

Computer Simulation of within Family Selection in Finite Populations

S.S.Y. Young and R.V. Skavaril

Department of Genetics, The Ohio State University, Columbus, Ohio (USA)

Summary. The relative efficiency, in terms of selection limits, between mass selection and within family selection was compared by computer simulation methods. A 20-locus additive model was used to simulate a quantitative trait under selection. It was assumed that 50-75 percent of the genetic variance in the base population was controlled by four major genes initially at low frequencies.

In populations of size $N = 100$ no loss of major genes was found when either method of selection was used. When $N = 50$ within family selection was generally superior to mass selection but when $N = 10$ the situation was reversed. For $N = 30$ within family selection was more efficient only under high selection intensity or high heritability situations.

Introduction

In his paper on artificial selection in plants and animals Robertson (1966) concluded that if a selection response was mostly due to fixation of genes which were at low frequencies in the initial population, then a restriction of the population size would lower the limit of selection. Such lowered limit is due to chance fixation, and the probability of fixation is dependent on the effective population size, N_e , and the effective intensity of selection, s . (Robertson 1960). When the value of the product of N_e and s is small, a favorable gene might become extinct by chance.

Since, in selection work, the population size is frequently small, it is of interest to explore selection methods which would tend to minimize chance fixation. One possible method is within family selection. This method is frequently used as a way to eliminate possible non-genetic variations between families (Falconer 1960). It is known that within family selection is less efficient, in terms of genetic gain per generation (Hill 1971), as compared with mass selection; but within family has the advantage of having twice the effective number, when the population size is held constant (Gowe, Robertson and Latter 1959). However, since within family selection utilizes only one half of the genetic variance, Robertson (1960) pointed out that the advantage of a larger N_e would be negated by the smaller effective selection pressure, s , and doubted that within family selection would affect the selection limit. Since Robertson's theory was based on a single locus model, he cautioned that,

in its application to a quantitative trait, it was necessary to consider the assumptions involved; these include the assumption of the constancy of the phenotypic variance.

Recently, Dempfle (1975) investigated the effect of within family selection on selection limits, based on an additive model. In this model it was assumed that the quantitative trait was controlled by many loci each with a small effect. He showed that within family selection was more efficient than mass selection when the heritability was very high. The theoretical result was verified by a simulation study using a model of 10 to 256 loci and small population size of 4 to 12 individuals. The results showed that within family selection was more efficient when the gene frequencies were low and when the correlation between the selection criterion and the additive genetic value was high. The main reason for the superiority of within family selection was due to a relatively lower decay of the additive variance during selection.

If a quantitative trait were controlled by a large number of genes, each with a small effect, then the loss of a small fraction of these should not lead to an appreciably lower selection limit. However, experiments by Thoday and his colleagues (Thoday 1961, Spiekett and Thoday 1966), Robertson (1966), Jones (1967) in *Drosophila* and Wehrhahn and Allard (1965) in wheat have suggested that a high proportion of genetic variabilities in quantitative traits might be controlled by a few loci with large effects. In these cir-

cumstances, the loss of even a single major gene would be expected to lead to a sizable reduction in the ultimate genetic gain.

In a theoretical consideration James (1971) showed that when the major allele was of low frequency, the effect of population size varied with $N_e s$. Thus, it seems possible that when a quantitative trait is controlled by a few major genes of low frequencies, the efficiencies of mass selection and within family selection on selection limits may be dependent also on genetic models. This paper reports an investigation, using computer simulation methods, of the relative efficiency of the two selection methods in terms of limits of selection using a model incorporating genes with large effects on quantitative traits.

The Parameters and the Computer Runs

It is assumed that the quantitative trait under selection is controlled by 20 loci, with two alleles segregating at each locus. The 20 loci are distributed equally in four linkage groups. Of the five loci in each group, one is designated as a major locus where a gene with a large effect is located. The genes are assumed to be in linkage equilibrium and have only additive effects. Interactions between loci are assumed to be absent.

Under the additive model, the amount of the genetic variance controlled by major loci can be calculated simply as a function of the gene effects and frequencies. In the present work, two additive models are used; and these are:

A. When 75 percent of the genetic variance in the initial population is controlled by the 4 major loci. In this case, the initial frequencies of the minor genes are 0.5 and those of the major genes are 0.05 for all linkage groups. The effect of the major gene, β , is assumed to be eight times that of a minor gene.

B. When 50 percent of the genetic variance in the initial population is controlled by the major loci. The frequencies for the major and minor genes are the same as in A but the effect of the major gene is assumed to be only five times that of a minor gene.

Hence, the two models simulate the situation when a quantitative trait is mainly controlled by a few major genes, initially present at low frequencies. Other parameters used in the simulation programs are (1) four population sizes: $N = 100, 50, 30$ and 10 ; (2) two initial levels of heritabilities: $h^2 = 0.3$ and 0.1 ; (3) two recombination probabilities: $r = 0.3$ and 0.05 (the r values are assumed to be the same for all adjacent loci in the linkage groups); (4) three selection intensities: $I = 10\%$ ($2/20$), 20% , ($2/10$) and 50% , ($2/4$). For example, when $I = 20\%$, each mating produces 10 offspring and during mass selection, the best 20 percent of the total population are selected as parents of the next generation of individuals. During within family selection for the same selection intensity two best individuals within each family are selected. The selected individuals are allowed to mate at random. The computer programs written for this study were written in PL/1 for the IBM 370/165 and

are essentially the same as that described by Young (1966) with two exceptions and these are: (1) meiosis is simulated by the method of binary masking (Fraser and Burnell 1970); (2) mass selection or within family selection is carried out in each generation within each population. In mass selection, the best individuals are selected by truncation, while in within family selection, the best two individuals (a "male" and a "female") are selected within each family regardless of the family mean.

The computer runs were organized into different series. Each series consisted of 10 replicates and was characterized by a set of parameters (N, I, r, h^2), one gene model, and one selection method. Each replicated run, using a different random number sequence, simulated a population under selection for 15 generations.

The first 70 runs involving smaller populations showed that the loss of major genes almost always occurred before generation 8 (with a mean generation number of about 3). This was true even when the population was small ($N = 10$) and the selection pressure low ($h^2 = 0.1, I = 50\%$). In later generations, the frequencies of major genes were always reasonably high. Nevertheless, as a conservative measure, selection was continued up to generation 15.

During the course of this study, it became clear, that in some situations, for example when $N = 100$, further, additional series of runs would not be expected to provide new information. For this reason, some combinations involving $N = 100$ and those involving different r values have been omitted.

Results

The large population

Initial runs of several series for $N = 100$ were made. As expected, the genetic gain per generation is higher in mass selection than in within family selection. Some typical results are presented in Fig. 1.

After 12 generations of selection, the population means attained by mass selection are higher than those from within family selection. However, in all runs, no major genes in any of the populations were lost, this was true, even when the heritability was low (0.1); hence the simulations were terminated at generation 12. Since the results were uninteresting from the point of view of selection limits, no further runs involving the large population size were made.

Intermediate populations

Preliminary results for $N = 50$ and $N = 30$ showed that in a number of populations, under both methods of selection, one or more major genes had been lost during 15 generations of selection. Figure 2 presents four sets of typical results showing effects of

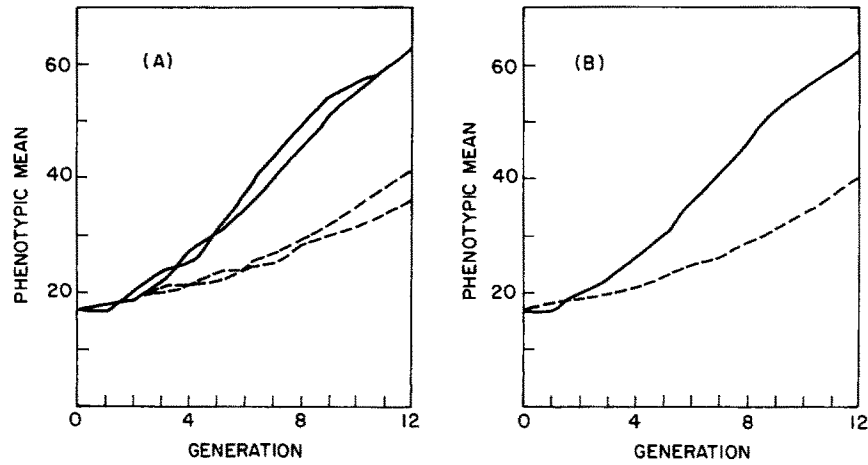


Fig. 1. Genetic advance under mass selection and within family selection when $N = 100$, $\beta = 5$, $r = 0.3$, $I = 0.10$, and $h^2 = 0.1$. (A) Results of two replicated populations. (B) Mean values of ten replications. --- Mass selection. --- Within family selection

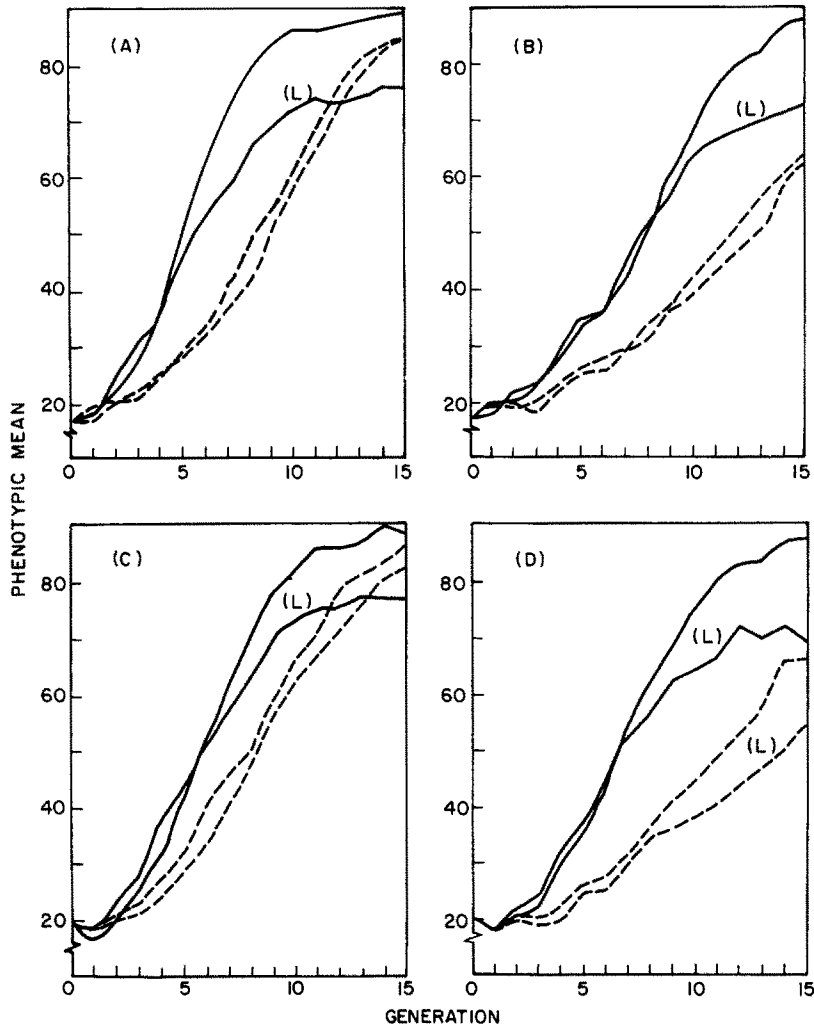


Fig. 2. Typical replicates showing effect of the loss of a major gene during selection. For all populations $I = 20$, $\beta = 8$. (A) $N = 50$, $h^2 = 0.3$. (B) $N = 50$, $h^2 = 0.1$. (C) $N = 30$, $h^2 = 0.3$. (D) $N = 30$, $h^2 = 0.1$. — Mass selection. --- Within family selection. (L) Major gene lost

Table 1. Average numbers and standard errors of major genes lost in 15 generations of selection for $N = 30$

Selection method	Selection intensity (%)	Heritability	Recombination probability	
			0.3	0.05
Mass selection	10	0.1	0.1 ± 0.10	0.1 ± 0.10
	20	0.1	0	0.4 ± 0.22
	50	0.1	0.6 ± 0.16	0.5 ± 0.17
	10	0.3	0.1 ± 0.10	0
	20	0.3	0.1 ± 0.10	0.1 ± 0.10
	50	0.3	0.5 ± 0.17	0.2 ± 0.13
Within family selection	10	0.1	0.2 ± 0.13	0
	20	0.1	0.1 ± 0.10	0.2 ± 0.13
	50	0.1	0.9 ± 0.28	1.2 ± 0.20
	10	0.3	0	0
	20	0.3	0	0
	50	0.3	0.4 ± 0.22	0.5 ± 0.22

Table 2. Analysis of variance of the number of genes lost in 15 generations of selection for $N = 30$

Source of variation	Degrees of freedom	Mean square
Selection methods (M)	1	0.204
Selection intensities (I)	2	6.862***
Heritabilities (H)	1	2.204***
Recombinations (R)	1	0.038
MI	2	0.879*
MH	1	0.504
IH	2	0.679*
MR	1	0.037
IR	2	0.162
HR	1	0.204
MIH	2	0.179
MIR	2	0.462
MHR	1	0.104
IHR	2	0.179
MIHR	2	0.029
Residual	216	0.196

the loss of a single major gene. The drastic effect due to a loss of major gene is, of course, a direct consequence of the models used in this study; and it can be calculated that a single loss represents a decrease of 14-17 percent in selection limits. Therefore, it is pertinent to present the data in terms of the average gene loss, in 10 replicates, during a fixed period of selection.

The first block of runs was designed to investigate the effect of linkage between major and minor genes on the relative efficiency of the selection methods. From the preliminary runs, it was found that frequent losses occurred in both selection schemes when $N = 30$. The effect of linkage was therefore investigat-

ed in this population size. The average numbers and standard errors of major genes which have become extinct in 15 generations of selection are presented in Table 1. The data summarized in Table 1 were obtained from computer runs specified by $\beta = 8$, $h^2 = 0.1$ and 0.3, and $I = 10\%$, 20%, and 50%. The recombination values of 0.3 and 0.05 were chosen to represent two fairly extreme situations.

The means listed in Table 1 show that both selection intensities and levels of heritability have a major influence on gene loss. As expected, low heritability or low selection intensity results in a greater loss. The effect of the selection intensity is particularly striking when a low intensity ($I = 50\%$) is coupled with within family selection. Results of individual runs, in terms of the number of major genes lost during selection, were analyzed by a 3-way ANOVA assuming a fixed model (Table 2). The main effects due to selection intensities, heritabilities, and the interactions between selection intensities and methods and selection intensities and heritabilities are all significant statistically. Tables 1 and 2 also show that the effect due to recombination probabilities between major and minor loci, is small and negligible. In addition, all interaction terms involving linkage are small and non-significant. Hence, it may be concluded that, under the present model, linkage has a negligible effect on selection limits. Another set of computer runs with $\beta = 5$, $h^2 = 0.1$, $N = 30$, $r = 0.5$, 0.3 and 0.05 gave essentially similar results. These data are, therefore, not presented. For these rea-

Table 3. Average and standard error of number of major genes lost during 15 generations of selection in 10 replicates

Selection method	Selection intensity (%)	Heritability	(A) N = 50		(B) N = 30		(C) N = 10	
			Major gene $\beta = 5$	effect $\beta = 8$	Major gene $\beta = 5$	effect $\beta = 8$	Major gene $\beta = 5$	effect $\beta = 8$
Mass Selection	10	0.1	0.2±0.13	0.2±0.13	0.3±0.15	0.1±0.10	1.0±0.26	0.6±0.27
	20	0.1	0.2±0.13	0	0.3±0.21	0	0.8±0.25	0.9±0.23
	50	0.1	0.4±0.22	0.2±0.13	0.8±0.25	0.6±0.16	0.8±0.25	2.4±0.30
	10	0.3	0	0	0.1±0.10	0.2±0.13	0.4±0.27	0.1±0.10
	20	0.3	0	0	0.4±0.16	0.1±0.10	0.4±0.22	0.4±0.22
	50	0.3	0.3±0.21	0.1±0.10	0.4±0.22	0.5±0.17	1.3±0.26	1.0±0.30
Within family selection	10	0.1	0.1±0.10	0	0.2±0.13	0.2±0.13	1.3±0.21	2.7±0.30
	20	0.1	0.1±0.10	0	0.2±0.13	0.1±0.10	2.5±0.37	1.7±0.33
	50	0.1	0.2±0.21	1.0±0.10	1.4±0.30	0.9±0.28	2.6±0.34	1.1±0.23
	10	0.3	0	0	0	0	2.3±0.33	0.5±0.17
	20	0.3	0	0	0	0	0.5±0.17	1.3±0.21
	50	0.3	0	0	0.6±0.22	0.4±0.22	0.7±0.26	0.3±0.15

Table 4. Analyses of variance of the number of major genes lost during 15 generations of selection

Source of variation	Degrees of freedom	Mean Square		
		(A) N = 50	(B) N = 30	(C) N = 10
Selection Methods (M)	1	0.704*	0.038	22.817***
Selection Intensities (I)	2	0.267	8.629***	0.987
Heritabilities (H)	1	0.704*	2.604**	35.267***
Gene Effects (β)	1	0.204	1.204*	1.067
MI	2	0.067	0.787	10.454***
MH	1	0.037	0.938*	4.817**
IH	2	0.017	0.954*	0.579
M β	1	0.004	0.004	3.750*
I β	2	0.067	0.129	0.454
H β	1	0.104	0.338	2.400
MIH	2	0.050	0.113	1.029
MI β	2	0.067	0.404	4.587**
MH β	1	0.004	0.004	0.817
IH β	2	0.017	0.088	6.612***
MIH β	2	0.017	0.029	13.829***
Residual	216	0.111	0.263	0.669

sons, subsequent simulation studies were made using only one, arbitrary linkage parameter, $r = 0.3$.

The means and standard errors of the numbers of genes lost during 15 generations of selection for $N = 50$ and $N = 30$ are shown in Tables 3(A) and (B), respectively. The ANOVA of results of individual runs are presented in Tables 4(A) and (B).

When populations are as large as 50, the overall loss in major genes is small. In general, within family selection is significantly superior to mass selection; the average gene loss in the former method is

about one third that of the latter. Indeed, when within family selection was used in a high heritability situation, no genes were lost in all replicates even when the selection intensity was low. The analysis (Table 4A) shows that the level of heritability is important; high heritability leads to a significantly smaller loss. No other effects appear to be important.

In the smaller populations of 30 individuals, the situation is quite different. Here the overall difference between selection methods is negligible, but selection intensity and heritability levels assume major

roles. Table 3(B) also shows that mass selection is superior when the heritability is low, while the relative merit is reversed when the heritability is high and the difference is significant statistically. The effect of major gene also has an influence on gene loss; a higher effect leads to a lower loss.

The small population

In populations of 10 individuals, Table 3(C), the situation is again different. In these small populations the overall loss is great for both schemes but mass selection is much superior over within family selection. The mean loss is 0.84 genes under mass selection but 1.46 under within family selection. The difference is very highly significant (Table 4C). In this population size, the level of heritability also has a major influence on selection limits. A high level of heritability results in about one half of the average loss as compared with the low heritability situations.

Of some interest to this study is the highly significant MI and MH interaction terms. The loss is relatively higher under intense selection in within family selection; but, under weak selection pressure, the loss in mass selection is slightly more severe. When heritability is low, mass selection is superior; but its superiority decreases when heritability is high. There are also highly significant second and third order interaction terms in Table 4C, suggesting changes in the relative efficiency with the various combinations of parameters. For example, when gene effect is large, within family selection appears to be more efficient when the selection intensity is low; and the situation is reversed when the intensity is high. Also, within family selection is more efficient when heritability is high and selection low. However, in the small population size mass selection is clearly superior in most genetic situations investigated here.

Discussion

The present results provide some comparisons on the relative efficiency, in terms of selection limits, between mass selection and within family selection. In addition, the information obtained in this study, which is based on a multi-loci quantitative model, may be used to compare with the theoretical conclusions derived by Robertson (1960) and James (1971) from a

single locus model. In general, there are good agreements between the simulated results and the theoretical predictions. In the simulated populations, larger population size and high effective selection pressure can each, and in combination, lead to higher selection limits. However, under the present model, if the population is as large as 100, then it appears that the limits of selection will be reached regardless of the selection pressure. But in intermediate and small populations, the effective selection pressure plays a major role in determining limits of selection.

The effective selection pressure is jointly determined by selection intensity, heritability, and gene effect. The level of heritability appears to be consistently important over all intermediate and small populations, while selection intensity and gene effect are only important in intermediate populations.

The relatively neutral role of linkage found in this study must be qualified by the fact that the present model ensures no linkage between major genes. Hill and Robertson (1966), using a two loci model, have shown that the chance of fixation was linearly related to linkage distance. Hence, if the present model assumed tight linkage between major genes, then the effect of linkage would be expected to be large.

The relative efficiency of selection methods is affected by population size. For population size of 50 individuals, within family selection is superior; but, in the small population of size 10, mass selection is clearly more efficient. For the intermediate population of size 30, the overall difference in efficiency between the two methods is small and non-significant. In this population size, within family selection appears to be the method of choice only when heritability is high. Thus, for populations of intermediate size, the relative efficiency is also affected by the effective selection pressure.

It should be reiterated that the efficiency discussed above refers only to limits of selection. The present results suggest that for a population of size 30-50, within family selection may lead to higher selection limits, particularly when the quantitative trait under selection is highly heritable or when the intensity of selection is high.

It would be pertinent to compare the current results obtained from a major gene model with those reported by Dempfle (*loc. cit.*) which were based on

a polygenic model. However, direct comparisons are difficult to make for several reasons. Firstly, relative to the current work, the simulation study based on the polygenic model, involved smaller populations and higher initial gene frequencies. Secondly, the current work uses the heritability as one of the parameters, while in the polygenic study a related function, the correlation between the selection criterion and additive genetic value (ρ) was used. Hence, in order to compare the results, it is first necessary to relate the two parameters explicitly.

It is our understanding that under mass selection $\rho^2 = h^2$. But under within family selection, if there were no common environmental effects for members of the same family (Hill 1971)

$$\rho_w^2 = \frac{(2-p)h^2}{4(2-h^2)}$$

where p is the portion selected. If our interpretation of ρ_w is correct, then the values of ρ_w^2 used in this study ranged from 0.02 to 0.08, which are much lower than the values of 0.40 to 0.90 used in the polygenic simulations. In fact, under the scheme of within family selection used here, ρ_w^2 has a limiting value of 0.5.

Despite these difficulties, it seems worthwhile to compare the results, at least qualitatively. One reasonable comparison is that between our results for $N = 10$ and those obtained from the polygenic model, since the latter involved only small populations. In the polygenic work when $\rho^2 = 0.4$ and when the gene frequencies are low or intermediate, within family selection is shown to be more efficient. But under the major gene assumption, mass selection is more efficient when the gene frequency is low. However, such a difference in efficiency may be due to the different levels of heritability used in the two studies, since both sets of results show that the relative efficiency of within family selection increases with higher

values of heritability. Despite this apparent difference both studies suggest that mass selection, in small population, is in general more efficient than within family selection when the heritability is low.

Acknowledgements

We wish to acknowledge the support of the Instruction and Research Computer Center of the Ohio State University for providing free computer time for this work.

Literature

- Dempfle, L.: A note on increasing the limit of selection through selection within families. *Genet. Res. Camb.* **24**, 127-135 (1975)
- Falconer, D.S.: *Introduction to Quantitative Genetics*. Edinburgh: Oliver and Boyd 1960
- Fraser, A.S.; Burnell, D.: *Computer Models in Genetics*. New York: McGraw-Hill 1970
- Gowe, R.S.; Robertson, A.; Latter, B.D.H.: Environment and poultry breeding problems. V. The design of poultry control strains. *Poult. Sci.* **38**, 462-471 (1959)
- Hill, W.G.: Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* **27**, 293-311 (1971)
- Hill, W.G.; Robertson, A.: The effect of linkage on limits to artificial selection. *Genet. Res. Camb.* **8**, 269-294 (1966)
- James, J.W.: The founder effect and response to artificial selection. *Genet. Res. Camb.* **16**, 214-250 (1971)
- Jones, L.P.: Effects of X-rays on response to selection for a quantitative character of *Drosophila melanogaster*. *Genet. Res. Camb.* **9**, 221-231 (1967)
- Robertson, A.: A theory of limits in artificial selection. *Proc. Roy. Soc. (B)* **153**, 243-249 (1960)
- Robertson, A.: Artificial selection in plants and animals. *Proc. Roy. Soc. (B)* **164**, 341-349 (1966)
- Spickett, S.G.; Thoday, J.M.: Regular responses to selection. 3. Interactions between polygenes. *Genet. Res. (Camb.)* **7**, 96-121 (1966)
- Thoday, J.M.: Location of polygenes. *Nature* **191**, 368-370 (1961)
- Wehrhahn, G.; Allard, R.W.: The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. *Genetics* **51**, 109-119 (1965)
- Young, S.S.Y.: Computer simulation of directional selection in large populations. I. The program, the additive and the dominance models. *Genetics* **53**, 189-205 (1966)

Received July 14/December 28, 1975
Communicated by L.D. Van Vleck

S.S.Y. Young
R.V. Skavaril
Department of Genetics
The Ohio State University
Columbus, Ohio 43210 (USA)