THE RESPONSE TO ARTIFICIAL SELECTION DUE TO AUTOSOMAL GENES OF LARGE EFFECT

II. THE EFFECTS OF LINKAGE ON LIMITS TO SELECTION IN FINITE POPULATIONS

By B. D. H. LATTER*

[*Manuscript received March* 16, 1965]

Summary

The limits to artificial selection for pairs of linked additive loci of equal proportionate effect, starting from a base population in linkage equilibrium, have been studied by means of a simulation technique on a CDC 3600 computer. Particular attention has been paid to the case of genes of large effect, making use of the definition of the selective value of a genotype given in the first paper of this series.

It has been shown that the expected total response is progressively reduced as the degree of linkage is intensified, the effect being most pronounced when the response due to unlinked genes is expected to be 60–70% of the maximum possible advance. The magnitude of the reduction observed in these experiments is appreciable only at recombination values less than 0·10. It is concluded that if the effects of population size alone are such as to reduce the expected response by 40% or more, the linkage effect may be relatively unimportant for genes separated by as little as five map units.

I. INTRODUCTION

The expected total response in a quantitative character under artificial selection depends upon the effective size of the breeding population, N, and on the intensity of selection imposed. If two alleles of additive effect, A_1 and A_2, occur at a given locus with initial frequencies p_1 and p_2 respectively, and if a/σ represents the proportionate effect of the locus on the quantitative trait concerned, the probability of ultimate fixation of the favoured homozygote A_2A_2 under selection is given approximately by

$$u(p_2) = [1-\exp(-2vp_2)]/[1-\exp(-2v)], \tag{1}$$

where $v = N\bar{\imath}a/\sigma$, and $\bar{\imath}$ denotes the standardized selection differential (Robertson 1960). The approximation cannot be expected to hold if $\bar{\imath}a/\sigma$ is of a higher order of magnitude than $1/N$, and can only be applied to a locus segregating independently of other loci contributing to the additive genetic variance (Kimura 1957).

Since the partial differential coefficient of $u(p_2)$ with respect to v is positive for non-zero v, the probability of fixation of the desired genotype increases as $N\bar{\imath}$ increases. The greater the selection differential involved in the reproduction of the population, the greater will be the directed change in gene frequency in each generation, which has expectation $\frac{1}{2}\bar{\imath}(a/\sigma)q(1-q)$ for a gene of frequency q. In addition, an increase in the effective population size reduces the magnitude of random changes in gene frequency from generation to generation, the change due to genetic sampling having expectation zero and variance $\frac{1}{2}q(1-q)/N$. Theoretical considerations of this nature are of obvious importance in the determination of breeding strategy.

* Division of Plant Industry, CSIRO, Canberra.

Though it is to be expected *a priori* that linkage between loci affecting a quantitative character will reduce the expected total response from a population initially in equilibrium under random mating, no quantitative study of this phenomenon has as yet been reported. It is the purpose of this paper to examine the limits to selection for a pair of linked additive loci of equal proportionate effect on a quantitative trait, assuming the base population to be in linkage equilibrium, and giving particular attention to the case of genes of large effect. The results to be presented have been generated by means of a simulation technique on a CDC 3600 computer.

TABLE 1

PARAMETERS OF THE BASE POPULATION

Genotype	Frequency	Value	Gamete Frequencies			
			A_1B_1	A_1B_2	A_2B_1	A_2B_2
$(A_1B_1)(A_1B_1)$	$(1-p_2)^4$	$-a$	1	0	0	0
$(A_1B_1)(A_1B_2)$	$2(1-p_2)^3p_2$	$-\frac{1}{2}a$	$\frac{1}{2}$	$\frac{1}{2}$	0	0
$(A_1B_2)(A_1B_2)$	$(1-p_2)^2p_2^2$	0	0	1	0	0
$(A_1B_1)(A_2B_1)$	$2(1-p_2)^3p_2$	$-\frac{1}{2}a$	$\frac{1}{2}$	0	$\frac{1}{2}$	0
$(A_1B_1)(A_2B_2)$	$2(1-p_2)^2p_2^2$	0	$\frac{1}{2}(1-y)$	$\frac{1}{2}y$	$\frac{1}{2}y$	$\frac{1}{2}(1-y)$
$(A_1B_2)(A_2B_1)$	$2(1-p_2)^2p_2^2$	0	$\frac{1}{2}y$	$\frac{1}{2}(1-y)$	$\frac{1}{2}(1-y)$	$\frac{1}{2}y$
$(A_1B_2)(A_2B_2)$	$2(1-p_2)\ p_2^3$	$+\frac{1}{2}a$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$(A_2B_1)(A_2B_1)$	$(1-p_2)^2p_2^2$	0	0	0	1	0
$(A_2B_1)(A_2B_2)$	$2(1-p_2)\ p_2^3$	$+\frac{1}{2}a$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
$(A_2B_2)(A_2B_2)$	p_2^4	$+a$	0	0	0	1

II. THE COMPUTER PROGRAMME

The method of approach has not been that of strict simulation of genetic processes by means of logical operations with individual binary digits in a computer, as reported by Fraser (1957) and Martin and Cockerham (1959). Instead a method has been adopted which is particularly suited to the study of a small number of loci under very general conditions, all operations in the passage from one generation to the next being algebraically defined, apart from a process involving random sampling of genotypes from a conceptual population of potential parents.

The programme is concerned with two linked loci A and B initially in linkage equilibrium with recombination value y, each carrying two alleles denoted by A_1, A_2, and B_1, B_2. It has also been assumed in the present studies that each locus has the same proportionate effect, a/σ, on the quantitative character under selection, and that the favoured allele at each locus (A_2 or B_2) is initially present in the population with the same frequency p_2. The loci have been assumed to show additive gene action, the heterozygote at each being intermediate in effect between the alternative homozygotes, and the effects of the two loci being mutually independent. Any of these assumptions may be relaxed by minor alterations of the programme. The two linked genes are also assumed to be segregating independently of all other loci affecting the character.

The base population is generated as a set of 10 genotypic frequencies with corresponding genotypic values as set out in Table 1. The frequencies are then modified to those expected following truncation selection in a large population, by means of selective values

$$w = 1 + i(d/\sigma)[1 + \tfrac{1}{2}dx_0/\sigma^2], \tag{2}$$

where d/σ represents the relevant genotypic value and x_0/σ the value of the character at the point of truncation, both expressed as deviations from the population mean and measured in phenotypic standard deviations (Latter 1965). A finite set of N parental individuals is next chosen from the truncated population by the use of a pseudo-random number generator due to Rotenberg (1960), each of the 10 genotypes being weighted according to its frequency in the population of potential parents.

TABLE 2

MAXIMUM ERRORS (%) INVOLVED IN THE CALCULATION OF CHANGES IN GENE FREQUENCY BY THE METHOD OF SELECTIVE VALUES

Selection Intensity (P)	Proportionate Effect of Each Locus			
	0·25	0·50	0·75	1·00
0·80	1·0	3·9	8·7	15·4
0·40	0·5	1·9	4·1	7·1
0·20	0·2	0·7	1·7	3·4
0·10	0·9	3·4	7·2	12·2

Since N has been defined to be the *effective* size rather than the actual size of the breeding population, it is appropriate to produce the offspring generation by a process involving random union of gametes without regard to parental origin (Wright 1951): the resulting genotypic frequencies represent the population produced by a single generation of artificial selection. The complete cycle may thereafter be repeated until an end-state is reached in which all chosen parents are of the same homozygous genotype.

There are obvious limitations to the use of the method of approximate selective values adopted in this programme, comparable to those outlined in the first paper of this series for selection among genotypes at a single locus. The range of genotypic values involved in the present study, and the contribution of the loci under examination to the additive genetic variance, are both increased by a factor of two by comparison with the single locus case, and errors due to the approximate nature of the selective values are thereby magnified. Changes in gene frequency at loci of proportionate effect less than or equal to 0·25 are calculated with great accuracy by the programme, the errors involved being no greater than 1·0% over a range of selection intensities from 80% to 10%, i.e. a range of values of i extending from 0·350 to 1·755 (Table 2). For $a/\sigma \leq 0{\cdot}50$ the comparable errors are less than 4·0%

over the same range of selection intensities. In the assessment of the maximum error involved at a given selection intensity for loci of a given proportionate effect, a range of gene frequencies extending from 0·001 to 0·999 has been tested.

III. Experimental Designs and Analyses

In the course of these studies, three experiments have been conducted. Experiment 1 was designed to document in considerable detail the effects of linkage on limits to selection for genes of large effect, under selection of moderate to high intensity in small populations. Three regimes with almost identical values of $N\bar{\imath}a/\sigma$ have been studied at each of six initial gene frequencies (Table 3), each combination

Table 3
regimes involved in experiment 1

Regime	N	a/σ	P	$\bar{\imath}$
1	5	0·50	0·10	1·7550
2	10	0·25	0·10	1·7550
3	10	0·50	0·45	0·8796

being tested at recombination values of 0·50, 0·10, 0·05, and 0·01. The experiment therefore involves a total of $3\times6\times4 = 72$ treatments, and the expected total response for each treatment has been estimated from the frequencies observed over 1000 replicate computer runs.

Experiment 2 was conducted to explore a wider range of effective population sizes ($N = 10, 20, 40$) in combination with genes of either small or large proportionate effect ($a/\sigma = 0{\cdot}1,\ 0{\cdot}5$). Each of the six combinations has been tested at each of two selection intensities ($P = 0{\cdot}10, 0{\cdot}25$) with three recombination values ($y = 0{\cdot}5,\ 0{\cdot}05,\ 0{\cdot}01$). For each of the $3\times2\times2\times3 = 36$ treatment combinations a total of 1000 replicates has been observed. The initial gene frequency was in each case specifically chosen to give an expected total response due to unlinked genes equal to 70% of the maximum possible advance, i.e. so that

$$R^* = [u(p_2)-p_2]/(1-p_2) = 0{\cdot}70. \tag{3}$$

This level of response was chosen on the basis of the results of experiment 1, as the region in which the effects of linkage on total response were most likely to be at a maximum.

Experiment 3 involves factorial combinations of two population sizes ($N = 10,\ 40$), two levels of proportionate effect ($a/\sigma = 0{\cdot}1,\ 0{\cdot}5$), and two recombination values ($y = 0{\cdot}5,\ 0{\cdot}01$), at each of two levels of expected response ($R^* = 0{\cdot}50,\ 0{\cdot}70$). The selection intensity for each treatment is the same at $P = 0{\cdot}10$. The purpose of the experiment was simply to check that inferences drawn from experiment 2 also were valid at other levels of expected response. Each of the 2^4 treatment combinations was assessed over 1000 replicate runs. However, no attempt was made to repeat treatments tested in experiment 2, i.e. those for which $R^* = 0{\cdot}70$, and the same data have been used in the analyses of both experiments.

For each treatment in the above experiments the following statistics have been calculated.

(*a*) *Observed Response*

If we denote the observed frequency of end-state $(A_iB_j)(A_iB_j)$ by (ij), the estimated probability of fixation of the desired allele at locus A is (21)+(22), and the corresponding probability for locus B is (12)+(22). Since the programme we are discussing assumes both loci to have identical values of a/σ and p_2, the two probabilities have the same expectation and we may take

$$U = (22)+\tfrac{1}{2}[(12)+(21)] \tag{4}$$

as an estimate of the required probability of fixation. The expected total response to selection as a fraction of the maximum possible advance may then be estimated to be

$$R = (U-p_2)/(1-p_2). \tag{5}$$

(*b*) *Interaction between the Two Loci*

The interaction between the two loci *at the selection limit* can be defined in terms of

$$\Delta = (22)(11)-\tfrac{1}{4}[(12)+(21)]^2, \tag{6}$$

which measures the departure of the observed frequencies (ij) from an expectation of $(1-U)^2 : 2U(1-U) : U^2$. The most useful statistic is

$$I = \Delta/[U(1-U)], \tag{7}$$

since if the number of computer runs involved is n, the distribution of nI^2 is that of χ_1^2 if the two loci are independent at the limit. Since $n = 1000$ in all experiments discussed in this paper, a value of I may be judged significant at the 5% level if it falls outside the range $\pm 0{\cdot}0620$, at the 1% level if it falls outside $\pm 0{\cdot}0815$, and at the 0·1% level if it falls outside $\pm 0{\cdot}1041$.

The standard error of estimation of U can readily be shown from multinomial distribution theory to be

$$\text{S.E.}(U) = \{[U(1-U)+\Delta]/2n\}^{\frac{1}{2}}. \tag{8}$$

The standard error associated with the statistic R is therefore

$$\text{S.E.}(R) = \text{S.E.}(U)/(1-p_2). \tag{9}$$

(*c*) *Frequency of End-state,* $(A_1B_1)(A_1B_1)$

It has been found of considerable interest in the above experiments to analyse the frequency with which both desired alleles A_2 and B_2 are eliminated by chance from the population under selection. We have previously denoted this frequency by (11), which is given in terms of the other variables by the relationship

$$(11) = (1-U)^2+\Delta. \tag{10}$$

Since (11) has a binomial distribution, the angular transformation has been used in analyses of variance of the variable, the theoretical error variance then being $820{\cdot}7/n$ (Mather 1951).

IV. Experimental Results

The results of experiments 1, 2, and 3 are presented in Tables 4–8, together with the theoretical expectation for each tested combination given by

$$R^* = [u(p_2) - p_2]/(1 - p_2).$$

The agreement between observed response (R) at a recombination value of 0·50, and the corresponding value of R^*, is remarkably good throughout the three experiments when one considers the assumptions on which the theory is based, viz. that differences in selective value are constant and of small magnitude.

Table 4

Limits to selection for $N = 5$, $a/\sigma = 0{\cdot}5$, and $P = 0{\cdot}10$ (regime 1 of experiment 1)

Initial Frequency (p_2)	Recombination Value (y)	Theoretical Response (R^*)	Observed Response (R)	S.E. (R)	Interaction (I)
0·300	0·50	0·897	0·885	0·008	−0·060
	0·10	—	0·839	0·010	−0·058
	0·05	—	0·829	0·010	−0·117***
	0·01	—	0·772	0·011	−0·152***
0·200	0·50	0·784	0·764	0·011	−0·037
	0·10	—	0·726	0·011	−0·118***
	0·05	—	0·699	0·011	−0·121***
	0·01	—	0·649	0·011	−0·251***
0·150	0·50	0·685	0·685	0·011	−0·065*
	0·10	—	0·645	0·012	−0·101**
	0·05	—	0·605	0·011	−0·183***
	0·01	—	0·548	0·011	−0·315***
0·125	0·50	0·618	0·615	0·012	0·006
	0·10	—	0·567	0·012	−0·096**
	0·05	—	0·546	0·011	−0·245***
	0·01	—	0·506	0·010	−0·326***
0·100	0·50	0·538	0·514	0·012	−0·046
	0·10	—	0·497	0·011	−0·193***
	0·05	—	0·474	0·011	−0·167***
	0·01	—	0·452	0·010	−0·336***
0·075	0·50	0·440	0·432	0·012	−0·037
	0·10	—	0·423	0·011	−0·117***
	0·05	—	0·398	0·011	−0·102**
	0·01	—	0·377	0·010	−0·306***

*, **, *** Significant at 5, 1, and 0·1% levels, respectively.

For genes of proportionate effect equal to 0·10 in breeding populations of effective size 20 or greater, the agreement with expectation for unlinked genes is excellent (Tables 7 and 8). However, the observed response for genes of large effect

segregating in small populations under selection of moderately high intensity is consistently less than expected (Tables 4–6). Averaging over all tested combinations with $y = 0{\cdot}5$ in experiment 1, the mean response is $0{\cdot}648 \pm 0{\cdot}003$ compared with a mean R^* value of $0{\cdot}661$.

TABLE 5

LIMITS TO SELECTION FOR $N = 10$, $a/\sigma = 0{\cdot}25$, AND $P = 0{\cdot}10$ (REGIME 2 OF EXPERIMENT 1)

Initial Frequency (p_2)	Recombination Value (y)	Theoretical Response (R^*)	Observed Response (R)	S.E. (R)	Interaction (I)
0·300	0·50	0·897	0·881	0·009	−0·038
	0·10	—	0·861	0·009	−0·017
	0·05	—	0·842	0·010	−0·094**
	0·01	—	0·791	0·011	−0·123***
0·200	0·50	0·784	0·774	0·011	−0·005
	0·10	—	0·769	0·010	−0·087**
	0·05	—	0·704	0·011	−0·088**
	0·01	—	0·662	0·011	−0·218***
0·150	0·50	0·685	0·658	0·012	0·006
	0·10	—	0·645	0·012	−0·076*
	0·05	—	0·627	0·011	−0·141***
	0·01	—	0·582	0·011	−0·280***
0·125	0·50	0·618	0·617	0·012	−0·006
	0·10	—	0·591	0·011	−0·127***
	0·05	—	0·566	0·011	−0·167***
	0·01	—	0·508	0·011	−0·303***
0·100	0·50	0·538	0·535	0·012	−0·034
	0·10	—	0·521	0·012	−0·101**
	0·05	—	0·490	0·012	−0·063*
	0·01	—	0·454	0·011	−0·236***
0·075	0·50	0·440	0·436	0·012	0·012
	0·10	—	0·418	0·012	−0·088**
	0·05	—	0·411	0·011	−0·123***
	0·01	—	0·394	0·010	−0·321***

*, **, *** Significant at 5, 1, and 0·1% levels, respectively.

For each of the tested combinations in experiments 1, 2, and 3 the observed response R, expressed as a proportion of the total possible advance, falls off progressively as the recombination value between the two loci is reduced, with only one non-significant exception. Again almost without exception, the reduction in total response is associated with significant negative values of I, as defined by equations (6) and (7).

To assist in an understanding of the mechanics of the process leading to such negative measures of interaction between the two loci at the selection limit, the

frequency with which *both* desired alleles A_2 and B_2 have been *eliminated* has been analysed separately for each of the three experiments (Table 9). The main effects corresponding to differences in linkage intensity are non-significant in each of the three analyses, as are all first- and second-order interactions involving y, with the exception of the $y \times p_2$ interaction in experiment 1 (Table 9). A breakdown of the 15 degrees

TABLE 6

LIMITS TO SELECTION FOR $N = 10$, $a/\sigma = 0{\cdot}5$, AND $P = 0{\cdot}45$ (REGIME 3 OF EXPERIMENT 1)

Initial Frequency (p_2)	Recombination Value (y)	Theoretical Response (R^*)	Observed Response (R)	S.E. (R)	Interaction (I)
0·300	0·50	0·898	0·887	0·009	−0·017
	0·10	—	0·869	0·009	−0·028
	0·05	—	0·851	0·009	−0·105***
	0·01	—	0·765	0·011	−0·139***
0·200	0·50	0·785	0·768	0·011	0·042
	0·10	—	0·764	0·010	−0·088**
	0·05	—	0·717	0·011	−0·122***
	0·01	—	0·659	0·011	−0·265***
0·150	0·50	0·686	0·673	0·012	−0·036
	0·10	—	0·652	0·012	−0·080*
	0·05	—	0·616	0·011	−0·194***
	0·01	—	0·546	0·011	−0·302***
0·125	0·50	0·620	0·584	0·013	0·054
	0·10	—	0·579	0·012	−0·084**
	0·05	—	0·573	0·011	−0·153***
	0·01	—	0·514	0·011	−0·282***
0·100	0·50	0·539	0·533	0·012	−0·028
	0·10	—	0·477	0·012	−0·066*
	0·05	—	0·454	0·011	−0·140***
	0·01	—	0·437	0·010	−0·298***
0·075	0·50	0·441	0·432	0·012	−0·071*
	0·10	—	0·430	0·012	−0·039
	0·05	—	0·373	0·011	−0·162***
	0·01	—	0·378	0·010	−0·256***

*, **, *** Significant at 5, 1, and 0·1% levels, respectively.

of freedom concerned in this interaction has shown the only significant single degrees of freedom to be the cubic $y \times$ linear p_2, the cubic $y \times$ cubic p_2, and the cubic $y \times$ quartic p_2 components. Such a pattern of interaction is extremely difficult to interpret, and may well be a chance rather than a real phenomenon. The observed frequency (11) in no instance departed by more than $\pm 0{\cdot}04$ from the corresponding observation for unlinked loci.

TABLE 7

LIMITS TO SELECTION FOR REGIMES INVOLVED IN EXPERIMENT 2

Population Size (N)	Proportionate Effect (a/σ)	Recombination Value (y)	Observed Response (R)	S.E. (R)	Interaction (I)
		(a) $P = 0{\cdot}10$, $R^* = 0{\cdot}70$			
10	0·1	0·50	0·706	0·016	−0·017
		0·05	0·696	0·015	−0·060
		0·01	0·666	0·016	−0·062*
	0·5	0·50	0·692	0·011	−0·048
		0·05	0·647	0·010	−0·145***
		0·01	0·571	0·009	−0·361***
20	0·1	0·50	0·694	0·012	0·026
		0·05	0·683	0·011	−0·130***
		0·01	0·656	0·012	−0·111***
	0·5	0·50	0·686	0·010	−0·045
		0·05	0·683	0·010	−0·097**
		0·01	0·634	0·010	−0·227***
40	0·1	0·50	0·691	0·011	−0·038
		0·05	0·666	0·011	−0·034
		0·01	0·620	0·010	−0·193***
	0·5	0·50	0·703	0·010	−0·021
		0·05	0·678	0·010	−0·069*
		0·01	0·656	0·010	−0·131***
		(b) $P = 0{\cdot}25$, $R^* = 0{\cdot}70$			
10	0·1	0·50	0·694	0·025	0·024
		0·05	0·694	0·024	−0·050
		0·01	0·685	0·024	−0·039
	0·5	0·50	0·698	0·011	0·019
		0·05	0·632	0·011	−0·117***
		0·01	0·566	0·010	−0·351***
20	0·1	0·50	0·709	0·013	0·012
		0·05	0·683	0·013	−0·056
		0·01	0·639	0·013	−0·145***
	0·5	0·50	0·705	0·011	0·003
		0·05	0·653	0·010	−0·108***
		0·01	0·602	0·010	−0·268***
40	0·1	0·50	0·695	0·012	0·024
		0·05	0·690	0·011	−0·069*
		0·01	0·634	0·011	−0·130***
	0·5	0·50	0·676	0·011	0·027
		0·05	0·673	0·010	−0·051
		0·01	0·639	0·010	−0·210***

*, **, *** Significant at 5, 1, and 0·1% levels, respectively.

Despite this slight difficulty of interpretation, the general conclusion from the analyses is quite clear, viz. that the degree of linkage between the loci has virtually no effect on the probability of fixation of the homozygote $(A_1B_1)(A_1B_1)$ over the range $0{\cdot}01 \leq y \leq 0{\cdot}50$. A reduction in total response due to linkage must therefore be due simply to an increase in the probability of fixation of homozygotes $(A_1B_2)(A_1B_2)$ and $(A_2B_1)(A_2B_1)$, without alteration in the probability of end-state $(A_1B_1)(A_1B_1)$. It is this phenomenon which leads to negative values of the parameter I.

The Reduction in Response due to Linkage

An unweighted analysis of the variation in total response observed in experiment 1 is given in Table 10, together with a theoretical error mean square calculated as the mean of the individual error variances of R. The effect of linkage on total response, averaged over the three regimes and six values of p_2, is highly significant,

TABLE 8

LIMITS TO SELECTION FOR REGIMES INVOLVED IN EXPERIMENT 3 ($P = 0{\cdot}10$, $R^* = 0{\cdot}50$)

Population Size (N)	Proportionate Effect (a/σ)	Recombination Value (y)	Observed Response (R)	S.E. (R)	Interaction (I)
10	0·1	0·50	0·467	0·015	0·047
		0·01	0·446	0·015	−0·061
	0·5	0·50	0·502	0·011	−0·046
		0·01	0·438	0·010	−0·304***
40	0·1	0·50	0·505	0·012	0·050
		0·01	0·464	0·011	−0·138***
	0·5	0·50	0·521	0·011	−0·035
		0·01	0·481	0·011	−0·069*

*, *** Significant at 5 and 0·1% levels, respectively.

and the magnitude of the effect depends on the initial gene frequency concerned. The three regimes differ in mean response, despite the fact that all three have virtually the same value of $N\bar{i}a/\sigma$, but the interaction between regimes and linkage intensity is non-significant.

We may therefore consider the effects of linkage averaged over the three regimes separately for each value of p_2, as characteristic of pairs of additive genes of large proportionate effect, under selection of moderate to high intensity in breeding populations of effective size 5–10. In Figure 1 is plotted the *reduction* in response due to linkage for $y = 0{\cdot}10$, $0{\cdot}05$, and $0{\cdot}01$, calculated as a percentage of the observed response for $y = 0{\cdot}50$. Each point on the graph is based on 3000 replicates, including two sets of points which are additional to the data presented in Tables 4–6.

The percentage reduction in response due to linkage is greatest when the expected response due to unlinked genes is roughly 60–70% of that possible in very large populations. However, the effect is of a very minor order for a pair of

TABLE 9

ANALYSIS OF VARIANCE OF THE FREQUENCY (11) OF END-STATE $(A_1B_1)(A_1B_1)$ IN EXPERIMENTS 1, 2, AND 3

Experiment No.	Source of Variation	Degrees of Freedom	Mean Square
1	Recombination value (y)	3	0·5094
	Initial frequency (p_2)	5	1185·6467***
	Regimes	2	1·5604
	$y \times p_2$	15	2·2360***
	$y \times$ regimes	6	0·3811
	$p_2 \times$ regimes	10	0·7457
	$y \times p_2 \times$ regimes	30	0·8903
	Theoretical error	—	0·8207
2	Recombination value (y)	2	1·7356
	Population size (N)	2	98·3174***
	Regimes	3	90·2361***
	$y \times N$	4	0·1369
	$y \times$ regimes	6	0·7764
	$N \times$ regimes	6	21·4971***
	$y \times N \times$ regimes	12	0·8701
	Theoretical error	—	0·8207
3	Recombination value (y)	1	0·0930
	Theoretical limit (R^*)	1	670·8100***
	Population size (N)	1	58·3696***
	Proportionate effect (a/σ)	1	43·7582***
	$y \times R^*$	1	0·0306
	$y \times N$	1	0·8556
	$y \times a/\sigma$	1	0·7922
	$R^* \times N$	1	1·2996
	$R^* \times a/\sigma$	1	3·3672*
	$N \times a/\sigma$	1	28·4622***
	$y \times R^* \times N$	1	0·1981
	$y \times R^* \times a/\sigma$	1	0·9216
	$y \times N \times a/\sigma$	1	0·1600
	$R^* \times N \times a/\sigma$	1	1·1557
	$y \times R^* \times N \times a/\sigma$	1	0·2208
	Theoretical error	—	0·8207

*, *** Significant at 5 and 0·1% levels, respectively.

loci separated by more than 10 map units, by comparison with the reduction in response due to restricted population size alone. With $y = 0{\cdot}10$, for example, reductions in response of 20, 30, and 40% due to finite population size alone become reductions of 22·6, 32·7, and 42·5% respectively under the combined influence of genetic sampling and linkage.

For a pair of loci of large effect separated by less than five map units, the reduction in response due to linkage may clearly be quite appreciable in small

TABLE 10

ANALYSIS OF VARIANCE OF OBSERVED RESPONSE (R) IN EXPERIMENTS 1, 2, AND 3

Experiment No.	Source of Variation	Degrees of Freedom	$10^3 \times$ Mean Square
1	Recombination value (y)	3	28·9260***
	Initial frequency (p_2)	5	295·3817***
	Regimes	2	0·9486***
	$y \times p_2$	15	0·4419***
	$y \times$ regimes	6	0·1338
	$p_2 \times$ regimes	10	0·1301
	$y \times p_2 \times$ regimes	30	0·1508
	Theoretical error	—	0·1206
2	Recombination value (y)	2	13·1602***
	Population size (N)	2	0·1631
	Selection intensity (P)	1	0·1013
	Proportionate effect (a/σ)	1	4·5518***
	$y \times N$	4	0·1231
	$y \times P$	2	0·0388
	$y \times a/\sigma$	2	0·8502**
	$N \times P$	2	0·0371
	$N \times a/\sigma$	2	2·8747***
	$P \times a/\sigma$	1	0·6184
	$y \times N \times P$	4	0·3010
	$y \times N \times a/\sigma$	4	1·0476***
	$y \times P \times a/\sigma$	2	0·1139
	$N \times P \times a/\sigma$	2	0·1201
	$y \times N \times P \times a/\sigma$	4	0·0432
	Theoretical error	—	0·1676
3	Recombination value (y)	1	12·4322***
	Theoretical limit (R^*)	1	136·7520***
	Population size (N)	1	1·4364**
	Proportionate effect (a/σ)	1	0·0001
	$y \times R^*$	1	0·8180*
	$y \times N$	1	0·1254
	$y \times a/\sigma$	1	0·5905*
	$R^* \times N$	1	0·4368
	$R^* \times a/\sigma$	1	0·9364*
	$N \times a/\sigma$	1	1·6892***
	$y \times R^* \times N$	1	0·0980
	$y \times R^* \times a/\sigma$	1	0·0102
	$y \times N \times a/\sigma$	1	1·3988**
	$R^* \times N \times a/\sigma$	1	1·4822**
	$y \times R^* \times N \times a/\sigma$	1	0·2402
	Theoretical error	—	0·1446

*, **, *** Significant at 5, 1, and 0·1% levels, respectively.

populations. With $y = 0{\cdot}05$, for instance, reductions in response of 20, 30, and 40% due to finite population size alone become reductions of 26·0, 36·3, and 45·5%

respectively when the effects of linkage are introduced, while at $y = 0{\cdot}01$ the corresponding reductions are 31·8, 41·2, and 49·5%.

In Table 10 is presented the analysis of variance of observed response in experiment 2. As previously outlined, the purpose of this experiment was to compare regimes involving genes of either small or large effect in populations of small or moderately large size ($N = 10$, 20, 40). Specifically the aim was to determine whether linkage effects on total response due to genes of *small* effect in moderately *large* populations, were of the same order of magnitude as those observed in experiment 1. All tested combinations corresponded to an R^* value of 0·70 [see equation (3)].

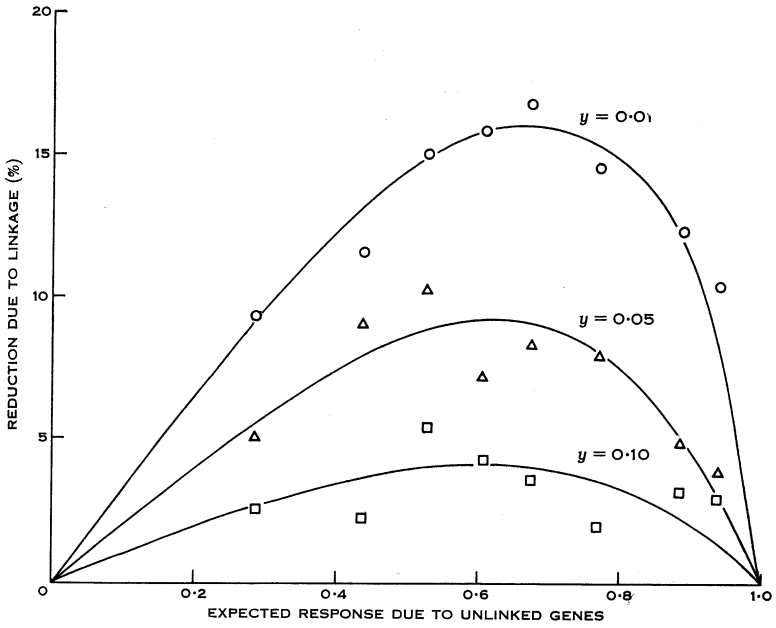

Fig. 1.—Reduction in expected total response due to linkage, averaged over three regimes involving loci of proportionate effect 0·25–0·50, with selection of moderate to high intensity in breeding populations of effective size 5–10.

The analysis shows the important interactions involving linkage to be $y \times a/\sigma$ and $y \times N \times a/\sigma$, so that the reduction in response due to linkage depends on the particular combination of N and a/σ concerned. Averaging over the two selection intensities and three population sizes in Table 7, the reduction in response at $y = 0{\cdot}01$ for genes of proportionate effect equal to 0·5 is $0{\cdot}082 \pm 0{\cdot}007$ or 11·8%. The corresponding reduction for genes of proportionate effect equal to 0·1 is considerably less, viz. $0{\cdot}048 \pm 0{\cdot}007$ or 6·9%.

Note especially from the data in Table 7 that the maximum reduction in response due to linkage in experiment 2, viz. $0{\cdot}126 \pm 0{\cdot}013$ or 18·2% at a recombination value of 0·01, corresponds to the combination $N = 10$, $a/\sigma = 0{\cdot}5$, which is typical of the regimes tested in experiment 1. The effects of linkage in all other tested combinations of experiment 2 were considerably less marked.

The purpose of experiment 3 was simply to check that the conclusions from experiment 2 were applicable at another level of expected response, viz. $R^* = 0{\cdot}50$. The analysis presented in Table 10 shows the $y \times a/\sigma$ and $y \times N \times a/\sigma$ mean squares to be again significant. In addition the $y \times R^*$ interaction is also significant, due to the fact that the absolute reduction in response at $R^* = 0{\cdot}50$ (Table 8) is in each case less than the corresponding reduction at $R^* = 0{\cdot}70$ (Table 7). The maximum linkage effect for $R^* = 0{\cdot}50$ was again observed with $N = 10$, $a/\sigma = 0{\cdot}5$.

V. Discussion

Possibly the most interesting outcome of this study of the effects of linkage on the expected total response due to a pair of additive genes has been the demonstration that the theory developed by Kimura (1957) has a wide range of applicability to populations under artificial selection, provided close linkage is not involved. The observed responses for unlinked genes in each of the three experiments conducted (Tables 4–8) are all remarkably close to the predictions given by equation (1), despite the fact that genes of proportionate effect as large as 0·5 are involved, in populations ranging in effective breeding size from 5 to 40. It has been shown that the expected response due to a pair of linked genes is progressively reduced as the degree of linkage is intensified, but the magnitude of the effect observed in these experiments has been appreciable only at recombination values less than 0·10.

One would anticipate *a priori* that linkage between loci would have its greatest influence on the expected limit to selection (*a*) when the gamete carrying both favoured alleles (i.e. A_2B_2) is initially rare, either because of linkage disequilibrium in the base population, or because of the low frequency of alleles A_2 and B_2; and (*b*) when the expected time span of the selection process is short, thereby reducing the total probability of the formation of the desired gamete by repulsion-phase heterozygotes. The experiments discussed in this paper have shown linkage to have virtually no effect on the probability of fixation of the least favoured gamete A_1B_1 (Table 9), the reduction in response being due to the fixation of gametes A_1B_2 and A_2B_1 more frequently than would be expected in the case of unlinked genes. Those replicates which reach a stage wherein the gamete A_2B_2 can arise only by means of a recombinational event in repulsion heterozygotes, are then presumably responsible for the reduced response.

Figure 1 summarizes the extensive data available on the magnitude of the reduction in response due to linkage, for genes of fairly large effect in small populations (Tables 4–6). The effect is most pronounced when the response due to unlinked genes is in the vicinity of 60–70% of the maximum possible advance, which corresponds to a probability of fixation of the gamete A_2B_2 close to 0·5. The reduction in response is quite spectacular for genes located one map unit apart on the chromosome, but it may be relatively unimportant even for a recombination value as low as 0·05, if the effects of population size alone are such as to reduce the expected response by 40% or more.

The results of experiments 2 and 3 (Tables 7 and 8) indicate that at the same level of expected response R^*, the reduction in response due to linkage is less pronounced for genes of smaller proportionate effect, or for populations of greater

effective breeding size, than those involved in experiment 1. It is therefore probable that for pairs of additive genes of proportionate effect less than 0·5 in populations of effective size greater than 5, the results summarized in Figure 1 represent *maximum* linkage effects, provided the proportion of individuals selected for breeding is not less than 10%.

VI. References

FRASER, A. S. (1957).—Simulation of genetic systems by automatic digital computers. II. Effects of linkage on rates of advance under selection. *Aust. J. Biol. Sci.* **10**: 492–9.

KIMURA, M. (1957).—Some problems of stochastic processes in genetics. *Ann. Math. Stat.* **28**: 882–901.

LATTER, B. D. H. (1965).—The response to artificial selection due to autosomal genes of large effect. I. Changes in gene frequency at an additive locus. *Aust. J. Biol. Sci.* **18**: 585–98.

MARTIN, F. G., and COCKERHAM, C. C. (1959).—High speed selection studies. In "Biometrical Genetics". (Ed. O. Kempthorne.) pp. 35–45. (Pergamon Press: New York.)

MATHER, K. (1951).—"Statistical Analysis in Biology." (Methuen: London.)

ROBERTSON, A. (1960).—A theory of limits in artificial selection. *Proc. R. Soc.* B **153**: 234–49.

ROTENBERG, A. (1960).—A new pseudo-random number generator. *J. Ass. Comp. Mach.* **7**: 75–7.

WRIGHT, S. (1951).—The genetical structure of populations. *Ann. Eugen.* **15**: 323–54.