

Genetic Selection Strategies: Computer Modeling¹

WILLIAM M. MUIR

Department of Animal Sciences, Purdue University, West Lafayette, Indiana 47907-1151

ABSTRACT There are four primary factors to consider in genetic selection strategies: 1) accuracy of selection, 2) selection intensity, 3) effective population size, and 4) mating system. Current theory indicates that optimum response to selection is achieved by maximizing the first three factors and using a mating systems that allows optimization of reproductive characteristics in dam lines and production characteristics in sire lines. However, with limited resources, compromises among the first three factors are needed. Simulations are useful for examining those compromises. Unrealistic simplifying assumptions are necessary for analytic theoretical results and thus do not address real world breeding problems.

Using simulations, the relationship between selection accuracy, which is increased by use of family selection indices or Best Linear Unbiased Prediction (BLUP), and response to selection was examined. Results show that those procedures place a great restriction on effective population size, which offsets most of their advantage, i.e., there is too little emphasis on effective population size. A revision of the methodology and a reappraisal of the results of selection theory for optimization of genetic response is required.

(Key words: selection intensity, infinitesimal model, gene model, simulation, family selection)

1997 Poultry Science 76:1066–1070

INTRODUCTION

There are four primary factors to consider in genetic selection strategies: 1) accuracy of selection, 2) selection intensity, 3) effective population size, and 4) mating system. Current breeding theory indicates that optimum response to selection can be achieved by maximizing the first three factors and using a mating systems that allows for optimization of reproductive characteristics in dam lines and production characteristics in sire lines. However, with limited resources, not all of these factors can be maximized. For example, increasing selection intensity decreases effective population size, which results in a decrease in response to selection (Robertson,

Another relationship that is of fundamental importance in breeding programs is that between selection intensity and effective population size. Analytical results for the additive case have been developed but have never been extended to heterotic traits. A gene level simulation program was developed to examine that relationship. Results show that the optimal selection strategy depends on the trait being selected. For additive traits and in the short term (20 generations), one should maximize selection intensity. For heterotic traits, an intermediate proportion (25% of each sex) gives optimal response.

In all breeding strategies, primary attention must be given to the rate of inbreeding, which is increased by increasing either accuracy of selection or selection intensity. Inbreeding reduces response to selection in two ways. First, for both additive and nonadditive traits, inbreeding is a measure of the amount of random genetic drift that has occurred. Genetic drift causes loss of favorable alleles. Once lost, those alleles can never be recovered and thus genetic drift lowers the selection limit. Second, for heterotic traits, inbreeding results in a depression of the mean caused by directional dominance.

1960). Similarly, increasing the accuracy of selection by use of family selection indices also reduces effective population, which results in a more rapid increase in the rate of inbreeding (Robertson, 1961). Those results were arrived at by mathematical modeling.

When mathematical derivations are available, they are preferred to simulations. The major disadvantage of mathematical modeling is that the equations quickly become intractable except if very simplistic, and often unrealistic, assumptions are made. As a result, theoretical guidelines for optimum combinations of selection intensity, accuracy, and population size have not been derived, even for additive gene action. However, by modeling the selection process, such complex interrelationships can be examined using simulation. Simulations are based on a set of assumptions from which data is generated and alternative breeding schemes are compared. Simulations always includes a stochastic element and are generally refereed to as Monte Carlo simula-

Received for publication August 9, 1996.

Accepted for publication February 21, 1997.

¹Journal Paper Number 15125 of the Purdue University Agriculture Research Program.

tions, implying that chance or gaming theory is involved. With Monte Carlo simulation, many assumptions can be relaxed, but only a limited number of specific cases can be examined. The generality of the findings depends on how many combinations of parameters were examined, how similar those are to actual animal breeding situations, and remaining assumptions.

Computer simulations of genetic systems can be divided into two categories, those based on genes, i.e., there are a specified number of loci, with a specified number of alleles, etc., and those based on an infinite number of loci, in which means and variance structures are specified. The latter has been termed the *infinitesimal model* (Bulmer, 1971), whereas the former will be referred to as the *gene model*. The assumption of an infinite number of loci in the infinitesimal model is necessary so as selection will not change gene frequency. The infinitesimal model also assumes additivity because changes in the variance structure with selection can only be predicted in that limited case. With those assumptions, the changes in disequilibrium and additive variance as a result of selection and inbreeding can be predicted (Bulmer, 1971; Sorensen and Kennedy, 1984). The validity of the infinitesimal model is clearly limited, but in the short term, if selection is not too intense and there are a large number of additive loci influencing the trait (> 500), predictions of the model are reasonable as verified by a gene level simulation (Muir, unpublished observations).

Although the classic paper of Robertson (1960) gave a thorough examination of the relationship between effective population size and selection intensity, the results were limited to the case of additive gene action. Most traits of economic importance are heterotic in nature and affected by inbreeding depression, indicating that dominant gene action is important for those traits. A second factor to consider in this relationship is the time scale. Poultry breeders are not only concerned with selection limits but must stay in business long enough to achieve those limits. Thus, optimizing short-term response to selection is also crucial to breeders. Little attention has been given to the relationship between these factors and short- to intermediate-term response to selection.

The objective of this paper is, first, to examine the relationship between accuracy of selection, and selection response and, second, to examine the relationship between selection intensity, effective population size, and selection response in the short to intermediate term for additive and heterotic traits.

ACCURACY OF SELECTION AND SELECTION RESPONSE

The Best Linear Unbiased Prediction (BLUP; then described as Henderson's, 1963 "best index") is a generalized procedure for combining information on the individual and all relatives in a selection program and

greatly increases the accuracy of selection, especially for traits with low heritabilities. Lin (1978) reviewed selection index, basic theory, assumptions, and limitations. He classifies limitations of selection index to: changes in genetic parameters due to selection, sampling errors of parameter estimation, evaluation of relative economic weights, and internal deterrents to index selection. Limitations of index selection have been demonstrated in a number of species. However, reasons for those failures were not understood until simulations were used to investigate possible causes. For example, a number of experiments failed to confirm the theoretical advantages of index selection (single trait, multiple sources) over mass selection (Kinney *et al.* 1970; Doolittle, *et al.*, 1972; Wilson, 1974; Campo and Tagarro, 1977; Garwood and Lowe, 1979; Garwood *et al.*, 1980; Perez and Toro, 1992). Wilson (1974:836) concluded that "there is no obvious explanation for the discrepancies that exist between these experimental results and the theoretical expectations". A possible explanation is inbreeding.

Robertson (1961) warned that selection based on family indices would result in smaller effective population sizes because relatives tend to be selected. However, it was not until the computer simulation results presented by Quinton *et al.* (1992) and Quinton and Smith (1994) that the extreme limitation this placed on BLUP was understood. Quinton *et al.* (1992) concluded that comparisons between alternative breeding schemes should be made at the same level of inbreeding rather than at the same selection intensity due to adverse effects of inbreeding depression. In addition to inbreeding depression, there are also consequences for long-term response caused by a loss of favorable alleles due to drift, which thereby lowers the selection limit (Robertson, 1960).

Using an infinitesimal simulation model, Quinton *et al.* (1992) showed that at the same level of inbreeding, differences between methods (BLUP vs Phenotypic) were much smaller than if inbreeding was unrestricted. In fact, if low to intermediate levels of inbreeding were targeted, as in a closed line of limited size, their results show that selection on phenotype can yield higher genetic response than selection on BLUP. When response of BLUP and phenotypic selection are graphed against inbreeding, they are similar to those observed by Wilson (1974).

Jeyaruban *et al.* (1995) compared index selection to BLUP using an additive infinitesimal computer simulation of a layer breeding program. They also concluded that BLUP always gave greater gains than classical selection index but at a cost in terms of inbreeding. They concluded that because commercial poultry production is based primarily on crossing programs, inbreeding in the parental lines can be recovered in the cross. However, as with the results of by Quinton *et al.* (1992) and Quinton and Smith (1994), the additive infinitesimal model ignores loss of favorable genes due to drift, which is directly proportional to the rate of inbreeding (Muir,

1991). The probability that a favorable allele will be lost in both lines is also proportional to the level of inbreeding in each and becomes a permanent loss that cannot be overcome by crossing lines.

Quinton *et al.* (1992) suggested that another strategy may be to select for maximum response, irrespective of inbreeding, in two or several subpopulations, then intercross the subpopulations periodically and start a new subset of subpopulations. This strategy has been termed the subdivide-merge scheme of breeding. Smith and Quinton (1993) examined that suggestion using an additive infinitesimal simulation and concluded that, in general, selecting in sublines and crossing gave less genetic response at the same level of inbreeding than selection in a single line.

From their results, Quinton *et al.* (1992) concluded that a revision of the methodology and a reappraisal of the results of selection theory and optimization of net genetic response rates are required. Maximizing accuracy of selection does not necessarily translate into greater overall response to selection. Quinton and Smith (1995) examined alternative methods to reduce the rate of inbreeding with BLUP selection, including substituting false high heritabilities to reduce selection of families. They concluded that the optimism about maintaining high response at low inbreeding by varying the heritability in BLUP or by adjusting the Expected Breeding Value (EBV) in selection for the relationship with individuals already selected is not warranted. Quinton and Smith (1995) found that the same result could be achieved by simply increasing the number of sires selected. Because they were using an infinitesimal model, they could not include the effects of drift on loss of favorable alleles. Therefore, the long-term effects of selection based on family indices or BLUP could be even worse than they predicted.

INTENSITY OF SELECTION AND SELECTION RESPONSE

In order to examine the relationship between selection intensity, population size, and gene action, a gene level simulation program was developed. This program allows for any number of loci, number of alleles per locus, gene action, proportion selected of each sex, initial gene frequency distribution, gene effect distribution, environmental effects distribution, population structure and size, and mutation rate. For these simulations, gene effects for each allele at a locus were assigned randomly from a uniform distribution with a mean zero and unit variance. With additive effects, the genotype was simply the sum of the allelic effects over both alleles and all loci affecting the trait. For dominance, allelic effects were assumed to be directional, in the direction of the favorable allele, and complete, i.e., the heterozygote was assigned the same genotypic effect as the favorable homozygote. The favorable homozygote was assumed to be in the positive direction. These effects were then summed over loci. Directional dominance was assumed

TABLE 1. Number (N), proportion (p) selected in each line, and selection intensity (i) for populations of size 720 (equal number of each sex selected)

Line	N	p (%)	i
1	36	5	2.06
2	72	10	1.76
3	120	17	1.49
4	180	25	1.27
5	360	50	.80

in order to simulate inbreeding depression. Inbreeding depression is the combined consequence of increased homozygosity and directional dominance (Falconer, 1981:227).

The genetic variance was computed as the variance among genotypes. Given the genetic variance, environmental effects were generated from a normal distribution with a mean of zero and a variance set to give the desired specified heritability. The phenotype was assumed to be the sum of the genotypic and environmental effects, i.e., no genotype-environment interactions.

The program only allows for three traits, but the number of loci affecting each trait, either independently or jointly (pleiotropy), can be specified, as can the degree of environmental correlation between traits. Selection methods are currently limited to index (multi-trait), tandem, independent culling, and multi-stage index. The program is currently under development to include information from relatives, both classical family index and BLUP.

For the research presented here, simulations were limited to one trait, the number of loci was set to 500, with 2 alleles per locus, and no mutations. Initial gene frequency was randomly set at each locus between 0 and 1 based on the uniform distribution. Gene effects for each allele at each locus were different and based on a unit normal distribution. Information was limited to the individual's own performance in each case. Each simulation was replicated 50 times starting over with different initial gene frequencies and genetic map. Two types of gene action were considered: 1) completely additive, and 2) highly heterotic with 100% of the loci showing directional dominance. The initial narrow sense heritability in both cases was 0.2 but the broad sense heritability of the latter was 0.3. Population sizes were set to 720, with an equal number of individuals of both sexes evaluated. It was assumed that the trait could be directly measured on the phenotype of both sexes, as with 7-wk weight in broilers. The number of breeders selected was the same for both sexes as shown in Table 1. The number of generations was set to 20, an intermediate time horizon achievable within the lifetime of the geneticist directing the program.

Robertson (1960) showed that for an additive trait 1) the expected limit of individual selection is proportional to the product of number selected (N) and selection intensity (i) (Ni), 2) for low values of Ni , the total

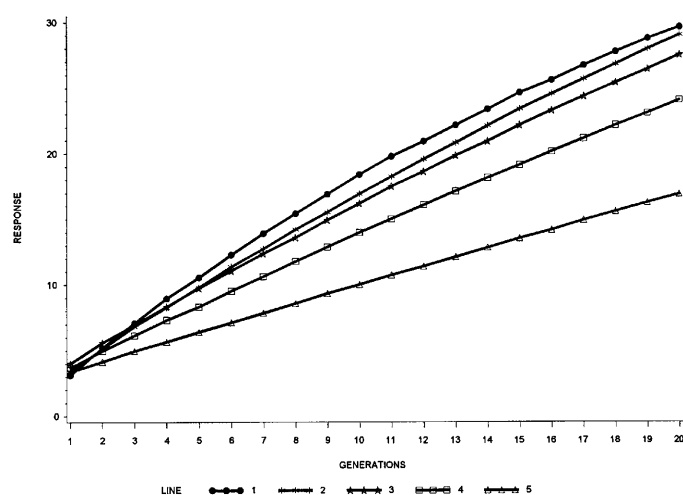


FIGURE 1. Response to selection for an additive trait in populations of size 720 with proportions selected of 5% (Line 1), 10% (Line 2), 17% (Line 3), 25% (Line 4), and 50% (Line 5).

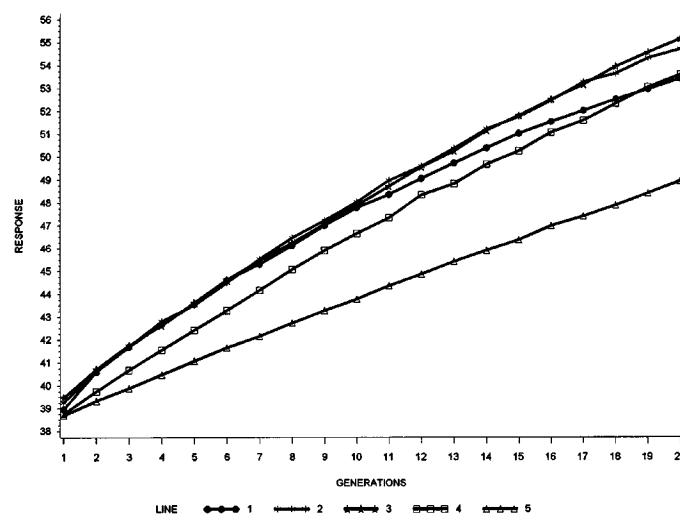


FIGURE 2. Response to selection for a heterotic trait in populations of size 720 with proportions selected of 5% (Line 1), 10% (Line 2), 17% (Line 3), 25% (Line 4), and 50% (Line 5).

advance by selection is $2N$ times the gain in the first generation (ΔG), and 3) the maximum response to selection is achieved by using a selection proportion of 50% each generation for each sex. Table 2 gives N_i and $2N\Delta G$ for each selection intensity. Robertson's equation would therefore predict that Lines 1 and 5 would, respectively, give the worst and best overall response to selection.

Figures 1 and 2 give simulated responses to selection for those lines for additive and dominant gene action respectively. Responses in the 20th generation are tabulated in Table 2. For an additive trait, actual responses were in the reverse ranking of the predictions given by Robertson (1960). This result does not indicate that the predictions of Robertson were wrong. With advanced generations, and additive gene action, the lines would eventually crossover and rank as predicted (not shown). However, in the short term, with additive gene effect, the optimum selection strategy would have been to maximize selection intensity. Thus, with additive traits, the only consequence of inbreeding is the loss of favorable alleles, which does not affect the breeding program appreciably in the short term.

Results with a heterotic trait were remarkably different. Optimal selection response was obtained with an intermediate selection intensity. With dominant gene action, inbreeding has both a long- and short-term consequence. The long-term consequence is the same as with additive gene action; however, in addition, mean performance is reduced through inbreeding depression with immediate effects in the short term.

It is expected that as population size increases, for a constant selection proportion, the rate of inbreeding will be less and thus also will be inbreeding depression. In poultry operations the maximum cull rate is approximately 90% for females and 97.5% for males because of reproductive limitations. If the usual size of a poultry layer operation is in the order of 10,000 (personal communication, Alan Emsley, ISA Babcock, Ithaca NY 14851-0280), the effective population size is approximately $N = 800$ and short-term consequences of inbreeding depression can be ignored. Thus, in the short to intermediate term, for large populations, the optimal selection intensity will be similar to that of additive traits.

TABLE 2. Response to selection in populations of size 720 for additive and heterotic traits after 20 generations (GEN), and the selection limit as predicted by $2N\Delta G$ and N_i ¹

Gene action							
Additive				Dominant			
Line	GEN 20	$2N\Delta G$	N_i	Line	GEN 20	$2N\Delta G$	N_i
1	29.7	142	74	1	53.8	97	74
2	29.0	244	126	2	54.7	175	126
3	27.5	355	179	3	55.9	274	179
4	24.1	442	228	4	53.6	342	228
5	16.9	565	288	5	49.5	432	288

¹ N = number selected; ΔG = gain in the first generation; i = selection intensity.

CONCLUSIONS

The optimal selection strategy partially depends on the trait being selected and is not necessarily that which either maximizes selection intensity or accuracy. In all programs, primary attention must be given to the rate of inbreeding, which is increased by increasing either accuracy of selection or selection intensity. Inbreeding reduces response to selection in two ways. First, for both additive and nonadditive traits, inbreeding is a measure of the amount of random genetic drift that has occurred. Genetic drift causes loss of favorable alleles. Once lost, those alleles can never be recovered and thus genetic drift lowers the selection limit. Second, for heterotic traits, inbreeding causes a depression of the mean caused by directional dominance.

Taking into account both factors, it is uncertain to what extent one should utilize family selection indices such as BLUP. Further research is needed in that area. However, the optimal selection intensity in the short to intermediate term for small populations is different for heterotic and additive traits. For additive traits, and in the short term (20 generations), one should maximize selection intensity. For heterotic traits, an intermediate proportion (25% of each sex) gives optimal response. For large effective population sizes, short-term consequences of inbreeding depression can be ignored.

REFERENCES

- Bulmer, M., 1971. Effect of selection on genetic variability. *Am. Naturalist* 105:201–211.
- Campo, J., and P. Tagarro, 1977. Comparison of three selection methods for pupal weight in *Tribolium castaneum*. *Ann. Genet. Sel. Anim.* 9:259–268.
- Doolittle, D., S. Wilson, and L. Hulbert, 1972. A comparison of multiple trait selection methods in the mouse. *J. Hered.* 63:366–372.
- Falconer, D. 1981. *Introduction to Quantitative Genetics*. 2nd ed. Longerman, London, UK.
- Garwood, V., and P. Lowe, 1979. Comparison of individual, sire family and index selection for short term rate of egg production in chickens. *Poultry Sci.* 58:751–753.
- Garwood, V. A., P. Lowe, and B. Bohren, 1980. An experimental test of the efficiency of family selection in chickens. *Theor. Appl. Genet.* 56:5–9.
- Jeyaruban, M., J. Gibson, and R. Gowe, 1995. Comparison of index selection and best linear unbiased prediction for simulated layer poultry data. *Poultry Sci.* 74:1566–1576.
- Kinney, T., Jr., B. Bohren., J. Craig., and P. Lowe, 1970. Responses to individual, family or index selection for short term rate of egg production in chickens. *Poultry Sci.* 49:1052–1064.
- Lin, C., 1978. Index selection for genetic improvement of quantitative characters. *Theor. Appl. Genet.* 52:49–56.
- Muir, W., 1991. Estimation of response to selection in non-replicated populations. Pages 95–121 in: *Proceedings of the 41st National Breeders Roundtable*, St. Louis, MO.
- Perez, M., And M. Toro, 1992. Classical and mixed-model analysis of an index selection experiment for fecundity in *Drosophila melanogaster*. *J. Anim. Sci.* 70:2673–2681.
- Quinton, M., and C. Smith, 1994. Choice of evaluation selection system to maximize genetic response at the same level of inbreeding. *Proc. 5th World Congr. Genet. Appl. Livest. Prod.* 18:147–151.
- Quinton, M., and C. Smith, 1995. Comparison of evaluation-selection systems for maximizing genetic response at the same level of inbreeding. *J. Anim. Sci.* 73:2208–2212.
- Quinton, M., C. Smith, and M. Goddard, 1992. Comparison of selection methods at the same level of inbreeding. *J. Anim. Sci.* 70:1060–1067.
- Robertson, A., 1960. A theory of limits in artificial selection. *Proc. R. Soc. B.* 153:234–249.
- Robertson, A., 1961. Inbreeding in artificial selection programmes. *Genet. Res.* 2:189.
- Smith, C., and M. Quinton, 1993. The effect of selection in sublines and crossing on genetic response and inbreeding. *J. Anim. Sci.* 71:2631–2638.
- Sorensen, D., and B. Kennedy, 1984. Estimation of response to selection using least-squares and mixed model methodology. *J. Anim. Sci.* 58:1097–1106.
- Wilson, S., 1974. An experimental comparison of individual, family and combination selection. *Genetics* 76:823–836.