Detection and Characterization of Amplified Fragment Length Polymorphism Markers for Clinical Mastitis in Canadian Holsteins

B. S. Sharma,*¹ G. B. Jansen,* N. A. Karrow,* D. Kelton,† and Z. Jiang‡

*Department of Animal and Poultry Science, and

†Department of Population Medicine, University of Guelph, Guelph, N1G 2W1, Canada

‡Department of Animal Sciences, Washington State University, Pullman 99164-6351

ABSTRACT

Mastitis is the most frequent, complex, and costly disease in dairy cattle. Genetic improvement of milk production traits has accompanied an increased susceptibility to mastitis. To determine genome-wide quantitative trait locus-linked markers for mastitis resistance, a total of 200 cows, comprising 100 top clinical mastitis- (CM) resistant and 100 top CM-susceptible cows, were screened by selective DNA pooling and amplified fragment length polymorphism (AFLP) technique. The AFLP analysis on resistant and susceptible pools using 89 selective primer combinations revealed 27 significant AFLP markers at a false discovery rate (FDR) of <5%. The most promising AFLP marker was then selected for further characterization. Individual AFLP genotyping of the marker on all selected animals confirmed a significant difference. Sequence analysis detected a single nucleotide polymorphism (A↔G) responsible for the AFLP polymorphism, which was named CGIL4. The PCR-RFLP analysis indicated that the frequency of allele A was significantly higher in the resistant group. The logistic regression analysis demonstrated that the marker was significantly associated with somatic cell score, CM residual values, and production traits. Radiation hybrid mapping assigned the marker to *Bos taurus* autosome 22. The present study provides promising markers for marker-assisted selection for CM resistance. Our results also demonstrated the capability of AFLP on selective DNA pools as a method for detection of genome regions containing quantitative trait loci.

Key words: amplified fragment length polymorphism, clinical mastitis, quantitative trait locus mapping, radiation hybrid mapping

INTRODUCTION

Mastitis is the most frequent disease in dairy cattle (Kossabati et al., 1998) and is believed to be influenced

by many genes and numerous environmental factors. In addition to causing animal distress, mastitis is estimated to cost dairy farmers approximately \$200 per cow annually, due to veterinary expenses, reduced milk production, and early culling (Schutz, 1994). To date, milk SCC or SCS have been integrated in breeding programs as indicators for assessing mastitis resistance (De Jong and Lansbergen, 1996), on the basis of the many studies indicating a positive and moderate to high genetic correlation between clinical mastitis (CM) and SCC with estimates averaging 0.7 (Pösö and Mäntysaari, 1996; Heringstad et al., 2000). Despite the suggestion of pleiotropy, recent studies have not identified common QTL for CM and SCC. For example, Klungland and colleagues (2001) detected a single QTL of genome-wide significance with an effect on CM on *Bos taurus* autosome (BTA) 6, and 4 suggestive QTL on BTA3, BTA4, BTA14, and BTA27. However, none of these QTL for CM was reproduced as QTL for SCC. Similarly, Schulman et al. (2004) reported QTL affecting SCS on BTA1, 3, 11, 18, 21, 24, 27, and 29, and 2 for CM on BTA14 and 18 in Finnish Ayrshire.

Complex genetic traits, such as resistance to mastitis, are likely determined by multiple genes that, when considered individually, may contribute only a modest effect on a phenotype (Prochazka et al., 2001). Although genome-wide linkage studies of complex traits, conducted by utilizing highly informative microsatellite markers, have proven to be a feasible means of detecting QTL, recent evidence suggests that linkage studies can detect only major genes with a large phenotypic effect and may have limited power to detect genes with a modest contribution to a trait (Long et al., 1997). To identify genes with modest effects on complex traits, current recommendations suggest pursuing genome-wide association studies in properly selected case-control groups with diallelic markers that should ideally represent every gene (Rish and Merikangas, 1996). However, genotyping several hundred individuals with thousands of markers covering the entire genome would obviously be costly and time-consuming.

The amplified fragment length polymorphism technique (AFLP) described by Vos et al. (1995) is a PCR-

Received October 25, 2005.

Accepted March 30, 2006.

¹Corresponding author: bhawani@uoguelph.ca

based DNA fingerprinting method that allows analysis of multiple DNA fragments of different length in a single reaction. The AFLP technique simultaneously screens large numbers of loci for polymorphisms and detects many more polymorphic DNA markers than any other PCR-based detection system. A single enzyme combination may permit the amplification of over 100,000 unique AFLP fragments, thus revealing a large number of polymorphisms. This efficiency raises the prospect of using the AFLP technique to perform whole-genome scanning of markers linked to QTL and thus improve our ability to map QTL and estimate their effects on complex traits like mastitis resistance.

The objectives of this study were, therefore, to use AFLP to identify markers closely linked to QTL for CM resistance, to develop a unique approach for characterization and localization of the QTL-linked AFLP markers, and to estimate the marker effect on functional and production traits.

MATERIALS AND METHODS

Resource Population, DNA Extraction, and Pooling

Animals were selected from a Sentinel Herd project, which was developed through a partnership among the University of Guelph, the Ontario Ministry of Agriculture and Food, the Ontario Dairy Herd Improvement, the Dairy Farmers of Ontario, and the Ontario Association of Bovine Practitioners (Kelton et al., 1999). This study followed 56 Holstein herds located in 23 counties across Ontario for a period of 18 mo from July 1997 to December 1998. Clinical mastitis was diagnosed as a red, swollen appearance of the gland and by appearance of flakes or clots in the milk. A total of 1,070 CM cases were recorded. Whole blood samples were collected from cows. Cows were divided into 3 lactation categories as first, second, and third lactation or greater based on their lactation number at the beginning of the 18-mo study period.

A total of 3,314 Holstein cows were analyzed. Information on CM cases, year in risk, herd, and lactation number were kindly provided by Ledwidge (2003). The factors affecting the number of CM cases were fit by SAS PROC GLM (SAS Institute, 1999) using the following model:

$$Y_{ijk} = H_i + L_j + \beta X_k + e_{ijk},$$

where Y_{ijk} = number of CM cases observed for the k th cow; H_i = fixed effect of i th herd ($i = 1-56$); L_j = fixed effect of j th number of lactation ($j = 1, 2, 3+$); X_k = number of years for which the k th cow remained in the i th herd; β = coefficient of regression on years for which

a cow remained in the herd; and e_{ijk} = random residual effects of CM cases.

Individuals were selected on the basis of the distribution of residual effects from this analysis, which are hereafter referred to as clinical mastitis residual (CMR). Two groups were formed by selecting 100 animals with the lowest CMR as the resistant group and 100 animals with the highest CMR as the susceptible group for CM (Figure 1). Genomic DNA extraction was carried out by standard phenol-chloroform protocol (Winfrey et al., 1997). The DNA concentration and quality were assessed based on the absorbance of UV light at 260 nm (A_{260}) and 280 nm (A_{280}) by micro plate spectrophotometer—SPECTRAMax (Molecular Devices, Sunnyvale, CA). A total of 10 subpools were created, 5 subpools in each group, with each pool containing 20 animals, maintaining average CMR as similar as possible among pools from each group. Equal amounts (200 ng) of DNA from each individual were used. In the CM-resistant group, CMR ranged from -1.19 to -0.79 with mean ($\pm$ SD) of -0.94 ± 0.12 . In the CM-susceptible group, the respective figures were 1.67 to 6.96 , and 2.77 ± 1.10 , respectively. The CM-resistant animals never suffered from CM, whereas the susceptible animals encountered 2 to 8 cases of CM.

The genetic relationship among animals was not considered in analysis because (a) the selected cows (except 5 whose sire identities were not available) were progeny of 121 different sires; the resistant group was composed of progeny of 70 sires, whereas the susceptible group was derived from 64 sires, (b) most of the sires contributed only 1 or 2 progeny, (c) the few sires with multiple progeny were represented in both groups, (d) the average inbreeding coefficient was observed as 0.0314, 0.03, and 0.0329 for total, resistant, and susceptible groups, respectively, and (e) the average number of known ancestors was 6.3 for the resistant as well as the susceptible group.

AFLP Fingerprinting and Analysis

The AFLP typing and analysis was done following the method of Vos et al. (1995) with minor modifications (Ajmone-Marsan et al., 1997). In brief, 200 ng of pooled DNA samples was digested first with *TaqI* and then with *EcoRI* restriction enzymes, followed by ligation to 5 pmol of double-stranded (ds) *EcoRI* adaptor and 50 pmol of ds *TaqI* adaptor in a final volume of 20 μ L, and incubated at 16°C overnight. Subsequently, the reaction mixtures were diluted 10-fold in $T_{10}E_{0.1}$ buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) and amplified in 2 consecutive PCR rounds (preamplification and selective amplification). Then, 1.5 μ L of selective amplified products was mixed with 1.5 μ L of sequencing loading dye

Figure 1. Distribution of residuals for clinical mastitis (CM) cases in the resource population. The boxes indicate fractions selected for creating pools.

with GeneScan-500 ROX-labeled internal molecular size marker of 35 to 500 bp (Applied Biosystems, Foster City, CA) and denatured at 95°C for 2 min. An equal amount of the selective PCR amplicons with internal molecular size marker (GeneScan-500 ROX) from all 10 subpools were loaded in 64-well plates and separated on 5% denaturing polyacrylamide gels on an ABI Prism 377 automated DNA sequencer (PE Applied Biosystems) for 4 h in 0.6× Tris-borate-EDTA buffer. Selective amplification was performed using 89 primer combinations with *EcoRI* and *TaqI* primers, including *EcoRI*+AAC (E32), *EcoRI*+AAG (E33), *EcoRI*+AAT (E34), *EcoRI*+ACA (E35), *EcoRI*+ACC (E36), *EcoRI*+ACG (E37), *EcoRI*+ACT (E38), *EcoRI*+AGA (E39), *EcoRI*+AGC (E40), *EcoRI*+AGG (E41), *EcoRI*+AGT (E42), *EcoRI*+ATA (E43), *EcoRI*+ATC (E44), or *EcoRI*+ATG (E45), and *TaqI*+AAC (T32), *TaqI*+AAG (T33), *TaqI*+ACA (T35), *TaqI*+ACT (T38), *TaqI*+CAC (T48), *TaqI*+CAG (T49), *TaqI*+CAT (T50), or *TaqI*+CCA (T51). All *EcoRI* primers were fluorescently labeled.

The resulting AFLP electropherograms in the range of 75 to 500 bp were analyzed by ABI Prism GeneScan Analysis version 2 followed by ABI Prism Genotyper version 2.5 (PE Applied Biosystems) for each primer pair. The fluorescence units (i.e., peak height; **PH**) of AFLP fragments from each electropherogram were imported separately into Microsoft Excel files for each primer pair set. Fragments showing fewer than 400 fluorescence units were considered either as gel noise or nonspecific PCR products and were not imported.

The PH of 14 standard fragments (GeneScan-500 ROX) of sizes from 75 to 500 bp of each electropherogram was imported independently. A scaling factor (**SF**) was calculated for each primer pair separately to normalize PH in lanes corresponding to a single primer pair. The sum of PH of 14 standard fragments was calculated for all 10 subpools. In each lane the SF was calculated by dividing the sum of PH of 14 standard fragments by a similarly calculated sum in one of the lanes (from one subpool), taken as reference. Finally, the data file for each primer set was prepared with the following columns: gel number, subpools (e.g., resistant pool 1, ..., susceptible pool 5), primer pair (e.g., E32-T32), fragment size (bp), raw PH, SF, and standardized PH (i.e., raw PH/SF). The data files were finally assembled into a single file for statistical analysis.

The standardized PH of AFLP fragments were natural log transformed prior to statistical analysis. An ANOVA was carried out using the following mathematical model and the SAS PROC GLM procedure (SAS Institute, 1999):

$$Y_{ijklmn} = G_i + P_{j(i)} + E_k + O_{l(k)} + F_{m(kl)} + E_k \times P_{j(i)} + P_{j(i)} \times O_{l(k)} + G_i \times F_{m(kl)} + e_{ijklmn},$$

where Y_{ijklmn} = natural log transformed PH; G_i = fixed effect of i th group ($i = 1, 2$ for CM resistant or CM susceptible); $P_{j(i)}$ = random effect of j th pool within i th group ($j = 1 \dots 5$); E_k = random effect of k th gel ($k = 1 \dots 16$); $O_{l(k)}$ = fixed effect of l th primer pair run on k th

Table 1. Primers for genome walking in both directions for 155-bp amplified fragment length polymorphism fragment

Primer name	Primer sequence (5' to 3')	T _{ann} , °C
M _{AFLP} F1	AGGCAAGAGATTCAGGCAAG	57
M _{AFLP} F2	CTGCAGGATGGAGCAAAATC	57
M _{AFLP} F3	ATGACTGTTGGGAGGAGGTG	57
M _{AFLP} R1	CACCTCCTCCCAACAGTCAT	57
M _{AFLP} R2	CTTGATTTTGCTCCATCCTG	56
M _{AFLP} R3	CTTGCTGAATCTCTTGCT	57

¹T_{ann} = Annealing temperature.

gel ($l = 1 \dots 89$); $F_{m(kl)}$ = fixed effect of m th fragment size amplified by l th primer pair run on k th gel; $E_k \times P_{j(i)}$ = interaction of k th gel by j th pool within i th group; $P_{j(i)} \times O_{l(k)}$ = interaction of j th pool within i th group by l th primer pair run on k th gel; $G_i \times F_{m(kl)}$ = interaction of i th group by m th fragment size amplified by l th primer pair run on k th gel; and e_{ijklmn} = random error associated with each observation, assumed normally and independently distributed with zero mean and equal variance (σ_e^2).

Multiple test comparisons were carried out between resistant and susceptible pools for each fragment amplified by various primer combinations by controlling the false discovery rate (FDR; Benjamini and Hochberg, 1995; SAS Institute, 1999).

AFLP Marker Characterization and Association Analysis

The most promising AFLP marker of 155 bp was observed with primer combination *Eco*RI+AGG/*Taq*I+CAG (E41-T49), which was further confirmed with individual AFLP genotyping and the χ^2 test. Two positive and 2 negative individuals for the AFLP 155-bp marker were selected, and their DNA corresponding fragments were amplified by the primer combination E41-T49. The PCR products were electrophoresed on an 8% bisacrylamide gel. The PCR band of 155 bp was carefully excised from the gel, reamplified twice, and confirmed on the ABI Prism 377 automated DNA sequencer before sequencing. Three forward and 3 reverse primers (Table 1) were designed to obtain flanking sequences as per APA Genome Walking kit instructions (Bio S&T, Montreal, Canada). The genomic sequence of 157 bp flanking the *Eco*RI site was obtained. To obtain the flanking sequences to the *Taq*I site, PCR products were cloned into the pCR4-TOPO vector (Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. The cloned DNA fragment of ~330 bp was sequenced with universal primers M13 Forward and M13 Reverse.

The flanking sequences obtained above were used to design externally oriented primers for revealing the

polymorphism in the AFLP marker: forward E₁₅₅F: 5'-TGA CGC AGA ATC CAA AGT TAA AAC A-3', and reverse T₁₅₅R: 5'-GAG GAG GTG GCC GGT TCA GA-3', using the online oligonucleotide design tool Primer3 (Rozen and Skaletsky, 1998). The DNA from 2 positive and 2 negative AFLP-typed individuals were amplified by PCR and contained 50 ng of genomic DNA, 20 ng of E₁₅₅F and T₁₅₅R, 1× PCR buffer, 1.5 mM MgCl₂, 200 μM of each dNTP, and 0.75 U of AmpliTaq Gold DNA polymerase (Applied Biosystems). Touchdown PCR was performed at an annealing temperature of 60°C. The PCR products of 399 bp were submitted for sequencing. Sequences were compared, and a single nucleotide A↔G mutation at the *Taq*I cut site was identified.

In fact, this A↔G mutation created a second *Taq*I restriction site within the AFLP fragment, which was used to develop a *Taq*I RFLP. The PCR products (5 μL) were incubated in a final volume of 10 μL with 1× *Taq*I buffer, 0.25 μg of BSA and 5 U of *Taq*I at 65°C for 90 min. The restriction fragments (125 bp/274 bp; 39 bp/125 bp/235 bp) were resolved on a 2% agarose gel from 4 sequenced animals. Finally, all 200 individuals were PCR-RFLP genotyped. Genotypic and allele frequencies of the marker were calculated for both CMR groups and tested for independence using a χ^2 test.

Cow EBV and ETA were obtained from the national genetic evaluation of February 2004 from the Canadian Dairy Network (Guelph, Canada) and were used for the estimation of marker effects on various production and functional traits. This information was accessed for 194 cows, which consisted of 6 AA, 118 AG, and 70 GG genotypes. A general trait-based analysis, which accounts for selective genotyping, was carried out by logistic regression (Henshall and Goddard, 1999). Because the frequency of genotype AA was only 3% (a total of only 6 individuals), only the most frequent genotype classes (AG and GG) were used for the estimation of marker-associated effects. The following model was used for predicting the genotype of an individual as a response variable with single EBV, ETA, and CMR as explanatory variables:

$$\text{logit } [p_{(AG)}] = \log[p_{(AG)}/1-p_{(AG)}] = a + bX,$$

where p = probability of AG genotype; a = intercept; b = slope of curve or regression coefficient; and X = EBV or ETA or CMR for trait of interest.

The effect of the AG genotype relative to the GG genotype on various traits was estimated by principles described by Henshall and Goddard (1999) as

$$2 \left[\left\{ -1 + \sqrt{1 + b^2 \sigma^2} \right\} / b \right],$$

where σ^2 is an estimate of variance for a trait from all the data from which individuals were selected. The significance of effects was determined on the basis of significance level of b , given their direct relationship (Henshall and Goddard, 1999).

To control false positives when testing the marker for association with multiple traits, a modified Bonferroni correction was applied. The effective number of multiple tests was determined through principal component analysis of the correlation matrix of traits under study. The first 5 principal components accounted for more than 91% of the total variation. Therefore, 5 effective tests were assumed when determining the level of significance for a single marker; hence, final level of significance was determined as 0.01 (the critical P -value $0.05/5 = 0.01$).

Chromosomal Assignment

The complete sequence of the marker was compared with known genes via BLASTn (Altschul et al., 1990) at NCBI (<http://www.ncbi.nlm.nih.gov/>), which showed some degree of homology with the bovine lactoferrin (**Lf**) gene. Primers based on bovine *Lf* were designed using NCBI and the online oligonucleotide design tool Primer3. The PCR was performed on 94 cell lines of a

3,000-rad bovine/hamster radiation hybrid (**RH**) panel (ResGen Invitrogen Life Technologies) in a 10- μ L reaction volume containing 50 ng of DNA of the hybrid cell clones, 20 ng of forward and reverse primers (LTF: 5'-CGT ACT TGA GCT GGA CAG AGT CAC-3'; LTR: 5'-GGG TAT GCT TGT CTA TCA ATG CAG-3'), 200 μ M of each dNTP, 2 mM $MgCl_2$, 1 $\times$ PCR buffer, and 0.75 U of AmpliTaq Gold DNA polymerase. Step-down PCR was performed at an annealing temperature of 60°C. The AFLP marker was typed on the RH panel using E₁₅₅F and T₁₅₅R primers. The RH typing data for the AFLP marker and bovine *Lf* were analyzed by RMAP 3.0 statistical software (Boehnke et al., 1996). Locus-specific retention probabilities and 2-point LOD scores for linkage were estimated by RH2PT (Boehnke et al., 1996).

RESULTS

QTL-Linked AFLP Markers for Mastitis Resistance

In this study, 89 primer pairs specific to *Eco*RI+ANN and *Taq*I+(A/C)NN were used in the AFLP analysis on 5 subpools of CM-resistant and CM-susceptible groups. A total of 2,829 scorable bands were generated in the range of 75 to 500 bp, with an average of 31.8 bands

Table 2. A list of amplified fragment length polymorphism markers showing most striking difference in peak height (PH) between clinical mastitis-resistant and clinical mastitis-susceptible groups

Primer pair	Size, bp	Mean log-transformed PH		<i>F</i> -value	<i>P</i> -value	FDR-value
		Resistant group	Susceptible group			
E41-T49	155	7.651	7.183	36.72	<0.00001	<0.0001
E32-T49	94	6.929	7.332	27.23	<0.00001	0.0002
E32-T49	421	6.471	6.859	25.17	<0.00001	0.0005
E32-T35	251	6.979	7.349	22.97	<0.00001	0.0011
E42-T48	109	6.786	6.496	20.96	0.00001	0.0024
E39-T49	384	6.161	6.513	20.82	0.00001	0.0024
E41-T32	352	6.845	6.514	18.29	0.00002	0.0077
E36-T48	96	6.231	6.556	17.77	0.00003	0.0088
E35-T50	267	6.091	6.405	16.57	0.00005	0.0148
E39-T50	214	6.744	7.052	15.94	0.00007	0.0160
E39-T51	93	6.313	6.619	15.73	0.00007	0.0160
E40-T49	173	6.025	6.331	15.72	0.00007	0.0160
E39-T35	206	6.508	6.809	15.19	0.00010	0.0197
E33-T32	273	6.785	6.483	14.59	0.00013	0.0248
E40-T48	263	6.394	6.100	14.47	0.00014	0.0248
E38-T32	209	6.208	6.500	14.34	0.00015	0.0248
E36-T48	252	6.486	6.778	14.28	0.00016	0.0248
E38-T50	395	6.726	6.437	14.00	0.00018	0.0273
E36-T33	113	6.230	6.517	13.81	0.00020	0.0273
E37-T49	240	7.551	7.838	13.81	0.00020	0.0273
E32-T33	211	6.697	6.982	13.55	0.00023	0.0300
E35-T32	110	6.469	6.188	13.31	0.00026	0.0325
E35-T32	99	7.434	7.154	13.19	0.00028	0.0333
E38-T50	179	7.212	6.935	12.88	0.00033	0.0376
E32-T32	357	6.499	6.775	12.77	0.00035	0.0384
E44-T50	184	6.362	6.635	12.54	0.00040	0.0417
E36-T33	223	6.525	6.796	12.32	0.00045	0.0455

Figure 2. Comparative electropherograms of clinical mastitis- (CM) resistant and CM-susceptible pools amplified by *EcoRI*+AGG/*TaqI*+CAG primer pair. The x-axis represents the size (bp) of the AFLP fragment, and the y-axis shows the amount of the amplified products in arbitrary units.

per primer pair combination and a range of 8 to 64 bands. Primer pair combinations using *EcoRI*+AG(A/T) and *EcoRI*+AG(G/T) as the *EcoRI*-specific primer generated more AFLP fragments, whereas *EcoRI*+AC(C/G) and *EcoRI*+AAC produced fewer fragments. Then ANOVA was carried out on log-transformed standardized PH of AFLP fragments, and contrasts between resistant and susceptible pools for each AFLP fragment were constructed from the group $\times$ fragment size interaction ($G_i \times F_{m(kl)}$). Multiple test comparisons were carried out by controlling the FDR, and a set of 27 significant ($FDR < 0.05$) AFLP fragments were identified (Table 2). These promising AFLP markers are potential candidates for further study and could be useful for marker-assisted selection programs for CM resistance.

Confirmation and Characterization of a QTL-Linked AFLP Marker

The most significant AFLP marker ($FDR < 0.0001$) was amplified by the primer combination of *EcoRI*+AGG/*TaqI*+CAG at size of 155 bp (Table 2), showing a marked difference in PH between resistant and susceptible groups (Figure 2). All 200 individuals were typed individually for the marker using AFLP technique. Frequency of the AFLP 155-bp marker appeared significantly higher in resistant animals than that in susceptible ones (79 vs. 51; $P < 0.0001$). The

155-bp fragment was sequenced, and a 618-bp genomic region was successfully obtained and submitted to GenBank (Accession no. AY741138). The comparison of sequence information revealed a single nucleotide polymorphism (A $\leftrightarrow$ G) at position bp 41 of the AFLP fragment toward *TaqI* site, which was named CGIL4.

PCR-RFLP Genotyping

A new primer pair (E_{155}^F and T_{155}^R) was designed to amplify a fragment of 399 bp for detection of the A $\leftrightarrow$ G substitution by restriction digestion with *TaqI*. Individuals having a G residue had a second 5'-TCGA-3' sequence, which was cleaved by the *TaqI*, thus producing 3 size fragments of 39, 125, and 235 bp; whereas individuals with an A residue yielded 2 fragments of molecular sizes of 125 and 274 bp (39 bp + 235 bp; Figure 3). The PCR-RFLP technique correctly identified the original AFLP polymorphism; however, discrepancies were observed in 4 animals. This problem might have been due to incomplete digestion of DNA in both the methods or failure in the ligation of adaptors in the AFLP method. These discrepancies were finally resolved by sequencing. The sequencing results showed that out of the 4 animals with discrepant results, 2 AFLP and 2 PCR-RFLP observations were correct.

Genotypic frequencies were compared between CM-resistant and CM-susceptible groups. The majority of the resistant animals were found to be heterozygous

Figure 3. The PCR-RFLP differentiating all 3 genotypes at the promising 155-bp amplified fragment length polymorphism marker.

for the mutation, and susceptible animals were more frequently homozygous for the G residue. The allele frequencies were quite different between resistant and susceptible groups (Table 3), suggesting that a QTL for CM resistance is closely linked to the marker.

Estimation of EBV and ETA on Predicting Genotypes and Marker-Associated Effects

The marker was tested for its association with production and functional traits using 14 different EBV and 8 different ETA values. The marker was also tested for CMR. The EBV and ETA information were accessed for 194 cows, of which 6 animals were homozygous AA, 118 were heterozygous, and 70 animals were homozygous GG. Given the small number of homozygous AA cows, these were excluded from further analysis. In a trait-based approach, marker effects were estimated by a logistic regression model for a binary response variable (i.e., genotype).

The maximum likelihood estimates of regression coefficients (b) for the various traits and their significance, where the AG genotype was considered as success (i.e., 1) and the GG genotype as failure (i.e., 0), were obtained. The most significant effect was found for ETA for SCS in the first lactation ($P < 0.001$), followed by CMR ($P < 0.003$), ETA for SCS ($P < 0.007$), EBV fat percentage ($P < 0.018$), EBV for milk in first lactation ($P < 0.028$), and SCS in the second lactation ($P < 0.039$).

Probability of an individual carrying the GG genotype significantly increased with high CMR, ETA for SCS in the first and second lactations, and overall, and EBV for milk in the first lactation, whereas it significantly decreased with high EBV for fat percentage. A transformation of the logistic regression coefficients yielded estimates of the marker genotype on various traits.

Our primary interest in the CGIL4 marker lies in the health traits; that is, CMR and SCS. Although QTL linked to CGIL4 had a very significant effect on CMR ($P < 0.003$), the average difference between GG animals and AG animals was small in standard deviation units compared with other traits (0.19 SD, Table 4). The coefficient of variation was very high for CMR compared with other traits. This is due to skewed distribution of CMR values in combination with the mean near zero (Figure 1). Interestingly, the marker-associated effects on SCS were just as strong as for CMR. The ETA for SCS of GG cows was close to 0.1 units and 0.5 SD higher than AG cows for first lactation and overall performance. The ETA for SCS in the third lactation was affected at a lower level of significance ($P < 0.12$) with GG cows being higher than AG cows. In general, cows with GG genotypes had higher SCS during their lactations as compared with cows carrying AG genotypes (Table 4).

The GG genotypes were superior to AG by a similar magnitude for milk yield in the first lactation. In contrast, a similar effect was also observed for fat percentage; however, AG genotypes were superior. Given the lack of significance of the logistic regression for the rest of the 12 production trait EBV and persistency ETA, the QTL-linked marker effects were not estimated for these traits. However, prior to Bonferroni correction, EBV for protein yield in the first lactation was affected significantly at the 2% level, and the difference between GG and AG genotypes was 8.13 kg.

Chromosomal Assignment

A BLAST (<http://ncbi.nih.gov/BLAST/>) search using the marker sequence (AY741138) as a query against the GenBank nonredundant databases revealed its homology to partial 14th intron sequences of the bovine

Table 3. Genotypic and allele frequencies of the most promising amplified fragment length polymorphism marker in clinical mastitis-resistant and clinical mastitis-susceptible animals

Group	Genotypic frequency			Allele frequency	
	AA	AG	GG	A	G
Resistant	3	77	20	83 (42%)	117 (58%)
Susceptible	3	46	51	52 (26%)	148 (74%)
Test-statistic		$\chi^2 = 20.16$		$\chi^2 = 10.74$	
P-value		$P < 0.0001$		$P < 0.001$	

Table 4. Average difference between amplified fragment length polymorphism marker genotypes for significant traits

Trait ¹	Difference GG vs. AG		Corrected significance ²
	General units	SD units	
EBV Milk1, kg	345	0.44	$P < 0.028$
EBV Fat percentage	-0.11	0.44	$P < 0.018$
ETA SCS	0.10	0.49	$P < 0.007$
ETA SCS1	0.11	0.53	$P < 0.001$
ETA SCS2	0.09	0.42	$P < 0.039$
CMR ³	0.13	0.19	$P < 0.003$

¹Numbers following traits indicate lactation number.

²Bonferroni correction for 5 independent tests.

³CMR = Residuals for clinical mastitis cases after accounting for effects of herd, lactation number, and years on study.

Lf gene (L19990). Three aligned blocks were identified, one with 84% sequence identity over a 73-bp sequence, second with 87% sequence identity over a 57-bp sequence, and a third with 96% sequence identity over a 26-bp sequence. Obviously, the AFLP marker does not represent the bovine *Lf* gene, but we thought they might be correlated in some way. Fortunately, genotyping both the *Lf* gene and the AFLP marker on 94 cell lines of the 3,000-rad bovine-hamster RH panel demonstrated their close linkage with a logarithm of the odds (LOD) score of 8.90 based on the analysis using the RHMAP (3.0) program. The retention rate was 0.245 for the AFLP marker and 0.340 for the bovine *Lf* gene, respectively. Recently, Griffin et al. (2005) placed this AFLP marker in between the *Lf* and tetranectin (*TNA*) genes and assigned them with 9 other markers to the BTA22q24 region. In addition, we have obtained the same gene sequence (AAFC02089984) for the AFLP marker by using a BLAST search against the 6× bovine genome sequence database (<http://www.hgsc.bcm.tmc.edu/projects/bovine/>). The 98% sequence identity between AY741138 and AAFC02089984 indicated that our AFLP flanking walking was very successful.

DISCUSSION

The aim of the present study was to evaluate the usefulness of AFLP typing of selective DNA pools for QTL detection and characterization for CM in Canadian Holsteins. Because of the very low heritability of CM (Pösö and Mäntysaari, 1996), this study can provide useful markers for selection against CM susceptibility. In the future, such markers may be used to track QTL and to reveal the complex molecular mechanism(s) responsible for the CM resistance.

Breeding for increased resistance to mastitis is usually based on indirect selection. Somatic cell count is considered as the most useful measure of CM resis-

tance, but it also includes incidences of subclinical mastitis. The genetic correlation between CM and SCC averages 0.7 (Pösö and Mäntysaari, 1996; Heringstad et al., 2000), indicating that although both are expressions of udder health, they are not expressions of the same trait. Variation in this correlation among lactations also suggests that CM and SCC monitor different characteristics of udder health (Pösö and Mäntysaari, 1996). Many studies have noted that CM is mainly due to environmental pathogens (Schukken et al., 1997). Hence, a study of clinical infections in cows would require direct observation of their occurrence. In the present study, selection of cows was based on extreme values of the residuals of CM cases after accounting for the effects of herd, lactation number, and number of years for which a cow remained in the herd. This sampling strategy took into account environmental as well as lactation differences, allowing us to look more clearly at the occurrence of CM of an animal within a herd.

The AFLP selective genotyping of DNA pools of 2 groups constructed from extreme individuals represents a case-control (i.e., association) study. In the present study, a high number of fragments ($n = 27$) was identified when comparisons were made between the resistant and susceptible DNA pools. This high number of polymorphic fragments might be attributed to the use of *TaqI* restriction enzyme. The *TaqI* site contains a CpG dinucleotide within its sequence, which is prone to mutation in vertebrate DNA due to the transition of methylated cytosine to thymine (Ajmone-Marsan et al., 1997). Combinations of *EcoRI* (5'-G/AATTC) and *TaqI* (5'-T/CGA-3') appear to be the most suitable for producing AFLP markers for the animal genome (Ajmone-Marsan et al., 1997). This high number of identified markers may also be because of the threshold that was set for the multiple test comparisons. When the FDR was set to 1%, instead of 5%, only 9 fragments were identified. Finally, 2 or more bands may have been generated by 1 primer pair representing the same region of the genome. Hence, characterization and mapping of each of these markers on the bovine genome is needed to confirm their possible segregation with CM-resistance QTL.

Several studies have reported a large number of QTL for health, SCS, and CM on many different chromosomes (e.g., BTA3, 4, 6, 7, 8, 12, 13, 14, 16, 18, 19, 21, 23, 27; Zhang et al., 1998; Schrooten et al., 2000; Klungland et al., 2001; Ashwell et al., 2004; Schulman et al., 2004). Pan (2000) reported 4 AFLP fragments associated with SCS in Canadian Holsteins. Generally, quantitative traits are multifactorial and are influenced by several genes as well as environmental conditions; hence, one or many QTL can affect a trait or a phenotype.

Among these significant AFLP fragments, individual genotyping of all 200 cows was performed for the most promising 155-bp fragment. The occurrence of this fragment between the 2 groups appeared in accordance with the PH. The 155-bp fragment was observed in 79 individuals from the CM-resistant group and only in 51 individuals from the CM-susceptible group. The log-transformed standardized PH in the 5 subpools of the resistant group were obtained as 7.83, 7.89, 7.54, 7.61, and 7.38, and individual AFLP typing revealed the presence of the 155-bp AFLP among 18, 19, 16, 14, and 12 of the cattle, respectively. In the 5 subpools of the susceptible group, the PH were 7.15, 7.40, 7.20, 7.19, and 6.97, and individual typing showed the 155-bp AFLP present in 8, 11, 12, 9, and 11 individuals, respectively. These findings, along with similar observations obtained in another study (Sharma et al., 2003), strengthen the view that the appearance of AFLP bands in banding patterns derived from DNA pools reflect the relative band frequency.

With the exception of Pan (2000), no other study has combined selective DNA pooling and AFLP. In Pan's (2000) study, however, further characterization of AFLP fragments was not carried out. Several studies have reported a direct relationship between allele frequency and band intensity or PH in DNA pools for microsatellite markers (Breen et al., 1999; Lipkin et al., 2002) and for single nucleotide polymorphisms (Chen et al., 2002). In the present study, presumably the contrast between standardized PH obtained from all replicates of extreme DNA pools, which were processed and electrophoresed at the same time, provides a close estimate of relative frequency of an AFLP marker; however, this was not explicitly validated.

Although AFLP is a novel and very powerful multilocus DNA fingerprinting technique with high multiplex ratio, it is limited by its dominant and anonymous nature. It is also too expensive and laborious for practical genotyping of large numbers of animals. To confirm the presence of a QTL in a region surrounding a potential marker, a test for linkage is needed. To facilitate its wider applicability and linkage analysis, we derived a simple PCR-based genotyping method for the most prominent AFLP marker. Simple PCR markers are more suitable for large-scale genotyping and can be easily applied in marker-assisted selection. Through genome walking and cloning methods, this AFLP marker was identified as an A $\leftrightarrow$ G transition single nucleotide polymorphism. Few other studies have reported isolation of single nucleotide polymorphisms from AFLP fragments by band-specific genome walking and cloning of genomic region harboring a portion of AFLP sequence (Wimmers et al., 2002). Brugmans et al. (2003) presented a flow chart for the procedure to

convert AFLP markers into allele-specific PCR-based markers, which involves many sequential steps. However, they reported low success rates in searching for polymorphisms internal to the AFLP fragment.

The AFLP marker CGIL4 showed significant association with mastitis-related traits. Effects of the QTL linked to the marker were high for CMR and SCS in the first and second lactations and on cumulative SCS. This suggests that some gene variants are affecting both SCS and CM. The high genetic correlation between SCS and CM is well established. Milk yield was also affected by a nearby QTL, which is consistent with the unfavorable positive correlation between both the traits. The decreased fat percentage in susceptible cows could be explained to some extent by an increase in milk yield without any increase in fat yield. It has also been reported that total protein content may undergo little change in cows with mastitis (Harmon, 1994).

The majority of cows were heterozygous for CGIL4, and the overall heterozygosity was 61%. The cows with GG genotypes had high CMR, SCS, and milk yield in their first lactation. The Canadian dairy industry has generally been selecting dairy cattle for high yield of milk and protein. The findings of the present study suggest that selection for milk yield might have led to the selection of a QTL allele that has antagonistic effects on udder health and disease resistance.

In the present study, CM cases were identified irrespective of whether a microbial pathogen was isolated from the sample. There is some evidence that susceptibility to specific pathogens may be due to different genetic traits (Wilton et al., 1972). Hence, it is suggested that in future, CGIL4 should be studied on CM cases diagnosed on the basis of microbiological examination.

The ultimate objective of any QTL study is mapping and locating QTL and positional cloning of causative gene(s). The first step is to know the chromosomal location of promising markers and their neighborhoods. Radiation hybrid panels have been shown to be very useful for physical mapping of markers in numerous species. Recent advances in genome sequence analysis support the view that genes of related function, or those that work in the same pathway, are sometimes physically close to each other and share some degree of homology (Cohen et al., 2000; Geraghty, 2002). The *Lf* gene has been shown to be associated with the orderly expression of the genetic program leading to an antibacterial action during inflammation (Sanchez et al., 1992; Seyfert et al., 1994); hence, linkage analysis was performed between markers and bovine *Lf* by RH mapping. This method assigned the CGIL4 marker to BTA22, in close proximity to the *Lf* gene. Because of a paucity of data on CM, few QTL searches for this trait have been published to date. In these studies, few chromosomes have

been shown to carry QTL, including BTA3, 4, 6, 14, 27 (Klungland et al., 2001), and BTA14 and 18 (Schulman et al., 2004). To our knowledge, the present study is the first report showing QTL affecting CM on BTA22. The findings also suggest linkage of this QTL to the bovine *Lf* gene. This chromosome can be exploited for candidate genes through fine mapping of the region, and expression studies can be carried out for its confirmation.

In conclusion, the present study has demonstrated the feasibility of assaying AFLP on DNA pools to reveal markers segregating with QTL in extreme CM groups of individuals. In contrast to QTL detection and mapping by genome scans performed with microsatellites, this approach provides the opportunity to generate new markers and to reveal new candidate genes surrounding QTL regions. A set of 27 AFLP markers was found to be associated with mastitis resistance. The characterized marker CGIL4 may be utilized in marker-assisted selection for mastitis resistance, particularly in breeds or populations lacking recording of CM. A detailed study of the genome around the CGIL4, including the bovine *Lf* gene, may provide more insight into coexpression of genes and a better understanding of the putative QTL in this region.

ACKNOWLEDGMENTS

Financial assistance was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Dairy Cattle Genetics Research and Development Council (DairyGen; CRD245792-01 to Z.J.), Canada. The Sentinel Herd Project was funded by Dairy Farmers of Ontario and OMAF by a one-time grant called Grow Ontario. Jalal Fatehi and Margaret Quinton (Dept. Anim. Poult. Sci., Univ. Guelph, Canada) are acknowledged for their help in pedigree construction and statistical analyses.

REFERENCES

- Ajmone-Marsan, P., A. Vlaentini, M. Cassandro, G. Vecchiotti-Analdi, G. Bertoni, and M. Kuiper. 1997. AFLP™ markers for DNA fingerprinting in cattle. *Anim. Genet.* 28:418–426.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
- Ashwell, M. S., D. W. Heyen, T. S. Sonstegard, C. P. Van Tassell, Y. Da, P. M. VanRaden, M. Ron, J. I. Weller, and H. A. Lewin. 2004. Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. *J. Dairy Sci.* 87:468–475.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate—A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.* 57:289–300.
- Boehnke, M., K. Lunetta, E. Hauser, K. Lange, J. Uro, and J. Vandersloot. 1996. RHMAP: Statistical package for multipoint radiation hybrid mapping (version 3.0). Available <http://csg.sph.umich.edu/boehnke/rhmap.php> Accessed 2003.
- Breen, G., P. Sham, T. Li, D. Shaw, D. A. Collier, and D. St. Clair. 1999. Accuracy and sensitivity of DNA pooling with microsatellite repeats using capillary electrophoresis. *Mol. Cell. Probes* 13:359–365.
- Brugmans, B., R. G. M. Van der Hulst, R. G. F. Visser, P. Lindhout, and H. J. Van Eck. 2003. A new and versatile method for the successful conversion of AFLP markers into simple single locus markers. *Nucleic Acids Res.* 31:e55.
- Chen, J., S. Germer, R. Higuchi, G. Berkowitz, J. Godbold, and J. G. Wetmur. 2002. Kinetic polymerase chain reaction on pooled DNA: A high-throughput, high-efficiency alternate in genetic epidemiological studies. *Cancer Epidemiol. Biomarkers Prev.* 11:131–136.
- Cohen, B. A., R. D. Mitra, J. D. Hughes, and G. M. Church. 2000. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat. Genet.* 26:183–186.
- De Jong, G., and L. M. T. E. Lansbergen. 1996. Udder health index: Selection for mastitis resistance. International Workshop on Genetic Improvement of Functional Traits in Cattle. *Interbull Bull.* 12:42–47.
- Geraghty, D. E. 2002. Genetic diversity and genomics of the immune response. *Immunol. Rev.* 190:5–8.
- Griffin, K. B., J. J. Michal, L. K. Fox, C. T. Gaskins, and Z. Jiang. 2005. Fine mapping of the bovine chromosome 22q24 region that harbours antimicrobial genes and a QTL for somatic cell score. *Anim. Genet.* 36:448–450.
- Harmon, R. J. 1994. Physiology of mastitis and factors affecting somatic cell counts. *J. Dairy Sci.* 77:2103–2112.
- Henshall, J. M., and M. E. Goddard. 1999. Multiple-trait mapping of quantitative trait loci after selective genotyping using logistic regression. *Genetics* 151:885–894.
- Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: A review with focus on the situation in the Nordic countries. *Livest. Prod. Sci.* 64:95–106.
- Kelton, D., A. Godkin, D. Alves, K. Lissemore, K. Leslie, N. Smart, C. Church, and P. Meadows. 1999. Prevalence of major mastitis pathogens on Ontario Sentinel dairy farms. Pages 257–258 in *Proc. 32nd Annu. Conf. Am. Assoc. Bovine Pract.*, Nashville, TN.
- Klungland, H., A. Sabry, B. Heringstad, H. G. Olsen, L. Gomez-Raya, D. I. Vage, I. Olsaker, J. Odegard, G. Klemetsdal, N. Schulman, J. Vilkki, J. Ruane, M. Aasland, K. Ronningen, and S. Lien. 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. *Mamm. Genome* 12:837–842.
- Kossaibati, M. A., M. Hovi, and R. J. Esslemont. 1998. Incidence of clinical mastitis in dairy herds in England. *Vet. Rec.* 143:649–653.
- Ledwidge, S. A. 2003. Associations between specific bovine leukocyte antigen DRB3 alleles and mastitis in Canadian Holsteins. M. S. Thesis, Univ. Guelph, Guelph, Canada.
- Lipkin, E., H. Fulton, H. Cheng, N. Yonash, and M. Soller. 2002. Quantitative trait locus mapping in chickens by selective DNA pooling with dinucleotide microsatellite markers by using purified DNA and fresh or frozen red blood cells as applied to marker-assisted selection. *Poult. Sci.* 81:283–292.
- Long, A. D., M. N. Grote, and C. H. Langley. 1997. Genetic analysis of complex diseases. *Science* 275:1328.
- Pan, Y. 2000. Identification and examination of molecular genetic markers associated with somatic cell scores of Ontario Holstein cattle. Ph.D. Thesis, Univ. Guelph, Guelph, Canada.
- Pösö, J., and E. A. Mäntysaari. 1996. Relationship between clinical mastitis, somatic cell score, and production for the first three lactations of Finnish Ayrshire. *J. Dairy Sci.* 79:1284–1291.
- Prochazka, M., K. Walder, and J. Xia. 2001. AFLP fingerprinting of the human genome. *Hum. Genet.* 108:59–65.
- Rish, N., and K. Merikangas. 1996. The future of genetic studies of complex human diseases. *Science* 273:1516–1517.
- Rozen, S., and H. J. Skaletsky. 1998. Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html Accessed 2003.
- Sanchez, L., M. Calvo, and J. H. Brock. 1992. Biological role of lactoferrin. *Arch. Dis. Child.* 67:657–661.
- SAS Institute. 1999. SAS User's Guide. Statistics, Version 8 Developer's Edition. SAS Inst. Inc., Cary, NC.

- Schrooten, C., H. Bovenhuis, W. Coppieters, and J. A. Arendonk. 2000. Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *J. Dairy Sci.* 83:795–806.
- Schukken, Y. H., T. J. G. M. Lam, and H. W. Barkema. 1997. Biological basis for selection on udder health traits. *Interbull Bull.* 15:27–32.
- Schulman, N. F., S. M. Viitala, D. J. de Koning, J. Virta, A. Maki-Tanila, and J. H. Vilkkii. 2004. Quantitative trait loci for health traits in Finnish Ayrshire cattle. *J. Dairy Sci.* 87:443–449.
- Schutz, M. M. 1994. Genetic evaluation of somatic cell scores for United States dairy cattle. *J. Dairy Sci.* 77:2113–2129.
- Seyfert, H. M., A. Tuckoricz, H. Interhal, D. Koczan, and G. Hobom. 1994. Structure of the bovine lactoferrin-encoding gene and its promoter. *Gene* 143:265–269.
- Sharma, B. S., Z. Jiang, and G. B. Jansen. 2003. Identification and characterization of an AFLP marker for protein yield in Canadian Holsteins. *J. Anim. Sci.* 81(Suppl. 1):319. (Abstr.)
- Vos, P., R. Hogers, M. Bleekers, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Wilton, J. W., L. D. Van Vleck, R. W. Everett, R. S. Guthrie, and S. J. Roberts. 1972. Genetic and environmental aspects of udder infections. *J. Dairy Sci.* 55:183–193.
- Wimmers, K., E. Murani, S. Ponsuksili, M. Yerle, and K. Schellander. 2002. Detection of quantitative trait loci for carcass traits in the pig by using AFLP. *Mamm. Genome* 13:206–210.
- Winfrey, M. R., M. A. Rott, and A. T. Wortman. 1997. *Unraveling DNA: Molecular Biology for the Laboratory*. Prentice Hall, Upper Saddle River, NJ.
- Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen, B. Murkve, M. Pfister-Genskow, L. A. Witte, F. E. Grignola, P. Uimari, G. Thaller, and M. D. Bishop. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* 149:1959–1973.