Theor Appl Genet (1987) 73:419–427

Biochemical polymorphism in relation to performance in horses

L. Andersson, Th. Arnason and K. Sandberg

Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, S-750 07 Uppsala, Sweden

Accepted August 14, 1986
Communicated by R. Riley

Summary. Investigations on relationships between biochemical polymorphism and variation in quantitative traits are of interest from the perspectives of both theoretical quantitative genetics and practical animal breeding. This subject was studied by using racing performance records of more than 25,000 horses of the Swedish Trotter breed born in the period 1970–1979. For all horses data on six blood group and nine electrophoretic loci were available. Two different performance traits were investigated. A racing performance index value was calculated for all individuals which had started in at least five races. Horses which had not started at all or less than five times were pooled in an unstarted class and the proportion of started horses was analysed as an all-or-none trait. The relationships between the marker genes and these two performance traits were analysed statistically by using linear models. Analysis within sires revealed a very highly significant association between variation at the serum esterase locus (*Es*) and the proportion of started horses. In addition, four weakly significant associations were found. A striking feature of the highly significant association involving the esterase locus was that the effect of different alleles showed a good fit to an additive genetic model as the value of each heterozygous type was intermediate to the two corresponding homozygotes. In addition to the association tests, the possibility of genetic linkage between marker genes and genes affecting performance was tested as well as the influence on performance of heterozygosity at marker loci. No significant relationships were revealed in these latter tests.

Key words: Horse – Racing performance – Biochemical polymorphism

Introduction

In farm animals, much interest has focused on the possibility to detect relationships between genetic markers showing simple Mendelian inheritance and traits of economic importance. One impetus for such studies is of course the hope to find markers which might be used for selection of breeding animals. These investigations have also a theoretical interest as they may increase our knowledge on the genes which determine important traits in farm animals. Most of these traits show a continuous variation and are, in quantitative genetics theory, assumed to be determined by a large number of genes, each with a small effect. The knowledge on the underlying genetic basis for quantitative variation is still very limited, however.

There is a considerable number of reports on relationships between marker genes and quantitative traits in farm animals. One problem in the evaluation of these investigations is that the statistical power in most studies has been weak because of the small sample sizes used. Furthermore, it is apparent that some authors have not been aware of the risk of obtaining false associations if family relationships in the material are not taken into account. Another problem is the occurrence of chance associations when multiple tests are carried out. It is therefore not surprising that these studies have given contradictory results. This inconsistency may, though, in part have a genetic explanation. Some of the associations reported may have been caused by linkage disequilibrium between marker genes and genes influencing a quantitative trait. In such cases the sign and the magnitude of a given association may differ between populations. However, some of the more extensive studies have indicated that relationships between marker genes and production traits can be found, but that they in general have little predictive value in practical animal breeding (Neimann-Sørensen and Robertson 1961; Jensen et al. 1968; Kennedy et al. 1973).

In farm animals there are notably two examples of relationships between genetic markers and economically important traits that have proved to be of great significance; the

strong association between one haplotype (B^{21}) in the chicken major histocompatibility complex (MHC) and resistance to Marek's disease (Hansen et al. 1967; Briles et al. 1977) and the linkage between the halothane locus, which influences productive as well as reproductive traits in pigs, and no less than five biochemical loci (Andresen and Jensen 1977; Rasmusen 1981; Juneja et al. 1983). These examples encourage further search for relationships between genetic markers and economically important traits.

The purpose of the present study was to utilize a very extensive material to investigate whether any relationships can be found between biochemical polymorphisms and the most important trait in breeding of trotting horses i.e. the ability to compete on the racecourse. Racing performance of a horse is a quantitative trait of a different nature compared with the characters normally dealt with in animal breeding such as milk yield, growth rate and egg production. However, racing performance is also to a certain extent genetically determined. Heritability estimates in the range 0.20–0.40 have generally been reported for this trait (for review see Hintz 1980; Langlois 1982; Tolley et al. 1985). Similar heritability estimates have been obtained in studies on racing performance in the Swedish Trotter breed (Arnason et al. 1982). These previous analyses have constituted the basis for the statistical treatment of this trait in the present study.

Relationships between quantitative traits and biochemical marker loci can be tested for according to three different main principles, namely by association tests, linkage tests or tests for the influence of heterozygosity. An association between a marker locus and a quantitative trait simply implies that the mean of the trait differs between different genotypes at the marker locus. An association may either be due to pleiotropic effects of variation at the marker locus or to linkage disequilibrium between the marker locus and one or more genes influencing the quantitative trait in question. Association tests have been by far the most common way to test for relationships between marker loci and quantitative traits in farm animals.

A number of theoretical studies have dealt with the possibility to detect genetic linkage between marker loci and loci affecting quantitative traits (Thoday 1961; Jayakar 1970; Soller 1974; Geldermann 1975, 1976; Smith 1975; Soller et al. 1976; Soller and Genizi 1978; Elston 1979; Soller and Beckmann 1982, 1983; Beckmann and Soller 1983; Cockerham and Weir 1983). The most favourable condition for detecting linkage by this approach is when widely divergent strains are available. Such linkages have also been detected for instance in *Drosophila* (Thoday 1979; and references therein), in crosses between inbred strains of mice (Kluge and Geldermann 1982) and in interspecific backcrosses of tomato (Tanksley et al. 1982). The power for detecting linkage between marker loci and loci affecting quantitative traits is much reduced for segregating populations (Soller 1974; Soller and Genizi 1978); very large numbers of offspring are needed if the effect on the quantitative trait is not very large. In farm animals, such studies have previously been carried out on various production traits in cattle (Rendel 1959; Neimann-Sørensen and Robertson 1961; Zwiauer 1980). The lack of evidence for genetic linkage was not surprising as the amount of data available in those studies was quite limited. However, in a recent study on cattle by Geldermann et al. (1985) offspring in three large sire families were divided according to which marker allele they had received from their heterozygous sire; significant effects on milk production were reported in particular for some chromosome segments marked by polymorphic milk proteins.

A third way of testing for a relationship between the variation at biochemical loci and the variation in a quantitative trait is to compare the mean of the trait among groups of individuals classified with respect to their heterozygosity i.e. the number of biochemical loci heterozygous.

A preliminary report on the present study was given at the 19th Conference of the International Society for Animal Blood Group Research (Andersson et al. 1985).

Materials and methods

Animals

The material used in this study has accumulated in our laboratory during more than ten years of routine blood typing service for parentage control in the Swedish Trotter breed. All horses registered in the period 1970–1979, altogether 27,101 individuals, were available for study. The material has previously been used for extensive linkage studies (Andersson and Sandberg 1982, 1984; Andersson 1983; Sandberg and Andersson 1984).

Genetic markers

Data were available on the six equine blood group systems A, C, D, K, P and Q and the nine electrophoretic systems serum albumin (Al), red cell acid phosphatase (AP), carbonic anhydrase (CA), serum carboxylesterase (Es), haemoglobin α-chain (Hb), 6-phosphogluconate dehydrogenase (PGD), phosphoglucomutase (PGM), phosphohexose isomerase (PHI) and transferrin (Tf). The *K* and *PGD* loci both belong to equine linkage group (LG) I and the male recombination distance between these loci has been estimated to about 26 centimorgan. The *Al* and *Es* loci are also loosely linked, male recombination distance about 37 centimorgan, and belong to LG II. The *A* and *PHI* loci have been assigned to LG III and IV, respectively. The remaining nine loci appear to segregate independently of each other and of the four known linkage groups (see Weitkamp and Andersson 1984). The techniques for the analysis of these genetic markers are given by Sandberg and Andersson (1984).

Performance traits

Racing performance records of the horses were made available by the Swedish Trotting Association. For all individuals which had started in at least five races, an index value was calculated. The index value representing the deviation of an individual's performance from the mean of its sex and year class was based on best racing time, total earnings, earnings per start and percent races placed among the three first horses in the race. All the variables were transformed for improved distributional properties according to Arnason et al. (1982), where also the

construction of the index was described. All sire families having at least ten offspring with an index were included. Altogether 11,712 offspring of 247 sires were involved in the analysis; the average size of the sire progeny groups was thus 47 (range 10-370). The great majority of offspring had been blood typed at an age of 4–18 months whereas they begin to compete on the race-course at an age of two to four years. The index values were scaled such that the mean became 100.0. The standard deviation of index equals 3.26 in this material.

The fact that only about 50% of all horses involved had started in at least five races suggested that it may be worthwhile to also utilize the information on whether a horse had started or not. The ability to become started has not previously been examined by quantitative genetics analyses in the Swedish Trotter breed. It is reasonable to assume, however, that there is genetic variation for this trait also since the variation between paternal half-sib groups is substantial. It is likely that the genetic correlation between this trait and the performance index is positive although the correlation may not be strong. It is also a reasonable assumption that this discontinuous variable reflects an underlying continuous variation.

Horses which had not started at all or started in less than five races were pooled in an "unstarted class". The decision to pool horses which had started in 1, 2, 3 or 4 races with those which had not started at all was of course quite arbitrary but was justified by the fact that horses in both these classes must be considered to have failed in getting established as race horses. The proportion of started horses was analysed as an all-or-none trait. To reduce the binomial error variance for this trait, only sire families comprising at least 50 tested offspring were included. Despite this restriction 17,467 offspring of 142 sires were available for the analysis; the average size of the progeny groups was 123 (range 51-377). The proportion of started horses was 0.544.

Statistical analysis

Association. The blood marker phenotype was the independent variable of interest in the association test. In the electrophoretic systems, which are all codominant, each phenotype represents a single genotype whereas in the blood group systems, which all involve dominance, some phenotypes represent more than one genotype. To avoid several small phenotype classes for the electrophoretic systems involving multiple alleles, rare alleles having a frequency less than five percent were pooled with the least common allele having a frequency greater than five percent. For the same reason some rare blood group factors were ignored in the analyses of the three complex blood group systems A, D and Q.

The analysis was performed for each marker locus separately and the linear models used are given in Table 1. The use of these models is equivalent to comparing marker phenotypes within sire families. In the analysis the effects of sex and year were also eliminated.

Linkage. In the linkage test the mean value of the performance traits was compared within half-sib groups which had been divided according to the blood marker allele they had received from their heterozygous sire. The data on the allelic contribution of heterozygous sires used in the present study are almost the same as previously used for an examination of the linkage relationships among the marker loci (Sandberg and Andersson 1984). Some further data for the blood group systems were obtained by using, in addition to all offspring from matings to recessive homozygotes, all homozygous recessive offspring from other matings; it is evident that the heterozygous sire must have transmitted the recessive allele to these progeny.

Table 1. Linear models used in the analysis

Test	Model
Association	$y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \sigma_l + e_{ijklm}$
Linkage	$y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \sigma_{il} + e_{ijklm}$

In the models the symbols have the following meaning:

y	index value of animal m *or* value showing if animal m have started (1) or not started (0)
μ	mean
α_i	random effect of sire i
β_j	effect of sex j
γ_k	effect of birth year k
σ_l	effect of blood marker phenotype l
σ_{il}	effect of allele l received from the ith sire
e_{ijklm}	random residual effects

In the statistical test for linkage, we used a linear model in which the effect of marker alleles was nested within sires (Table 1). This model was chosen to allow the effect of marker alleles to vary among sire families both with regard to magnitude and sign.

Heterozygosity. The heterozygosity at the electrophoretic loci, which are all codominant, can unequivocally be determined by counting the number of heterozygous loci for each animal. This is not the case for the blood group loci as for these dominant systems some phenotypes represent both homozygous and heterozygous genotypes. For this reason data on the relationship between heterozygosity and performance traits will only be presented for electrophoretic loci whereas the result obtained on blood group loci will only be discussed. In the analysis of the blood group loci, a given phenotype representing both homozygous and heterozygous genotypes was counted as the expected Hardy-Weinberg proportion of heterozygous genotypes in that particular phenotype class. All animals included in the analysis of the heterozygosity had complete records for the electrophoretic and blood group loci, respectively. The electrophoretic locus *CA* was excluded from this analysis because the data on this locus are rather incomplete. (*CA* is only weakly expressed in very young horses and therefore difficult to type.)

The relationships between heterozygosity and performance traits were examined by performing curvilinear regressions. The analyses were carried out as described by Sokal and Rohlf (1981, p. 671) in a stepwise manner by fitting consecutively first a linear regression to the data and then adding higher order terms. The procedure was continued up to adding the cubic power of heterozygosity. The effect of sex and birth year was included in the linear model used in the analysis.

Computer analyses. The General Linear Model (GLM) procedure of the Statistical Analysis System (SAS; Goodnight et al. 1982) and the LSML76 program of Harvey (1977) was utilized for the statistical examination of the data.

Results

Association

The results obtained in the test for association are given in Table 2 for both performance traits. A highly sig-

Table 2. Probability values obtained in the test for association between blood marker phenotypes and performance in horses

Locus	No. of phenotypes	Performance index		Proportion of started horses	
		n	*P*	n	*P*
A	4	11,712	0.24	17,467	0.42
C	2	11,712	0.83	17,467	0.87
D	11	11,709	0.52	17,460	0.89
K	2	11,712	0.39	17,467	0.23
P	2	11,706	0.59	17,456	0.04
Q	2	11,712	0.77	17,467	0.08
Al	3	11,712	0.02	17,467	0.75
AP	3	10,633	0.77	16,056	0.04
CA	3	10,723	0.08	15,961	0.04
Es	6	11,676	0.56	17,403	0.000002
Hb	3	11,711	0.37	17,465	0.28
PGD	3	11,709	0.21	17,460	0.56
PGM	3	10,676	0.54	16,104	0.33
PHI	3	10,648	0.20	16,075	0.17
Tf	10	11,712	0.16	17,467	0.21

n = No. of individuals

Table 3. Least-squares estimates for phenotypes at the loci for which significant probability values were obtained in the association test

Trait	Locus	Phenotype	n	L-S Means	SE
Performance index	*Al*	FF	2,876	99.671	0.069
		FS	5,902	99.649	0.049
		SS	2,934	99.454	0.068
Proportion of started horses	*P*	Pb –	16,825	0.534	0.004
		Pb +	631	0.488	0.022
	AP	FF	149	0.482	0.041
		FS	2,822	0.513	0.011
		SS	13,085	0.538	0.006
	CA	FF	6	0.230	0.199
		FI	1,216	0.504	0.014
		II	14,739	0.536	0.005
	Es	FF	1,393	0.554	0.014
		FI	6,077	0.535	0.007
		FS	996	0.590	0.016
		II	6,460	0.507	0.007
		IS	2,268	0.557	0.011
		SS	209	0.603	0.035

n = No. of individuals

nificant association between the *Es* locus and the proportion of started horses was revealed. Further, four weakly significant associations were found. The possibility of interaction between the effects of sire and blood marker phenotype was also tested (results not shown). A significant interaction was only indicated in one case, blood group Q with respect to the performance index. Since this is not more than expected by chance only, this effect was not further considered. Furthermore, additional analyses did not indicate any significant interaction between the effects of marker loci belonging to the same linkage group.

Least-square estimates were calculated, using the procedure GLM of SAS, for all cases showing significant associations (Table 3). It should be noted that the only unbiased estimates which can be obtained in this type of analysis are the differences between phenotypes. The highly significant association between genotypes at the esterase locus and the proportion of started horses is also illustrated in Fig. 1.

Genetic linkage

No significant indication of genetic linkage was found for any combination of performance trait and marker locus, except for a weak significance for the esterase locus with respect to the proportion of started horses (Table 4). This significance was expected due to the highly significant association between this trait and variation at the esterase locus itself (cf. Table 2). Additional analyses did not reveal any significant interactions between marker loci belonging to the same linkage group (results not shown).

Heterozygosity

Means of the performance index values and least-square estimates of the proportion of started horses are given in Table 5 according to the number of heterozygous loci per individual. The reason for estimating least-square means for the proportion of started horses is that it was necessary to make corrections for sex and year of birth. There was no obvious trend in mean value for any trait with increasing heterozygosity. The results of the curvilinear regression analyses applied to the data are compiled in Table 6. This analysis revealed no significant relationship between heterozygosity and variation of the performance index whereas an indication of a cubic relationship between heterozygosity and the proportion of started horses approached significance ($P = 0.07$). This indication was not supported by an analysis using data on blood group loci (data not shown) which gave nonsignificant results for both traits.

The distribution of the number of heterozygous loci per individual based on all electrophoretic loci except *CA* was investigated by a test described by Chakraborty (1981) in order to facilitate the interpretation of the lack of relationship between heterozygosity and performance in this study. In this test the observed distribution of the number of heterozygous loci per individual is compared, by a chi-square test, with the one expected on the basis of estimated allele frequencies at individual loci. This test was applied to a sample of 24,702 individuals comprising all horses registered in the

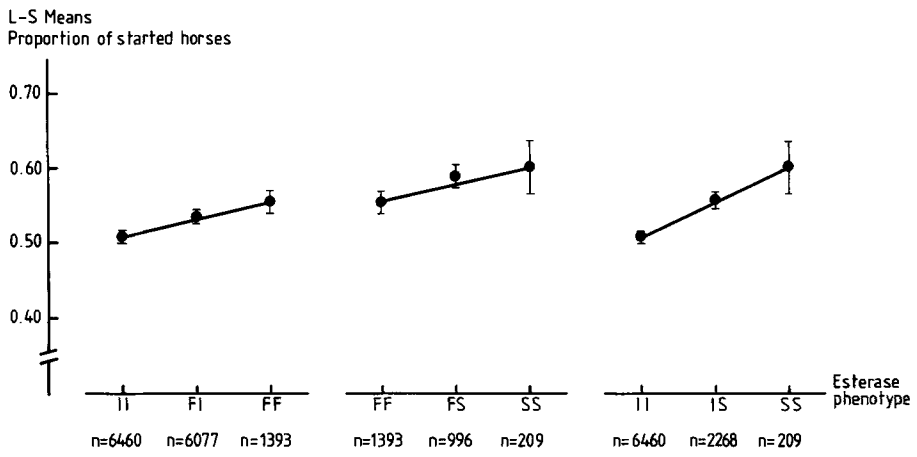

Fig. 1. Estimates of least-square means together with standard errors for the proportion of started horses among horses classified according to their phenotype at the esterase locus. The degree of dominance is illustrated by comparing alleles pairwise and by drawing a line between the point estimates for the homozygous phenotypes. n = sample size

Table 4. Test for genetic linkage between blood marker loci and performance in horses

Locus	Performance index			Proportion of started horses		
	No. of sire families	No. of offspring	*P*	No. of sire families	No. of offspring	*P*
A	56	2,125	0.77	31	3,085	0.15
C	21	504	0.82	5	330	0.72
D	44	1,416	0.68	12	1,454	0.84
K	83	3,349	0.66	44	4,587	0.29
P	11	407	0.16	6	569	0.53
Q	63	2,499	0.41	31	3,434	0.58
Al	102	4,376	0.51	54	6,085	0.79
AP	29	1,029	0.67	15	1,531	0.23
CA	20	435	0.91	3	213	0.97
Es	126	6,014	0.74	68	8,677	0.02
Hb	28	1,756	0.25	17	2,592	0.69
PGD	85	4,048	0.23	45	5,833	0.41
PGM	43	1,479	0.96	22	1,985	0.43
PHI	12	404	0.36	5	400	0.69
Tf	145	6,360	0.55	75	9,148	0.42

Table 5. Means of performance index values and proportion of started horses according to the number of heterozygous loci per individual[a]

No. of heterozygous loci	Mean of index values ± SE (n)	Proportion of started horses[b] ± SE (n)
0	99.948 ± 0.191 (326)	0.548 ± 0.024 (431)
1	99.876 ± 0.080 (1,659)	0.539 ± 0.010 (2,270)
2	100.055 ± 0.057 (3,129)	0.532 ± 0.007 (4,462)
3	100.028 ± 0.057 (3,355)	0.540 ± 0.007 (4,750)
4	99.972 ± 0.073 (1,986)	0.551 ± 0.009 (2,839)
5	100.138 ± 0.126 (643)	0.534 ± 0.016 (959)
6	99.841 ± 0.341 (126)	0.529 ± 0.037 (186)
7	99.571 ± 1.192 (7)	0.304 ± 0.098 (23)

n = No. of individuals
[a] The analysis is based on eight electrophoretic loci
[b] Estimated by least-squares means

period 1970–1979 having complete records for all electrophoretic loci except *CA*. Despite the very large sample size and the fact that year classes from a ten year period have been pooled, the probability value in the chi-square test only approached significance ($\chi^2_7 = 13.86$; $P = 0.054$). Thus, it must be concluded that there was a good agreement between expected and observed distributions in this population.

Discussion

The purpose of the present study was to investigate whether any relationship can be found between variation at biochemical marker loci and two racing performance traits (racing performance index and proportion of started horses) in horses.

Table 6. Regression of performance index values and proportions of started horses on the number of heterozygous loci per individual[a]

Trait	Source of variation	d.f.	MS	F	*P*	Regression coefficients
Performance index	Explained	3	5.562	0.52	0.67	
	Linear	1	7.402	0.70	0.40	0.115
	Quadratic	1	9.187	0.86	0.35	−0.021
	Cubic	1	0.098	0.01	0.92	0.001
	Unexplained	11,227				
Proportion of started horses	Explained	3	0.571	2.37	0.07	
	Linear	1	0.107	0.45	0.50	−0.043
	Quadratic	1	0.127	0.52	0.47	0.020
	Cubic	1	1.478	6.13	0.01	−0.002
	Unexplained	15,907	0.241			

[a] The analysis is based on eight electrophoretic loci

Racing performance is a complex trait depending on many different biological and environmental factors. The traits investigated can therefore hardly be considered ideal for studies with the aim of detecting associations between quantitative variation and marker genes as the more complex a trait is, the more genes we expect to influence it (Thompson 1975). Nevertheless, the results of the present study clearly indicate that such relationships exist. This is in good agreement with similar studies on cattle (Rendel 1959; Neimann-Sørensen and Robertson 1961; Zwiauer 1980; Geldermann et al. 1985) and pigs (Jensen et al. 1968; Kennedy et al. 1973). It is evident that this type of studies should be based on an extensive amount of data to be meaningful.

Three different forms of relationships between marker genes and racing performance were tested for, namely association, linkage and influence of heterozygosity. Statistically significant relationships were only revealed in the association test (cf. Table 2). It is evident that there was a very highly significant association between the esterase locus and the proportion of started horses. In addition four weakly significant associations were obtained. This is more than expected by chance only. However, it is not possible to judge which of these weakly significant associations, if any, are true and which are spurious. At present we can only note their presence and let future studies confirm or reject the existence of these associations.

Three esterase alleles Es^F, Es^I and Es^S are distinguished in the Swedish Trotter breed by the routine methods employed in our laboratory. Our results indicated that the effect on the proportion of started horses were different for all the three alleles. Es^S was evidently associated with a more positive effect than Es^I, while Es^F appeared to have an intermediate effect. A striking feature of the results was that the effects of different alleles showed a very good fit to an additive genetic model as the values of the heterozygous types were intermediate to the two corresponding homozygous types (Table 3; Fig. 1). The fact that the observed effect fit a plausible genetic model supports the conclusion that a true association has been revealed for this system. The variance components for the effects of sire and the esterase locus were estimated by standard means in order to quantify the relative importance of the latter system. It was found that the esterase variance component constituted about 7% of that of sire. In a half-sib analysis the sire variance is expected to estimate 25% of the genetic variance for the trait investigated. However, in this case it is very likely that the sire variance is somewhat inflated. Highly ranked stallions are more often mated to mares having high breeding values and more effort is usually invested to get offspring of famous stallions started.

It is possible to increase the number of alleles distinguished at the esterase locus by employing other electrophoretic methods than those used in our routine analyses (Scott 1972; Fisher and Scott 1978). In the Swedish Trotter breed, the allele Es^F can be divided into two alleles, designated Es^F and Es^G, having about equal frequency. This fact does not diminish the validity of the results in the present study but suggests that some information may have been lost. Despite that only breeder stallions had been classified with respect to Es^F and Es^G it was possible to get an indication of the effect of these alleles on the proportion of started horses. The linear model used for this purpose was almost identical with the one used in the general association test (cf. "Materials and methods") with the exception that only heterozygous sires (Es^F/Es^I, Es^F/Es^S, Es^G/Es^I, Es^G/Es^S and Es^I/Es^S) were used and that the independent variable of interest was the allele transmitted from the heterozygous sires to their offspring. This means that the data set was the same as that employed in the test for genetic linkage. By using the alternative model and by utilizing the information on the genotype of sires with respect to Es^F and Es^G, it

was possible to compare four alleles at the esterase locus (Es^F, Es^G, Es^I and Es^S). This analysis indicated that there were no significant differences between Es^F and Es^I or between Es^G and Es^S whereas all other pairwise comparisons of alleles were significant. Thus, the alleles at the esterase locus appear to be divided into two groups. Namely, one group (Es^G and Es^S) with an apparent positive effect on the proportion of started horses compared with the other group (Es^F and Es^I).

The highly significant association between variation at the esterase locus and the proportion of started horses suggests that esterase allele frequencies may change over time because of artificial selection. The allele frequencies of each year class born in the period 1970–1983 were therefore analysed. The alleles Es^F, Es^I and Es^S occurred with frequencies of about 0.25, 0.65 and 0.10, respectively, in each year class. The lack of changes is hardly surprising as the time period is rather short compared with the generation interval which is about ten years in this breed. Furthermore, it is uncertain whether the ability to get started is selected for in the Swedish Trotter breed. Most of the selection pressure is put on the selection of stallions with excellent performance. The selection of mares is very weak as most mares are used for breeding. The application of this procedure implies that unstarted mares will be used for breeding at a younger age than those which successfully compete on the race-track. The effect of this breeding practice appears to be that the ability to get started is selected for among stallions but against among mares. The combined effect is likely to be no or weak selection only for this trait.

Which type of analysis is best suited to detect relationships between marker genes and quantitative traits in segregating populations? As regards relationships to single marker genes, tests for association and tests for linkage may be carried out. Most recent studies have focused on the latter type of tests (e.g. Soller and Beckmann 1982, 1983; Geldermann et al. 1985). The difference between the two types of test may be described as follows. The association test is designed to detect a relationship, between a marker gene and a quantitative trait, which is consistent in all families examined in a given study. In the linkage test, on the other hand, the relationship may vary both in magnitude and direction since the marker may, in different families, be linked to different alleles at a postulated linked locus affecting the quantitative trait. These differences in statistical outline imply that the association test is always the most sensitive to detect a pleiotropic effect of a marker gene. The association test may also detect a relationship caused by linkage provided that there is linkage disequilibrium between the marker gene and the linked gene; the stronger disequilibrium the more powerful the association test will be to detect a relationship. The linkage test is less powerful to detect an association (caused by pleiotropy or strong linkage disequilibrium) for two reasons. Firstly, the material is much reduced in the linkage test since it only includes those offspring for which the marker allele inherited from a heterozygous parent can be deduced. Secondly, in the association test a small, consistent effect of the same sign will accumulate to a highly significant effect when all families are considered. On the other hand, when there is no disequilibrium only the linkage test may reveal a relationship.

A major impetus for carrying out the present investigation was the strong need for empirical data on the suitability of different methods for detecting relationships between marker genes and quantitative traits in farm animals. According to the result of the present study the test for association appears to be the most promising one. This finding indicates that either variation at biochemical loci contributes to quantitative variation (i.e. pleiotropy) or linkage disequilibrium between marker genes and genes affecting quantitative traits occurs more frequently than previously thought. With available data, it is not possible to determine which of these possibilities (or both) is valid for the observed associations. However, the negative result in the linkage test (cf. Tables 2 and 4) gives no support for the interpretation that they were caused by linkage although it did not, of course, exclude this possibility. It should be noted that evidence of pleiotropic effects must be based on experimental data showing the physiological significance of the variation at the marker loci.

It is a well established fact that reduced heterozygosity due to inbreeding leads to reduced fitness. This fact gives raise to the important question whether the degree of heterozygosity is correlated with fitness among individuals from randomly breeding populations also. Data on heterozygosity at marker loci offer a possibility to test this relationship. There is in fact a number of reports, based on electrophoretic data from natural populations, which indicates a positive relationship between heterozygosity and fitness (for review see Soulé 1980). Furthermore, several investigations have tested Lerner's (1954) hypothesis that increased heterozygosity should reduce morphological variability. The finding in this respect is conflicting, however, since significant correlations have been found in some species (Mitton 1978; Eanes 1978; Leary et al. 1983) but not in others (Handford 1980; McAndrew et al. 1982; Ryman et al. 1984). It is therefore necessary to investigate a wide range of species before any general conclusions can be drawn concerning the relationship between heterozygosity and fitness. The present investigation contributes to this issue.

It is obvious that racing competition must be a rather stringent measure of the physiological condition of horses. Despite the very large amount of data analysed, we found no significant correlation between heterozygosity and racing performance. There was no indication at all that a high number of heterozygous loci was associated with a better racing performance. A relationship to heterozygosity may either be due to variation at the biochemical loci themselves or to variation at other loci affecting the trait provided that the heterozygosity at the biochemical loci reflects that of those other loci. The lack of relationship between heterozygosity and racing performance is entirely consistent with other findings in the present study. Firstly, the association test indicated that heterozygosity at

individual loci should not influence performance since the effect of alleles at marker loci was either negligible or additive. Secondly, the analysis of the distribution of the number of heterozygous loci did not indicate that the heterozygosity at marker loci should reflect that at other loci since there was no apparent deviation from a random association of alleles among different loci in this population.

In our previous studies on genetic linkage in the horse we found it remarkable that among the approximately 30 loci tested for linkage so far, no less than seven loci have been assigned to the same linkage group (Andersson 1983; Sandberg and Andersson 1984). This apparent non-random distribution of marker loci is similar to observations in other species. It is therefore particularly interesting to note that the most significant association detected in the present study involved the esterase locus, which in fact belongs to this large linkage group. There is a parallel situation in the pig which has one major linkage group comprising five marker loci and in addition the halothane locus which has a major effect on both productive and reproductive traits (see Juneja et al. 1983). The observation in both pigs and horses of a clustering of polymorphic loci in a chromosome region containing important quantitative trait loci is worth paying attention to. However, there appears to be no theoretical ground for expecting such relationships.

The most important finding in the present study was the highly significant association between esterase genotypes and the proportion of started horses. The finding of such a strong association should be the starting point for further studies with the aim of disclosing the causal basis for the observation. The most obvious topic for future research is to study specific reasons why horses are unstarted and then investigate if one of those reasons may explain the association to the esterase locus. An earlier investigation in the Swedish Trotter breed indicated that lesions, diseases, bad mentality, bad talent and growth problems were the most important reasons for horses to be unstarted (Bendroth 1981). With improved knowledge on the basis of the association it will hopefully be possible to elucidate whether the effect is caused by the esterase locus itself or by a linked locus.

Acknowledgements. We are indebted to J. Rendel, P. Seeger and C. Smith for kindly reviewing a preliminary version of the manuscript. We thank N. Ryman and G. Ståhl for the possibility to use their computer program for testing the distribution of heterozygous loci. The investigation was financially supported by the Swedish Agricultural Research Council and the Swedish Racing Board.

References

Andersson L (1983) Studies on genetic linkage in domestic animals with special reference to the horse. PhD Thesis, Swedish University of Agricultural Sciences, Uppsala

Andersson L, Sandberg K (1982) A linkage group composed of three coat color genes and three serum protein loci in horses. J Hered 73:91–94

Andersson L, Sandberg K (1984) Genetic linkage in the horse. 2. Distribution of male recombination estimates and the influence of age, breed and sex on recombination frequency. Genetics 106:109–122

Andersson L, Arnason Th, Sandberg K (1985) Studies on possible associations between genetic markers and racing performance in horses. Anim. Blood Groups Biochem Genet (Suppl 1) 16:90–91

Andresen E, Jensen P (1977) Close linkage established between the HAL locus for halothane sensitivity and the PHI (phosphohexose isomerase) locus in pigs of the Danish Landrace breed. Nord Vet Med 29:502–504

Arnason Th, Darenius A, Philipsson J (1982) Genetic selection indices for Swedish trotter broodmares. Livest Prod Sci 8:557–565

Beckmann JS, Soller M (1983) Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theor Appl Genet 67:35–43

Bendroth M (1981) Orsaker till låg startfrekvens hos 2–4-åriga varm- och kallblodiga svenska travare. Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala (mimeograph in Swedish)

Briles WE, Stone HA, Cole RK (1977) Marek's disease: effects of *B* histocompatibility alloalleles in resistant and susceptible chicken lines. Science 195:193–195

Chakraborty R (1981) The distribution of the number of heterozygous loci in an individual in natural populations. Genetics 98:461–466

Cockerham CC, Weir BS (1983) Linkage between a marker locus and a quantitative trait of sibs. Am J Hum Genet 35:263–273

Eanes WF (1978) Morphological variance and enzyme heterozygosity in the monarch butterfly. Nature 276:263–264

Elston RC (1979) Major locus analysis for quantitative traits. Am J Hum Genet 31:655–661

Fisher RA, Scott AM (1978) Isoelectric focusing of horse serum esterase isozymes and detection of new phenotypes. Anim Blood Groups Biochem Genet 9:207–213

Geldermann H (1975) Investigations on inheritance of quantitative characters in animals by gene markers. 1. Methods. Theor Appl Genet 46:319–330

Geldermann H (1976) Investigations on inheritance of quantitative characters in animals by gene markers. 2. Expected effects. Theor Appl Genet 47:1–4

Geldermann H, Pieper U, Roth B (1985) Effects of marked chromosome sections on milk performance in cattle. Theor Appl Genet 70:138–146

Goodnight JH, Sall JP, Sarle WS (1982) The GLM procedure. In: SAS user's guide: statistics. Cary NC, SAS Institute, pp 139–199

Handford P (1980) Heterozygosity at enzyme loci and morphological variation. Nature 286:261–262

Hansen MP, Van Zandt JN, Law GRJ (1967) Differences in susceptibility to Marek's disease in chickens carrying two different *B* locus bloodgroup alleles. Poult Sc 46:1268

Harvey WR (1977) User's guide for LSML76. The Ohio State University, Columbus, Ohio, 76 pp

Hintz RL (1980) Genetics of performance in the horse. J Anim Sci 51:582–594

Jayakar SD (1970) On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus. Biometrics 26:451–464

Jensen EL, Smith C, Baker LN, Cox DF (1968) Quantitative studies on blood group and serum protein systems in pigs. 2. Effects on production and reproduction. J Anim Sci 27:856–862

Juneja RK, Gahne B, Edfors-Lilja I, Andresen E (1983) Genetic variation at a pig serum protein locus, *PO-2* and its assignment to the *Phi, Hal, S, H, Pgd* linkage group. Anim Blood Groups Biochem Genet 14:27–36

Kennedy BW, Moxley JE, Saison R (1973) Some relationships between blood group factors and economic traits in swine. Can J Anim Sci 53:389–397

Kluge R, Geldermann H (1982) Effects of marked chromosome sections on quantitative traits in the mouse. Theor Appl Genet 62:1–4

Langlois B (1982) L'héritabilité des performances chez le *Trotteur.* Une revue bibliographique. Ann Genet Sel Anim 14:399–414

Leary RF, Allendorf FW, Knudsen KL (1983) Developmental stability and enzyme heterozygosity in rainbow trout. Nature 301:71–72

Lerner IM (1954) Genetic homeostasis. Oliver and Boyd, Edinburgh

McAndrew BJ, Ward RD, Beardmore JA (1982) Lack of relationship between morphological variance and enzyme heterozygosity in the plaice, *Pleuronectes platessa.* Heredity 48:117–125

Mitton JB (1978) Relationship between heterozygosity for enzyme loci and variation of morphological characters in natural populations. Nature 273:661–662

Neimann-Sørensen A, Robertson A (1961) The association between blood groups and several production characteristics in three Danish cattle breeds. Acta Agric Scand 11:163–196

Rasmusen BA (1981) Linkage of genes for PHI, halothane sensitivity, A-O inhibition, H red blood cell antigen and 6-PGD variants in pigs. Anim Blood Groups Biochem Genet 12:207–209

Rendel J (1959) A study on relationships between blood group and production characters in cattle. In: Rep 6th Int Congr Blood Group Res in Animals. Munich, pp 8–23

Ryman N, Lagercrantz U, Andersson L, Chakraborty R, Rosenberg R (1984) Lack of correspondence between genetic and morphologic variability patterns in Atlantic herring (*Clupea harengus*). Heredity 53:687–704

Sandberg K, Andersson L (1984) Genetic linkage in the horse. 1. Linkage relationships among 15 blood marker loci. Herditas 100:199–208

Scott AM (1972) Improved separation of polymorphic esterases in horses. In: Proc 12th Eur Conf Anim Blood Groups Biochem Polymorphism. Budapest, pp 551–553

Smith CAB (1975) A non-parametric test for linkage with a quantitative character. Ann Hum Genet 38:451–460

Sokal RR, Rohlf FJ (1981) Biometry, 2nd edn. WH Freeman, San Francisco

Soller M (1974) Methods for the detection of linkage between a marker gene and a linked polygene affecting a quantitative trait. In: Proc 1st World Congr Genet Applied to Livestock Prod 3:284–294

Soller M, Beckmann JS (1982) Restriction fragment length polymorphisms and genetic improvement. In: Proc 2nd World Congr Genet Applied to Livestock Prod 6:396–404

Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67:25–33

Soller M, Genizi A (1978) The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. Biometrics 34:47–55

Soller M, Brody T, Genizi A (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor Appl Genet 47:35–39

Soulé ME (1980) Thresholds for survival: maintaining fitness and evolutionary potential. In: Soulé ME, Wilcox BA (eds) Conservation biology. Sinauer Associates, Massachusetts

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:1–25

Thoday JM (1961) Location of polygenes. Nature 191:368–370

Thoday JM (1979) Polygene mapping: uses and limitations. In: Thompson JN, Thoday JM (eds) Quantitative genetic variation. Academic Press, New York, pp 219–233

Thompson JN Jr (1975) Quantitative variation and gene number. Nature 258:665–668

Tolley EA, Notter DR, Marlowe TJ (1985) A review of the inheritance of racing performance in horses. Anim Breed Abstr 53:163–185

Weitkamp LR, Andersson L (1984) Linkage groups of the horse (*Equus Caballus*). Genet Maps 3:390–391

Zwiauer D (1980) Beziehungen zwischen biochemischen Markergenen und Leistungseigenschaften beim Rind. Fortschr Tierzücht Züchtungsbiol, vol 2. Parey, Hamburg Berlin