

# Effects of Milk Protein Genotypes on the Variation for Milk Production Traits of Holstein and Jersey Cows in California

MATTI OJALA,\* THOMAS R. FAMULA,† and JUAN F. MEDRANO†

\*Department of Animal Science, University of Helsinki, PO Box 28, FIN-00014 Helsinki University, Finland

†Department of Animal Science, University of California, Davis 95616

## ABSTRACT

The objectives of this study were to form appropriate composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes in order to assess which genotypes were favorably associated with first lactation milk production, fat and protein percentages, and fat and protein production for data of 916 Holstein and 116 Jersey cows. Multiple-trait animal models were used with assumed fixed effects for herd, year and season of calving, age at calving, days open, composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes, and  $\beta$ -LG genotypes. The differences between the  $\beta$ -LG genotypes for production traits were not statistically significant for either breed. The proportion of phenotypic variance that was due to the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes was 5% for milk production, 4% for protein production, and 3% for fat percentage. The  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotype ABA<sub>1</sub>A<sub>2</sub>BB was superior to the comparable AAA<sub>1</sub>A<sub>2</sub>BB and ABA<sub>1</sub>A<sub>1</sub>BB genotypes by 252 and 338 kg for first lactation milk yield, respectively, and 8.7 and 11.5 kg for protein yield, respectively. Thus, neither the  $\beta$ -CN A<sub>2</sub> allele nor the  $\kappa$ -CN B allele alone had a positive effect on milk and protein production, but the joint effect was strongly positive. These results may be explained by epistatic effects between the  $\kappa$ -CN and  $\beta$ -CN loci and possibly by closely linked quantitative trait loci with favorable alleles on the  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN BA<sub>2</sub>B haplotype.

(**Key words:** genetic variation, milk production traits, milk protein genotype, multivariate animal model)

**Abbreviation key:** QTL = quantitative trait loci.

## INTRODUCTION

The milk proteins that have a significant effect on milk production traits vary widely in terms of the genotype, trait, and magnitude of the effect in differ-

ent dairy cattle populations [e.g.,(14)]. Variations may be caused, for example, by the model assumptions, by the statistical procedure used for analyses, and by the size and structure of the available data files. In order to detect the contribution of the milk protein alleles to the variation of milk production traits, the variability has to be properly accounted for in the model. In addition to milk protein genotype, fixed effects usually included in the model are herd, calving year and season, and age at calving.

The milk protein genotypes have been included in the model in different ways. If the individual genotypes for the  $\kappa$ -,  $\beta$ -, and  $\alpha_{s1}$ -CN loci and  $\beta$ -LG have been included individually in the model as fixed effects [e.g., (16)], then the observed effect for a casein genotype is not necessarily correct because of the influence from other closely linked casein loci. This deficiency may be overcome, in part, if the individual milk protein genotypes are included simultaneously as separate fixed effects in the model, as in the majority of studies reviewed. In this case, the observed effect of a casein genotype may be due to the additive effects of the two alleles at a single locus and the possible dominance effect between them. It is also possible that loci exist that are closely linked with the milk protein locus and that have similar allelic effects, in which case, all of these loci, including the milk protein locus, can be considered to be quantitative trait loci (QTL).

Close linkage between the caseins was observed in early studies (9), and physical mapping techniques later confirmed that the casein loci reside on chromosome 6 within a region of less than 200 kb in the order of  $\alpha_{s1}$ -CN,  $\beta$ -CN, and  $\kappa$ -CN (6, 21). Because of the close linkage, the allelic effects of these loci are confounded in statistical analysis even when all of these loci are included simultaneously in the model. This problem can be overcome by treating all casein loci as composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and using the smallest subcells in the analysis. Consequently, the observed effect of a composite casein genotype would include all of the previously mentioned genetic effects

Received July 1, 1996.

Accepted December 18, 1996.

within each locus in addition to the possible interaction effects between the single loci (i.e., epistatic effects). The composite casein genotypes (8) or interaction terms between the casein loci (1, 7, 15, 22) were alternatively included in the model, but the contribution of interaction effects for the variation in milk production traits was small or not statistically significant in those studies. To allow for meaningful statistical analyses, the rare composite genotypes must be grouped or deleted.

Most results are from studies analyzed with the least squares procedure assuming a fixed model and not accounting for polygenic effects through the relationships among the animals (e.g., 1, 2, 12, 13, 16, 17, 18, 22). In some cases, the relationships among the sires of the animals with records have been considered (7, 8). The mixed model methodology (10), which includes both fixed and random effects in the model, is the method of choice when both single locus and polygenic effects are involved (11).

A mixed linear model, specifically a sire model, that accounted for paternal half-sib relationships (15) was initially used. This model was followed by the use of an individual animal model including all parents (5) or including all parents and grandparents (4). Because of the pedigree information available, only in the last study was it possible to derive the full benefit of the individual animal model. Milk production traits are known to be correlated; thus, accounting for the correlations among the traits by assuming a multivariate mixed model is expected to improve the efficiency of analyses over those based on univariate models.

The objective of this study was to find which composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes or individual alleles are favorably associated with the milk production traits of California Holstein and Jersey cows assuming univariate and multitrait animal models. A parallel objective was to estimate the contribution of milk protein genotypes on the additive genetic variation and the heritability for milk production traits for the Holstein data.

## MATERIALS AND METHODS

Records on 916 first lactation Holstein and 116 Jersey cows from registered dairy herds in California were used in this study. These data were extracted from the data collected and described by Van Eenennaam and Medrano (22). For the purposes of this study, the parental pedigrees for cows with records (5) were supplemented with up to at least grandparents or great grandparents, if known. The fixed ef-

fects in the model and the classes of fixed effects also differed from the models previously applied on these data (5, 22).

The following 305-d milk production traits of cows in first lactation were used in the statistical analyses: milk production, fat percentage, protein percentage, fat production, and protein production. For preliminary analyses, a model of fixed effects corresponding to Model [1] was assumed, and the data were analyzed separately within the two breeds using the general linear models procedure of SAS (19).

To analyze the data within the Holstein and Jersey breeds, univariate, bivariate, and trivariate models were assumed under the following linear model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Qg} + \mathbf{Za} + \mathbf{e} \quad [1]$$

where

- $\mathbf{y}$  = vector of observations for the milk production traits (e.g., 305-d milk production, fat and protein percentages, and fat and protein production) on the genotyped cows;
- $\mathbf{X}$  = known incidence matrix relating observations in  $\mathbf{y}$  to the classes of fixed effects in  $\mathbf{b}$ ;
- $\mathbf{b}$  = vector of unknown contributions of fixed effects for herd, year and season of calving, age at calving, and days open;
- $\mathbf{Q}$  = known incidence matrix relating observations in  $\mathbf{y}$  to the classes of fixed effects in  $\mathbf{g}$ ;
- $\mathbf{g}$  = vector of unknown contributions of fixed effects for composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and  $\beta$ -LG genotypes;
- $\mathbf{Z}$  = known incidence matrix relating observations in  $\mathbf{y}$  to the classes in the random vector  $\mathbf{a}$ ;
- $\mathbf{a}$  = vector of unknown contributions of additive genetic effects; and
- $\mathbf{e}$  = vector of unknown residual effects associated with the observations in  $\mathbf{y}$ .

Both  $\mathbf{a}$  and  $\mathbf{e}$  are assumed to have null means. The following variance-covariance structure between the traits 1 and 2 for a bivariate model was also assumed:

$$\text{Var} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_{a_1}^2 & \mathbf{A}\sigma_{a_1a_2} & 0 & 0 \\ \mathbf{A}\sigma_{a_1a_2} & \mathbf{A}\sigma_{a_2}^2 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{e_1}^2 & \mathbf{I}\sigma_{e_1e_2} \\ 0 & 0 & \mathbf{I}\sigma_{e_1e_2} & \mathbf{I}\sigma_{e_2}^2 \end{bmatrix}$$

where

$\mathbf{A}$  = numerator relationship matrix among animals represented in vector  $\mathbf{a}$ ,

$\sigma_a^2$  = additive genetic variance,

$\mathbf{I}$  = identity matrix,

$\sigma_e^2$  = the residual variance, and

$\sigma_{a_1a_2}$  and  $\sigma_{e_1e_2}$  = the additive genetic and residual covariances between the traits 1 and 2.

The variance-covariance matrix is analogously reduced by one variable for univariate models and extended by one variable for trivariate models. After the known ancestors for the cows with records are added, the orders of the  $\mathbf{a}$  vector and the  $\mathbf{A}$  matrix were 3088 for Holsteins, of which 916 cows had recorded milk protein genotypes and milk production traits; for Jerseys, the corresponding numbers were 384 and 116, respectively.

For fixed effects of  $\mathbf{b}$ , the data were collected from 10 Holstein and 5 Jersey herds in California. The year and season of calving was classified into seven groups for the Holstein data file; the classification was in 2-mo periods from September 1988 through June 1989 and included two additional classes at both ends. For Jerseys, the year and season of calving was classified into five groups; the majority of observations were from the first half of 1989. For Holsteins, the age at calving was classified into eight groups: the classification interval was 1 mo from the age of 25 to 31 mo; two additional extreme classes were included. For Jerseys, the age at calving was classified into four groups; the innermost classes were from 24 to 26 mo at calving. For Holsteins, the effect of days open was classified into 10 groups with unequal class intervals, and a 15-d period was used from 60 to 120 d after calving; the same classification principle was applied to Jerseys in a total of six groups.

In vector  $\mathbf{g}$ , the classes for casein loci were derived from the individual composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes (Table 1). To allow for comparison of the effects of the  $\beta$ -CN genotypes within the  $\kappa$ -CN genotypes, the order of nesting resulted in using the reversed alphabetical order for the casein loci. The classes for statistical analyses included the individual composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes; the largest number of observations and groups of rare  $\alpha_{s1}$ -CN and  $\beta$ -CN genotypes within the  $\kappa$ -CN genotypes resulted in a total of nine classes for Holsteins and eight for Jerseys (Table 2). The  $\beta$ -LG genotype classes were AA, AB, and BB for both breeds.

Statistical analyses were performed initially for each of the five traits separately for Holstein and

Jersey data. Milk production, fat percentage, and protein percentage were also analyzed assuming a trivariate model. A bivariate model was deemed feasible for fat and protein production because those composite traits are formed from two of the three original traits: milk production, fat percentage, and protein percentage.

Estimates of variance and covariance components for the milk production traits have REML properties (20) and were estimated using the MTDFREML program (3). To assess the amount of additive genetic variation that was contributed by the milk protein genotypes on milk production traits, Model [1], which included milk protein genotypes, and Model [1a], which excluded milk protein genotypes, were used on the Holstein data file.

Statistical significance for contrasts, a difference between two classes of a factor in this study, was evaluated with standard errors for univariate fixed and mixed models and for multivariate mixed model analyses. The effects of individual milk protein genotypes for milk production traits were expressed as the difference from the genotype with the largest number of observations. This procedure resulted in the smallest possible standard errors for the contrasts and also provided a direct comparison with the most common milk protein genotype in the data file. To supplement comparison of the effects of composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes for milk production traits within the two breeds, analyses were also performed within the largest Holstein herds with the highest and lowest milk yields.

## RESULTS AND DISCUSSION

### Groups for Composite $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN Genotypes

Use of the same subclasses for both breeds very clearly showed the differences in composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes between Holstein and Jersey cows (Table 1). Most Holstein cows at the  $\kappa$ -CN locus were of the AA genotype; a small minority was of the BB genotype. The opposite was true for Jersey cows. For the Holstein data, the majority of observations at the  $\beta$ -CN locus was of the  $A_1A_2$  genotype and the two corresponding homozygotes. For the  $\alpha_{s1}$ -CN locus, most cows were of the BB genotype.

To allow for meaningful statistical analyses, the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes with the most observations were used as defined, and the rare genotypes were merged within the  $\kappa$ -CN genotypes (Tables 1 and 2). Relative to the other genotypes within the  $\kappa$ -

CN AA genotype, the AAA<sub>2</sub>A<sub>2</sub>BB genotype (group 3) is much more frequent than the comparable genotype (group 7) within the  $\kappa$ -CN AB genotype from the Holstein data in this study (Table 2). This relationship was also observed in Italian data (1). These results indicate that the  $\beta$ -CN A<sub>2</sub> allele is associated more often with the  $\kappa$ -CN A allele than the  $\kappa$ -CN B allele, as also has been observed in the data file of crosses between Holsteins and Dutch-Friesians (4). These observations are also supported by the previously estimated frequencies for the two  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN haplotypes, 0.46 for AA<sub>2</sub>B and 0.08 for BA<sub>2</sub>B in the data for Holstein cows (22). The group (group 4) of rare genotypes within the  $\kappa$ -CN AA genotype (Table 2) includes mainly the composite genotypes with the C allele at the  $\alpha_{s1}$ -CN locus and either the A<sub>2</sub> or A<sub>3</sub> allele at the  $\beta$ -CN locus, as shown in Table 1 of this study and in the Dutch data (4). The group (group 8) of rare genotypes within the  $\kappa$ -CN AB genotype represents mainly the genotypes with the B allele at the  $\beta$ -CN locus (Tables 1 and 2).

The frequencies for 6 of the individual composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes were  $\geq 0.09$  for Jersey cows, and 4 of these were within the  $\kappa$ -CN BB genotype (Table 2). There were only four composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes

with frequency  $\geq 0.09$  for Holstein cows, implying concentration on fewer genotypes in this breed. These observations may be explained, in part, by the small data file for Jersey cows. The most frequent composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotype, BBA<sub>2</sub>BBC, for Jersey cows in this data was the same as in previous studies (2, 16).

### Effect of Milk Protein Genotypes on Milk Production Traits

**General observations.** The differences ( $P < 0.05$ ) between the milk protein genotypes were the same from the fixed model and the corresponding univariate, bivariate, and trivariate mixed models for Holstein and Jersey data. The magnitude of the estimates for the differences between the milk protein genotypes tended to be larger for fixed than for multivariate mixed models. The largest differences between the fixed and trivariate models deviated by 84 kg; differences between univariate and trivariate models deviated by 9 kg for milk production for the Holstein data for genotypes with few observations.

The differences between the  $\beta$ -LG genotypes for milk production traits were not statistically significant for the Jersey and Holstein data (Tables 3 and 4). The  $\beta$ -LG A allele has been reported to have a

TABLE 1. Number of observations for the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes in 916 Holstein and 116 Jersey cows in California.

$\beta$ -CN Genotype	Holstein $\alpha_{s1}$ -CN genotypes				Jersey $\alpha_{s1}$ -CN genotypes			
	AB	BB	BC	CC	AB	BB	BC	CC
	Within $\kappa$ -CN genotype AA (n = 620)				Within $\kappa$ -CN genotype AA (n = 2)			
A <sub>1</sub> A <sub>1</sub>	...	102	1	...	...	...	...	...
A <sub>1</sub> A <sub>2</sub>	4	297	2	...	...	...	...	...
A <sub>1</sub> A <sub>3</sub>	...	1	2	...	...	...	...	...
A <sub>1</sub> B	...	...	...	...	...	...	...	...
A <sub>2</sub> A <sub>2</sub>	...	203	5	...	...	1	1	...
A <sub>2</sub> A <sub>3</sub>	...	...	2	...	...	...	...	...
A <sub>2</sub> B	...	...	1	...	...	...	...	...
BB	...	...	...	...	...	...	...	...
	Within $\kappa$ -CN genotype AB (n = 267)				Within $\kappa$ -CN genotype AB (n = 24)			
A <sub>1</sub> A <sub>1</sub>	1	56	...	...	...	...	...	...
A <sub>1</sub> A <sub>2</sub>	2	133	...	...	...	1	4	...
A <sub>1</sub> A <sub>3</sub>	...	...	...	...	...	...	...	...
A <sub>1</sub> B	...	15	...	...	...	1	...	...
A <sub>2</sub> A <sub>2</sub>	...	44	1	...	...	2	11	1
A <sub>2</sub> A <sub>3</sub>	...	...	...	...	...	...	...	...
A <sub>2</sub> B	...	15	...	...	...	3	1	...
BB	...	...	...	...	...	...	...	...
	Within $\kappa$ -CN genotype BB (n = 29)				Within $\kappa$ -CN genotype BB (n = 90)			
A <sub>1</sub> A <sub>1</sub>	...	8	...	...	...	4	...	...
A <sub>1</sub> A <sub>2</sub>	...	10	...	...	...	2	13	...
A <sub>1</sub> A <sub>3</sub>	...	...	...	...	...	...	...	...
A <sub>1</sub> B	...	3	...	...	...	15	...	...
A <sub>2</sub> A <sub>2</sub>	...	3	...	...	...	1	7	14
A <sub>2</sub> A <sub>3</sub>	...	...	...	...	...	...	...	...
A <sub>2</sub> B	...	4	...	...	...	3	20	1
BB	...	1	...	...	...	10	...	...

TABLE 2. The groups in statistical analyses for composite  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotypes and for  $\beta$ -LG genotypes with the number of observations and frequencies for 916 Holstein and 116 Jersey cows in California.

Holstein			Jersey		
Group no. and genotype	Cows	Frequency	Group no. and genotype	Cows	Frequency
	(no.)			(no.)	
$\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotypes					
1. AAA <sub>1</sub> A <sub>1</sub> BB	102	0.11	1. ABA <sub>2</sub> A <sub>2</sub> BB	11	0.09
2. AAA <sub>1</sub> A <sub>2</sub> BB	297	0.32	2. ABrere <sup>2</sup>	15	0.13
3. AAA <sub>2</sub> A <sub>2</sub> BB	203	0.22	3. BBA <sub>1</sub> A <sub>2</sub> BC	13	0.11
4. AArere <sup>1</sup>	18	0.02	4. BBA <sub>1</sub> BBB	15	0.13
5. ABA <sub>1</sub> A <sub>1</sub> BB	56	0.06	5. BBA <sub>2</sub> A <sub>2</sub> CC	14	0.12
6. ABA <sub>1</sub> A <sub>2</sub> BB	133	0.15	6. BBA <sub>2</sub> BBC	20	0.17
7. ABA <sub>2</sub> A <sub>2</sub> BB	44	0.05	7. BBBBBB	10	0.09
8. ABrere	34	0.04	8. BBrere	18	0.16
9. BBrereBB	29	0.03			
$\beta$ -LG genotypes					
1. AA	153	0.17	1. AA	8	0.07
2. AB	471	0.51	2. AB	67	0.58
3. BB	292	0.32	3. BB	41	0.35

<sup>1</sup>re = A group of rare genotypes at the locus.

<sup>2</sup>The composite ABrere genotype includes two observations of the AArere genotypes.

favorable and statistically significant effect on protein percentage (5, 17, 18) and protein production (1, 4, 15, 17) in populations of purebred or crossbred Holstein cows. The  $\beta$ -LG B allele has been favorable for fat percentage (1, 4, 15, 17) and fat production (12, 17).

**Differences between the composite  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotypes.** Only in a few cases were the estimates of differences for the composite  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotypes for milk production traits large enough to be statistically significant for Jersey and Holstein data (Tables 3 and 4). The differences between the genotypes for milk production were fairly large in the Jersey data, especially relative to the rather low mean production (Table 3). The number of observations within the composite  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotypes was, however, too small to permit specific comparisons within this data. The sample sizes have also been small in other studies with Jersey cows (2, 16).

Compared with the most common  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotype of AAA<sub>1</sub>A<sub>2</sub>BB for the Holstein data, the best genotypes for milk yield of first lactation, had the  $\kappa$ -CN B and  $\beta$ -CN A<sub>2</sub> alleles in addition to the common  $\alpha_{S1}$ -CN B allele in the composite  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotype (Table 4). These genotypes were also among the best for protein production, although only the  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotype ABA<sub>1</sub>A<sub>2</sub>BB deviated significantly ( $P < 0.05$ ) for both milk and protein production from the most common genotype for the Holstein data. The  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN genotype ABA<sub>1</sub>A<sub>2</sub>BB was superior to the comparable AAA<sub>1</sub>A<sub>2</sub>BB and ABA<sub>1</sub>A<sub>1</sub>BB genotypes by 252

and 338 kg for milk production and 8.7 and 11.5 kg for protein production. Thus, neither the  $\beta$ -CN A<sub>2</sub> allele nor the  $\kappa$ -CN B allele alone had a positive effect on milk and protein production, but their joint effect was strongly positive. These results may be explained by epistatic effects between the  $\kappa$ -CN and  $\beta$ -CN loci and possibly by closely linked QTL with favorable alleles on the  $\kappa$ - $\beta$ - $\alpha_{S1}$ -CN BA<sub>2</sub>B haplotype.

The  $\kappa$ -CN B allele has been reported to be significant and favorable for milk production (13, 15), protein production (1, 15), and protein percentage (1, 4, 7, 15, 17, 18) in populations of purebred or crossbred Holstein cows. The  $\beta$ -CN A<sub>2</sub> allele has been reported to have favorable effects for milk (4, 12, 18), fat (12), and protein (4, 12) production. In the previous studies, the model included the milk protein loci as separate effects. Favorable and significant joint effects of the  $\kappa$ -CN B allele and the  $\beta$ -CN A<sub>2</sub> allele on milk and protein production have not been previously reported.

**Relative magnitude of the fixed effects.** The differences in milk production traits caused by the effects of milk protein genotypes (Table 4) were approximately of the same magnitude as the differences between the classes for age at calving effect (Table 5). Age at calving, followed by year and season of calving, were the least significant factors among the fixed effects in the Holstein data. Days open and herd were the most important factors contributing to the differences in milk production traits. The magnitude

of differences between the herds was about 10-fold compared with those between genotypes for milk protein. Herd effects of this magnitude in research data may affect the results obtained for the other effects, in this case, milk protein genotypes. To remove the herd effects effectively, analyses that were similar to those used for the entire data were performed within the largest Holstein herds with high and low milk production.

**Differences between milk protein genotypes within herds.** Higher than average frequencies for the composite genotypes within the  $\kappa$ -CN AB locus were observed in the herd with the second highest milk yield (Table 6). These genotypes were not, however, among the best genotypes for milk, fat, and protein production in this herd, which implies that the increased frequency was obviously not due to selection for increased milk production.

Compared with the most frequent genotypes in the Holstein data, statistically significant differences were sporadic between the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes for the five traits, and no significant differences were observed between  $\beta$ -LG genotypes within the three herds with high and low milk production (Table 6). No composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotype was uniformly the best within the three herds, regardless of the level of milk production. Sporadic results may be explained by insufficient sample sizes within the herds.

### Contribution of Milk Protein Genotypes to Additive Genetic Variance

Univariate, bivariate, and trivariate models were compared for their effectiveness at estimating genetic parameters. The estimates from multitrait analyses were biologically well-founded, especially for the estimates of heritability from the analyses of small data files, for the Jersey data, and for the within-herd analyses of Holstein data. For small samples, parameter estimates are likely to be on or close to the boundary of the parameter space for a particular point estimate.

After the milk protein genotypes were removed from the model, an increase in the additive genetic variance for the milk production traits, except fat production, was observed in the Holstein data (Table 7). This result implies that the alleles at the milk protein loci, or closely linked QTL, or both, contribute to the additive genetic variation in milk production traits. A parallel decrease was observed in residual variances, except for fat production. The increase in the additive genetic variance relative to the phenotypic variance resulted in an increase in the heritability estimates of 0.05 units for milk and protein production and 0.03 for fat percentage; no increase occurred for protein percentage or fat production. The increases in the additive genetic variance relative to the phenotypic variance were contributed almost exclusively by the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes. These results agree with and support previous find-

TABLE 3. Estimates of differences for composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and  $\beta$ -LG genotypes for milk production traits of first lactation Jersey cows in California.

Genotype	Cows (no.)	Trivariate model			Bivariate model	
		Milk production (kg)	Fat (%)	Protein (%)	Fat production (kg)	Protein production (kg)
Mean	116	5224	5.09	4.01	264	208
$\kappa$ - $\beta$ - $\alpha_{s1}$ -CN						
ABA <sub>2</sub> A <sub>2</sub> BB	11	-219	0.20	0.04	0.7	-6.6
ABrere <sup>1</sup>	15	-279	0.18	-0.03	-0.2	-9.1
BBA <sub>1</sub> A <sub>2</sub> BC	13	86	-0.06	-0.15*	3.8	-5.0
BBA <sub>1</sub> BBB	15	48	0.00	0.02	10.1	7.4
BBA <sub>2</sub> A <sub>2</sub> CC	14	-144	0.00	0.05	-5.5	-1.3
BBA <sub>2</sub> BBC <sup>2</sup>	20	0	0.0	0.0	0.0	0.0
BBBBBB	10	-294	0.12	0.07	-5.5	-6.0
BBrere	18	22	0.10	0.10	8.0	6.2
$\beta$ -LG						
AA	8	119	-0.06	0.06	2.4	8.6
AB <sup>2</sup>	67	0	0.0	0.0	0.0	0.0
BB	41	-82	-0.04	-0.01	-5.4	-3.0

<sup>1</sup>re = A group of rare genotypes at the locus.

<sup>2</sup>The genotype of comparison for the paired contrasts.

\* $P < 0.05$ .

ings and conclusions about the effects of the composite casein genotypes in the Holstein data.

The proportion of phenotypic variance that was due to milk protein genotypes was moderate for milk and protein production and for fat percentage in this study compared with the corresponding estimates from other studies. In this study, the variance components for additive genetic and residual effects were estimated using a multiple-trait animal model. When estimates were based on gene frequencies and gene substitution effects (12), the contribution of the milk protein genotypes accounted for almost 9% of a phenotypic variance in milk and protein production and 5% of that in fat production. The corresponding proportions, estimated based on sire variances, were close to 0 for milk, fat, and protein production (8) and for the five milk production traits (15). The large variability of available estimates is obviously due to different methods of estimation, models assumed, and the size and structure of data files.

### CONCLUSIONS

The use of an animal model in place of a fixed model reduced the inflated estimates of differences for milk protein genotypes. Estimates were again reduced slightly after the genetic correlations among the milk production traits were accounted for with

bivariate and trivariate animal models. Because of a parallel reduction in the standard errors of the estimates, no difference in statistically significant effects between the models was observed for these data files. Realistic estimates are necessary for the differences between milk protein genotypes if the genotypes are to be used to make selection decisions. Thus, analyses using a multiple-trait animal model should be preferred when computationally feasible.

The composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes, in place of separate effects for the three loci, was used in the model in accordance with the well-established close linkage between the casein loci. This method should generally be preferred, for estimation of the effects of closely linked loci on quantitative traits. Grouping the  $\beta$ -CN and  $\alpha_{s1}$ -CN genotypes within the  $\kappa$ -CN genotypes resulted in a reasonable number of well-defined composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and groups of genotypes in both Holstein and Jersey data. Compared with the  $\beta$ -LG genotypes, the composite casein genotypes had a major contribution, 3 to 5%, to the phenotypic variance in fat percentage and milk and protein production in the Holstein data. The  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotype  $ABA_1A_2BB$  was superior to the comparable  $AAA_1A_2BB$  and  $ABA_1A_1BB$  genotypes by 252 and 338 kg for first lactation milk production and 8.7 and 11.5 kg for protein production. Thus, neither the  $\beta$ -CN  $A_2$  allele nor the  $\kappa$ -CN B allele alone had a positive effect

TABLE 4. Estimates of differences for composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and  $\beta$ -LG genotypes for milk production traits of first lactation Holstein cows in California.

Genotype	Cows	Trivariate model			Bivariate model	
		Milk production	Fat	Protein	Fat production	Protein production
	(no.)	(kg)	(%)	(%)	(kg)	(kg)
Mean	916	8539	3.74	3.27	317	278
$\kappa$ - $\beta$ - $\alpha_{s1}$ -CN						
AAA <sub>1</sub> A <sub>1</sub> BB	102	-35	0.03	0.00	1.5	-1.3
AAA <sub>1</sub> A <sub>2</sub> BB <sup>1</sup>	297	0	0.0	0.0	0.0	0.0
AAA <sub>2</sub> A <sub>2</sub> BB	203	44	-0.05	-0.02	-2.8	-0.7
AAre <sup>2</sup>	18	185	-0.07	0.03	1.8	8.8
ABA <sub>1</sub> A <sub>1</sub> BB	56	-86	-0.02	0.00	-5.5	-2.8
ABA <sub>1</sub> A <sub>2</sub> BB	133	252*	-0.03	0.01	6.5	8.7*
ABA <sub>2</sub> A <sub>2</sub> BB	44	289	-0.11 <sup>†</sup>	-0.03	1.2	5.8
ABre	34	99	-0.07	-0.00	-0.3	3.8
BBreBB	29	191	-0.07	0.01	0.5	6.5
$\beta$ -LG						
AA	153	61	0.01	0.02	3.4	3.7
AB <sup>1</sup>	471	0	0.0	0.0	0.0	0.0
BB	292	35	0.04	-0.01	4.9	0.5

<sup>1</sup>The genotype of comparison for the paired contrasts.

<sup>2</sup>re = A group of rare genotypes at the locus.

<sup>†</sup> $P < 0.10$ .

\* $P < 0.05$ .

on milk and protein production, but the joint effect was strongly positive. These findings were not consistent, however, across herds.

The results from this study may be interpreted by epistatic effects between the  $\kappa$ -CN and  $\beta$ -CN loci and possibly by closely linked QTL with favorable alleles on the  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN BA<sub>2</sub>B haplotype. Previous findings may partially explain some of the contradictory results reported in the literature for the separate

effects of the alleles at the  $\kappa$ -CN and  $\beta$ -CN loci. The  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN BA<sub>2</sub>B haplotype has been reported to be the fourth most common, with a frequency of 0.08, for Holstein cows in California (22). Thus, before the composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes or the corresponding haplotypes are used for selection to improve milk production traits, these results should be confirmed using sufficiently large data for more than one breed and possibly within large herds as well.

TABLE 5. Estimates of differences between classes for herd, year and season of calving, age at calving, and days open effects for milk production traits of first lactation Holstein cows in California.

Classes of fixed effects	Cows	Trivariate model			Bivariate model	
		Milk production	Fat	Protein	Fat production	Protein production
		(no.)	(kg)	(%)		(kg)
Mean	916	8539	3.74	3.27	317	278
Herd						
1	9	-57	-0.01	-0.16**	-6.3	-15.3
2	123	777***	-0.45***	0.06*	-9.2	31.8***
3 <sup>1</sup>	294	0	0.0	0.0	0.0	0.0
4	19	1402***	0.10	-0.09*	64.3***	39.7***
5	18	667*	-0.01	-0.12**	22.8	10.5
6	42	2363**	0.05	-0.14***	95.2***	64.6***
7	169	311*	0.02	-0.10***	13.8**	2.4
8	133	1904***	-0.33***	-0.16***	41.1***	47.3***
9	23	410	-0.34***	-0.01	-12.9	12.4
10	86	888***	-0.07	-0.09***	28.6***	22.3***
Year and season of calving						
Nov 1987–Aug 1988	70	256	0.01	-0.05*	9.9	3.1
Sep–Oct 1988	77	131	0.10*	-0.01	13.0*	3.5
Nov–Dec 1988	131	229	0.09*	0.01	15.0**	8.1
Jan–Feb 1989 <sup>1</sup>	240	0	0.0	0.0	0.0	0.0
Mar–Apr 1989	234	-58	0.05	0.05**	1.7	2.2
May–Jun 1989	114	-88	0.12**	0.04	7.1	0.1
Jul–Sep 1989	50	-261	0.09	0.05	-3.7	-5.0
Age at calving						
<751 d	61	47	-0.09	-0.02	-6.5	0.0
751 to 780 d	156	-218	0.00	-0.00	-8.0	-7.3
781 to 810 d <sup>1</sup>	196	0	0.0	0.0	0.0	0.0
811 to 840 d	159	106	-0.01	0.00	2.6	3.4
841 to 870 d	114	-119	0.11*	0.01	5.7	-2.6
871 to 900 d	85	-10	0.05	0.01	4.4	0.4
901 to 930 d	56	-6	0.02	0.03	1.3	2.0
>930 d	89	141	0.04	-0.01	9.0	4.8
Days open						
<61 d	61	-1087***	0.11	0.06*	-32.6***	-31.5***
61 to 75 d	114	-807***	0.10*	0.03	-21.0***	-23.9***
76 to 90 d	117	-263	0.01	0.02	-8.7	-6.3
91 to 105 d	93	46	0.08	-0.02	9.8	0.6
106 to 120 d	73	-55	0.03	-0.01	-0.4	-2.9
121 to 150 d <sup>1</sup>	121	0	0.0	0.0	0.0	0.0
151 to 180 d	96	73	0.03	-0.01	5.8	2.3
181 to 240 d	92	28	-0.02	0.00	-2.5	0.8
241 to 304 d	65	137	-0.04	-0.02	1.3	3.2
>304 d	84	315	0.00	-0.02	12.1	9.0

<sup>1</sup>The class of comparison for the paired contrasts.

\* $P < 0.05$ .

\*\* $P < 0.01$ .

\*\*\* $P < 0.001$ .



TABLE 6. Estimates of differences for composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and  $\beta$ -LG genotypes for milk production traits within three Holstein herds in California.

Genotypes within herds	Cows	Frequency	Trivariate model			Bivariate model	
			Milk production	Fat	Protein	Fat production	Protein production
	(no.)		(kg)	(%)		(kg)	
Herd 6 (the highest milk production)							
Mean	42		10650	3.90	3.21	412	340
$\kappa$ - $\beta$ - $\alpha_{s1}$ -CN							
AAA <sub>1</sub> A <sub>1</sub> BB	4	0.10	126	-0.07	-0.01	2.4	4.4
AAA <sub>1</sub> A <sub>2</sub> BB <sup>1</sup>	11	0.26	0	0.0	0.0	0.0	0.0
AAA <sub>2</sub> A <sub>2</sub> BB	13	0.31	861	-0.30 <sup>†</sup>	-0.05	-1.3	23.8
AAre <sup>2</sup>	...	...	...	...	...	...	...
ABA <sub>1</sub> A <sub>1</sub> BB	3	0.07	1757 <sup>†</sup>	-0.58*	-0.05	5.6	48.8 <sup>†</sup>
ABA <sub>1</sub> A <sub>2</sub> BB	2	0.05	1192	-0.48	-0.06	5.4	33.5
ABA <sub>2</sub> A <sub>2</sub> BB	1	0.02	2698	-0.72	-0.13	28.1	67.3
ABre	5	0.12	237	0.05	0.23 <sup>†</sup>	15.2	30.2
BBreBB	3	0.07	-279	-0.05	0.02	-5.7	-1.3
$\beta$ -LG							
AA	10	0.24	-183	-0.02	-0.05	-8.4	-8.7
AB <sup>1</sup>	20	0.48	0	0.0	0.0	0.0	0.0
BB	12	0.29	-662	0.11	-0.05	-11.6	-25.5
Herd 8 (the second highest milk production)							
Mean	133		9826	3.51	3.15	344	309
$\kappa$ - $\beta$ - $\alpha_{s1}$ -CN							
AAA <sub>1</sub> A <sub>1</sub> BB	9	0.07	-61	-0.07	0.02	-8.0	-1.6
AAA <sub>1</sub> A <sub>2</sub> BB <sup>1</sup>	31	0.23	0	0.0	0.0	0.0	0.0
AAA <sub>2</sub> A <sub>2</sub> BB	31	0.23	-40	-0.05	-0.04	-5.2	-5.5
AAre	...	...	...	...	...	...	...
ABA <sub>1</sub> A <sub>1</sub> BB	13	0.10	-440	-0.02	-0.01	-19.0 <sup>†</sup>	-14.6
ABA <sub>1</sub> A <sub>2</sub> BB	25	0.19	-269	-0.04	-0.01	-15.2 <sup>†</sup>	-10.3
ABA <sub>2</sub> A <sub>2</sub> BB	13	0.10	-67	-0.08	-0.01	-12.9	-3.6
ABre	7	0.05	28	0.06	0.04	3.8	5.6
BBreBB	4	0.03	-49	-0.15	0.04	-17.8	3.8
$\beta$ -LG							
AA	20	0.15	-221	0.12	0.04	2.7	-3.4
AB <sup>1</sup>	70	0.53	0	0.0	0.0	0.0	0.0
BB	43	0.32	-11	0.06	-0.04	4.1	-4.1
Herd 3 (the second lowest milk production)							
Mean	294		7731	3.86	3.34	297	257
$\kappa$ - $\beta$ - $\alpha_{s1}$ -CN							
AAA <sub>1</sub> A <sub>1</sub> BB	38	0.13	-188	0.08	0.04	0.3	-2.9
AAA <sub>1</sub> A <sub>2</sub> BB <sup>1</sup>	92	0.31	0	0.0	0.0	0.0	0.0
AAA <sub>2</sub> A <sub>2</sub> BB	59	0.20	-350	0.07	0.04	-7.0	-7.9
AAre	10	0.03	20	0.06	0.06	8.4	7.3
ABA <sub>1</sub> A <sub>1</sub> BB	22	0.07	-355	0.08	0.05	-5.4	-6.5
ABA <sub>1</sub> A <sub>2</sub> BB	36	0.12	-61	0.01	0.03	0.5	1.5
ABA <sub>2</sub> A <sub>2</sub> BB	9	0.03	404	-0.21	-0.16**	0.4	1.3
ABre	17	0.06	-60	-0.07	-0.02	-4.6	-3.1
BBreBB	11	0.04	77	0.07	0.03	9.5	4.7
$\beta$ -LG							
AA	40	0.14	206	0.05	0.01	12.4	6.8
AB <sup>1</sup>	154	0.52	0	0.0	0.0	0.0	0.0
BB	100	0.34	180	0.04	0.01	10.8	6.8

<sup>1</sup>The genotype of comparison for the paired contrasts.

<sup>2</sup>re = A group of rare genotypes at the locus.

<sup>†</sup> $P < 0.10$ .

\* $P < 0.05$ .

\*\* $P < 0.01$ .

TABLE 7. The REML estimates of additive genetic ( $\sigma_a^2$ ) and residual ( $\sigma_e^2$ ) variance components and heritability for milk production traits of first lactation Holstein cows based on Models [1] and [1a].

Model and trait	Model [1]			Model [1a] <sup>1</sup>				
	$\sigma_a^2$	$\sigma_e^2$	$h_1^2$	$\sigma_a^2$	$\sigma_e^2$	$h_{1a}^2$	CCN <sup>2)</sup>	CLG <sup>3)</sup>
Trivariate model								
Milk production	394,406	1,066,957	0.27	466,463	999,287	0.32	0.05	0.00
Fat percentage	0.042	0.087	0.32	0.045	0.084	0.35	0.03	0.00
Protein percentage	0.017	0.016	0.51	0.017	0.016	0.51	-0.01	0.01
Bivariate model								
Fat production	562	1550	0.27	544	1557	0.26	0.01	-0.02
Protein production	656	739	0.47	731	675	0.52	0.04	0.01

<sup>1</sup>The composite  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes and  $\beta$ -LG genotypes were not included in Model [1a].

<sup>2</sup>Contribution of the  $\kappa$ - $\beta$ - $\alpha_{s1}$ -CN genotypes to the phenotypic variance and to  $h_{1a}^2$ .

<sup>3</sup>Contribution of the  $\beta$ -LG genotypes to the phenotypic variance and to  $h_{1a}^2$ .

## ACKNOWLEDGMENTS

Appreciation is extended to Curt Finley for his assistance and advice in collecting and editing the data files and for writing a Fortran program to renumber the animals in the pedigree file in the order of their putative age. The leave of absence of the first author at the Department of Animal Science, University of California, Davis was supported by the grant from the Academy of Finland.

## REFERENCES

- Aleandri, R., L. G. Buttazzoni, J. C. Schneider, A. Caroli, and R. Davoli. 1990. The effects of milk protein polymorphisms on milk components and cheese-producing ability. *J. Dairy Sci.* 73:241.
- Bech, A.-M., and K. R. Kristiansen. 1990. Milk protein polymorphism in Danish dairy cattle and the influence of genetic variants on milk yield. *J. Dairy Res.* 57:53.
- Boldman, K. G., L. A. Kriese, L. D. Van Vleck, C. P. Van Tassel, and S. D. Kachman. 1995. A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances. ARS-USDA, Lincoln, NE.
- Bovenhuis, H., J.A.M. van Arendonk, and S. Korver. 1992. Associations between milk protein polymorphisms and milk production traits. *J. Dairy Sci.* 75:2549.
- Famula, T. R., and J. F. Medrano. 1994. Estimation of genotype effects for milk proteins with animal and sire transmitting ability models. *J. Dairy Sci.* 77:3153.
- Ferretti, L., P. Leone, and V. Sgaramella. 1990. Long range restriction analysis of the bovine casein genes. *Nucleic Acids Res.* 18:6829.
- Gonyon, D. S., R. E. Mather, H. C. Hines, G.F.W. Haenlein, C. W. Arave, and S. N. Gaunt. 1987. Associations of bovine blood and milk polymorphisms with lactation traits: Holsteins. *J. Dairy Sci.* 70:2585.
- Graml, R., J. Buchberger, H. Klostermeyer, and F. Pirchner. 1986. Pleiotrope Wirkungen von  $\beta$ -Lactoglobulin- und Casein-Genotypen auf Milchlaktation und Milchproteinmengen des bayerischen Fleckviehs und Braunviehs. *J. Anim. Breed. Genet.* 103:33.
- Grosclaude, F., J. C. Mercier, and B. Ribadeau Dumas. 1973. Genetic aspects of cattle casein research. *Neth. Milk Dairy J.* 27:328.
- Henderson, C. R. 1984. Applications of Linear Models in Animal Breeding. Univ. Guelph Press, Guelph, ON, Canada.
- Kennedy, B. W., M. Quinton, and J.A.M. van Arendonk. 1992. Estimation of effects of single genes on quantitative traits. *J. Anim. Sci.* 70:2000.
- Lin, C. Y., A. J. McAllister, K. F. Ng-Kwai-Hang, and J. F. Hayes. 1986. Effects of milk protein loci on first lactation production in dairy cattle. *J. Dairy Sci.* 69:704.
- Lin, C. Y., A. J. McAllister, K. F. Ng-Kwai-Hang, J. F. Hayes, T. R. Batra, A. J. Lee, G. L. Roy, J. A. Vesely, J. M. Wauthy, and K. A. Winter. 1989. Relationships of milk protein types to lifetime performance. *J. Dairy Sci.* 72:3085.
- Lin, C. Y., M. P. Sabour, and A. J. Lee. 1992. Direct typing of milk proteins as an aid for genetic improvement of dairy bulls and cows: a review. *Anim. Breed. Abstr.* 60:1.
- Mao, I. L., L. G. Buttazzoni, and R. Aleandri. 1992. Effects of polymorphic milk protein genes on milk yield and composition traits in Holstein cattle. *Acta Agric. Scand., Sect. A, Anim. Sci.* 42:1.
- McLean, D. M., E.R.B. Graham, R. W. Ponzoni, and H. A. McKenzie. 1984. Effects of milk protein genetic variants on milk yield and composition. *J. Dairy Res.* 51:531.
- Ng-Kwai-Hang, K. F., J. F. Hayes, J. E. Moxley, and H. G. Monardes. 1984. Associations of genetic variants of casein and milk serum proteins with milk, fat, and protein production by dairy cattle. *J. Dairy Sci.* 67:835.
- Ng-Kwai-Hang, K. F., H. G. Monardes, and J. F. Hayes. 1990. Association between genetic polymorphism of milk proteins and production traits during three lactations. *J. Dairy Sci.* 73:3414.
- SAS/STAT<sup>®</sup> User's Guide, Version 6, Fourth Edition, Volume 2. 1989. SAS Inst., Inc., Cary, NC.
- Searle, S. R., G. Casella, and C. E. McCulloch. 1992. Variance Components. John Wiley & Sons, Inc., New York, NY.
- Threadgill, D. W., and J. E. Womack. 1990. Genomic analysis of the major bovine milk protein genes. *Nucleic Acids Res.* 18: 6935.
- Van Eenennaam, A. L., and J. F. Medrano. 1991. Milk protein polymorphisms in California dairy cattle. *J. Dairy Sci.* 74:1730.