

SHORT PAPER

A simple method for increasing the response to artificial selection

BY MIGUEL ANGEL TORO* AND BLANCA MARIA NIETO**

* *Departamento de Genética Cuantitativa, I.N.I.A., Carretera La Coruña KM7, Madrid-35, Spain*

** *Colegio Universitario Integrado, Universidad Complutense, Arcos de Jalón s/n, Madrid-17, Spain*

(Received 31 October 1983 and in revised form 5 July 1984)

SUMMARY

Selection theory usually assumes an equally probable contribution of each selected individual to a large 'gene pool' from which the individuals to be measured in the next generation are sampled. With unequal contributions it is possible to find several sets of values for N (the number of selected individuals) and f_i (probability of contribution of the i th individual) such that the same selection intensity is attained. It is suggested that the set of values producing minimum genetic drift should be chosen in order to increase the long-term response without any reduction in the short-term advance.

1. INTRODUCTION

A general relationship between the selection imposed on a measured character and the selective advantage of an allele has been derived by Kimura & Crow (1978). They compared the efficiency of truncation selection with less extreme alternatives, namely one in which fitness is represented by an S-shape curve as in the integrated normal distribution. They concluded that although truncation is the most efficient form of directional selection, substantial departures from strict truncation do not greatly reduce efficiency.

In this note it is suggested that this result can be used to improve the efficiency of artificial selection. If several fitness functions result in the same selection differential we should choose that leading to the largest effective population size (N_e) because the long-term response will then be increased without any reduction in the short-term advance.

2. THEORY

Consider a line undergoing artificial selection for an additive character with discrete generations. Assume that the population is monoecious and that the M individuals to be measured are sampled from a large gene pool with a gene frequency q for the desirable allele. The top N phenotypes are selected as parents to form the gene pool of the next generation. Suppose that the probability of contribution of a parent is f_i ($\sum^N f_i = 1$) and the frequency of the desirable allele in that individual is q_i , where $q_i = 0, 0.5$ or 1 . The mean and the variance of the gene frequency in the newly formed pool, q' , ignoring quadratic terms, will be given by

$$E(q') = E(\sum^N f_i q_i) = \sum^N f_i E(q_i) = q + q(1-q) \sum^N f_i x_i, \quad (1)$$

$$V(q') = \frac{q(1-q)}{2} \sum^N f_i^2, \quad (2)$$

Table 1. Number of selected parents (N') out of 20 scored and their probability of contribution (f_i , $i = 1$ to 10)

	$N' = 5$	$N' = 6$	$N' = 7$	$N' = 8$	$N' = 9$	$N' = 10$
f_1	0.20	0.24	0.26	0.27	0.27	0.27
f_2	0.20	0.19	0.20	0.20	0.20	0.20
f_3	0.20	0.17	0.16	0.16	0.15	0.15
f_4	0.20	0.15	0.13	0.12	0.12	0.12
f_5	0.20	0.13	0.11	0.10	0.10	0.10
f_6	—	0.12	0.09	0.07	0.07	0.07
f_7	—	—	0.07	0.05	0.05	0.05
f_8	—	—	—	0.03	0.03	0.03
f_9	—	—	—	—	0.01	0.01
f_{10}	—	—	—	—	—	0.00
N_e	5.00	5.67	5.71	5.84	5.94	5.94

where $2u$ is the difference in genotypic value between the two homozygotes at the locus considered measured in standard deviation units and x_i is the expected value of the i 'th order statistics of a standard normal distribution, $\Sigma^N f_i x_i$ being the selection intensity (i).

The effective population size for the genes affecting the metrical character, using formula (2) in reverse, is given by

$$N_e = 1/\Sigma f_i^2 \tag{3}$$

Two selection methods with the same selection intensity can now be compared. Typically the top N individuals are selected and their probable contributions are made identical. Alternatively a larger number of individuals N' are selected, but their probable contributions are made unequal depending on performance. The optimal value for N' and the f_i 's will be calculated minimizing $\Sigma^N f_i^2$ (i.e. maximizing N_e), subject to the restriction $\Sigma^N f_i x_i = i$. This can be easily done using quadratic programming techniques.

An example will illustrate the method. Suppose that five individuals out of 20 in each generation are selected. If the probability of contribution of each selected individual to the next generation is the same, then the distribution of family sizes at measurement will be Poisson, with a selection intensity of 1.214. When the probability of contribution of each parent f_i is not the same, the distribution of family sizes at measurement will be multinomial, with the selection intensity now $\Sigma^N f_i x_i$. Table 1 shows for $N' = 5, 6, 7, 8, 9$ and 10, the appropriate f_i values for a selection intensity of 1.214. In the bottom line effective population sizes (3) are given. The set of values of N' and f_i leading to the largest N_e must be chosen in order to increase the selection limit. Although in principle the best strategy would be to select all individuals measured ($N' = 20$), the probabilities of contribution of other individuals than the top nine would be negligible. In practice, the optimal contribution of individuals with phenotypes below the mean should be zero, because, as suggested by Smith (1969), it will be better to use unscored individuals than individuals which are below average.

Although the sampling method of progeny from selected individuals has been assumed to be multinomial with probability f_i for parent i , in practice one would not sample by specifying probabilities but by controlling offspring numbers. In this case a fixed distribution of family sizes as close as possible to $w_i = 2Mf_i$ will be imposed. Although formula (3) will not then be strictly valid, the general principles remain unchanged.

3. DISCUSSION

In the previous analysis, the population was assumed to be monoecious. With separate sexes, when only one of them is scored and selected, the application of above methodology is straightforward. When both sexes are considered it would be desirable to introduce

assortative mating in order to increase the difference among the average values of the parents for the metric trait.

The comparison between the two methods has also been investigated by Monte Carlo simulations showing that, for a wide range of genetic parameters, long-term response improvements of the order of 5% to 20% can be achieved.

There is an important factor that has not been taken into account. In a population under artificial selection the effective population size for genes which do not affect the character under selection may be less than the actual number because parents do not have equal probability of contributing to the next generation (Robertson, 1960). This effect will be greater if the *a priori* expected contributions (f_i) of the selected parents are intentionally made unequal, but it is difficult to quantify the magnitude of that effect. Simulation results show that, in general, the expected larger reduction in population effective size for neutral genes due to Robertson's effect will be overcompensated by the difference between N' and N . This will therefore imply that if the different expected contributions of offspring are taken into account in the design of a breeding programme an increase in the long term response together with a lower fitness deterioration caused by inbreeding could be expected.

We are grateful to Luis Silvela and Carlos López-Fanjul, for critically going through the script and giving valuable suggestions. We also thank an unknown referee for valuable comments.

REFERENCES

- KIMURA, M. & CROW, J. (1978). Effect of overall phenotypic selection on genetic change at individual loci. *Proceedings of the National Academy of Sciences U.S.A.* **75**, 6168–6171.
- ROBERTSON, A. (1960). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.
- SMITH, C. (1969). Optimum selection procedures in animal breeding. *Animal Production* **11**, 433–442.