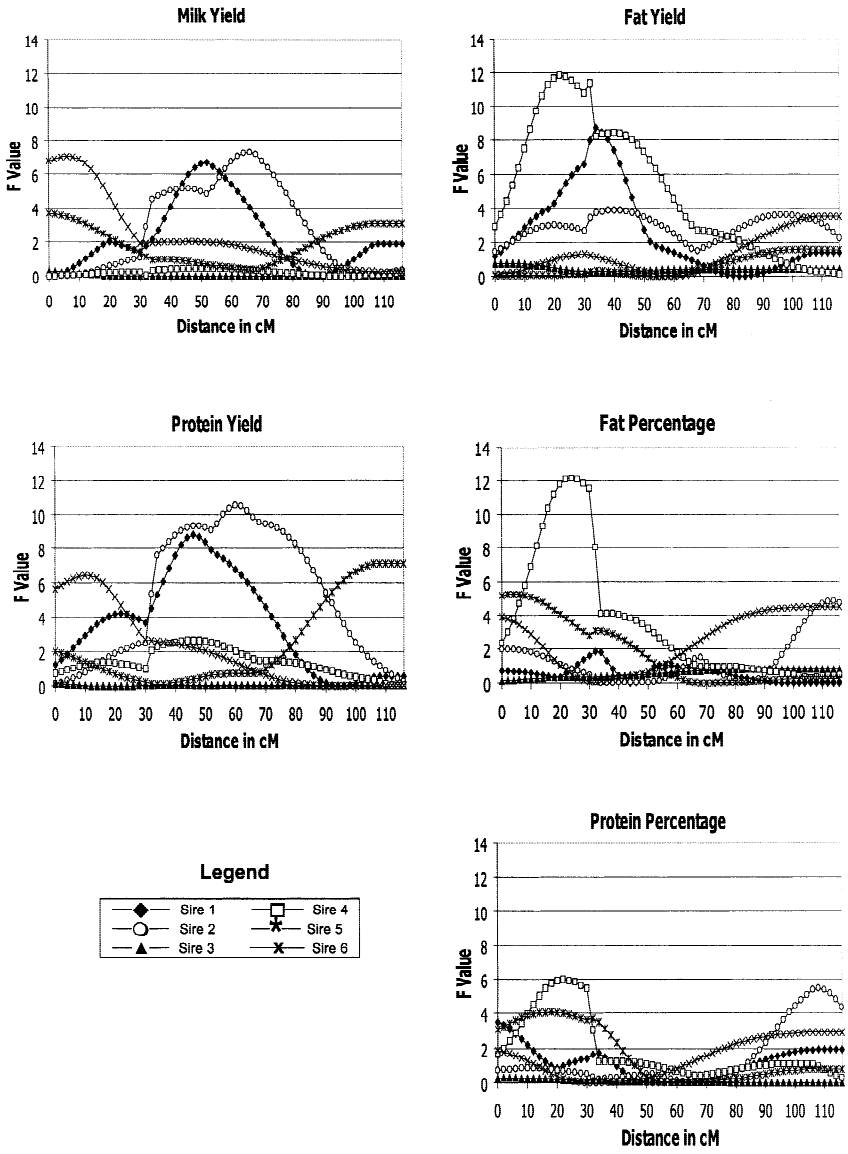

Fig. 2. Test statistics for QTL effects within six sire families on five milk production traits along Chr 1.

In most cases, the estimated effects of the putative QTL conform to the generally observed genetic correlations among milk component traits. Thus, most QTL with a positive effect on one yield trait had positive effects on other yield traits. A possible exception is the QTL at about 24 cM on Chr 1 in sire family 4, which has a large negative effect on fat yield with a small positive effect on milk yield and a large negative effect on fat percentage. While the QTL at 0 cM on Chr 6 in sire family 5 seems also to have opposite effects on fat percentage and milk yield, the lack of clear significance of this effect makes interpretation uncertain.

Given the relatively large errors attached to estimates of QTL location and effect in experiments of this size and design, the most parsimonious interpretation of the results presented in Fig. 2 and 3 and in Table 2 is that in many cases there is a single QTL that affects several traits simultaneously and that is segregating in more than one family. We have assumed this to be the case where QTLs are located within 30 cM of each other and have similar effects in more than one sire family, and/or where QTLs are located within 30 cM of each other within a single family and have similar estimates of effects on the five traits. Confidence intervals were not calculated for locations estimated here, but comparisons with computer simulations and previous publications of results from similar experimental designs indicate that most QTLs will have confidence intervals of at least 30 cM. The choice of 30 cM for deriving putative consensus locations was designed to err on the conservative side, the data likely supporting the combining of information over even broader chromosome regions. The consensus location for the single QTL was taken as being the average of the locations of the peak F value for each trait approaching statistical significance ($p < 0.1$). When combining putative QTLs in this way, the data indicate the presence of four QTLs on Chr 1 and one on Chr 6. Estimates of their effects at the consensus locations are provided in Table 3.

There seem to be three QTLs on Chr 1 that have a general effect on yield of milk, fat, and protein. The distances were fairly large between these QTLs, and there was no consensus location between these QTLs where the estimated effects of the QTL remained consistent with the estimates at the location of the individual peak F values. Nevertheless, there exists a possibility that two of these QTLs might be the same QTL. The chance that all three could be the same QTL seems extremely remote. The fourth QTL on Chr 1 has a clear effect on fat yield and fat percentage that is entirely different from the flanking QTL in other families that affect milk, fat, and protein yield.

On Chr 6, the most parsimonious explanation of the data is a single QTL lying at about 25 cM, with one allele having a positive

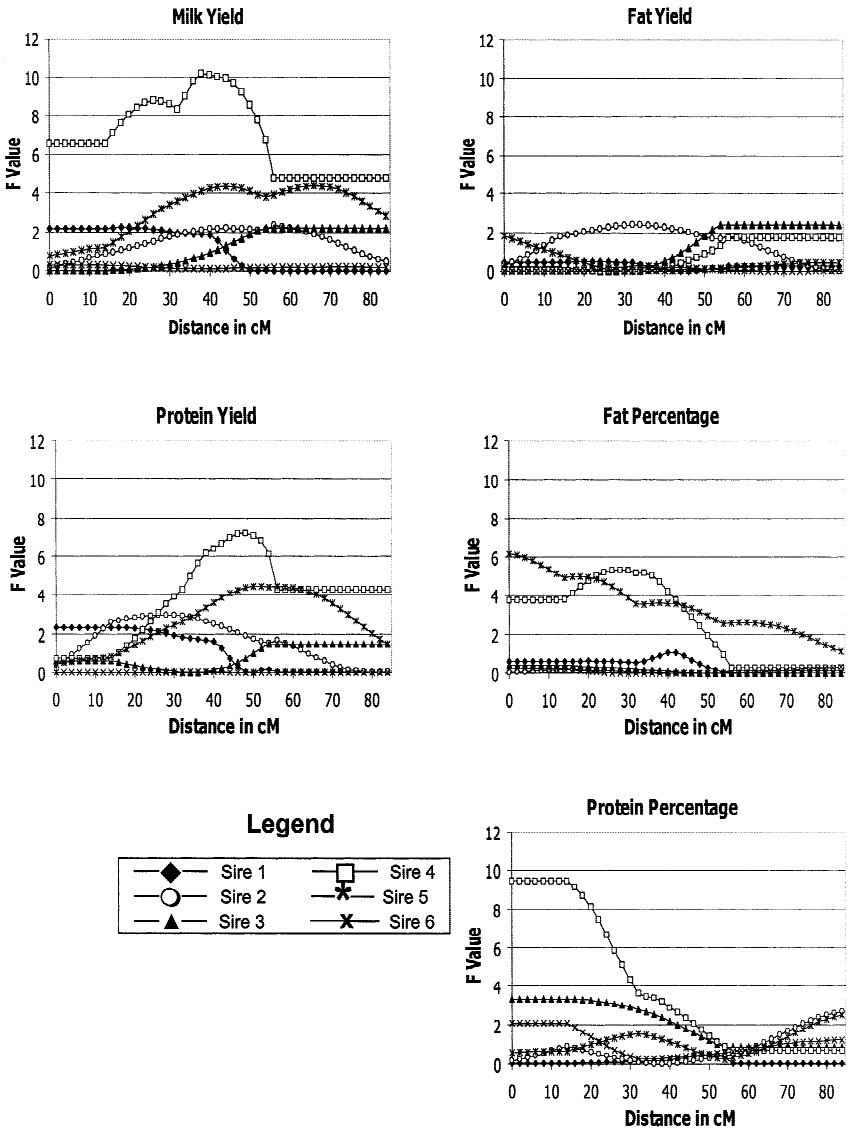

Fig. 3. Test statistics for QTL effects within six sire families for five milk production traits along Chr 6.

effect on milk yield and a lesser effect on protein yield, and possibly a negative effect on fat yield, so that fat and protein percentages are lowered.

Discussion

A critical question in QTL mapping studies is how many of the statistically significant effects represent real QTLs rather than false positive results. Several observations indicate that most of the seven consensus QTLs by sire by chromosome effects detected (Table 3) represent real QTLs. Firstly, even if one considers all five traits as statistically independent of each other, one would expect only $6 \times 5 \times 2 \times 0.05 = 1.5$ significant estimates at $p < .05$. However, because of the high correlations among milk, fat, and protein, one would expect considerably less than 1.5 significant consensus QTLs.

Secondly, two of the QTL by sire combinations on each of Chrs 1 and 6 occurred at very similar locations, with very similar effects in two different pairs of sires. In addition, sires 1 and 2, who have an additive genetic relationship of approximately 0.25, share a four-marker haplotype associated with the increasing allele of the shared QTL on Chr 1.

Thirdly, several of the QTLs observed here are similar in lo-

Table 2. Estimates of QTL effects[c] on all traits at the position of the maximum F value for each trait and sire combination.

Chromosome	Location	Sire	Milk	Fat	Protein	Fat %	Protein %
1	34	1	137	10.2[a]	6.2	0.05	0.02
1	46	1	260	8.9	9.1[a]	−0.01	0.01
1	52	1	246[b]	5.4	7.8	−0.04	−0.00
1	61	2	256	6.1	8.7[a]	−0.04	0.01
1	66	2	248[a]	4.8	8.0	−0.05	0.00
1	22	4	50	−14.5[a]	−3.2	−0.17	−0.05[b]
1	24	4	50	−13.9	−3.0	−0.16[a]	−0.05
1	4	5	192	−0.7	3.7	−0.09[b]	−0.03
1	106	5	184	5.2	7.3[a]	−0.02	0.01
1	6	6	306[a]	2.1	7.7	−0.10	−0.02
1	11	6	333	3.3	8.7[b]	−0.10	−0.02
6	13	4	221	−0.1	2.0	−0.09	−0.05[a]
6	39	4	270[a]	1.3	5.6	−0.09	−0.03
6	47	4	268	3.1	6.1[b]	−0.07	−0.02
6	0	5	91	−5.5	1.9	−0.09[b]	−0.01

[a] Indicates that trait effect was statistically significant ($p < .05$) at the location of the original peak F value.

[b] Indicates effect approached statistical significance ($0.05 < p < 0.1$).

[c] QTL effects are allele substitution effects in kg or % DYD.

Table 3. Estimates of QTL effects[a] at consensus locations across traits and/or sires.[b]

Chromosome	Location (cM)	Sire	Milk (kg)	Fat (kg)	Protein (kg)	Fat %	Protein %
1	8	6	319.55[c]	2.56	8.16[c]	−0.10	−0.02
1	23	4	50.26	−14.21[c]	−3.08	−0.17[c]	−0.05[d]
1	48	1	261.89[d]	7.77[c]	8.84[c]	−0.02	0.00
1	48	2	218.34[c]	7.36	8.14[c]	−0.01	0.01
1	106	5	183.74	5.20	7.25[c]	−0.02	0.01
6	25	4	269.63[c]	−0.16	4.14[d]	−0.11	−0.05[c]
6	25	5	172.16	−1.45	3.90	−0.08[d]	−0.02

[a] QTL effects are allele substitution effects in kg or % DYD. Since allele designation is arbitrary, all effects have been given for the allele that increases milk yield.
[b] The consensus location was the average position across traits and/or sires for peak F values that approached statistical significance ($p < 0.1$).
[c] Indicates that trait effect was statistically significant ($p < .05$) at the location of the original peak F value.
[d] Indicates effect approached statistical significance ($0.05 < p < 0.1$).

cation and effect to those located in other studies. Zhang et al. (1998) and Georges et al. (1995) observed a QTL affecting protein yield on Chr 1 at a location consistent with that observed in sires 1 and 2 here. The fitted effects for the QTL in Georges et al. (1995) indicated a general effect on milk production traits. Zhang et al. (1998) did not report the estimated effects of their QTLs on other milk production traits. Previous studies have also indicated the presence of a QTL variously affecting milk yield, fat %, and protein % at about 15 cM on Chr 6 in Holsteins (Georges et al., 1995; Kuhn et al. 1996; Spelman et al. 1996; Lipkin et al. 1998; Zhang et al. 1998) and Ayrshires (Velmala et al. 1999). This QTL is commonly thought of as affecting protein %, but the estimated effects here and those published by Georges et al. (1995), Zhang et al. (1998), and Velmala et al. (1999) suggest that the primary effect is on milk yield, with little effect on fat and protein yield, such that fat % and protein % are both affected.

Overall, we conclude that at least four of the significant sire by QTL combinations represent real QTL effects. Of the remaining three (for sires 4, 5, and 6 on Chr 1), statistical probability argues that at least one, and perhaps all three, represent real QTLs.

It is a consistent finding across studies that the peak of the test statistic for the QTL on Chr 6 occurs somewhere in the region of BM 143. This seems to rule out the casein genes as being candidates for this QTL, since they map about 40 cM from BM 143.

The estimates of effects of each QTL (Tables 2 and 3) are given as allele substitution effects on DYD, which is equivalent to a son's expected progeny difference (EPD) or one-half of a son's breeding value (BV). These estimates thus represent very large predicted effects on performance. These estimates are likely to be biased upward, perhaps substantially, because only large estimates will pass the stringent threshold for significance. Nevertheless, these estimates clearly represent effects that would be of commercial significance in selection programs.

Acknowledgments. This research was funded by the Cattle Breeding Research Council, the National Research Council Industrial Research Assistance Program, the Natural Sciences and Engineering Research Council, and the Ontario Ministry of Agriculture, Food and Rural Affairs. Daughter yield deviation and pedigree data were kindly supplied by Gerald Jansen and Janusz Jamrozik of the Canadian Dairy Network. S. Knott, P. Visscher, C. Haley, D. Nielson, and P. Green kindly supplied us with copies of software. We thank L.R. Schaeffer for statistical advice; H. Mehrabani-Yeganeh, S. Lefebvre, and G. VanderVoort for data preparation and ancillary analyses; and J.L. Atchison, X. Wu, and M. Lu for laboratory work.

References

Barendse W, Armitage SM, Kossarek LM, Shalom A, Kirkpatrick BW et al. (1994) A genetic linkage map of the bovine genome. Nat Genet 6, 227–235

Barendse W, Vaiman D, Kemp SJ, Sugimoto Y, Armitage SM et al. (1997) A medium-density genetic linkage map of the bovine genome. Mamm Genome 8, 21–28

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SL et al. (1994) A genetic linkage map for cattle. Genetics 136, 619–639

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138, 963–971

Georges M, Nielsen D, Mackinnon M, Mishra A, Okimoto R et al. (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139, 907–920

Gomez-Raya L, Gibson JP (1993) Within-family selection at an otherwise unselected locus in dairy cattle. Genome 36, 433–439

Green P, Falls K, Crooks S (1990). CRI-MAP documentation: version 2.4. Washington University, St. Louis, Mo., USA.

Kashi Y, Hallerman E, Soller M (1990) Marker-assisted selection of candidate bulls for progeny testing programmes. Anim Prod 51, 63–74

Knott SA, Elsen JM, Haley CS (1996) Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. Theor Appl Genet 93, 71–80

Kuhn C, Weikadrd R, Goldammer T, Grupe S, Olasaker I et al. (1996) Isolation and application of chromosome 6 specific microsatellite markers for detection of QTL for milk production traits in cattle. J Anim Breed Genet 113, 355–362

Lipkin E, Mosig MO, Darvasi A, Ezra E, Shalom A et al. (1998) Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. Genetics 149, 1557–1567

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44, 397–401

Nadesalingam J (1999) Mapping and detection of quantitative trait loci (QTL) for milk production traits based on linkage analysis using microsatellite markers in Canadian Holstein cattle. Ph.D. thesis, University of Guelph.

Spelman RJ, Coppieters W, Karim L, van Arendonk JA, Bovenhuis H (1996) Quantitative trait loci analysis for five milk production traits on chromosome 6 in the Dutch Holstein-Friesian populations. Genetics 144, 1799–1808

Velmala RJ, Vikki HJ, Elo KT, de Koning DJ, Maki-Tanila AV (1999). A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. Anim Genet 30, 136–143

Zhang Q, Boichard D, Hoeschele I, Arnst C, Eggen A et al. (1998) Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics 149, 1959–1973