GENETICS AND BREEDING

Effects of Maternal Lineages Grouped by Mitochondrial Genotypes on Milk Yield and Composition[1]

M. M. SCHUTZ,[2] A. E. FREEMAN, G. L. LINDBERG, and D. C. BEITZ[3]
Department of Animal Science
Iowa State University
Ames 50011

ABSTRACT

Maternal lineages grouped by several methods of classification using mitochondrial DNA sequence characteristics were evaluated with animal models. Maternal pedigrees for cattle from a selection experiment begun in 1968 were traced to the earliest female member in the *Holstein-Friesian Herd Book*, and foundation females were assigned to maternal lineages. Mitochondrial DNA displacement loop sequence data were available for 36 lineages; all cows within lineages were considered to be identical for useful DNA sequence polymorphisms. Maternal lineages were grouped according to base pair substitution at a single locus, clustering based on displacement loop sequences, or genotypes defined by sequence and restriction site differences. Base pair substitution (adenine to guanine) at nucleotide 169 defined two maternal lineage groups that differed significantly for fat yield and percentage and estimated milk energy. Clustering the 36 maternal lineages by using 16 mitochondrial DNA displacement loop sequence differences produced 24 groups that significantly influenced fat percentage and energy concentration in milk. Decreasing the number of clustered groups from 24 to 14 produced groups that differed at some polymorphic sites but remained identical for most. The *F* statistics for 14 groups were larger, but significant outcomes were observed for the same traits. Genotype groups previously defined by sequence and restriction fragment length differences did not have a significant effect on yield or composition of cows in 29 maternal lineages.

(**Key words**: maternal lineage, cytoplasmic inheritance, mitochondrial inheritance, mitochondrial deoxyribonucleic acid clustering)

Abbreviation key: **D** = displacement, **mtDNA** = mitochondrial DNA, **nt** = nucleotide.

INTRODUCTION

Several reports have suggested the existence of maternal lineage effects on yield or composition (2, 5, 11, 23) and reproduction (2, 4, 20, 22); those results indicate the possibility of cytoplasmic inheritance of genetic effects. Mitochondrial DNA (**mtDNA**) is a probable source of such cytoplasmic effects because mitochondria in metazoans are thought to be transmitted predominantly from female parents to offspring (7), although limited biparental inheritance of mtDNA has been established in mussels (10) and mice (6).

Some reports have challenged the significance of previous estimates of cytoplasmic inheritance by using simulated data (14) or analysis of trios of daughters, dams, and grandams (21). Nevertheless, the ability of well-defined animal models to partition maternal influences into cytoplasmic and additive genetic components has been demonstrated (25). Southwood et al. (25) used simulated data and true or incorrect models containing additive direct, additive maternal, cytoplasmic, and error variances and concluded that certain animal models correctly partition these components of variance. Schutz et al. (23) found that additive maternal variation of nuclear origin was negligible for yield and composition traits of dairy cattle.

Received June 19, 1992.
Accepted September 29, 1992.
[1]Journal Paper Number J-14915 of the Iowa Agriculture and Home Economics Experiment Station, Ames. Project Number 1053.
[2]Present address: Animal Improvement Programs Laboratory, USDA-ARS, Beltsville, MD 20705-2350.
[3]Department of Biochemistry and Biophysics.

Mitochondrial DNA sequence differs among dairy cattle (3, 17). Koehler (15) used restriction enzymes to detect 11 polymorphisms among maternal lineages, and two additional polymorphisms occurred within lineages. One polymorphism within a single lineage seemed to result from a single mutational event, but the other polymorphism occurred within several lineages and is considered to be heteroplasmic for lineages (16). No other such sites of heteroplasmy have been documented. Johnston et al. (1991, unpublished data) identified 10 nucleotide (**nt**) substitutions in the displacement (**D**) loop of mtDNA from five breeds of dairy cattle using restriction analysis. Lindberg et al. (19) sequenced entire mtDNA D loops of 36 maternal lineages and identified 48 sites of nt substitution (plus one deletion) and two variable length regions. In addition, Johnston (13) observed 11 different sequences in ribosomal RNA subunits of mtDNA in 38 lineages of Holsteins.

Studies associating maternal lineages with yield and composition traits have defined maternal lineage sources as foundation cows in the herd studied (2, 11, 20) or as maternal lineage matriarchs identified by tracing maternal pedigrees to the beginning of a herd book (5, 23). Defining maternal lineages this way has several disadvantages. In either definition, several lineages may have a single common maternal ancestor before earliest recorded pedigree information. Also, misidentification of maternal ancestors or erroneously recorded registration numbers lead to incorrect assignment to lineages. Most commonly, branches of a single maternal pedigree are inadvertently defined as separate maternal lineages. These disadvantages suggest that previous studies may have underestimated the true impact of cytoplasmic effects on milk yield and reproduction.

Lindberg et al. (19) demonstrated the potential for 27 restriction endonuclease target sites in the bovine mtDNA D-loop region to be used to describe Holstein cytoplasmic genotypes. Single base pair substitutions in bovine mtDNA D loops have been associated with yield and composition differences in Holsteins (3, 23). The objective of this study was to define maternal lineages according to several different mtDNA molecular characterizations and to associate these new lineage groupings with yield and composition traits in a herd of Holsteins.

MATERIALS AND METHODS

Cattle studied were part of a selection experiment begun with the Iowa State University Breeding Research Herd in 1968. Foundation females were mated to Holstein AI sires with high or average transmitting abilities for milk yield to form two divergent genetic lines. Records initiated through 1988 were included in analyses.

Genetic backgrounds of original foundation females in the herd were quite diverse, but frequencies of bovine lymphocyte antigen genotypes were similar to those in the US Holstein population (28), strongly suggesting that nuclear genes are representative of that population. Cows were bred artificially to sires from commercial AI, allowing continuous introduction of nuclear genes.

The herd was assembled through purchases of 158 foundation females acquired from 38 Holstein breeders located throughout Iowa. Because maternal heritage was determined by tracing maternal lineage to the first female member recorded in the *Holstein-Friesian Herd Book* (27), only the 133 foundation registered females were assigned to 81 separate maternal lineages. Thirty-six registered lineages had members remaining in the herd when restriction and sequence analysis of the mtDNA genome was conducted. There were 71 foundation females in these 36 lineages, and they were, on average, 19 generations removed from their matriarchs first registered in the herd book. Among the 36 lineages were 6 foundation females in the herd for lineages with the most foundation females and 1 for the lineage with the fewest. There was a total of 728 cows across generations in the 36 maternal lineages. These same lineages have been analyzed in previous studies (3, 22, 23).

Mature equivalent (twice daily milking, 305-d lactation) milk, fat, and SNF yields and fat and SNF percentages were the yield and composition traits available for analysis. Milk net energy is affected by maternal lineages (23). Milk net energy concentration in kilocalories per kilogram was calculated by using lactation average of test day fat and SNF per-

centages according to Tyrrell and Reid (26) as follows:

$$\text{net energy} = 41.84\ (\text{fat percentage}) + 22.29\ (\text{SNF percentage}) - 25.58.$$

Lactation net energy was calculated by multiplying net energy concentration by mature equivalent milk yield. Records were assigned to year-season of calving subclasses. Seasons were October to April and May to September to account for winter and summer feeding and management differences; there were an average of 54 records per year-season. Up to 7 records per cow were included for analyses.

The 36 maternal lineages were grouped in three ways, based on characterization of mtDNA. First, maternal lineages were grouped according to their genotype at one particular nt site. Second, cluster analysis grouped lineages based on genotypes defined by 16 polymorphic sites in the mtDNA D loop. Finally, lineages were grouped according to sequence genotypes from a previous report by Lindberg et al. (19). The mtDNA D-loop sequences of these 36 lineages have been well documented, and details of the molecular characterizations have been reported (15, 16, 18, 19).

Base Pair 169

Lindberg et al. (19) used mtDNA D loop nt sequence data from 28 of these 36 maternal lineages to construct a phylogenetic tree by parsimony analysis. Their analysis generated a two-part evolutionary tree based on division of the population at nt169, which was numbered according to the nomenclature of Anderson et al. (1). Nucleotide information for this study was available for cows from all 36 maternal lineages. All members of a single maternal lineage were assumed to have an identical mtDNA genotype with respect to nt169. Spot checking of more cows in several lineages supported this assumption. Lineages were classified as identical or polymorphic compared with the first published mtDNA sequence (1) at nt169. The effect of these binomial data on each yield or composition trait was analyzed using the following single-trait mixed animal model:

$$y_{ijklp} = ys_i + p_j + x_k + b_l(\text{nt169}) + pe_p + a_p + e_{ijklp}, \quad [1]$$

where y_{ijklp} = milk, fat, or SNF yield; fat or SNF percentage; energy concentration; or lactation energy; ys_i = effect of calving year-season i (i = 1 to 33); p_j = effect of parity j (j = 1 to 7); x_k = effect of selection line k (high or average); b_l = effect of mtDNA D-loop sequence l (identical or polymorphic) at nt169; pe_p = permanent environmental effect of cow p (p = 1 to 728); a_p = additive genetic value of cow p and was composed of sire and dam breeding values and a Mendelian sampling effect; and e_{ijklp} = residual error.

Year-season, parity, selection line, and D-loop sequence at nt169 were considered to be fixed effects. Confounding was not detected between polymorphism at nt169 and other fixed effects. Permanent environment, animal additive genetic, and residual effects were considered to be random and independently distributed with expectations equal to 0. Variance among permanent environments was var(pe) = $\mathbf{I}\sigma^2_{pe}$, where **I** is an identity matrix, and variance among animals was var(a) = $\mathbf{A}\sigma^2_a$, where **A** is the numerator relationship matrix. Matrix **A** included sires and dams of all cows with records in the herd, sire and paternal grandsire relationships among AI bulls represented by daughters with records, and 25 unknown-parent groups for a total of 950 animal equations. Estimates of variances for each yield or composition trait were obtained from Schutz et al. (23).

The effect of nt substitution at nt169 was of primary interest. Significance of this substitution effect (under the null hypothesis that effect of all mtDNA D-loop sequences are the same) was tested by solution of mixed model conjugate normal equations (8, 9, 23). Such tests of significance are exact under the assumption of normality if estimates of variances are assumed to be accurate.

Maternal Lineage Clusters

The same 36 maternal lineages were grouped by clustering techniques that were based on mtDNA D-loop sequence information. The 17 most frequent sequence substitution sites in maternal lineages in this herd were

described by Schutz et al. (22). One of those sites, nt363, was hypervariable and, therefore, not useful for clustering of lineages. For each of the remaining 16 nt positions, maternal lineages were classified as identical or polymorphic with respect to the first published bovine mtDNA sequence (1). Each lineage had a total of 16 values, and lineages with the same polymorphisms were equal in their combination of identical or polymorphic indicators.

A matrix of distances between maternal lineages based on binomial sequence data was computed as 1 − Jaccard coefficient for each pair of lineages. Jaccard coefficients (12) are measures of similarity calculated as

$$\text{Jaccard coefficient } (i,j) = n_i/(n_1 + n_2 + n_3)$$

where n_1 = number of nt sites polymorphic in both maternal lineages i and j, n_2 = number of nt sites polymorphic only in maternal lineage i, and n_3 = number of nt sites polymorphic only in maternal lineage j. Number of nt sites not polymorphic in either maternal lineage does not enter the equation because this occurrence would be far more common and because inclusion would decrease the impact of polymorphic sites in determining similarities or distances between lineages. Thus, a matrix of distances was created that was of the order of the number of maternal lineages. The matrix was symmetric with zeros on diagonals and 1 − Jaccard coefficients on the off-diagonals. Off-diagonals for two maternal lineages with identical polymorphisms also were zero.

Based on the matrix of pairwise distances, maternal lineages were organized into the most homogeneous groups possible by average linked cluster analysis (24). Average linkage clustering of groups is based on average distance between pairs of observations (one in each group) and on making all possible comparisons of each member of one group with each member of the group being compared. This clustering method tends to form clusters with small variances and avoids extreme results inherent to other methods that consider only the nearest or farthest members from each group compared.

Two lineages were not polymorphic at all 16 nt sites. These lineages were joined to form a single cluster but were not so assigned by average linkage analysis, which only joined clusters on the basis of polymorphic nt sites. The 36 lineages were grouped into maternal lineage clusters of either 24 groups (clustering 1) or 14 groups (clustering 2) for separate analyses. The term "clustering" is used instead of "cluster" to avoid confusion with the individually clustered groups. With cluster analysis, any number of clusters may be chosen. Clustering into 24 groups was the first clustering for which the normalized distance between groups joined was greater than 0. Clustering into 14 groups was somewhat arbitrarily chosen as the smallest number of groups for which the joined groups had normalized distances of less than .6.

Clusterings of maternal lineages were analyzed with the following animal model:

$$y_{ijkmp} = \mu + ys_i + p_j + x_k + mlc_m + pe_p + a_p + e_{ijkmp}, \quad [2]$$

where all effects are as previously defined, except that mlc_m was the fixed effect of maternal lineage cluster group m (m = 1 to 24 for clustering 1 and m = 1 to 14 for clustering 2). Assumptions, expectations, and variances of random effects are as defined previously. Confounding was not detected between cluster groups and other fixed effects. Residuals (e_{ijkmp}) obtained with either clustering 1 or clustering 2 from this model were examined with a model consisting of actual maternal lineages traced by pedigree. Size of maternal lineage effects on the residuals were used to compare the ability of the two clusterings to account for maternal lineage differences.

Sequence Genotypes

Lindberg (18) reported that restriction fragment length polymorphism and sequence data could accurately describe mitochondrial genotypes of bovine cytoplasmic lineages. In his report [Table 2 (18)], 29 lineages were assigned to 21 genotypes according to restriction enzyme and sequence analyses. Sequence information was obtained on two or more members of each lineage, and sequence differences were confirmed, if detectable, by restriction enzyme analysis. One of the 29 lineages was not registered, and maternal ancestry could not be verified through registry information, but cows in this lineage were included in this

TABLE 1. Yield and composition traits for 1800 records of 728 cows in 36 maternal lineages.

Trait	$\bar{X}$	SD
ME[1] Milk, kg	8085	1771
ME Fat, kg	288	60
Fat percentage	3.63	.44
SNF, kg	745	165
SNF, %	9.21	.40
Energy concentration, kcal/kg	732	50
Lactation energy, Mcal	5888	1225

[1]Mature equivalent (twice daily milking, 305-d lactation).

TABLE 2. Effect of mitochondrial DNA nucleotide 169 type on yield and composition traits.

Trait	Effect[1]	t^2
ME[3] Milk, kg	349	.96
ME Fat, kg	39	2.77**
Fat percentage	.16	3.30**
SNF, kg	49	1.48
SNF, %	.06	1.56
Energy concentration, kcal/kg	8	3.05**
Lactation energy, Mcal	277	2.30*

[1]Polymorphic minus identical effects for which sequence was classified as identical or polymorphic with respect to the first published sequence by Anderson et al. (1).

[2]The null hypothesis is that effects of identical and polymorphic sequences are the same; df = 1759.

[3]Mature equivalent (twice daily milking, 305-d lactation).

*$P > t \leq .05$.

**$P > t \leq .01$.

portion of this study because genotype as defined by mtDNA D-loop sequence was known (19).

The model employed to examine effects of mtDNA genotypes on yield and composition traits was

$$y_{ijknp} = ys_i + p_j + s_k + g_n + pe_p + a_p + e_{ijknp}, \quad [3]$$

where all effects again were as previously defined, except that g_n was the fixed effect of mtDNA genotype n (n = 1 to 21). Numbers of fixed effect classes remained the same; however, because the number of lineages was fewer, there were only 1407 records of 572 cows (pe_p, p = 572) and 787 animal equations in these 29 maternal lineages.

RESULTS AND DISCUSSION

Base Pair 169

Overall means and standard deviations of records of all cows in the available 36 maternal lineages are in Table 1. Of the 728 cows in these lineages, 583 (80.1%) were polymorphic at mtDNA D loop nt169 with respect to the first published sequence (1). Probably, that first cow had a sequence that was less common than that of the entire Holstein population. These 583 cows had 1468 records versus 332 records for 145 cows identical to the reference sequence.

Table 2 has effect of nt169 sequence type on yield and composition traits. Adenine to guanine transition on heavy strand mtDNA at nt169 had highly significant effects on fat yield and percentage and on energy concentration ($P > t \leq .01$). The effect on lactation energy concentration also was significant ($P > t \leq .05$). Previous reports (2, 23) based on cytoplasmic lineages suggested a larger impact on fat and energy than on milk volume. Schutz et al. (22) also reported a larger effect on fat yield than on milk yield when adenine to guanine transition at nt169 was analyzed concurrently with other mtDNA D-loop sequence differences in the same lineages.

Lindberg (18) determined that sequence polymorphism at nt169 marked an evolutionarily important bifurcation in phylogenetic trees based on mtDNA sequence information. Such distinct lineages may have evolved while they were geographically separated. Furthermore, separation may have been accompanied by differences in artificial selection, potentially leading to significant effects on yield or composition that are similar to those observed in this study.

No known gene products are coded by the D loop of mtDNA, but the D loop is a site of important transcriptional and replicational controls. Differences in yield or composition associated with sequence polymorphism in that region of mtDNA may relate to control of mtDNA function or may serve as markers for important sequence variation elsewhere in the mtDNA genome, which is inherited in its entirety. Therefore, mtDNA D-loop polymorphism may become established in artificially

selected populations when affected traits depend on control of transcription and replication of mtDNA. Alternatively, sequence variants may by chance alone become fixed in subpopulations with differences in mtDNA gene-coding regions.

Maternal Lineage Clusters

Data for analysis of maternal lineage clusters were the same as for the previous analysis (Table 1). Assignment of lineages to cluster groups and number of cows in each cluster group are shown in Table 3. Lineages in a cluster group are separated by double spaces. With 24 cluster groups, only a single cluster (maternal lineages 14 and 18) had members with differing genotypes. All other clusters were identical for these 16 nt indicated in Table 3.

With 14 cluster groups, members of a single group obviously had more diverse genotypes,

TABLE 3. Assignment of 36 maternal lineages to 24 (clustering 1) or 14 (clustering 2) groups by clustering.[1]

Clustering 1			Clustering 2		
Maternal lineage	Number of cows	Binomial sequence codes[2]	Maternal lineage	Number of cows	Binomial sequence codes
20	47	0001000000000000	20	47	0001000000000000
29		0001000000000000	29		0001000000000000
3	**133**	**0011000000000000**	**2**	**203**	**0011000000000001**
32		**0011000000000000**	**3**		**0011000000000000**
33		**0011000000000000**	**32**		**0011000000000000**
41		**0011000000000000**	**33**		**0011000000000000**
51		**0011000000000000**	**37**		**0011000000000011**
61		**0011000000000000**	**40**		**0011000000000001**
77		**0011000000000000**	**41**		**0011000000000000**
2	45	0011000000000001	**51**		**0011000000000000**
40		0011000000000001	**58**		**0011000000000011**
17	**42**	**0010000000000000**	**61**		**0011000000000000**
52		**0010000000000000**	**77**		**0011000000000000**
37	25	0011000000000011	17	42	0010000000000000
58		0011000000000011	52		0010000000000000
14	**43**	**0111001000010000**	**13**	**99**	**0111000000000000**
18		**0111000000010000**	**14**		**0111001000010000**
10	30	1011000100001000	**18**		**0111000000010000**
55	**23**	**1011000100000000**	**66**		**0111000000001000**
13	9	0111000000000000	10	53	1011000100001000
66	**47**	**0111000000001000**	55		1011000100000000
8	15	0011100000000000	**8**	**45**	**0011100000000000**
67	**30**	**0011100000100000**	**67**		**0011100000100000**
5	12	0000110000000000	5	21	0000110000000000
57	**9**	**0001110000000000**	57		0001110000000000
39	19	0001000010000010	**39**	**41**	**0001000010000010**
74	**7**	**0001000010000000**	**69**		**0001000011000000**
45	9	0011001100000000	**74**		**0001000010000000**
60	**22**	**0011000101000000**	45	31	0011001100000000
69	15	0001000011000000	60		0011000101000000
4	**7**	**0010000010000000**	**4**	**7**	**0010000010000000**
75	45	0011000000000110	75	45	0011000000000110
71	**35**	**0011111000010000**	**71**	**35**	**0011111000010000**
22	23	0010000100000001	22	23	0010000100000001
35	**36**	**0000000000000000**	**35**	**36**	**0000000000000000**
46		**0000000000000000**	**46**		**0000000000000000**

[1]Boldfacing marks every other cluster group.

[2]Binomial sequence codes are for nucleotides 8, 106, 169, 216, 16022, 16049, 16057, 16058, 16074, 16085, 16111, 16113, 16141, 16230, 16231, and 16247, respectively, and 1s correspond to sites that are polymorphic with respect to the first published bovine mitochondrial DNA sequence (1).

TABLE 4. Tests of significance and ranges of estimates for effects of clusters for 36 maternal lineages on yield and composition traits.

Trait	Clustering 1[1]			Clustering 2[2]		
	F	$P > F$	Range of estimated effect	F	$P > F$	Range of estimated effect
ME[3] Milk, kg	.93	.549	1402	1.03	.412	1154
ME Fat, kg	1.15	.282	75	1.33	.189	71
Fat percentage	1.68	.023	.64	2.44	.002	.58
SNF, kg	.98	.490	135	1.08	.375	107
SNF, %	1.23	.208	.46	1.00	.444	.37
Energy concentration, kcal/kg	1.58	.041	81	2.19	.008	71
Lactation energy, Mcal	1.07	.375	1245	1.21	.267	1182

[1]24 cluster groups; numerator df = 23, and denominator df = 1737.

[2]14 cluster groups; numerator df = 13, and denominator df = 1747.

[3]Mature equivalent (twice daily milking, 305-d lactation).

but maternal lineages within a single cluster never differed at more than 3 base pair sites and that occurred in only one cluster (maternal lineages 13, 14, 18, and 66). Clustering proved to be useful in grouping maternal lineages with similar mtDNA genotypes.

The 24 groups in clustering 1 had from 17 records of 7 cows to 311 records of 133 cows. Clustering 1 averaged 30.3 cows and 75 records per group, and clustering 2 averaged 52 cows and 128.6 records per group. Clustering 2 (with 14 groups) had from 19 records of 7 cows to 509 records of 203 cows. Significance tests of the effect of cluster groups on yield and composition traits and ranges of cluster group estimates are in Table 4. Clustering 1 had a significant ($P \leq .05$) effect on fat percentage and energy concentration. An effect of further clustering is to increase the variability among groups. Combining maternal lineages into 14 groups increased F statistics for all traits, although clustering 2 had statistically significant effects on the same traits (Table 4). Ranges of estimates were larger for clustering 1, probably because of greater sampling variance associated with smaller groups in clustering 1.

The F statistics and probabilities were consistent with those previously reported for maternal lineage effects (23). Also, correlations between solutions for the 24 groups in clustering 1 and solutions for the original 36 maternal lineages of which the 24 groups were composed (23) were moderately high (.80 to .90). The correlations between solutions for the 14 groups of clustering 2 and solutions for the 36 maternal lineages were smaller (.57 to .74). That clustering 1 was in somewhat better agreement with maternal lineage results was not surprising because combining fewer lineages than with clustering 2 was more similar to that analysis.

Coefficients of determination were .806 and .802, respectively, for fat percentage and energy concentration for Model [2] with clustering 1. The coefficient of determination was only slightly less for each trait with clustering 2 than with clustering 1 and was .805 for fat percentage and .801 for net energy concentration. Analysis of residuals from the model with either clustering 1 or clustering 2 by using a model including only maternal lineage demonstrated the ability of clustered line-

TABLE 5. Overall means and standard deviations of yield and composition traits for 1407 records by 572 cows of 21 mitochondrial DNA genotypes.

Trait	$\bar{X}$	SD
ME[1] Milk, kg	7975	1760
ME Fat, kg	286	59
Fat percentage	3.65	.44
SNF, kg	736	164
SNF, %	9.23	.39
Energy concentration, kcal/kg	734	49
Lactation energy, Mcal	5830	1215

[1]Mature equivalent (twice daily milking, 305-d lactation).

TABLE 6. Tests of significance and ranges of estimates for effects of mitochondrial DNA genotypes on yield and composition traits.[1]

Trait	F	$P > F$	Range of estimated effect
ME[2] Milk, kg	.99	.465	1334
ME Fat, kg	.87	.628	51
Fat percentage	1.42	.105	1334.38
SNF, kg	.95	.528	125
SNF, %	1.01	.453	.36
Energy concentration, kcal/kg	1.46	.085	76
Lactation energy, Mcal	.90	.584	1016

[1]Numerator df = 20; denominator df = 1347.

[2]Mature equivalent (twice daily milking, 305-d lactation).

ages to account for the underlying maternal lineage effects. Maternal lineage effects on residuals could still have been appreciable if they were not accounted for by clusterings. Even with clustering of 36 lineages into 14 groups, effects of maternal lineages on residuals from Model [2] were negligible ($P \geq .99$). Maternal lineages accounted for more variation in the residuals from Model [2] with clustering 2 than with clustering 1. Again, this result is expected because clustering 1 is more analogous to maternal lineages as originally defined.

Clustering lineages into groups based on sequence differences in mtDNA D loops likely grouped maternal lineages that had similar mtDNA sequences but were defined as distinct lineages because of the limited time available for definition based on pedigrees. Because the entire mtDNA genome is transmitted to the cow's offspring, lineages that are similar for D-loop sequences are probably more nearly identical for gene-coding regions of mtDNA, too. Therefore, they should express any biological effects similarly.

Sequence Genotypes

The purpose of this analysis was to determine whether economic traits could be associated with lineages assigned to homogeneous groups by sequence and restriction enzyme analysis. Overall means and standard deviations of records by cows in the 29 maternal lineages assigned to 21 genotypes by Lindberg et al. (19) are in Table 5. Means and standard deviations are similar to those of the 36 lineages from the prior analyses. The 21 genotypes had from 17 records of 7 cows to 197 records of 70 cows. On average, genotypes had 67 records and 27.2 cows.

Table 6 has tests of significance of mtDNA genotypes on yield and composition traits. Genotype approached significance ($P \leq .085$) for energy concentration and fat percentage ($P \leq .105$) but was not significant at $P \leq .10$ for any other trait. The F statistics were as large or nearly as large as those from testing maternal lineages in a previous report by Schutz et al. (23), but associated probabilities were much larger because of fewer numerator and denominator degrees of freedom. Ranges of estimates tended to be smaller than for cluster analysis. For Model [3], coefficients of determination were .816 for fat percentage and .807 for net energy concentration. These values were somewhat greater than for either clustering in Model [2]. Comparisons must be made cautiously because data were not identical.

CONCLUSIONS

The intent of this work was not so much to determine optimal methods of grouping cows by mtDNA differences as to examine the effects on yield and composition of groups formed by previously proposed methods. The significance of a transition from adenine to guanine at nt169 in the mtDNA D loop on fat and energy produced by dairy cattle may have evolutionary implications, as proposed by Lindberg et al. (19). This present work also indicates the benefit of mixed models to associate potentially important mtDNA sequence polymorphisms with milk yield and composition traits. Maternal lineage groups defined by clustering based on D-loop sequence polymorphism or by mtDNA genotype identification had the greatest effect on fat percentage and energy concentration. The magnitude of these effects was similar to that of maternal lineage effects reported previously (23). Further research may determine mechanisms by which mtDNA sequence variation affects precise assignment of cows to maternal groups with similar mitochondrial genetic value. Optimal definition of mitochondrial lineages also will incorporate sequence variation in gene-coding regions of mtDNA, where mitochondrial effects more likely originate.

ACKNOWLEDGMENTS

Appreciation is expressed to the National Association of Animal Breeders and Eastern Artificial Insemination Cooperative for partial financial support. Funding was also provided by US-Israeli Binational Agricultural Research and Development Grant Number US-1519-88R.

REFERENCES

1 Anderson, S., M.H.L. de Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger, and I. G. Young. 1982. Complete sequence of bovine mitochondrial DNA. J. Mol. Biol. 156:683.

2 Bell, B. R., B. T. McDaniel, and O. W. Robison. 1985. Effects of cytoplasmic inheritance on production traits of dairy cattle. J. Dairy Sci. 68:2038.

3 Brown, D. R., C. M. Koehler, G. L. Lindberg, A. E. Freeman, J. E. Mayfield, A. M. Myers, M. M. Schutz, and D. C. Beitz. 1989. Molecular analysis of cytoplasmic genetic variation in Holstein cows. J. Anim. Sci. 67:1926.

4 Faust, M. A., O. W. Robison, and B. T. McDaniel. 1989. The effects of cytoplasm on reproduction and production in Holsteins. J. Dairy Sci. 72(Suppl. 1): 52.(Abstr.)

5 Faust, M. A., O. W. Robison, and B. T. McDaniel. 1990. Animal model estimates of cytoplasmic line constants for yield in Holsteins. J. Anim. Breed. Genet. 107:401.

6 Gyllensten, U., D. Wharton, A. Josefsson, and A. C. Wilson. 1991. Paternal inheritance of mitochondrial DNA in mice. Nature (Lond.) 352:255.

7 Gyllensten, U., D. Wharton, and A. C. Wilson. 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. J. Hered. 76:321.

8 Harville, D. A. 1979. Some useful representations for constrained mixed model estimation. J. Am. Stat. Assoc. 74:200.

9 Henderson, C. R. 1974. General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 57: 963.

10 Hoeh, W. R., K. H. Blakely, and W. M. Brown. 1991. Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial DNA. Science 251:1488.

11 Huizinga, H. A., S. Korver, B. T. McDaniel, and R. Politiek. 1986. Maternal effects due to cytoplasmic inheritance in dairy cattle. Influence on milk production and reproduction traits. Livest. Prod. Sci. 15:11.

12 Jacquard, A. 1974. The Genetic Structure of Populations. Springer-Verlag, Berlin, Germany.

13 Johnston, S. D. 1991. Sequence heterogeneity of bovine mitochondrial ribosomal RNA genes. M.S. Thesis, Iowa State Univ., Ames.

14 Kennedy, B. W. 1986. A further look at evidence for cytoplasmic inheritance of production traits in dairy cattle. J. Dairy Sci. 69:3100.

15 Koehler, C. M. 1989. Nucleotide variation in bovine mitochondrial DNA. M.S. Thesis, Iowa State Univ., Ames.

16 Koehler, C. M., G. L. Lindberg, A. E. Freeman, A. M. Myers, J. E. Mayfield, and D. C. Beitz. 1991. Complete replacement of a bovine mitochondrial genotype in a single generation. Genetics 129:247.

17 Laipis, P. J., M. J. Van de Walle, and W. W. Hauswirth. 1988. Unequal partitioning of bovine mitochondrial genotypes among siblings. Proc. Natl. Acad. Sci. USA 85:8107.

18 Lindberg, G. L. 1989. Sequence heterogeneity of bovine mitochondrial DNA. Ph.D. Diss., Iowa State Univ. University Microfilm Order No. DA-9014925, Ann Arbor, MI.

19 Lindberg, G. L., C. M. Koehler, S. D. Johnston, D. R. Brown, A. M. Myers, A. E. Freeman, and D. C. Beitz. 1989. High frequency of sequence variation in the displacement-loop region of bovine mitochondrial DNA. J. Dairy Sci. 72(Suppl. 1):53.(Abstr.)

20 Nibler, T., F. Reinhardt, and F. Pirchner. 1990. Einfluss Zytoplasmatischer Vererbung auf Milchleistung und Fruchtbarkeit bei Milchvieh. Zuechtungskunde 62:179.

21 Reed, P. D., and L. D. Van Vleck. 1987. Lack of evidence for cytoplasmic inheritance of production traits of dairy cattle. J. Dairy Sci. 70:837.

22 Schutz, M. M. 1991. Cytoplasmic and mitochondrial genetic effects on economic traits in dairy cattle. Ph.D. Diss., Iowa State Univ. University Microfilm Order No. 9207254. Ann Arbor, MI.

23 Schutz, M. M., A. E. Freeman, D. C. Beitz, and J. E. Mayfield. 1992. The importance of maternal lineage on milk yield traits of dairy cattle. J. Dairy Sci. 75: 1331.

24 Sneath, P.H.A., and R. R. Sokal. 1973. Numerical Taxonomy. W. H. Freeman and Co., San Francisco, CA.

25 Southwood, O. I., B. W. Kennedy, K. Meyer, and J. P. Gibson. 1989. Estimation of additive maternal and cytoplasmic genetic variances in animal models. J. Dairy Sci. 72:3006.

26 Tyrrell, H. F., and J. T. Reid. 1965. Prediction of the energy value of cow's milk. J. Dairy Sci. 48:1215.

27 Wales, T. P., Jr. 1885. Holstein Friesian Herd Book. Egbart, Fidlar, and Chambers, Davenport, IA.

28 Weigel, K. A., A. E. Freeman, M. E. Kehrli, Jr., M. J. Stear, and D. H. Kelley. 1990. Association of class I bovine lymphocyte antigen complex alleles with health and production traits in dairy cattle. J. Dairy Sci. 73:2538.