Theoretical and Applied Genetics 40, 88–94 (1970)

Effects of Selection and Drift on the Dynamics of Finite Populations

III. Times to Fixation or Loss of an Allele in the Case of Multiple Loci and Variable Population Size

R. F. NASSAR and L. SCOTT

Department of Statistics and Computer Science, Kansas State University, Manhattan (USA)

Summary. Selection, in the case of a variable finite population size and a two-allelic locus with overdominance, caused an acceleration in the time to fixation or loss of the favorable allele (i.e. time with selection was less than that with no selection) when the deterministic gene frequency equilibrium was above 0.8. The acceleration was over a range of initial gene frequencies, dependent on the selection intensity and the overdominance parameter.

In the case of multiple loci and a small, diploid population of fixed size derived from a large population in initial linkage equilibrium, an acceleration in the time to fixation or loss occurred over a range of initial gene frequencies (as in the one locus case) for strong selection intensity ($N s > 14$) and weak overdominance effect. For a large number of overdominant loci, acceleration did not occur under linkage. Initial coupling or repulsion disequilibrium with independent assortment had no effect on the observed acceleration. Repulsion with linkage, however, caused a retardation in the time to fixation or loss.

Some populations in nature and many in the laboratory as well as in breeding programs contain a small number of individuals. It is important to note that under these conditions the success or failure of a gene is determined not only by its fitness but also by the random sampling of gametes due to the finite gene pool (random genetic drift). This concept was first investigated by Fisher in 1922 and by Wright (1931 and later). Wright was the first to use the Fokker-Plank equation (Kolmogorov forward equation) in his solution of gene frequency distributions under the joint effect of selection and random genetic drift. Exact solutions for the entire process of change of gene frequencies were formulated by Kimura (1955 and later) using a continuous Markov chain in time and space.

Carr and Nassar (1970), using a finite Markov chain to describe the stochastic change in gene frequencies in a finite population of fixed size, solved for the effects of selection and random sampling of gametes on the expected time to fixation or loss of a favorable allele in the one locus, two allele case. Of particular relevance to this work is the finding that selection for overdominance accelerated instead of retarded the time to fixation or loss when the equilibrium gene frequency in an infinite population was above 0.8 or below 0.2, equivalent findings were also reported by Robertson (1962). This means that we can no longer rely under all conditions on overdominance as a mechanism for retarding gene fixation in finite populations.

It is essential to know whether the same results would be true under the more complicated, but realistic assumptions of (1) variable population size and (2) multiple loci. Hence it is the purpose of this paper to examine for assumptions 1 and 2 the effects of small population size and selection on the time to fixation or loss of a favorable allele in the case of overdominance.

1. Variable Population Size

Consider a random mating population of size N with nonoverlapping generations. Let the relative fitness of the three genotypes AA, Aa and aa be $1 + s$, $1 + hs$ and 1, respectively. (s is the selection coefficient and h is the dominance factor.) Assume also that the $2N$ gametes in adults of a given generation are a random sample of the infinite number produced (or potentially produced) by adults of the previous generation. If N is fixed from generation to generation, then the process of change in gene frequency is a finite Markov chain to which solutions are known. When N is a random variable, however, the process of change is a branching process to which analytical solutions are yet to be determined.

In this formulation we shall let N be a Poisson random variable and J (the number of A alleles) be a binomial random variable. In generation t_0 let

$$n_0 = \text{population size}$$
$$j_0 = \text{number of } A \text{ genes}$$

then

$$P[N_1 = n_1, J_1 = j_1 | N_0 = n_0, J_0 = j_0] = \Phi(n_1, n_0\lambda)\, B(j_1, n_1; j_{0/2n_0}) = P_1 \qquad (1.1)$$

is the probability that the population size in generation t_1 is n_1 and the number of A alleles is j_1 given that they were n_0 and j_0 in generation t_0 respectively.

in (1.1) $$\Phi(n_1, n_0\lambda) = \frac{\bar{e}^{n_0\lambda}(n_0\lambda)^{n_1}}{n_1!}, \quad \lambda = 1$$

and

$$B\left(j_1, n_1; \frac{j_0}{2n_0}\right) = \frac{n_1!}{(n_1 - j_1)!\, j_1!}\left(\frac{j_0}{2n_0}\right)^{j_1}\left(1 - \frac{j_0}{2n_0}\right)^{n_1 - j_1}.$$

With selection, $j_0/2\,n_0$ is replaced by

$$\frac{2\,n_0 j_0\,(1 + h\,s) + j_0^2\,s\,(1 - h)}{4\,n_0^2 + 4\,n_0 j_0\,h\,s + j_0^2\,s\,(1 - 2\,h)}. \qquad (1.2)$$

Expression (1.2) is the frequency of allele A in the infinite population after selection. In generation t_2 the probability that the population size is n_2 and the number of A alleles j_2 given that they were n_1 and j_1 in t_1 is

$$P_2 = \sum_{n_1=0}^{M} \sum_{j_1=0}^{2n_1} \Phi\,(n_2, n_1\,\lambda)\; B\left(j_2, n_2; \frac{j_1}{2\,n_1}\right) P_1. \qquad (1.3)$$

M in the summation is a finite upper limit for population size taken such that $\Phi\,(M, n_1\,\lambda) \to 0$.

In t_3, P_3 becomes

$$P_3 = \sum_{n_2=0}^{M} \sum_{j_2=0}^{2n_2} \Phi\,(n_3, n_2\,\lambda)\; B\left(j_3, n_3; \frac{j_2}{2\,n_2}\right) P_2. \qquad (1.4)$$

After a finite number of generations the A allele will with certainty attain the frequency 0 or 1.

It is clear that expressions (1.1), (1.3) and (1.4) get more complicated every generation and that no analytical solution is possible. For this reason a numerical solution was obtained on the computer. To avoid difficulties in solution it was assumed that in any generation the population does not go to extinction and does not also go beyond an upper limit L which was set to be four times n_0, $0 < n \leq L$. It was felt that this assumption of an upper and lower limit on a population that has reached equilibrium ($\lambda = 1$) is not an unrealistic assumption.

Solutions on the computer was found for initial population sizes $n_0 = 10$ and $n_0 = 20$ and for different initial gene frequencies. The cumulative probabilities of $n_1 = 1, 2, \ldots, L$ was computed stepwise from the Poisson $\Phi\,(n_1, n_0\,\lambda)$, corrected for the probabilities $P(0)$ and $P(n_1 > L)$ and compared with a uniform random number between 0 and 1.

If the random number was equal or less than the cumulative probability for an n_1 value, that value was chosen as the population size in generation t_1.

The next step was to determine the number of A alleles in t_1. This was similarly done by generating the cumulative probabilities stepwise for $j_1 = 0, 1, \ldots, 2\,n_1$ from the binomial $B\,(j_1, n_1; j_0/2\,n_0)$. For selection $j_0/2\,n_0$ was replaced with expression (1.2). The cumulative frequencies were compared with a uniform random number between 0 and 1. When the random number was equal or less than the cumulative frequency for a j_1 value, that value was chosen as the number of A alleles in t_1. This iteration was continued until the frequency of the A allele was 0 or 1. One hundred replications were run and the average number of generations until fixation or loss was computed based on the 100 replications.

2. Multiple Loci

A program was written for a digital computer to simulate a diploid population of finite fixed size N with nonoverlapping generations. The genetic model described earlier was used for 2, 3, 5 and 15 loci. An individual genotypic value was obtained as the sum of the appropriate values for the n-loci, $G = \sum_{i=1}^{n} G_i \cdot G_i$ was assumed equal for all loci and was 0 for aa, 1.125 for Aa and 1 for AA. The phenotype of an individual was given by the expression

$$P_j = \sum_{i=1}^{n} G_i + e_j, \qquad (2.1)$$

where

n = the number of loci

e_j = a random environmental effect simulated by adding twelve uniform random deviates ($-1/2$, $1/2$) and multiplying by a constant C. Thus each e_j is distributed as an independent normal random variable with mean zero and variance C^2.

The phenotypic variance in the population was given by

$$\sigma_P^2 = \sigma_G^2 + \sigma_e^2. \qquad (2.2)$$

The genotypic value at a locus was determined such that the equilibrium gene frequency (q) in an infinitely large population would be 0.9 ($\hat{q} = h/2\,h - 1$, $h = 1.125$), since from previous work (Carr and Nassar, 1970) with one locus, strong selection in the case of overdominance ($\hat{q} = 0.9$, or $h = 1.125$) caused the expected time of fixation or loss of the favorable allele to be smaller than that for the case of no selection ($s = 0$). This difference was also of relatively large magnitude such that, if present in the case of multiple loci, it can be easily detected by simulation.

Populations of different sizes were generated and pair mating was at random with one sire mated to one dam. Selection was directional ($s > 0$) or random ($s = 0$) and the size of the population was kept constant. To illustrate a 4-1-4 population stands for four pair matings with four offsprings (2 males and 2 females) per mating. For the population size to remain constant, four males and four females were selected as parents of the next generation. A population under consideration was regarded as

1. being initially a random sample from a hypothetically large population in linkage equilibrium.

2. being a cross between two homozygous lines in coupling or repulsion phase.

The first assumption was justified on the basis that the frequency of the favorable allele at each locus was determined independently of other loci. The second assumption was met by generating each individual with a genotypic array (111.../000...), coupling or (1010.../0101...), repulsion. As a consequence, the gene frequency was $1/2$ in a population.

The effective population number ($N\,e$) for a population under selection was computed from Kimura and Crow (1963) as

$$N\,e = \frac{(N_{t-1} - 1)\,\bar{k}}{1 + V_k/\bar{k}}$$

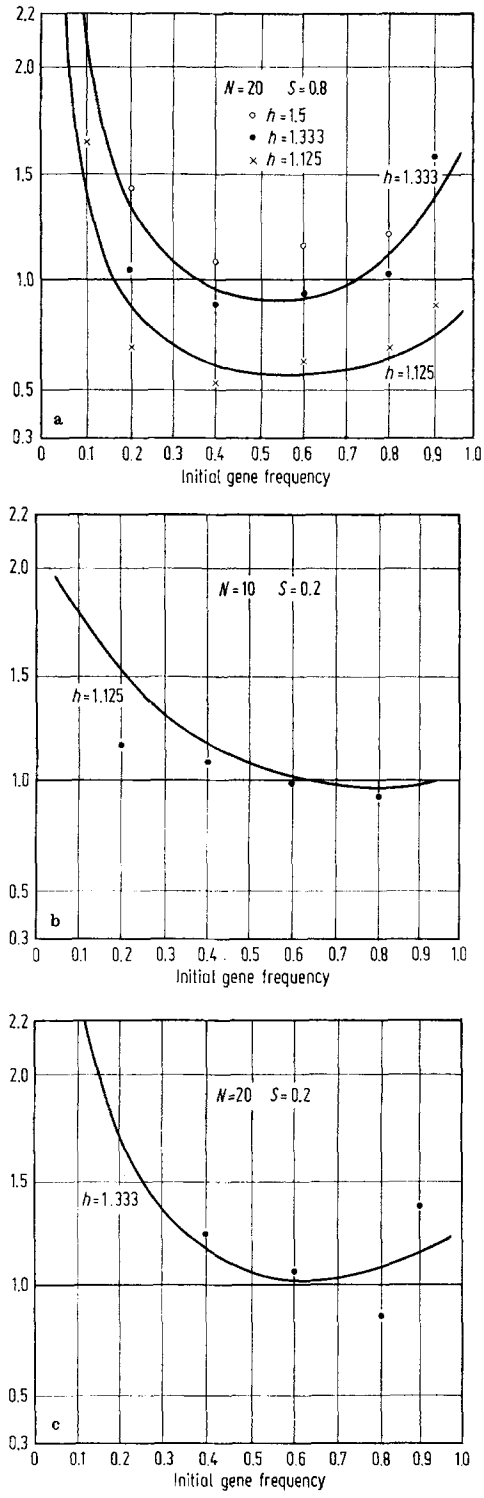

Fig. 1 a, b, c. The ordinate is the ratio of the time to fixation or loss with selection ($s > 0$) to that with no selection ($s = 0$). The abscissa is the initial gene frequency. The curves are for different overdominant (h) values. In each graph portions of the curves below the horizontal line represent acceleration in the time to fixation or loss. Those above the line represent retardation in the time to fixation or loss. Solid curves are for the case where N, the population size, is assumed constant from generation to generation. Solid points are for N variable

where $\bar{k}$ = the mean progeny number per parent and V_k = the variance in progeny number.

The effective population number, as defined, is for genes under no selection. *Ne* would decrease with directional selection, but probably not enough to be considered consequential.

Selection was practiced in each population until fixation or loss of the favorable allele. The selection coefficient was determined as follows: Assuming normality of the phenotypic distributions of genotypes with equal variances, then

$$1 + s = \frac{\frac{1}{\sqrt{2\pi\sigma_p^2}} \int_{\mu}^{\infty} e^{-1/2 \frac{(x-\mu_1)^2}{\sigma_p^2}}}{\frac{1}{\sqrt{2\pi\sigma_p^2}} \int_{\mu}^{\infty} e^{-1/2 \frac{(x-\mu_2)^2}{\sigma_p^2}}}, \qquad (2.3)$$

where

μ = the mean of the genotype values in the population and can be easily determined from G_i and from the allelic frequency at each locus

μ_1 = genotype value of $1/1 = 1$

μ_2 = genotype value of $0/0 = 0$

The integration is from μ to ∞ since 50% of the population was selected. From expression (2.3) it is seen that, for fixed genotypic values, different values for s can be obtained by varying σ_p^2. The selection coefficient, as tabulated, is for the first generation before selection. It should be realized that s or $N\,s$ will not be constant during selection, but will vary due to the finiteness of the population size and due to the fact that σ_p^2 is not constant. This variation, however, is probably more random than systematic and $N\,s$, as tabulated, can perhaps be considered as an average value over generations.

A population under selection was replicated and the time to fixation or loss was the average number of generations over replications. Runs were made under independent assortment and linkage. For linkage all loci were equally spaced on a chromosome that varied in length according to whether linkage was tight or loose. For two, three and five loci, one cross-over was allowed. For fifteen loci, the loci were divided into three groups of five loci each. One cross-over was allowed per group and only second cross-over between groups. If a first cross-over occurred in one group, the probability of a second cross-over occurring in another group was determined as the product of the first cross-over probabilities in the two groups in question. Cross-overs were directional, proceeding from one end to the other on a chromosome.

3. Results

Figure 1 a, b, c have as ordinate the ratio of the time to fixation or loss with selection ($s > 0$) to that with no selection ($s = 0$). The abscissa is the initial

gene frequency in the population. Portions of the graphs below the horizontal line of ratio 1 represent acceleration in the time to fixation or loss. Those above the line represent retardation in the time to fixation or loss. The solid curves are from previous results by Carr and Nassar (1970) obtained under the assumption of constant population size from generation to generation. They are presented here for comparisons with the solid points in the graphs. The latter are from this study and are obtained under the assumption that the population size is allowed to vary from generation to generation in accord with the Poisson law. Carr and Nassar (1970) reported that for overdominance gene action, acceleration and retardation in the time to fixation or loss occurred above and below the deterministic gene frequency equilibrium of 0.8 ($h = 1.333$) respectively. It was important to find also that in all cases an acceleration occurred over an interval of initial gene frequencies.

From the graphs in figure 1 a, b, c it is clearly seen that the same phenomenon of acceleration or retardation in the time to fixation or loss as compared with no selection occurs when the population size is allowed to vary. It is difficult in the case of variable population size to determine the exact range of initial gene frequencies over which acceleration occurs, since the results are only estimates and are further limited to few points. However, it appears that for N variable the results are in good agreement with the exact results (solid curves) obtained for a constant N. It is also interesting to note that the time to fixation or loss with no selection ($s = 0$) was longer for population size fixed than for population size variable (table 1).

Table 1. *The time to fixation or loss of an allele under no selection and with population size (N) variable or fixed*

Initial Gene Frequency	N			
	Constant 10	Variable 10	Constant 20	Variable 20
0.6	25.44	16.25±1.47	52.14	35.92±3.6
0.7	23.02	16.77±1.68	47.23	35.75±3.2
0.8	18.74	15.59±1.75	38.52	35.35±4.6
0.9	11.98	9.42±1.2	24.75	17.05±2.51

In the case of two loci, at an initial gene frequency of 0.8, in linkage equilibrium and under independent assortment (table 2a, b), the time to fixation or loss, $E(t)$, for $N_e s = 2$ and 8 was not significantly different from $E(t)$ for no selection, $s = 0$. Linkage also had no significant effect on the time to fixation or loss. In all cases, an increase in the selection intensity ($N_e s = 2$ to $N_e s = 8$) caused a reduction in the time to fixation or loss as compared to the case of no selection. This reduction was not significant. For stronger selection intensity ($N_e s = 14, 16$, table 2b), however, the reduction was significant causing an acceleration in the time to fixation or loss (note that

Table 2a, b, c, d. *Time to fixation or loss, $E(t)$, of a favorable allele in the multi-locus case with overdominance ($h = 1.125$) and initial linkage equilibrium. P_0 is the initial frequency of the favorable allele and is equal at all loci, I.A. is independent assortment. Linkage is presented as the probability of a first cross-over between the two end loci on a chromosome*

Table 2a. *Population is 4-1-4 ($N e = 10$), number of loci = 2, $P_0 = .8$*

Linkage	Reps.	Selection s	$N_e s$	$E(T)$
I.A.	56	0.0	0	26.64±2.99
	40	.2	2	27.85±3.96
	40	.8	8	23.32±2.93
.05	40	0.0	0	25.88±2.92
	70	.2	2	31.13±2.94
	40	.8	8	20.72±2.42
.25	40	.2	2	34.32±4.90
	40	.8	8	19.82±3.29

Table 2b. *Population is 8-1-4 ($N_e = 20$), number of loci = 2*

P_0	Linkage	Reps.	Selection s	$N_e s$	$E(T)$
.8	I.A.	56	0.0	0	56.96±5.44
		47	.4	8	50.36±4.71
		40	.7	14	41.03±3.29
		40	.8	16	33.05±3.89
	.05	50	0.0	0	53.27±5.95
		48	.8	16	38.42±4.20
	.25	40	.8	16	48.58±2.84
.6	I.A.	16	0.0	0	86.31±12.57
		27	.8	16	42.96±4.35
	.25	25	.2	4	69.96±10.45
		49	.8	16	51.14±4.43

Table 2c. *Population is 8-1-4 ($N_e = 20$), $P_0 = .8$*

Number of Loci	Linkage	Reps.	Selection s	$N_e s$	$E(T)$
3	I.A.	16	0.0	0	79.56±13.86
		20	.8	16	45.90±5.74
5	I.A.	24	0.0	0	90.54±13.60
		13	.8	16	59.23±6.80
	.05	20	.8	16	53.75±6.82
	.30	19	.2	4	83.84±9.42
		24	.8	16	52.13±2.69

Table 2d. *Number of loci = 15, $P_0 = .8$*

Population	Linkage	Reps.	Selection s	$N_e s$	$E(T)$
8-1-4	I.A.	10	0.0	0	64.70±9.60
		10	.2	4	60.70±5.87
4-1-4	I.A.	17	0.0	0	42.70±2.85
		10	.2	2	46.60±4.50
		10	.4	4	48.70±3.89
		13	.8	8	31.08±2.33
	.15	24	.4	4	68.08±9.04
		20	.8	8	56.50±6.54

the time to fixation or loss increases with N_e as is well known from theory). This acceleration was still in effect with tight linkage (.05) and $N_e s = 16$, but disappeared with moderate linkage (.25). At an initial gene frequency of 0.6, selection with and without linkage also caused an acceleration in the time to fixation or loss. With 3 and 5 loci (table 2 c) an acceleration in the time to fixation or loss occurred for an $N_e s$ of 16. Here the degree of linkage intensity had no effect on the acceleration or retardation. In the case of 15 loci (table 2d), an acceleration occurred for an $N_e s$ of 8 in variant with the case of 2, 3 and 5 loci. Another important thing to notice is the fact that linkage with an $N_e s$ of 4 did retard the time to fixation or loss. With more intense selection ($N_e s = 8$) the retardation effect was rendered insignificant in accord with the previously observed trend that in the region of acceleration the time to fication or loss decreased as the intensity of selection increased (Carr and Nassar, 1970).

With regard to coupling and repulsion, table 3a clearly shows that the time to fixation or loss in the two locus case was the same regardless of the linkage disequilibrium phase in the population at the outset and regardless of linkage or no linkage. An exception was the case of coupling and tight linkage which reduced the time to fixation or loss. With 5 loci (table 3b) at an initial gene frequency of 0.5, an acceleration occurred for independent assortment and all linkage disequilibrium phases. This was so, since for $s = 0$ the time to fixation or loss was found to be 100.85 ± 7.5. The time to fixation or loss under independent assortment was less for the repulsion phase than for coupling and linkage equilibrium phases. With linkage and repulsion, however, the time to fixation or loss was prolonged and significantly exceeded that under no selection. This relative increase was greatest at the initial gene frequency of 0.5 and had no effect when the time to fixation or loss was measured after the population had reached a gene frequency of 0.8. For two loci in the coupling phase and with tight linkage, the time to fixation or loss was accelerated, while with five loci no such effect was shown. For 15 loci (table 3c), as in the five - locus case, there seemed to be no difference in the time to fixation or loss between coupling, repulsion or linkage equilibrium in the presence of independent assortment. Linkage with coupling, as in

Table 3a, b, c. *Time to fixation or loss of a favorable allele in the multi-locus case with overdominance ($h = 1.125$) and with initial linkage equilibrium (L.E.), coupling or repulsion. Population size = 8-1-4 ($N_e = 20$), $N_e s = 16$. P of 0.5 is the initial frequency of the favorable allele. That of 0.6 and 0.8 are the frequencies of the favorable allele attained by the population through selection. Linkage and I.A. have the same meaning as in table 2*

Table 3a. *Number of loci = 2*

Phase	Linkage	Reps.	P	$E(T)$
L.E.	I.A.	38	.5	50.92±4.64
			.6	48.18±4.68
			.8	39.02±4.33
Repulsion	I.A.	35	.5	45.03±4.46
			.6	42.48±4.43
			.8	34.80±4.50
	.05	35	.5	50.71±3.83
			.6	42.57±3.69
			.8	32.91±3.26
	.25	20	.5	48.80±5.21
			.6	44.70±5.31
			.8	37.45±5.11
Coupling	I.A.	29	.5	49.31±4.60
			.6	45.52±4.36
			.8	34.67±4.07
	.05	20	.5	32.45±3.96
			.6	30.85±3.91
			.8	26.50±3.55
	.25	20	.5	49.45±5.98
			.6	46.85±6.05
			.8	37.45±5.51

Table 3b. *Number of loci = 5*

Phase	Linkage	Reps.	P	$E(T)$
L.E.	I.A.	18	.5	66.61±6.01
			.6	63.56±6.10
			.8	53.94±6.12
Repulsion	I.A.	10	.5	48.70±3.81
			.6	45.90±3.78
			.8	39.30±3.93
	.05	11	.5	137.72±9.12
			.6	100.63±6.79
			.8	54.00±8.32
	.25	7	.5	85.57±6.78
			.6	76.28±6.70
			.8	60.14±8.87
Coupling	I.A.	7	.5	67.71±7.29
			.6	63.86±7.23
			.8	55.29±6.28
	.05	20	.5	59.75±9.27
			.6	58.50±9.24
			.8	53.55±9.16
	.25	20	.5	54.70±5.99
			.6	53.55±5.99
			.8	49.00±5.92

Table 3c. *Number of loci = 15*

Phase	Linkage	Reps.	P	$E(T)$
L.E.	I.A.	10	.5	84.00± 6.78
			.6	80.25± 7.52
			.8	68.25± 7.26
Repulsion	I.A.	10	.5	81.25± 3.17
			.6	77.25± 1.66
			.5	66.75± 2.87
Coupling	I.A.	10	.5	92.00± 2.27
			.6	88.50± 2.33
			.8	80.75± 3.09
	.15	10	.5	85.90±14.28
			.6	84.60±14.29
			.8	78.80±14.58

the five-locus case, did not cause an acceleration in the time to fixation or loss.

4. Discussion

In the case of one locus and a variable population size it appears that acceleration in the time to fixation or loss of a favorable allele occurs for overdominance when the deterministic gene frequency equilibrium is above 0.8 or below 0.2. The acceleration is also over a range of initial gene frequencies. It seems that what was substantially changed by relaxing the assumption of a constant population size is that the time to fixation or loss was reduced. This can be explained by the fact that, for a constant N, the time to fixation or loss is not a linear function of N. The time approaches asymptotically a finite value as N approaches infinity. For such a curvilinear function the time to fixation or loss will be reduced as N is allowed to vary.

It seems clear, in the case of multiple loci, that an acceleration in the time to fixation or loss can occur with and without linkage under relatively high $N\,s$ values (higher than required for the case of one locus) and when overdominance is relatively weak (deterministic gene frequency equilibrium in the neighborhood of 0.9).

If the number of overdominant loci is large, then linkage can prevent an acceleration in the time to fixation or loss from occurring. Linkage disequilibrium can be the determining factor in this regard. It is known that linkage in a small population can give rise to linkage disequilibrium which inflates the degree of overdominance. From studies of the one locus case (Carr and Nassar, 1970), it is known that stronger overdominance than $h = 1.125$ decreases the interval of acceleration and eventually causes retardation in the time to fixation or loss. Thus linkage disequilibrium can be the cause of retardation in these results. The magnitude of linkage disequilibrium is no doubt larger in the 15 locus case than in 2, 3 or 5 locus cases. This can explain why retardation occurred for linkage and 15 loci only. The above argument can be presented mathematically as follows. Consider two loci A_1 and A_2 with two alleles per locus. Let the average values of the three genotypes A_1A_1, A_1a_1 and a_1a_1 be $2u_1$, $a_1u_1 + u_1$ and 0 and of A_2A_2, A_2a_2 and a_2a_2 be $2u_2$, $a_2u_2 + u_2$ and 0 respectively. a_1 and a_2 are measures of the degree of dominance. Let the gametic frequency of $A_1A_2 = p$, $A_1a_2 = r$, $a_1A_2 = s$ and $a_1a_2 = t$. In the 3×3 table of genotypes it is easy to determine the frequency and value for each genotype assuming no epistasis. $A_1A_1A_2A_2$, for example, has a value $2u_1 + 2u_2$ and a frequency p^2. Let $Y_{2\cdot}$, $Y_{1\cdot}$ and $Y_{0\cdot}$ be the marginal values for A_1A_1, A_1a_1 and a_1a_1 respectively. Then it is easy to show that the degree of dominance at the A_1 locus is

$$(2\,Y_{1\cdot} - Y_{2\cdot} - Y_{0\cdot})/2 = a_1u_1 + [a_2u_2(pt - rs)^2/(p + r)^2(s + t)^2]. \quad (4.1)$$

Note that $(pt - rs)$ is a measure of linkage disequilibrium between the two loci in question and $(p + r)^2(s + t)^2$ = the square of one-half the probability of the A_1a_1 genotype. Thus if $(pt - rs)$ is not equal to zero, or the two loci are not independent, the value a_1u_1 is inflated by the amounted $a_2u_2(pt - rs)^2/(p + r)^2(s + t)^2$. Generalizing to n loci the right hand side of (4.1) becomes

$$a_ju_j + \left(\sum_{i\neq j,\,i=1}^{n} a_iu_i(pt - rs)_{ij}^2\right)\Big/(1/2\,P_{Aja_j})^2.$$

It is clear that the inflation in the degree of dominance at the j^{th} locus increases with an increase in the number of loci that are in association with the j^{th} locus. Although the populations discussed so far are regarded as finite random samples from a hypothetically large population in linkage equilibrium, the finiteness of the sample is apt to cause $(pt - rs)$ to be different from zero due to sampling error. Also, for small populations with many loci (n), it was interesting to observe from unpublished results that the sum of squares of $(pt - rs)$ per locus $\left[\sum_{i=1}^{n=1}(pt - rs)^2\right]$ increases instead of decreasing with random mating and selection.

For the conditions that were investigated it was certain that linkage disequilibrium was at a minimum due to the fact that populations were started as a random sample from a hypothetically large population in linkage equilibrium. As a result, the effects of linkage disequilibrium in preventing an acceleration in the time to fixation or loss were only manifested with linkage and a large number of loci. It should be noted, however, that due to the genetic load involved a population may not be able to maintain a large number of overdominant loci, and hence the role of linkage disequilibrium may not be so likely to arise.

It was of interest, therefore, to determine whether a retardation in the time to fixation or loss would take place in the case of a smaller number of loci if the population at the outset was in linkage disequilibrium. For that purpose the two extreme initial situations of linkage disequilibrium, coupling and repulsion, were investigated. The conclusions to be drawn is that initial coupling or repulsion under independent assortment did not cause any deviations from results obtained for populations with initial linkage equilibrium. It seems that under the conditions of independent assortment, random mating and no epistasis a population in initial repulsion or coupling disequilibrium can rapidly approach linkage equilibrium. With tight linkage, however, linkage equilibrium is not attained as rapidly and genes in repulsion can cause a retardation in the time to fixation or loss. This retardation can occur for a relatively small number of loci at intermediate initial gene frequencies due to the large magnitude

of linkage disequilibrium generated under these conditions.

It was interesting to observe that only for two tightly linked genes in coupling or in linkage equilibrium an acceleration in the time to fixation or loss occurred when the selection intensity was relatively strong, $N s = 16$. This phenomenon may be explained by the possibility that two genes can go to fixation in a group as a whole and that this becomes less likely as the number of loci increases.

Acknowledgement

This work was supported in part by a Senior Research Fellowship from the Alexander von Humboldt-Stiftung while one of us (R. Nassar) was at the Institut für Forstgenetik, 207 Schmalenbeck, Germany (BRD) from September, 1969 till June, 1970.

Zusammenfassung

Im Falle eines begrenzten, jedoch variablen Populationsumfangs verursacht die Selektion bei Vorliegen eines Locus mit 2 Allelen und Superdominanz eine Akzeleration in der Zeit bis zur Fixierung oder dem Verlust des begünstigten Allels (d. h. die Zeit mit Selektion war niedriger als ohne Selektion), wenn das deterministische Genfrequenz-Gleichgewicht über 0,8 war.

Die Akzeleration erstreckt sich über einen Bereich ursprünglicher Genfrequenzen, sie hängt ab von der Selektionsintensität und dem Parameter der Superdominanz.

Im Falle multipler Loci und einer kleinen diploiden Population fixierten Umfangs, die aus einer großen Population mit ursprünglichem Koppelungsgleichgewicht abgeleitet wurde, trat eine Akzeleration in der Fixierungs- bzw. Eliminationszeit über einen Bereich ursprünglicher Genfrequenzen (wie beim unilokalen Fall) bei hoher Selektionsintensität ($N s > 14$) und schwachem Superdominanzeffekt auf. Für eine große Zahl superdominanter Loci tritt unter Koppelungsbedingungen keine Akzeleration auf. Ein ursprüngliches Attraktions- oder Repulsions-Ungleichgewicht mit unabhängiger Genverteilung hat auf die beobachtete Akzeleration keinen Einfluß. Repulsion mit Koppelung verursacht dagegen eine Retardierung der Fixierungs- bzw. Eliminationszeit.

References

1. Carr, R. N., Nassar, R. F.: Effects of selection and drift on the dynamics of finite populations. II. Expected time to fixation or loss of a favorable allele. Biometrics (1970, in press). — 2. Fisher, R. A.: On the dominance ratio. Proc. Roy. Soc. Edin. **42**, 321—341 (1922). — 3. Kimura, M.: Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbor Symp. **20**, 33—53 (1955). — 4. Kimura, M., Crow, J. F.: The measurement of effective population numbers. Evolution **17**, 279—288 (1963). — 5. Robertson, A.: Selection for heterozygotes in small population. Genetics **47**, 1291—1300 (1962). — 6. Wright, S.: Evolution in Mendelian populations. Genetics **16**, 97—159 (1931).

Received January 21, 1970

Communicated by W. Seyffert

R. F. Nassar
Institut für Forstgenetik
Sieker Landstr. 2
207 Schmalenbeck (Germany, BRD)

L. Scott
Department of Statistics and Computer Science
Kansas State University
Manhattan 66502 (USA)