Theor Appl Genet (1992) 85:403–406

Responses to selection on genotypic or phenotypic values in the presence of genes with major effects

E. Sehested[1] **and I. L. Mao**[2]

[1] Department of Animal Science, Norwegian Pig Breeders Association, Agricultural University of Norway, N-1432 Aas, Norway
[2] Department of Animal Science, Michigan State University, E. Lansing, MI 48824-1225, USA

Received March 9, 1992; Accepted June 3, 1992
Communicated by G. Wenzel

Summary. Average genotypic responses were compared after selection for genotypic values and for phenotypic values on the basis of single-gene models and multigene models in simulated livestock populations. Single-gene models dealt with single gene control of the genetic differences between animals, while multigene models considered a collection of genes with various magnitudes of effects on a trait. In each case, selection lasted through discrete generations until the fixation of the gene frequencies occurred. Generations to reach fixation were used to compare various models, and the two criteria for selection, for their efficiency in selection. Implications of using these models versus using infinitesimal models for selection in practice are presented.

Key words: Stochastic simulation – Selection response – Single genes – Genotype – Phenotype

Introduction

In selection methods applied in livestock improvement programmes, various forms of the infinitesimal model for genetic effects have been used. This model assumes an infinite number of genes each with trivial effect but collectively determining the breeding value of an animal. However, Robertson (1967) suggested that a more suitable model may involve a small number of major loci: a larger number of moderate ones and a very large number of minor loci, where major, moderate and minor describe the relative magnitude of influences constituting the additive genetic effect on a quantitative trait of interest. Therefore, although applications of infinitesimal models have been successful in different species and populations of livestock (Smith 1984), the exploration of major genes that is now possible as a result of advances in gene technology may present opportunities for furthering the accuracy of genetic selection.

If and when a gene is proven to have a major effect, incorporating that gene in a selection programme by directly selecting its genotypic value would be an improvement over traditional selection programmes. This has prompted several authors to announce the arrival of a new era in animal breeding (e.g. Smith and McMillan 1989; Kennedy et al. 1990). However, little has been reported on the efficiency of selection on the genotypic value of major genes relative to the traditional selection method from an infinitesimal model. In fact, it is possible that genes with large effects on economically important livestock traits may have been in existence and undergone selection for generations. It would be interesting to examine how successful a selection programme could have been had the existence of such genes been recognized. Therefore, the main objectives of the present study were: (1) to compare the efficiency of alternative methods of selection in a population of medium size in which genotypic values were determined by a finite number of loci, and (2) to evaluate the likelihood of the existence of major genes in a population in which selection on phenotypes was practiced.

Methods

The following models were used in the simulation. Only two types of alleles, superior and neutral for a hypothetical trait, were assumed at each locus. In the base population at generation zero,

---

Correspondence to: E. Sehested

the genotypic value (G) of each animal was

$$G = \Sigma_i b_i f(A_i) \qquad (1)$$

where b_i was the additive effect of the superior allele in locus i. The additive effect of the neutral allele was set to zero. Variable $f(A_i)$ took on a value of 0, 1 or 2 for the number of superior alleles at locus i. The number of superior alleles at each locus was assigned according to a likelihood function of a stochastic variable simulating the Hardy-Weinberg equilibrium. The initial frequency of the superior allele was 0.01 for all loci. There was no linkage between loci.

The phenotypic value (P) of each animal was

$$P = G + a\sigma_e \qquad (2)$$

where a was the stochastic variable that followed a standardized normal distribution with zero expectation and unity variance, and σ_e was the environmental standard deviation.

Two versions of model [1] for genetic effects were applied: (1) a single locus model for $i = 1$ and $b_1 = k$, and (2) a multigene model in which five loci had effects on the phenotype with assigned magnitude of $b_i = 1/i$ for $i = 1, 2, 3, 4$ or 5. In order to achieve the same maximum genotypic values in both versions of model (1), k in the single locus version was chosen to equal Σb_i in the version with five loci.

In both versions, an additional locus with two alleles with equal frequencies of 0.5, which had no effect on the phenotype, was included. This was done to monitor possible inbreeding and genetic drift.

Based on each of the two versions of model (1), selection was carried out on genotypic values in genotypic selection and on phenotypic values (2) in phenotypic selection. As can be seen from the definitions of G and P, the response to selection on genotypic values would be independent of σ_e, while the response to selection on phenotypic values would not be. Hence, in cases of selection for phenotypes, the ratio of the genetic quadratic term over the phenotypic quadratic term would have to be defined. A parameter g^2 was used to denote such ratio, which is analogous to the heritability, h^2, in the infinitesimal model:

$$g^2 = V_G/(V_G + \sigma_e^2)$$

where V_G, which denotes the genetic quadratic term, is dependent on the frequencies of the superior alleles in each locus. These frequencies would change in the process of selection from generation to generation. In order to standardize, the g^2 value was defined for a Hardy-Weinberg equilibrium state at which the frequencies of superior alleles were 0.5 in all loci, hence representing the maximum g^2 throughout the selection process. Three levels of g^2 were chosen: 0.25, 0.50 and 0.75. This was achieved by adjusting σ_e, while holding V_G constant, for each of the g^2 values (Table 1).

The populations were simulated with the intention to mimic the structures in important livestock species in numbers and selection intensity. However, the generations in the simulated populations were discrete, i.e. no animals survived across generations. For each generation, 100 males and 500 females were selected as parents for the next generation and mated randomly. The number of offspring per female was determined by a stochastic variable with a probability distribution of

(0.04, 0.08, 0.12, 0.16, 0.20, 0.16, 0.12, 0.08, 0.04)

corresponding to 'litter sizes' of 0, 1, ..., 8, respectively. This gave an expected litter size of 4, and maintained an average population size of 2000 animals born per generation. The sex of each animal was determined by another stochastic variable, which gave equal probabilities for the two sexes. The effects of inbreeding on phenotypic performance were ignored.

For the multigene model, genotypic selection was also done for the practical situation in which the genetic effects were not known. This was simulated by using G_j instead of G as selection criterion, where G_1 considered only genetic effect of locus number 1, G_2 of locus numbers 1 and 2, G_3 of locus numbers 1, 2 and 3, and G_3 of locus numbers 1, 2, 3 and 4.

Table 1. Effects of superior genes relative to σ_e in the case of selection on phenotypic values

Effects of		Model					
		Single locus	Multiple loci				
			1	2	3	4	5
Superior gene		2.28	1.00	0.50	0.33	0.25	0.20
σ_e	$g^2 = 0.75$	0.93			0.49		
	$g^2 = 0.50$	1.62			0.86		
	$g^2 = 0.25$	2.80			1.48		

Selection continued until superior alleles in all loci had reached fixation or up to 30 generations, whichever occurred first. Each case was replicated 10 times.

Results and discussion

The gene frequencies of alleles in the neutral locus stayed close to equilibrium in all cases, and frequencies of heterozygotes did not deviate significantly from 2pq either. This indicated that both genetic drift and inbreeding were negligible in all cases studied.

Selection responses under single-gene models

Average genotypic values as response to selection by single-gene models are presented in Fig. 1. Fixation of the favourable gene was reached in generation 6 when selection was based on genotypes. When selection was based on phenotypes, the favourable gene reached fixation at generations 7, 10 and 16 for g^2 values of 0.75, 0.50 and 0.25, respectively.

With $g^2 = 0.75$, the difference between genotypic and phenotypic selection with respect to rates of reaching fixation was negligible. This may mean that once the effect of an allele falls beyond a threshold, selection on genotypes no longer gives a greater response than selection on phenotypes. This threshold point most likely lies between g^2 values of 0.75 and 0.5 under a single-gene model. In the first few generations of selection, when the frequencies of the superior allele were low, the number of parents needed to maintain the population size was much greater than the number of carriers of the superior allele in the population. With high g^2 the probability of failing to include all carriers among the animals selected on phenotypes was small.

When the effect of the major gene fell short of that threshold, selection on genotypes was more efficient

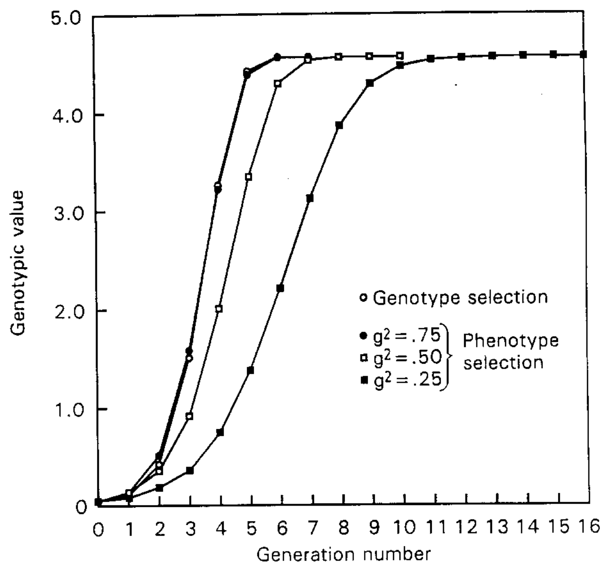

Fig. 1. Response to selection by the single-gene model

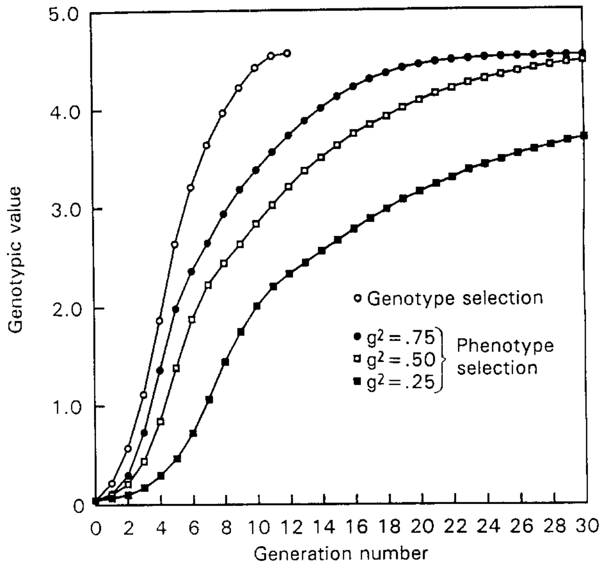

Fig. 2. Response to selection by the multigene model

than selection on phenotypes. However, the results showed that if a gene existed with a major effects on the phenotype in a population that was undergoing selection on that phenotype, this gene was likely to become fixed within only a few generations.

Selection responses under multigene models

Average genotypic values responding to selection by multigene models are presented in Fig. 2. Favourable genes at all five loci reached fixation by the 12th generation under genotypic selection, whereas fixation at all of the loci was not reached by the 30th generation in any of the three phenotypic selection schemes. The distance from generation 30 to the generation when fixation occurred was relatively large with $g^2 = 0.25$. The generation number at fixation and gene frequencies of genes that did not reach fixation at generation 30 are summarized in Table 2. No favourable genes were lost in any of the replicates.

The multigene model used in this study was a simple approximation to the more realistic model of gene effects according to Robertson (1967). Figure 2 shows that selection on genotypes was consistently superior to selection on phenotypes, even when g^2 was as high as 0.75. The relative superiority of genotypic selection over phenotypic selection appeared to be negatively correlated to the magnitude of g^2, which would imply that it is more effective to conduct genotypic selection when g^2 is low. However, a low g^2 means a low ratio of b_i over σ_e and hence a low accuracy in the estimation of b_i. This issue is closely linked to the problems in identifying single genes and in the estimation of their effects involving designs and power of tests that have been discussed in recent literature (Elsen et al. 1991; Knott et al. 1990; Weller 1990).

Table 2. Generation number at which fixation occurred for selection under the multigene model, and gene frequencies at generation 30 when fixation did not occur (in brackets)

Locus number	Relative gene effect	Selection on			
		Genotype	Phenotype		
			$g^2 = 0.75$	$g^2 = 0.50$	$g^2 = 0.25$
1	1	8	9	12	19
2	1/2	11	16	25	(0.89)
3	1/3	11	22	(0.99)	(0.66)
4	1/4	11	29	(0.94)	(0.50)
5	1/5	13	(0.94)	(0.88)	(0.29)

Not surprisingly, alleles with large effects reached fixation sooner than alleles with smaller effects under phenotypic selection. After some generations of selection on phenotypes, the remaining segregating loci will be those with small effects on the phenotype. In extrapolating this finding, one may conclude that the infinitesimal model would become more realistic than the model suggested by Robertson (1967) as phenotypic selection was applied over generations.

The results for genotypic selection under the multigene model when the gene effects were not known, i.e. selecting on G_1, G_2, G_3 and G_4 instead of selecting on G, are shown in Fig. 3. The selection limit was 4.57 under both phenotypic and genotypic selection when selecting all five loci. When selecting on genotypic values of less than five loci, i.e. when selecting on G_1, G_2, G_3 and G_4, the selection limits were 2.00, 3.00, 3.67

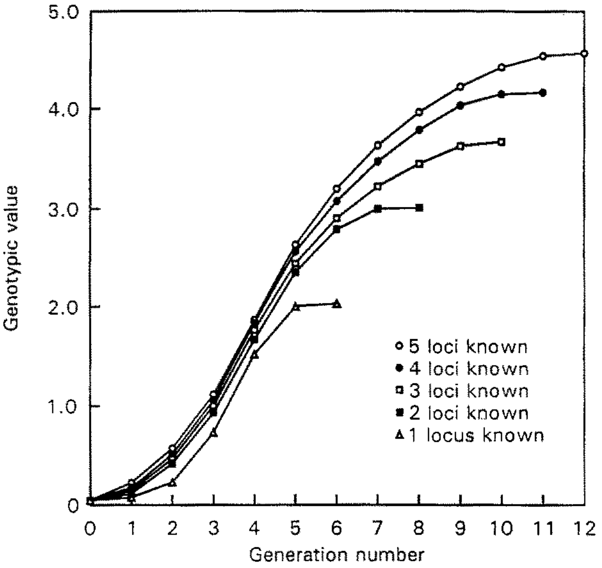

Fig. 3. Response to gene selection by the multigene model when knowledge of the gene effects varies

and 4.17, respectively. The fixation of favourable genes was reached after 6, 8, 10 and 11 generations when selection was based on one, two, three and four loci, respectively.

Gene effects are most probably unknown in practice, and it is likely that particular loci of major importance are ignored when selecting on genotypes. The results presented in Fig. 3 are therefore relevant for a realistic situation. If single gene effects are not completely known, the general conclusion of selection on genotypes being superior over selection on phenotypes may not necessarily apply in all cases. In such cases, genotypic selection may be regarded as a supplement, but not as an alternative, to phenotypic selection. Combining the two can be done by utilizing mixed model methodology as described by Hoeschele (1988).

The simulation results indicated that the probability of finding genes with major effects on phenotypes is very low for traits that have undergone phenotypic selection for several generations. However, linkage, mutation, genetic correlations between traits of interest and non-additive genetic effects such as overdominance could keep major genes segregating in a population that has undergone more generations of selection than those populations studied here.

Conclusions

Under a single-gene model, the difference in selection results from selecting genotypic values versus phenotypic values is small if the genetic differences account for a large proportion of the total variation. However, selection on genotypic values is superior if that proportion is moderate or small in magnitude. In any case, single genes with major effects become fixed after only a few generations under selection. Under a multigene model, selection based on genotypic values is consistently superior over that on phenotypic values. Genes with large effects reach fixation sooner than those with smaller effects, and after a population has undergone selection for several generations, the likelihood of identifying segregating single genes with major effects is very small.

Traits that have undergone intense selection for several generations and still show substantial genetic variation are likely to be determined by a large number of loci each with a trivial effect. Hence, selection based on an infinitesimal model seems to be an appropriate approach for these traits. The traits in various livestock species that are economically important may mainly belong to this category, since they all have undergone intense selection in most developed populations. Thus, the search for major genes should focus on those traits that have not undergone meaningful selection. When single genes with major effects are found, the most efficient approach to utilize this knowledge in a breeding programme is probably to incorporate the genotypic values of these loci into the prediction of breeding values based on cumulated effects from the genes on quantitative loci.

References

Elsen JM, Le Roy P, Goffinet B (1991) Comparison of four statistical tests for genotype determination at a major locus of progeny-tested sires. J Anim Breed Genet 108:167–173

Hoeschele I (1988) Genetic evaluation with data presenting evidence of mixed major gene and polygenic inheritance. Theor Appl Genet 76:81–92

Kennedy BW, Verrinder Gibbins AM, Gibson JP, Smith C (1990) Coalescence of molecular and quantitative genetics for livestock improvement. J Dairy Sci 73:2619–2627

Knott SA, Haley CS, Thompson R (1990) Approximations to segregation analysis for the detection of major genes. In: Hill WG, Thompson R, Wooliams JA (eds) 4th World Congr Genet Appl Livestock Prod, vol 13. Edinburgh, pp 504–507

Robertson A (1967) The nature of quantitative variation. In: Brink RA, Styles ED (eds) Heritage from Mendel. University of Wisconsin Press, Madison, p 265

Smith C (1984) Rates of genetic change in farm livestock. Res Dev Agric 1:79–85

Smith C, McMillan I (1989) Use of identified genes in animal breeding. In: Hill WG, Mackay TFC (eds) Evolution and animal breeding. Reviews on molecular and quantitative approaches in honour of Allan Robertson. CAB Int, Wallingford, pp 237–241

Weller JI (1990) Experimental design for mapping quantitative trait loci in segregating populations. In: Hill WG, Thompson R, Wooliams JA (eds) 4th World Congr Genet Appl Livestock Prod, vol 13. Edinburgh, pp 113–116