Impacts of using inbred animals in studies for detection of quantitative trait loci

G. Freyer,^{*1} N. Vukasinovic,[†] and B. Cassell[‡]

^{*}FBN Research Unit Genetics and Biometry, D-18196 Dummerstorf, Germany

[†]Newsham Choice Genetics, 701 Crown Industrial Ct., Ste. H, Chesterfield, MO 63005

[‡]Virginia Tech, Dairy Science Department, Blacksburg 24061

ABSTRACT

Effects of utilizing inbred and noninbred family structures in experiments for detection of quantitative trait loci (QTL) were compared in this simulation study. Simulations were based on a general pedigree design originating from 2 unrelated sires. A variance component approach of mapping QTL was applied to simulated data that reflected common family structures from dairy populations. Five different family structures were considered: FS0 without inbreeding, FS1 with an inbred sire from an aunt-nephew mating, FS2 with an inbred sire originating from a half-sib mating, FS3 and FS4 based on FS2 but containing an increased number of offspring of the inbred sire (FS3), and another extremely inbred sire with its final offspring (FS4). Sixty replicates each of the 5 family structures in 2 simulation scenarios each were analyzed to provide a praxis-like situation of QTL analysis. The largest proportion of QTL position estimates within the correct interval of 3 cM, best test statistic profiles and the smallest average bias were obtained from the pedigrees described by FS4 and FS2. The approach does not depend on the kind and number of genetic markers. Inbreeding is not a recommended practice for commercial dairy production because of possible inbreeding depression, but inbred animals and their offspring that already exist could be advantageous for QTL mapping, because of reduced genetic variance in inbred parents.

Key words: quantitative trait locus mapping, inbreeding, dairy cattle, variance components model

INTRODUCTION

Inbreeding is increasing in dairy populations. Selection for a relatively consistent breeding objective by many breeders has reduced the number of highly influential sires in dairy populations and increased the

average relationship between prospective mates in the process. Hansen (2006) showed that 2 bulls born in the 1960s (Elevation and Chief) were the source of about 30% of the genes segregating in the US Holstein population in 2005. Known offspring of these 2 sires have an enormous impact on the present Holstein breed. The annual increase in inbreeding is 0.15 and 0.20% for Holsteins and Jerseys in the United States, respectively (Hansen, 2006). Increasing inbreeding poses problems to dairy farmers because of reduced fitness, fertility, and longevity of dairy cows—traits with relatively low heritability as shown by many studies (e.g., Thompson et al., 2000). Increased inbreeding is also believed to be a problem in QTL mapping, because of increasing homozygosity in the mapping population and reduction in allelic variance at markers and QTL sites. The fundamental conditions necessary for QTL mapping are parental heterozygosity at markers and QTL, and detectable QTL effects.

Experimental populations in dairy cattle designed exclusively for QTL mapping purposes have not been widely used because of costs and time requirements. Instead, QTL studies are performed using data from commercial dairy populations. The most popular design for QTL mapping in dairy cattle has been the grand-daughter design as described first by Weller et al. (1990). Each grandsire family is treated separately, thus ignoring relationships that exist between individual families, and potentially dismissing useful pedigree information, because individuals descending from related parents are assumed to be unrelated. Several approaches have been proposed to cope with inbreeding in pedigrees such as cutting inbreeding (or marriage) loops in a pedigree (e.g., Fernando et al., 1993), since direct processing of inbred individuals is computationally demanding.

Methods for QTL mapping based on identity by descent (IBD) probability as adapted to general pedigrees have been developed in recent years (e.g., Meuwissen and Goddard, 2000; Pong-Wong et al., 2001). Hernández-Sánchez et al. (2006) stated that “Genetic relatedness between individuals is a key parameter in QTL mapping via variance component estimation.”

Received June 21, 2007.

Accepted October 15, 2008.

¹Corresponding author: freyer@fbn-dummerstorf.de

However, advanced solutions for fine mapping that consider combined linkage and linkage disequilibrium analysis in dairy cattle do not distinguish between IBD from inbreeding and co-ancestry (Hernández-Sánchez et al., 2006). To date, there is no comparative study of the performance of methods for QTL mapping in practical applications in dairy cattle considering inbreeding. There is a need to evaluate different approaches with differing intrinsic pedigree structure involving inbreeding resulting from mating related individuals under practical conditions.

In this study, we consider 5 family structures, differing with regard to relationship of single mates and increased load of inbreeding, for QTL mapping. Specifically, data sets with 1 or 2 inbred animals and their offspring for QTL mapping were examined. Pedigrees of 4 to 5 generations and typical of real dairy data were of interest. The effect of using inbred sires on QTL mapping in dairy cattle is targeted by a simulation study.

MATERIALS AND METHODS

Simulated Mapping Population

A general pedigree structure with 5 generations including 850 individuals in total was simulated. Final offspring descended from 2 great-grandires (**GGS**; GGS1 and GGS2), followed by 4 grandsires (**GS**; GS7 to GS10), and 8 sires (**S**; S20 to S27). Overlapping generations were also included in that one great-grandire, GGS1, was the sire of 69 final offspring. For simplicity, we call both male founders “great-grandires,” regardless of overlapping generations. A total of 544 offspring descended from 9 sires, with 42 to 78 offspring per sire. Different family structures were created by 4 single-mating designs, determining the inbreeding level of particular individuals in the pedigree of the 5 family structures (**FS**) studied: FS0 included no inbreeding, as no parents were related; family structures FS1, FS2, FS3, and FS4 contained one inbred grandsire, GS9, originating from a sire-daughter mating, who produced 2 noninbred sires (sires S24 and S25) with a total of 89 final (noninbred) offspring. Sire-daughter matings are rare in dairy cattle, but do occur. This sire was included to force the analysis approach to cope with more complicated situations.

Family structures 1 and 2 differed in the way the origin of sire S20, with 78 offspring, was simulated,

- in FS1, from an aunt-nephew mating, considered mildly inbred, Wright's inbreeding coefficient ($F_x = 0.0625$),
- in FS2, from mating half-sibs, considered strongly inbred, $F_x = 0.125$,

- in FS3, the same structure as in FS2, but the number offspring of S20 increased to 138, while reducing offspring in other sires,
- FS4 is an extension of FS2, with additional strongly inbred sire S26 (from a brother-sister mating) with 73 offspring.

All females in the pedigree except 2 granddames (**GD**) GD5 and GD12 were assumed to be unrelated with unknown ancestors. Differences in maternal structure follow the discriminating criteria of family structures. A stronger level of inbreeding, as simulated in FS4, is not expected to happen in real dairy cattle practice. We used it to compare the results of analyzing family structures with increasing inbreeding up to an unusually high level.

As the parents of founder individuals are unknown, their genotypes were sampled according to specified allele frequencies and their distribution was modeled to be in Hardy-Weinberg equilibrium (Schelling et al., 1998), assuming allelic frequencies in equal proportion. Specifically, diploid offspring genotypes composed by a maternal and a paternal gamete were assumed in Mendelian inheritance. Both great-grandires have equal preconditions in all family structures, in terms of number informative markers and their trait values. Thus, the differences in results of the various family structures can be attributed to the differences in inbreeding level. Recombination events were simulated on the basis of a binomial map function with fewer than 10 crossovers, with any odd number of crossovers denoting a recombination. QTL effect was simulated by random allocation of genotypes of founders. Trait values (phenotypes) were generated in 2 scenarios.

In scenario A, as usually assumed in dairy cattle breeding, individual trait observation y is based on a normally distributed QTL effect q , polygenic effect g (here, as a sum of 20 single gene effects) plus residual effect e , according to the model

$$y_i = q_i + \sum_{20} g_i + e_i.$$

Scenario B reflects an extreme situation, for example after strong single-trait selection, where the individual trait observation y is based on a normally distributed QTL effect q plus residual effect e , according to model $y_i = q_i + e_i$, where i denotes a single individual. Whereas scenario A is the commonly assumed model, scenario B was explicitly generated in order not to obscure the behavior of QTL estimates by the polygenic component under different inbreeding structures, as it was the major focus of this study. If not explicitly distinguished, the discussion of results is based on scenario A. The

Table 1. Simulation parameters used to create 60 replicates for each family structure (FS0, FS1, FS2, FS3, and FS4) in scenario A and B, and their symbols¹

Simulation parameters	Versions of simulation parameters in detail (marked by their code in parentheses)
Maternal structure in final offspring (general)	- 176 dams with only one offspring - 83 pure maternal half-sib groups, 2 ... Three offspring - 33 mixed maternal half- and full-sib groups, 2 ... Eight offspring
Sets of marker alleles	(2_A) 2 marker alleles, (4_A) 4 marker alleles, (6_A) 6 marker alleles
Marker position on the map (cM)	(M1) 0, 13.7, 32.8, 35.7, 40.5 , 42.5 , 43.5, 44.5, 45.5, 48.5, 53.2 (M2) 0, 10.9, 17.8, 25.9, 33.6 39.8 , 46.3 , 48.3, 49.3, 51.2, 54.4 (M3) 0, 13.7, 32.8, 35.5, 37.5, 39.3, 40.3 42.4 , 43.3, 44.3, 48.4 (M4) 0, 13.7, 32.8, 35.7, 37.7, 39.7 , 43.5 , 44.5, 45.5, 46.5, 49.4 Bold script marks the marker bracket harboring the QTL at 41.5 cM.
General information content	(IC_0) original data with all individuals genotyped, (IC_1) 20% marker genotypes missing, (IC_2) 20% dam phenotypes missing, (IC_3) 20% markers and 20% dam phenotypes missing, (IC_4) increased error variance

¹Family structures: FS0 = without inbreeding; FS1 = with an inbred sire from an aunt-nephew mating; FS2 = with an inbred sire originating from a half-sib mating; FS3 and FS4 = based on FS2 but containing an increased number of offspring of the inbred sire from 15% to 28% (FS3), and another extremely inbred sire with its final offspring (FS4). Scenarios: A = based on the full genetic model including polygenes, effect of one QTL and residual effect; B = based on a simple QTL model including residual effect.

QTL effect was assumed to be normally distributed with mean zero and QTL variance, contributing 15% of the total trait variance, based on additive and dominance effects of the QTL (a and d) and allelic frequencies (m and n) according to Falconer (1989),

$$\sigma_q^2 = m^2[2n(a - md)]^2 + 2mn[a(m - n) + d(1 - 2nm)]^2 + m^2[-n(a + md)]^2.$$

The polygenic component, contributing 25% of the total variance in scenario A, is assumed to be normally distributed with mean zero and additive polygenetic variance, of single genes (g) with small effects of alleles l and k and allelic frequencies (m and n) at each single locus,

$$\sigma_g^2 = [2mnl(1 + k(m - n))]^2.$$

A random deviate e was normally distributed with mean zero and variance σ_{res}^2 (i.e., 60 percent of the total variance in scenario A, and 85 percent in scenario B, respectively). The ratio expressed between the QTL variance and the non-QTL variance, hereafter called the QTL variance ratio, was 0.176. Phenotypes were simulated assuming equal accuracy for all individuals and no previous sire or dam selection.

This study was initiated before introduction of new dense SNP technology, which generates much larger dimensions of marker information than we use here. That is why the assumed budget for a pre-fine-mapping QTL study was limited to 11 markers within a 55-cM chromosomal segment, which was expected to harbor one QTL by a preceding genome scan. For each

family structure, 60 replicates were simulated each in scenario A and B, based on varying and equally distributed simulation parameters (Table 1). Trait values and marker genotypes were simulated in an identical manner for all family structures, using the approach PEDSIM (Schelling et al., 1998). A linear model analysis using the GLM procedure from the SAS package (version 9.1, SAS Institute Inc., Cary, NC) was applied to test effects of simulation parameters on QTL mapping results. Significance of these effects was accepted in cases of $P < 0.05$.

QTL Analysis

The QTL effect associated with each individual phenotype at a given genome position was assumed random, having the variance-covariance structure which is a function of the probabilities that 2 QTL alleles (either in the same or in 2 different individuals within the pedigree) at a particular location are IBD. The IBD probabilities between each pair of gametes were calculated using a recursive algorithm by following transmission of marker alleles from parents to offspring (Fernando and Grossman, 1989). If a QTL position overlaps the position of a fully informative marker, the IBD probability of 2 QTL alleles can only take values of zero (non-IBD) and one (IBD). For positions between markers, IBD probabilities take values between zero and one, depending on the distance between the putative QTL position and the linked (informative) markers. If no marker information is available, 2 QTL alleles have equal chance of being IBD and non-IBD. An inbred individual will have greater probability of IBD between paternal and maternal alleles. That is, IBD matrices for a pedigree including inbred individuals will contain

more values close to or equal to one than matrices with no inbred individuals or that ignore inbreeding.

Uncorrelated haplotypes between founders were assumed. In this study, a rapid deterministic algorithm by Pong-Wong et al. (2001) was adapted to obtain IBD probabilities. This algorithm only uses the closest informative (phase known) marker bracket. Markers that were not informative, missing, or where the linkage phases could not be determined, were skipped. To avoid loss of information, a method by Knott and Haley (1998) was used to determine IBD probabilities among sibs' gametes, whereas a method by Vukasinovic and Martinez (2002) was used to determine IBD probabilities among half-sibs. These additional methods were applied to offspring of base individuals where linkage phases could not be determined, and to all full- and half-sib families for which average number of informative markers in the family (across both maternal and paternal gametes) using the original recursive method was smaller than the average number of informative markers used by the (half)-sib method. IBD probabilities were calculated for each pair of gametes independently to obtain a matrix $\mathbf{G}$ of gametic IBD probabilities. Then, the following mixed model was applied:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{g} + \mathbf{H}_p\mathbf{q}_p + \mathbf{e},$$

where $\mathbf{y}$ is an $(n_1 \times 1)$ vector of phenotypes (trait observation), $\mathbf{X}$ is an $(n_1 \times s)$ design matrix, $\mathbf{Z}$ is an $(n_1 \times n_0)$ incidence matrix relating animals to their phenotypes, $\mathbf{H}_p$ is a $(n_1 \times 2n_0)$ incidence matrix relating animals to paternal and maternal QTL gametes at position p . Here, β is an $(s \times 1)$ vector of fixed effects, $\mathbf{g}$ is an $(n_0 \times 1)$ vector of random polygenic effects. $\mathbf{q}_p$ is a $(2n_0 \times 1)$ vector of the effect of a QTL at position p , and $\mathbf{e}$ is an $(n_1 \times 1)$ residual vector with expectation and covariance matrix $(0, \mathbf{E} \otimes \mathbf{I})$, where $\mathbf{E}$ was the unknown (co) variance matrix of the residual effects. $\mathbf{I}$ denotes the identity matrix. Here, n_1 refers to the number of animals with phenotypes and n_0 to the total number of animals in the pedigree. $\mathbf{X}$ is equal to one, because all phenotypes were assumed preadjusted for fixed effects, and thus $s = \text{one}$ and $\beta = \text{mean}$. The random polygenic effects $\mathbf{g}$, and QTL effects $\mathbf{q}_p$, are assumed to follow a normal distribution with mean zero and variances $\mathbf{A}\sigma_g^2$ and $\mathbf{G}_p\sigma_p^2$, respectively. Matrix $\mathbf{A}$ is the additive relationship matrix. Matrix $\mathbf{G}_p$ contains IBD probabilities at each chromosomal position p , obtained as described above.

The model was fitted for each single position p (in steps of 1 cM) on the chromosomal segment. Scenarios A and B were analyzed by means of the same model to mimic analyses of real dairy data, where the true model

underlying the trait of interest is unknown. The data was analyzed using a random model variance component approach. The REML procedure implemented in the ASReml software (Gilmour et al., 2002) was used to maximize the likelihood for computing the likelihood ratio $[\mathbf{LR} = -2(\ln L_{H_0} - \ln L_{H_A})]$, where $\ln L_{H_0}$ is the logarithm of the likelihood computed for the pure polygenic model with no QTL (null hypothesis) and $\ln L_{H_A}$ is the logarithm of the likelihood obtained from the model including a QTL (alternative hypothesis). The likelihood ratio was calculated at each position p , for estimating the most likely QTL position in each single analysis by means of the highest LR. In total, 600 analyses based on the conditions described were run.

RESULTS

Inbreeding Coefficients, IBD Probabilities, and Computational Aspects

By definition, mating an inbred individual to an unrelated mate yields noninbred offspring. Mild inbreeding is common in dairy cattle, whereas inbreeding $>12.5\%$ is much less common. Casual relationships between mates, unrecognized or unavoidable at time of mating, lead to most inbreeding in Holsteins and Jerseys. Inbreeding coefficients and within-locus IBD probabilities averaged over all parameter combinations and over all testing positions are given in Table 2 for all family structures including inbreeding. Inbreeding coefficients did not deviate from zero in FS0. The average overall IBD probability is relatively small in FS1 and highest in the extremely inbred family structure FS4.

Identity by descent probabilities affect the results of estimation. Average IBD probabilities increased gradually as level of inbreeding increased from FS0 to FS1 over FS2 to FS4. The shapes of profiles plotting IBD averages against map position were flat in FS0, but peaked sharply from 38 to 45 cM for FS1, FS2, and FS4 in parallel (plots of IBD parameters not shown). Variances of IBD probability increased gradually from FS0 to FS1, FS2, and FS4. Correlations between IBD probabilities at the true QTL position and all other positions along the map generally decreased as distance from the true position increased. But FS0 showed the greatest variation in correlation profiles and seemed to be most sensitive to simulation parameters. The change between IBD probabilities along the map is increasingly homogenous in family structures with increasing level of inbreeding, indicating less sensitivity to information content of markers.

In the majority of cases, computations with ASReml progressed smoothly, and convergence was reached in

Table 2. Inbreeding coefficients (F_x) and overall average within-locus identity by descent probabilities due to inbreeding¹ of individuals that affected family structures FS0, FS1, FS2, FS3, and FS4²

Item	Family structure	Overall	Inbred sires		
			S20	GS9	S26
F_x	FS1	0.0394	0.063	0.250	—
	FS2	0.0465	0.125	0.250	—
	FS4	0.0790	0.125	0.250	0.375
Within-locus ³ IBD probability	FS1	0.0006	0.018	0.487	—
	FS2	0.0011	0.487	0.484	—
	FS4	0.0017	0.465	0.484	0.443
Total offspring number ⁴		544	78 or 138	2 (89) ⁵	73

¹Additional contribution, since the within-locus IBD probability is equal to 1 for noninbred individuals (within-locus IBD probabilities refer to correlations between paternal and maternal allele at the same locus of an individual).

²Family structures: FS0 = without inbreeding; FS1 = with an inbred sire from an aunt-nephew mating; FS2 = with an inbred sire originating from a half-sib mating; FS3 and FS4 = based on FS2 but containing an increased number of offspring of the inbred sire from 15% to 28% (FS3), and another extremely inbred sire with its final offspring (FS4).

³In FS0, F_x and within-locus IBD probability deviations from 1 are all zero and in FS3 they are the same as in FS2 and therefore not listed.

⁴Total number final offspring is equal in all family structures, except for offspring number of S20 which is 138 in FS3 instead of 78 in others.

⁵Only 2 direct offspring from GS9.

fewer than 20 iterations. Computational time was 2 to 10 h per analysis. In a few cases, mainly in scenario B, convergence problems occurred within segments of 1 to 5 cM, primarily in the first part of the 55-cM chromosomal segment with no effect on mapping results. The total number of 850 simulated individuals was relatively small for a variance component analysis, but it meets practical conditions of a large number of reports on gene mapping.

Indicators of QTL Mapping Success

The following indicators were used to evaluate results of QTL mapping: (1) overall frequency of estimates within a 3-cM interval targeting the true QTL position at 41.5 cM, (2) bias of estimates as an absolute deviation of the estimated QTL position from the true QTL position (in centimorgans), hereafter called bias, (3) the QTL variance ratio, and (4) test statistic and test statistic profile (Table 3). Profiles with a score of 7 were most desirable, as sharper profiles with a clearly distinguishable peak usually indicates more precise mapping.

The significant effect of FS on bias ($P = 0.0285$ and 0.0159) and on QTL variance ratio ($P = 0.0039$ and 0.046) in scenario A and B, respectively, and also on the test statistic level ($P < 0.0001$), suggests an impact of using inbred sires on QTL parameter estimates. Marker allele number had a significant effect on QTL variance ratio ($P < 0.0001$) and LR ($P = 0.0029$), whereas remaining simulation parameters did not affect the results.

Effects of FS on Mapping Results

Polygenic variance means more genetic noise for the analysis than a simulation based on QTL and residual effects only and explains the different precision of QTL position estimates in scenario A and B. The average bias and the span of estimated QTL position was highest in FS0 in scenario A. Outliers occurred in all family structures, but FS0 yielded the largest number of such extremes (Figure 1). The 3-cM interval between 40 and 43 cM harboring the true QTL was reached by 16.7% of the estimates in FS0, by 21.7% in FS1, by 42.7% in FS2, by 23.3% in FS3, and by 51.7% in FS4 (scenario A). A high frequency of scores 2, 3, and 5, reached by FS3 (scenario A), denoted flat or multi-peaked profiles (Table 3) and very little difference in LR estimates of several positions. Thus, some adjacent positions could be equally likely. We always and strictly used the highest LR to find the position estimate even if differences were very small. A multi-peaked test statistic profile is difficult to interpret under experimental conditions, because the true position of one (or more) QTL is unknown. The FS3 results are an exception to the almost linear increase in frequency of correct QTL position estimates with increasing inbreeding (Figure 1, scenario A). The results from inbred family structures FS1, FS2, and FS4 reflect an increased precision in QTL position estimates with increased inbreeding. The distribution of the QTL position estimates showed that FS0 was significantly different from inbred structures FS2 and FS4 ($P < 0.001$ according to the chi-squared test). These

Table 3. Subjective scores of test statistic profiles and their relative frequency obtained from analyses of family structures FS0, FS1, FS2, FS3, and FS4 in scenarios A and B¹

Family structure and scenario												
Score	Description of test statistic profile	FS0		FS1		FS2		FS3		FS4		
		A	B	A	B	A	B	A	B	A	B	
1	Experiencing convergence problems	0.033	0.283	0	0.100	0.017	0.033	0	0.150	0	0.217	
2	Very flat profile along the whole chromosomal segment	0.450	0.017	0.117	0.017	0.267	0.033	0.183	0.017	0.100	0.067	
3	Flat profile around the highest LR peak	0.133	0.167	0.083	0.233	0.050	0.117	0.15	0.050	0.233	0.020	
4	More than four distinct peaks	0.083	0	0	0.034	0	0.033	0.033	0.067	0	0.007	
5	More than one peak >10 cM apart	0.150	0.200	0.350	0.183	0.117	0.100	0.200	0.050	0.300	0.050	
6	Peaked over the region ±10 cM from the position with the highest LR	0.050	0.266	0.117	0.183	0.217	0.450	0.167	0.217	0.217	0.260	
7	One single, sharp and well-defined peak ±5 cM	0.100	0.067	0.333	0.250	0.333	0.233	0.267	0.450	0.150	0.383	

¹Family structures: FS0 = without inbreeding; FS1 = with an inbred sire from an aunt-nephew mating; FS2 = with an inbred sire originating from a half-sib mating; FS3 and FS4 = based on FS2 but containing an increased number of offspring of the inbred sire from 15% to 28% (FS3), and another extremely inbred sire with its final offspring (FS4). Scenarios: A = based on the full genetic model including polygenes, effect of one QTL and residual effect; B = based on a simple QTL model including residual effect.

results suggest that the main reason for positive effects of including inbred sires on estimated QTL position was a reduction in genetic noise.

We observed an increase in QTL variance ratio as the number of marker alleles increased in all family structures. These results are not shown in detail, however, as the question of number of marker alleles has been rendered moot by SNP technology that is currently being introduced. Variance component estimation in scenario A was best for FS0, but this advantage was not accompanied by sufficiently accurate QTL estimates. FS0 yielded the smallest number of correctly estimated QTL positions, largest bias, and worst test statistic profiles. Correct position estimates are judged by the within-analysis LR, for which the average was greatest in FS3 (21.83) and lowest for FS0 (6.75). The highest difference in the average test statistic level among the family structures based on the model without assuming any QTL ($\ln L_{H0}$) was less important (−419.68 in FS2 vs. −410.55 in FS3).

A preliminary chromosome wide permutation test (based on an *F*-test, according to Doerge and Churchill, 1996) yielded minor differences in significance thresholds at $P < 0.01$ for scenario A ($F_{\text{threshold}} = 2.68$ in FS0 to $F_{\text{threshold}} = 2.91$ in FS1, with no specific range among the inbred structures). But inbred structures exceeded this level ($F_{\text{realized}} = 3.85$ to 4.30), whereas FS0 was below its threshold ($F_{\text{realized}} = 2.13$).

Thresholds used to maintain type I error depend upon the parental relationship (Leutenegger et al., 2002). Correlation structures within individuals of inbred and outbred populations are different (Abney et al., 2000). In scenario A, there was no significant difference in FS1, FS2, and FS4 with respect to $\ln L_{H0}$. More precise estimation of polygenic variance in FS0 in comparison to inbred family structures (Table 4) could be an explanation for lower $\ln L_{HA}$, based on assuming a QTL. The estimated polygenic variance differed among inbred family structures, an observation that could be caused by inbreeding loops in interaction with the polygenic variance present within family. Our study did not formally address the power of QTL detection.

Family structure 0 displayed the smallest number of score 7 in LR profiles, and inbred pedigree structures yielded significantly higher frequencies of desirable LR profiles (scores 6 and 7) than FS0 in both scenarios (Table 3). However, the ideal shape of the LR profile (score 7) was in many cases accompanied by overestimated QTL variance ratio estimates.

DISCUSSION

The idea for this study predates dense marker maps. Dense SNP marker array technology had been discussed,

Figure 1. Estimated QTL positions summarized as relative frequency in intervals around the targeted QTL region of 40 to 43 cM for family structures FS0, FS1, FS2, FS3, and FS4 [FS0 = without inbreeding; FS1 = with an inbred sire from an aunt-nephew mating; FS2 = with an inbred sire originating from a half-sib mating; FS3 and FS4 = based on FS2 but containing an increased number of offspring of the inbred sire from 15% to 28% (FS3), and another extremely inbred sire with its final offspring (FS4)] in simulated scenarios A (based on the full genetic model including polygenes, effect of one QTL and residual effect), and B (based on a simple QTL model including residual effect).

but was still far from application in dairy cattle. SNP markers will definitely enrich QTL mapping possibilities to a great degree. However, correct pedigrees and well-suited family designs will be necessary to take optimum advantage of data from dense SNP chips. It remains important to design QTL mapping studies in ways that obtain maximum information from the investment in genotyping. Considering inbreeding can be an advantage both in genomic selection and in basic research.

The existence of inbred animals in dairy populations appears to offer advantages for QTL mapping. Offspring of single inbred sires seem to stabilize analysis to estimate a QTL position. Offspring groups descending from an inbred sire are expected to be genetically more similar than sibs of noninbred parents. Therefore, we assume that when inbred animals are mated to “average” animals, more inherent “homogeneous” information from a pedigree is present for tracing inheritance

of marker and QTL alleles, provided that complete and accurate pedigrees are available. Moreover, occurrence of inbreeding (e.g., offspring from an inbred sire) in a highly selected population means an accumulation of favorable alleles, whereas unwanted alleles tend to disappear. Mating such individuals to partners from different families increases information for the mapping process—a situation that may become even more advantageous as more marker information is available. Historically, considerations of inbreeding seems computationally demanding, because the whole mapping population has to be considered at once without breaking the pedigree into individual families. However, computational impediments in animal breeding have tended to disappear as computing power increased. Utilization of inbred sires in a pedigree could enable more precise QTL mapping with fewer animals to be genotyped.

In this simulation, we did not focus on a specific genetic trait-related effect of inbreeding. In future studies,

Table 4. Quantitative trait loci (QTL) position bias, QTL variance ratio, QTL variance ratio distribution (minimum and maximum) and estimated average polygenic variance in family structures FS0, FS1, FS2, FS3, and FS4¹

Family Structure	Scenario A					Scenario B				
	Bias (cM) average ²	QTL variance ratio ²	QTL variance ratio distribution		Polygenic variance ²	Bias (cM) average ²	QTL variance ratio ²	QTL variance ratio distribution		
			min	max				min	max	
FS0	15.72	0.167	0.029	0.383	0.256	2.37	0.302	0.079	0.998	
FS1	4.53	0.225	0.106	0.452	0.203	1.77	0.270	0.049	0.999	
FS2	4.54	0.248	0.099	0.990	0.225	1.55	0.246	0.071	0.990	
FS3	6.02	0.275	0.102	0.690	0.222	1.55	0.568	0.050	0.690	
FS4	3.53	0.260	0.097	0.999	0.236	1.63	0.354	0.081	0.999	

¹Family structures: FS0 = without inbreeding; FS1 = with an inbred sire from an aunt-nephew mating; FS2 = with an inbred sire originating from a half-sib mating; FS3 and FS4 = based on FS2 but containing an increased number of offspring of the inbred sire from 15% to 28% (FS3), and another extremely inbred sire with its final offspring (FS4). Scenarios: A = based on the full genetic model including polygenes, effect of one QTL and residual effect; B = based on a simple QTL model including residual effect.

²Averages of 60 replicates.

this could be done by adapting infinitesimal or finite-locus models.

To allow for comparison of QTL estimates within various levels of inbreeding, the phenotypic variance was the same in all of our family structures although a diminishing genetic variance with increasing inbreeding is expected (Kristensen and Sorensen, 2005). Furthermore, those studies on model animals suggest a higher environmental sensitivity in inbred individuals. As we experienced, increased residual variance and simultaneously reduced genetic variance affect the process of computation and estimation positively. Related considerations demonstrate the high complexity of this challenging subject.

CONCLUSIONS

The dairy breeding literature documents that inbreeding can be associated with negative effects on both production and fitness of animals. It reduces genetic variability that normally is a basis of breeding success in populations. However, mating of closely related individuals, usually through artificial insemination, happens either through ignorance of existing relationships, remoteness of several common ancestors in pedigrees of sires and dams, or, more rarely, intentionally. In practice, a net of ancestry and relationship exists within large multi-generation pedigrees involving both sexes. Dairy field databases on relatively complete pedigree information will include complicated family structures with closer relationships of mates and a higher percentage of inbred individuals than were considered in this study. The majority of extreme outliers in QTL position estimates and highest bias of QTL position estimates was found in the noninbred family structure. Quantitative trait loci variance ratio estimates were closer to the simulated value with the use of inbred family structures in scenario B only, where interaction of polygenic variance with inbreeding loops was missing.

The primary benefit of using inbred sires for QTL mapping studies is that it reduces genetic noise and positively affects inference of parental origin of the offspring haplotypes. Quality of genetic markers increases with the use of SNP technology that is currently the method of choice. A well designed animal basis for gene mapping is still an advantage in optimizing returns on research investment.

ACKNOWLEDGMENTS

The authors are grateful to unknown reviewers for constructive and critical comments that helped us very much to focus and to improve the manuscript.

REFERENCES

- Abney, M., M. S. McPeck, and C. Ober. 2000. Estimation of variance components of quantitative traits in inbred populations. *Am. J. Hum. Genet.* 66:629–650.
- Doerge, R. W., and G. A. Churchill. 1996. Permutation tests for multiple loci affecting a quantitative character. *Genetics* 142:285–294.
- Falconer, D. S. 1989. *Introduction to Quantitative Genetics*. 3rd ed. Longman Sci. Tech., Harlow, UK.
- Fernando, L. R., C. Stricker, and R. C. Elston. 1993. An efficient algorithm to compute the posterior genotypic distribution for every member of a pedigree without loops. *Theor. Appl. Genet.* 87:89–93.
- Fernando, R. L., and M. Grossman. 1989. Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.* 21:467–477.
- Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2002. *ASReml User Guide Release 1.0*. VSN International Ltd., Hemel Hempstead, UK.
- Hansen, L. B. 2006. Monitoring the worldwide genetic supply for dairy cattle with emphasis on managing crossbreeding and inbreeding. In *Proc. 8th World Congr. Genetics Applied to Livestock Production*, Belo Horizonte, Brazil, CD-Rom communication no. 01–01, Instituto Prociencia.
- Hernández-Sánchez, J., C. S. Haley, and J. A. Woolliams. 2006. Prediction of IBD based on population history for fine gene mapping. *Genet. Sel. Evol.* 38:231–252.
- Knott, S. A., and C. S. Haley. 1998. Simple multiple-marker sib-pair analysis for mapping quantitative trait. *Heredity* 81:48–54.
- Kristensen, T. N., and A. C. Sorensen. 2005. Inbreeding—Lessons from animal breeding, evolutionary biology and conservation genetics. *Anim. Sci.* 80:121–133.
- Leutenegger, A.-L., E. Génin, E. A. Thompson, and F. Clerget-Darpoux. 2002. Impact of parental relationships in maximum lod score affected sib-pair method. *Genet. Epidemiol.* 23:413–425.
- Meuwissen, T. H., and M. E. Goddard. 2000. Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. *Genetics* 155:421–430.
- Pong-Wong, R., A. W. George, J. A. Woolliams, and C. S. Haley. 2001. A simple and rapid method for calculating identity-by-descent matrices using multiple markers. *Genet. Sel. Evol.* 33:453–471.
- Schelling, M., C. Stricker, R. L. Fernando, and N. Künzi. 1998. PEDSIM—A simulation program for pedigree data. Pages 475–476 in *Proc. 6th World Congr. Genetics Applied to Livestock Production*, Vol. 27, Armidale, Australia.
- Thompson, J. R., R. W. Everett, and N. L. Hammerschmidt. 2000. Effects of inbreeding on production and survival in Holsteins. *J. Dairy Sci.* 83:1856–1864.
- Vukasinovic, N., and M. L. Martinez. 2002. Mapping quantitative trait loci in general pedigrees using the random model approach. In *Proc. Joint Statistical Meetings*, CD-ROM, American Statistical Association, New York, NY.
- Weller, J. I., Y. Kashi, and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* 73:2525–2537.