USING A COMPLETE MICROSATELLITE MAP AND THE GRAND-DAUGHTER DESIGN TO LOCATE POLYGENES CONTROLLING MILK PRODUCTION.

Michel Georges[&], Dahlia Nielsen, Margaret Mackinnon, Anuradha Mishra, Ron Okimoto, Alan T. Pasquino[#], Leslie S. Sargeant, Anita Sorensen, Michael R. Steele, Xuyun Zhao, James E. Womack[$] & Ina Hoeschele[#].

Genmark Inc., 421 Wakara Way, Salt Lake City, UT 84108. [#]Department of Dairy Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0315. [$]Department of Pathobiology, College of Veterinary Medicine, Texas A&M University, College Station.

SUMMARY

The objective of this study was to identify quantitative trait loci underlying the genetic variation of milk production in an elite dairy cattle population. A total of 1518 progeny-tested sires, belonging to 14 paternal half-sib families, were genotyped with 159 autosomal microsatellites composing a marker map bracketing 1645 centimorgan or approximately two thirds of the bovine genome. Using multilocus linkage analysis, five chromosomes gave very strong evidence (lodscore $\geq$ 3) for the presence of a quantitative trait locus: chromosomes 1, 6, 9, 10 and 20. These findings demonstrate that loci with considerable effects on milk production are still segregating in highly selected populations, and pave the way towards marker assisted selection in dairy cattle breeding.

INTRODUCTION

With the advent of AI, the impact of superior sires on the genetic progress of the herd is potentially enormous. Widespread use of a given sire, however, is only justified when its breeding value is estimated with sufficient reliability. Until now this has required "progeny-testing": young sires, resulting from planned matings of sires and dams with highest breeding value estimates, are tested based on the milking performances of 50-100 of their daughters. Progeny testing is a very time-consuming and expensive procedure. Approximately 5-6 years elapse between the time of selection of parents of a candidate sire and the estimation of its breeding value including its daughter records. The costs of progeny testing are estimated at $45,000 per sire, and only 10% of the tested sires will be selected after progeny test for large scale use in breeding programs. Although this procedure has proven very effective in improving the genetic merit of the herd, its tedious character has spurred research to develop faster and cheaper methods to predict the genetic value of an animal. One such possibility consists of identifying the "Quantitative Trait Loci (QTL)" contributing to the genetic variance of the production traits of interest in the relevant populations. This information could then be used to select animals based on their genotype at the QTL in a procedure called "Marker Assisted Selection" (Soller qnd Beckman, 1982; Smith and Simson, 1986).

In this paper we describe the use of a bovine genetic map composed of 159 microsatellite markers, to locate QTL with large effects on milk production that segregate in an elite Holstein dairy cattle population selected intensely for these traits for several generations.

MATERIALS AND METHODS

The "grand-daughter design" (Weller et al., 1990):

We identified 14 paternal half-brother US. Holstein pedigrees, with between 33 and 208 progeny-tested sons per founder sire (mean: 108), for a total of 1518 sons. None of the dams were available for analysis.

Sire:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
Sons:	200	45	103	160	57	111	33	44	115	208	50	129	173	90	1518

Microsatellite genotyping:

All genetic markers used in this study were previously described microsatellite markers (Georges and Massey, 1991). Microsatellite genotyping was performed using standard procedures.

Map construction:

All linkage analyses were performed with the ANIMAP programs (D. Nielsen, unpublished). These programs were designed to perform linkage studies in half-sib pedigrees. They can be used to 1. generate lodscore tables between pairs of markers with codominant alleles; 2. perform multipoint linkage analysis with up to 16 markers (maximum likelihood recombination rates between adjacent markers are determined for all or a subset of marker orders); 3. generate lodscores between a QTL whose position can be varied with respect to a set of up to 16 markers whose relative positions are held fixed.

QTL mapping:

Five milk production traits were analyzed in this study: milk yield, fat yield, protein yield, fat percentage and protein percentage. The quantitative measurements used in the linkage analysis were sires' Daughter Yield Deviations (DYDs) obtained from the sire summary data base of January 1993 of the U.S. Department of Agriculture.

We used a multilocus maximum likelihood method, related to interval mapping, for the identification of QTL. Our method differs from conventional interval mapping as described by Lander and Botstein (1986) in that information from all markers composing the linkage group is used in computing the likelihood at a putative QTL location, instead of information from flanking markers only, and in the fact that we analyzed phenotypic averages in an unbalanced design instead of individual phenotypes, requiring us to account for variance heterogeneity of the phenotypes. The analysis was performed *within* half-sib ship. Evidence for a QTL at the corresponding map position was expressed as a lodscore, i.e. the $\log_{10}$ of the likelihood ratio. Following Lander and Botstein (1986) and knowing that we explored approximately 16 Morgan with brackets of approximately 15 cM, we chose a very stringent lodscore threshold of 3 to reduce the chance of a false positive occurring anywhere in the genome to less than 5%.

RESULTS

Construction of a primary DNA marker map:

The 14 founder sires were genotyped for 181 previously described bovine microsatellite markers. Informative families, i.e. sib ships for which the founder sire was heterozygous, were genotyped with the respective markers. The 104,523 resulting genotypes were used to construct a microsatellite map spanning a total of 1,645 bracketted

autosomal centimorgans (mean bracket size = 14.8 cM) or 2/3 of the bovine genome.

Mapping QTL controlling milk production:

We identified five chromosomes giving lodscores ≥ 3 within bracketed segments: chromosomes 1 (U10), 6 (U15), 9 (U2), 10 (U5) and 20 (U20). (Table 1)

Each of the mapped QTL affects the different milk production traits in a distinct manner. The QTL on chromosome 9 increased the amount of milk produced without significantly altering its fat and protein composition: fat and protein yield were increased concomitantly. The two QTL on chromosomes 6 and 20, on the other hand, appeared to increase milk yield but not fat or protein yield: fat and protein percentage were both reduced. The QTL on chromosomes 1 and 10 seemed to have differential effects on milk composition: while the higher milk yield was accompanied by a stronger increase in fat yield than in protein yield for the QTL on 10, the opposite was observed for the QTL on 1.

For the significant linkages (lodscores ≥ 3), the maximum likelihood estimates of the QTL effects ranged from .67 to 1.45 standard deviations σ_{DYD}, explaining from 11% to 52% of the total variance of DYD within a half-sib family, σ_{DYD}^2. It should be realized, however, that these estimates are determined by the power - or lack of - characterizing the experimental design, as much as by the actual effect of the mapped QTL: given the size of our pedigrees, the effects need to be of a given magnitude to yield the imposed threshold lodscores. In consequence, these QTL effects are very likely overestimated. This allso explains why some of the mapped QTL explain more than 100% of the expected mendelian segregation variation of the sire. As σ_{DYD}^2 approximately equals $[.1875\sigma_A^2 + (.5\sigma_A^2 + \sigma_E^2)/n]$, and assuming a trait with 30% heritability and a progeny test based on 100 daughters (=n), 87% of σ_{DYD}^2 is genetic in nature. One third of this genetic component, or 29% of σ_{DYD}^2 correspond to the sire's mendelian sampling variance. Consequently, the identified QTL would explain between 38% and 179% of the expected mendelian segregation variation of the sire.

DISCUSSION

The identification of QTL segregating in elite dairy cattle populations, is the first step towards the application of marker assisted selection for milk production. The selection of young dairy sires, which presently relies on the expensive and time-consuming progeny-testing procedure, offers a unique opportunity for the utilization of genetic markers in livestock production. Young dairy bulls result from planned matings of "bull-sires" and "bull-dams" with the highest BVs or PTAs (Predicted Transmitting Ability). The predicted BV of the offspring correspond to the average of the parental BVs. The actual BV of the offspring will, however, deviate from the predicted, because 1. the estimates of the parental BV is not fully accurate, and mainly because 2. mendelian sampling effects, or the fact that different offspring receive a different sample of genes from their parents. Progeny-testing has been implemented for that very reason.

Several years ago, the use of genetic markers in the selection of young dairy bulls was proposed (Soller and Beckman, 1982). The benefit of markers has often been analyzed in terms of improved accuracy of selection. The gain to be made following this approach is generally accepted to be marginal (Smith and Simpson, 1986; Meeuwissen and Van Arendonk, 1992). Furthermore, to be effective this approach requires a very detailed understanding of the identified QTL in terms of number of segregating alleles and their respective effects, which may be very difficult to achieve in the near future.

Table 1: Identified QTL effects.

For each synteny group - pedigree combination yielding a lodscore ≥ 3 for at least one of the five traits studied, we report for each trait the maximum lodscore (Lods), the Maximum Likelihood estimate of .5α, with α corresponding to the average effect of a QTL allele substitution, and the standard deviation for the respective DYDs in the corresponding half-sib family (σ_{DYD}).

DYD	Lods	$\alpha/2$	σ_{DYD}
1. U2(9)-pedigree 3: MY(kg)	2.58	+230	286
FY(kg)	4.00	+ 0	10.9
PY(kg)	3.38	+ 5.9	6.4
F%	0.00		0.4
P%	0.00		0.18
2. U5(10)-pedigree 3:MY(kg)	2.21	+ 336	286
FY(kg)	3.67	+12.0	10.9
PY(kg)	1.17	+ 5.2	6.4
F%	0.00		0.4
P%	2.27	- 0.21	0.18
3. U10(1)-pedigree 1:MY(kg)	3.15	+ 266	299
FY(kg)	0.34	+ 5	10.4
PY(kg)	3.19	+ 8.2	9.1
F%	0.90	- 0.32	0.4
P%	2.31	- 0.32	0.22
4. U15(6)-pedigree 9:MY(kg)	3.42	- 244	291
FY(kg)	0.59	+ 3.2	10
PY(kg)	0.10	+ 1.4	7.7
F%	4.66	+ 0.73	0.5
P%	3.60	+ 0.5	0.27
5. U20(20)-pedigree 10:MY(kg)	0.41	- 78	294
FY(kg)	0.83	+ 3.6	10.9
PY(kg)	0.00		8.2
F%	2.62	+ 0.36	0.59
P%	3.96	+ 0.18	0.27
6. U20(20)-pedigree 3:MY(kg)	1.61	- 171	286
FY(kg)	0.00		10.9
PY(kg)	0.84	+ 4.1	6.4
F%	2.35	+ 0.36	0.4
P%	3.20	+ 0.18	0.18

In the short term the major advantage of markers will likely result from predicting parts of the mendelian sampling effects at a stage where current selection schemes do not provide any information to differentiate among full-sibs. As Multiple Ovulation and Embryo Transfer (MOET) enables to produce larger numbers of full-sibs in dairy cattle, markers will allow to preselect among full-sib brothers prior to progeny-testing, and to test only those more likely to have BVs superior to the parental mean. To implement such a scheme, it must be determined for which of the identified QTL "bull-sire" and "bull-dam" are heterozygous. Indeed, these are the QTL contributing to differentiation among siblings due to mendelian sampling in a given mating. In the short term, this analysis seems difficult to achieve for bull-dams. However, segregation analysis (involving markers linked to identified QTL) using the progeny-test daughters may allow to determine heterozygosity in the bull-sire. The feasibility of such a scheme has been examined by Hoeschele and Romano (1994) and is presently under further study (Mackinnon et al., in preparation). Kashi et al. (1990) proposed to use information from paternal and maternal grand-sire to select amongst QTL alleles from bull-sire and bull-dam respectively. While in the short term, this might be the only feasible alternative to select amongst the QTL alleles from the bull-dam, this approach has the disadvantage that selection potential is wasted on QTL for which the bull-dam is homozygous and which consequently do not contribute to mendelian sampling.

In the long term, a better understanding of QTL parameters may lead to more complex strategies combining phenotypic and QTL data into a single analysis (Hoeschele, 1993).

Our demonstration that QTL can be mapped in highly selected segregating dairy populations should strongly encourage efforts to develop selection schemes incorporating marker information.

REFERENCES.

Georges, M.; Massey, J. (1992). Polymorphic DNA markers in Bovidae (World Intellectual Property Org., Geneva), WO Publ. No. 92/13102.

Hoeschele, I. (1993). J. Dairy Science 76: 169

Kashi, Y.; Hallerman, E.; Soller, M. (1990).. Anim. Prod. 51:63-74.

Lander, E.S.; Botstein, D. (1989). Genetics 121:185-199.

Meeuwissen, T.H.E.; Van Arendonk, J.A.M. (1992). J. Dairy Sci.75:1651-1659.

Smith, C.; Simpson, S.P. (1986).. J. Anim. Breed. Genet. 103:205-217.

Soller, M.; Beckman, J.S. (1982). Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production 6:396-404.

VanRaden, P.M.; Wiggans, G.R. (1991). J. Dairy Sci. 74:2737-2746.

Weller, J.L.; Kashi, Y.; Soller, M. (1990). J. Dairy Sci. 73:2525-2537.