ESTIMATION OF LIFE CYCLE COMPONENTS OF SELECTION IN AN EXPERIMENTAL PLANT POPULATION[1]

M. T. CLEGG[2], A. L. KAHLER AND R. W. ALLARD

Department of Genetics, University of California, Davis, California 95616

Manuscript received August 23, 1976
Revised copy received February 22, 1978

ABSTRACT

Viability and fertility components of selection associated with linkage blocks marked by four electrophoretically detectable loci were estimated in an experimental population of barley [Composite Cross V (CCV)]. The intensity of selection affecting the distribution of pollen types in the outcross pool was also estimated and comparisons were made between the selective values of genes in the pools of uniting ovules and pollen. The estimates show that selection was intense at various stages of the life cycle and that viability and fertility components often opposed one another. Estimates of viability and fertility components of selection were also extended to the three-locus level. The multilocus estimates reveal large differences in viability and fertility among homozygous genotypes. It is concluded that strong selection operates at all life cycle stages in CCV, although often in differing directions.

DIFFERENTIAL selection has been demonstrated repeatedly in studies of natural and experimental plant populations. For example, in Composite Cross V, an experimental population of barley (*Hordeum vulgare* L.), large changes in genotype frequencies have been detected between single generations and over longer term intervals, at loci governing morphological polymorphisms (JAIN and ALLARD 1960), measurement characters (ALLARD and JAIN 1962) and enzyme polymorphisms (ALLARD, KAHLER and WEIR 1972; WEIR, ALLARD and KAHLER 1972, 1974). A major issue, yet to be resolved, is the mode of selection and the life cycle stages at which selection is manifested.

There have been three general approaches to partitioning components of selection in animal species. The first method (using Drosophila as the experimental organism) has involved the estimation of components of selection through a series of essentially separate experiments (*e.g.*, PROUT 1971a; BUNDGAARD and CHRISTIANSEN 1972) and checking the predictive value of the estimates in population experiments (PROUT 1971b). The second approach has been to define a two-component model of selection (viability and fertility components) and to obtain maximum likelihood estimates of the parameters over several generations directly from the population of interest (ANDERSON 1969). The third method, pioneered

[1] Supported in part by National Science Foundation BMS 73-01113-A01.
[2] Present address: Department of Botany, University of Georgia, Athens, Georgia 30602.

by CHRISTIANSEN and FRYDENBERG (1973, 1976), has utilized information contained in the family structure of populations by taking mother-offspring samples together with samples from the male population. Examination of the structure of the data permits an hierarchical system of hypothesis testing for selection components and for nonrandom mating. This method has been applied to eelpout (*Zoarces viviparus*) populations (*e.g.*, CHRISTIANSEN and FRYDENBERG 1976).

In plant populations, CLEGG and ALLARD (1973) have estimated viability and fertility components of selection in a natural population of the slender wild oat, *Avena barbata*. In this case, the approach involved obtaining multiple census data within each generation from the population. The results indicated significant changes in genotypic frequencies due to viability selection. The present investigation extends this method in two ways; first, factors which affect pollen pool frequencies are considered separately from viability and fertility estimates; and second, the method of estimating components of selection is extended to the three-locus joint distribution in order to investigate the multilocus effects of selection components.

MATERIALS AND METHODS

The population studied, Composite Cross V (CCV), was developed from intercrosses among thirty varieties of barley representing the major barley growing regions of the world. Details of the method by which the population was propagated have been reported elsewhere (KAHLER and ALLARD 1970; ALLARD, KAHLER and WEIR 1972). For present purposes, it is sufficient to note that the first generation of the newly synthesized population was grown in 1940 at Davis, California, and that it has subsequently been grown each year at Davis following standard agricultural practices, with no conscious selection. In addition to growing a random sample of the seed of every generation in the following year, a random sample of seed from each generation was stored for five- to ten-year intervals and then advanced one generation to provide rejuvenated seed, which was similarly saved. At present, seed stocks are available for generations 5 to 10, 15 to 21, and 25 to 30.

The present experiments involved growing two plots, each containing approximately 10,000 adult plants, of generations 8, 19 and 28 at Davis in 1972. Plots were allocated at random in a uniform field nursery. One plot of each generation, chosen at random, was assayed to determine adult genotypic frequencies. The census involved taking a single spike from each of 1,100 randomly chosen adult plants in each of the three generations, germinating nine seeds per spike in the laboratory, and assaying the resulting seven-day-old seedlings (28,950) in total). Electrophoretic assays were for esterase loci *EA*, *EB*, *EC*, and *ED* (*EA*, *EB* and *EC* are very tightly linked and *ED* is located in an independent linkage group), following the procedures of KAHLER and ALLARD (1970). The genotypic array of the progeny was used to estimate adult genotypic frequencies and mating system parameters following the methods of APPENDIX A.

The second plot of each generation was harvested *en masse* at maturity to provide bulk seed of generations 9, 20 and 29. The first plot could not be used for this seed census because sampling of adult plants causes substantial damage to the population. Thus, the zygotic census data were obtained from the original seed lots of generations 8, 19 and 28 harvested in 1971 and the seed lots of generations 9, 20 and 29 from the second replicate harvested *en masse* in 1972. Eleven-hundred random seeds from each of generations 8, 9, 19, 20, 28, and 29 were germinated in the laboratory and the resulting seven-day-old seedlings (6,570 in total) were assayed electrophoretically for the four esterase loci. It should be noted that more than 99% of the seeds germinated and produced assayable seedlings in a germination chamber at 20°; hence, the assays are representative of the zygotic stage in the life cycle (see Figure 1 for a schematic representation of the relevant life cycle stages).

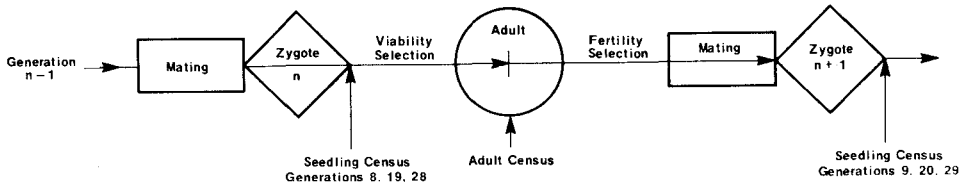

FIGURE 1.—A schematic diagram of the life cycle in *Hordeum vulgare* in which are indicated the points at which census data were obtained.

The data base for the analysis consists of the numbers of each genotype observed at each zygotic census and also the numbers of genotypes observed within the progeny of each spike (*i.e.*, the family produced by a single maternal parent). Maximum likelihood estimates of maternal genotypic frequencies in the adult population were made following the methods presented in APPENDIX A.

RESULTS

Single-locus analysis

There were 4, 3, 3 and 4 alleles at loci *EA*, *EB*, *EC* and *ED*. The fourth allele at locus *EA* (*EA 2.6*) was rare (< 3%) and it was therefore combined with allele *EA 0.2* into a single class [see KAHLER and ALLARD (1970) for allelic designations]. A recessive null allele at the *ED* locus was present in moderate frequency (heterozygous genotypes involving this allele cannot be distinguished from homozygous banded genotypes). All other banding expressions at the four esterase loci are codominant, and genotypic frequencies can be obtained directly from the electrophoretic results. Mating system parameters [probability of an outcross is denoted by $t = (1 - s)$, where s is the probability of selfing] and gene frequencies in the pool of pollen involved in outcross events (denoted p_{im} for the i^{th} allele at a locus) are given in Table 1. The mating system analysis was not applied to the *ED* locus because of complications presented by the null allele; hence, we will restrict discussion of single-locus frequencies to the *EA*, *EB* and *EC* loci.

In two instances (loci *EB* and *EC* generation 8) it was necessary to consolidate the data into diallelic classes, one class representing the most frequent allele and the second representing a composite complementary class. This further consolidation was required because the estimation program did not converge, due to very low frequencies of the third allele at these loci. Hence, there is one degree of freedom of the χ^2 goodness-of-fit test under the diallelic consolidation and 15 degrees of freedom for the triallelic data (APPENDIX A). All nine χ^2 values indicate an acceptable fit of the mating system model to the observed data.

Table 2 reports triallelic, single-locus genotypic frequencies for the zygote-adult-zygote transitions at each locus. A noteworthy feature of these data is that genotypic frequency changes are often large, even between different life cycle stages within a generation. For example, in generation 19, the A_1A_1 genotype increased from a frequency of 0.547 at the zygote stage to a frequency of 0.697 at the adult stage, and there was a concomitant change of frequency from 0.362 to 0.224 at zygote and adult stages for genotype A_2A_2. It is appropriate to test whether the vectors of relative genotypic frequencies have changed from the

TABLE 1

Outcrossing rate (t), *gene frequencies in the pollen pool* (p_i; i=*1, 2 for triallelic loci*) *and chi square goodness-of-fit statistics*

Locus	t	p_{1m}	p_{2m}	χ^2
		Generation 8		
	0.00517	0.4552	0.4027	8.10
EA	(0.00097)	(0.0944)	(0.0915)	
	0.00491	0.8816	——	1.66
EB	(0.00071)	(0.0471)		
	0.00537	0.6722	——	0.80
EC	(0.00075)	(0.0658)		
		Generation 19		
	0.01215	0.6307	0.2870	23.60
EA	(0.00181)	(0.0637)	(0.0564)	
	0.00557	0.7986	0.1210	6.53
EB	(0.00183)	(0.0825)	(0.0586)	
	0.01127	0.4870	0.0918	16.43
EC	(0.00146)	(0.0652)	(0.0301)	
		Generation 28		
	0.01038	0.4429	0.4865	8.63
EA	(0.00155)	(0.0774)	(0.0740)	
	0.00998	0.6350	0.2832	22.36
EB	(0.00168)	(0.0717)	(0.0635)	
	0.01052	0.3906	0.1014	14.90
EC	(0.00134)	(0.0646)	(0.0360)	

$\chi^2_{[1]} = 3.84$ at $P = 0.05$, $\chi^2_{[15]}$ 25.00 at $P = 0.05$.

Standard errors are given in parentheses.

zygotic stage (denoted $\mathbf{F}^n$ in generation n) to the adult stage (denoted $\mathbf{A}^n$). A likelihood ratio test of the hypothesis, H_0:$\mathbf{F}^n = \mathbf{A}^n$, against the alternative, H_A:$\mathbf{F}^n \neq \mathbf{A}^n$, was performed and the resulting statistic is reported in Table 2. The test statistic was computed assuming multinomial sampling. This assumption is not met by the sampling scheme at the adult stage; however, due to the very low outcrossing rates and the moderate sized progenies, the sampling approximates the multinomial scheme. Reference to Table 2 shows that H_0 is rejected at the 5% level for all loci and generations and that all but two of the tests are significant beyond the 1% level. The data therefore provide evidence of a strong viability component of selection.

The genotypic frequencies can also be represented in terms of gene frequencies and fixation indices as,

$$A_iA_i : p_i^2 + p_i \sum_{j \neq i}^{k} p_j F_{ij} \qquad i \neq j;\ i, j = 1, 2, \ldots, k\ .$$

$$A_iA_j : 2p_ip_j(1-F_{ij})$$

Changes in the genotypic frequency distribution might arise either from changes

TABLE 2

Single-locus relative genotypic frequencies with χ^2 statistics

	Generation								
	8–9			19–20			28–29		
Genotype	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
EA locus									
A1A1	0.496	0.470	0.510	0.547	0.697	0.567	0.582	0.650	0.600
A1A2	0.002	0.007	0.012	0.001	0.011	0.013	0.016	0.012	0.008
A1A3	0.007	0.002	0.008	0.001	0.0	0.003	0.0	0.002	0.002
A2A2	0.350	0.388	0.335	0.362	0.224	0.324	0.354	0.306	0.332
A2A3	0.002	0.0	0.001	0.002	0.0	0.006	0.001	0.0	0.003
A3A3	0.143	0.134	0.134	0.086	0.067	0.088	0.047	0.030	0.056
$\chi^2_{[5]}$*	12.01			69.14			16.46		
N†	1243	1077	1001	1007	1097	1078	1102	1049	1139
EB locus									
B1B1	0.854	0.899	0.887	0.870	0.851	0.881	0.749	0.697	0.711
B1B2	0.002	0.0	0.001	0.002	0.004	0.001	0.003	0.004	0.001
B1B3	0.006	0.004	0.002	0.0	0.002	0.002	0.004	0.006	0.003
B2B2	0.044	0.048	0.047	0.080	0.074	0.061	0.175	0.231	0.202
B2B3	0.002	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
B3B3	0.093	0.049	0.063	0.048	0.069	0.055	0.070	0.062	0.083
$\chi^2_{[5]}$	23.38			24.51			27.55		
N	1243	1097	1001	1007	1094	1078	1102	1058	1139
EC locus									
C1C1	0.615	0.646	0.636	0.661	0.533	0.650	0.519	0.476	0.533
C1C2	0.0	0.001	0.003	0.0	0.001	0.001	0.009	0.003	0.001
C1C3	0.010	0.007	0.007	0.002	0.005	0.010	0.005	0.004	0.010
C2C2	0.056	0.044	0.058	0.058	0.096	0.058	0.205	0.256	0.215
C2C3	0.0	0.0	0.001	0.002	0.0	0.001	0.002	0.0	0.0
C3C3	0.318	0.302	0.295	0.277	0.364	0.279	0.259	0.260	0.240
$\chi^2_{[5]}$	21.31			43.76			13.50		
N	1241	1084	1001	1007	1096	1078	1102	1062	1139

* $\chi^2_{[5]} = 11.07$ at $P = 0.05$, $\chi^2_{[5]} = 15.09$ at $P = 0.01$.

† N is the total number of individuals assayed per locus per generation.

in gene frequencies or from changes in the distribution of genes in zygotes. The genotypic frequency distribution can be resolved into these two components through the estimation of fixation indices and gene frequencies. In the present instance, heterozygous frequencies are sometimes zero, necessitating some reduction in the parameterization. Consequently, maximum likelihood estimates of gene frequencies and a single fixation index, $F = F_{12} = F_{13} = F_{23}$, were undertaken. These estimates, given in Table 3, show that most of the change in genotypic frequencies from zygote to adult to zygote stages can be accounted for in terms of changes in gene frequencies. The changes in F from one stage to the next are usually nonsignificant. However, several features of the data require comment. The first point to note is that for loci *EA* and *EC*, F for the adult stage is

TABLE 3

Gene frequencies and average F *statistics with standard errors in parentheses*

	Generation 8–9			Generation 19–20			Generation 28–29		
Statistic	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
EA locus									
P_1	0.501	0.475	0.522	0.547	0.703	0.573	0.590	0.658	0.604
	(0.014)	(0.015)	(0.016)	(0.016)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)
P_2	0.350	0.391	0.340	0.364	0.231	0.334	0.363	0.312	0.337
	(0.013)	(0.015)	(0.015)	(0.015)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)
F	0.981	0.986	0.965	0.993	0.976	0.961	0.967	0.972	0.976
	(0.005)	(0.005)	(0.008)	(0.004)	(0.007)	(0.008)	(0.008)	(0.008)	(0.006)
EB locus									
P_1	0.857	0.901	0.889	0.871	0.854	0.883	0.752	0.703	0.713
	(0.010)	(0.009)	(0.010)	(0.011)	(0.011)	(0.010)	(0.013)	(0.014)	(0.013)
P_2	0.046	0.047	0.047	0.082	0.076	0.061	0.176	0.231	0.202
	(0.006)	(0.006)	(0.007)	(0.009)	(0.008)	(0.007)	(0.011)	(0.013)	(0.012)
F	0.965	0.980	0.985	0.991	0.979	0.987	0.984	0.979	0.992
	(0.010)	(0.010)	(0.008)	(0.006)	(0.009)	(0.007)	(0.006)	(0.007)	(0.004)
EC locus									
P_1	0.620	0.650	0.641	0.662	0.537	0.655	0.527	0.481	0.540
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)
P_2	0.055	0.044	0.061	0.059	0.096	0.059	0.211	0.257	0.213
	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)	(0.007)	(0.012)	(0.013)	(0.012)
F	0.979	0.983	0.978	0.992	0.989	0.975	0.973	0.990	0.981
	(0.006)	(0.006)	(0.007)	(0.004)	(0.004)	(0.007)	(0.006)	(0.004)	(0.005)

usually greater than F at the subsequent zygotic stage. The expected behavior of F, for neutral loci in infinite populations practicing a mixture of selfing and random outcrossing is $F_n = \frac{s}{2}(1 + F_{n-1})$, where n is time measured in generations. Comparison of the expected values of F_n (where the adult values were used for F_{n-1} and estimates of s for the appropriate locus and generation are taken from Table 1) to the actual estimates of F_n reveals that five of the six *EA* and *EC* comparisons have estimated values of F_n that are less than the value expected based upon the estimated selfing rate. (In the sixth case, the two estimates are equal.) In one instance (generation 8–9, locus *EA*) the contrast is significant.

This result suggests an excess of heterozygotes following the mating cycle. The excess of heterozygotes is usually small and it can, at least in part, be accounted for by the fact that gene frequencies differ among the pool of pollen and ovules involved in outcross events, as will be shown below. Surprisingly, the pattern at the *EB* locus is different. Two of the three estimated values are greater than the expected value, while in the remaining case the two are equal. The three loci are very tightly linked, and in linkage disequilibrium, so that a different pattern of change among loci is unexpected.

The second noteworthy feature of the F statistics is that no strong evidence exists for heterozygous advantage in viability. Of the nine comparisons between

estimates deriving from the zygotic stage generation n and the adult stage generation n, five show an increase in F and four show a decrease from zygote to adult stages. One contrast indicates a significant decrease in F (generation 19, EA) and one contrast reveals a significant increase (generation 28, EC).

The third and most striking feature of the data in Table 2 is the very large and usually opposing gene frequency changes observed among the zygote to adult to zygote transitions. In addition to the within generation changes, there are also substantial changes in gene frequency over the 21 generation period studied. For example, the frequency of B_2 changes from 0.046 to 0.202, and the frequency of C_2 exhibits a similar change from 0.055 to 0.213. Gene frequency changes at the EA locus are less pronounced. Nevertheless, the results indicate that the genetic composition of the population has changed substantially during the period spanned by these data.

The model of selection presented in APPPENDIX B permits the estimation of ovule frequencies in the pool of uniting gametes (denoted p_{if}), following fertility selection. Comparison of the ovule pool frequencies, given in Table 4, with the corresponding pollen pool frequencies (Table 1) reveals some instances where pollen and ovule gene frequencies differ significantly (*e.g.*, C_1 generation 19, C_1 generation 28 and C_2 generation 28). Evidently selection has affected the gene frequencies among male and female gametes differently.

TABLE 4

Ovule pool frequencies together with ovule and pollen selective values

Generation	Locus	Gene frequency		Selective values			
		P_{1f}	P_{2f}	$1m$	$1f$	$2m$	$2f$
8	EA	0.52	0.34	0.96	1.10	1.02	0.87
		(0.02)	(0.01)	(0.20)	(0.05)	(0.24)	(0.05)
	EB	0.89	0.05	0.98	0.99	1.21	0.98
		(0.01)	(0.01)	(0.05)	(0.01)	(1.03)	(0.19)
	EC	0.64	0.06	1.03	0.99	0.94	1.32
		(0.01)	(0.01)	(0.10)	(0.03)	(1.79)	(0.25)
19	EA	0.57	0.33	0.90	0.82	1.25	1.45
		(0.02)	(0.01)	(0.09)	(0.03)	(0.25)	(0.10)
	EB	0.88	0.06	0.94	1.03	1.59	0.81
		(0.01)	(0.01)	(0.10)	(0.02)	(0.79)	(0.13)
	EC	0.66	0.06	0.91	1.22	0.95	0.61
		(0.01)	(0.01)	(0.12)	(0.04)	(0.38)	(0.09)
28	EA	0.60	0.34	0.67	0.92	1.56	1.08
		(0.01)	(0.01)	(0.12)	(0.03)	(0.25)	(0.07)
	EB	0.71	0.20	0.90	1.02	1.22	0.87
		(0.01)	(0.01)	(0.10)	(0.03)	(0.28)	(0.07)
	EC	0.54	0.22	0.81	1.13	0.39	0.84
		(0.01)	(0.01)	(0.14)	(0.05)	(0.14)	(0.06)

Standard errors are in parentheses.

We can parameterize the effects of selection on the pool of uniting gametes by comparing gene frequencies in the adult population contributing gametes (Table 3) with the frequencies in the pool of uniting gametes. Thus, we estimate the weight ("selective value") that transforms the gene frequency among adults into the frequency among uniting gametes as $\hat{l}_{if} = p_{if}/p_i$ and $\hat{l}_{im} = p_{im}/p_i$. (It should be noted that the weights $\hat{l}_{if}$ and $\hat{l}_{im}$ are frequency dependent because they are estimated relative to $\sum_i p_i l_{if} = \sum_i p_i l_{im} = 1$.) Values of these maximum likelihood estimators are also given in Table 4. An estimate of unity implies no change from adult frequencies, while values greater than unity imply an increase and values less than one imply a decrease in gene frequency, as compared to the adult population. The estimates for pollen and ovule pool frequencies have a slight negative correlation because they derive from the same sample. Nevertheless, there is a tendency for pollen and ovule pool frequencies to differ, and this is reflected in their respective selective values. This tendency is most pronounced for the C_1 allele, where the weighted average selective values are 0.941 ± 0.067 and 1.084 ± 0.022 for the pollen and ovule pools, respectively.

The factors that could account for differences between pollen and ovule frequencies are: (1) differential production of male and female gametophytes according to genotype; this alternative seems highly unlikely because the main determinants of gamete production are number of spikes and number of florets per spike, and each of these affects ovule and pollen production in the same direction; (2) differential contribution of pollen to the outcross pool independent of actual pollen production due to unequal release of pollen by different plants; there is no suggestion that this occurs in Composite Cross V; (3) differential gametophytic selection; this is the most plausible alternative, particularly because gametophytic selection may occur among the pollen involved in outcross events.

To quantify the impact of selection at each of the life cycle stages, we have estimated viability and fertility components of selection (Table 5). The models of selection adopted, together with the maximum likelihood estimators for the selection parameters, are presented in APPENDIX B. Viability selection is assumed to occur through differential survival, according to genotype, from zygote to adult stages. Fertility selection estimates apply to the average number of zygotes produced by each maternal genotype. (See Figure 1 for a schematic representation of stages in which viability and fertility selection was measured.) In computing fertility selection estimates, the mating cycle must be properly accounted for by using the appropriate mating system estimates from Table 1. Both viability and fertility selection estimates are defined relative to the most frequent genotype which is assigned fitness of unity. In interpreting viability and fertility estimates for a particular genotype within a generation, it should be borne in mind that these estimates are negatively correlated (in a statistical sense) relative to the overall fitness of the genotype in question, and relative to the genotype to which they are normalized. This negative correlation arises because the two estimates share the adult census data (ANDERSON 1969).

TABLE 5

Single-locus estimates of selective values with standard errors in parentheses

Genotype	Generation 8–9 $\hat{v}$	8–9 $\hat{l}$	19–20 $\hat{v}$	19–20 $\hat{l}$	28–29 $\hat{v}$	28–29 $\hat{l}$	χ^2 $\hat{v}$	χ^2 $\hat{l}$
EA locus								
A1A1	1.00	1.00	1.00	1.00	1.00	1.00		
A1A2	2.99	2.99	8.64	2.37	0.64	1.03		
	(2.05)	(1.23)	(9.00)	(0.80)	(0.24)	(0.38)		
A1A3	0.27	7.63	—	—	—	1.40		
	(0.21)	(5.77)				(1.11)		
A2A2	1.17	0.79	0.49	1.73	0.77	1.17	53.4	28.2
	(0.09)	(0.08)	(0.04)	(0.17)	(0.06)	(0.11)		
A3A3	0.98	0.91	0.61	1.53	0.56	1.99	7.4	9.1
	(0.12)	(0.12)	(0.10)	(0.26)	(0.13)	(0.45)		
EB locus								
B1B1	1.00	1.00	1.00	1.00	1.00	1.00		
B1B2	—	—	1.86	0.26	1.54	0.00		
			(1.62)	(0.28)	(1.17)	(0.29)		
B1B3	0.62	1.02	—	1.60	1.67	0.70		
	(0.38)	(0.65)		(1.33)	(1.08)	(0.39)		
B2B2	1.04	0.97	0.94	0.79	1.42	0.85	5.9	0.5
	(0.20)	(0.20)	(0.15)	(0.14)	(0.14)	(0.09)		
B3B3	0.50	1.26	1.49	0.74	0.95	1.29	16.0	6.3
	(0.08)	(0.24)	(0.28)	(0.14)	(0.16)	(0.22)		
EC locus								
C1C1	1.00	1.00	1.00	1.00	1.00	1.00		
C1C2	—	5.11	—	0.75	0.39	0.08		
		(5.64)		(1.12)	(0.24)	(0.26)		
C1C3	0.63	1.58	3.43	2.33	0.79	3.87		
	(0.29)	(0.81)	(2.80)	(1.07)	(0.50)	(1.97)		
C2C2	0.75	1.31	2.06	0.49	1.36	0.75	17.1	11.5
	(0.14)	(0.27)	(0.35)	(0.08)	(0.14)	(0.08)		
C3C3	0.90	0.98	1.63	0.62	1.09	0.82	21.9	10.7
	(0.07)	(0.10)	(0.14)	(0.06)	(0.10)	(0.08)		

* $\chi^2_{[2]}$ 5.99 at $P = 0.05$; $\chi^2_{[2]} = 9.21$ at $P = 0.01$.

Heterogeneity chi-squares for viability and fertility estimates over generations for 22 and 33 genotypes are also given.

Selection estimates for heterozygous genotypes are not particularly informative because of their high standard errors. Among homozygous genotypes the components of selection are significantly different from unity in many cases; thus viabilities of B_3B_3 in generation 8, A_2A_2, C_2C_2 and C_3C_3 in generation 19, and A_2A_2, A_3A_3, B_2B_2 and C_2C_2 in generation 28 all differ significantly from unity. The impact of selection is greatest in the intermediate generation, as judged by the likelihood ratio test (Table 2). Estimates of fertility and viability selection among the 22 and 33 homozygous genotypes are usually heterogeneous over generations (Table 5); the only exception is the B_2B_2 estimate. Evidently the

pattern of viability and fertility selection is not constant over time. Differences in the pattern of selection are especially noteworthy because the three generations were replicated in the same environment and year. Changes in the architecture must reflect changes in the distribution of selective effects over the evolving genetic organization of this population. In addition, fertility differences, which are also intense among the homozygous genotypes, tend to cancel the viability effects. Consequently, overall selective values among homozygous genotypes (given approximately by the product of viability and fertility estimates) are small, but usually in accord with long-term directions of genotypic frequency change. A further point to note is the absence of any clear-cut evidence of frequency dependent selection at either of the two life cycle stages. However, failure to detect frequency dependent selection is not surprising because a very strong dependence of fitness on gene frequency would be necessary to produce detectable effects. This is a consequence of the restricted range of gene frequencies over which the estimates were obtained and also of the limited number of samples.

Multilocus analysis

We turn now to consideration of the multilocus effects of the intense selection thus far demonstrated. Tables 6 to 9 give the three-locus genotypic frequency distribution for each census point where heterozygous frequencies have been consolidated into single classes. These data are based on diallelic frequency classes. The diallelic consolidation was achieved by taking the most frequent allele as one class and combining all other alleles into a single synthetic class, following the procedures of WEIR, ALLARD and KAHLER (1972). The reduction of the data is essential because even three triallelic loci would produce 378 different genotypes; however, a consequence of combining genotypic classes with

TABLE 6

Three-locus joint genotypic frequencies for the ABC *combination and* χ^2 *statistics*

	Generation 8–9			Generation 19–20			Generation 28–29		
Genotype	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
111/111	0.276	0.274	0.292	0.309	0.318	0.314	0.191	0.170	0.180
211/211	0.294	0.325	0.284	0.300	0.211	0.279	0.310	0.283	0.295
121/121	0.020	0.014	0.032	0.010	0.003	0.022	0.010	0.026	0.049
112/112	0.101	0.129	0.124	0.161	0.233	0.161	0.142	0.186	0.142
221/221	0.018	0.032	0.019	0.042	0.004	0.022	0.001	0.0	0.002
212/212	0.170	0.158	0.164	0.097	0.081	0.105	0.085	0.055	0.083
122/122	0.096	0.054	0.057	0.066	0.139	0.068	0.230	0.259	0.226
222/222	0.004	0.001	0.001	0.010	0.001	0.004	0.003	0.001	0.004
Hets.†	0.021	0.014	0.028	0.006	0.010	0.025	0.028	0.022	0.020
$\chi^2_{[8]}$*	30.41			111.00			28.57		

† All heterozygous genotypes were combined into a single class.
* $\chi^2_{[8]} = 15.51$ at $P = 0.05$, $\chi^2_{[8]} = 20.09$ at $P = 0.01$.

TABLE 7

Three-locus joint genotypic frequencies for the ABD *combination and* χ^2 *statistics*

	Generation 8–9			Generation 19–20			Generation 28–29		
Genotype	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
111/111	0.200	0.172	0.212	0.264	0.311	0.302	0.203	0.190	0.194
211/211	0.252	0.204	0.240	0.233	0.174	0.267	0.223	0.176	0.239
121/121	0.085	0.033	0.060	0.060	0.117	0.080	0.220	0.267	0.240
112/112	0.175	0.215	0.201	0.206	0.228	0.174	0.133	0.159	0.132
221/221	0.018	0.022	0.010	0.024	0.004	0.016	0.001	0.001	0.004
212/212	0.211	0.266	0.206	0.165	0.112	0.119	0.172	0.157	0.137
122/122	0.030	0.034	0.025	0.015	0.024	0.010	0.021	0.022	0.036
222/222	0.005	0.007	0.009	0.028	0.002	0.009	0.0	0.0	0.002
Hets.†	0.023	0.043	0.038	0.005	0.028	0.024	0.028	0.029	0.018
$\chi^2_{[8]}$ *	54.05			110.30			14.59		

† All heterozygous genotypes were combined into a single class.
* $\chi^2_{[8]} = 15.51$ at $P = 0.05$, $\chi^2_{[8]} = 20.09$ at $P = 0.01$.

potentially unequal selective values is to render the selective values of synthetic genotypes frequency dependent. The frequency dependency arises because the selective value of a synthetic genotype is simply the mean selective value taken over all genotypes contained in the synthetic class. Changes of genotype frequencies within these classes will therefore cause changes in the selective value attributed to the class as a whole. The *ED* locus data have also been included to provide comparisons between combinations involving the tightly linked *EA-EB-EC* triplet and triplets involving a locus that is inherited independently of the other two (KAHLER and ALLARD 1970). Inclusion of the *ED* locus leads to errors

TABLE 8

Three-locus joint genotypic frequencies for the ACD *combination and* χ^2 *statistics*

	Generation 8–9			Generation 19–20			Generation 28–29		
Genotype	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
111/111	0.155	0.136	0.172	0.198	0.216	0.227	0.112	0.114	0.132
211/211	0.168	0.152	0.146	0.204	0.126	0.210	0.173	0.142	0.182
121/121	0.132	0.074	0.097	0.129	0.211	0.154	0.312	0.341	0.298
112/112	0.140	0.142	0.148	0.122	0.095	0.110	0.090	0.077	0.096
221/221	0.102	0.075	0.103	0.052	0.052	0.071	0.049	0.035	0.060
212/212	0.146	0.198	0.154	0.137	0.084	0.088	0.135	0.136	0.111
122/122	0.065	0.108	0.079	0.099	0.154	0.074	0.064	0.105	0.070
222/222	0.070	0.076	0.060	0.056	0.030	0.037	0.037	0.020	0.028
Hets.†	0.021	0.045	0.038	0.004	0.031	0.029	0.028	0.031	0.024
$\chi^2_{[8]}$ *	54.66			103.70			24.23		

† All heterozygous genotypes were combined into a single class.
* $\chi^2_{[8]} = 15.51$ at $P = 0.05$, $\chi^2_{[8]} = 20.09$ at $P = 0.01$.

TABLE 9

Three-locus joint genotypic frequencies for the BCD *combination and* χ^2 *statistics*

	Generation 8–9			Generation 19–20			Generation 28–29		
Genotype	Zygote	Adult	Zygote	Zygote	Adult	Zygote	Zygote	Adult	Zygote
111/111	0.297	0.262	0.292	0.378	0.338	0.408	0.283	0.240	0.286
211/211	0.026	0.026	0.032	0.025	0.004	0.034	0.004	0.017	0.031
121/121	0.155	0.119	0.164	0.119	0.147	0.166	0.143	0.128	0.144
112/112	0.273	0.323	0.289	0.231	0.176	0.191	0.220	0.206	0.188
221/221	0.076	0.030	0.039	0.060	0.116	0.061	0.215	0.247	0.213
212/212	0.012	0.017	0.017	0.027	0.003	0.009	0.005	0.008	0.020
122/122	0.113	0.157	0.122	0.139	0.165	0.101	0.085	0.112	0.081
222/222	0.022	0.025	0.017	0.016	0.024	0.010	0.015	0.014	0.017
Hets.†	0.025	0.042	0.029	0.004	0.026	0.018	0.029	0.028	0.021
$\chi^2_{[8]}$ *	51.74			94.54			20.79		

† All heterozygous genotypes were combined into a single class.
* $\chi^2_{[8]} = 15.51$ at $P = 0.05$, $\chi^2_{[8]} = 20.09$ at $P = 0.01$.

in the estimated genotypic frequency distribution due to the presence of the null allele. While the error is small, its effect is to bias estimates of heterozygous frequencies among zygotes downwards.

Many of the features noted for the single-locus distributions are also apparent at the three-locus level. The likelihood ratio test yields highly significant χ^2 statistics for all three-locus combinations except one (*ABD*, generation 28) and there are substantial changes in genotypic frequencies between life-cycle stages among the eight homozygous genotypes for each three-locus combination. Changes in frequency between zygotic and adult stages within a generation were most pronounced in generation 19. Some three-locus genotypes show a long-term increase in frequency and others decrease. The increase was especially large for homozygous genotypes involving the 1221 gametic type (for loci *EA*, *EB*, *EC* and *ED*, respectively), the gametic type shown to be in great excess by CLEGG, ALLARD and KAHLER (1972).

Quantification of the relative viabilities is achieved in the same manner as described for single loci in APPENDIX B. Three-locus fertility estimates are obtained by comparing relative genotypic frequencies among the adult population to the distribution of post-selection frequencies estimated by the methods of APPENDIX C. In computing post-selection frequencies the average estimate of $\hat{t}$ obtained over loci within a generation was applied to the data of the appropriate generation. A complete parameterization of the mating cycle, as done for single loci, is not feasible for the joint three-locus distribution. This results from the fact that there are now eight independent parameters to estimate and, because outcrossing is infrequent, the eight-dimensional likelihood surface is often too poorly defined to locate maxima. We have therefore estimated gametic frequencies in the effective pollen pool from the census data using methods described in APPENDIX C. Even though the single-locus mating system analysis assures us

that these estimates are different from the true values, the error introduced is less than the order of the outcrossing rate (< 0.01) and it thus has only trivial effects on the estimates of fertility selection.

The selection estimates, normalized to the selective value of the 111/111 homozygous genotype, are reported in Tables 10 to 13. Selective values often differ significantly from unity among the eight homozygous genotypes, *e.g.*, among the 24 values for homozygotes estimated for the *EA-EB-EC* combination, ten are significant and six of the significant values occur in generation 19, in which there is a 24-fold difference between the viabilities of genotypes 221/221 and 122/122. The maximum values are commonly four to six or more times greater than the minimum values among homozygous genotypes. Although fertility estimates tend to diminish the overall effect of these drastic frequency changes, large differences in selective values remain. Infrequent homozygous genotypes tend to be characterized by large and erratic fluctuations in fitness over generations, probably due to sampling variation. There is also no clear-cut evidence for frequency-dependent selection at the three-locus level; again, however, these data provide only a weak test of frequency dependency. Because heterozygotes were rare, components of fitness are reported only for the observed singly heterozygous and doubly heterozygous classes (triply heterozygous classes were not

TABLE 10

ABC *combination selection estimates relative to* $w_{111}^{111} = 1$

	Generation 8–9			Generation 19–20			Generation 28–29		
Genotype	$\hat{v}$	$\hat{l}$	$\hat{w}$	$\hat{v}$	$\hat{l}$	$\hat{w}$	$\hat{v}$	$\hat{l}$	$\hat{w}$
111/111	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
211/211	1.111 (0.112)	0.823	0.914	0.684 (0.074)	1.324	0.905	1.031 (0.125)	0.990	1.020
121/121	0.693 (0.231)	2.193	1.520	0.268 (0.177)	8.300	2.225	2.893 (1.069)	1.870	5.411
112/112	1.285 (0.181)	0.880	1.131	1.407 (0.168)	0.707	0.994	1.472 (0.208)	0.722	1.062
221/221	1.758 (0.487)	0.552	0.971	0.106 (0.051)	4.952	0.527	—— (——)	——	2.072
212/212	0.936 (0.115)	0.970	0.908	0.812 (0.130)	1.279	1.038	0.727 (0.139)	1.420	1.033
122/122	0.936 (0.098)	1.001	0.563	2.058 (0.331)	0.505	1.040	1.272 (0.161)	0.841	1.070
222/222	0.231 (0.254)	0.854	0.197	0.089 (0.094)	4.217	0.377	0.393 (0.455)	3.508	1.378
Single hets.	0.513 (0.216)	5.604	2.862	0.447 (0.325)	12.603	5.800	0.737 (0.261)	1.887	1.230
Double hets.	0.722 (0.408)	1.564	1.123	0.004 (0.002)	0.118	0.000	0.725 (0.287)	1.009	0.648

Standard errors of viability estimates are in parentheses.

TABLE 11

ABD *combination selection estimates relative to* $w_{111}^{111} = 1$

Genotype	Generation 8-9 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 19-20 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 28-29 $\hat{v}$	$\hat{l}$	$\hat{w}$
111/111	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
211/211	0.913 (0.113)	0.993	0.907	0.636 (0.076)	1.558	0.991	0.843 (0.110)	1.324	1.115
121/121	0.441 (0.093)	1.498	0.661	1.642 (0.280)	0.714	1.171	1.299 (0.160)	0.893	1.160
112/112	1.388 (0.178)	0.783	1.088	0.945 (0.110)	0.795	0.751	1.274 (0.182)	0.819	1.044
221/221	1.355 (0.412)	0.347	0.469	0.130 (0.071)	4.268	0.557	1.114 (1.579)	3.515	3.917
212/212	1.422 (0.173)	0.658	0.936	0.580 (0.080)	1.082	0.627	0.979 (0.135)	0.848	0.830
122/122	1.298 (0.322)	0.562	0.729	1.356 (0.450)	0.458	0.621	1.114 (0.343)	1.649	1.838
222/222	1.731 (0.947)	1.006	1.742	0.056 (0.041)	5.115	0.286	—— (——)	——	——
Single hets.	1.943 (0.414)	1.793	3.091	3.912 (2.246)	1.587	7.306	0.930 (0.200)	1.292	0.883
Double hets.	3.895 (3.978)	——	——	0.005 (0.002)	——	——	8.915 (8.848)	——	——

Standard errors of viability estimates are in parentheses.

observed); even then the components are often not estimable (indicated by a dash in the tables) due to observed frequencies of zero. When estimable, heterozygous viabilities fluctuated over a wide range of values; however average viabilities over generations usually exceeded one.

One way of examining the distribution of selective effects at the three-locus level is to compare marginal single- locus estimates to joint three-locus estimates. If we denote joint three-locus frequencies by $a^n(x,y,z)$ and $f^n(x,y,z)$ for adult and zygotic stages in generation n, where the arguments x, y and z denote the genotypic state at the first, second and third loci, respectively, then the viability component, $v(x,y,z) = a^n(x,y,z)/f^n(x,y,z)$. Marginal frequencies can be obtained from appropriate sums, *e.g.*, $a(x,y) = \sum_z a(x,y,z)$ and $a(x) = \sum_y \sum_z a(x,y,z)$. Similarly, marginal selective values can be obtained by comparing marginal frequencies so that $v(x) = a^n(x)/f^n(x)$, etc. The same arguments apply to obtaining marginal fertility estimates from marginal adult and post-selection frequencies. A natural comparison of joint to marginal values is $v(x,y,z) - v(x)v(y)v(z)$. It can be shown that this quantity is equal to

$$\frac{f(x)f(y)f(z)D_a(x,y,z) - a(x)a(y)a(z)D_f(x,y,z)}{f(x)f(y)f(z)f(x,y,z)},$$

TABLE 12

ACD *combination selection estimates relative to* $w_{111}^{111} = 1$

Genotype	Generation 8–9 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 19–20 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 28–29 $\hat{v}$	$\hat{l}$	$\hat{w}$
111/111	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
211/211	1.025 (0.149)	0.763	0.782	0.565 (0.079)	1.548	0.875	0.811 (0.132)	1.106	0.897
121/121	0.637 (0.108)	0.997	0.635	1.498 (0.209)	0.698	1.045	1.081 (0.154)	0.765	0.827
112/112	1.148 (0.172)	0.836	0.960	0.719 (0.114)	1.106	0.795	0.845 (0.163)	1.098	0.929
221/221	0.833 (0.145)	1.088	0.906	0.924 (0.194)	1.273	1.176	0.708 (0.174)	1.476	1.045
212/212	1.544 (0.221)	0.624	0.964	0.562 (0.090)	0.958	0.539	0.999 (0.168)	0.701	0.701
122/122	1.887 (0.333)	0.563	1.062	1.425 (0.217)	0.467	0.666	1.615 (0.313)	0.590	0.953
222/222	1.231 (0.227)	0.624	0.769	0.497 (0.117)	1.147	0.570	0.554 (0.161)	1.183	0.656
Single hets.	2.539 (0.641)	1.233	4.990	6.746 (4.498)	1.786	13.151	1.337 (0.432)	1.421	1.919
Double hets.	1.143 (0.519)	0.666	0.669	8.432 (9.646)	0.313	2.881	0.689 (0.316)	0.328	0.228

Standard errors of viability estimates are in parentheses.

where $D_a(x,y,z) = a(x,y,z) - a(x)a(y)a(z)$ and $D_f(x,y,z) = f(x,y,z) - f(x)f(y)f(z)$ are the deviations of the joint genotypic frequency distributions from independence. If complete independence prevails at both census points (*i.e.*, $D_a = D_f = 0$), the joint estimates are identical to the products of the marginal estimates. If $D_a = D_f \neq 0$ then the difference $v(x,y,z) - v(x)v(y)v(z)$ is directly proportional to the difference in products of marginal genotypic frequencies at the two census points. Therefore, the quantity of interest in comparing marginal to joint estimates is the change in genotypic identity distributions (D_a, D_f) from zygotic to adult stages. Changes in these distributions reflect how selection has altered the structure of the joint genotypic frequencies. Table 14 gives the ratios D_a/D_f for some representative homozygous genotypes. A ratio of one implies $D_a = D_f$, while ratios greater than one mean that the departure of the adult distribution from independence (randomness) is greater than that of the zygotic distribution; conversely, values less than unity imply an adult distribution that approaches independence more closely than the zygotic distribution. Negative values imply a reversal of the direction of deviation from deficiency to excess or excess to deficiency.

The trend among homozygous genotypes is noteworthy. In generation eight all ratios for the *ABC* combination are smaller than unity (in absolute value),

TABLE 13

BCD *combination selection estimates relative to* $w_{111}^{111} = 1$

Genotype	Generation 8–9 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 19–20 $\hat{v}$	$\hat{l}$	$\hat{w}$	Generation 28–29 $\hat{v}$	$\hat{l}$	$\hat{w}$
111/111	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
211/211	1.106	1.080	1.195	0.206	6.098	1.256	4.439	1.564	6.944
	(0.294)			(0.102)			(2.267)		
121/121	0.876	1.200	1.051	1.382	0.935	1.291	1.054	0.934	0.984
	(0.116)			(0.185)			(0.145)		
112/112	1.342	0.808	1.084	0.853	0.899	0.767	1.106	0.765	0.847
	(0.136)			(0.097)			(0.131)		
221/221	0.439	1.159	0.509	2.144	0.444	0.952	1.358	0.735	0.998
	(0.094)			(0.357)			(0.157)		
212/212	1.565	0.855	1.338	0.114	2.745	0.314	1.850	2.037	3.767
	(0.557)			(0.070)			(0.984)		
122/122	1.583	0.676	1.070	1.331	0.509	0.678	1.548	0.614	0.951
	(0.210)			(0.169)			(0.242)		
222/222	1.257	0.551	0.693	1.673	0.365	0.611	1.008	1.065	1.159
	(0.349)			(0.541)			(0.393)		
Single hets.	1.889	1.381	2.298	8.238	1.069	7.875	0.822	1.549	1.078
	(0.396)			(4.498)			(0.199)		
Double hets.	2.608	——	——	4.119	——	——	0.008	——	——
	(2.813)			(4.113)			(0.003)		

Standard errors of viability estimates are in parentheses.

while in generations 19 and 28 only two ratios are smaller than one. Viability selection has thus restructured the joint genotypic frequency distribution for the tightly linked triplet from the earlier to the later generations. Combinations involving those triplets that include the unlinked *ED* locus show a similar trend, as illustrated by the *ABD* combination in Table 14; however, in these cases there is a more pronounced tendency for $|D_a/D_f|$ to exceed one. It is therefore apparent

TABLE 14

D_a/D_f *ratios for* ABC *and* ABD *combinations*

Genotype	Generation 8 *ABC*	*ABD*	19 *ABC*	*ABD*	28 *ABC*	*ABD*
111/111	—0.13	0.23	—0.12	3.30	1.44	1.17
211/211	0.69	—0.02	1.85	5.65	1.09	1.02
121/121	0.77	0.27	1.44	2.80	1.02	1.17
112/112	0.29	1.92	—54.78	1.19	0.70	—0.45
221/221	0.06	0.04	—5.43	2.31	0.95	0.97
212/212	0.03	0.64	0.97	6.67	1.17	1.05
122/122	0.53	18.55	2.20	1.00	1.00	1.63
222/222	0.78	0.83	2.00	—3.65	1.13	1.04

that the multilocus selective effects *are not predictable* from knowledge of marginal effects alone.

DISCUSSION

The data of this experiment demonstrate that selection operates during each of the life cycle stages studied and that the intensity of selection, and even the direction, is often not the same in different stages. The demography of natural selection is thus not simple in annual plant populations and studies that describe only changes in allelic frequencies will perforce fail to describe this diversity of selective modes. This issue is not a minor one. Overall selective values taken by themselves would, in the present population, lead to a serious underestimate of that fraction of the reproductive potential of the population that is absorbed in differential mortality; this in turn bears directly upon arguments involving the concept of genetic load and the maintenance of genetic variability.

It is necessary to confront the problem of assignment of selective effects to gene substitutions at particular loci. To approach the problem, consider a pair of loci, one of which is selected and the other neutral, according to the following model:

$$a(x,y) = w(x)f(x,y) .$$

As before, the arguments x and y denote genotypic state and the census data are assumed to come from two stages in the life cycle separated by an episode of selection. The apparent selection at the neutral locus is given by

$$w(y) = 1 + \frac{\sum_x w(x)D_f(x,y)}{f(y)} ,$$

which, in general, is not equal to one. For two neutral loci associated with a selected locus, the joint fitness at the neutral loci is

$$w(y,z) = \frac{f(y)f(z) + \sum_x w(x)D_f(x,y,z)}{f(y,z)} .$$

WEIR and COCKERHAM (1973) have shown that genotypic identity [$D_f(x,y) = 0$ for all x,y] does not generally occur at equilibrium for mixed mating systems, even without selection. Consequently, selective effects are transmitted throughout the genome in these species; thus, even if the enzyme loci were selected, an issue that cannot be resolved with population data of this kind, measurements of selection at these loci would be confounded with the forces of selection playing over the entire genome. What is measured at the marker loci is the selective flux transmitted throughout the genome by the correlational structure of the entire multilocus distribution. We can conclude that there is substantial selection, but we can isolate neither the phenotypic structures nor the genetic loci causally related to the selection. This point is further reinforced by the changes in the architecture of selection observed between generations for both the single-locus

and the three-locus data. Because these generations were replicated and randomized in the same environment, the observed reversals in the structure of selection are most easily explained by noting that the genetic organization of the population has changed and, therefore, that the distribution of selective effects has also been altered.

Predominantly inbreeding populations like CCV can be thought of as a polymorphic collection of largely homozygous lines. In the case of CCV, genetic exchanges among different lines occur relatively infrequently, so that each line maintains its genetic integrity through many generations. Electrophoresis allows us to identify many of the lines that exist in the population. In general, each line associated with a particular electrophoretic identification will still be a genetically heterogeneous assemblage, because the technique most probably detects only a small fraction of the genetically based variation in the population. We can, however, ask how the variance in selection changes as we divide the population into more and more classes. We expect that if only one locus (or very few loci) is under selection in the population, most of the variance in fitness will be revealed by a few classes. If, on the other hand, many loci are selected, we expect the variance in fitness to increase as we increase the partition of the population. Unfortunately, it is impossible to quantify these hypotheses in the absence of any information on linkage disequilibrium between the observed genetic markers and selected factors that we cannot observe. It is, nevertheless, instructive to examine the variance in selection associated with fitness components.

Table 15 reports the variance of the distribution of fertility estimates for

TABLE 15

Variance in fertility selection

Generation	Triplets		Pairs		Singles	
	ABC	0.075	*AB*	0.115	*A*	0.079
	ABD	0.053	*AC*	0.026	*B*	0.001
8	*ACD*	0.130	*AD*	0.261	*C*	0.013
	BCD	0.025	*BC*	0.047	*D*	0.032
			BD	0.049		
			CD	0.121		
	ABC	0.407	*AB*	0.209	*A*	0.095
	ABD	0.129	*AC*	0.147	*B*	0.009
19	*ACD*	0.118	*AD*	0.222	*C*	0.065
	BCD	0.282	*BC*	0.331	*D*	0.030
			BD	0.047		
			CD	0.091		
	ABC	0.071	*AB*	0.049	*A*	0.011
	ABD	0.108	*AC*	0.035	*B*	0.012
28	*ACD*	0.087	*AD*	0.059	*C*	0.010
	BCD	0.045	*BC*	0.202	*D*	0.015
			BD	0.052		
			CD	0.051		

triplets of loci, pairs of loci and marginal single-locus estimates. Three-locus estimates were made relative to $\bar{l} = \sum_x \sum_y \sum_z a(x,y,z) \cdot \hat{l}(x,y,z) = 1$; marginal one- and two-locus fertility estimates were obtained from marginal one- and two-locus frequency distributions in the manner described earlier. Thus the variance for a particular triplet, say ABC, is $V(ABC) = \sum_x \sum_y \sum_z a(x,y,z) \hat{l}^2(x,y,z) - 1$. Similarly, the variance for a marginal two-locus pair, say AB, is $V(AB) = \sum_x \sum_y a(x,y) \hat{l}^2(x,y) - 1$.

The single-locus distributions partition the population into three distinct classes, two of which are the frequent homozygous classes and the third is the infrequent heterozygous class (recall that the three-locus distributions were based on diallelic classes, so that the marginal single-locus distributions are also diallelic). The two-locus distributions consist of four relatively frequent homozygous classes and six infrequent heterozygous classes, while the three-locus distributions partition the population into eight homozygous classes with most of the remaining classes having observed frequencies of zero. Hence, the number of frequent classes increases from 2^1 to 2^2 to 2^3 as the level of the distribution increases. Examination of Table 15 shows that the variance of the fertility estimates more than doubles from the single-locus to the two-locus level; however, there is no apparent increase thereafter. Failure of the variance to continue to grow can in part be explained by the intense multilocus associations already know to characterize this population (CLEGG, ALLARD and KAHLER 1972). For example, allele C_1 is strongly associated with the diatype A_2B_1; hence, adding C partitions the population further only to the extent that C is uncorrelated with AB. It may also be that nearly all the variance in fertility selection has been revealed by a partition into four classes, although this seems implausible to us.

A second feature brought out by Table 15 is the very large variance in fertility associated with the estimates. This is particularly striking for generation 19. Part of the larger variance observed in generation 19 can be accounted for by the fact that linkage disequilibrium is somewhat lower in intermediate generations than in the early or late generations (CLEGG, ALLARD and KAHLER 1972). However, we have no explanation for the bulk of the increase in variance associated with this generation.

The variance in the viability component of selection follows a pattern similar to the fertility component. The size of the variance in selection is striking, particularly if we consider the sum of the viability and fertility components. We take these data to indicate that the variance in fitness among the lines that compose the population is large, and, therefore, that a substantial fraction of the reproductive potential of the population is absorbed in differential survival and differential fertility.

We thank JEANNETTE DAVIES for carrying out the technical aspects of the electrophoretic assays. We also thank TIMOTHY PROUT and WYATT ANDERSON for helpful criticisms of an earlier version of the manuscript.

LITERATURE CITED

ALLARD, R. W. and S. K. JAIN, 1962 Population studies in predominantly self-pollinated species. II. Analysis of quantitative genetic changes in a bulk-hybrid population of barley. Evolution **26**: 90–101.

ALLARD, R. W., A. L. KAHLER and B. S. WEIR, 1972 The effect of selection on esterase allozymes in a barley population. Genetics **72**: 489–503.

ANDERSON, W. W., 1969 Selection in experimental populations. I. Lethal genes. Genetics **62**: 653–672.

BAILEY, N. J. T., 1951 Testing the solubility of maximum likelihood equations in the routine application of scoring methods. Biometrics **7**: 268–274.

BROWN, A. H. D. and R. W. ALLARD, 1970 Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. Genetics **66**: 133–145.

BUNDGAARD, J. and F. B. CHRISTIANSEN, 1972 Dynamics of polymorphisms: I. Selection components in an experimental population of *Drosophila melanogaster*. Genetics **71**: 439–460.

CEPPELLINI, R., M. SINISCALCO and C. A. B. SMITH, 1955 The estimation of gene frequencies in a random-mating population. Ann. Hum. Genet. **20**: 97–115.

CHRISTIANSEN, F. B. and O. FRYDENBERG, 1973 Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations. Theor. Pop. Biol. **4**: 425–445. ——, 1976 Selection component analysis of natural polymorphisms using mother-offspring samples of successive cohorts. In: *Population Genetics and Ecology*. Edited by S. KARLIN and E. NEVO. Academic Press, New York.

CLEGG, M. T. and R. W. ALLARD, 1973 Viability versus fecundity selection in the slender wild oat, *Avena barbata* L. Science **181**: 667–668.

CLEGG, M. T., R. W. ALLARD and A. L. KAHLER, 1972 Is the gene the unit of selection? Evidence from two experimental plant populations. Proc. Nat. Acad. Sci. U.S. **69**: 2474–2478.

COCKERHAM, C. C. and B. S. WEIR, 1973 Descent measures for two loci with some applications. Theoret. Pop. Biol. **4**: 300–330.

FELDMAN, M. W., I. FRANKLIN and G. J. THOMSON, 1974 Selection in complex genetic systems: I. The symmetric equilibria for the three-locus symmetric viability model. Genetics **76**: 135–162.

JAIN, S. K. and R. W. ALLARD, 1960 Population studies in predominantly self-pollinated species: I. Evidence for heterozygote advantage in a closed population of barley. Proc. Nat. Acad. Sci. U.S. **46**: 1371–1377.

KAHLER, A. L. and R. W. ALLARD, 1970 Genetics of isozyme variants in barley: I. Esterases. Crop Science **10**: 444–448.

KAHLER, A. L., M. T. CLEGG and R. W. ALLARD, 1975 Evolutionary changes in the mating system of an experimental population of barley (*Hordeum vulgare* L.). Proc. Nat. Acad. Sci. U. S. **72**: 943–946.

PROUT, T., 1971a The relation between fitness components and population prediction in *Drosophila*: I. The estimation of fitness components. Genetics **68**: 127–149. ——, 1971b The relation between fitness components and population prediction in Drosophila: II. Population prediction. Genetics **68**: 151–167.

STROBECK, C., 1973 The three-locus model with multiplicative fitness values. Genet. Res. **22**: 195–200.

TEMPLETON, A. R., C. F. SING and B. BROKAW, 1976 The unit of selection in *Drosophila mercatorum*. I. The interaction of selection and meiosis in parthenogenetic strains. Genetics **82**: 349–376.

WEIR, B. S., R. W. ALLARD and A. L. KAHLER, 1972 Analysis of complex allozyme polymorphisms in a barley population. Genetics **72**: 505–523. ——, 1974 Further analysis of complex allozyme polymorphisms in a barley population. Genetics **78**: 911–919.

WEIR, B. S. and C. C. COCKERHAM, 1973 Mixed self and random mating at two loci. Genet. Res. **21**: 247–262.

WORKMAN, P. L. and S. K. JAIN, 1966 Zygotic selection under mixed random mating and self-fertilization: theory and problems of estimation. Genetics **54**: 159–171.

Corresponding editor: J. F. KIDWELL

APPENDIX A

Estimation of mating system parameters and adult genotypic frequencies: The sampling procedure at the adult stage involved drawing random spikes from separate mature plants in the field. Each spike contained a large number of seeds; an average of nine seeds per spike were germinated and the resulting seedlings were assayed electrophoretically when seven days old. Suppose there are k alleles at a locus, and therefore, $g = 0.5\ k(k+1)$ genotypes. Each seedling is classified by genotype, resulting in a vector, $\mathbf{N} = [n_1, n_2, \ldots, n_g]$, of the numbers of each genotype observed within a family, where $R = \sum_{i=1}^{g} n_i$ is the total number of seeds assayed per family (R need not be constant over families, although we will assume so for simplicity). A total of m families are drawn from the population so that the entire data set consists of a $g \times m$ array classified by family and genotype.

Every member of a family will have received one gamete from the same maternal parent. We assume that the second gamete derives from self-fertilization with probability s or from a random outcross with probability $t(=1-s)$. It is further assumed that pollen type frequencies (denoted p_{im}) are uniformly distributed over the population of maternal plants and that the probability of an outcross does not depend on the maternal genotype. Under these assumptions, the conditional probabilities of observing the various progeny genotypes, given the maternal genotype are:

Maternal genotype	Progeny genotype A_iA_i	A_iA_j	A_jA_j	A_jA_l	A_iA_l
A_iA_i	$s + tp_{im}$	tp_{jm}	0	0	tp_{lm}
A_iA_j	$\frac{s}{4} + \frac{t}{2}p_{im}$	$\frac{s}{2} + \frac{t}{2}(p_{im}+p_{jm})$	$\frac{s}{4} + \frac{t}{2}p_{jm}$	$\frac{t}{2}p_{lm}$	$\frac{t}{2}p_{lm}$

where $i \neq j \neq l$; $i, j, l = 1, 2, \ldots, k$.

In applying the model to data, it is also necessary to assume that selection has not intervened between mating and the determination of progeny genotype distributions.

For ease of discussion denote the above conditional probabilities by θ_{ij} ($i, j = 1, 2, \ldots, g$); *i.e.*, θ_{ij} is the probability that a maternal parent of genotype j gives

rise to a progeny of genotype i. Now the probability of observing a particular vector of progeny for the γ^{th} family given the maternal genotype, is

$$P(\mathbf{N}_\gamma|j)\ R!\prod_{i=1}^{g}\frac{\theta_{ij}\,n_{i\gamma}}{n_{i\gamma}!}, \qquad \gamma = 1, 2, \ldots, m.$$

Suppose P_j is the probability of drawing the j^{th} maternal genotype; then the joint probability of drawing a particular maternal genotype and observing the progeny vector $\mathbf{N}_\gamma$ is,

$$P(\mathbf{N}_\gamma, j) = R!\,P_j\prod_{i=1}^{g}\frac{\theta_{ij}\,n_{i\gamma}}{n_{i\gamma}!}\ .$$

To estimate maternal genotypic frequencies we compute the probability of the j^{th} maternal genotype given the observed vector $\mathbf{N}_\gamma$ as,

$$P(j|\mathbf{N}_\gamma) = \frac{R!P_j\prod_{i=1}^{g}\frac{\theta_{ij}\,n_{i\gamma}}{n_{i\gamma}!}}{\sum_{j=1}^{g}P(\mathbf{N}_\gamma|j)P_j} = \pi_{j\gamma}\ .$$

Because we do not know the P_j, we must begin with a provisional estimate. New estimates are obtained from

$$P_j = \frac{\sum_{r=1}^{m}\Pi_{jr}}{m} \quad \text{and}$$

the procedure is repeated until convergence is obtained. This is the gene counting method of CEPPELLINI, SINISCALCO and SMITH (1955).

Our second objective is to estimate the parameters which enter the mixed mating model. There is a $g \times g$ array of the probabilities θ_{ij} and we seek to classify the progeny distributions into a corresponding $g \times g$ array, whose i, j^{th} element, e_{ij}, is the total number of progeny of genotype i whose maternal parent was of type j. Because we do not know the maternal parent of each array, the classification is performed as,

$$e_{ij} = \sum_{r=1}^{m} n_{ir}\pi_{jr}\ ,$$

where e_{ij} is the expected number of progeny of type i produced by maternal parents of type j. Maximum likelihood estimates of the mating system parameters $(s, p_{im}, \ldots, p_{km})$ are obtained by finding the maximum of the log likelihood equation,

$$lnL = C + \sum\sum_{ij} e_{ij} ln\theta_{ij}\ ,$$

where C is a constant. Estimates are obtained following standard numerical techniques.

After the mating system estimates are obtained the procedure for estimating maternal frequencies is repeated. This modifies the array $[e_{ij}]$, hence new estimates of the mating system parameters are obtained. This two-stage process is repeated until convergence has occurred at both steps.

A test of goodness of fit of the estimated parameters to the data was also performed after each iteration. The test statistic is

$$\chi^2 = \sum_{ij}\sum \frac{(E_{ij}-e_{ij})^2}{E_{ij}}$$

where $E_{ij} = MP_j\theta_{ij}$ and where M is the total number of seeds assayed. The test statistic is approximately distributed as Chi square with $k(k^2-k-1)$ degrees of freedom for $k > 2$; for $k = 2$ there is one degree of freedom. All the analyses indicated an acceptable fit of the data to the model (Table 1).

Variance estimates attached to the P_j estimates have two components. One component is due to multinomial sampling of families and the second component arises from the variance in probable maternal parents associated with each progeny array. In the present case, this second variance component is small because the vast majority of progeny arrays consisted of a single homozygous genotype. Consequently, only the variance component associated with multinomial sampling of families has been used in subsequent calculations.

The estimation procedure outlined above is a modification of the procedure of BROWN and ALLARD (1970) and KAHLER, CLEGG and ALLARD (1975). The estimates given in Table 1 differ from those reported earlier by KAHLER, CLEGG and ALLARD (1975); however, these differences result from assigning correct genotypes to a few progeny that were misclassified in the original data set. The modification avoids the classification step of the earlier procedures; actual estimates for this heavily self-fertilizing population are, however, virtually the same using either method.

Estimates of three-locus maternal genotypic frequencies were made using the single-locus outcrossing estimates and gametic frequencies were estimated as described in APPENDIX 3. Thus, only the stage involving the estimation of the π_{jy} was employed in determining three-locus maternal frequencies. The problem is still formidable though, because the array of θ_{ij} involve 1,296 elements. A complete listing of all 1,296 conditional probabilities can be obtained from M. T. CLEGG on request.

APPENDIX B

Single-locus selection estimates: In deriving single-locus selection estimates, we seek to take account of information on pollen pool gene frequencies (denoted p_{im} for the frequency of the i^{th} allele in the pollen pool), and also to estimate viability and fertility components of selection separately. We, therefore, adopt a model of fertility selection which assumes that maternal genotypes differ in expected numbers of viable seed produced, and denote the average number of viable seed produced by a maternal genotype carrying the i^{th} and j^{th} alleles at a locus by l_{ij}. The usual assumptions of the mixed mating model are also invoked; specifically, it is assumed that the probability of an outcross $t(=1-s)$ is constant and independent of maternal genotype and that pollen and ovules are united randomly in the outcross pool. Denoting the frequency of homozygotes carrying

the i^{th} among k alleles by a_{ii} and f_{ii} for adult and subsequent zygotic stages, respectively, and the frequency of heterozygotes carrying the i^{th} and j^{th} alleles by $2a_{ij}$ and $2f_{ij}$, the general recurrence formulae connecting adult to zygotic stages are,

$$f_{ii} = \frac{s}{\bar{l}}\left(l_{ii}a_{ii} + 0.5\sum_{j\neq i} l_{ij}a_{ij}\right) + t(p_{if}p_{im}) \,, \qquad i = 1, 2, \ldots, k \tag{1}$$

$$2f_{ij} = \frac{s}{\bar{l}} l_{ij}a_{ij} + t(p_{if}p_{jm} + p_{jf}p_{im}) \,, \qquad \begin{matrix} i \neq j \\ i, j, = 1, 2, \ldots, k \end{matrix}$$

where $\bar{l} = \sum_{ij} l_{ij}a_{ij}$ and $p_{if} = \frac{1}{\bar{l}} \sum_{j} l_{ij}a_{ij}$ is the frequency of ovules carrying the i^{th} allele *after* fertility selection. A further assumption implicit in the equations (1) is that ovule production is the only factor that limits fertility among selfed genotypes. This assumption is in accord with the reproductive biology of plants because vastly more pollen than ovules are produced by adults.

To estimate p_{if}, we define $\bar{l} = 1$ so that l_{ij} will be estimated relative to unit mean fitness and consider the quantity $p_i = \sum_j f_{ij}$,

$$p_i = sl_{ii}a_{ii} + \frac{s}{2}\sum_{j\neq i} l_{ij}a_{ij} + t\,p_{if}p_{im} + \frac{s}{2}\sum_{j\neq i} l_{ij}a_{ij}$$

$$+ \frac{t}{2}\sum_{j\neq i}(p_{if}p_{jm} + p_{jf}p_{im}).$$

Collecting terms gives

$$p_i = p_{if}\left(s + \frac{t}{2}\right) + \frac{t}{2}p_{im}; \text{ therefore } p_{if} = \frac{2\,p_i - t\,p_{im}}{2 - t} \,.$$

The quantities p_{im} are known from the mating system analysis (APPENDIX A) and the p_i are observable from the zygotic census; hence, the p_{if} are estimable. Moreover, because the p_{if} are functions of maximum likelihood estimates, they also are maximum likelihood estimates.

Solutions to equations (1) for the l_{ij} are,

$$l_{ii} = \frac{2\,f_{ii} - t\,p_{if}(s + 2t\,p_{im})}{sa_{ii}} \qquad i = 1, 2, \ldots, k$$

$$l_{ij} = \frac{2\,f_{ij} - t(p_{jm}p_{if} + p_{im}p_{if})}{sa_{ij}} \qquad \begin{matrix} i \neq j \\ i, j = 1, 2, \ldots, k \end{matrix}$$

These solutions are the maximum likelihood estimators for the l_{ij} (BAILEY 1951).

A consequence of estimating the l_{ij} relative to $\bar{l} = 1$ is that, even if the true selective values are constants, the estimates will give the appearance of frequency-dependent selection. To remove this frequency dependency, we express the selection estimates relative to some standard genotype, say l_{11} as $l'_{ij} = l_{ij}/l_{11}$.

Viability selection estimates were obtained by the method described by CLEGG and ALLARD (1973). This method rests on finding a set of weights that transform the observed zygotic frequencies into the observed adult frequencies. Thus, if f_{ij}^n and a_{ij}^n denote relative zygotic and adult genotypic frequencies in generation n then,

$$a_{ij}^n = v_{ij} f_{ij}^n \quad , \qquad i, j = 1, 2, \ldots, k$$

and therefore

$$v_{ij} = a_{ij}^n / f_{ij}^n \quad .$$

These estimates are general for any number of alleles per locus and they also may be extended to any number of loci. As in the case of the fertility estimates, the v_{ij} may be expressed relative to some standard genotype. The total selection over the transition from zygotic stage generation n to the same zygotic stage in generation $n+1$ (denoted w_{ij}) is equal to the product $v_{ij} l_{ij}$ only when $p_{im} = \sum_j a_{ij}$; however for $t \approx 1\%$, as in the present experiment, the w_{ij} are closely approximated by $v_{ij} l_{ij}$.

Variance estimates: There are three sets of random variables involved in computing the $\hat{l}_{ij}$. These are the adult frequencies, zygotic frequencies and mating system parameters. To take proper account of these sources of variation, we have computed large sample variance estimates which, to the order of accuracy of the variances of genotypic frequencies and mating system parameters, are,

$$\text{Var}\,(\hat{l}_{ij}) = \sum_q \sum_r \sum_x \sum_y \left(\frac{\partial \hat{l}_{ij}}{\partial a_{qr}}\right)\left(\frac{\partial \hat{l}_{ij}}{\partial a_{xy}}\right) \text{Cov}\,(a_{qr}, a_{xy}) +$$

$$\sum_q \sum_r \sum_x \sum_y \left(\frac{\partial \hat{l}_{ij}}{\partial f_{qr}}\right)\left(\frac{\partial \hat{l}_{ij}}{\partial f_{xy}}\right) \text{Cov}\,(f_{qr}, f_{xy}) +$$

$$\sum_q \sum_r \left(\frac{\partial \hat{l}_{ij}}{\partial p_{qm}}\right)\left(\frac{\partial \hat{l}_{ij}}{\partial p_{rm}}\right) \text{Cov}\,(p_{qm}, p_{rm}) + \frac{\partial \hat{l}_{ij}^2}{\partial t} \text{Var}\,(t).$$

$$q,r,x,y = 1,2,\ldots,k$$

When $q=x$, $r=y$ or $q=y$, $r=x$ then, for instance, $\text{Cov}(a_{qr}, a_{qr}) = \text{Var}(a_{qr})$. We have neglected covariance terms between mating system parameters and adult frequencies which are small relative to other sources of variation. As noted by ALLARD, KAHLER and WEIR (1972), the stochastic nature of the outcrossing rate, $\hat{t}$, does not affect estimates of standard errors to two decimal place accuracy; hence, the variance of $\hat{t}$ was neglected in the actual computation. Variances and covariances of genotypic frequencies are obtained from multinomial sampling of genotypes. Variances and covariances of mating system parameters are maximum likelihood estimates and they are obtained from the information matrix (KAHLER, CLEGG and ALLARD 1975).

The v_{ij} involve only two sources of variation, the a^n_{ij} and the f^n_{ij}, because they are computed independent of the mating system. Thus, approximate variance formulae for the v_{ij} are,

$$\text{Var}\,(v_{ij}) = \sum_q \sum_r \sum_x \sum_y \left(\frac{\partial v_{ij}}{\partial a^n_{qr}}\right)\left(\frac{\partial v_{ij}}{\partial a^n_{xy}}\right) \text{Cov}(a^n_{qr}, a^n_{xy})$$

$$+ \sum_q \sum_r \sum_x \sum_y \left(\frac{\partial v_{ij}}{\partial f^n_{qr}}\right)\left(\frac{\partial v_{ij}}{\partial f^n_{xy}}\right) \text{Cov}(f^n_{qr}, f^n_{xy}). \qquad q,r,x,y = 1, 2, \ldots, k$$

APPENDIX C

Estimation of joint three-locus selective values: Although there has been recent interest in investigating equilibrium states for three-locus models (e.g. FELDMAN *et al.*, 1974; STROBECK 1973) and TEMPLETON *et al.* (1976) have developed a method for estimating three-locus fitness values in parthenogenetic populations of *Drosophila mercatorum*, no attempt has yet been made to estimate three-locus selective values in plant populations. Here we develop the estimation theory for the three-locus case under the mixed mating model. The methods involve a straigtforward extension of the estimation procedure given by WEIR, ALLARD and KAHLER (1972).

The assumptions include the usual postulates of the mixed mating model listed in APPENDIX B and we also assume no interference between loci, although this assumption is not necessary. In particular, we assume that the probability of recombination between loci A and B is r_1, between loci B and C is r_2 and between loci A and C is r_3 ($= r_1 + r_2 - 2r_1r_2$ for $r_3 < 0.5$, otherwise $r_3 = 0.5$) where the loci are arranged in the order $A - B - C$. It is further assumed that census data are obtained from zygotes in generation n followed by selection, mating cycle and a zygotic census prior to any selection in generation $n + 1$ [*i.e.*, the assumptions of Model II, WORKMAN and JAIN (1966)]. Following WEIR, ALLARD and KAHLER (1972), we seek to obtain a new vector of relative genotypic frequencies, **M** which represents the frequencies prior to the uniting of gametes but after all selection has occurred. In order to estimate **M** consider the transition equations which connect **M** to the vector of observed $n + 1$ generation relative frequencies, **F**,

Triple homozygote

$$f^{ijk}_{ijk} = s\,\{m^{ijk}_{ijk} + \frac{1}{2}\,[\sum_{l\neq i} m^{ijk}_{ljk} + \sum_{n\neq j} m^{ijk}_{ink} + \sum_{p\neq k} m^{ijk}_{ijp}] + \frac{1}{2}\sum_{l\neq i}\sum_{n\neq j}\,[(1-r_1)^2$$

$$m^{ijk}_{lnk} + r_1^2\, m^{ink}_{ljk}] + \frac{1}{2}\sum_{l\neq i}\sum_{p\neq k}\,[(1-r_3)^2\, m^{ijk}_{ljp} + r_3^2\, m^{ijp}_{ljk}] + \frac{1}{2}\sum_{n\neq j}\sum_{p\neq k}$$

$$[(1-r_2)^2\, m^{ijk}_{inp} + r_2^2\, m^{ink}_{ijp}] + \frac{1}{2}\sum_{l\neq i}\sum_{n\neq j}\sum_{p\neq k}\,[(1-r_1)^2\,(1-r_2)^2\, m^{ijk}_{lnp} +$$

$$r_2^2\,(1-r_1)^2\, m^{ijp}_{lnk} + r_1^2\, r_2^2\, m^{ink}_{ljp} + r_1^2\,(1-r_2)^2\, m^{ljk}_{inp})]\} + t\,(g_{ijk})^2 \qquad \ldots \quad (3)$$

Single heterozygote

$$2f^{ijk}_{ljk} = s\{m^{ijk}_{ljk} + r_1(1-r_1)\sum_{l\neq i}\sum_{n\neq j}(m^{ijk}_{lnk} + m^{ink}_{ljk}) + r_3(1-r_3)\sum_{l\neq i}\sum_{p\neq k}$$

$$(m^{ijk}_{ljp} + m^{ijp}_{ljk}) + \sum_{l\neq i}\sum_{n\neq j}\sum_{p\neq k}[r_1(1-r_1)(1-r_2)^2(m^{ijk}_{lnp} + m^{ljk}_{inp})$$

$$+ r_1(1-r_1)r_2^2(m^{ijp}_{lnk} + m^{ink}_{ljp})]\} + 2t\, g_{ijk}\, g_{ljk} \qquad \text{for } i\neq l$$

Double heterozygote

$$2f^{ijk}_{ljp} = s\{\sum_{l\neq i}\sum_{p\neq k}[(1-r_3)^2 m^{ijk}_{ljp} + r_3^2 m^{ijp}_{ljk}] + r_1r_2(1-r_1)(1-r_2)$$

$$\sum_{l\neq i}\sum_{n\neq j}\sum_{p\neq k}(m^{ijk}_{lnp} + m^{ijp}_{lnk} + m^{ink}_{ljp} + m^{ljk}_{inp})\} + 2t\, g_{ijk}\, g_{ljp} \qquad \text{for } i\neq l,\ p\neq k$$

$$\ldots \quad (3)$$

Triple heterozygote

$$2f^{ijk}_{lnp} = s\{\sum_{l\neq i}\sum_{n\neq j}\sum_{p\neq k}[(1-r_1)^2(1-r_2)^2 m^{ijk}_{lnp} + r_2^2(1-r_1)^2 m^{ijp}_{lnk} +$$

$$r_1^2 r_2^2 m^{ink}_{ljp} + r_1^2(1-r_2)^2 m^{ljk}_{inp}]\} + 2t\, g_{ijk}\, g_{lnp} \qquad \text{for } i\neq l,\ j\neq n,\ k\neq p,$$

where $g_{ijk} = \sum_l \sum_n \sum_p f^{ijk}_{lnp}$ is the gametic input into generation $n+1$. We have adopted the notation of COCKERHAM and WEIR (1974) denoting the relative frequency of heterozygous individuals receiving a gamete bearing genes $A_i B_j C_k$ from one parent and a gamete bearing genes $A_l B_n C_p$ for the other by $2f^{ijk}_{lnp}$ $(2m^{ijk}_{lnp})$ and the relative frequency of triple homozygous individuals receiving a gamete bearing genes $A_i B_j C_k$ from both parents by f^{ijk}_{ijk} (m^{ijk}_{ijk}). The transformation $h^{ijk}_{lnp} = \frac{1}{s}(f^{ijk}_{lmp} - t\, g_{ijk}\, g_{lnp})$, $s \neq 0$, produces a new system of linear equations. Because h^{ijk}_{lnp} is a function of observable quantities, it too is known and the new system of linear equations can be solved for the m^{ijk}_{lnp}. In matrix notation the new system becomes,

$$\mathbf{H} = \mathbf{RM}.$$

It should be noted that the general system of recurrence equations is not completely defined by the equations (3) because actual recombination probabilities depend on the particular loci held heterozygous. The elements of the matrix **R** are therefore also not explicitly defined. In general, if there are x, y and z alleles at the first, second and third loci, respectively, there are xyz triple homozygotes, $0.5\ xyz\ (x+y+z-3)$ single heterozygotes, $0.5\ xyz\ (xy+xz+yz-2x-2y-2z+3)$ double heterozygotes and $0.5\ xyz\ (x-1)(y-1)(z-1)$ triple heterozygotes. Hence, **H** and **M** are of dimension $0.5\ xyz\ (xyz+1)$. A listing of the elements of the matrix **R** for the diallelic case may be obtained from M. T. CLEGG upon request.

In the present investigation, there were three vectors of observations spanning each generation ($\mathbf{F}^n$, $\mathbf{A}^n$, $\mathbf{F}^{n+1}$, zygote, adult and zygote census, respectively, for generations n to $n+1$). The elements of **A** were estimated by the methods described in APPENDIX A and **A** is therefore of dimension 36. The elements of **F** were estimated from direct electrophoretic assay of seedlings. Because the true frequency of multiple heterozygotes which differ only in their phase relationships cannot be determined without progeny testing, **F** is of dimension 27. It is therefore necessary to impose some additional assumptions on the system of equations (3); in particular it is assumed that multiple heterozygotes which differ only in phase relationships have equal frequency (*e.g.*, $f_{lnp}^{ijk} = f_{ijp}^{ink} = f_{lnk}^{ijp} = f_{inp}^{ljk} = 0.25f(il, jn, kp)$ where $f(il, jn, kp)$ denotes half the observed frequency of triple heterozygotes, etc.). The consequence of this assumption is an error in the estimation of the g_{ijk} and therefore an error in *all* elements of **M**. The estimates can be improved by computing the gametic output of **M** and using the new g_{ijk} to compute a new vector **H** and solving the system of equations successively. In the present instance, however, multiple heterozygotes are rare ($< 2\%$ and $t \simeq 1\%$, so that errors in the estimation of g_{ijk} are negligible. Moreover, successive iterations will not converge to the true post-selection frequencies because the initial assumption about **F** (*i.e.*, phase equality) remains dominant. Hence, the first solution was used in subsequent selection estimates. WEIR, ALLARD and KAHLER (1974) have demonstrated that the assumption of phase equality has negligible effects in the two-locus case and their arguments carry over to the three-locus case as well.

Estimates of selective values are obtained by comparing pre- and post-selection frequencies as described for the viability estimates in APPENDIX B.