Genet. Res., Camb. (1972), **20**, *pp.* 75–99
With 6 text-figures
Printed in Great Britain

Population structure in artificial selection programmes: simulation studies

By F. E. MADALENA† and W. G. HILL
Institute of Animal Genetics, Edinburgh EH9 3JN

(*Received 11 January 1972*)

SUMMARY

A simulation study was undertaken of methods of subdividing populations into several small sublines and utilizing the variances generated between lines by selecting among them. Crosses of chosen lines were made, and either selection was continued in a single large population (single cycle) or the population was subdivided again (repeated cycles). As a control for the efficiency of these schemes, a single large population was maintained and selected at the same intensity from the outset. Simple models were used of additive or completely dominant genes, usually of equal effect and equally spaced on a single chromosome.

The single and repeated cycle structures give similar results, but the repeated cycle structure is more extreme.

With additive models intense selection between lines gives short-term advances, but causes a reduction in the limit when compared with a single population. The effect on the limit is greatest with free recombination, very small with complete linkage. If no selection is practised between lines the limit is unaffected, but takes longer to attain.

With complete dominance, and the recessive allele initially at low frequency, greater responses from selection are obtained within sublines than in the large population, large gains are made from selection between sublines, and a higher limit can be reached. If the recessive allele is at high initial frequency the subdivision is not beneficial.

Some simple theory is developed to explain these results. It is concluded that subdivision and crossing schemes are unlikely to be very useful except for elimination of deleterious recessive genes.

1. INTRODUCTION

In the ideal selection programme rapid response would be made from the outset, and would continue until all the useful genetic variation in the source material had been incorporated. Unfortunately these objectives are partly incompatible since selected populations are necessarily of finite size. Rapid short-term gains can be made by selecting a very small proportion of the population for breeding the next generation, but many favourable genes will be lost by chance and the limit will be reduced. Dempster (1955) and Robertson (1960) showed theoretically that for single genes the limit is maximized when 50 % of the population are selected each

† Present address: Facultad de Agronomia, Universidad de la Republica, Paysandu, Uruguay.

generation. When linkage effects are important, rather more than 50% should be chosen (Hill & Robertson, 1966; Robertson, 1970*a*). More intense selection should be practised if the total advance is to be maximized in a specified, finite, number of generations (Robertson, 1970*b*), or if higher economic weight is given to early response (James, 1972). But in an attempt to avoid the conflict between short-term and long-term gains we should look at other breeding systems, such as structured or subdivided populations.

The structure of Mendelian populations has long been recognized as an important factor in evolution (Wright, 1951). Its effects on the progress from artificial selection have received less attention, except in breeding plans designed to exploit non-additive variation for improvement of line crosses. However Baker & Curnow (1969) considered populations divided into small sublines, and compared the rates of response and variance between lines for different sizes of the sublines and for alternative genetic models. They predicted that useful gains could be made even with small sublines, and then considerable further response could be obtained by selection between lines. Wright (1939) proposed a structure of repeated cycles of subdividing the population and practising within and between-line selection and crossing. He considered this method would be effective in preventing the loss by recombination of favourable epistatic combinations in cross-fertilizing species, and with a model of multiple 'peaks' of desirability in relation to gene frequencies, drift could allow the population as a whole to move to new peaks after crossing (Wright, 1951). Baker & Curnow (1969) did not investigate the effects of reselection from line crosses.

Some relevant theory is known however. With a model of independent additive genes Robertson (1960) showed that if m replicate lines were selected to fixation with size N each, crossed together and selected as a single population with size Nm, the same final limit would be attained as in a single population selected throughout at the same intensity with size Nm. Maruyama (1970) generalized these results for additive genes by showing that any subdivision of the total population gives the same selection limit, regardless of when crossing or migration occurs, so long as this happens without a change in mean gene frequency in the total population, i.e. without selection between lines. This generalization can also be derived from a formula given by Pollak (1966). Robertson's (1960) result for crosses of fixed lines holds approximately with dominance, but the subdivision structure gives a slightly higher limit when the recessive allele is favoured, a slightly lower limit when the dominant allele is favoured.

However, in structures in which the population is subdivided into lines of smaller size, the additive genetic variance within lines and consequently the response to selection are reduced by random drift. Thus unless selection between lines is practised the limit will take longer to reach in a subdivided population, except perhaps if the variability derives from low-frequency recessive genes when the additive variance may increase with initial inbreeding (Robertson, 1952). Since inbreeding increases variability between lines which can be utilized accurately by selection of the lines on mean performance it may be possible to design subdivided systems to obtain higher rates of advance and perhaps limits than by selection in a single population.

Experimental studies of gains from artificial selection in population structures involving between-line selection have been made by Bowman & Falconer (1961), Hill (1963), Madalena (1970) and Goodwill (1971). While the results obtained in these experiments with different traits of various species are not the same, in no case are large gains obtained from between-line selection and crossing, relative to selection in single populations.

In this paper a theoretical study has been made of structures utilizing between-line selection similar to those proposed by Wright (1939), and a preliminary report has already appeared (Hill & Madalena, 1969). Although we have not considered epistatic loci, linkage has been included, so that we can carry further the results of Robertson (1960) and Maruyama (1970). Monte Carlo simulation techniques have been used throughout; simple approximations using selective values at a single locus are not adequate, for the selective value at the locus during between-line selection is very much affected by segregation at the other loci.

In all comparisons which we make between selection schemes, the same total number of individuals (Q) are recorded each generation, either in one population with Q measured, or, say, 8 with $Q/8$ measured in each. Only in this way can a fair comparison between alternatives be made in terms of expense of measurement or utilization of facilities. However, we ignore biological difficulties, such as a decline in reproductive performance due to inbreeding.

2. METHODS

(i) *Design of population structures*

The structures studied are shown diagrammatically in Fig. 1. These are the *single-cycle structure* (Fig. 1*a*) in which one cycle of subdivision into small lines and intercrossing of selected lines is followed by selection thereafter in a single large population; and the *repeated-cycle structure* (Fig. 1*b*) in which a new set of lines are started from the intercross of the initial lines and the same procedure of inbreeding and crossing repeated.

In both systems the first cycle started at generation 0 with sampling of M individuals at random into each of m replicate lines from a base population in Hardy–Weinberg and linkage equilibrium. These M individuals were scored for a quantitative trait which was a function of their genotype and environmental error. The best N were chosen by truncation selection to be parents of the next generation and M progeny were bred.

Selection at this intensity (N/M) was continued for T generations. At generation T between-line selection was practised on the mean phenotype of the M individuals in the line, and the best v from the m lines chosen. In these v lines, within-line selection was again practised at the same intensity as before to give N individuals in each, a total of Nv, for crossing. These Nv individuals were randomly mated and selfed as if they were a single population to give a total of Q progeny. Thus both cross and 'pure' line progeny were formed, with the total number of chromosomes sampled from any line following a multinomial distribution. To allow recombination among

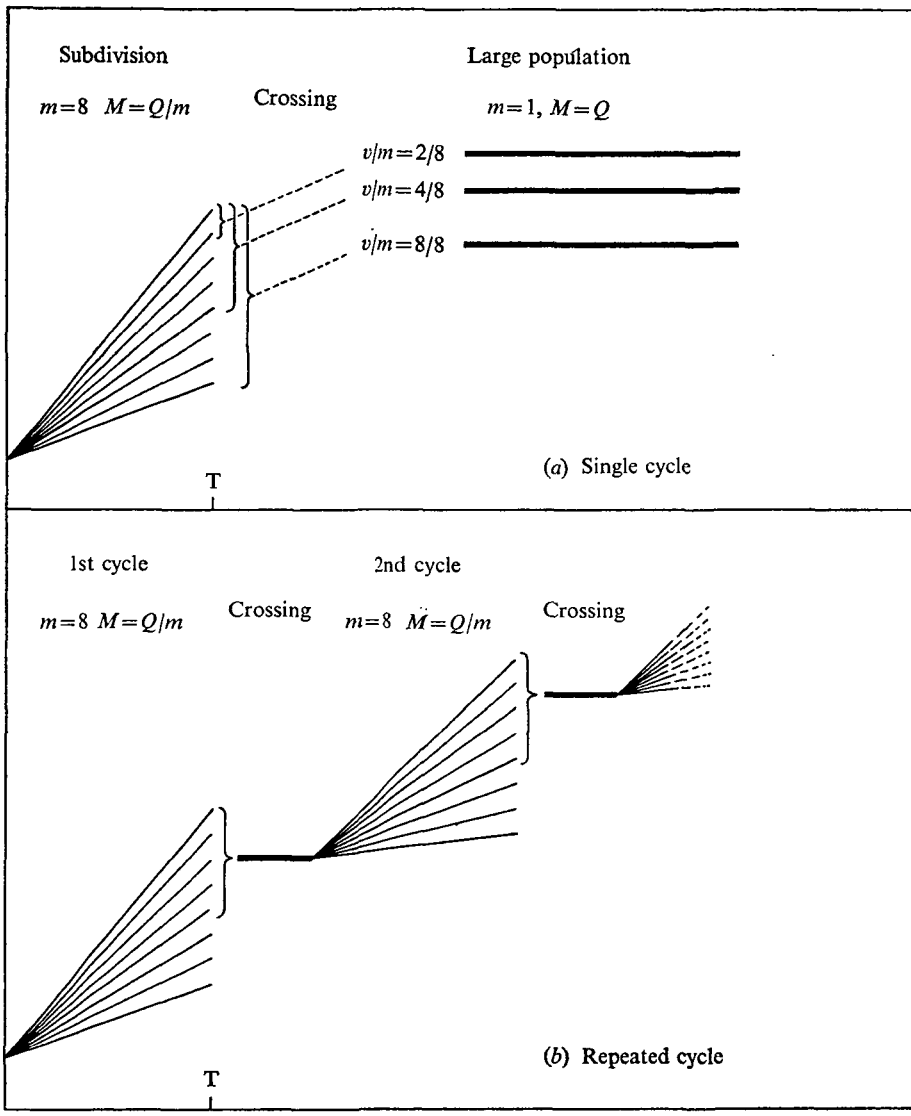

Fig. 1. The structures studied. (*a*) Single cycle: subdivision of the Q individuals measured into m lines of $M = Q/m$ individuals each, selection within lines of N individuals (a proportion N/M) for T generations. At generation T, selection between lines and crossing v selected lines to form a single population with Q individuals recorded and Nm selected (again a proportion N/M) until fixation. (*b*) Repeated cycle: repetition of cycles each of subdivision, selection within lines, selected between lines and crossing.

genes from different parent lines, these Q individuals were mated at random, without selection, and gave Q progeny at generation $T+2$.

A new cycle could therefore start at generation $T+2$. In the one-cycle structure, however, the cross population was maintained as a single large population of size Q and selected with intensity $Nm/Q(= N/M)$. In the repeated cycle structure the

https://doi.org/10.1017/S0016672300013604 Published online by Cambridge University Press

Q individuals at generation $T+2$ were subdivided randomly into m lines, again of size M, and the process repeated. Thus each cycle (including the first) lasted $T+2$ generations, with T generations of within-line selection preceding the between-line selection, 1 generation of within-line selection in the chosen lines, and 1 generation without selection following crossing.

The symbols are summarized below:

Q = total number of individuals measured per generation ($Q = Mm$), and is the same for all systems;
m = number of replicate lines;
M = number of individuals measured per line;
N = number of parents selected in each line, so intensity of within line selection $= N/M$;
v = number of lines selected, so intensity of between-line selection $= v/m$;
T = number of generations of sublining before between-line selection.

A single large population (denoted L) in which mass selection was practised without subdivision was maintained as a control selection system. Each generation Nm individuals were chosen from a total of Q recorded, so that the L line had a size m times as large as the sublines, but had the same selection intensity as that used within lines. It was thus maintained in the same way as the large population after line crossing in the single cycle structure.

(ii) *Genetic model*

Individuals were assumed to be monecious diploids, in which random mating was accompanied by random selfing. The following parameters describe the genetic model:

n = number of loci affecting the character;
a = difference between the homozygotes at a locus in their effect on the character, with all loci having two alleles and additive or completely dominant genes, but no epistasis;
q = initial frequency of favourable allele;
c = recombination fraction between adjacent loci, with all loci equally spaced on a single chromosome;
σ^2 = variance of normally distributed environmental error.

For additive genes the initial heritability of the trait, h^2, is given by

$$h^2 = \tfrac{1}{2}na^2q(1-q)/[\tfrac{1}{2}na^2q(1-q)+\sigma^2].$$

In our runs we have typically taken $n = 5$, $a/\sigma = 0{\cdot}5$ so $h^2 = q(1-q)/[q(1-q)+1{\cdot}6]$. With an initial frequency of $q = 0{\cdot}2$, then $h^2 = 0{\cdot}1/1{\cdot}1 \sim 0{\cdot}1$, which changes during the course of selection, tending to increase initially due to selection but finally to decrease due to inbreeding. We have generally used heritabilities of this order; although they are low, they refer to single chromosomes.

In any generation chromosomes were paired in the order they were produced, to form genotypes. Their genotypic value was computed, an environmental deviation

added and truncation selection practised. The first chromosome for the next generation was obtained by choosing one of the selected parents at random, and performing a random walk (conceptually) along its chromosomes to permit recombination. This process was repeated until the required number of chromosomes were obtained. The whole experiment of sublining, selection, crossing, etc., was replicated 100 or so times for each set of parameters. In each replicate, lines were carried for 80 generations or until fixation, which usually occurred earlier, although limits are denoted '∞' in the tables.

Simulation was carried out on the Edinburgh Regional Computing Centre's KDF 9 computer. Inner loops in machine language were kindly written for us by Dr J. A. Burns.

3. RESULTS

In most of the genetic models which we have studied, where we have found a difference in rate of response or limit to selection between the large population and single cycle structure, we have also found a difference of the same direction, but not size, in rates or limits between the large population and repeated cycle structure. Most of our results therefore refer to the single cycle structure, since by using the same set of sublines to originate the subsequent large lines after different times and intensities of between-line selection, a greater range of parameters could be investigated with the single cycle than the repeated cycle structure for a given computing cost. For example, a set of $M = 8$ sublines was generally used to initiate 9 subsequent large lines, comprising three values of T (usually 1, 3 and 7), each with three values of v (usually 2, 4 and 8). In addition, a positive correlation is induced between the responses in the populations started from the same set of single cycle lines, so that the variance in response between them is reduced.

We shall investigate in turn those 'structural' parameters, such as the number of sublines, which can be controlled by the breeder. In each case we consider how the comparisons between alternative schemes are affected by the genetic model, which is outside the breeder's control. But since the results differ markedly for additive and non-additive models, we shall discuss these separately.

(i) *Single-cycle structure: additive model*

(*a*) *Between-line selection.* A typical result is shown in Fig. 2 for a simple model of five loci of equal effects and initial frequency 0·2 at each. The mean of the selected trait is then a linear function of the mean gene frequency, which is plotted. Prior to crossing, the figure shows the mean performance of all replicate sublines, which soon falls behind that of the large population as the within-line variance of the small lines is reduced. When all sublines are used at generation 3 to make the cross ($v/m = 8/8$) the mean advance lags behind that of the single population, and is furthest behind immediately following line crossing. However, the new synthetic population reaches about the same limit, within the range of sampling error. From Maruyama's (1970) theory we would expect this result for independent loci, but it seems to hold even for those which are tightly linked. Similarly, for other runs we

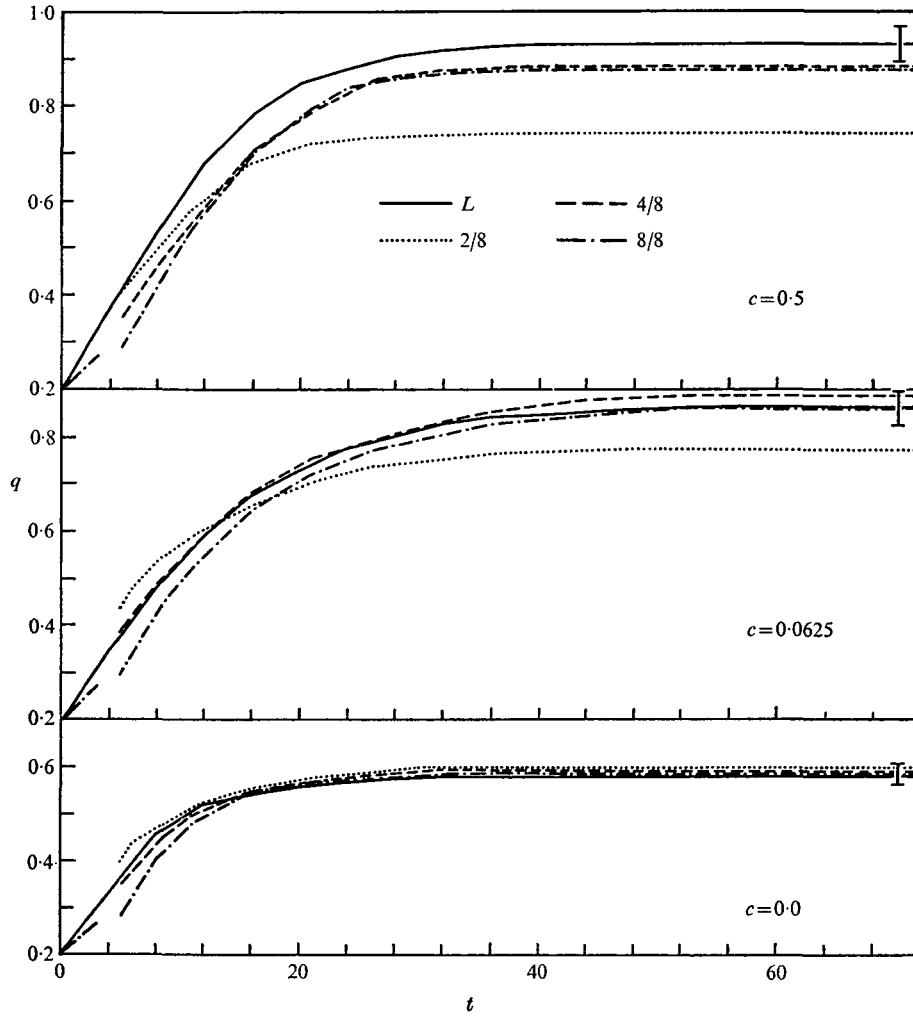

Fig. 2. Comparison between selection in a large population (L) and alternative intensities of between line selection (v/m = 2/8, 4/8 and 8/8) in the single cycle structure with $Q = 40$, $m = 8$, $M = 5$, $N = 2$ for an additive model with $n = 5$, $a/\sigma = 0{\cdot}5$, $q = 0{\cdot}2$ and recombination fraction, c. The mean gene frequencies are shown; these are for the mean of all sublines prior to between-line selection at generation 3 ($= T$) and the blank at generation 4 denotes the random mating following crossing. A range of length approximately 2 standard errors is shown for the difference between L and alternative structures at the limit.

have made for additive models, there is never an important difference between the limits obtained in the single population and in the two-cycle structure when there is no between-line selection. When selection is practised between lines we see (Fig. 2) that following crossing the mean of the cross may exceed that of the large population and remain ahead for a few generations. However, with intense between-line selection (v/m = 2/8) the limit for the single-cycle structure is lower than the limit for the single large population (L), except when the genes are very tightly linked, when there is no difference.

https://doi.org/10.1017/S0016672300013604 Published online by Cambridge University Press

Table 1. *Effect of selection intensity between lines, initial gene frequency and linkage in a single-cycle structure with an additive model*

($Q = 40$, $m = 8$, $N/M = 2/5$, $T = 3$, relative response ± S.E. or mean frequency in $L(\bar{q}_L)$.)

q	v	$c = 0{\cdot}5$ $t = 5$	10	20	∞	$c = 0{\cdot}0625$ $t = 5$	10	20	∞	$c = 0{\cdot}0$ $t = 5$	10	20	∞
$n = 5$, $a/\sigma = 0{\cdot}5$													
0·1	2	6 ± 19	−6	−27	−35 ± 4	24 ± 14	−1	−15	−19 ± 5	−8 ± 6	−8	−18	−19 ± 8
	4	−18 ± 14	−6	−6	−11 ± 6	−12 ± 13	−12	−9	−5 ± 6	−22 ± 10	−13	−8	−11 ± 8
	8	−50 ± 10	−22	−4	−7 ± 5	−60 ± 9	−24	−14	−4 ± 5	−56 ± 9	−23	−7	−6 ± 8
	$\bar{q}_L$	0·21	0·36	0·60	0·73	0·24	0·37	0·56	0·68	0·23	0·34	0·41	0·42
0·5	2	−9 ± 8	−14	−6	−5 ± 3	8 ± 13	−10	−10	−7 ± 4	17 ± 10	4	−3	−3 ± 6
	4	−22 ± 7	−9	−3	0 ± 0	−7 ± 10	−9	1	1 ± 1	−5 ± 9	−4	−4	−3 ± 6
	8	−59 ± 7	−18	−5	0 ± 0	−44 ± 6	−19	−1	−1 ± 2	−50 ± 5	−20	−7	−5 ± 6
	$\bar{q}_L$	0·75	0·88	0·98	1·00	0·74	0·88	0·97	1·00	0·73	0·85	0·91	0·92
0·7	2	15 ± 11	5	2	0 ± 0	30 ± 11	2	−6	−3 ± 2	28 ± 9	2	1	−1 ± 3
	4	−16 ± 10	−6	1	0 ± 0	4 ± 11	−4	−6	0 ± 0	1 ± 9	0	−1	−1 ± 3
	8	−43 ± 7	−9	−2	0 ± 0	−34 ± 10	−9	−2	0 ± 0	−40 ± 8	−17	−3	−4 ± 4
	$\bar{q}_L$	0·87	0·95	0·99	1·00	0·86	0·94	0·99	1·00	0·87	0·96	0·99	1·00
$n = 10$, $a/\sigma = 0{\cdot}35$													
0·1	2	32 ± 23	13	−8	−20 ± 8	6 ± 21	−10	−22	−32 ± 7	19 ± 24	−1	−10	−13 ± 10
	4	−15 ± 18	−9	−4	−5 ± 9	−24 ± 21	−21	−14	−14 ± 7	−12 ± 22	−7	0	2 ± 11
	8	−63 ± 10	−29	−10	4 ± 8	−52 ± 9	−30	−12	−5 ± 6	−60 ± 6	−25	−5	−3 ± 11
	$\bar{q}_L$	0·18	0·28	0·43	0·59	0·18	0·27	0·40	0·55	0·17	0·24	0·28	0·29

Table 2. *Effect of time of crossing in a single cycle structure with an additive model*

($Q = 40$, $M = 8$, $N/M = 2/5$, $n = 5$, $a/\sigma = 0{\cdot}5$, $c = 0{\cdot}5$. Relative response with approx. S.E. for each entry in column.)

		$q = 0{\cdot}1$		$q = 0{\cdot}2$		$q = 0{\cdot}3$		$q = 0{\cdot}4$	
v	T	$t = 10$	∞	$t = 10$	∞	$t = 10$	∞	$t = 10$	∞
2	1	−26	−28	−9	−16	−9	−2	−2	0
	3	−6	−35	−14	−27	−14	−5	5	0
	7	−4	−31	−18	−19	—	—	—	—
4	1	−12	−6	3	−1	1	0	−1	0
	3	−6	−11	−20	−9	−9	0	−6	0
	7	−18	−13	−36	−7	—	—	—	—
8	1	−25	−4	−15	−1	−11	0	−7	0
	3	−22	−7	−22	−9	−18	0	−9	0
	7	−41	−10	−53	−5	—	—	—	—
	Approx. S.E.	15	7	10	2	4	1	4	0

In Table 1 results are given to show the effect of initial gene frequency for a model with other parameters remaining the same as in Fig. 2. Here, and in later tables, the structures are compared in terms of their *relative response*, R_t. Denoting the initial mean by μ_0, the mean of the large population by L_t and that of the other structure by Y_t at generation t, then

$$R_t = 100(Y_t - L_t)/(L_t - \mu_0).$$

Values of R_t are given at intermediate generations and at the limit ($t \to \infty$). With the lowest gene frequency ($q = 0{\cdot}1$) the results in Table 1 are essentially the same as in Fig. 2 ($q = 0{\cdot}2$) in that intense between-line selection has most effect on the limit when there is free recombination. At the higher gene frequencies shown, the chance of fixation of individual genes in the single population approaches 1·0. Then there is little reduction in the limit with between-line selection, and the mean performance with the single cycle structure may be higher for several generations following crossing. Also included in Table 1 is a model with a low initial frequency, a larger number of loci (10) and smaller gene effects than the other models in the Table. The chance of fixation in population L is now only 0·59 for free recombination and 0·29 for complete linkage. However, the results are very similar to those of the model with five loci and $q = 0{\cdot}1$ or 0·2.

(*b*) *Length of the first cycle.* In Table 2 comparisons are made of alternative times (T) of selection between lines (after 1, 3 or 7 generations in sublines) using the same models as in Fig. 2 and Table 1. Only free recombination is included since greater differences are likely to be found than with linkage. In these results, and others not shown, we find that the limit is scarcely and inconsistently affected by the time of crossing, since sampling errors are large relative to the differences we observe. The time of crossing does, of course, affect the mean at intermediate generations (Table 2). When all sublines are chosen the line cross mean is higher at generation 10

Table 3. *Effect of number of sublines in a single cycle structure for an additive model with* $T = 3$, $c = 0{\cdot}5$

(Relative response, or mean gene frequency in $L(\bar{q}_L)$.)

						t			
n	q	a/σ	Q	v/m	N/M	5	10	20	∞
5	0·2	0·5	40	4/8	2/5	−24	−20	−11	−9
				2/4	4/10	3	−7	−8	−9
					$\bar{q}_L$	0·42	0·61	0·85	0·93
10	0·1	0·5	80	4/16	2/5	3	−6	−11	−15
				2/8	4/10	18	−5	−16	−22
				8/16	2/5	−23	−18	−8	−5
				4/8	4/10	−3	−2	−8	−7
				2/4	8/20	−3	−2	−2	−3
					$\bar{q}_L$	0·23	0·40	0·72	0·81
10†	0·5	0·1	80	4/8	2/10	−46	−18	−3	−2
				2/4	4/10	−41	−24	−19	−11
					$\bar{q}_L$	0·58	0·65	0·73	0·88

† Simulation terminated at $t = 80$.

if the crossing is made early since no use is made of the between-line variance. However, with intense between-line selection, temporarily higher means may be obtained with later between-line selection since a larger selection differential can be attained as the variance between lines increases with drift.

(*c*) *Number of sublines.* If the total facilities are kept constant, an increase in the number of sublines must be accompanied by a decrease in the size of each. Thus, at a given time, the variance within lines is reduced and that between lines increased, so the relative efficiencies of within-line and between-line selection may be altered. Results for several models are given in Table 3, each for free recombination. When no selection is practised between lines the limit is independent of the number of sublines (Maruyama, 1970) and no results are included in the table. However, even when selection is practised between the lines, the effect of changing the number of sublines on the limit is small and not significant if the proportion selected within and between lines is not altered. There is one exception in Table 3: $v/m = 2/8$ is poorer than 4/16 for a model with low initial frequency and $a/\sigma = 0{\cdot}5$. However, both schemes are poorer than the single population. At intermediate generations the number of sublines has more effect; higher responses are obtained when the size of the individual sublines is increased.

(*d*) *Total size of the programme.* The relative efficiency of the single cycle and large population structures are compared in Table 4 for different total population sizes (Q). In both schemes the chances of fixation are, of course, increased at larger Q values since the same within-line selection intensities are used. Therefore, although we find smaller differences between the structures at the higher Q values, this is probably solely because the probabilities of fixation approach unity, and we have the same effect as with increase in initial frequency (Table 1). But from the practical

Table 4. *Effect of total number recorded in a single cycle structure with an additive model*

($c = 0{\cdot}5$, $m = 8$, $T = 3$ and $N/M = 0{\cdot}4$. Relative response ± S.E. or mean frequency in $L(\bar{q}_L)$.)

				t				t			
n	q	a/σ	v	5	10	20	∞	5	10	20	∞
				$Q = 40$				$Q = 80$			
5	0·1	0·5	2	6 ± 19	−6	−27	−35 ± 4	0 ± 12	−12	−22	−27 ± 5
			4	−18 ± 14	−6	−6	−11 ± 6	−32 ± 9	−24	−16	−12 ± 5
			8	−50 ± 10	−22	−4	−7 ± 5	−48 ± 7	−23	−8	0 ± 3
			$\bar{q}_L$	0·22	0·36	0·60	0·73	0·23	0·43	0·77	0·92
				$Q = 80$				$Q = 160$			
10	0·1	0·5	2	18 ± 10	−5	−16	−22 ± 3	21 ± 10	6	−3	−7 ± 2
			4	−3 ± 10	−2	−8	−7 ± 3	−11 ± 5	−11	−5	−1 ± 1
			8	−29 ± 11	−14	−2	4 ± 3	−33 ± 4	−19	−6	−1 ± 1
			$\bar{q}_L$	0·22	0·40	0·72	0·81	0·23	0·42	0·78	1·00
				$Q = 80$				$Q = 160$			
10†	0·5	0·1	2	10 ± 19	1	2	−3 ± 4	11 ± 15	3	4	−1 ± 2
			4	7 ± 15	1	7	5 ± 3	−23 ± 19	−17	−1	1 ± 2
			$\bar{q}_L$	0·56	0·61	0·69	0·93	0·57	0·62	0·71	0·96

† Simulation terminated after 80 generations.

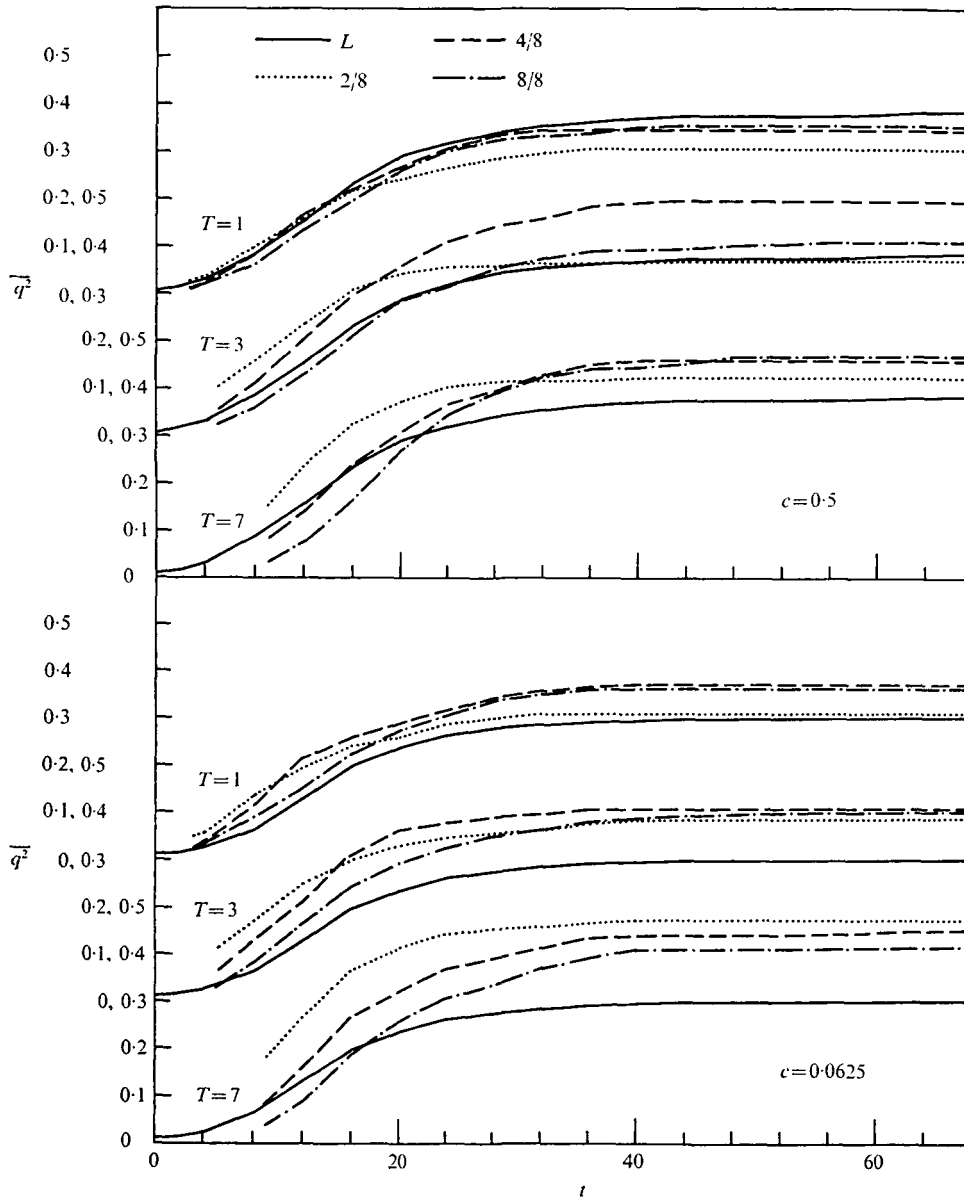

Fig. 3. Comparison between selection in a large population (L) and alternative intensities of between-line selection in the single cycle structure with $Q = 40$, $m = 8$, $M = 5$, $N = 2$ for a recessive model with $n = 5$, $a/\sigma = 0{\cdot}5$, $q = 0{\cdot}1$ and recombination fraction, c, for three times (T) of crossing after subdivision. The population mean is a function of $\overline{q^2}$.

viewpoint this is important, since we have schemes where the mean of the single cycle structure exceeds that of the single populations for a long period with little sacrifice at the limit – for example, when $q = 0{\cdot}1$, $a/\sigma = 0{\cdot}5$, $v = 2$ and $Q = 160$ (Table 4).

https://doi.org/10.1017/S0016672300013604 Published online by Cambridge University Press

Table 5. *Effects of selection intensity between lines, initial gene frequency and linkage in a single-cycle structure with a recessive model*

($Q = 40$, $m = 8$, $N/M = 2/5$, $T = 3$, $n = 5$, $a/\sigma = 0{\cdot}5$. Relative response ± S.E., or mean performance of $L(\bar{q}_L^2)$.)

		$c = 0{\cdot}5$				$c = 0{\cdot}0625$				$c = 0{\cdot}0$			
q	v	$t = 5$	10	20	∞	$t = 5$	10	20	∞	$t = 5$	10	20	∞
0·1	2	201 ± 66	62	18	−4 ± 11	361 ± 89	154	44	29 ± 12	170 ± 96	70	6	4 ± 11
	4	46 ± 29	36	26	30 ± 12	141 ± 45	78	58	37 ± 14	42 ± 65	11	1	7 ± 11
	8	−41 ± 19	−33	−1	6 ± 8	−13 ± 20	25	26	34 ± 12	−34 ± 23	−11	−5	−4 ± 13
	$\bar{q}_L^2$	0·04	0·12	0·29	0·38	0·03	0·10	0·23	0·30	0·05	0·13	0·28	0·30
0·4	2	0 ± 7	−20	−11	−9 ± 3	28 ± 12	−6	−12	−11 ± 4	45 ± 11	−16	5	1 ± 5
	4	−38 ± 7	−15	−6	−6 ± 3	−21 ± 7	−19	−8	−9 ± 3	−5 ± 8	14	13	8 ± 4
	8	−69 ± 6	−32	−11	−1 ± 2	−72 ± 6	−40	−20	−5 ± 4	−51 ± 8	−9	7	5 ± 5
	$\bar{q}_L^2$	0·48	0·74	0·91	0·94	0·44	0·68	0·87	0·90	0·44	0·65	0·73	0·76

Table 6. *Effect of intensity and time of between line selection and initial gene frequency in a single-cycle structure with a recessive model*

($Q = 40$, $m = 8$, $N/M = 2/5$, $n = 5$, $a/\sigma = 0{\cdot}5$, $c = 0{\cdot}5$. Relative response, or mean performance of $L(\bar{q}_L^2)$.)

		$q = 0{\cdot}1$				$q = 0{\cdot}4$				$q = 0{\cdot}7$			
v	T	$t = 5$	10	20	∞	$t = 5$	10	20	∞	$t = 5$	10	20	∞
2	1	49	0	−18	−21	−12	−15**	−9**	−5	2	−3	1	0
	3	201**	62**	18	−4	0	−20**	−11**	−9*	−3	−4	−1	−2
	7	—	54*	30	11	—	−22**	−8	−4	—	−10**	−4	−3
4	1	−8	2	−9	−11	−26**	−6	0	1	−4	−2	0	0
	3	46	26	26	30*	−38**	−15*	−6	−6	−24**	−3	0	0
	7	—	−25	5	19	—	−4	1	3	—	−23**	0	0
8	1	−33	−21	−9	−9	−43**	−20**	−2	2	−24**	−8**	−2	−2
	3	−41	−33	−1	6	−69**	−32**	−11**	1	−48**	−16**	−3	−2
	7	—	−72**	−7	21	—	−38**	−4	5*	—	−40**	−1	0
	$\bar{q}_L^2$	0·04	0·12	0·29	0·38	0·48	0·74	0·91	0·94	0·85	0·97	1·00	1·00

* $0{\cdot}01 < P < 0{\cdot}05$; ** $P < 0{\cdot}01$ relative to L.

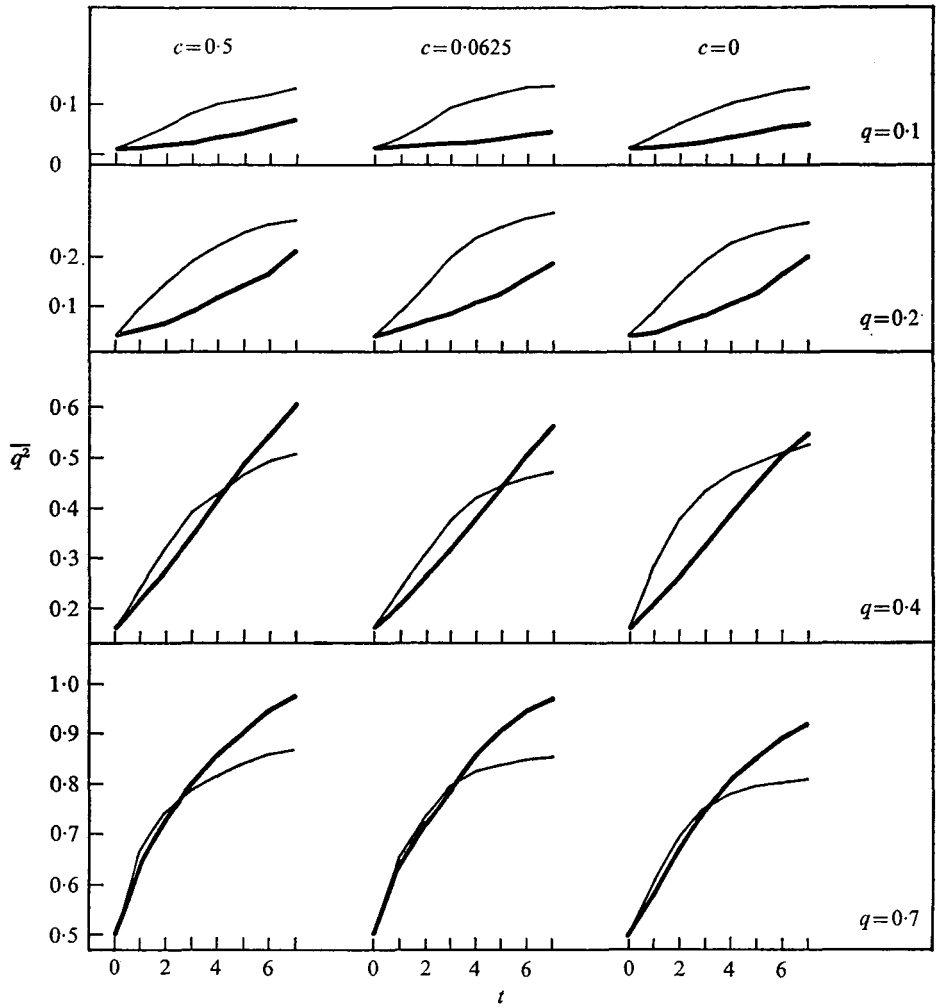

Fig. 4. Comparison between the mean performance (expressed as $\overline{q^2}$) of the large population (thick lines) and the mean of the sublines (thin lines) prior to crossing, for a recessive model with $n = 5$, $a/\sigma = 0{\cdot}5$ and specified initial frequencies (q) and recombination fraction (c). The numbers selected/recorded are 16/40 in the large population and 2/5 in the sublines.

(ii) *Single-cycle structure: recessive model*

We shall use the term 'recessive model' when, at each locus, there is complete dominance and the recessive allele is favoured by selection. If all loci have the same effect on the quantitative trait, the mean performance is a linear function of $\overline{q^2}$, where q is the gene frequency at a single locus in a single replicate. This statistic is used in the figures and tables.

Results for recessive models are given in Tables 5 and 6 and Fig. 3. The structures used are similar to those investigated earlier for the additive model, but the results differ considerably. We find that immediately following between-line selection the mean may be higher than in the single population and can remain ahead at the

https://doi.org/10.1017/S0016672300013604 Published online by Cambridge University Press

limit. These effects are seen most markedly with low initial gene frequencies such as 0·1 (Fig. 3). At higher initial frequencies, such as 0·7 (Table 6) when all favourable alleles are fixed, the mean of the single cycle structure does not exceed that of the single population and the same limit is reached. In general we see that the different intensities of between-line selection have rather small effect on the limit, but, of course, large effects at intermediate generations.

In the recessive model the length of the cycle of sublining has an important influence on the limit. We see in Table 6 that where the schemes differ appreciably in efficiency at low initial frequencies, the highest limits are attained when the between-line selection is delayed. But if no between-line selection is practised the intermediate generations are poorer when crossing is delayed for the lines have ceased to respond to within-line selection. With very tight linkage we find, as in the additive model, that the different intensities of between-line selection do not influence the limit markedly (Table 5).

In Fig. 4 the responses in the initial generations of sublines are compared with those of the single population. In contrast with the additive model, higher rates of gain may be made in the very small lines if the initial frequency is low. In these situations the additive variance actually increases up to intermediate levels of inbreeding (Robertson, 1952). In addition, when the recessive alleles are favoured, there is an inbreeding, enhancement' as homozygotic frequency increases. This is lost in crossing and we see (Fig. 3) that with no between-line selection the line cross is at first poorer than the single population.

(iii) *Single-cycle structure: dominant models*

Some results are given in Table 7 for a model of equal effects and initial gene frequencies with free recombination, in which there is complete dominance with the dominant allele favoured by selection. If it has a low initial frequency the response is less in the single-cycle structure than in the large population throughout the selection period. However, at the limit the difference is small if no between-line selection is practised. In addition, prior to crossing, the sublines perform much more poorly than the single population since the lines exhibit inbreeding depression. At higher initial frequencies of the dominant allele the pattern alters, for as we have seen in the previous section the efficiency of within and subsequently between-line selection is enhanced if the lines are small. However, in our example the chance of fixation is very high and only small differences are observed at the limit. We consider these models further in the repeated cycle scheme.

(iv) *Repeated-cycle structure*

All repeated cycle studies were undertaken with the intermediate cycle length $T = 3$. A typical run with an additive model is shown in Fig. 5, in which the parameters are the same as those used in Fig. 2, and further results are given in Table 8. In each case comparison is made with the large population system.

The repeated subdivision with no between-line selection gives essentially the same limit as the single population (or single cycle) structure, but the limit is reached

Table 7. *Effect of intensity and time of selection between lines in a single cycle structure with a dominant model*

($Q = 40$, $m = 8$, $N/M = 2/5$, $n = 5$, $a/\sigma = 0{\cdot}5$, $c = 0{\cdot}5$. Relative response, or mean performance of L expressed as $1-\overline{(1-q_L)^2}$.)

		$q = 0{\cdot}1$				$q = 0{\cdot}4$				$q = 0{\cdot}7$			
v	T	$t = 5$	10	20	∞	$t = 5$	10	20	∞	$t = 5$	10	20	∞
2	1	−12	−25**	−19**	−30**	−2	−3	1	0	−2	−7	6	2
	3	−21*	−35**	−38**	−39**	−22*	−5	−3	−5	−27	9	11	3
	7	—	−39**	−39**	−38**	—	−18**	−4	−5	—	16	8	3
4	1	−5	−4	−5	−7	1	−1	2	−1	21	4	2	3
	3	−29**	−20**	−14**	−14**	−26**	−6	−1	0	−4	−7	0	3
	7	—	−42**	−22**	−18**	—	−14*	1	0	—	12	6	3
8	1	−27**	−11	0	0	−10	−2	1	0	8	4	4	3
	3	−59**	−24*	−12	−10	−36**	−10**	−1	0	−39	−10	−13	2
	7	—	−55**	−16**	−7	—	−29**	−4	0	—	−34	−11	3
$1-\overline{(1-q_L)^2}$		0·49	0·70	0·84	0·90	0·84	0·92	0·97	1·00	0·94	0·96	0·99	1·00

* $P < 0{\cdot}05$, **$P < 0{\cdot}01$ of zero relative response.

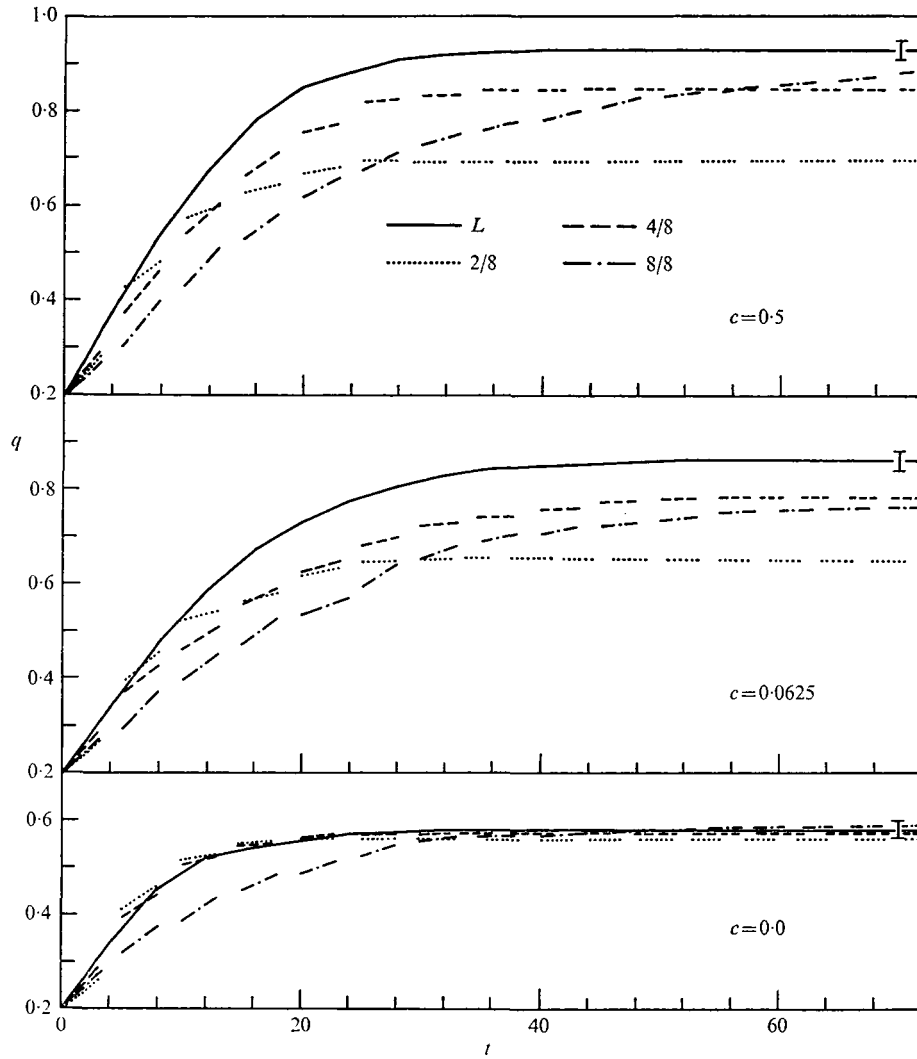

Fig. 5. Comparison between selection in a large population (L) and alternative intensities of between line selection in a repeated cycle structure with an additive model as for Fig. 2: $Q = 40$, $m = 8$, $M = 5$, $N = 2$, $n = 5$, $q = 0{\cdot}2$, $a/\sigma = 0{\cdot}5$.

at a much slower rate. There are, of course, a large number of generations in which no within-line selection is practised following each cross and these both reduce the rate of advance and also the limit to a small extent. With intense selection between the lines the rate of advance is increased, such that in the example shown in Fig. 5 when $v/m = 2/8$ and linkage is complete, the repeated cycle is superior to the large population for the greater part of the two cycles after first crossing, and finally a similar limit is reached. However, with free recombination or partial linkage, the response soon drops below that of the large population, and a lower limit is attained. It is clear that the single cycle and multiple cycle schemes give essentially the same results.

https://doi.org/10.1017/S0016672300013604 Published online by Cambridge University Press

Table 8. *Repeated cycle structure with an additive model*

($T = 3$, $c = 0{\cdot}5$. Response relative to L.)

							t			
Q	m	N/M	n	a/σ	q	v	5	10	20	∞
40	8	2/5	5	0·5	0·1	2	18	3	−25	−36
						4	4	6	−5	−11
						8	−33	−39	−25	−2
					0·5	4	−24	−16	−4	0
					0·7	4	2	−1	−2	0
80	16	2/5	5	0·5	0·1	2	52	13	−28	−38
						4	26	13	−8	−15
						8	−8	−15	−11	1
						16	−61	−64	−57	−9†
80	8	2/10	10	0·1	0·5	4	−13	−18	−11	−9†
						8	−31	−37	−33	−18†

† Simulation terminated prior to fixation (after 80 generations).

Table 9. *Repeated cycle structure with a recessive and dominant model*

($Q = 40$, $T = 3$, $c = 0{\cdot}5$. Response relative to L.)

				Recessive				Dominant			
a/σ	n	v	q	$t = 5$	10	20	∞†	$t = 5$	10	20	∞†
				$m = 8$, $N/M = 2/5$							
0·5	5	4	0·1	34	42	23	6	−27	−24	−15	−17
			0·7	−11	−4	−2	0	−10	19	9	3
0·35	10	4	0·1	63	48	24	24	−41	−38	−36	−29
			0·4	−4	−8	−6	−1	−24	−20	−12	−5
				$m = 20$, $N/M = 2/2$							
0·5	5	5	0·1	−25	−35	−42§	4	—	—	—	—
		10	0·1	−41	−58	−60§	53‡	—	—	—	—

† $t = 60$ for $a/\sigma = 0{\cdot}35$.
‡ $t = 100$.
§ At $t = 60$ relative response is $+2$ for $v = 5$, $+10$ for $v = 10$.

A few results for non-additive models with repeated cycles are given in Table 9. With the recessive allele initially at low frequency, whether at a selective advantage of disadvantage, greater advances may be made both in the early generations and at the limit. Table 9 also includes a model with a low-frequency-favoured recessive in which no selection is practised within sublines, but with 5/20 or 10/20 sublines selected after $T = 3$ generations each cycle. The rate of advance is very slow, but a much higher limit is reached with the less intense between-line selection scheme than with the large population control system.

When the dominant allele is favoured the pattern of response is very irregular with the repeated cycle scheme, since there are intermittent periods of inbreeding

followed by line crossing to restore heterozygosis. Only after several generations do the sublines become fixed sufficiently for their performance not to fall below that of the single population before they are crossed.

4. DISCUSSION

We have studied a restricted range of genetic models with a rather small number of genes of equal effect and initial frequency, and we must ask whether we are entitled to generalize beyond them. We may be justified in doing so if it can be explained why the alternative schemes performed in the way they did. Most of the discussion will be restricted to additive models, for which the theory has been developed furthest.

(i) *Additive genes*

The important item of existing theory is that any subdivision structure, including one of no subdivision, in a single locus additive model gives the same limit so long as there is no between-line selection and the selection intensity is the same in each population (Maruyama, 1970). Our results show that this generalization holds for multiple loci which recombine freely. Now when selection is practised between lines the mean level of inbreeding in the subsequent single population or second cycle sublines is increased and, at least for an additive model, the genetic variance correspondingly reduced. If the inbreeding level in each subline is F_{T+1} at the generation the crosses are made, then the cross of v lines has inbreeding coefficient F_{t+1}/v. For example, with $N = 2$, $T = 3$ and random mating, $F_4/v = 34{\cdot}2\,\%$, $17{\cdot}1\,\%$ and $8{\cdot}5\,\%$ for $v = 2$, 4 or 8. It is clear from our results that the gain from between-line selection is more than compensated by a reduction in subsequent response. This simple argument can be quantified for an additive model with a large number of independent loci each with genes of small effect, as we now show.

Let us assume that the variances change in proportion to the level of inbreeding, since the populations are mated at random and the mean changes in gene frequency are small (Robertson, 1960). Let the heritability of the trait be h^2 and the phenotypic variance σ_p^2. The response to selection with lines in the first cycle, including the selection within each line for crossing, is

$$\begin{aligned}\mu_{T+1} - \mu_0 &= \sum_{t=0}^{T} (1 - 1/2N)^t i h^2 \sigma_p \\ &= 2NiF_{T+1} h^2 \sigma_p, \end{aligned} \tag{1}$$

where i is the standardized within-line selection differential (which we shall assume depends only on the proportion selected, although it is also marginally affected by the total number scored). The genetic variance between lines at generation T when selection is practised between lines is $2F_T h^2 \sigma_p^2$. The within-line phenotypic variance is then

$$[(1 - F_T) h^2 + 1 - h^2] \sigma_p^2,$$

so if M individuals are recorded, the variance of an observed line mean is

$$2F_T h^2 \sigma_p^2 + [(1 - F_T) h^2 + 1 - h^2] \sigma_p^2 / M,$$

where, as in our simulation model, we assume there is no environmental variance common to all members of a line. Thus with a standardized selection differential of i_B, the response, B, to between-line selection is expected to be

$$B = 2i_B F_T h^2 \sigma_p^2 \{2F_T h^2 \sigma_p^2 + [(1-F_T)h^2 + 1 - h^2]\sigma_p^2/M\}^{-\frac{1}{2}}$$
$$= 2i_B h^2 \sigma_p F_T \{M/[(2M-1)F_T h^2 + 1]\}^{\frac{1}{2}}.$$

In the first t^* generations of within-line selection in a population of size Nm in a single cycle structure, subsequent to crossing and random mating (in a sufficiently large population that drift can be ignored at generation $T+1$), the response is

$$\mu_{T+2+t^*} - \mu_{T+1} = 2NmiF_{t^*} h^2 \sigma_p (1 - F_{T+1}/v),$$

where $F_{t^*} = 1-(1-1/2mN)^{t^*}$ is the inbreeding level relative to that after crossing. Thus the total advance from t^* generations of selection after crossing is

$$\mu_t - \mu_0 = 2h^2\sigma_p[iNF_{T+1} + iNmF_{t^*}(1-F_{T+1}/v) + i_B F_T\{M/[(2M-1)F_T h^2+1]\}^{\frac{1}{2}}]$$

and as $t \to \infty$, the limit is

$$\mu_\infty - \mu_0 = 2h^2\sigma_p(iNm - iNF_{T+1}[(m/v)-1] + i_B F_T\{M/[(2M-1)F_T h^2+1]\}^{\frac{1}{2}}).$$

If there is no between-line selection, i.e. $v = m$ and $i_B = 0$, then $\mu_\infty - \mu_0 = 2Nmih^2\sigma_p$, which is the total advance expected in the large population (L) without any subdivision, with this simple model in which the genetic variance is directly proportional to the level of inbreeding.

Using the above formulae we have calculated the advance for the structures used in our simulation studies, and have assumed that $h^2 = 0{\cdot}2$ and line means are normally distributed. This heritability is slightly larger than those used in the simulation (e.g. Table 1, Fig. 2). The results are shown in Fig. 6, using two different scales for time: either generations (t) or $F = 1-(1-1/32)^t$, which is the inbreeding coefficient in L at generation t. On the latter scale the responses in both L and the other large populations after crossing of sublines are linear. Since these results strongly resemble those obtained earlier for additive models with free recombination (Fig. 2), they illustrate the utility of the simple model. Only when between-line selection is practised early and is intense does the response in the single-cycle structure exceed that in the large single population, but then the limit is reduced. The limit is least affected when between-line selection and crossing is done as early as possible, thereby minimising inbreeding in the subsequent population. However, with early crossing less response is made directly from the between-line selection. In our simulation studies we were unable to detect which effect was larger, but presumably would have shown that short cycles of inbreeding gave the highest limits if sufficient replicate computer runs had been made. With cycles of length of only one generation the repeated cycle structure degenerates into a family selection scheme, and since with comparable selection intensities family selection gives a lower limit than mass selection (Robertson, 1960) our results could be anticipated.

A less precise argument on the effects of between-line selection can be used and then extended to include linkage. Imagine the trait under selection is controlled by a few, say 8, independent genes of low initial frequency and that selection is continued in sublines until all the loci are fixed, with the probability of fixation

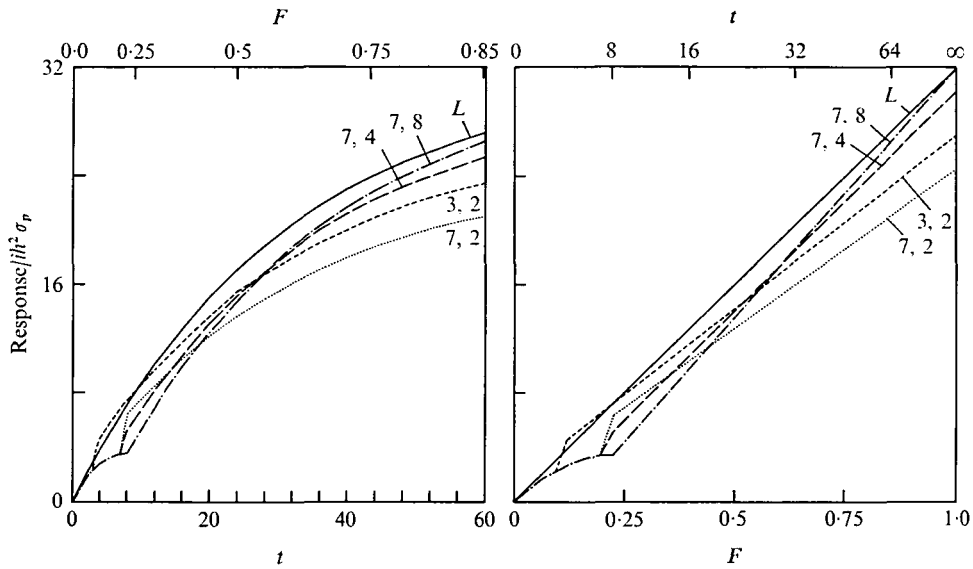

Fig. 6. Responses predicted for an additive model of small gene effects in a single large population (L) or in a single cycle structure with (T, v) generations of inbreeding and lines selected, and $Q = 40$, $m = 8$, $M = 5$, $N = 2$, $h^2 = 0{\cdot}2$. The response is shown as the coefficient of $ih^2\sigma_p$, where i is the within-line selection intensity and σ_p the phenotypic standard deviation for two time-scales: generations, t, and inbreeding coefficient in L, $F = 1-(1-1/32)^t$.

of the favourable allele being 0·25 at each locus. Thus the probability that any line contains 0, 1, 2, 3, 4, $\geqslant 5$ favourable alleles at fixation is 0·10, 0·27, 0·31, 0·21, 0·09, 0·03 respectively, from the binomial distribution. Imagine also that there are eight sublines, so this is also the frequency distribution of the number of sublines which contain the favourable allele at a specified locus. When no selection between lines is practised, there is thus a 90 % chance of having at least one favourable allele at this locus, which, with an initial frequency of at least 1/8, would have a fairly high chance of fixation in the new, larger, population. By contrast, imagine only the best two sublines are chosen. The probability that a line contains at least 4 favourable alleles is 0·12 (or, more precisely, 0·1138), so the probability that at least 2 of 8 lines have 4 or more favourable alleles is $1-(0{\cdot}88)^8-8\times 0{\cdot}12\times(0{\cdot}88)^7=0{\cdot}23$. Thus, even if the two sublines were chosen without error, in only 23 % of samples would these both contain 4 or more favourable alleles, and even if both contain 4 favourable alleles, the probability that the allele at a specific locus is present is only 75 %.

With free recombination the crucial requirement is that at least one representative of the favourable allele at each locus should occur in the cross of selected lines, for subsequent recombination will permit formation of the best possible chromosomes. At the other extreme, if all genes affecting the trait under selection are completely linked on a single chromosome, the most desirable outcome is to retain the best chromosome, initially sampled at the start of the experiment in one subline, during selection between lines and subsequent selection. Now since the between-

line selection has a high accuracy, the line containing the best chromosome has a high chance of being selected, even if only two or so lines are chosen. (Even if it is missed, the next best chromosome will probably be chosen.) Thus the probability of fixing the best chromosome should be little affected by the intensity of between-line selection, and this is the result we obtain. Further, we do not expect to find large differences between sublined structures and a single large population when linkage is complete since the best, or nearly the best, initial chromosome is fixed in either case. The relevance of this kind of genetic model in selection limits in single populations is discussed further by Robertson (1970*a*).

Of course, in nature we have neither independent loci nor complete linkage on single chromosomes, but a mixture of linkage relationships on individual chromosomes together with independence of genes from different chromosomes. Our results show that, with some recombination, the selection between lines has an effect intermediate between that of independence and complete linkage. Thus even for species with few chromosomes we must expect that selection between sublines in the structures we have considered could markedly reduce the limit if the trait is affected mostly by additive genes.

We have undertaken a small number of computer runs with the restriction of equal gene effects or frequencies removed. Using the same structural parameters as in Figure 2, a model was simulated of five additive loci with equal initial frequency and effects $a/\sigma = 0{\cdot}875, 0{\cdot}5, 0{\cdot}375, 0{\cdot}25$ and $0{\cdot}177$, such that the genetic variance is the same as in a model of five loci of effect $a/\sigma = 0{\cdot}5$. The general pattern was found to be similar to that of equal effects, but between-line selection had rather less effect at the limit, presumably because those genes with the largest effect have a high chance of being selected and these contribute most to the total advance. With a more extreme additive model of one locus with $a/\sigma = 1$ and $q = 0{\cdot}025$ and nine loci of $a/\sigma = 0{\cdot}25$ and $q = 0{\cdot}4$, the probability of fixing the gene of large effect was little influenced by the structure, whereas between-line selection reduced the probability of fixing those genes of smaller effect.

In an attempt to utilize the immediate response from selection between lines but to minimize the somewhat drastic effects of truncation selection between lines on the limit we tested a scheme whereby a high proportion of chromosomes to form the line cross pool were taken from the best lines, but some were allowed to enter from the poorer ones. However, we were not successful in this attempt: in order to attain large gains from between-line selection the limit had to be sacrificed.

(ii) *Intermediate generations*

We have concentrated our attention on the mean performance and selection limits after crossing the replicate sublines. However in practice it might be possible to utilize the variation between the sublines by choosing one for multiplication and commercial use, if only on a temporary basis. Baker & Curnow (1969) have estimated this variance between lines for a range of genetic models and shown that the best sublines are likely to be very superior to a large contemporaneous population. Using the model of small gene effects described above, A. Robertson (personal com-

munication) has derived formulae for the relative merits of the best subline and a single large population. He has kindly let us present his analysis, which is based on some approximations appropriate for sizes of sublines rather larger than those used in this study (say $N \geqslant 8$). From equation (1), the expected gain in the sublines of size N after t generations is

$$\begin{aligned} \mu_t - \mu_0 &= 2Nih^2\sigma_p[1-(1-1/2N)^t] \\ &= ih\sigma_A(t-t^2/4N) \quad \text{approximately,} \end{aligned} \tag{2}$$

where σ_A^2 is the additive variance, and provided N is not too small. Thus the reduction in response due to inbreeding up to the tth generation, relative to using a very large population in which inbreeding effects are negligible in this period, is $ih\sigma_A t^2/4N$. At the same time, the genetic variance between lines will be $2F_t\sigma_A^2 = t\sigma_A^2/N$ approximately. If the expected superiority of the best line of the set is k times the standard deviation between them (i.e. $k = i_B$ when one line is chosen), the expected superiority of the best line over the large population may be written as

$$D = \sigma_A[-iht^2/4N + k\sqrt{(t/N)}], \tag{3}$$

which passes through a maximum when $t^3 = Nk^2/i^2h^2$, giving $D = (k^4/Nih)^{1/3}3\sigma_A/4$.

In the selection experiment with *Drosophila melanogaster* of Madalena (1970) there were eight sublines of $N = 10$, with $ih = 0{\cdot}8$. Then the greatest difference between the best subline and the large population is expected at generation 3 when $D = 0{\cdot}58\sigma_A$, about 25 % of the response in the large population at that time. The actual difference was smaller, but could be explained by sampling. In our example of Fig. 2 we have $N = 2$, $i = 1$, $k = 1$, $h^2 = 0{\cdot}1$, approximately, and the above formulae predict that $t_{\max}$ lies between 3 and 4, and that at generation 3, $D = 1{\cdot}3\sigma_A$, whereas in the large population at this time the response would be $0{\cdot}9\sigma_A$. Although formulae such as (2) and (3) do not hold exactly in our example, since N is so small, the prediction is essentially correct for direct calculation gives a maximum D of $1{\cdot}35\sigma_A$ at generation 4. As Baker & Curnow (1969) have shown numerically, the best subline is likely to be much superior to a large population for only a short time. Eventually the large population is likely to be best.

However, the above analysis requires that the line of best genotype be identified. In our simulation experiments small samples were measured each generation, but accuracy of choosing lines could have been improved by recording line means for several generations. For example, the correlation of line means for the model of Fig. 2 was only 0·36 between generations 3 and 7. In addition, these gains from selecting the best line last only a few generations, and although a consequent loss at the limit need not be incurred since all lines can be crossed, the mean of crosses is then poorer than that of a single population selected throughout. There is probably need for further study of methods of structuring populations to make the best use of short-term benefits.

All our comparisons have been made at the same selection intensity within sublines and the large population. Higher response in the initial generation, at the expense of the limit, can be obtained by selecting more intensely within the large

population. This is a much simpler scheme, and can give essentially the same results as a period of subdivision followed by between-line selection and crossing.

(iii) *Dominant genes*

When there is dominance we have seen that the effects of subdivision and between-line selection may differ markedly from those with additive genes. Firstly there is an *increase* in the additive variance in small random mating populations if the recessive alleles are at low frequency (Robertson, 1952). Therefore, as we have seen in Fig. 4, the response in the cycle of subdivision may be higher in the sublines than in the large population. More important, perhaps, the variance between lines at fixation is a function of $q(1-q)$, whereas the initial additive variance is proportional to $q^3(1-q)$ (where q is the frequency of the recessive allele, which is not assumed to change much during selection, i.e. we adopt a small effects model for illustration). Thus the between-line variance and response can be of a different order of magnitude to that within a single large population if the recessive allele is at low frequency. The between-line variance increases in proportion to F^3, where F is the inbreeding coefficient, so it becomes much more efficient if between-line selection is delayed, as our simulations results show. At these later times both the single and repeated cycle schemes give higher responses both in intermediate generations and at the limit than does the single large population. However, we see from our results that intense between-line selection depresses the limit (at least below that for no selection between lines in the same structure) for the same reasons as given in the additive model, and that favourable alleles at some loci are lost during this restriction of population size.

When the recessive alleles are at intermediate or high frequency we have found that a structured scheme is not of benefit, and, as predicted in the additive case, delaying between-line selection gives lower responses in intermediate generations, as well as lower limits. The arguments of the previous section on low-frequency recessives now act in reverse. We have simulated some models with both additive and completely dominant genes (with the recessive favoured) and found that between-line selection influences response in a manner roughly intermediate between that for additive and recessive models taken separately.

Few, if any, quantitative traits of economic importance show negative heterosis. Therefore it is unlikely that much useful variation is expressed at loci in which the recessive alleles are favoured, so we can suggest that the kind of structured systems discussed here are only likely to be useful for removing deleterious recessive genes initially at low frequency. In other genetic situations it seems unlikely that there will be sufficient extra gain in initial generations from between-line selection to compensate for the potential loss at the limit when between-line selection is practised and similar gains can be obtained simply by using more intense selection within single populations. Small and temporary benefits can be obtained, however, from using the best sublines prior to crossing. We have not investigated epistatic models, for which these line crossing systems were originally proposed by Wright, and some studies with these models could be rewarding.

REFERENCES

BAKER, L. H. & CURNOW, R. N. (1968). Choice of population size and use of variation between replicate populations in plant breeding selection programs. *Crop Science* **9**, 555–560.

BOWMAN, J. C. & FALCONER, D. S. (1960). Inbreeding depression and heterosis of litter size in mice. *Genetical Research* **1**, 262–274.

DEMPSTER, E. R. (1955). Genetic models in relation to animal breeding problems. *Biometrics* **11**, 535–536.

GOODWILL, R. (1971). Effect of population structure in selection experiments with *Tribolium castaneum*. *Genetics* **68**, s 24.

HILL, W. G. (1963). Cyclical inbreeding with selection in *Drosophila melanogaster*. Unpublished M.S. thesis. University of California, Davis.

HILL, W. G. & MADALENA, F. E. (1969). Optimum population structure in selection programs. *Genetics* **61**, s 26–27.

HILL, W. G. & ROBERTSON, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269–294.

JAMES, J. W. (1972). Optimum selection intensity in breeding programmes. *Animal Production* **14**, 1–9.

MADALENA, F. E. (1970). Studies on the limits to artificial selection. Unpublished Ph.D. thesis, University of Edinburgh.

MARUYAMA, T. (1970). On the fixation probability of mutant genes in a subdivided population. *Genetical Research* **15**, 221–227.

POLLAK, E. (1966). On the survival of a gene in a subdivided population. *Journal of Applied Probability* **3**, 142–155.

ROBERTSON, A. (1952). The effects of inbreeding on the variation due to recessive genes. *Genetics* **37**, 189–207.

ROBERTSON, A. (1960). A theory of limits in artificial selection. *Proceedings of the Royal Society of London* B **153**, 234–249.

ROBERTSON, A. (1970*a*). A theory of limits in artificial selection with many linked loci. In *Mathematical Topics in Population Genetics* (ed. K. Kojima), pp. 246–288. Berlin: Springer.

ROBERTSON, A. (1970*b*). Some optimum problems in individual selection. *Theoretical Population Biology* **1**, 120–127.

WRIGHT, S. (1939). Genetic principles governing the rate of progress of livestock breeding. *Proceedings of the American Society of Animal Production*, **32**, 18–26.

WRIGHT, S. (1951). The genetical structure of populations. *Annals of Eugenics* **15**, 323–354.