Multiple Quantitative Trait Locus Analysis of Bovine Chromosome 6 in the Israeli Holstein Population by a Daughter Design

Micha Ron,* David Kliger,* Esther Feldmesser,* Eyal Seroussi,* Ephraim Ezra† and Joel Ira Weller*

**Institute of Animal Sciences, ARO, The Volcani Center, Bet Dagan 50250, Israel and †Israel Cattle Breeders Association, Caesaria 38900, Israel*

Manuscript received March 26, 2001
Accepted for publication July 3, 2001

ABSTRACT

Nine Israeli Holstein sire families with 2978 daughters were analyzed for quantitative trait loci effects on chromosome 6 for five milk production traits by a daughter design. All animals were genotyped for 2 markers. The three families with significant effects were genotyped for up to 10 additional markers spanning positions 0–122 cM of BTA6. Two sires were segregating for a locus affecting protein and fat percentage near position 55 cM with an estimated substitution effect of 0.18% protein, which is equivalent to one phenotypic standard deviation. This locus was localized to a confidence interval of 4 cM. One of these sires was also heterozygous for a locus affecting milk, fat, and protein production near the centromere. The hypothesis of two segregating loci was verified by multiple regression analysis. A third sire was heterozygous for a locus affecting milk and protein percentage near the telomeric end of the chromosome. Possible candidates for the major quantitative gene near position 55 cM were determined by comparative mapping. IBSP and SSP1 were used as anchors for the orthologous region on human chromosome 4. Twelve genes were detected within a 2-Mbp sequence. None of these genes have been previously associated with lactogenesis.

MANY studies have shown that individual quantitative trait loci (QTL) can be detected and mapped in commercial dairy cattle populations with the aid of genetic markers by application of daughter or granddaughter designs. The granddaughter design has the advantages that it is more powerful per individual genotyped and it is logistically easier to collect genetic material from artificial insemination sires located at a few studs, as opposed to cows, which are scattered over a much large number of herds (Weller *et al.* 1990). However, the required population structure for the granddaughter design (several sires, each with many progeny-tested sons) is available only in the largest commercial populations. The appropriate population structure for the daughter design (several sires, each with hundreds of milk-recorded daughters) can also be found in moderately sized populations, such as the Israeli Holsteins, or U.S. breeds other than Holstein. Although daughter designs are less powerful than granddaughter designs per individual genotyped, potentially many more daughters are available for analysis, even in moderately sized populations. Furthermore, with multiple records per cow, the advantage of the granddaughter design decreases, especially for high heritability traits (Weller *et al.* 1990).

Corresponding author: Joel I. Weller, Institute of Animal Sciences, ARO, The Volcani Ctr., P.O. Box 6, Bet Dagan 50250, Israel. E-mail: weller@agri.huji.ac.il

Segregating QTL for milk production traits on bovine chromosome 6 have been found in U.S. Holsteins (Georges *et al.* 1995; Zhang *et al.* 1998), Canadian Holsteins (Nadesalingam *et al.* 2001), Dutch Holsteins (Spelman *et al.* 1996), German Holsteins (Kühn *et al.* 1999), British Black and White cattle (Wiener *et al.* 2000), Israeli Holsteins (Lipkin *et al.* 1998), and Finnish Ayrshires (Velmala *et al.* 1999). All these analyses, except Lipkin *et al.* (1998), were based on granddaughter design analyses. Lipkin *et al.* (1998) used a daughter design with sample pooling, but only two markers on chromosome 6 were genotyped.

To apply marker-assisted selection efficiently, the QTL should be localized to a relatively short chromosomal segment (Smith and Smith 1993). Darvasi and Soller (1997) demonstrated that, with a saturated genetic map, QTL resolving power is an inverse function of the squared QTL effect relative to the residual standard deviation and the sample size. In all of the granddaughter design studies, the number of sons genotyped in the families with significant effects was <100. Thus the estimated 95% confidence interval (CI95) for QTL location was always >20 cM (Zhang *et al.* 1998). Decreasing the CI95 by increasing the number of sons genotyped is generally not a valid option because the number of progeny-tested sons in the families with significant effects is limited.

Several of the previous studies have presented evidence for two separate segregating QTL affecting pro-

duction traits on this chromosome (Spelman *et al.* 1996; Zhang *et al.* 1998; Velmala *et al.* 1999). In all these experiments sample sizes were too small to statistically reject the single QTL hypothesis. Furthermore, the most likely QTL location varied greatly among families and analyses (Kühn *et al.* 1999; Velmala *et al.* 1999).

The goals of this study were to confirm the presence of at least one segregating QTL on BTA6 affecting production traits found previously in the Israeli and other dairy cattle populations by a large daughter design analysis, to more accurately map these QTL to determine the number and effects of the segregating QTL, and to construct a list of candidate genes for the most well-defined QTL.

MATERIALS AND METHODS

Population sample: Blood samples were collected from over 13,000 Israeli Holstein cows, daughters of 11 sires from 233 herds. Semen samples were collected from the 11 sires. A total of 6047 cows were analyzed for microsatellite genetic markers in a genome scan for QTL that will be presented elsewhere. All cows were genotyped for at least five microsatellites to confirm paternity. Cows that did not inherit either paternal allele for at least two loci were considered to be not daughters of the sire listed and were therefore deleted from further analysis. Cows without genetic evaluations for all five production traits, milk, fat, and protein production and fat and protein percentage, were also deleted from the analysis.

The 12 genetic markers analyzed on chromosome 6 are listed in Table 1. Cows from nine sire families were genotyped for microsatellites BM143 and BM415. Eight of the nine sires were heterozygous for each locus. Significant effects ($P < 0.01$) associated with either locus for at least one of the traits analyzed were found for three sire families. Daughters of these sires were genotyped for all the additional markers listed in Table 1 for which their sires were heterozygous. The total number of cows genotyped from each family and the number of informative daughters for each marker are given in Table 1. Daughters are considered informative if the daughter genotype was different from her sire's genotype (Ron *et al.* 1996). The map locations of the markers genotyped, based on the Clay Center genetic map, (http://sol.marc.usda.gov/genome/cattle/cattle.html) are given in Table 2.

Genotyping methods: DNA from frozen blood or semen was extracted by the salting out procedure (Ma *et al.* 1996). DNA was diluted to 7 ng/μl and 5 μl was aliquoted to 96-well and 384-well plates using Hydra robotic system (Robbins Scientific, Sunnyvale, CA). DNA in plates was dried and stored at room temperature. The PCR protocols for DNA isolated from semen and blood cells were as described by Ron *et al.* (1995) using a DNA engine thermocycler (MJ Research, Watertown, MA). Annealing temperatures of PCR ranged from 55° to 64°, with 30 cycles of amplification.

PCR reactions were run on the ABI 377 DNA sequencer (Applied Biosystems, Foster City, CA). Automated fragment analysis, size calling, and binning were then used by GeneScan (version 3.1) and Genotyper (version 2.0) genetic softwares (Applied Biosystems) to identify the alleles of each of the microsatellite loci.

Phenotypic records: The official Israeli Holstein genetic evaluations are computed twice yearly at the Agricultural Research Organization. Milk, fat, and protein production over 305 days, preadjusted for calving age and month, are analyzed

TABLE 1

Number of informative genotypes by sire and marker

Sire	Total daughters	Informative daughters per marker											
		ILSTS093	INRA133	ILSTS090	BM1329	BMS2508	BM143	BMS518	BMS483	BMS360	BM415	BM8124	BMS739
2278	683	507		280	343	396	393	326	335	562	364		482
2283	241						179				177		
2357	168						—				102		
3070	601	350	387				385	210	211	354	290	298	399
3089	172						132				116		
3099	233	136	84			145	163			146	—		95
3208	171						164				125		
3212	384						335				210		
3241	325						228				197		
Total	2978	993	471	280	343	541	1979	536	546	1062	1581	298	976

TABLE 2

The markers genotyped on chromosome 6 and their map location

Marker	Map location	
	Clay Center[a]	CRIMAP
ILSTS093	0	0
INRA133	8.2	6.9
ILSTS090	11.8	16.4
BM1329	35.5	38.9
BMS2508	44.2	47.7
BM143	49.4	55.4
BMS518	55.2	61.0
BMS483	64.0	67.5
BMS360	66.5	73.5
BM415	76.3	80.8
BM8124	94.2	99.8
BMS739	113.4	122.5

[a] http://sol.marc.usda.gov/genome/cattle/cattle.html

by a repeatability animal model (Weller *et al.* 1994). Genetic evaluations for fat percentage for each cow are derived by

$$BV_{FP} = (BV_F + M_F)/(BV_M + M_M) - M_{FP},$$

where BV_{FP}, BV_F, and BV_M are the cow's estimated breeding values for fat percentage, fat yield, and milk and M_F, M_M, and M_{FP} are mean adjusted first parity fat yield, milk, and fat percentage of cows born in 1995. Genetic evaluations for protein percentage are computed similarly, with protein yield and percentage instead of fat yield and percentage. The October 2000 evaluations were analyzed. Means, standard deviations, and minimum and maximum values of genetic evaluations of the cows genotyped for the five traits analyzed are given in Table 3, and the correlations among the evaluations are given in Table 4.

Statistical methods: Preliminary QTL analysis for markers BM143 and BM415 was by the linear model

$$BV_{ijkl} = S_{ij} + M_{ijk} + e_{ijkl},$$

where BV_{ijkl} is the estimated breeding value for trait i of cow l, daughter of sire j, that received paternal allele k; S_{ij} is the effect of sire j on trait i; M_{ijk} is the effect of paternal allele k of sire j on trait i; and e_{ijkl} is the random residual associated with each record. A significant paternal allele effect is indicative of a segregating QTL linked to the genetic marker.

A cow's estimated breeding value is a function of her sire's and dam's genetic evaluations, in addition to her own production. Nearly all of the sire effect on the daughter evaluations should be absorbed by the S_{ij} effect, while the dam's effect on the daughter evaluation is included in the residual. In this analysis, the dam's effect can be considered virtually random, because, of all the cows genotyped, only 70 dams had more than a single daughter, and there were only 15 sets of full sibs.

TABLE 4

Correlations among the cow estimated breeding values for the traits analyzed (2978 cows)

	Fat yield	Protein yield	Fat %	Protein %
Milk	0.28	0.70	−0.66	−0.69
Fat yield	1	0.33	0.53	−0.06
Protein yield		1	−0.37	0.03
Fat %			1	0.56

Map distances between the 12 markers analyzed were computed with the "fixed" option of CRIMAP (http://linkage.rockefeller.edu/soft/crimap/) using the daughters of the three sires that were genotyped for more than two markers. The map locations of the loci as computed by CRIMAP are also listed in Table 2. Generally there was good correspondence between the two maps. The CRIMAP results were used for QTL interval mapping. Information content of the markers genotyped on chromosome 6 was computed as described by Spelman *et al.* (1996). Information content was computed separately for each of the three families with significant effects.

For these three families interval mapping based on nonlinear regression was performed by the method of Knott *et al.* (1996), using the program developed by R. J. Spelman (Spelman *et al.* 1996). The test statistic was the ratio of the model to residual sums of squares at each point along the chromosome. Under the null hypothesis of no segregating QTL, this test statistic has an approximately central F distribution. The test statistic and QTL effects were evaluated at 1-cM intervals. The dependent variables were the daughter estimated breeding values. All daughter evaluations were weighted equally. Each family was analyzed separately. In addition, sire families 2278 and 3099 were analyzed jointly, because there was evidence that the same QTL was segregating in both families.

The presence of multiple QTL segregating on chromosome 6 was tested by the linear multiple regression method of Whittaker *et al.* (1996), as modified by Kadarmideen and Dekkers (1999) to account for uncertain paternal allele transmission. Each family was analyzed separately. All the alleles from one paternal haplotype were arbitrarily assigned a value of zero,

TABLE 3

Means, standard deviations, and minimum and maximum of the estimated cow breeding values for the traits analyzed (2978 cows)

Trait	Mean	Standard deviation	Minimum	Maximum
Milk	324	494	−1131	2144
Fat yield	8.34	13.90	−44.28	54.94
Protein yield	3.20	10.53	−28.26	39.81
Fat %	−0.016	0.165	−0.588	0.522
Protein %	−0.058	0.095	−0.350	0.195

TABLE 5

Significant effects on production traits associated with markers BM143 and BM415

Marker	Heterozygous sires	Informative daughters	Trait	Prob. F^a	Sire	Cows[b]	Effect[c]	Prob. T^d
BM143	8	1979	Milk	0.0001	2278	372	179	10^{-5}
					3099	163	−251	10^{-5}
			Protein	0.0399	2278	372	−2.9	0.0024
			% fat	10^{-11}	2278	372	−0.084	10^{-10}
					3099	163	0.108	10^{-6}
			% protein	10^{-14}	2278	372	−0.074	10^{-14}
					3099	163	0.064	10^{-10}
BM415	8	1483	Milk	0.0004	2278	343	144	0.0004
					3070	270	163	0.0029
			% fat	0.0350	2278	343	−0.039	0.0028
			% protein	10^{-7}	2278	343	−0.041	10^{-8}
					3070	270	−0.025	0.0066

[a] Significance of the effect of paternal marker allele computed over all heterozygous sires.
[b] Number of informative daughters per sire.
[c] Marker allele substitution effect in kilograms for milk, fat, and protein, and percentages for fat and protein.
[d] Significance of the within-family effect of paternal marker allele.

and all the alleles from the other paternal haplotype were assigned a value of unity. Markers were tested for inclusion in the final model by "stepwise" and "backward" regression analysis, including all markers genotyped in each family. Whittaker *et al.* (1996) explain that parameters for a segregating QTL can be estimated between two markers only if the two paternal alleles derived from the same chromosome have coefficients of the same sign. In most cases this criterion was not met by either the stepwise or backward methods. No model was uniformly best for all five traits. The final model therefore included the positions of markers ILST93, BMS2058, and BMS518. With this model, markers BMS2058 and BMS518 had the same sign for all five traits, and a QTL effect and position could be estimated. The second putative QTL could not be bracketed by markers because it is apparently located at the end of the chromosome.

The CI95 for QTL position and effect were estimated by the nonparametric bootstrap method (Visscher *et al.* 1996). Two hundred bootstrap samples were generated from the data, and interval mapping as described previously was performed for each of the five traits. The shortest interval including 95% of the bootstrap samples was selected as the CI95. Thus some of the CI95 were asymmetrical with respect to either the distribution mode or the fraction of excluded samples. CI95 were computed for the map location with the highest model-to-residual variance ratio and for the estimated QTL effect at this position. As in the previous analyses, each family was analyzed separately, and sire families 2278 and 3099 were analyzed jointly for the production traits, because there was evidence that the same QTL was segregating in both families.

Bioinformatics: A map of candidate genes on BTA6 in the vicinity of BM143 was constructed using human genomic clones. Clones related to SSP1 and IBSP genes were detected using the BLAST programs on the National Center for Biotechnology Information/ National Institutes of Health server (http://www.ncbi.nlm.nih.gov/BLAST). Sequence of clones and draft contig sequences were downloaded and assembled using the *GAP4* program (Staden *et al.* 2000) on an XP1000 Unix workstation. The database consisted of bacterial artificial chromosome (BAC) clones with the following accession numbers: AC013762, AC019007, AC019279, AC021183, AC021836, AC021959, AC022718, AC023334, AC023521, AC083829, AC084732, and AC087106. The sequence of scaffold assembly GA_x2HTBL5H6RW (Venter *et al.* 2001; http://www.celera.com/) that represents positions 87334301–88990861 on the axis of human chromosome 4 was incorporated into this database. Public map of the human genome was accessed through (http://genome.ucsc.edu/goldenPath/hgTracks.html).

RESULTS

Significant effects on production traits associated with markers BM143 and BM415 are given in Table 5 for the ANOVA analysis across all families with heterozygous sires. The within-family effects and *t*-values for the families with significant contrasts are also listed. Although the sign of the effect is arbitrary, the sign was consistent throughout all the analyses. For example, if one sire haplotype had a positive effect on both milk and fat, relative to the alternative haplotype, then the sign was the same for both traits in all analyses.

The effect associated with BM143 was significant for all traits ($P < 0.05$) except fat yield. The effects associated with milk and protein percentage were highly significant for both loci. Highly significant within-family contrasts were found for sires 2278, 3070, and 3099. There was marginal significance for the effects of locus BM415 on milk and protein percentage in family 3212 ($0.01 < P < 0.05$), but the sample size was relatively small. Sire 3099 was heterozygous only for BM143, while the other two sires were heterozygous for both loci. Effects associated with sire 2278 were significant for both loci, while only the effect associated with BM415 was significant for sire 3070.

Marker information content for these three families including all 12 loci genotyped on chromosome 6 is plotted in Figure 1. There is a major reduction in information content between positions 0 and 40 for all three

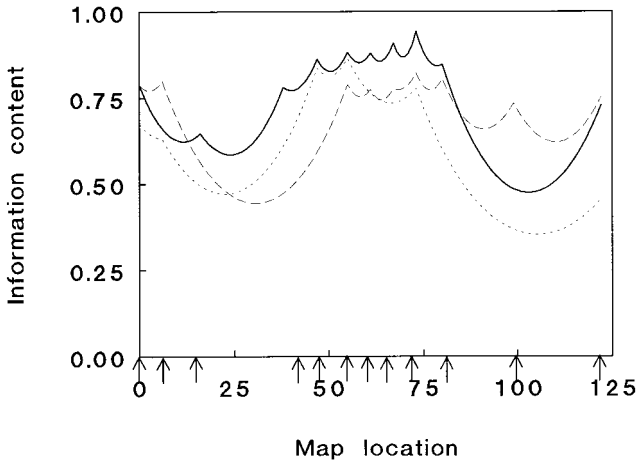

FIGURE 1.—Information content for the three families genotyped for more than two markers. Information content was estimated separately for each family. —, family 2278; - - -, family 3070; ⋯, family 3099. The positions of the markers are indicated by arrows.

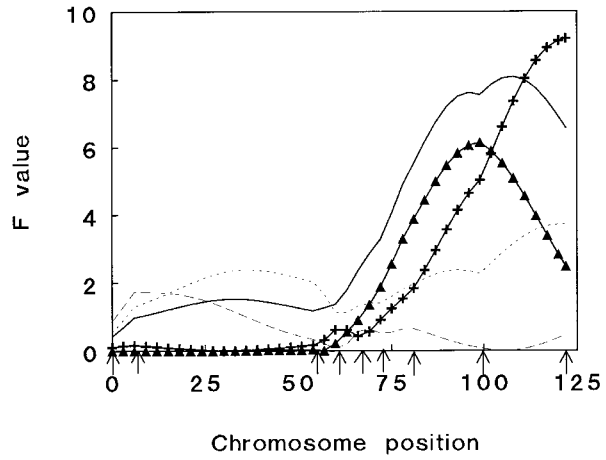

FIGURE 3.—Interval mapping for sire family 3070 for the five production traits. —, milk yield; - - -, fat yield; ⋯, protein yield; +, fat percentage; ▲, protein percentage. The positions of the markers are indicated by arrows.

families and between positions 80 and 120 for sires 2278 and 3099. The interval mapping results for all five traits are given in Figures 2–4, separately for each family. The locations of the test statistic peaks, the test statistic values at the peaks, and the estimated substitution effects at the peaks are given in Table 6.

In general the effects estimated by interval mapping are smaller than the effects associated with the individual marker with the greatest effect. The same QTL near position 55 cM appears to be segregating in the daughters of sires 2278 and 3099. This locus has effects in the opposite direction on milk and fat yield and, therefore, very large effects on fat and protein percentage. In sire 2278 the effect on protein yield is in the same direction as fat yield, while in sire 3099 the effect on protein yield is not significant. Generally, positions of the maximum test statistic were similar for these two families, but there is a difference of 11 cM in the position of the maximum test statistic for milk yield. The profile of effects for sire 3070 is radically different from the other two sires. Significant effects on milk and fat and protein percentage are found between positions 100 and 120 cM, but no effects are found near position 55 cM.

Apparently, a second QTL is segregating in family 2278 close to the centromere at position 0 cM. This locus affects all three production traits in the same direction but does not significantly affect fat or protein percentage. Therefore, the effects of the two loci are in the same direction for fat and protein yield, and the test statistic is relatively high over the entire range of 0–55 cM. However, for milk yield the effects of the two QTL are in repulsion. Therefore the test statistic approached zero near position 15 cM, between the two loci.

For sire 2278 the maximum of the test statistic for milk is shifted toward the telomere, relative to fat and protein percentage. This is to be expected if the ob-

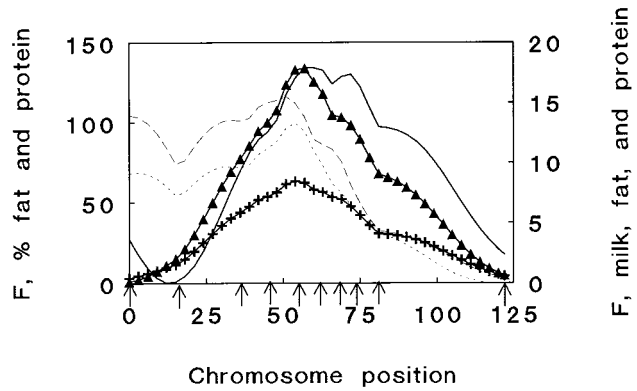

FIGURE 2.—Interval mapping for sire family 2278 for the five production traits. —, milk yield; - - -, fat yield; ⋯, protein yield; +, fat percentage; ▲, protein percentage. The positions of the markers are indicated by arrows.

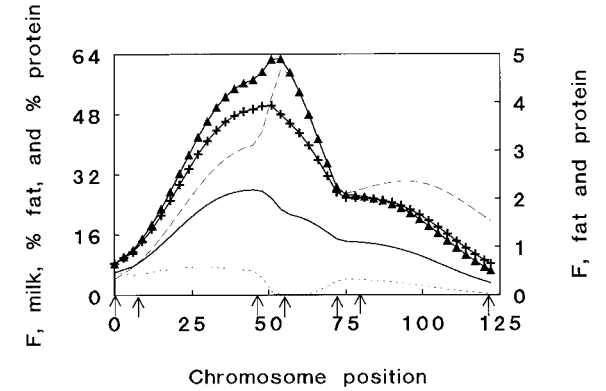

FIGURE 4.—Interval mapping for sire family 3099 for the five production traits. —, milk yield; - - -, fat yield; ⋯, protein yield; +, fat percentage; ▲, protein percentage. The positions of the markers are indicated by arrows.

TABLE 6

The locations of the test statistic peaks, the test statistic values at the peaks, and the estimated substitution effects at the peaks

Sire	Trait	Peak location	Test statistic[a]	Substitution effect
2278	Milk yield	58	18.0	138
	Fat yield	51	15.5	−4.0
	Protein yield	54	13.3	−2.8
	Fat %	54	63.3	−0.073
	Protein %	56	134.6	−0.061
3070	Milk yield	108	8.1	112
	Fat yield	7	1.7 NS	1.6
	Protein yield	121	3.7 NS	1.7
	Fat %	122	9.2	−0.035
	Protein %	99	6.1	−0.015
3099	Milk yield	47	27.9	−272
	Fat yield	55	4.8	3.4
	Protein yield	28	0.6 NS	−1.2
	Fat %	50	50.6	0.108
	Protein %	53	63.1	0.067

[a] Values marked NS were not significant, at $P < 0.05$, for central F with numerator d.f. = 1 and denominator d.f. = no. of daughters − 1. All other test statistics were significant.

served test statistic profile is due to the joint effects of two QTL (Martinez and Curnow 1992). This possibility was investigated by multiple regression analysis of family 2278. Results are presented in Table 7. Positions of 0, 48, and 61 cM were included in the model, which correspond to markers ILSTS093, BMS2508, and BMS518. For all five traits the signs of the coefficients for markers BMS2508 and BMS518 were the same. Therefore the effects of a segregating QTL between these two loci could be estimated by formula of Whittaker *et al.* (1996) as modified by Kadarmideen and Dekkers (1999) for half-sib families. The QTL effect and location for each trait are also presented in Table 7. The effects were generally similar to the interval mapping effects, but the positions were shifted toward the centromere for all five traits. The multiple regression QTL positions for sire family 2278 were very close to the interval mapping positions for sire family 3099, except for protein yield, which was not significant for sire 3099. Thus, these results support the hypothesis that the same QTL is segregating in both families and that this QTL affects all five traits.

The results of the bootstrap analyses are presented in Table 8. Sire families 2278 and 3099 were analyzed jointly. Since the peak location and confidence interval are common to both families, these values are presented only for family 2278. In general the means of the bootstrap analyses for both QTL effect and location were close to the interval mapping estimates derived from the actual data. For sire 3070 the CI95 for substitution effect included 0 for all five traits, even though a test statistic of 9.2 was obtained for fat percentage. The probability of obtaining this central F-value with 1 numerator d.f. and 606 denominator d.f. is 0.0025. For sire 3099 the CI95 for the substitution effects of fat and protein yield included 0. All test statistic values for these traits in the analyses of the actual data were <10. In general CI95 for QTL location were quite large, except for the effects of fat and protein percentage in families 2278 and 3099. The CI95 for protein percentage was only 4 cM in these families. For family 3070 the CI95 for all traits spanned nearly the entire chromosome. Therefore, despite the major difference in the estimated QTL positions from both the analyses of the actual data and the bootstrap means, it was still not possible to prove that the QTL location is different in this family.

Candidate genes within the 4-cM CI95 were determined by comparative mapping. BM143 is adjacent to the SSP1 and IBSP genes on BTA6 in a region syntenic to human chromosome 4 (Band *et al.* 2000). One-half of the genomic clones found were initially mapped to other chromosomes, mostly to HSA7. Using the Blast program, we identified 12 genes in ~2 Mbp of sequence on HS4 in the region syntenic to SSP1 and IBSP on BTA6: KIAA0914, HERC3, CEB1, FLJ20637, BCRP, PKD2, SSP1, MEPE, IBSP, DMP1, DSPP, and SPARCL1. None of these genes have been previously associated with lactogenesis. MEPE, IBSP, DMP1, DSPP, and SPARCL1 are a cluster of genes related to bone formation (Rowe *et al.* 2000).

DISCUSSION

Several previous studies presented evidence for two QTL affecting production traits segregating on chromosome 6, one close to the middle of the chromosome and a second QTL more distant from the centromere (Spelman *et al.* 1996; Zhang *et al.* 1998; Velmala *et al.* 1999). The QTL close to position 50 cM was definitely identified in this study and found to be segregating in families 2278 and 3099. Evidence was presented for the second QTL, which appears to be segregating in family 3070. A test statistic of 9.2 was obtained for fat percentage. However, the bootstrap analyses were inconclusive with respect to presence of a segregating QTL in this family. It should also be noted that information content at the telomeric end of the chromosome was higher for sire 3070 than for the other two sires. This would result in an increased test statistic for sire 3070 if a QTL is segregating in this chromosomal region.

Most recent studies used the permutation test of Churchill and Doerge (1994) to determine chromosome-wide or genome-wide significance levels (*e.g.*, Nadesalingam *et al.* 2001). We believe that the empirical bootstrap is more useful in most cases because it can be used to derive a confidence interval for both QTL location and effect. If the CI95 for QTL effect of the scanned chromosome includes zero, the null hypothesis

TABLE 7

Multiple regression analysis for sire family 2278

Trait	Coefficients			Substitution effect	Map location
	ILST93	BMS2058	BMS518		
Milk yield	−109	128	41	170	51
Fat yield	−3.18	−1.94	−1.27	−3.2	53
Protein yield	−1.79	−2.21	−0.08	−2.3	48
Fat %	0.0017	−0.0539	−0.0230	−0.077	52
Protein %	0.0121	−0.0534	−0.0115	−0.065	50

of no segregating QTL cannot be rejected. In this study, estimates of CI95 for QTL effect and location were generally consistent with respect to rejecting the null hypothesis. If the CI95 for QTL effect included zero, then the CI95 for QTL location included nearly the entire chromosome. The empirical bootstrap has the added advantage that much fewer simulations must be computed. However, if more than a single QTL is segregating on the chromosome, this method can give erroneous results and is overly conservative in most cases (Bennewitz *et al.* 2000). Furthermore, it is not clear how the bootstrap analysis performs with biased QTL estimates.

A third QTL close to position 0 cM was also identified with a different profile of effects. This QTL is apparently segregating only in sire 2278. Most of the previous studies did not examine this region of the chromosome. Therefore no conclusion as to whether this polymorphism is unique to the Israeli population can be made. However, the Israeli Holstein population is closely related to the U.S., Canadian, and Dutch populations.

The number of individuals genotyped in this study was much greater than all previous analyses of chromosome 6. Therefore, even though a daughter design was employed, it was possible to more accurately map the segregating QTL as compared to the previous studies. The effect on protein percentage was localized to a CI95 of 4 cM. No previous study has been able to obtain this level of accuracy for mapping a segregating QTL in a commercial animal population.

Darvasi and Soller (1997) found that the CI95 with a saturated genetic map, CI95*M*, could be estimated as

$$\mathrm{CI95}M = 3000/(mN\delta^2),$$

where m is the number of informative meioses per individual (for the daughter and granddaughter designs, $m = 1$); N is the number of individuals genotyped; and δ is the substitution effect in units of the residual standard deviation. For the effect of protein percentage in sire families 2278 and 3099, $\delta = 0.065/0.062 = 1.05$, where 0.065 is the mean QTL effect as estimated from the estimated breeding values and 0.062 is the root residual mean squares from the interval mapping. There were 914 informative daughters of these two sires. Thus $\mathrm{CI95}M = 3000/[(914)(1.05)^2] = 3.0$ cM, as compared to CI95 of 4 cM from the bootstrap analysis. Thus, genotyping additional markers on the same sample of daughters should not significantly decrease the CI95 for QTL location. Genotyping additional daughters from the heterozygous families could decrease the CI, and additional daughters of sire 2278 are available for analysis.

As noted previously, the magnitude of the QTL effect presented is clearly underestimated, because the analyses were based on estimated cow breeding values. However, the power of QTL detection based on analysis of genetic evaluations is not reduced relative to other alternatives (Israel and Weller 1998). Furthermore, Israel and Weller (1998) found in simulation studies that estimates of QTL effects derived from analysis of daughter yield deviations in granddaughter designs and yield deviations in daughter designs were also biased.

Israel and Weller (1998) found that for a candidate gene affecting a trait with a heritability of 0.25, the effect estimated by analysis of cow genetic evaluations is equal to approximately one-half of the actual substitution effect. Recent results (C. Israel and J. I. Weller, unpublished results) indicate that for the daughter design, the estimated effect from analysis of cow breeding values is only about one-third of the actual effect. In this case the substitution effects for the QTL in the center of BTA6 on protein percentage would be ~0.18% or ~1 phenotypic standard deviation. The magnitude of the effect found is therefore very similar to the effects of 0.09% for the U.S. grandsire family with the greatest effect, where the estimated effect is one-half of the substitution effect (Zhang *et al.* 1998). The granddaughter design effects of ~0.12% protein found by Spelman *et al.* (1996) and Velmala *et al.* (1999) are somewhat larger but are probably biased upward (Georges *et al.* 1995).

Similar to most previous studies, the majority of the sire families analyzed did not display any evidence of a segregating QTL near the middle of the chromosome. Although there was marginal significance for family 3212 for locus BM415, as noted previously, the sample size was relatively small, and this result must be consid-

TABLE 8

Bootstrap means and confidence intervals for QTL effect and location for the production traits

		Substitution effect		Peak location	
Sire	Trait	Mean	CI95	Mean	CI95
2278	Milk yield	142	74–218	54	38–84
	Fat yield	−4.5	(−6.5)–(−2.6)	37	0–67
	Protein yield	−3.1	(−4.8)–(−1.7)	36	0–67
	Fat %	−0.075	(−0.090)–(−0.058)	53	49–59
	Protein %	−0.061	(−0.073)–(−0.051)	55	53–57
3070	Milk yield	113	(−30)–(183)	92	0–122
	Fat yield	1.2	(−3.5)–(4.6)	54	0–122
	Protein yield	2.1	(−1.4)–(3.8)	79	0–122
	Fat %	−0.034	(−0.060)–(0.026)	110	4–122
	Protein %	−0.014	(−0.030)–(0.013)	89	7–117
3099	Milk yield	−272	(−370)–(−165)		
	Fat yield	2.9	(−1.1)–(6.9)		
	Protein yield	−0.7	(−4.1)–(2.5)		
	Fat %	0.107	0.077–0.138		
	Protein %	0.066	0.052–0.082		

Sire families 2278 and 3099 were analyzed jointly.

ered inconclusive. This tends to indicate that one of the QTL alleles has a frequency ⩾50% throughout the population. If 75% of the sires are homozygous for the QTL (six out of eight), and only two QTL alleles are segregating in the population, then the frequency of the more frequent allele assuming random mating is ~0.85. Assuming that one-quarter of the individuals are in fact heterozygous for this QTL, this locus explains about one-quarter of the phenotypic variance for protein percentage or 40% of the genetic variance (Weller *et al.* 1990). If the economically favorable allele is at a relatively low frequency in the population, then the potential gain due to selection for this allele could be quite significant.

None of the three Israeli sires with segregating QTL were closely related. Sires 2278 and 3099 share no common known ancestors in the three previous generations. It is not surprising that the same QTL close to position 55 cM was segregating in two unrelated sires, because this polymorphism was found to be segregating even in different cattle breeds (Velmala *et al.* 1999). Furthermore, sires 2278 and 3099 had four different alleles for BM143. Therefore, if these two sires share a common identical-by-descent segment including the rare QTL allele (Riquet *et al.* 1999), it no longer includes this marker, even though its location is very close to the center of the CI95.

We used comparative mapping to uncover possible candidates for the major QTL in the vicinity of the BM143 marker. The human genomic maps of the region were inconsistent. Several genomic clones contained genes that were mapped to other chromosomes (*e.g.*, AC019007 contain the gene PRO0813 mapped to HSA6 as well as the HERC3 gene mapped to HSA4). The source of this ambiguity is not clear, since chimerism is considered rare in BAC clones, and contamination of clones during preparation can be avoided easily. The list of 12 genes on BTA6 should include most of the genes in the critical region. Besides the cluster of genes related to bone formation, there are genes involved in cellular transport and regulation. In the public map, a larger view of 5 Mbp around IBSP and SSP1 genes revealed only 4 additional known genes (PTP13, MLLT2, SNCA, and MMRN). Mapping of further candidates should await publication of better human genome maps and can be readily updated using the site http://genome.ucsc.edu/goldenPath/hgTracks.html.

Now that the CI95 has been reduced to 4 cM and most of the genes within this segment have been identified, two approaches can be applied to determine the actual QTL: the candidate gene approach and the common identical-by-descent segment approach (Kim and Park 2001).

We thank R. J. Spelman for the use of his interval-mapping program. This research was supported by a grant from the Israel Milk Marketing Board and the U.S.-Israel Binational Agricultural Research and Development fund (BARD).

LITERATURE CITED

Band, M. R., J. H. Larson, M. Rebeiz, C. A. Green, D. W. Heyen *et al.*, 2000 An ordered comparative map of the cattle and human genomes. Genome Res. **10:** 1359–1368.

Bennewitz, J., N. N. Reinsch and E. Kalm, 2000 Proposals for improved bootstrap confidence intervals in QTL. Proceedings of the 51st Annual Meeting of the European Association for Animal Production, The Hague, The Netherlands, p. 6.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971.

Darvasi, A., and M. Soller, 1997 A simple method to calculate resolving power and confidence interval of QTL map location. Behav. Genet. **27:** 125–132.

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

Israel, C., and J. I. Weller, 1998 Estimation of candidate gene effects in dairy cattle populations. J. Dairy Sci. **81:** 1653–1662.

Kadarmideen, H. N., and J. C. M. Dekkers, 1999 Regression on markers with uncertain allele transmission for QTL mapping in half-sib designs. Genet. Sel. Evol. **31:** 437–455.

Kim, J. J., and Y. I. Park, 2001 Current status of quantitative trait locus mapping in livestock species—review. Asian-Aust. J. Anim. Sci. **14:** 587–596.

Knott, S. A., J. M. Elsen and C. S. Haley, 1996 Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. Theor. Appl. Genet. **93:** 71–80.

Kühn, C., G. Freyer, R. Weikard, T. Goldammer and M. Schwerin, 1999 Detection of QTL for milk production traits in cattle by application of a specifically developed marker map of BTA6. Anim. Genet. **30:** 333–340.

Lipkin, E., M. O. Mosig, A. Darvasi, E. Ezra, A. Shalom *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.

Ma, R. Z., J. E. Beever, Y. Da, C. A. Green, I. Russ *et al.*, 1996 A male linkage map of the cattle (Bos taurus) genome. J. Hered. **87:** 261–271.

Martinez, O., and R. N. Curnow, 1992 Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. **85:** 480–488.

Nadesalingam, J., Y. Plante and J. P. Gibson, 2001 Detection of QTL for milk production on chromosomes 1 and 6 of Holstein cattle. Mamm. Genome **12:** 27–31.

Riquet, J., W. Coppieters, N. Cambisano, J. J. Arranz, P. Berzi *et al.*, 1999 Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. Proc. Natl. Acad. Sci. USA **96:** 9252–9257.

Ron, M., Y. Blank and M. Band, 1995 Determination of the optimal tissue source and number of microsatellites for detection of zygotic origin of cattle twins. Anim. Biotech. **6:** 27–37.

Ron, M., Y. Blank, M. Band, E. Ezra and J. I. Weller, 1996 Misidentification rate in the Israeli Dairy cattle population and its implications for genetic improvement. J. Dairy Sci. **79:** 676–681.

Rowe, P. S. N., P. A. de Zoysa, R. Dong, H. R. Wang, K. E. White *et al.*, 2000 MEPE, a new gene expressed in bone marrow and tumors causing osteomalacia. Genomics **67:** 54–68.

Smith, C., and D. B. Smith, 1993 The need for close linkages in marker-assisted selection for economic merit in livestock. Anim. Breed. Abstr. **61:** 197–204.

Spelman, R. J., W. Coppieters, L. Karim, J. A. M. van Arendonk and H. Bovenhuis, 1996 Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. Genetics **144:** 1799–1808.

Staden, R., K. F. Beal and J. K. Bonfield, 2000 The Staden package, 1998. Methods Mol. Biol. **132:** 115–130.

Velmala, R. J., H. J. Vikki, K. T. Elo, D. J. de Koning and A. V. Mäki-Tanila, 1999 A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. Anim. Genet. **30:** 136–143.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural *et al.*, 2001 The sequence of the human genome. Science **291:** 1304–1351.

Visscher, P. M., R. Thompson and C. S. Haley, 1996 Confidence intervals in QTL mapping by bootstrapping. Genetics **143:** 1013–1020.

Weller, J. I., Y. Kashi and M. Soller, 1990 Power of "daughter" and "granddaughter" designs for genetic mapping of quantitative traits in dairy cattle using genetic markers. J. Dairy Sci. **73:** 2525–2537.

Weller, J. I., C. Israel and E. Ezra, 1994 A simple procedure for obtaining approximate interim cow solutions from an animal model. J. Dairy Sci. **77:** 1126–1131.

Whittaker, J. C., R. Thompson and P. M. Visscher, 1996 On the mapping of QTL by regression of phenotype on marker-type. Heredity **77:** 23–32.

Wiener, P., I. Maclean, J. L. Williams and J. A. Wooliams, 2000 Testing for the presence of previously identified QTL for milk production traits in new populations. Anim. Genet. **31:** 385–395.

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

Communicating editor: C. Haley