Association Between SSCP Haplotypes at the Bovine Growth Hormone Gene and Milk Protein Percentage

A. Lagziel, E. Lipkin and M. Soller

Department of Genetics, Silberman Life Sciences Institute, Hebrew University of Jerusalem, 91904 Jerusalem, Israel

Manuscript received April 18, 1995
Accepted for publication November 25, 1995

ABSTRACT

The bovine Growth Hormone gene (*bGH*) is an attractive candidate gene for milk production in cattle. Single-strand conformation polymorphisms at *bGH* were identified and used to define haplotype configurations at this gene in the Israeli Holstein dairy cattle population (*Bos taurus*) and in the parent animals of the International Bovine Reference Family Panel (a collection of *B. taurus* and *B. indicus* crosses). *B. taurus* and *B. indicus* haplotypes at the *bGH* gene differed qualitatively, confirming the previously proposed long evolutionary separation of these cattle subraces. Only a small number of *bGH* haplotypes were present in the Israel Holstein population. One of the haplotypes, apparently of *B. indicus* origin, was found to have a highly significant positive effect on milk protein percentage. This illustrates the utility of the haplotype approach for uncovering candidate gene involvement in quantitative genetic variation in agricultural populations. The strong effect of an indicine haplotype in a taurine background raises the possibility that indicine alleles at other candidate genes may comprise a genetic resource for improvement of taurine populations. It is proposed that haplotype analysis may be a useful adjunct to measures of genetic distance for evaluating rare breeds with respect to gene conservation.

IN human (STENGARD *et al.* 1995), agricultural (BOVENHUIS and WELLER 1994; ROTHSCHILD *et al.* 1994) and experimental populations (MACKAY and LANGLEY 1990), polymorphisms at candidate genes have been associated with quantitative genetic variation. LAI *et al.* (1994) showed that intragenic haplotypes were a powerful means of associating quantitative effects on bristle number with DNA level polymorphisms at wild-type alleles of the scabrous locus (a neurogenic locus at which major mutations cause bristle absence). This makes it of interest to explore a haplotype approach to the search for associations of polymorphic variation at candidate genes and quantitative variation in traits of economic importance in agricultural populations. In the present study, single-strand conformation polymorphisms (SSCPs) were used to define intragenic haplotypes at the the bovine Growth Hormone (*bGH*) gene in the Israeli Holstein dairy cattle population (*Bos taurus*) and in the parent animals of the International Bovine Reference Family Panel (a mixture of *B. taurus* and *B. indicus* crosses). *B. taurus* and *B. indicus* haplotypes at the *bGH* gene were found to differ markedly; only a small number of *bGH* haplotypes were present in the Israel Holstein population, and one of these haplotypes, apparently of *B. indicus* origin, had a highly significant positive effect on milk protein percentage.

MATERIALS AND METHODS

SSCPs at the *bGH* gene: The complete sequence of the *bGH* gene (from −649 bp upstream of the first exon to a point 412 bp downstream of the last exon) was obtained (GORDON *et al.* 1983), and primers were constructed (WOJCIECH and RHOADS 1989) to almost completely cover the gene sequence (Table 1). There are nine amplification fragments, ranging in size from 129 to 347 bp.

PCR amplification was carried out for each primer pair. The reaction mixture contained a total of 25 μl and included 100 ng of bovine genomic DNA, 0.5 μM of each primer, 200 μM of each dNTP, 0.1 unit of Taq DNA Polymerase (Appligene), 2.5 μl of 10 × PCR buffer (Appligene) and 0.1 μl ^{32}P-labeled dCTP (3000 mc/ml). All PCR reactions were run for 30 cycles.

For SSCP analysis, 3 μl of each amplification product were added to 9 μl of Stop solution (95% formamide, 20 mM EDTA and tracking dyes). The samples were denatured for 3 min in 90° and cooled on ice for 3 min. Each sample (4.5 μl) was loaded onto a Mutation Detection Enhancement Gel (MDE) (AT BioChem. Inc. Malvern, PA), according to manufacturer's instructions with added 6% glycerol. The gel was run at constant 7–8 W for 17 hr. After the run, the gel was removed from the apparatus, dried in vacuum for 20 min at 80°, and autoradiographed.

DNA samples: DNA was obtained from peripheral blood leukocytes of five sires in active service in Israeli Holstein A. I. Centers, from a semen sample of a sixth sire, and from daughters of these sires. Samples of blood totaling ~30 ml (with EDTA to final concentration of 0.5%) were centrifuged, and the serum was drawn off. Samples were stored at −70°. DNA extraction followed standard procedures (WYMAN and WHITE 1980). DNA samples were also obtained for 22 parent individuals of the International Texas/CSIRO/ILRAD bovine reference family panel (BARENDSE *et al.* 1994) and their progeny. These included pure and crossbred animals of *B. indicus*, *B. taurus* and mixed origin (Table 4). However, PCR amplification was successful for 21 of these animals only.

Notation: Individual amplification fragments are denoted by uppercase letters and alleles of individual fragments by italicized number, *e.g.*, fragment A, allele *1*. Genotypes are

Corresponding author: Morris Soller, Department of Genetics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel.
E-mail: soller@vms.huji.ac.il

TABLE 1

SSCP fragments at the bGH gene

Fragment	Type[a]	Primer pair	Start point[b]	Primer sequence	Length[c]	T[d]
A	U, E1, I1	Bov1	444	5′ TGG TGG CAG TGG AGA CGG GA	314	63
		Bov2	739	5′ GGA CAC GCG AAT GGA GGG GA		
B	I1	Bov7	773	5′ GGG GGG CCC TGC CGA TGG AT	186	63
		Bov8	940	5′ CCA GGA GGG CCG CGA CGG AG		
C	E2, I2	Bov9	1000	5′ GCC CTG CTC TGC CTG CCC TG	282	68
		Bov10	1263	5′ CCC CAC ACA CCC CCG TTT CT		
D	E3, I3	Bov7	1401	5′ ACA CCC AGG TTG CCT TCT GC	347	NA[e]
		Bov8[f]	1729	5′ CGA GGA GTA GGT CAG CAC CG		
E	E4, I4	Bov9[f]	1729	5′ GCT CCT CAT CCA GTC GTG GC	317	65
		Bov10	2027	5′ GGA GAA GGG CGA GGA AGG AG		
F	I4, E5	Bov11	2089	5′ GCT GCT CCT GAG GGC CCT TC	146	63
		Bov12	2216	5′ TTC TTG AGC AGC GCG TCG TC		
G	E5, D	Bov13	2252	5′ TGC TTC CGG AAG GAC CTG CA	229	60
		Bov14	2462	5′ CCA CCC CAC CCC CCA GAA TA		
H	D	Bov15	2473	5′ GTG GGG TGG GGC AGG ACA GC	173	65
		Bov16	2627	5′ ACC AGG GGC GTG GAC AGG GT		
I	D	Bov17	2632	5′ GTC CAC GCC CCT GGT TCT TA	129	58
		Bov18	2742	5′ TGG TTT GGT GGG CTG ATG AG		

[a] Fragment types: U, upstream; E1, E2, . . ., exon 1, exon 2, . . .; I1, I2, . . ., intron 1, intron 2, . . .; D, downstream.

[b] According to GORDON *et al.* (1983).

[c] Length of product in bp.

[d] Optimum annealing temperature in °.

[e] NA, not available. A smear was obtained on amplification, at all conditions employed.

[f] Bov 8 and Bov 9 are complementary sequences.

denoted by a series of italicized allele numbers, in which place number represents fragment type, *e.g.*, genotype *1/2, 1/1, 1/1, 1/3, 1/1, 1/2, 1/2, 1/2* denotes fragment A, heterozygous for alleles *1* and *2;* fragment B, homozygous for allele *1;* fragment C, homozygous for allele *1,* etc. Note that fragment D is not included, as this fragment did not migrate as a discrete band in the gel.

Haplotypes are denoted by italicized uppercase letters, *e.g,* haplotype *A*. Haplotype composition is denoted by a series of italicized numbers, for which place represents fragment type, as above (*e.g.*, haplotype composition *21131222* represents the haplotype having allele *2* at fragment A, allele *1* at fragment B, allele *1* at fragment C, etc.). Genotype composition, in terms of haplotypes, is denoted by a pair of italicized uppercase letters, *e.g.*, genotype *AE*.

Inferring haplotypes: Haplotypes were inferred as follows:

1. From homozygous individuals. For example, from an animal showing genotype *1/1, 1/1, 1/1, 3/3, 1/1, 1/1, 1/1, 1/1,* the presence of haplotype *11131111* was inferred.

2. From heterozygous individuals showing only a single heterozygous site. For example, from an animal having the genotype *1/1, 1/1, 1/1, 1/1, 1/2, 1/1, 1/1,* haplotypes *11111111* and *11112111* were inferred.

3. From consistent joint inheritance of a set of alleles in a series of daughters of the same sire. For example, a Holstein sire having genotype *1/2, 1/1, 1/1, 1/3, 1/1, 1/2, 1/2, 1/2* had 14 daughters, of which nine carried alleles *2, 1, 1, 3, 1, 2, 2, 2,* and none could be interpreted as having received from the sire any one or more of allele *2* at fragment A, allele *3* at fragment E, or allele *2* at fragments G, H, I, without all of the others. From this, haplotype *21131222* was inferred.

4. In some instances, a genotype could be decomposed into two well documented haplotypes, into one well documented haplotype and one novel haplotype, or into one or more pairs of novel haplotypes. In this case, the more conservative decomposition was postulated. For example, a number of animals had the genotype *1/2, 1/1, 1/1, 1/3, 1/1, 1/2, 1/2, 1/2.* Since haplotypes *11111111* and *2113122* are well documented, genotype composition 11111111/2113122 was postulated. Similarly, one animal had genotype *3/2, 1/1, 1/1, 3/3, 1/1, 1/2, 1/2, 1/2.* Since haplotype *21131222* was well documented, its presence was postulated, and haplotype *31131111* was inferred by subtraction. In both instances, other genotype decompositions were theoretically possible but required inferring two novel haplotypes.

Haplotypes of homozygous reference panel animals were confirmed by showing that all progeny examined (four to 10) carried at least one parental allele at the diagnostic fragments defining the haplotype.

Effects of *bGH* haplotypes on milk protein percentage in the Israel Holstein population: Because of the widespread use of artificial insemination in dairy cattle reproduction, the bulk of a dairy cattle population consists of a relatively small number of large half-sib families, each produced by a single progeny-tested sire. Since each sire will have transmitted his own candidate gene haplotypes to his daughters, this produces a population-wide confounding of candidate gene haplotype and overall sire breeding value. To avoid this confounding, *bGH* haplotype effects in the Israel Holstein dairy cattle population were investigated on a within-sire basis. For this purpose, recourse was had to a DNA sample set consisting of the highest and lowest daughters, with respect to milk protein percentage, of each of the six sires examined in this study. For each sire, 80 high and 80 low daughters were selected from among 1000 or more milk-recorded daughters (960 selected daughters in all). Blood samples were obtained from each of the selected daughters, and DNA extracted as above.

Five of the six sires were heterozygous for SSCP *bGH* haplo-

types. In each case, the two haplotypes of a heterozygous sire could be distinguished by genotyping a single fragment only. Hence it was possible to identify daughters receiving one or other of the haplotypes from their sire, by genotyping the daughters with respect to the diagnostic fragment only. Inasmuch as the two daughter groups of each sire are drawn from the same sire and from the same dam population, confounding effects of population structure are eliminated, and the two daughter groups are expected to differ systematically only at the chromosomal region surrounding the candidate gene.

Because dams were not genotyped, when a daughter had the same heterozygous genotype as her sire, it was not possible to determine unequivocally which of the sire alleles the daughter received from her sire. In this case, the procedure followed depended on the relative frequency in the population of the particular haplotypes involved. To wit: two of the sires were heterozygous for a pair of rather frequent haplotypes. In this case, only homozygous daughters were assigned to one or the other of the sire haplotype groups. Of the remaining three heterozygous sires, one was heterozygous for a very rare haplotype (none of the daughters of the sire were homozygous for the rare allele, nor did the haplotype turn up in heterozygous state among daughters of two other sires in which it could have been distinguished). For this sire, when a daughter had the same genotype as her sire, she was assumed to have received the rare-sire haplotype. Two other sires were heterozygous for a relatively rare haplotype (four of a total of 180 daughters of these two sires were homozygous for the rare haplotype, giving it a population frequency of ~0.05). For these sires, in an equivalent proportion (0.05) of sire-type heterozygotes, the rare haplotype was assumed to have originated from the maternal parent. To avoid possible bias, the number of daughters in the various sire groups were adjusted by proportionally subtracting this proportion of daughters from the groups assumed to have received the rare sire alelle, and adding them to the groups assumed to have received the common sire allele.

Statistical analyses: A chi-square contingency test (WALPOLE and MYERS 1972) was carried out to test for differential frequency of sire *bGH* haplotypes among the high protein- and low protein-percentage daughter groups, pooling daughters over sires heterozygous for the same SSCP haplotypes. Following DARVASI and SOLLER (1992) the actual genetic effect of SSCP haplotype on milk protein percentage, δ, was calculated as follows:

$$\delta = D_I/2\gamma_P,$$

where δ is the actual genetic effect in the unselected population, D_I is the difference in mean trait value of daughters receiving the alternative sire haplotypes, taken over all daughters in both of the selected daughter groups, $\gamma = 1 + Z_{1-P/2}i_{P/2}$, P is the total proportion of the sample analyzed ($P/2$ at each tail), and $i = x_t/t$ = the mean of an upper tail of a standard normal distribution, t is the area under the tail, and x_t is the ordinate of the standard normal distribution at the point Z_t.

RESULTS

SSCP polymorphisms: Considering the six Israeli Holstein sires and their daughters, and the 21 informative animals of the International Bovine Reference Panel, the following SSCP polymorphisms were found (Table 2). Fragment C was monomorphic. Fragment D gave a smear on electrophoresis, and individuals bands could not be distinguished. Two alleles each were present for fragments B, F, G, and I; three alleles were present for fragments E and H. Most alleles were found in at least two individuals, except for allele *2* of fragment B and allele *3* of fragment H that were each found in one reference panel individual only. However, allele *2* of fragment B was also observed in a beef animal of unknown taurine origin. All rare alleles were genotyped a number of times in reference panel individuals and were confirmed by genotyping their offspring (data not shown). Figure 1 shows SSCP polymorphisms at consecutive fragments of the *bGH* gene, as uncovered among the six Israeli Holstein sires examined.

TABLE 2

Inferred haplotypes at the bGH gene, based on SSCP fragments A–C, E–I

Haplotype designation	Composition ABCEFGHI[a]	Method[b]	Source animals[c]
A	*11111111*	1	H24, H25
		2	H21, H26
B	*11121111*	1	H6, H10
		2	H1, H9
C	*11112111*	2	H21, H14
D	*11122111*	2	H7, H14
E	*21131222*	2	R7
		3	H18 to H32
F	*21121222*	1	R4, R12
		2	R7
G	*21131232*	4	R11
H	*31131222*	4	R17
I	*21131122*	4	R5
J	*11131111*	1	R1, R2
K	*21131111*	1	R3, R15
L	*31131111*	4	R8
M	*12111111*	4	R20

[a] Numbers represent alleles at SSCP fragments A, B, C, E, F, G, H, I, respectively.

[b] Method used for inferring haplotype composition. See text for details.

[c] Exemplar animals from which the given haplotype was inferred. See Table 3 for haplotype details.

In genotyping the animals for fragment A, a distinct and consistent polymorphism was observed in some runs, but not in others. Subdividing fragment A into two smaller subfragments did not improve consistency (data not shown). In the informative runs, two relatively common alleles were found, while a third allele was found in two individuals only.

SSCP haplotypes at the bGH gene: Table 2 shows various haplotypes present at the *bGH* gene, as inferred from the genotypes of the Holstein sires and their daughters, and from the genotypes of the reference panel animals. Three distinct groups of haplotypes are observed. Group I comprises haplotypes *A, B, C, D,* and *M.* Haplotypes *B, C,* and *M* differ from haplotype *A* by a single mutation, and haplotype *D* can be obtained by recombination between haplotypes *B* and *C.* Group II comprises haplotypes *E, F, G, H, I.* All of these can be derived by a single step from haplotype *E* but differ at four or five

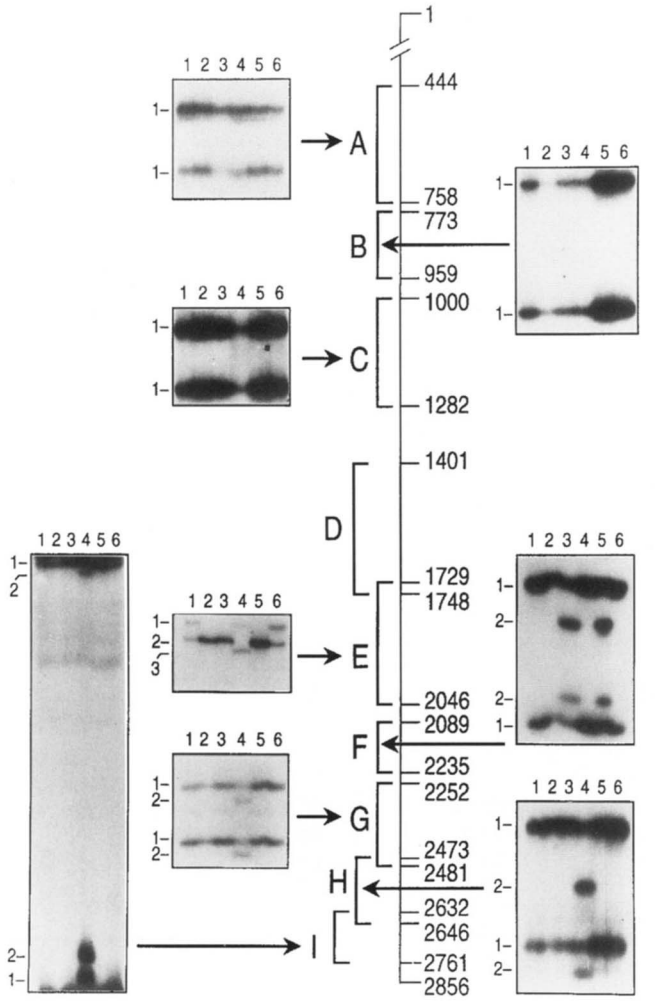

FIGURE 1.—SSCP polymorphisms at the *bGH*. The vertical line represents the nucleotide sequence of *bGH* (GORDON *et al.* 1983). PCR amplification fragments are indicated by letters on the left side of the vertical line (A–I) and by nucleotide numbers on the right side (1–2856). SSCP bands for each fragment are shown in boxes with arrows leading to the corresponding PCR fragment. Numbers at the left of the boxes designate SSCP alleles. Lane numbers correspond to sire identification numbers for the six sires analyzed.

points from any of the haplotypes in group I. Group III comprises haplotypes *J, K, L*. These share alleles *2* and *3* at fragment A and allele *3* at fragment E, with group II, and allele *1* at fragments A, G, H and I, with group I.

Table 3 shows the inferred haplotypes of the Israeli Holstein and reference panel individuals. With the exception of sire 4 and his daughters (animals H18–H32), the Israeli Holstein animals showed only group I haplotypes (*A, B, C, D*); sire 4 and daughters also showed haplotype *E* of group II. Among the reference panel animals, those of pure indicine origin (R1, R2, R3, R6, and R10) showed group II or group III haplotypes only; each of two animals (R11 and R21) that were first generation progeny of a cross between indicine and taurine parents, showed one group I and one group II haplotype, while animals that were F_2 or later generation progeny of indicine and taurine parents (R7, R8, R9, and R12–R19) showed haplotypes of all three groups. The two N'Dama animals, generally classed as taurine types, can be interpreted as showing either group II and III haplotypes only, or a mixture of all three groups.

Effect of *bGH* haplotypes in milk protein percentage in Israeli Holstein cattle: Table 4 shows the distribution of haplotypes of the daughters of Israeli Holstein sires 1, 3, 4, 5, and 6 (sire 2 was homozygous) among the high and low milk percentage daughter groups. Haplotypes *A* and *B* are found in almost identical frequencies among the high and low daughters of sires 1 and 6. Haplotype *B* shows a small excess relative to haplotype *D*, among the high daughters of sire 3, but the effect is not statistically significant and is not found among the daughters of sire 5. Among the high and low daughters of sire 4, however, a statistically significant excess of haplotype *A* was found among the daughters in the low protein-percentage group, while haplotype *E* was in excess among the daughters in the high protein-percentage group ($P < 0.005$).

Among the pooled "high and low" daughters of sire 4, milk protein percentage of daughters receiving haplotype *E* from their sire was 0.078% greater than that of the daughters receiving haplotype *A*. From this, the effect of haplotype *E* in heterozygous state on milk protein percentage in the population as a whole was estimated as described in MATERIALS AND METHODS, setting $P = 0.16$, and was found to equal 0.010%.

The absence of a difference in the frequency of haplotypes *A* and *B* in the high and low daughters of sires 1 and 6 implies that these two haplotypes have equivalent effects on milk protein percentage. Similarly, the absence of an appreciable difference in frequency of haplotypes *B* and *D* in the high and low daughters of sires 3 and 5 implies that these two haplotypes also have roughly equivalent effects on milk protein percentage. From this it follows that the positive effect of haplotype *E* on milk protein percentage is expressed more or less equally, relative to haplotypes *A, B,* and *D*.

DISCUSSION

Polymorphisms and haplotypes at the *bGH* gene: The fragments examined in this study cover the total extent of the bGH gene, including some of its upstream and downstream sequences, with the exception of some 300 bp included in fragment *D*. Thus, the haplotypes defined in this study can be taken to represent all or most of the DNA sequence alleles present at the *bGH* gene in the surveyed population (HAYASHI 1992; SHEFFIELD *et al.* 1993). Among the animals surveyed, seven polymorphic fragments were identified. Assuming only one polymorphic site per fragment, this implies a polymorphic site every 200 bp or so along the gene.

A total of 14 different haplotypes were inferred. Of

TABLE 3

Genotypes at the bovine growth hormone gene for Israeli-Holstein and International Reference panel animals

Des[a]	Id[b]	Genotype[c]	Type[d]	Des[a]	Id[b]	Genotype[c]	Type[d]
H1	S1	*AB*	H	R1	0101	*JJ*	**G × B**
H2	S1-1	*AB*	H	R2	1202	*JJ*	**I × GB**
H3	S1-2	*AA*	H	R3	1102	*KK*	**I × B**
H4	S1-3	*AD* or *BC*	H	R4	ND7	*FF*	N
H5	S1-4	*AB*	H	R5	ND8	*BE* or *FJ*	N
H6	S2	*BB*	H	R6	2102	*BE* or *FJ*	**B**
H7	S2-1	*BD*	H	R7	5191	*EF*	*Br*
H8	S2-2	*BD*	H	R8	S2308	*EL*[e]	*Br*
H9	S2-3	*AB*	H	R9	P0981	*AE*[e]	*Br*
H10	S2-4	*BB*	H	R10	1419	*JJ*	**Bo**
H11	S2-5	*AB*	H	R11	0802	*AG*[e]	Si × **B**
H12	S3	*BD*	H	R12	1502	*FF*	**S** × F (F_2)
H13	S3-1	*AD* or *BC*	H	R13	2002	*BE* or *FJ*	**S** × F (F_2)
H14	S3-2	*CD*	H	R14	0502	*DE*[e]	**S** × F (F_2)
H15	S3-3	*AD* or *BC*	H	R15	1401	*KK*	**S** × F (F_2)
H16	S3-4	*AD* or *BC*	H	R16	1402	*KK*	**S** × F (F_2)
H17	S3-5	*AD* or *BC*	H	R17	2002	*DH*[e]	**S** × F (F_2)
H18	S4	*AE*	H	R18	1702	*BE* or *FJ*	**S** × F (F_2)
H19	S4-1	*AE*	H	R19	0702	*BE* or *FJ*	**S** × F (F_2)
H20	S4-2	*AE*	H	R20	0202	*EM*[e]	**B** × HS
H21	S4-3	*AC*	H	R21	0201	*DE*[e]	Si × **B**
H22	S4-4	*AE*	H				
H23	S4-5	*AE*	H				
H24	S4-6	*AA*	H				
H25	S4-7	*AA*	H	H30	S4-12	*BE*	H
H26	S4-8	*AB*	H	H31	S4-13	*DE*	H
H27	S4-9	*AE*	H	H32	S4-14	*BE*	H
H28	S4-10	*AE*	H	H33	S5	*BD*	H
H29	S4-11	*AA*	H	H34	S6	*AB*	H

[a] Des, designation; H, Israel-Holstein animals; R, International Reference panel animals.

[b] Id, identification number. S1, S2, etc., Israel Holstein sires; S1-1, daughter no. 1 of Sire S1, etc.

[c] *AD* or *BC*, either of the pair of haplotypes are possible; *BE* or *FJ*, either pair of haplotypes is possible, assuming previously documented haplotypes only.

[d] Type of animal: H, Israel-Holstein; B, Brahma; Br, Brangus; Bo, Boran; F, Friesian; G, Gir; GB, GirxBrahma (F_1); I, Indus; HS, Hereford × Shorthorn (F_1); N, N'Dama; S, Sahiwal; Si, Simmenthal. Bold letters indicate animals of indicine origin; roman indicates animals of taurine origin; *italic*, animals of mixed taurine/indicine origin.

[e] Haplotypes inferred on the principle that a genotype composed of one or two previously documented haplotypes is more likely than a genotype composed of a pair of novel haplotypes.

these, only five were observed in 40 independent chromosomes of Israeli Holstein origin. This accords well with the proposition that founder, drift or selection effects have reduced the number of DNA sequence alleles present at any given locus in this closed population. If generally true, this implies that polymorphism at functional sites, when present, will be found embedded within unique DNA sequence haplotypes that are amenable to evaluation and population-wide selection (LANDE and THOMPSON 1990; SMITH and SMITH 1993; SOLLER 1994).

Taurine and indicine haplotypes at the *bGH* gene: Generally speaking, group I haplotypes were found only in the Israeli Holstein animals and in reference family animals that had at least one taurine ancestor, while group II and III haplotypes were found only among purebred indicine animals or among animals that had at least one indicine ancestor. These observations lead to the conclusion that group I haplotypes are characteristic of taurine breeds, while group II and III haplotypes are characteristic of indicine breeds. A major difference between taurine and indicine *bGH* haplotypes is consistent with results of a comparison of mitochondrial DNA in taurine and indicine races, which indicate a very long divergence of these types (LOFTUS *et al.* 1994).

There were two exceptions to the proposed differentiation of taurine and indicine haplotypes. Haplotype *E*, belonging to group II was found among some of the Israeli Holstein animals, and the two animals in the reference panel, belonging to the taurine N'Dama breed, carried group II and group III haplotypes. Both of these exceptions may be due to admixture of indicine genetic material into basically taurine breeds. The Israeli Holstein dairy cattle population was produced by continued topcrossing of the local "Baladi" and "Damascus" Landraces to Dutch Friesian and United States

TABLE 4

Association between SSCP haplotypes at the *bGH* gene and milk protein percentage[a]

Sire	Haplotype	High	Low
1[d]	*A*	10	13
	B	8	9
3[b]	*B*	35 (36.45)	39 (40.05)
	D	29 (27.55)	21 (19.95)
4[c]	*A*	27	54**
	E	39	11
5[b]	*B*	13 (13.75)	13 (13.75)
	D	15 (14.25)	15 (14.25)
6[d]	*A*	15	8
	B	12	12

** $P < 0.005$.

[a] Frequency (no.) of Israeli-Holstein haplotypes among daughters from the high milk protein percentage tail (high) and low milk protein percentage tail (low) of the total daughter distribution by fragment and sire. Sire 2 was not included as he was homozygous for haplotype *B*.

[b] Not in parentheses: heterozygous daughters were assumed to have received haplotype *D* from their sire. In parentheses: heterozygous daughters were assumed to have received haplotype *D* from their dam, in proportion to the estimated overall frequency (0.05) of this haplotype in the population.

[c] Heterozygous daughters were assumed to have received haplotype *E* from their sire.

[d] Heterozygous daughters were not included.

and Canadian Holstein bulls, with some importation of purebred Dutch, U.S. and Canadian cows. Haplotype *E* was found only in the single sire studied here (sire 4), the grandson of an elite Holstein sire of U.S. origin and a local high-grade dam. Among the daughters examined in the context of the milk protein-percentage study, homozygotes for allele *3* at fragment E, diagnostic for this haplotype, were not found among the 131 daughters of sire 4, nor were heterozygotes for this haplotype found among 87 daughters of sires 3 and 5. This shows that haplotype *E* is indeed exceedingly rare in the Israeli Holstein population. Thus, we believe that it is appropriate to consider haplotype *E* as a residue of the original Landrace ancestors specific to the Israel Holstein population, rather than as a normal constituent of the Holstein gene pool. Similarly, the exceptional status of the two N'Dama animals is in accord with measures of genetic distance between the N'Dama and European cattle based on microsatellite allele frequencies (MACHUGH *et al.* 1994) and may reflect the very long period of physical contiguity of the N'Dama and the West African Zebu, which is found adjacent to the main N'Dama rearing areas.

***bGH* haplotypes and milk protein percentage:** Haplotype *E* was associated with high milk protein percentage relative to haplotypes *A*, *B*, and *D*. This can be attributed to functional site variation at *bGH* itself or at a locus in tight linkage disequilibrium with it, or to looser linkage between the *bGH* locus and a linked locus affecting milk protein percentage. The powerful effects of *bGH* on milk production (COLLIER *et al.* 1984) support the possibility of a direct effect of functional site polymorphism at the *bGH* locus on milk protein percentage. However, since the exact extent of the indicine-derived chromosomal region surrounding haplotype *E* is not known, it is certainly possible that linked loci other than *bGH* are involved. Indeed, if the introgressed region is large enough, general heterotic effects may be operative.

It will be possible to partially resolve this question by using additional polymorphic markers in the *bGH* region to determine the extent of haplotype *E;* by using statistical methods for determining quantitative trait loci (QTL) location with respect to a single locus on a larger sample of daughters of sire 4 (BOVENHUIS and WELLER 1994); and by considering the main effect and interaction components in a hierarchal analysis of variance involving additional sires heterozygous at the *bGH* locus. A direct effect of the *bGH* locus or of a tightly linked locus in strong linkage disequilibrium with it would be supported by mapping of the QTL close to or at the *bGH* markers and by absence of an interaction effect for markers nested within sires.

The differential expression of the indicine haplotype *E* on milk protein percentage, as compared to the taurine haplotypes, hints at the possibility that *B. indicus* may comprise a valuable genetic resource for improvement of the taurine races. In this event, a systematic exploration of indicine candidate gene alleles in a generally taurine genetic background may uncover other instances of agriculturally important effects. Unless *bGH* is exceptional, the results of this study imply that in other instances it may be possible to define specifically indicine or taurine haplotypes at candidate genes. This would permit comparison of effects of taurine and indicine alleles on trait of economic importance, without the necessity for planned matings and pedigreed offspring populations. Such "DNA sequence allele" comparisons could be carried out in long standing "creole" populations, such as the Brazilian Criollo breed, produced in the past by crossing taurine and indicine races; in taurine populations that appear to carry some proportion of indicine ancestry, such as the Israeli Holstein or N'Dama; or in relatively recent synthetic breeds, such as the Texan Brangus breed, produced by crossing indicine Brahman and taurine Angus beef cattle, and the Brazilian Girolanda breed, initiated some 15 years ago as a cross between indicine Gir and taurine Holstein dairy cattle.

Genetic distances and haplotype composition: Because of the outbreeding nature of animal species, even a relatively rare favorable allele present in a major commercial breed can be expected to eventually reach high frequency as a result of selection. For this reason, resource populations are of interest in animal agriculture, primarily to the extent that they contain unique functional alleles that are not present in the main commer-

cial populations. Clearly, the greater the genetic distance between resource and commercial populations, the greater the likelihood that the resource population carries unique functional alleles affecting expression of some given trait. But at what level of genetic distance do unique functional alleles begin to appear? And, how does the frequency of unique functional alleles increase with increase in genetic distance? Because the presence of unique *sequence* alleles is a necessary condition for the presence of unique *functional* alleles, comparison of gene haplotypes (representing sequence alleles) in resource and commercial populations may provide a useful approach to these questions.

This research was supported by the joint biotechnology program of the Israel Ministry of Science and Technology and the German Bundesministerium fur Biotechnologie, and by the the Israel Dairy Council. We thank the referees for their thoughtful comments.

LITERATURE CITED

BARENDSE, W., S. M. ARMITAGE, L. KOSSAREK, B. W. KIRKPATRICK, A. M. RYAN *et al.*, 1994 A preliminary map of the bovine genome. Nature Genet. **6:** 227–235.

BOVENHUIS, H., and J. I. WELLER, 1994 Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. Genetics **136:** 267–275.

COLLIER, R. J., J. P. MCNAMARA, D. R. WALLACE and M. J. DEHOFF, 1984 A review of endocrine regulation of metabolism during lactation. J. Anim. Sci. **59:** 498–510.

DARVASI, A., and M. SOLLER, 1992 Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genetics **85:** 353–359.

GORDON, D. F., D. P. QUICK, C. R. ERWIN, J. E. DONELSON and R.A. MAURER, 1983 Nucleotide sequence of the bovine growth hormone chromosomal gene. Mol. Cell. Endoc. **22:** 81–95.

HAYASHI, K., 1985 PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. PCR Meth. Appl. **1:** 34–38.

LAI, C., R. F. LYMAN, A. D. LONG, C. H. LANGLEY and T. F. C. MACKAY, 1994 Naturally occurring variation in bristle number and DNA polymorphism at the *scabrous* locus of *Drosophila melanogaster.* Science **266:** 1697–1702.

LANDE, R., and R. THOMPSON, 1990 Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

LOFTUS, R. T., D. E. MACHUGH, D. G. BRADLEY, P. M. SHARP and P. CUNNINGHAM, 1994 Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. USA **91:** 2757–2761.

MACKAY, T. F. C., and C.H. LANGLEY, 1990 Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster.* Nature **348:** 64–66.

MACHUGH, D. E., R. T. LOFTUS, D. G. BRADLEY, P. M. SHARP and P. CUNNINGHAM, 1994 Microsatellite DNA variation within and among European cattle breeds. Proc. R. Soc. London B **256:** 25–31.

ROTHSCHILD, M. F., C. JACOBSON, D. A. VASKE, C. K. TUGGLE, T. H. SHORT *et al.*, 1994 A major gene for litter size in pigs. 5th World Congress Genet. Appl. Livestock Prod. **21:** 225–228.

SHEFFIELD, D. V., S. J. BECK, E. A. KWITECK, W. D. SANDSTORM and M. E. STONE, 1993 The sensitivity of single strand conformation polymorphism analysis for the detection of single base substitutions. Genomics **16:** 325–332.

SMITH C., and D. B. SMITH, 1993 The need for close linkages in marker-assisted selection for economic merit in livestock. Anim. Breed. Abst. **61:** 197–203.

STEMGARD, J. H., K. E. ZERBA, J. PEKKANEN, C. EHNHOLM, A. NISSENEN *et al.*, 1995 Apolipoprotein E polymorphism predicts death from coronary heart disease in a longitudinal study of elderly Finnish men. Circulation **91:** 265–269.

SOLLER, M., 1994 Marker-assisted selection, an overview. Anim. Biotech. **5:** 193–208.

WALPOLE, R. E., and R. H. MYERS, 1972 *Probability and Statistics for Engineers and Scientists. Ed. 2.* Macmillan Publishing Co., Inc., New York.

WOJCIECH, R., and R. E. RHOADS, 1989 A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. Nucleic Acids Res. **17:** 8543–8551.

WYMAN, A. R., and R. WHITE, 1980 A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. USA **77:** 6754–6758.

Communicating editor: T. F. C. MACKAY