The Maintenance of Single-Locus Polymorphism.
IV. Models With Mutation From Existing Alleles

Hamish G. Spencer*†,‡ and R. William Marks§

* Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, † Department of Mathematics and Statistics, University of Waikato, Private Bag 3105, Hamilton, New Zealand, and § Department of Biology, Villanova University, Villanova, Pennsylvania 19085

Manuscript received February 6, 1991
Accepted for publication September 26, 1991

ABSTRACT

The ability of viability selection to maintain allelic polymorphism is investigated using a constructionist approach. In extensions to the models we have previously proposed, a population is bombarded with a series of mutations whose fitnesses in conjunction with other alleles are functions of the corresponding fitnesses with a particular allele, the parent allele, already in the population. Allele frequencies are iterated simultaneously, thus allowing alleles to be driven to extinction by selection. Such models allow very high levels of polymorphism to evolve: up to 38 alleles in one case. Alleles that are lethal as homozygotes can evolve to surprisingly high frequencies. The joint evolution of allele frequencies and viabilities highlights the necessity to consider more than the current morphology of a population. Comparisons are made with the neutral theory of evolution and it is suggested that failure to reject neutrality using the Ewens-Watterson test cannot be regarded as evidence for the neutral theory.

The problem of how genetic variation is maintained in natural populations of diploid individuals is a continuing one in theoretical population genetics. Many previous studies investigating selective models have suggested that selection is unlikely to be an explanation for polymorphisms with more than three or so alleles because the proportion of randomly generated fitness matrices affording stable, feasible polymorphisms is extremely small [see Lewontin, Ginzburg and Tuljapurkar (1978) and, for a review, Spencer and Marks (1988) and Marks and Spencer (1991)]. Neutral models, however, have often fared no better. The distributions of allele frequencies reported by Keith (1983) and Keith et al. (1985), for example, do not fit the neutral expectation if it is assumed that the frequencies have reached a steady state.

The work of Spencer and Marks (1988) and Marks and Spencer (1991), however, suggested a possible resolution of the problem. In their models they bombarded a population with a series of mutations while simultaneously iterating the allele frequencies according to the difference equations arising from the allelic viabilities. The fitnesses of the genotypes formed by these mutations were drawn from a normal distribution centered slightly below the viability of the genotype possessing the parent allele, and moreover, the polymorphisms were usually stable, i.e., none of the alleles present were iterating to extinction. In other words, although the proportion of the parameter space corresponding to stable, feasible polymorphisms is small [the result of Lewontin, Ginzburg and Tuljapurkar (1978)], it is not hard to reach via a natural process. Kingman (1988) arrived at a similar conclusion in an analytic study.

The models of Spencer and Marks, however, never generated the higher levels of polymorphism found in natural populations. For example, in the studies referred to previously, one population of Drosophila pseudoobscura from California had 33 electropheretically distinguishable esterase-5 alleles (Keith 1985), and 12 xanthine dehydrogenase alleles (Keith et al. 1985). Marks and Spencer (1991), however, showed that the distribution from which the mutant fitnesses were drawn did affect both the numbers and frequencies of alleles. A natural (and different) way to generate the mutant fitnesses is to make them some function of the fitness of the allele from which the mutant has arisen. Unfortunately, most of the available experimental measures of viabilities are for alleles that have already been subject to selection (and drift), i.e., they are a biased sample of newly arising mutations (Bodmer and Cavalli-Sforza 1972). The data of Mukai, Yoshikawa and Sano (1966) suggest that a reasonable method is to make some proportion lethal as homozygotes and to draw the remaining viabilities from distributions centered slightly below the viability of the genotype possessing the parent allele.
allele. The present paper reports a study of such models. These new models are analogous to the "continuum of alleles" models in the same way that our previously reported ones are to the "house of cards" models (see Turelli 1984). As we have reported previously (Spencer and Marks 1988; Marks and Spencer 1991), our models parallel several in theoretical ecology [see Taylor (1989) and Nee (1990) for overviews]. The new way of generating mutations is analogous in ecological community construction to generating invading species by the speciation of current species rather than by immigration from a large external pool of species (Ginzburg, Akçakaya and Kim 1988; Akçakaya and Ginzburg 1991).

THE MODEL

Monte Carlo simulations were carried out as follows. Every generation one allele already existing in the population was picked as the parent allele, \( A_i \), for the new mutant, \( A_{i+1} \). The probability that a particular allele, \( A_i \), became the parent allele was given by its current frequency. The fitness, \( w_i_{n+1} \), of the \( (i, n+1) \)th heterozygote, \( A_i A_{i+1} \) (\( i \leq n \)) was given by \( (1 - \alpha_i) w_i \rho \), where \( \alpha_i \) was drawn from a \( N(\mu_\alpha, \sigma_\alpha^2) \) distribution (i.e., a normal distribution with mean \( \mu_\alpha \) and variance \( \sigma_\alpha^2 \), where the \( \mu_\alpha \) and \( \sigma_\alpha^2 \) are positive constants). The homozygote \( A_{i+1} A_{i+1} \) had viability 0 with probability \( l \) and with probability \( 1 - l \) the viability was given by \( (1 - \beta) w_{i+\rho} \), where \( \beta \) was drawn from a \( N(\mu_\beta, \sigma_\beta^2) \) distribution, where \( \mu_\beta \) and \( \sigma_\beta^2 \) are positive constants. Any negative viabilities were rejected and another \( \alpha_i \) or \( \beta \) was generated. Note that the correlation between \( \alpha_i \) and \( \alpha_j \) (\( i \neq j \)) is zero (which is not to say that \( w_{i+1} \) and \( w_{j+1} \) are uncorrelated). In biological terms this means that if the mutant allele is such that it is fitter than the parent allele in one genotypic combination, that fact does not affect its other fitnesses. This important assumption was later removed (see below). Uniform pseudorandom numbers were generated using a lagged Fibonacci algorithm [an earlier version of that in Marsaglia and Zaman (1990)] and were transformed into normal variates using the polar method (Knuth 1981).

The mutant allele was given an initial frequency of \( 1 \times 10^{-5} \); the allele frequencies were then all iterated one generation. The new frequency, \( p_i' \) of the \( i \)th allele, \( A_i \), is given by

\[
p_i' = w_i \frac{p_i}{\hat{w}}
\]

in which \( p_i \) is the current frequency,

\[
w_i = \sum_j w_{ij} p_j
\]

is the current marginal fitness of \( A_i \), \( w_{ij} \) is the viability of \( AA_j \) genotypes and

\[
\hat{w} = \sum_i p_i w_i.
\]

is the current mean viability of the population. An allele whose frequency fell below \( 1 \times 10^{-6} \) was deemed to have become extinct.

The usual result of such a dynamic was that the mutant allele immediately became extinct and the population iterated to a position closer to the equilibrium point of the already existing \( n \) alleles. Occasionally, however, a new mutant successfully invaded the population, increasing in frequency from one generation to the next. A successful invasion could lead to either a permanent increase in the number of alleles present or a subsequent reduction in allele number, as the mutant pushed one or more of the existing alleles out of the population. In this way the population evolved through time, gaining and losing alleles.

Our model differs clearly from several previous ones (e.g., Eshel 1971, 1973; Bodmer and Cavalli-Sforza 1972; Guess 1974) in that the population is diploid, there are few restrictions on the viabilities (since we want to study their evolution) and drift is ignored.

Each simulation was initiated with one allele having viability 0.5. After \( 10^4 \) generations the simulation was stopped and the resulting allele frequency and fitness distributions examined. For each initial parameter set (i.e., values assigned to the constants \( l, \mu_\alpha, \sigma_\alpha^2, \mu_\beta \) and \( \sigma_\beta^2 \)) 200 replicate simulations were run, differing only in the random-number generator seeds.

RESULTS AND ANALYSIS

The number of different sets of parameter values which could be investigated is clearly confusingly large. We therefore chose to investigate in detail one such set, which we have called the standard set. Following this analysis, we explored the results of varying some of the parameters, making comparisons with the results of runs with the standard set.

**Standard runs:** Preliminary investigation of the behavior of this model was made using a standard set of parameter values, viz. \( l = 0.05, \mu_\alpha = 0.05, \sigma_\alpha^2 = 5 \times 10^{-4}, \mu_\beta = 0.10 \) and \( \sigma_\beta^2 = 2 \times 10^{-3} \). Hence approximately 1.27% of the \( w_{i+1} \) are expected to be greater than the corresponding \( w_i \) and, similarly, 1.21% (= 0.95 \times 1.27%) of the \( w_{i+1} +1 \) are expected to be greater than the corresponding \( w_{i+1} \).

The levels of polymorphism over time for four replicate standard runs are shown in Figure 1. Comparison with Figure 2 of Spencer and Marks (1988) shows that the overall pattern was similar to their model 2, i.e., where the mutation fitnesses were drawn from a \( U[0, 1] \) distribution (i.e., a uniform distribution on \([0, 1]\)). Again we found that extinctions occurred in clusters, often soon after a successful invasion, and that many multiallele polymorphisms were extremely resistant to invasions. The overall level of polymorphism fluctuated a little more than in model 2, and certainly this new model did on occasion allow many
more alleles than model 2 (see Tables 1 and 2). Again, however, monomorphisms, whether with respect to all the alleles or just the common alleles (those with frequencies >1%), were rare. At generation 10^4, in fact, there were no total-allele or common-allele monomorphisms.

The mean fitness, \( \bar{w} \), over time is also shown in Figure 1. It is interesting to note that it was often approximately constant for long periods of time, even if new mutations were successfully invading the population. Since the model we used was simple enough to obey FISHER's "Fundamental Theorem," the mean fitness necessarily increased when an invasion was successful. What we are seeing, then, are many "quasineutral" invasions. Larger increases in \( \bar{w} \) often occurred quickly and usually when the number of common alleles changed. Since \( \bar{w} \) is a function weighted by the allele frequencies this correlation is not surprising.

The statistics recorded at generation 10,000 for 200 replicates are summarized in Tables 3 and 4. Several comparisons between these results and those of MARKS and SPENCER (1991) are appropriate. First, \( \bar{n} \), the mean number of alleles here (7.755) was significantly greater than in either the uniform (\( \bar{n} = 5.438; t = 15.903, P < 0.001 \)) or the normal (\( \bar{n} = 3.925; t = 33.816, P < 0.001 \)) of MARKS and SPENCER (1991). Indeed, \( \bar{c} \), the mean number of common alleles (4.150) was greater than the total number of alleles in the normal model (\( t = 2.663, P < 0.005 \)). Nevertheless, the number of alleles is still nowhere near the level found in nature at highly polymorphic loci.

We can also look at the fitness matrices and compare them with the result of our previous model 2. The most obvious difference was the complete lack of total heterosis (i.e., all the \( w_{ij} \)'s being greater than the largest \( w_i \)). Even pairwise heterosis (i.e., \( w_{ij} > w_{ii}, w_{jj} \) for all \( i \) and \( j \)) was not as prevalent (only one of the four example runs exhibited pairwise heterosis). Nevertheless, it is clear that heterosis of a kind did naturally evolve in this system. The left hand side of Table 4 shows the means (\( \bar{w}_a \) and \( \bar{w}_g \), respectively) of the within-run mean homozygote and heterozygote fitnesses, as well as the means (\( \bar{w}_h \) and \( \bar{w}_p \), respectively) of the within-run variances of these values. The right-hand side of Table 4 shows these same values, except they have been normalized within each run by the largest viability. [It is the latter numbers that are comparable to those in Table 1 of MARKS and SPENCER (1991)].

The raw and normalized data both showed heterosis in a general sense, i.e., \( \bar{w}_{ij} > \bar{w}_{ii} \). Of course, some of this result was caused by the mutant \( w_{ij} \)'s being larger on average than their \( w_i \)'s. The raw data also showed a decrease in both \( \bar{w}_{ij} \) and \( \bar{w}_{ii} \) as \( n \) the number of alleles in the population, increased: the respective correlations were -0.740 and -0.398 (\( P < 0.001 \) for both). In part, this was due to the presence of rare, presumably less fit, alleles at low frequencies.

Table 3 shows the mean fraction of the alleles at generation 10,000 with a homozygote fitness of greater than 0.5, the initial population mean fitness. This fraction weighted by the frequency of the allele is also shown. In 17 runs (out of 200) did at least one allele in the population have its \( w_i \) greater than 0.5 (16 runs had one fitter allele, one run had two of its four alleles fitter). In other words most of the time, the initial allele was the fittest homozygote and this certainly contributed to this persistence until (at least) generation 10,000 in 195 of those runs. Of the 5 runs where the initial allele became extinct, all had at least one allele, \( A_i \), with \( w_{ii} > 0.5 \). The unweighted proportion of fitter homozygotes (1.5%) was about what would be expected, being somewhat greater than the frequency at which they were generated from the initial allele. The higher value for the weighted frequency (4.3%) reflects the fact noted by SPENCER and MARKS (1988) that alleles with higher homozygote viabilities are likely to be commoner.

Between-model comparisons of normalized fitness
values may be made as well. The decreases in both the normalized \( \bar{w}_i \) and \( \bar{w}_u \) as \( n \) increased was at least in part an artifact of the normalization by the largest \( w_y \) or \( w_u \). Unlike the uniform model, however, there are no clear expected values for (all) the mutations' normalized fitnesses, so we cannot say how much of the decrease was caused by the normalization. Nevertheless, several features of the uniform and normal models were preserved. First, the \( \bar{w}_u \) were less than the \( \bar{w}_y \) for all \( n \), as was necessary if all alleles were to be maintained (Lewontin, Ginzburg and Tuljapurkar 1978). Second, the \( \bar{w}_y \) were greater than the \( \bar{w}_u \) this time for all \( n \), not just for the most frequent. In part this was a consequence of the presence of lethal alleles in the populations which increased the within-run variance in \( w_y \)'s and barely affected the variance of \( w_u \)'s. In turn this reflected the irrelevance of the value of \( w_{y+1} \) to the mutant allele's \( \lambda \) chance of invading successfully (Spencer and Marks 1988).

Table 3 also reveals that more polymorphic populations had in general been subject to more successful

<table>
<thead>
<tr>
<th>Run</th>
<th>( n )</th>
<th>( c )</th>
<th>No. of lethal alleles</th>
<th>No. of successful invasions</th>
<th>( \bar{w} )</th>
<th>Mean</th>
<th>Variance</th>
<th>Raw ( w_y )</th>
<th>Raw ( w_u )</th>
<th>No. ( w_y )'s &lt; Max. (( w_u ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>45</td>
<td>0.50738</td>
<td>0.3100</td>
<td>0.0333</td>
<td>0.4936</td>
<td>0.0010</td>
<td>35</td>
</tr>
<tr>
<td>b</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>25</td>
<td>0.50669</td>
<td>0.4490</td>
<td>0.0014</td>
<td>0.4930</td>
<td>0.0003</td>
<td>21</td>
</tr>
<tr>
<td>c</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>20</td>
<td>0.50756</td>
<td>0.4389</td>
<td>0.0011</td>
<td>0.4844</td>
<td>0.0017</td>
<td>23</td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>29</td>
<td>0.50606</td>
<td>0.4511</td>
<td>0.0015</td>
<td>0.5054</td>
<td>0.0004</td>
<td>2</td>
</tr>
</tbody>
</table>

\( n = \text{total number of alleles} \); \( c = \text{number of common alleles} \) (i.e., those with frequencies > 0.01); \( \bar{w} = \text{mean fitness} (\sum_i p_i w_i/n) \); raw \( w_u \) mean = unweighted mean of homozygote viabilities (\( \bar{w}_u \)); etc.

**Table 2**

Summary statistics for generation 10,000 in 200 replicate standard runs

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of alleles (( n ))</td>
<td>7.755</td>
<td>0.189</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>No. of common alleles (( c ))</td>
<td>4.150</td>
<td>0.084</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>No. of lethal alleles</td>
<td>0.395</td>
<td>0.046</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total frequency of lethal alleles</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.059</td>
</tr>
<tr>
<td>Proportion of fitter alleles</td>
<td>0.015</td>
<td>0.004</td>
<td>0.000</td>
<td>0.500</td>
</tr>
<tr>
<td>Total frequency of fitter alleles</td>
<td>0.043</td>
<td>0.017</td>
<td>0.000</td>
<td>0.964</td>
</tr>
<tr>
<td>No. of successful invasions</td>
<td>28.485</td>
<td>0.363</td>
<td>10</td>
<td>53</td>
</tr>
</tbody>
</table>

**Table 3**

Summary statistics by total number of alleles for allele types at generation 10,000 in 200 replicate standard runs

<table>
<thead>
<tr>
<th>( n )</th>
<th>( N )</th>
<th>Mean No. (( t ))</th>
<th>Standard error</th>
<th>( T )</th>
<th>( \bar{s}[I] )</th>
<th>Mean No. Total freq.</th>
<th>Mean No. Total freq.</th>
<th>No. successful invasions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>6</td>
<td>2.667</td>
<td>(0.211)</td>
<td>0.111</td>
<td>0.167</td>
<td>0.000</td>
<td>0.000</td>
<td>0.167</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>3.333</td>
<td>(0.126)</td>
<td>0.218</td>
<td>0.150</td>
<td>0.133</td>
<td>0.000</td>
<td>0.200</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>3.560</td>
<td>(0.154)</td>
<td>0.298</td>
<td>0.133</td>
<td>0.160</td>
<td>0.001</td>
<td>0.040</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>3.852</td>
<td>(0.190)</td>
<td>0.314</td>
<td>0.119</td>
<td>0.148</td>
<td>0.000</td>
<td>0.185</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>4.125</td>
<td>(0.181)</td>
<td>0.354</td>
<td>0.017</td>
<td>0.333</td>
<td>0.003</td>
<td>0.125</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
<td>4.167</td>
<td>(0.250)</td>
<td>0.390</td>
<td>0.097</td>
<td>0.583</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>9</td>
<td>32</td>
<td>4.593</td>
<td>(0.229)</td>
<td>0.409</td>
<td>0.089</td>
<td>0.531</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>4.400</td>
<td>(0.214)</td>
<td>0.454</td>
<td>0.082</td>
<td>0.267</td>
<td>0.001</td>
<td>0.133</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>4.250</td>
<td>(0.429)</td>
<td>0.486</td>
<td>0.076</td>
<td>0.750</td>
<td>0.001</td>
<td>0.085</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>4.600</td>
<td>(0.427)</td>
<td>0.487</td>
<td>0.071</td>
<td>0.900</td>
<td>0.007</td>
<td>0.000</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>5.167</td>
<td>(0.477)</td>
<td>0.463</td>
<td>0.066</td>
<td>1.000</td>
<td>0.009</td>
<td>0.000</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>4.500</td>
<td>(1.500)</td>
<td>0.476</td>
<td>0.062</td>
<td>0.000</td>
<td>0.000</td>
<td>0.500</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>5.500</td>
<td>(1.500)</td>
<td>0.555</td>
<td>0.058</td>
<td>1.000</td>
<td>0.001</td>
<td>0.500</td>
</tr>
</tbody>
</table>

\( n = \text{total number of alleles} \); \( N = \text{number of replicate runs with} n \) alleles; common alleles in run \( r \) have frequency, \( p^r_i > 1\% \); \( T = \text{mean over} N \) runs of the \( I(=\sum_i (p_i/p)^r_i -1/n)^2\); \( \bar{s}[I] = (n-1)/(n(n+1)) \); Lethal alleles in run \( r \) are those \( A_i \) such that \( w_i^r = 0 \), the mean number being calculated as the mean over \( N \) runs of the number of lethal alleles in each run; total frequency of lethal alleles is the mean over \( N \) runs of the sum of the frequencies of lethal alleles in each run \( (=\sum_i (p_i/p)^r_i p^r_i)/N \), in which \( \delta^r = 1 \) if \( w_i^r = 0 \), or \( \delta^r = 0 \) otherwise; Fitter alleles in run \( r \) are those \( A_i \) such that \( w_i^r > 0.5 \); the mean and total frequencies being calculated analogously to those statistics for lethal alleles.
ties) may be quite persistent. One aspect of the turnover of alleles which we have not yet investigated here is their phylogeny. It would be interesting to see how invasions. The correlation coefficient between $n$ and the number of successful invasions was 0.296 ($P < 0.001$). This was not altogether surprising, but it should be noted that the minimum number of invasions (over all 200 replicates) was ten, enough to have produced the level of polymorphism seen in 90% of the replicates (i.e., those with $n < 11$). Hence the number of successful invasions did not impose a strict upper bound on the level of polymorphism. In fact, the number of successful invasions was usually far in excess of the number of alleles present at generation 10,000: the smallest difference observed in the 200 replicates was 5, in one run where $n$ was 8. This revealed an important feature of the model: the constant turnover of alleles. Indeed, the ultimate fate of all alleles was extinction because of the way the viabilities of mutants were calculated. There is never a perfect allele. In every replicate and even in 10 generations, more than half the number of alleles which had at one time been present in the population had been lost. The one possible exception to the constant turnover of alleles, was the original allele, $A_1$ used to start the simulations. Of the 200 replicates, 195 still maintained $A_1$ at generation 10,000, albeit often at a low frequency. Hence although all alleles eventually become extinct, some (with higher viabilities) may be quite persistent. One aspect of the turnover of alleles which we have not yet investigated here is their phylogeny. It would be interesting to see how alleles present at one time were related, and whether successful invaders drove their parent or other alleles to extinction. In their simulation of a haploid model, BODMER and CAVALI-SFORZA (1972) suggested that the gaps in an allelic phylogeny could indicate the strength of the evolutionary forces acting on a population.

The proximity of the normalized $\bar{w}_n$ and $\bar{w}_y$'s to their maximum and the low values for the $\bar{w}_n$ and $\bar{w}_y$'s suggested that the viabilities may have been very close to neutral. We can test this in the same way as MARKS and SPENCER (1991). For each vector of allele frequencies we took 200 multinomial samples of size 50, and recorded the number of these that the EWENS-WATTERSON test was rejected as not coming from a neutral vector, at the 5% level of significance. Now, if the vector really did consist of neutral allele frequencies, we would expect about 10 ($= 0.05 \times 200$) of the samples to be rejected. In order for us to say the vector is nonneutral, we require that we have significantly more than 10 rejections. We thus have a binomial experiment with 200 trials with the probability of "success," $\pi$, 0.05. Using, again, the 5% level of significance, the null hypothesis that $\pi = 0.05$, will be rejected if the number of "successes" is 18 or more. Hence Table 5 records the number of vectors with 18 or more neutrality rejections. This whole procedure was repeated for sample sizes of 100, 200 and 500 as well and the results are also shown in Table 5.

It is obvious from the table that the overwhelming majority of the time, the EWENS-WATTERSON test was
unable to distinguish our vectors from neutral ones: only 8 (4%) of the vectors were ever detected as different. The level of detection did not increase as the sample size increased, in fact, quite the reverse occurred: with a sample size of 500, only 5 nonneutral vectors were discovered. Another trend was that all the vectors revealed as heterotic contained rather few alleles (n \leq 5).

This failure to detect heterozygote advantage could be interpreted as a convergence toward neutrality. However, the mean heterozygote fitness was significantly lower than the mean heterozygote fitness for all values on n, and \( \tilde{w} \) always increased from \( w_i = 0.5 \). This selection in favor of heterozygotes should result in the observed homozygosities or \( F \) values (i.e., those of samples drawn from our populations) being too small. Indeed, all but one of the (admittedly few) neutrality rejections were due to small \( F \)’s.

The number of lethal alleles (i.e., those \( A_i \) with \( w_{i,j} = 0 \)) also increased with \( n \) (Spearman’s coefficient of rank correlation, \( r_s = 0.382, P < 0.001 \), but these were not usually common (the sum of frequencies of lethals in a replicate was at most 3.5%). One exception was a run where two of the three lethals were common (frequencies of 0.0109 and 0.0225). The high correlation between the number of lethals and their total frequency in a population (\( r_s = 0.981 \)) suggested that higher lethal frequencies were caused by more lethal alleles, rather than individually more frequent alleles.

The position of the allele-frequency vector in the simplex of possible vectors, \((p_1, p_2, \ldots, p_n)\): 0 \leq p, for all i, \( \Sigma p_i = 1 \), can be quantified by \( I = \Sigma (p_i - 1/n)^2 \). I may be interpreted as either the square of the (Euclidean) distance of the vector from the center of the simplex, or the variance of the \( p_i \)’s within the vector. The mean of the \( I \) values for vectors with the same number of alleles is shown in Table 3, along with the value expected if the vectors were uniformly distributed within the simplex. GINZBURG and BRAU-MANN (1980) have shown this latter value to be given by \( \Sigma[I] = (n - 1)/(n(n + 1)) \). Table 3 shows that for \( n \geq 4 \) the observed means were greater than the expected means, a reflection of the uneven allele-frequency distributions with one or two very common alleles and a larger number with low frequencies.

**Varying the parameters of the standard set:** The number of permutations of possible values of \( l, \mu_0, \sigma_n^2, \mu_5, \) and \( \sigma_n^2, \) is clearly enormous, but for our purposes we were interested only in the trends that changes in these parameters produce. In all our discussions of fitnesses we discuss the raw viability values, since the normalizing process within runs vitiates comparisons between runs with different numbers of alleles (n).

**Heterozygote variance (Figure 2a):** Increasing \( \sigma_n^2 \) while keeping all the other parameters constant at the values used in the standard runs directly caused an increase in the proportion of fitter mutant heterozygote viabilities. We used three parameter sets differing only in the value of \( \sigma_n^2 \) in our simulations: the standard set (with \( \sigma_n^2 = 0.0005 \)) and sets with \( \sigma_n^2 = 0.0010 \) and 0.0015. From approximately 1.27% of the \( w_{i,m} \)’s greater being than the corresponding \( w_i \) when \( \sigma_n^2 = 0.0005 \), this fraction increased to about 9.83% when \( \sigma_n^2 = 0.0015 \). Hence the number of successful invasions, the total number of alleles and the number of lethal alleles in the population also increased. In fact the maximum value for \( n \) in one run (when \( \sigma_n^2 = 0.0015 \)) was 38, a result clearly contradicting the assertion (e.g., KIMURA 1983: 282) that viability selection cannot maintain quite large numbers of alleles in large populations. A corollary to the increase in \( \tilde{w} \) with \( \sigma_n^2 \) was a slight decrease in \( \tilde{c} \) and a change from a positive correlation between these two variables to a negative one: runs with the largest values of \( n \) had no room for too many common alleles.

Another change in a correlation was that between \( n \) and \( \tilde{w} \), which increased from -0.250 in the standard runs (\( \sigma_n^2 = 0.0005 \)) to +0.273 by the time \( \sigma_n^2 = 0.0015 \). In the standard runs the \( w_i \) values were tightly clustered (the mean over all 200 runs of the within-run variances in heterozygote viabilities was 0.00096), whereas in the runs with \( \sigma_n^2 = 0.0015 \) they were more variable (respectively, 0.06090). When \( w_{i,j} \) values are similar, differences in \( \tilde{w} \) among runs will depend more on differences in the \( w_{i,j} \) values. Since, in general, \( w_{i,j} \) values were less than \( w_{i,m} \) values, a smaller \( \tilde{w} \) meant a greater level of overall heterosis and more alleles were maintained. Thus standard runs had a negative correlation between \( n \) and \( \tilde{w} \) (even though the correlation between \( \tilde{w} \) and the number of invasions was positive at 0.156). When the \( w_{i,j} \) values are more variable, however, a single high \( w_{i,j} \) value may dominate (and increase) \( \tilde{w} \). (In runs with \( \sigma_n^2 = 0.0015 \) the correlation between \( \tilde{w} \) and the within-run variance in heterozygote viabilities was 0.944 compared to 0.294 in standard runs.) The presence of such a singularly fit heterozygote simultaneously eases the conditions for an invasion, since a mutant derived from either of the two fit (and probably common) alleles of this heterozygote will have its marginal fitness depend mostly on this one high viability. Thus the larger \( \sigma_n^2 \) led to a strongly positive correlation between the number of invasions and \( \tilde{w} \) (0.913), and hence also between \( n \) and \( \tilde{w} \).

**Heterozygote variance and mean (Figure 2b):** We next explored three parameter sets with the same proportion (1.27%) of fitter newly arising mutant heterozygotes, but with different values of \( \mu_0 \) and \( \sigma_n^2 \) (the only differences between the sets): the standard set (\( \mu_0 = 0.05, \sigma_n^2 = 0.0005 \)), one set with \( \mu_0 = 0.10 \) and \( \sigma_n^2 = 0.0020 \) and one with \( \mu_0 = 0.20 \) and \( \sigma_n^2 = 0.0080 \).
Maintain the proportion of fitter mutant heterozygotes constant by simultaneously decreasing $p_a$ and increasing $\sigma_a^2$, still alters the distribution of the mutant $w_{i+n+1}$'s. In particular, the proportion of $w_{i+n+1}$'s greater than $k w_p$, where $k$ is between 0 and 1, decreases. For example, if $k = 0.99$ this proportion decreases from 3.67% in standard runs to 1.68% when $\mu_a = 0.20$ and $\sigma_a^2 = 0.0080$. (Note that if $k = 1$ the proportion is constant at 1.27%.) This means that the proportion of fitter and "almost fitter" $w_{i+n+1}$'s becomes smaller. Since the probability of a successful invasion depends on the $w_{i+n+1}$'s (SPENCER and MARKS 1988) we would thus expect fewer successful invasions and consequently fewer alleles as we decreased $\mu_a$ and increased $\sigma_a^2$. The number of common alleles would not be expected to be so drastically reduced since these are
likely to be the already rare mutants with particularly high \( w_{i+1} \)'s. All these expectations were borne out in our results.

**Homozygote variance and mean (Figure 2c):** We also altered the values for \( \mu_o \) and \( \sigma_o^2 \) as we did for \( \mu_a \) and \( \sigma_a^2 \), keeping constant the proportion of \( w_{i+1} \)'s greater than \( w_{pB} \). The parameter sets used differed only in the values of \( \mu_o \) and \( \sigma_o^2 \); one set with \( \mu_o = 0.05 \) and \( \sigma_o^2 = 0.0005 \), the standard set (\( \mu_o = 0.0010 \), \( \sigma_o^2 = 0.0020 \)) and one set with \( \mu_o = 0.20 \) and \( \sigma_o^2 = 0.0080 \). In this case, however, \( w_{i+1} \) is irrelevant to the success of the attempted invasion (Spencer and Marks 1988); it had a bearing only on the equilibrium allele frequency, \( \bar{p}_{i+1} \), toward which the mutant was iterating; the lower \( w_{i+1} \) the lower \( \bar{p}_{i+1} \), all other things being equal. Hence increasing \( \sigma_o^2 \) led to an increase in the variance of the allele frequencies in the population and such distributions are less resistant to invasions (Spencer and Marks 1988; Cannings and Vickers 1989). Thus increasing \( \sigma_o^2 \) (even while compensating by increasing \( \mu_o \)) led to greater numbers of invasions and alleles in the population.

**Heterozygote and homozygote variances and means (Figure 2d):** Figure 2d shows the results from altering \( \mu_o \), \( \mu_b \), \( \sigma_b^2 \) and \( \sigma_o^2 \), simultaneously so as to maintain the proportion of fitter mutant fitnesses. In all runs \( l = 0 \), \( \mu_o = \mu_b \) (the values used being 0.01, 0.05, 0.10 and 0.20) and \( \sigma_o^2 = \sigma_b^2 \) (the respective values used being 0.00002, 0.0005, 0.0020 and 0.0080). Essentially, these runs combined the effects of Figure 2, b and c. Where these effects were contrary (e.g., as they were on the number of invasions, \( \bar{p}, c, \bar{w}_i \) and \( \bar{w} \), the mean of the 200 \( \bar{w} \) values), the effect of changing \( \mu_o, \sigma_b^2 \) predominated, presumably because there are more heterozygote than homozygote viabilities, and the success of an invasion depended solely on the former.

**Proportion of lethals (Figure 2e):** Altering the proportion, \( l \), of mutants with zero homozygote fitness (through 0%, 1%, 5% (the standard runs), 20% and 50%) caused several statistically significant changes. The increase in the number of lethal alleles in the remaining population was to be expected, along with the decrease in \( \bar{w}_i \). There were also increases in the total number of alleles, the number common and the number of invasions, and a decrease in \( \bar{w}_i \). All these changes resulted from the increased number of lethals present: the stronger “heterozygote advantage” permitted more alleles to be present. The small but significant decrease in \( \bar{w}_i \) came from mutants deriving from lethal parents, since \( w_{pB} \) was necessarily zero. When \( l \) was small, the number of lethals was distributed according to a Poisson distribution (\( l = 0.01 \), \( G_2 = 0.983; l = 0.05 \), \( G_5 = 0.172 \)), but this was not true for larger values (\( l = 0.20 \), \( G_5 = 0.31, 222 \)). Given that we would expect many naturally arising mutants to be lethal as homozygotes, the increase in \( n \) may be quite significant. It should be noted, however, that this increase came entirely from the increase in the number of (presumably rare) lethals: the mean number of “normal” alleles in fact decreased slightly.

A slightly surprising result was the very small change in \( \bar{w} \). One might have thought that an increase in the proportion of lethal mutants from nil to a half would have severely decreased the fitness of the population. The result appears, however, to be another example of Haldane’s (1937) discovery that the effect of a mutation on \( \bar{w} \) is a function not of the mutation’s fitness but of the mutation rate. Both \( \bar{w}_i \) and \( \bar{w}_j \) decreased, however, because they are unweighted means.

**Mutation rate (Figure 2f):** The effects of increasing the number of mutations arising each generation were also investigated using sets with mutations rates of 1 mutant per generation (the standard runs), 2, 3 and 4 mutants per generation. All other parameters were identical to those of the standard runs. The correlations between mutation rate and all the measures shown, except the number of lethal alleles and \( \bar{w}_ij \), were significantly nonzero. Clearly increasing the mutation rate would be expected to increase the number of invasions and the total number of alleles present in the population, but the number of common alleles also increased significantly. The increase in the absolute number of fit mutants also led to an increase in \( \bar{w} \), whereas \( \bar{w}_i \) decreased slightly.

A MODEL WITH CORRELATED FITNESS DEVIATIONS

In the above model the important assumption was made that in the generation of a mutant’s viabilities \( \alpha_i \) and \( \alpha_j \) were independent (and thus uncorrelated). It seems reasonable to argue that such a property might not always hold. For example, since \( w_{i+1} \) is a diploid genotype property and the allele \( A_i \) is unchanged whether in the \( AA_0 \) or \( AA_{i+1} \) combination, it seems likely that the viabilities of these two combinations will be correlated to some extent. We have thus relaxed this assumption and investigated the results.

Using the results in the appendix, we introduced a new parameter, \( \rho \), the correlation in fitness deviations, the \( \alpha_i \)'s. For a given mutant, \( A_{i+1} \), we independently sampled \( \gamma \) from a \( N(0, \sigma^2) \) distribution, \( W \) from \( N(\mu_\delta, \sigma_\delta^2) \) and, for each \( i \leq n, X_i \) from \( N(\mu_\alpha, \sigma_\alpha^2) \), where

\[
\sigma^2 = (1 - \rho)\sigma^2_\alpha
\]

\[
\sigma^2_\alpha = \rho \sigma^2_\alpha
\]

and

\[
\sigma^2_\delta = \sigma^2_\alpha - \rho \sigma^2_\alpha
\]

As before, \( \mu_\alpha, \mu_\delta, \sigma_\alpha^2 \) and \( \sigma_\delta^2 \) were input constants, as
be seen in Figure 3c as a consequence of the correlation between invasions and the level of polymorphism, which was higher for both common and all alleles. Nevertheless, there were some important differences: there were many more invasions and extinctions, and the correlation between invasions and the number of common alleles was more likely to occur for the same \( A_i \) mutant, even though the overall proportion remains the same.

It should be noted, that the changes as \( \rho \) increases from 0.0 to 0.20 were relatively small (albeit statistically significant). Our results, therefore, appear to be fairly robust to changes in \( \rho \) and we suggest that many of the qualitative conclusions we have drawn about deterministic systems will apply whatever the actual details of the generation of mutant viabilities. As before, we note that a model with drift is a logical next step.

**RESULTS WITH CORRELATED DEVIATIONS**

The levels of polymorphism over time for two replicate runs with \( \rho = 0.5 \) are shown in Figure 3a and b. The overall pattern of all our models was preserved: a rapid initial buildup in numbers of alleles, the clustering of extinctions and the correlation between increases in \( \bar{w} \) and changes in the number of common alleles. Nevertheless, there were some important differences: there were many more invasions and extinctions, and the level of polymorphism was higher for both common and all alleles. These differences can be seen in Figure 3c as \( \rho \) was increased. The correlation between \( \rho \) and \( n \) was easily significant (0.591) as was that between \( \rho \) and \( c \) (0.519). This result and the other trends apparent in Figure 3c (e.g., increased \( \bar{w} \), number of lethals) were all the result of the strong correlation between \( \rho \) and the number of successful invasions (\( r = 0.620 \)). As we have argued before, the success of an invasion depends on the \( w_i \)'s of the \( A_i \)'s that are common. A larger value of \( \rho \) means that increases in fitesses \( (w_i) \) compared to \( w_o \) were more likely to occur for the same \( A_i \) mutant, even though the overall proportion remains the same.

The procedure described for the above model was followed to investigate the effect of the correlation: the parameters of the standard runs were held fixed and \( \rho \) was varied from 0 (i.e., standard runs) to 0.5.

**DISCUSSION**

The number of alleles that selection can maintain in a large population is much larger than previous workers have supposed. Although the proportion of the (viability) parameter space permitting stable, feasible polymorphisms becomes extremely small as the number of alleles increases, selection actively seeks out these parts of parameter space. As Taylor (1988, 1989) has eloquently argued, in an ecological context, it is as important to consider how a system is constructed as it is to examine its current state. In the case of simple viability selection these two views lead to radically different conclusions.

The realization that viabilities as well as allele frequencies evolve over time (Ginzburg 1979; Turelli and Ginzburg 1983) also comes more easily from a constructionist view. What is particularly interesting is how the final distribution of allelic viabilities differs from the input distribution of mutant viabilities. Naturally, the input distribution will affect the final distribution, but some of our results appear to be quite robust to changes in the former. For example, in all the above models, as well as those in Spencer and Marks (1988) and Marks and Spencer (1991) monomorphisms were extremely rare. Similarly, a sort of general heterozygote advantage always evolved: even in the models where \( \mu_\alpha < \mu_\beta \) (i.e., we started with heterozygote advantage), \( \bar{w}_j - \bar{w}_i > \mu_\beta - \mu_\alpha \). Investigating the evolution of ecological community structure, Ginzburg, Akçakaya and Kim (1988) showed how the interactions between members evolved to a lower average level than the input values. In a constructionist model of Lotka-Volterra communities, Case (1990) found community-level properties emerging that were not part of the original specifications for the model. It appears that the necessity of
viewing the development of a system and not just its current morphology is widespread, at least in biology.

The number of alleles found in the populations modeled above is larger than those in our previous papers (SPENCER and MARKS 1988; MARKS and SPENCER 1991). Indeed we have now produced models with allele numbers an order of magnitude greater than numbers previously considered the most selection could maintain and with appropriate manipulation of our models' parameters it is clear that almost any level of polymorphism could be produced. Suffice it to say that the higher numbers of alleles found in electrophoretic surveys of natural populations can be matched in the above models.

The ability of our models to maintain high levels of polymorphism, however, does not necessarily mean that viability selection can be invoked to explain the high levels found in nature. The above models, even as a representation of a one-locus system, fail to include several features important in evolution. Foremost among these missing features is probably drift. Although we have not assumed infinite populations (since there is a minimum allele frequency), we have ignored the stochastic effects of finite populations. These effects would be most important for alleles at low frequencies. In particular, some newly arising mutants, although favored by selection, would be eliminated by drift. Unfit mutants do not benefit from drift either: selection will prevent them becoming too frequent and they will eventually be eliminated by some combination of selection and drift. Drift may merely prolong the agony. In essence, then, the effect of drift is similar to that of lowering the mutation rate: fewer favorable mutations become available for selection. Inspection of Figure 2f suggests therefore, that including drift would lower the number of alleles in the population, although by just how much is not clear. BODMER and CAVALLI-SFORZA (1972) argued that introducing selection into a model with drift lowers the number of segregating alleles. Drift also slowed the increase in $\bar{w}$. We are currently examining this question.

Another unrealistic feature of our models is the assumption of constant selection for long periods of time. Temporal variation of selection pressures may temporarily lower the number of alleles present in a population. Nevertheless, we would not expect such variation to lead to long periods of little polymorphism in view of the rapid increases in $n$ at the start of every simulation (see Figures 1 and 3). The net effect of varying viabilities would appear to be an increase in allele turnover. Preliminary investigations (R. W. MARKS and H. G. SPENCER unpublished results) seem to confirm this conjecture.

One of the advantages of examining the way the viabilities evolve is the increased ability to be able to make predictions. One of the strengths of the neutralist model of evolution has been the predictions it makes, e.g., about the distributions of allele frequencies. This hypothesis can therefore be tested fairly easily. Conversely, the selectionist view is rightly criticized for being able to explain everything (and thus nothing): a viability scheme can be invented to explain any particular allele-frequency distribution, for example. Although the models we developed above require refinement before they can make testable predictions, certain patterns emerge from the results.

One of the constant features of the final fitness matrices was heterozygote advantage: $\bar{w}_{H} - \bar{w}_{h} > 0$. This is a necessary condition for a stable equilibrium (LEWONTIN, GINZBURG and TULAPURKAR 1978). The magnitude of the difference, however, is consistently larger than the input difference $(\mu_{F} - \mu_{H})$ and would appear to provide possibilities for predictions different from those of the neutralist view. One obvious problem in estimating $\bar{w}_{H}$ and $\bar{w}_{h}$ in natural populations (over and above those problems in measuring viabilities in general) is that both these measures are unweighted by allele frequencies and their unbiased estimation requires information about all the alleles in the population. In a stable polymorphism, the weighted means of the $w_{i}$'s and $w_{i}$'s also show heterozygote advantage (GINZBURG 1979). This property could be small and difficult to detect, however, since commoner alleles are likely to have higher homozygote viabilities.

A disconcerting feature of our simulation results is the low power of the EWENS-WATTERSON test for neutrality. In the light of the large values for $\bar{w}_{H} - \bar{w}_{h}$, we cannot say that our populations evolve toward neutrality and yet the standard test fails to find a difference. Even worse, increasing the size of the sample of alleles does not improve things. Clearly, in real populations, any failures to reject neutrality using the EWENS-WATTERSON test must be regarded as just that and not as evidence for neutrality.

Last, we should point out the ability of our models to maintain lethal alleles at non-negligible frequencies, as we predicted (MARKS and SPENCER 1991). The levels in our models match fairly well the levels in some natural populations of D. pseudoobscura (BRYANT 1976). It is possible, however, that the way in which we generate mutant viabilities is responsible for this level of lethal alleles, since it allows a lethal to have high heterozygous fitness. Unfortunately, we are almost completely ignorant of the fitness structure of newly arising mutants. Our models suggest that such knowledge could be invaluable to population genetics theory, a point made previously by TURELLI (1984).

We thank D. ADAM, C. CANNINGS, A. CLARK, L. GINZBURG, R. LEWONTIN, J. MCWHIRTER and an anonymous reviewer for their numerous helpful suggestions on this work. P. DANAHER and G.
MARSAGLIA kindly supplied us with the pseudorandom number generator. We are grateful to J. HARMAN-DUNN and J. SMITH for their expert typing and to F. BAILEY, P. ETHERIDGE and D. GILL-GRÉN for assistance with the figures. This work was supported in part by National Institutes of Health grant GM-21179 to R.C. LEWONTIN.

LITERATURE CITED


Communicating editor: B. S. WEIR

APPENDIX

If $X_1$ and $X_2$ are independent random variables from a $N(\mu_x, \sigma_x^2)$ distribution and $Y$ is also independent, but from a $N(0, \sigma_y^2)$ distribution, then $Z_1 = X_1 + Y$ and $Z_2 = X_2 + Y$ are distributed as $N(\mu_x, \sigma_x^2 + \sigma_y^2)$. The covariance between $Z_1$ and $Z_2$, $\sigma_{12}$, is not zero, however, being given by

$$\sigma_{12} = \varepsilon[Z_1Z_2] - \varepsilon[Z_1]\varepsilon[Z_2] \quad \text{(in which } \varepsilon \text{ denotes expectation)}$$

$$\varepsilon[(X_1 + Y)(X_2 + Y)] - \mu_x \mu_x \quad \text{since } \varepsilon[Z_1] = \varepsilon[Z_2] = \mu_x$$

$$= \varepsilon[X_1X_2] + \varepsilon[Y]\varepsilon[X_1 + X_2] + \varepsilon[Y]^2 - \mu_x^2$$

(since $Y$ is independent of $X_1$ and $X_2$)

$$= \mu_x^2 + 0(2\mu_x) + \varepsilon[Y]^2 - \varepsilon[Y]^2 - \mu_x^2$$

(since $X_1$ and $X_2$ are independent and $\varepsilon[Y] = 0$)

$$= \varepsilon[Y]^2 - \varepsilon[Y]^2$$

$$= \sigma_y^2.$$

Hence, the correlation between $Z_1$ and $Z_2$, $\rho$, is given by

$$\rho = \frac{\sigma_{12}}{\sigma_x^2 + \sigma_y^2}$$

$$= \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}.$$

[BHATTACHARYYA and JOHNSON (1977), for example, contains the relevant statistical machinery for these assertions.]