



## Random Drift, Uniform Selection, and the Degree of Population Differentiation

Michael Lynch

*Evolution*, Vol. 40, No. 3. (May, 1986), pp. 640-643.

Stable URL:

<http://links.jstor.org/sici?sici=0014-3820%28198605%2940%3A3%3C640%3ARDUSAT%3E2.0.CO%3B2-L>

*Evolution* is currently published by Society for the Study of Evolution.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ssevol.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## NOTES AND COMMENTS

*Evolution*, 40(3), 1986, pp. 640–643

### RANDOM DRIFT, UNIFORM SELECTION, AND THE DEGREE OF POPULATION DIFFERENTIATION

MICHAEL LYNCH

*Department of Ecology, Ethology, and Evolution, Sheldahl Vivarium,  
University of Illinois, Champaign, IL 61820*

Received June 13, 1985. Accepted January 16, 1986

Evolutionary biologists have long been interested in the forces that are responsible for the apparent degree of uniformity between conspecific populations. Since random drift will ultimately result in the differentiation of isolated populations, the prevention of divergence requires interpopulational gene flow and/or the operation of similar selection pressures on the isolates. Ehrlich and Raven (1969) argued that uniform selection, not gene flow, is the primary cohesive force in evolution. Many authors have subsequently sided with them, although there now appears to be greater appreciation for the subtle ways in which gene flow and selection may interact (Levin, 1979; Levin and Kerster, 1974; Endler, 1977).

A recent paper by Cohan (1984) challenges the very roots of the uniform-selection idea, arguing that instead of resulting in convergent evolution, uniform selection operating on finite populations may often cause more divergence than expected under drift alone. In short, the expected probability of two populations fixing alternate alleles under random drift alone is  $2p(1-p)$ . Directional selection increases the probability of fixation of the favored allele while decreasing the probability of fixation of the alternate allele. Thus, since  $2p(1-p)$  is maximized at  $p = 0.5$ , selection will always decrease the probability of fixation of alternate alleles when the initial frequency of the favored allele is greater than 0.5. However, if  $p$  is small enough (the critical frequency depending on the intensity of selection), the probability that two populations under uniform selection will ultimately become fixed for alternate alleles will be greater than  $2p(1-p)$ .

This is not the first suggestion that the dynamics of gene frequencies of finite populations under selection may be radically different than those of infinite populations. Robertson (1962), for example, demonstrated that selection for heterozygotes (which ensures a polymorphism in an infinite population) may actually accelerate the rate of fixation in small populations, provided the "equilibrium" gene frequency is sufficiently extreme. In contrast to Cohan's (1984) results for alleles that have an additive effect on fitness, however, uniform heterotic selection enhances the probability of fixation of the same allele in different populations.

If the phenomenon pointed out by Cohan (1984) occurs frequently, it has significant implications for our ability to demonstrate the operation of directional selection in natural (including fossil) populations. One way to test for the operation of uniform selection on conspecific populations is to ask whether the popula-

tion mean phenotypes are more similar to each other than expected for populations under random drift alone (see Lande [1977] for the case of divergent selection). This is not an easy task, since the latter process depends on effective population sizes and the time since isolation, neither of which is readily determined in nature. If Cohan (1984) is right, however, even when such tests are possible, there would be little point in performing them because the results would be ambiguous. In the case of directional selection, the uniform selection hypothesis might be evaluated with time-sequence data that allow one to test whether the directional changes in mean phenotypes are more similar than expected by chance, provided common environmental effects can be ruled out. However, estimates of interpopulational variances would be uninformative.

There are a few reasons why the mechanism proposed by Cohan (1984) may be of less general significance than he suggested. As can be seen in Figure 1, the total gain in probability mass caused by selection for low frequency alleles is always less than the loss of probability mass for high frequency alleles. Thus, in order for selection to cause a net enhancement of divergence at the allelic level, the initial frequencies of most favored alleles have to be less than 0.5 (considerably less than 0.5 for large  $N_s$ ). Such circumstances are not impossible, and they might be expected for the constituent loci of a character that has recently been subjected to a change in the direction of selection. Cohan (pers. comm.) has also pointed out how periods of stabilizing selection punctuated by directional selection would be conducive to his hypothesis. Under these conditions, and with no overdominance, purifying selection might normally keep the most favorable allele at a locus near fixation. A new phase of directional selection would then often favor alleles that were kept at low frequencies during the previous period of stabilizing selection. However, contrary to Cohan (1984), it cannot be assumed "that an allele never previously favored by selection will be in low frequency." Certainly, this will not be true in the case of effectively neutral alleles.

In order to see the importance of the initial gene frequency distribution, consider the case of initially neutral alleles. This is appropriate since the null model is the case of pure random drift, and it may be of relevance to some of the studies in Cohan's (1984) table 1 that concern selection responses to completely novel challenges (pesticide and heavy metal tolerance). Provided there are a large number of possible alleles at a

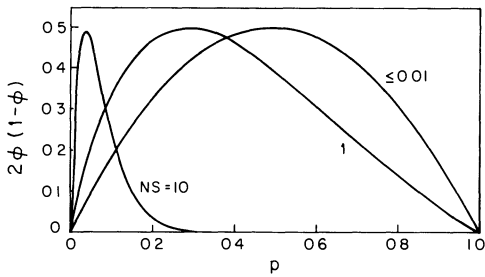


FIG. 1. Probability of fixation of alternate alleles in two populations as a function of the frequency of the selected allele ( $p$ ), the selection coefficient ( $s$ ), and the population size ( $N$ ). Note that the curve labeled  $\leq 0.01$  refers to all  $Ns \leq 0.01$  including the case of pure drift.

locus, the steady state probability distribution for the number of alleles with frequency in the range ( $p, p + dp$ ) under drift-mutation equilibrium is symmetrical, and approximately  $f(p) = 4N\mu(1 - p)^{4N\mu-1} p^{-1}$  (Kimura and Crow, 1964; Nei, 1975). (With the loss of a little generality, this example will not discriminate between effective and total population size.) Thus, for a locus with mutation rate  $\mu$ , the probability of two populations (starting with identical gene frequencies) ultimately becoming fixed for different alleles is

$$D = \int_{1/2N}^{1-(1/2N)} 2f(p)p(1 - p) dp$$

$$= \frac{8N\mu}{4N\mu + 1} \left[ \left(1 - \frac{1}{2N}\right)^{(4N\mu+1)} - \left(\frac{1}{2N}\right)^{(4N\mu+1)} \right]. \quad (1)$$

Since during the fixation process the original alleles may secondarily mutate,  $D$  is more precisely the probability of two populations ultimately becoming fixed by products of two different (in state) ancestral genes. Note that as  $N\mu$  becomes very large, essentially all pairs of populations become fixed for different alleles. This is because the effective number of neutral alleles under drift-mutation equilibrium is approximately  $4N\mu + 1$  (Kimura, 1968).

A similar derivation can be performed for the case of populations exposed to novel, identical selection pressures. Following Cohan (1984), I will assume that selection is operating directionally on loci with additive effects on fitness. The probability of fixation of alternate alleles by two populations is then

$$D = \int_{1/2N}^{1-(1/2N)} 2f(p)\phi(p)[1 - \phi(p)] dp \quad (2)$$

where  $\phi(p)$  is Kimura's (1957) fixation probability and

$$\phi(p) = \frac{1 - \exp(-2Nsp)}{1 - \exp(-2Ns)}$$

$s$  being the selection coefficient and  $p$  being the frequency of the favored allele. As  $Ns$  approaches 0,  $\phi(p)$  approaches  $p$ . I have been unable to find a general,

analytical solution to Equation (2), but have done the integration by Gauss-Laguerre quadrature.

Figure 2 illustrates the dependence of  $D$  (scaled by  $8N\mu$ ) on  $N$  for the case of random drift alone and for several different intensities of selection on finite populations.  $D$  is a rather complicated function of  $N$ ,  $s$ , and  $\mu$ , and although enhanced divergence under selection can occur even when the allele frequency distribution is accounted for, the conditions for it are fairly stringent. In general, the selection intensity must be very weak ( $s < 10^{-3}$ ) since this broadens the range of gene frequencies over which the enhancement effect is possible. In addition  $N\mu$  must be large ( $> 0.1$ ) since this results in more alleles with frequencies in the appropriate range. However,  $N$  cannot be too large or else the selection process becomes essentially deterministic. For most of the situations with initially neutral alleles in which selection enhances the rate of divergence, the magnitude of the effect is not very large, and possibly not detectable by conventional techniques. Moreover, those conditions that are most conducive to divergence, large  $N$  and small  $s$ , are precisely the circumstances that would require a very long time for fixation.

Nevertheless, it is also apparent from Figure 2 that there is a wide range of circumstances for which uniform selection will result in approximately the same level of divergence as random drift alone. Hence, we must remain cognizant that, at the level of individual loci, an increase in the intensity of selection need not result in a dramatic change in the rate of divergence of populations.

Whether the interaction of weak selection with drift can enhance the rate of divergence at the phenotypic level is another question. In the case of pure drift (neutral alleles), the fixation process proceeds without respect to the phenotypic effect of alleles. The between-population variance ultimately resulting from fixation of the  $i$ th locus is therefore  $D_i g_i^2$  where  $g_i^2$  is the variance of effects of alleles initially present at the locus. In the case of selection, however, the fixation process will not be independent of the effects of alleles since, in general, the selection coefficient for two alleles will be positively related to their differences in phenotypic effects. Since  $D$  declines with  $s$  under most circumstances (Fig. 2), the variance of effects among alternately fixed alleles will most often be  $< g_i^2$ . For diallelic loci alone, selection will result in fixation of alternate alleles most frequently at loci where allelic effects are very similar. This further reduces the likelihood of enhanced divergence under selection via Cohan's (1984) mechanism.

Too few directional selection experiments have been performed with sufficient controls and replication to allow a rigorous test of the idea that uniform selection reduces the extent of phenotypic differentiation among isolated populations. However, the experimental results of Clayton and Robertson (1957) and Enfield (1977) indicate an approximately equal level of inter-population variance for selected lines and controls after about 10-40 generations, whereas Jones et al. (1968) found a striking increase in between-population variance over controls after 50 generations of selection for bristle number in *Drosophila*. Similar results have recently been obtained by Cohan and Hoffmann (unpubl.) in experiments with knockdown resistance to ethanol in *Drosophila*. It is impossible to verify wheth-

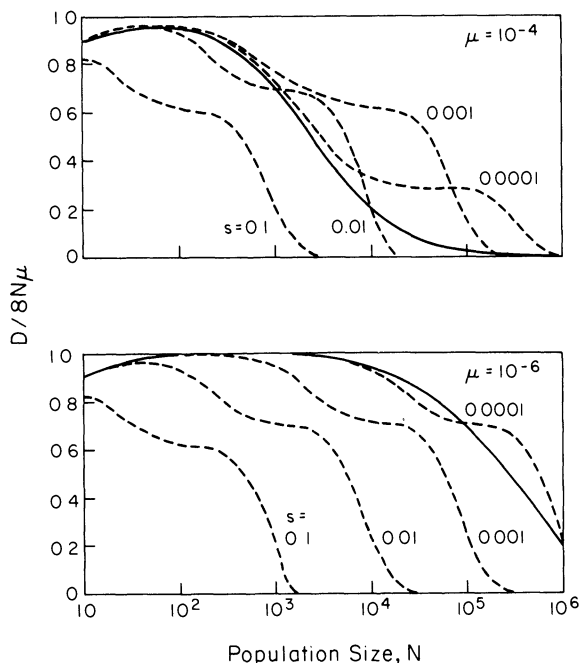


FIG. 2. Expected probability of fixation of alternate alleles in two populations for the case of selection (dashed lines) and pure drift (solid lines), scaled by  $8N\mu$ . The initial gene frequency is a function of population size ( $N$ ) and mutation rate ( $\mu$ ) as defined in the text.

er the relatively high variance between selected lines in these cases is due to the phenomenon envisioned by Cohan (1984). It may in large part result from initial differences in selected lines.

In the long term, new mutations will contribute to the differentiation of isolated populations, as may have been the case for many of the examples cited in Cohan's (1984) table 1. This further complicates the issue. For neutral characters, the rate of increase of between-population variance is approximately  $2V_M$ /generation, where  $V_M$  is the rate of input of new genetic variance via mutation (Chakraborty and Nei, 1982; Lynch and Hill, unpubl.). A lower degree of divergence might be expected for selected populations as favorable alleles lost by drift are ultimately replaced by new favorable mutations and all populations converge on the same mean phenotype. However, a formal theoretical analysis has not yet been performed in this area. Even if such a supposition is true, it might take a considerable amount of time before the effect would be noticeable, because of the stochastic nature of the mutational process. Moreover, mutations with epistatic effects might radically alter the evolutionary paths that different populations take in response to the same selection pressures.

In summary, the important contribution of Cohan's (1984) study is the attention it focuses on the frequent assumption that uniform selection enhances the degree of uniformity among population mean phenotypes. Cohan (1984) may be correct in stating that "weak,

uniform selection imposed on finite, identical populations will be at least as likely to hasten the random divergence between populations as to retard it," but the reasons for this are still far from clear. What is clear is that the power of uniform selection as a homogenizing factor should not be accepted uncritically without careful empirical scrutiny.

#### ACKNOWLEDGMENTS

I thank Jack Su for computational assistance and F. Cohan for helpful comments. Supported by NSF grants BSR 83-06072 and SUB U MICHX98674.

#### LITERATURE CITED

- CHAKRABORTY, R., AND M. NEI. 1982. Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. *Genet. Res.* 39:303-314.
- CLAYTON, G. A., AND A. ROBERTSON. 1957. An experimental check on quantitative genetic theory. II. The long term effects of selection. *J. Genet.* 55:152-170.
- COHAN, F. M. 1984. Can uniform selection retard random genetic divergence between isolated conspecific populations? *Evolution* 38:495-504.
- EHRlich, P. R., AND P. H. RAVEN. 1969. Differentiation of populations. *Science* 165:1228-1232.
- ENDLER, J. 1977. *Geographic Variation, Speciation, and Clines*. Princeton Univ. Press, Princeton, NJ.

- ENFIELD, F. D. 1977. Selection experiments in *Tri-bolium* designed to look at gene action issues, pp. 177-190. In E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.), Proc. Int. Conf. Quant. Genetics Iowa State Univ. Press, Ames.
- JONES, L. P., R. FRANKHAM, AND J. S. F. BARKER. 1968. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. II. Long-term response to selection. Genet. Res. 12:249-266.
- KIMURA, M. 1957. Some problems of stochastic processes in genetics. Ann. Math. Stat. 28:882-901.
- KIMURA, M. 1968. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. Genet. Res. 11:247-269.
- KIMURA, M., AND J. F. CROW. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725-738.
- LANDE, R. 1977. Statistical tests for natural selection on quantitative characters. Evolution 31:442-444.
- LEVIN, D. A. 1979. The nature of plant species. Science 204:381-384.
- LEVIN, D. A., AND H. W. KERSTER. 1974. Gene flow in seed plants. Evol. Biol. 7:139-220.
- NEI, M. 1975. Molecular Population Genetics and Evolution. North-Holland, Amsterdam, Neth.
- ROBERTSON, A. 1962. Selection for heterozygotes in small populations. Genetics 47:1291-1300.

Corresponding Editor: M. Slatkin

*Evolution*, 40(3), 1986, pp. 643-645

#### DEFINITION AND ESTIMATION OF FIXATION INDICES

MASATOSHI NEI

*Center for Demographic and Population Genetics,  
University of Texas at Houston, P. O. Box 20334, Houston, TX 77225*

Received June 28, 1985. Accepted January 16, 1986

Recently, Weir and Cockerham (1984) presented a simplified version of Cockerham's (1969, 1973) method for estimating Wright's (1943, 1951) fixation indices ( $F$ -statistics). They concluded that their method is more precise than the others currently in use. In this study, however, they did not consider the difference in assumptions between the different methods. In the following, I would like to clarify this point, particularly the difference between Cockerham and Weir's method and mine.

Cockerham's (1969) and Weir and Cockerham's (1984) method is based on the following assumptions. 1) Conceptually, there are infinitely many populations which are derived from the same ancestral population at the same time, and  $s$  populations are sampled from this ensemble population. All fixation indices are defined in terms of the ensemble population, and the fixation indices are estimated from genotype frequency data from the  $s$  populations sampled. 2) All populations are statistically independent, and no mutation, no migration, and no selection are assumed (Weir and Cockerham, 1984). However, certain types of migration and selection can be incorporated as long as the statistical independence is maintained (Cockerham, 1969, 1973). 3) The same population size ( $N$ ) is maintained for all populations and for all generations (Weir and Cockerham, 1984).

The above assumptions are essentially the same as those of Wright's (1951) inbreeding theory, and under these assumptions Weir and Cockerham's statistical method seems to be correct. However, the problem lies with the validity of the assumptions. Obviously, most natural populations do not satisfy any of these assumptions. Population sizes usually vary greatly, and

they are almost never constant over time. This is so even if we consider a relatively short evolutionary time. Migration between populations is also a serious problem as Cockerham (1973) himself realized; the migration rate usually varies greatly from population pair to population pair. The number of populations within a species also varies with time, and at a given time some populations are more closely related to each other than to others. In addition to these factors, mutation and selection are also expected to disturb Weir and Cockerham's formulation.

What should we do then? Can we do nothing with natural populations? My solution to this problem is to abandon the above assumptions and to reformulate fixation indices in a different way (Nei, 1977). I have done this by redefining all fixation indices in terms of existing populations for which genotype frequencies are studied. In this case, no assumption is required about the various factors that caused the differentiation of the populations. We are simply interested in the extent of genetic differentiation of the populations studied. Obviously, there is no need to consider replicate populations in this approach. However, since only a limited number of individuals are sampled from each population, it is necessary to estimate fixation indices (parameters), as in the case of estimation of gene frequencies or genetic distances (Nei and Chesser, 1983).

One might think that the fixation indices defined above are less useful than Weir and Cockerham's indices. This is not true. Certainly, my fixation indices refer only to the populations studied, but they can be used for comparing the extent of genetic differentiation among different sets of populations. For example, the extent of genetic differentiation of European popula-