Simulation of Models Involving Mixed Selfing and Random Mating

II. Effects of Selection and Linkage in Finite Populations*

S. K. JAIN

Department of Agronomy, University of California, Davis

Summary. The joint effects of linkage, inbreeding, and drift due to finite population size were investigated in terms of population changes under selection involving gene interaction. Six-locus models with the same amount of recombination between adjacent pairs of loci, mixed selfing and random mating, and selection of basically three forms (heterotic, optimizing and mixed optimum-heterotic) were used for Monte Carlo simulation. The results were primarily described in terms of certain measures of gene dispersion, genetic variability, gametic unbalance (linkage disequilibrium) and the approach to stable gene frequency equilibria. Under both cumulative and diminutive heterosis models, a steady state with polymorphisms could be attained with random gene dispersion being small and different replicate populations evolved high degrees of gametic unbalance in the direction of excess of either coupling or repulsion phase linkages depending on the random drift in gene frequencies. Under optimum models, on the other hand, all populations approached steady decay toward fixation at all loci although gene dispersion was governed by rather complex interactions between the parameters of selfing, linkage and selection intensity. Gene dispersion was not necessarily proportionately greater with the higher levels of inbreeding. An excess of repulsion linkages with mean population fitness approaching unity was noted in all runs with the optimum models, more so with tight linkage and heavy inbreeding. Any asymmetry in the sense of selection favoring one or the other hand, seemed to play a relatively greater role under inbreeding in terms of retaining heterozygosity. Mixed optimum-heterotic models provide a favorable compromise between these conflicting attributes of multilocus systems in terms of the maintenance of polymorphisms and the maximization of fitness in relation to certain optimal linked gene complexes. In general, for moderate to large population size these results are, as expected, in line with those reported previously

A major development in population genetic theory during the past decade has resulted from the analyses of complex polymorphisms involving selection and linkage in multilocus models. The work of Kimura (1956), Lewontin and Kojima (1960) and of Bodmer and Parsons (1962) on the two-locus, deterministic models for a random mating population clearly provided the framework in which the role of linkage and epistasis could be understood both in terms of the maintenance of complex polymorphisms and the evolution of linked gene systems. This was further extended for various selective processes by NEI (1964), KIMURA (1965), FELSENSTEIN (1965), and others. Recently, Bodmer and Felsenstein (1967) investigated a general model for determining analytically the conditions for stable equilibria with or without permanent linkage disequilibrium. A significant concluding statement was that "plausible situations do exist for which linkage can profoundly affect the selection and ultimate equilibrium of new genetic combinations". Wright (1952, 1965) has put forth a different viewpoint in which the role of linkage in gene systems with none or little overdominance is to make the saddles (unstable equilibria) on the adaptive topography shallower such that under conditions of changing environment and therefore, shifting positions of selective peaks, local populations can move to such peaks (stable equilibria) across the saddles. In any event, linkage and epistasis can significantly influence the rates and trajectories of gene frequency changes so that at least in relation to the short-term evolutionary changes these factors are likely to be important.

Several papers have appeared more recently on the rates of advance under selection in multilocus systems in which random drift due to finite population size and linkage together govern the rates of fixation, limits of response and the maintenance of genetic variability in a rather complex manner (GILL, 1965; Latter, 1966; Young 1966). Fraser, Miller and Burnell (1965) discussed briefly the evolution of polygenic balance in finite populations under optimizing selection and showed that population size and rate of recombination interact to yield optimal conditions for evolving the repulsion linkages characteristic of the so-called relational balance. Epistasis can promote or retard random gene dispersion depending on the model: in general, under a symmetrical optimum model where marginal combinations of gene frequencies represent unstable equilibria, it reinforces drift but vice versa holds for a symmetrical heterotic model (cf. Bodmer and Felsenstein, 1967).

Inbreeding occurs, in many plant species that practice mixed selfing and random mating. The essential properties of a single locus heterotic model for this system were first described by HAYMAN (1953) who showed that given sufficient overdominance, stable equilibria can be maintained by selection balanced against homozygosis due to inbreeding. JAIN and ALLARD (1965, 1966) investigated the role of linkage and epistasis under mixed mating by computer simulation of two-locus deterministic models. In general, inbreeding seems to accentuate epistatic effects so that relatively a wider range of conditions allow the maintenance of permanent linkage disequilibrium (gametic unbalance). The purpose of this paper is to report on the joint effects of finite

^{*} This work was supported by a grant (GM 10476) from the U.S. Public Health Service. I am indebted to Dr. D. G. Burnell for his generous help in the use of his computer program, and to Drs. R. W. Allard and A. S. Fraser for many helpful suggestions.

population size and partial selfing on certain properties of the multilocus systems (gene dispersion, heterozygosis, gametic unbalance). Specifically, results from simulation of intermediate optimum, heterotic and mixed optimum-heterotic models are described here in order to compare with our earlier findings from the two-locus deterministic runs.

Models and Procedures

(i) The program. — The computer program (GS-Dl), written in Fortran IV by Dr. D. G. BURNELL (see Fraser, et al., 1965; Fraser and Burnell 1968), allows Monte Carlo simulation of a six-locus, diallelic genetic system under zygotic selection in a diploid population of specified size using basically the same procedures as described by GILL and CLEMMER (1966) and Bellmann and Ahrens (1966). Random mating is simulated by sampling of parents of each sex with replacement, gametes are produced in the binary notation using a random walk along the genotype and the selective values (w_i) of various genotypes are defined by some function of a phenotypic score, the number of heterozygous loci, or some combination of these parameters as well as random environmental components. Epistasis can be introduced in the selection models by an appropriate choice of these functions. A mixed selfing and random mating system is obtained by combining the random mating and selfing (both gametes produced from the same parent) components. In this study, six-locus models were used with the same recombination value (c) between all pairs of adjacent loci and completely additive primary scale with metric value of phenotypes (x) ranging from 0 to 12 (at each locus genotypes 11, 10 and 00 have 2, 1 and 0 score respectively). Optimizing selection is defined in terms of a score (z) taken as negatively correlated with the deviations of metric values (x) from an intermediate optimum value (φ) (e.g. in quadratic deviation model, $z = 1 - (x - \varphi)^2/\varphi^2$). Under heterotic models, the relative fitness (w_i) is defined as a function of the number of heterozygous loci (y), irrespective of the phenotypic score; the combinations of optimum and heterotic models are then obviously given by some function of both y and z scales. For each model, a subprogram is written for defining the selective values. Individual offsprings are generated and selected one by one in order to maintain population size constant in each generation.

- (ii) Selection models. Two types of heterotic models involving epistasis were defined as follows: cumulative heterosis (model 1) with $w_i = y(y+1)/2 a$, and diminutive type (model 2) with $w_i = y \, a - y \, (y+1) a \, b$, where a and b are constants chosen to scale all w. values between 0 and 1 (cf. JAIN and ALLARD, 1966). Directional selection for, say, allele 1 is superimposed on them simply by taking w_i as some function of both x and y (model 3). Three well-known forms of optimum model were used: linear deviation from optimum φ (model 4), quadratic deviation from φ (model 5) and double truncation (model 6) against the extremes along the primary metric scale (x_1, x_2) being truncation points) (cf. Lewontin, 1964b; LEWONTIN and HULL, 1967). Models 5b, 5c and 6b are defined using asymmetrical optimum ($\varphi = 8$). Table 1 gives w_i values for some models to illustrate their numerical relationships in terms of selection intensity and asymmetry. Mixed optimum-heterotic models (7a, b, c) are given by the functions (a) $w_i = f(y + z)$, (b) $w_i = f(2y + z)$, and (c) $w_i = f(2y + z)$ $= f(y z)^{1/2}$, which involve increasingly stronger heterozygote advantage in this order.
- (iii) Input specifications. Besides the mode of selection, computer runs were specified by combinations of several input values of these parameters: proportion of selfing (s), recombination value (c), population size (N), initial genotypic frequencies, number of generations (n) and number of replicate runs for any given set of parameters. In all cases, initial composition of population was given by Hardy-Weinberg proportions with allelic frequency q=.5 at each of the six loci, the run length was 100 generations and the number of replications were 5 to 20 depending on the population size (5 for N = 1000, 10 for N = 100 and 20 for N = 20).
- (iv) Some output parameters. The results were studied in terms of the output from both individual replicate run and various averages over the replicates. Given initially $q_0 = .5$, an index of gene dispersion (di) is defined by the mean absolute deviations in gene frequencies (q_i) , i.e., $\overline{|q_t q_0|}$,

Table 1. Fitness values under some examples of selection models

(a) Heterotic Models

		Number	or netero	zygous 10	ci (y)		
Model	0	1	2_	3	4	5	6
(1) Cumulative (2) Diminutive	0	.048	.143	.286 .584	.476 .741	.714 .880	.1.0

(b) Optimizing Selection Models Metric value (x)

Model	0	1	2	3	4	5	6	7	8	9	10	11	12
(6) Double truncation	0	O	0	0	0	1.0	1.0	1.0	0	0	0	0	0
(5a) Quadratic deviation $(\varphi = 6.0)$ (5b) Quadratic deviation	О	.306	.556	.750	.889	.972	1.0	.972	.889	.750	.556	.306	0
$(\varphi = 8.0)$	0	.234	.438	.609	.750	.859	.938	.984	1.0	.984	.938	.859	.750
(5c) Optimum-directional	0	.207	.395	.562	.707	.827	.919	.978	1.0	.974	.884		

mum and heterotic models, $E(\Delta q) = 0$; alternatively, the proportion of loci fixed for allele 0 or 1 (l) can be used (di) and l are highly correlated). The variance of changes in gene frequencies, $\sigma_{\Delta q}^2$ is often used in studying the stochastic processes and can be used here as based on a given time interval or at the steady state. The amount of genetic variability at the segregating loci is conveniently measured by the level of heterozygosity (H) averaged over loci and replicates, or even better by Wright's fixation index $\left(F = 1 - \frac{H}{2 q (1 - q)}\right)$. To adjust partly for the bias due to gene dispersion we may use F =1 $-\frac{n}{2(q+di)(p-di)}$. Phenotypic variability on the metric scale is output in terms of σ_w . To evaluate the amount of coupling versus repulsion linkage unbalance in the gametic arrays, Lewontin (1964a) and Fraser (1967) used the matrix of D'_{ij} values where for the pair of i^{th} and j^{th} loci, $D = g_{11} g_{00} - g_{10} g_{01}$, adjusted by dividing with the product of appropriate gene frequencies. The usefulness of D'_{ij} matrix has been discussed in some detail by Fraser (1967). Further, Fraser, et al. (1965) introduced a measure of gametic unbalance in terms of the repulsion index (r i) which is based on the number of alternations in the allelic type (0 or 1) along a chromosome (0 in case of gamete 000000, 0.2 in case of gametes 100000 or 011111, etc. and maximum value of 1 in case of gametes 101010 or 010101). Thus, D'_{ij} -matrix and ritogether provide satisfactory measures of the gametic array unbalance. D > 0 and ri < 0.5 represent excess of coupling phase and D < 0, r i > 0.5 indicate an excess of repulsion phase linkages. Also, note that dispersion index (di) increasing at a steady rate indicates steady decay toward fixation of all loci (e.g. optimum models) and in contrast a more or less unchanging di signifies steady state (or stability) as met with under models of heterotic selection. In most cases the results will be given for the 100th generation and averages over reps and loci for a set of input values.

at any given time t, since under symmetrical opti-

Results

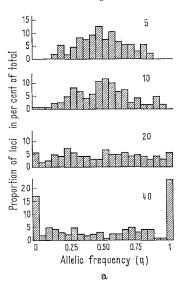
Random drift with none or weak selection. — Under increasing levels of inbreeding, the effective

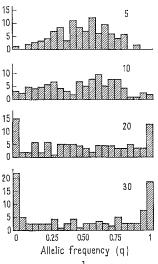
population size becomes smaller in proportion to the inbreeding coefficient (f) such that the rate of loss of heterozygosity, $\lambda=1-\frac{1+f}{4N}$ (Moran, 1962). Likewise, the variance of Δq , $\sigma_{\Delta q}^2$, becomes greater in linear relationship with $f\left(\sigma_{\Delta q}^2=\frac{p}{2N}(1+f)\right)$. Allard and Hansche (1964) studied random fixation at a single locus under mixed mating and drift due to finite size. To evaluate further the effect of linkage, runs were made with a six-locus model for s=0, .5, .75, .90, and for c=0.001, .01, .10 and .50, under (i) no selection, (ii) weak heterozygote advantage $(w_{11}=w_{00}=.9,w_{10}=1),$ and (iii) directional selection $(w_{11}=1,w_{10}=.95,w_{00}=.9)$ at each locus. Table 2 gives the values of di and $\sigma_{\Delta q}$ for some cases.

Table 2. Gene dispersion after 100 generations due to random drift (Case of no selection and c = .50; symbols as explained in the text)

	N =	20		N =	100	
S	di	Observed σ_{Aq}	Expected $\sigma_{\Delta q}$	di	Observed $\sigma_{\Delta q}$	Expected $\sigma_{\Delta q}$
0 .50 .90	.474 .493 .496	.0762 .0880 .0960	.0790 .0913 .1061	.185 .286 .336	.0319 .0363 .0428	.0352 .0408 .0476

Expected σ_{Aq} is based on $f = \frac{s}{2-s}$, the value of fat equilibrium. Gene frequency distribution is plotted against time in Fig. 1a—c for the case N=20and c = .5 and for three levels of selfing (s = 0, .5, .9). As expected, fixation proceeds more rapidly with inbreeding so that the state of steady decay (U-shaped distribution) is attained approximately in 2N/1 + fgenerations (where f is expected to be 0, .333 and .818 respectively for s = 0, .5, .9). The proportion of loci fixed (l) is given in Fig. 2a—c for s = 0, .5, .9and c = .01 and .50. The effect of linkage clearly depends on both the level of selfing or the form of selection and may change from an initially higher rate of dispersion to a lower value near the steady decay state (mean equilibrium fixation rate). Similarly, with N = 100, inbreeding and linkage seemed to have opposite effects in terms of the rate of fixation whereas in the case of N = 1000, the drift effects are too small





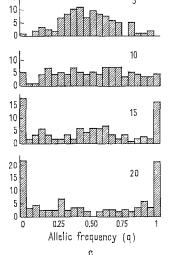


Fig. 1. Histograms showing the frequency distribution of allelic frequency (g) for no selection, N=20, c=.50, and (a) s=0, (b) s=.5, (c) s=.9, in the generations given by numerals

Table 3. Cumulative heterosis: Values of parameters after 100 generations (See text for explanation of symbols)

					<i>N</i> =	= 20						N =	100		
s	С	l	di	$\sigma_{arDelta q}$	H	F	r i	\overline{W}	σ_w	\overline{di}	H	F	r i	\overline{W}	$\sigma_{w_{_{j}}}$
	.001	.067	.060	.076	.441	.105	.418	.415	.439	.024	.509	020	.282	.503	.494
	.01 .10 .50	.016	.066 .082 .092	.078 .078 .077	.458 .483 .464	.069 .008 .038	.482 .496 .496	.403 .313 .286	.420 .224 .193	.020 .029 .037	.518 .498 .502	038 .001 010	.522 .540 .502	.474 .321 .325	.452 .212 .213
.50	.001	.016	.066	.077	.484	.015	.478 .416	.474 .452	.195 .485 .444	.022	.488	.022	.302	.483	.495
	.10	.016	.056	.077 .078 .081	.367	.239	.518	.432 .222 .212	.205	.037	.402	.024 .192 .218	.528	.253	.438
.90	.001	.117	.086	.072	.402 .469	.172	.650 .574	.360	.397	.022	.506	014 .026	.522	.503	.496 .471
	.10	.117	.162	.079	.274	.389	.500	.149	.166	.038	.318	.360 .482	.486	.192	.214

(Lewontin and Hull used N=1000 to approximate the deterministic case). In case of models with more intense selection to be described below, the effect of linkage in reducing gene dispersion is quite marked (see also Gill and Clemmer) whereas partial selfing did not consistently result in a proportionately higher rate of dispersion.

Heterotic models. — For cumulative heterosis (model 1), the population parameters based on the

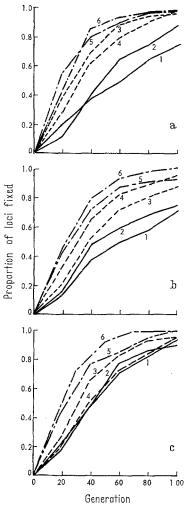


Fig. 2. Proportion of loci fixed (for allele 1 or 0) plotted against generation (n) for N=20 and (a) no selection, (b) heterozygote advantage, (c) directional selection. Numerals along the curves represent (1) s=0, c=.01, (2) s=0, c=.50, (3) s=.5, c=.01, (4) s=.5, c=.50, (5) s=.9, c=.01, and (6) s=.9, c=.50

replicate averages are given in Table 3. Steady state as judged by the values of dispersion index (di), fixation index (F), ri, σ_w , etc. was attained in all cases by the 100th generation. As expected from the strong heterozygote advantage involved in this model, there is little gene dispersion for N = 100 and in case of N=20, the values of σ_{Aq} were nearly same for various s and c sets. The main points to note are: (1) the level of heterozygosity is quite high even for s = .9 and is increasingly higher with tighter linkages, so that there seems to be effectively a greater departure from the expected F under high selfing than under random mating. In terms of the population structure and variation pattern under mixed mating system, this might pertain significantly to the case of predominantly selfing populations with heterozygote advantage, as evidenced by the experimental work on various species (for review see ALLARD, JAIN and WORKMAN, 1968). Further, note that the values of F are only slightly higher for N=20than in the case of N = 100. (2) The values of r_i averaged over replicates indicate only a small amount of gametic unbalance. However, the results from individual replicates, as illustrated by examples in Table 4, show that different reps can evolve toward extreme repulsion or coupling phase disequilibrium.

Table 4. Model 1 (cumulative heterosis): Examples of extreme replicate differences (N = 20)

s	С	Rep.	Gametes	Frequency	di	v i
0	.001	1	001110 110001	.5106 .4894	.011	.400
		2	101010 010101	.5000 .4762	.004	.995
		3	100000 011111	.5164 .4672	.107	.197
.50	.001	1	010100 101011	.4872 .5000	.006	.795
		2	101000 010101	.4911 .5089	.091	.804
		3	000000 111111	.5814 .4186	.081	0
.90	.001	1	100010 011101	.5686 .4314	.069	.600
		2	000100 011011	.5345 .4569	.114	.495
		3	010101 101011	.5000 .4919	.085	.898

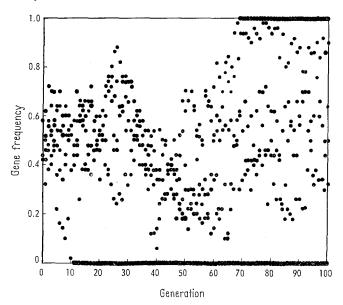


Fig. 3. Allelic frequency (q) at each of the six loci plotted against generation (n) for diminutive heterosis model, N=20, s=.9 and c=.50. Note a steady increase in gene dispersion toward fixation at all loci in a typical steady decay situation under large random drift effects

The variance of ri among reps is greater with tighter linkages. This follows from the findings of Lewontin (1964a) and Jain and Allard (1966) that under heterotic, symmetrical models there exist conjugate equilibria and different initial frequency sets would lead to them except in the case of metastable equilibrium ($q_i = .5$). Thus, with finite population size, different replicates may drift in to the alternative trajectories leading to the equilibria

with an excess of coupling or repulsion linkages. In studying the stochastic equivalents of deterministic models, in general, variation among replicate runs might be of far more interest than the mean over reps alone.

Under model 2 (diminutive heterosis), heterozygote advantage becomes smaller with each additional locus being heterozygous in an individual genotype and therefore in contrast to previous model, gene dispersion and homozygosis are expected to be higher than in the case of cumulative heterotic model. Table 5 summarizes the population parameters for model 2. The values of ri are close to .5 but again the between-replicate variation is quite large and ri can vary within a wide range between .1 to .9 in the 100th generation. With N=20, and s as high as .9, steady state was not attained under this model (Fig. 3). Note that under both models 1 and 2, the mean population fitness \overline{W} is lower with decreasing c, or greater amounts of selfing, and a similar pattern is noted for the variation in fitness (σ_{w}) .

Model 3 illustrates the results of heterotic model involving directional selection (Table 6) in which case except for very tight linkage and N=20, or in case of loose or no linkage for N=100, most loci are fixed for the favored allele $(\overline{p} \to 1)$ with $di \to 0.5$; particularly under inbreeding, any asymmetry due to directional selection markedly changes the stable equilibrium situations of heterotic selection. Note that under random mating, linkage can play an important role in terms of both the conditions for attaining steady state with polymorphisms, or as a factor delaying the process of eventual fixation at all loci.

Table 5. Model 2 (diminutive heterosis): Values of parameters in 100th generation (See text for explanation of symbols)

					N = 2	20				N	= 100	
s	С	1	di	Н	F	r i	\overline{W}	σ_w	\overline{di}	Н	F	r i
0	.001 .01 .10	0 0 .033 .050	.126 .104 .137 .132	.465 .465 .446 .447	.007 .028 .036 .039	.454 .562 .498 .484	.503 .524 .519 .542	.388 .320 .244 .209	.052 .057 .047 .064	.494 .494 .500 .487	.001 001 009	.438 .506 .482 .490
.50	.001 .01 .10 .50	.100 .050 .067 .067	.138 .128 .155 .171	.404 .380 .335 .317	.125 .187 .259 .282	.514 .514 .484 .516	.412 .427 .374 .372	.398 .350 .230 .233	.062 .062 .045 .070	.430 .392 .379 .361	.127 .204 .236 .264	.328 .480 .504 .498
.90	.001 .01 .10 .50	.100 .170 .380 .290	.081 .134 .268 .270	.444 .381 .181 .152	.088 .179 .492 .571	.528 .430 .468 .434	.414 .416 .196 .200	.469 .387 .171 .169	.029 .052 .058 .084	.483 .400 .248 .213	.031 .191 .497 .562	.482 .454 .498 .494

Table 6. Model 3 (heterotic-directional): Values of parameters in 100th generation

		N=20				$N \approx 100$)		
s	с	P	di	Н	v i	P	di	Н	r i
0	.001	.636	.173	.427	.368	.585	.111	.501	.480
	.01	.643*	.193	.375	.426	.616	.118	.472	.426
	.10	.879*	.390	.150	.298	.655	.456	.446	.448
.50	.001	.661	.261	.235	.412	.704	.211	.306	.306
	.01	.745*	.305	.175	.365	.775	.275	.267	.168
	.10	.978*	.494	.003	.033	.985*	.479	.027	.039
.90	.001	.745	.355	.124	.336	.646	.180	.300	.400
	.01	.820*	.414	.056	.223	.941*	.441	.049	.033
	.10	.983*	.500	0	.040	1.00*	.500	0	0

^{*} p and di increasing steadily. Other cases show steady state.

Intermediate optimum models. — The three symmetrical models (4a, 5a, 6a) used here represent an increasing order of selection intensity against the extreme phenotypes. Under double truncation model, for instance, all individuals with metric value outside the limits x_1 , x_2 (say, $x_1 = 5$ and $x_2 = 7$) are culled whereas models 4 and 5 (linear and quadratic deviation functions) allow with certain probability the reproduction by individuals with the phenotypic values in the wide range of 1 to 11 (cf. Table 1). Since models 4a and 5a gave qualitatively similar results differing only in the relative magnitude of quantities di, H, r i, etc., only the results of model 5a (so-called Wright's optimum model) are summarized in Table 7. Three time intervals (generations 20, 60 and 100) are given to illustrate briefly the rate of change toward fixation at all loci (di = .5).

In all cases the value of di steadily increases to its maximum value of .5 with H=0, and ri values are greater than .5 due to an excess of repulsion linkages. The effect of linkage for s=0 case is to lower the rate of gene dispersion whereas under inbreeding (particularly, note for s=.9), c=.001 and .5 gave lower values of di than c=.01 or .10. This trend could of course vary with the intensity of selection and specific values of N, etc. For instance, models 4a and 5a differ in the effect of s such that for s=0, di and ri are higher under model 5a, but lower for s=.9 than model 4a. Note that the variation in ri among replicates was significant and

relatively larger under inbreeding, which might in part account for the absence of any consistent relation between s or c and r i.

The values of D'_{ij} were used to compute the mean values of $D'_{(k)}$, where k = 1, 2, 3, 4, 5 is the unit of distance between loci i, j; $D'_{k=1}$ is based on the mean of $D'_{12}, D'_{23}, \ldots, D'_{56}$; $D'_{k=2}$ is based on the mean of D'_{13}, D'_{24} , etc., and so on. $D'_{(k)}$ indicate the different amounts of gametic unbalance as related to the relative position of (i, j)th pair of loci in a linkage block. Note from Table 7 that $D'_{(2)}$ tends to be increasingly larger with higher s which under looser linkages is associated with a change toward positive values of $D'_{(1)}$. The complexity of D'_{ij} values are given for some cases in Table 8 to illustrate this point. Further note that under these circumstances, terminal pair of loci did not seem to have consistently a larger gametic unbalance than the interstitial loci (viz. D'_{12} vs. D'_{34}). The values of D' plotted against time in Figs. 4 and 5 show changes over 100 generation period. These different characteristics of D'values for various k, as pointed out by Fraser (1967), deserve a detailed study of these interrelationships in order to measure effectively the role of optimizing selection in evolving polygenic balanced complexes.

The results for model 6a (double truncation, $x_1 = 5$, $x_2 = 7$) are summarized in Table 9 for N = 100 and 1000. There is a rapid increase in the frequency of gametes $(f_{(3)})$ in the potency class with 3 (1) and 3 (0) alleles (the additive score of a gamete

Table 7. Model 5a (WRIGHT's Optimum Model): Values of parameters in n^{th} (= 20, 60, 100) generation (N = 100, φ = 6)

		c = .00	01				c = .01				
s	n	\overline{di}	Н	r i	D'(1)	$D'_{(2)}$	\overline{di}	Н	r i	$D'_{(1)}$	D'(2)
0	20	.172	.412	.566	116	000	.148	.435	.652	054	059
	60	.255	.327	.510	214	009	.248	.321	.710	303	094
	100	.334	.216	.507	252	111	.269	.313	.734	353	058
.50	20	.142	.289	.584	133	267	.155	.275	.532	208	027
	60	.303	.191	.508	.031	443	.207	.238	.564	344	.042
	100	.291	.173	.532	047	401	.327	.167	.568	418	.068
.90	20	.158	.087	.673	289	.052	.164	.097	.676	132	266
	60	.280	.043	.692	427	.190	.244	.052	.668	377	140
	100	.326	.030	.721	539	.484	.417	.031	.670	247	318
		c = .10					c = .50)			
0	20	.156	.440	.658	050	030	.135	.447	.580	045	064
	60	.248	.343	.650	167	003	.259	.338	.582	025	147
	100	.340	.208	.686	108	040	.361	.206	.610	.068	136
.50	20	.166	.263	.602	156	118	.168	.278	.558	042	122
	60	.295	.184	.543	203	103	.280	.196	.518	076	182
	100	.422	.072	.546	135	217	.392	.090	.522	022	310
.90	20	.217	.084	.572	190	244	.163	.072	.636	199	122
	60	.362	.033	.600	.121	311	.296	.044	.602	.245	274
	100	.454	.008	.578	.104	387	.350	.034	.598	.297	250

Table 8. Values of D'_{ij} in 100th generation for model 5a, $\varphi = 6$, N = 100 (means over reps)

	k = 1					k=2				k = 3			k=4		k = 5
s, c	1,2	2,3	3,4	4,5	5,6	1,3	2,4	3,5	4,6	1,4	2,5	3,6	1,5	2,6	1,6
0, .001 0, .01 0, .50	354	303	230	371	506	.205	069	144	.240		.121 235 298	166	.218 .127 374	- 5	
.90, .001 .90, .01 .90, .50	191	.095	.498	667	968	960	632	347	.667	-1.00 650 443			.258 .463 444		

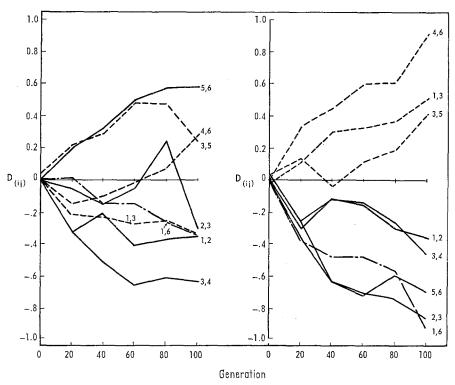


Fig. 4. Values of various D'_{ij} under model 5a for the case of N=100, s=0 and c=.001. Note that $D'_{1,2}$ and $D'_{1,6}$ are fairly close and that some of the D's changed sign

Fig. 5. Values of various D'_{ij} under model 5a for the case of N=100, s=.9 and c=.001. Note a stronger gametic unbalance than in the case of s=0 and that $D'_{(2)}$ are all positive, whereas $D'_{(1)}$ are negative

is defined as potency by Fraser, 1967) which indicates the strong effectiveness of this model of optimizing selection in changing the frequency distribution of potency classes toward an extreme kurtosis. In fact, a major part of these changes occurred during the first 20 or 30 generations. As $f_{(3)} \rightarrow 1$, and $di \rightarrow .5$, \overline{W} becomes very close to unity and in most cases all loci are fixed rapidly in a symmetric fashion for alleles 1 and 0 (Fig. 6). For random mating case, Lewontin (1964b) described similar results for a five-locus model in which by virtue of symmetrical pressures for fixation one locus is temporarily stalled near q=.5. However, it is significant to note that

linkage very markedly influenced the rate of gene dispersion (di), more so in the case of large population size (N =1000) and heavy inbreeding s = .90. This apparently results from the fact that in these cases the value of $f_{(3)}$ rises rapidly accompanied by a high rate of homozygosis so that drift becomes ineffective in giving gene dispersion due to sampling errors. Thus, these parameters are intricately balanced in terms of the relative rates of change in quantities di, H, ri, etc., and clearly gene dispersion need not always be proportionately greater under increasingly higher selfing. Further, it is interesting to note that parallel sets of results obtained with the truncation limits, $x_1 = 3$, $x_2 = 9$ (i.e., less intense selection) and N = 1000showed that while di tends to be smaller for s = .9 than s = 0 or 0.5, the difference is rather small and is not correlated with the linkage

intensities. Thus, it appears that selection intensity determines the relationship between the amount of gene dispersion and the parameters of inbreeding, linkage, or population size.

Several other points should be noted from these results: the level of heterozygosity is low and in general the values of fixation index (F) indicate little or no marginal overdominance, which is in line with our previous findings on the optimum models. Figs. 7 and 8 show that after an initial phase of rapid homozygosis under inbreeding, there is another period of rather sharp change in H around 40^{th} generation or so accompanied by an increase in $f_{(3)}$, which is in turn

Table 9. Model 6a (Double truncation): Values of parameters in 100th generation* (See text for explanation of symbols)

		N=10	00				N=10	000			
5	<i>c</i>	di	H .	f ₍₃₎	r i	σ_x	\overline{di}	Н	f ₍₃₎	r i	σ_x
)	.001	.261	.326	.979	.648	.183	.096	.469	.937	.600	.349
	.01	.303	.278	.867	.588	.508	.129	.455	.829	.700	.597
	.10	.372	.175	.686	.626	.795	.254	.323	.636	.728	.904
	.25	.490	.020	.951	.718	.265	.338	.229	.624	.664	.907
	.50	.487	.026	.929	.668	.329	.456	.072	.823	.580	.562
.50	.001	.307	.171	.994	.480	.104	.097	.322	.993	.620	.129
	.01	.296	.185	.973	.736	.224	.128	.299	.939	.650	.395
	.10	.456	.052	.949	.690	.311	.296	.184	.808	.788	.682
	.25	.477	.019	.951	.610	.149	.423	.073	.867	.754	.504
	.50	.499	.000	.999	.600	.020	.477	.028	.927	.550	.313
.90	.001	.280	.043	1.00	.640	0	.119	.085	.999	.600	.053
	.01	.343	.043	.991	.570	.124	.119	.081	.987	.652	.178
	.10	.354	.038	.973	.664	.227	.218	.068	.947	.622	.380
	.25	.467	.015	.976	.588	.167	.182	.072	.887	.568	.554
	.50	.412	.030	.957	.606	.264	.160	.080	.879	.608	.583

^{*} In all cases, di is increasingly steadily toward .5 (fixation at all loci) along with $f_{(3)} \rightarrow 1$ and a slow increase in ri.

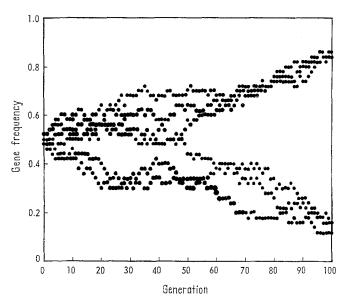
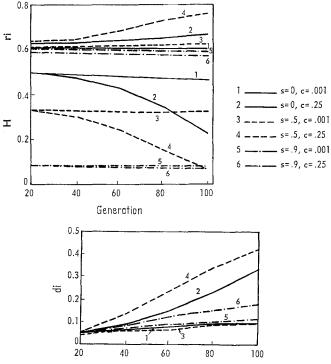


Fig. 6. Allelic frequency (q) at each of the six loci plotted against generation (n) for double truncation model (6a) and N=1000, s=.9, c=.5, showing a typical run with loci approaching fixation symmetrically for the two alleles

negatively correlated with the changes in the amount of phenotypic variability σ_x . Fraser, et al. (1965) postulated that di and σ_x might be highly correlated; however, note that σ_x in this case is more directly governed by $f_{(3)}$ values (e.g. highest σ_x and lowest $f_{(3)}$ for intermediate c and s=0 or .5). As pointed out by Lewontin (1964b), a low value of σ_x results from the presence of relationally balanced heterozygotes with high potential genetic variability in populations.

Models 5 b, 5 c and 6 b involve optimizing selection with asymmetrical optimum ($\varphi=8$) for which some of the population parameters in the 100th generation are given in Table 10. Under these models, with N=100 the variation among replicates was often large so that s and c did not seem to yield any consistent pattern in terms of di or ri. However, a point of some interest is that here again the values of ri and l are perhaps governed by complex interactions between the various parameters such that ri might be highest for various different combinations of s and c depending on the model. With completely additive metric scale, as shown by Lewontin (1964b)



Figs. 7, 8. Values of dispersion index (di), level of hetero zygosity (H) and repulsion index (ri) under model 6a, N=1000 and various combinations of s and c, showing the mode of change over time

Generation

and Jain and Allard (1965), there is no stable nontrivial equilibrium for any φ in its entire range. Partial dominance, linkage and inbreeding, on the other hand, allow a substantial region of stable equilibria for asymmetrical optima.

Mixed optimum-heterotic models are of considerable interest in view of some evidence for heterotic marker loci and optimizing selection for various quantitative traits within the same populations of barley (ALLARD and JAIN, 1962; JAIN and SUNESON, 1966) and lima bean (WÖHRMANN and ALLARD, in press). Moreover, such models might theoretically involve stable equilibria under a very wide range of optimum values (cf. JAIN and ALLARD, 1965). Some of the population parameters for models 7a—c are given in Tables 11 and 12. As expected, gene dispersion is increasingly smaller with the strong heterotic selection in model

Table 10. Optimum models with asymmetry ($\emptyset = 8.0$): Values of parameters in 100th generation (N = 100)

		Model 61 (double t		t = 7.0, 9.0	Model 5 (Wrigh	b т's optimun	$n, \varnothing = 8)$	Model 5c (optimum-directional)		
s	С	l	Н	r i	ı	H	r i	ī	Н	r i
0	.001	.233	.301	.564	.167	.314	.456	.267	.247	.556
	.01	.433	.111	.526	.233	.248	.450	.233	.239	.584
	.10	.600	.192	.540	.267	.266	.596	.133	.235	.514
	.50	.800	.015	.632	.367	.150	.442	.467	.009	.414
.50	.001	.467	.121	.554	.467	.125	.626	.267	.166	.524
	.01	.500	.115	.542	.167	.176	.510	.700	.059	.630
	.10	.800	.041	.612	.567	.065	.436	.467	.113	.672
	.50	1.00	0	.400	.500	.101	.546	.567	.080	.520
.90	.001	.533	.037	.425	-533	.026	.476	.733	.017	.568
•	.01	.467	.024	.503	.567	.028	.522	.667	.019	.670
	.10	.733	.024	.610	.667	.014	.640	.767	.023	.600
	.50	.733	.015	.496	.833	.012	.562	.433	.032	.470

Table 11. Models 7a, b: Values of parameters in 100th generation (N = 100)

		7a					7 b				
s	С	di	H	r i	D'(1)	D'(2)	\overline{di}	H	r i	D'(1)	$D'_{(2)}$
0	.01 .10	.107 .138	.470 .447	.504 .599	172 110	226 166	.091	.461 .482	.564 .622	125 090	116 001
	.50	.117	.467	.616	017	083	.097	.339	.604	062	028
.50	.01 .10	.161 .202	.329 .283	.654 .645	426 295	.145 102	.129 .142	.333 .304	.718 .610	240 115	046 030
	.50	.202	.244	.554	078	243	.144	.304	.580	055	088
.90	.01 .10 .50	.278 .338 .431	.076 .050 .018	.596 .554 .486	285 292 192	224 200 164	.146 .267 .342	.128 .067 .048	.648 .618 .586	490 189 097	101 177 088

Table 12. Optimum-heterotic models (7c): Values of parameters in 100th generation (N = 100)

s	С	N=20				N = 100			
		di	Н	F	r i	di	Н	F	ri
0	.001	.157	.412	.086	.526	.086	.482	.007	.586
	.01	.160	.416	.073	.498	.073	.492	~.005	.614
	.10	.181	.383	.118	.528	.066	.490	.003	.592
	.50	.185	.364	.157	.532	.065	.482	.020	.602
.50	.001	.118	.377	.202	.562	.059	.383	.221	.584
	.01	.136	.341	.264	.568	.053	.391	.207	.656
	.10	.192	.306	.282	.668	.075	.362	.260	.614
	.50	.222*	.259	-355	.602	.081	.351	.277	.616
.90	.001	.124	.418	.109	.564	.036	.493	.007	.618
	.01	.222*	.269	.330	.550	.042	.363	.267	.714
	.10	.325*	.134	.536	.592	.096	.211	.560	.630
	.50	.342*	.115	.568	.562	.086	.178	.633	.614

^{*} Steady state was not attained by n = 100.

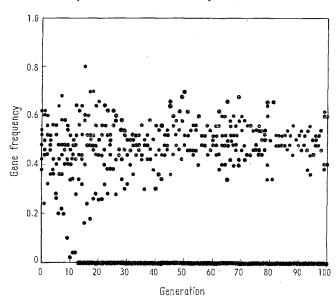


Fig. 9. Allelic frequency (q) at each of the six loci plotted against generation (n) under optimum-heterotic model $(7\,c)$ showing a typical run with the segregating loci at more or less steady state

7c (Fig. 9) and again, the values of F show that even with N=20, the excess of heterozygotes over the expected levels is proportionately greater under inbreeding, particularly in combination with tight linkages. The values of ri indicate a significant build-up of repulsion linkages and $D'_{(1)}$, $D'_{(2)}$ seem to follow more or less the same pattern as noted for model 5a. In general, these mixed models seem to

be quite favorable for linkage and gene interaction to help maintain variability in inbreeding populations.

Theoret. Appl. Genetics

General Implications of the Results

In this preliminary investigation of the dynamics of multigenic systems, all our models involved rather intense selection and only a few selected features of the total genetic system were considered. However, the main outcome of this study was to provide an overall picture of the relative magnitude of drift effects with population size in the range of 20 (small) to 1000 (large). With N = 1000 the results were in general agreement with those expected on the basis of equivalent deterministic runs. The pattern of gene dispersion, and the role of linkage and interaction in producing gametic unbalance varied only in some quantitative details rather than qualitatively. It appeared in general that while inbreeding and population size affect gene dispersion in linear relation with the coefficient of inbreeding (or fixation index F), linkage and selection could yield many interactive patterns. Linkage reduced the dispersive effects of random drift where selection models favor the evolution of gametic unbalance situations. Further work along the lines discussed, for instance, by GILL (1965), GILL and CLEMMER (1966) and KARLIN, et al. (1967) would seem very worthwhile. To extend this theory further to deal with the multiniche models, such as to examine the role of subdivision and interdeme selection, the measures of coancestry and group inbreeding are likely to be useful (Cockerham, 1967). Moreover, there are several interesting properties of the linked gene systems discovered recently for the case of random mating populations. For instance, Lewontin (1964b) showed that besides heterosis per se, homeostasis in terms of better buffering of heterozygotes against environment can lead to stable equilibria under optimum models. Lewontin and Hull (1967) found that a mixture of loosely linked and tightly linked blocks may result in synergistic effect on the evolution of higher population fitness plateaus. Fraser and Allan have shown in a paper in this volume that increasing the number of loci results in a significant decrease in the rate of loss of heterozygosity and the linkage effects on heterozygosis (see also Fraser and Burnell, 1968).

Both theoretical and experimental work so far provide evidence that linkage and gene interaction might be particularly important in the population dynamics and genetic structure of inbreeding populations. Data on the changes in the relative frequency of marked chromosome segments indicate an increasing amount of character associations in populations of lupins, oats, barley, lima bean and other species. The maintenance of genetic variability due to heterozygote advantage as observed at individual loci need to be re-examined in terms of multigenic systems where the rates of release of variation, pseudooverdominance and mixed models of selection should be considered in the estimation of variability and selection parameters. Wright (1967) has recently shown that the adaptive surfaces and therefore, dynamics of gene or genotypic frequency changes differ little due to linkage and epistasis causing gametic unbalance "if selective differences, including the interactive coefficients, are as small as is probably usually the case in nature, and the loci are in different chromosomes, as is usual in organisms with typical numbers, or if in the same chromosomes, are only loosely linked." Thus, more intensive experimental studies on these aspects of complex polymorphisms and selective forces alone can verify whether the plausible situations as shown by theory do exist in nature which are significantly governed by linkage, epistasis and drift.

Zusammenfassung

Die gemeinsamen Effekte der Koppelung, Inzucht und zufälligen genetischen Drift werden hinsichtlich der Populationsveränderung unter Selektion unter Einschluß von Geninteraktionen untersucht. Für die Monte-Carlo-Simulationen wurde ein 6-Locus-Modell mit einem einheitlichen Ausmaß der Rekombination zwischen benachbarten Paaren von Loci, gemischter Selbstung und Panmixie und dreier Grundtypen der Selektion (heterotisch, optimalisierend und gemischt optimalisierend-heterotisch) benutzt. Die Ergebnisse werden in erster Linie in Termini der Gendispersion, der genetischen Variabilität, der Gameten-Unbalance (Koppelungs-Ungleichgewicht) und der Näherung an stabile Genfrequenz-Gleichgewichte beschrieben. Sowohl unter kumulativen wie auch unter diminutiven Heterosis-Modellen kann ein stabiler Zustand des Polymorphismus erreicht werden, wobei die zufällige Gendispersion klein ist und verschiedene wiederholte Populationen einen hohen Grad gametischer Unbalance entwickeln, die sich in einem Überschuß entweder der Attraktionsoder der Repulsionsphase in Abhängigkeit von der zufälligen Drift in der Genfrequenz äußert. Auf der anderen Seite erleiden alle Populationen unter dem Optimum-Modell einen stetigen Verfall in Richtung auf Fixierung aller Loci, obwohl die Gendispersion durch ziemlich komplexe Interaktionen zwischen Parametern für Selbstung, Koppelung und Selektionsintensitäten beeinflußt wird. Die Gendispersion war bei höherem Inzuchtgrad nicht notwendigerweise höher. In allen Läufen mit Optimum-Modellen, in denen sich die mittlere Populations-Fitness dem Wert 1 nähert, wurde ein Überschuß von Typen mit Repulsionskoppelung beobachtet, vor allem bei enger Koppelung und starker Inzucht. Jede Asymmetrie in dem Sinne, daß Selektion das eine oder das andere Allel begünstigt, begünstigt zugleich die Genfixierung besonders bei Vorliegen von Inzucht. Auf der anderen Seite scheint ein Heterozygotenvorteil hinsichtlich der Erhaltung der Heterozygotie eine relativ größere Rolle bei Vorliegen von Inzucht zu spielen. Gemischte Optimum-heterotische-Modelle liefern einen Kompromiß zwischen den divergierenden Attributen multilokaler Systeme hinsichtlich der Erhaltung der Polymorphismen und der Maximalisierung der Fitness im Vergleich zu bestimmten optimal gekoppelten Genkomplexen. Im allgemeinen stimmen diese Ergebnisse, wie erwartet, bei mittlerem bis großem Populationsumfang mit denen früher für deterministische 2-Locus-Modelle berichteten über-

Literature

1. ALLARD, R. W., and P. E. HANSCHE: Some parameters of population variability and their implications in plant breeding. Adv. Agron. 16, 281-325 (1964). - 2. Allard, R. W., and S. K. Jain: Population studies in predominantly self-pollinated species. II. Analysis of quantitative genetic changes in a bulk-hybrid population of barley. Evolution 16, 90-101 (1962). - 3. Allard, R. W., S. K. Jain, and P. L. Workman: The genetics of inbreeding species. Adv. Genetics (in press). - 4. Bellmann, K., and H. Ahrens: Modellpopulationen in der Selektionstheorie und einige Ergebnisse aus Simulationsstudien. Der Züchter 36, 172-185 (1966). — 5. Bodmer, W. F., and J. Felsenstein: Linkage and selection: Theoretical analysis of the deterministic two locus random mating model. Genetics 57, 237-265 (1967). - 6. Bodmer, W. F., and P. A. Parsons: Linkage and recombination in evolution. Adv. Genetics 11, 1-99 (1962). — 7. COCKERHAM, C. C.: Group inbreeding and coancestry. Genetics 56, 89—104 (1967). — 8. Felsenstein, J.: The effect of linkage on directional selection. Genetics 52, 349-363 (1965). - 9. Fraser, A. S.: Gametic disequilibrium in multigenic systems under normalizing selection. Genetics 55, 507-512 (1967). — 10. Fraser, A. S., and D. G. Burnell: Computer Genetics (in press). New York: McGraw-Hill 1968. — 11. Fraser, A. S., D. MILLER, and D. BURNELL: Polygenic balance. Nature (London) 206, 114 (1965). — 12. GILL, J. L.: Effects of finite size on selection advance in simulated genetic populations. Austral. J. Biol. Sci. 18, 599-617 (1965). — 13. GILL, J. L., and B. A. CLEMMER: Effects of selection and linkage on degree of inbreeding. Austral. J. Biol. Sci. 19, 307-317 (1966). - 14. HAYMAN, B. I.: Mixed selfing and random mating when homozygotes are at a disadvantage. Heredity 7, 185—192 (1953).—15. Jain, S. K., and R. W. Allard: The nature and stability of equilibria under optimizing selection. Proc. Nat. Acad. Sci. U.S. 54, 1436—1443 (1965). — 16. Jain, S. K., and R. W. Allard: The effects of linkage, epistasis and inbreeding on population changes under selection. Genetics 53, 633-659 (1966). — 17. Jain, S. K., and C. A. Suneson: Increased recombination and selection in barley populations carrying a male sterility factor.

I. Quantitative variability. Genetics 54, 1215–1224 (1966). — 18. Karlin, S., J. McGregor, and W. F. Bodmer: The rate of production of recombinants between linked genes in finite populations. Proc. 5th Berkeley Symp. Probab. Stat. (in press). — 19. Kimura, M.: A model of a genetic system which leads to closer linkage by natural selection. Evolution 10, 278—287 (1956). — 20. Kimura, M.: Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics 52, 875—890 (1965). — 21. Latter, B. D. H.: The interaction between effective population size and linkage intensity under artificial selection. Genet. Res. 7, 313—323 (1966). — 22. Lewontin, R. C.: The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49, 49—67 (1964a). — 23. Lewontin, R. C.: The interaction of selection and linkage. II. Optimum models. Genetics 50, 757—782 (1964b). — 24. Lewontin, R. C., and P. Hull: The interaction of selection and linkage. III. Synergistic effect of blocks of genes. Der Züchter 37, 93—98 (1967). — 25. Lewontin, R. C., and K. Kojima: The evolutionary dynamics

of complex polymorphisms. Evolution 14, 458–472 (1960). — 26. Moran, P. A. P.: Statistical processes of evolutionary theory. Oxford: Oxford Univ. Press 1962. — 27. Narain, P.: Effect of linkage on homozygosity of a population under mixed selfing and random mating. Genetics 54, 303–314 (1966). — 28. Nei, M.: Effects of linkage and epistasis on the equilibrium frequencies of lethal genes. I. Linkage equilibrium. Japan. J. Genet. 39, 1–6 (1964). — 29. Wöhrmann, K., and R. W. Allard: Directional and stabilizing selection in a population of lima beans. Crop Sci. (in press). — 30. Wright, S.: The genetics of quantitative variability. In: Quantitative Inheritance, pp. 5–41, ed. E.C.R. Reeve et al., Agric. Res. Council, London (1952). — 31. Wright, S.: Factor interaction and linkage in evolution. Proc. Roy. Soc. London B 162, 80–104 (1965). — 32. Wright, S.: "Surfaces" of selective value. Proc. Nat. Acad. Sci. U.S. 58, 165–172 (1967). — 33. Young, S. S. Y.: Computer simulation of directional selection in large populations. I. The programme, the additive and the dominance models. Genetics 53, 189–205 (1966).