Chapter

9

Relative Fitness: Its Estimation and Its Significance for Environmental Applications of Microorganisms

Richard E. Lenski

Introduction

When considering the environmental application of some microorganism, one of the most important series of questions to ask concerns the opportunity for *persistence* of the population after it has been introduced into the target environment (Committee on Scientific Evaluation of the Introduction of Genetically Modified Microorganisms and Plants into the Environment, 1989; Tiedje et al., 1989). Is it desirable for the introduced population to be self-sustaining? Or is it better if the introduced population performs its intended function and then dies out, being reintroduced only as need arises? The answer will depend, of course, on a comparison of the magnitudes of the additional benefits that may derive from prolonged persistence with the possible costs, if any, that might arise from potential adverse effects caused by persistence.

Once this comparison has been made, it is appropriate to ask: What efforts, if any, have been made to enhance or limit the persistence of the introduced population, as so desired? Such efforts may involve deliberately disabling the microorganism, for example, by incorporating some restrictive nutritional requirement into its genome, in order to limit its persistence; or they may involve selecting a strain that is par-

ticularly well suited to conditions in the target environment in order to extend its persistence.

And finally: What empirical data are there concerning the likelihood of indefinite persistence of the introduced microbial population in the target environment? For many environmental applications of microorganisms, there may already exist a sufficient body of information to reasonably exclude the possibility of adverse effects or to predict with some certainty the persistence of the introduced population (Committee on Scientific Evaluation of the Introduction of Genetically Modified Microorganisms and Plants into the Environment, 1989). However, in cases where the environmental application is less familiar, it may be necessary to evaluate empirically the likelihood of persistence of the introduced population.

A critical factor in determining the likelihood of persistence of any introduced organism is its *fitness* in the new environment. Fitness is a broadly inclusive term that encompasses the combined effects of all biotic and abiotic interactions on an organism's capacity to survive and reproduce in a particular environment. If some organism's fitness in a particular environment is such that each individual, on average, leaves less than one progeny, then we can anticipate that a population of such organisms will not persist indefinitely in that environment. By contrast, if an organism's fitness in some environment is such that each individual, on average, leaves more than one progeny, then we can reasonably expect that a population of these organisms, upon introduction into the environment, will become established and may persist indefinitely. Of course, density-dependent factors, such as resource limitation, must eventually come into play, so that no population can continue to increase indefinitely.

In many cases, a target environment will contain an ecologically self-sustaining population of an indigenous organism that is closely related to the organism proposed for introduction. Ecological interactions between closely related introduced and indigenous populations are likely to be particularly significant for the fate of the introduction; even slight differences in the ability to exploit resources or to escape adversity may affect the opportunity for persistence of the introduced population. Thus, the fitness of an introduced organism *relative* to a closely related indigenous population is likely to be especially useful in predicting the fate of an introduced population.

In this chapter, I show how *relative fitness* can be estimated experimentally and used operationally to predict whether or not an introduced microbial population will persist in a target environment that contains a related indigenous population. I also discuss several assumptions of this approach that may limit its usefulness in certain circumstances.

I do *not* attempt to summarize evidence from laboratory and field studies that may bear on the question of whether there are empirical trends in relative fitness that may be useful in evaluating the likelihood of persistence of an introduced microorganism. It has often been suggested, for example, that genetically engineered microorganisms will usually be less fit than their wild-type counterparts, owing to the energetic and physiological costs associated with carriage and expression of recombinant genes. Various points of view on this subject can be found elsewhere (Brill, 1985; Colwell et al., 1985; Regal, 1986, 1988; Davis, 1987; Sharples, 1987; Lenski and Nguyen, 1988; Committee on Scientific Evaluation of the Introduction of Genetically Modified Microorganisms and Plants into the Environment, 1989; Tiedje et al., 1989).

Definitions and Principles

Fitness is a term that is most widely used in the fields of population genetics and evolutionary biology. A textbook definition of fitness is "The average contribution of one allele or genotype to the next generation or to succeeding generations, compared with that of other alleles or genotypes" (Futuyma, 1986, p. 552). Differences in fitness between alleles or genotypes may cause the frequency of some allele or genotype within the population to change systematically with time. Differences in fitness between alleles or genotypes reflect systematic differences in either mortality or reproduction, which in turn reflect systematic differences in ecological properties such as the ability to compete for limiting resources, susceptibility to predation, and so on. Therefore, *fitness must be viewed as a property of an allele or a genotype that depends upon the environmental circumstances.*

The process of systematic change in the frequency of alleles or genotypes due to differences in fitness is referred to as *selection*. Other processes may also cause changes in the frequency of alleles or genotypes, including *mutation* (which, broadly speaking, includes transposition and segregation of extrachromosomal elements), *recombination* (here, taken to mean intergenomic exchange by processes including gametic fusion, conjugation, transformation, and transduction) and *genetic drift*. Genetic drift can be defined as "Random changes in the frequencies of two or more alleles or genotypes within a population" (Futuyma, 1986, p. 552). Thus, drift differs from selection in that changes in the frequencies of alleles or genotypes are due to chance events, rather than to systematic differences in ecological properties such as competitive ability.

As implied by its earlier definition, fitness is best regarded as a relative property, not an absolute one. Therefore, it usually makes more

sense to state that "genotype j is more fit than genotype k under environmental conditions x, y, and z" than to say that "genotype j is fit" or that "genotype k is unfit." A *selection coefficient* is used to provide a quantitative measure of the difference in relative fitness between two genotypes or alleles. A selection coefficient has units of inverse time and indicates the rate at which one genotype or allele replaces another.

In this chapter, we are primarily interested in organisms that reproduce clonally, often with little or no intergenomic recombination. I shall use the term *strain* to refer to different clonally reproducing genotypes. If one strain is more fit than another in some particular environment, then it would seem reasonable to expect that the less fit strain would eventually be lost from that environment. Indeed, this may often be the case, even when the less fit strain could be sustained indefinitely in the absence of the more fit strain. However, various circumstances have been shown to permit the coexistence of two (or more) strains (or alleles). Several of the most important mechanisms that promote such *genetic polymorphism* are described below.

Selective neutrality: Two strains may coexist almost indefinitely if they are equally fit in a particular environment. In such cases, the strains are said to be selectively neutral. Genetic drift may nonetheless cause some change in the relative frequency of two selectively neutral strains, including even the extinction of one or the other, especially when the population size of one or both of the strains is very small.

Balance between selection and migration, mutation, or gene transfer: A strain that is less fit than another can be maintained by recurring mutation or by migration from another source population (including repeated releases into a target environment). Gene transfer can also maintain an allele or extrachromosomal element in a population in the face of opposing selection (Stewart and Levin, 1977; Levin and Rice, 1980).

Frequency-dependent selection: There are ecological circumstances in which the relative fitness of two strains depends upon their relative frequencies. If strain j is more fit than strain k when strain j is rare, and if strain k is more fit when it is rare, then this frequency-dependent selection actively promotes stable coexistence of the two strains. Consider, for example, the following situation. One strain degrades a substance in the environment that is toxic to a second strain, but there is some "cost" associated with the degradative function (Lenski and Hattingh, 1986). The degradative activities of the first strain reduce the concentration of toxin and thereby promote the growth of the second strain at the expense of the first.

When the frequency of the first strain becomes too low, however, toxins can accumulate to a level where the first strain is again more fit than the second. Such opposing ecological feedbacks thus promote a stable genetic polymorphism.

Spatial heterogeneity: A related situation may occur when two strains differ in their ability to exploit an environment that is spatially heterogeneous. If each strain is fitter than the other in some part of the environment, then they may coexist even though one strain may appear to be fitter "on average" than the other. Coexistence is not an inevitable outcome in these cases and may depend upon other factors such as rates of dispersal.

"Hitchhiking": An allele that is less fit than another may nevertheless be maintained in a population by virtue of its association with a favorable allele elsewhere in the genome. Such associations are termed *linkage disequilibrium* and are especially prevalent in organisms, such as bacteria, that reproduce asexually and where other forms of intergenomic recombination (such as conjugation, transformation, and transduction) usually occur at fairly low rates. Linkage disequilibrium may also be exploited deliberately to prevent a strain intended for release in the environment from persisting. For example, a gene that permits a bacterium to perform some new and useful function in the environment might be placed into a genetic background that contains a nonfunctional allele at another locus, which prevents the bacterium from synthesizing some product essential for growth in the target environment. Linkage patterns can change with subsequent mutation and selection, however, as might happen if a mutation occurred that restored function to the other locus.

Specific Methods

We are interested in determining whether an introduced strain is likely to persist in a certain environment, which already contains a closely related indigenous population. We will assume the availability of a suitable test system, such as a microcosm (Pritchard and Bourquin, 1984), into which the introduced strain can be inoculated and from which samples can be obtained that allow one to enumerate and distinguish the introduced and indigenous populations. (The choice of the test system and the methods for enumerating and distinguishing the two populations need not concern us here, as they are discussed in other chapters.) The basic logic of such an experiment is straightforward. If the *relative* abundances of the two populations change in a systematic fashion with time, then one can infer that one or the other type is fitter in that particular environment. But if the

relative abundances remain essentially constant, then one must conclude that the two types are equally fit, at least within the statistical resolution of the experiment.

A hypothetical data set is shown in Table 9.1; it comprises 11 samples, each containing a pair of densities corresponding to the introduced and indigenous populations. The time 0 sample corresponds to the time of inoculation of the introduced strain into the test environment. Table 9.1 also includes the ratio of the density of the introduced population to that of the indigenous population, computed for each sample.

In Fig. 9.1, the population densities have been transformed to a natural logarithmic scale and plotted against the time of the sample. The two lines indicate the least-squares linear regressions for the population densities; the slopes of these lines provide estimates of the rates of population increase or decrease. The estimated rate of change for the introduced population is -0.052 per unit time, while the estimated rate of change for the indigenous population is $+0.244$ per unit time.

Summary statistics for the regressions are given in Table 9.2; computational methods for regression statistics can be found in many texts (e.g., Kleinbaum and Kupper, 1978) and are widely available in computer software packages. A correlation coefficient provides one measure of the scatter of points about a regression line; it ranges from 0 when the fit of the data to the regression line is extremely poor, to $+1$ or -1 when the fit is extremely good. There is evidently considerable scatter in the data at hand. Standard errors provide a related measure of the statistical error in estimating a slope and permit hypothesis testing by means of a t-test. Neither of the slopes for the introduced and indigenous populations is significantly different from 0.

One can also compute the difference between two slopes, which indicates the difference between the rates of change for the two populations. This difference in slopes provides an estimate of the selection coefficient, or difference in relative fitness, between the introduced strain and the indigenous population. From these data, the estimated selection coefficient is -0.296. Using the standard error of the differ-

TABLE 9.1 Hypothetical Data Used to Compare Trends in Densities of Introduced and Indigenous Populations

Time	Density of introduced population	Density of indigenous population	Ratio of introduced to indigenous populations
0	1.55×10^5	1.91×10^6	8.12×10^{-2}
1	2.45×10^2	5.89×10^3	4.16×10^{-2}
2	4.47×10^5	1.55×10^7	2.88×10^{-2}
3	1.35×10^5	1.02×10^7	1.32×10^{-2}
4	8.71×10^3	1.35×10^6	6.45×10^{-3}
5	2.88×10^3	6.76×10^4	4.26×10^{-2}
6	3.39×10^4	1.23×10^6	2.76×10^{-2}
7	1.51×10^5	9.33×10^7	1.62×10^{-2}
8	8.32×10^3	1.78×10^6	4.67×10^{-3}
9	3.31×10^4	4.68×10^6	7.07×10^{-3}
10	3.09×10^3	3.02×10^6	1.02×10^{-3}

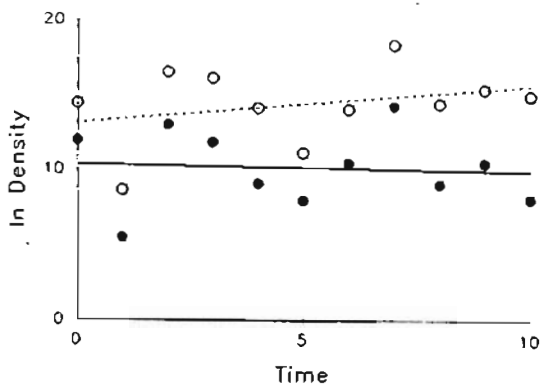

Figure 9.1 Natural logarithm of the densities of the introduced (solid circles) and the indigenous (open circles) populations versus time, using the data from Table 9.1. The lines indicate the least-squares linear regressions for the introduced (solid line) and indigenous (dashed line) populations, as given in Table 9.2. Neither slope is statistically significant from zero, nor are the slopes significantly different one from the other.

TABLE 9.2 Summary of Statistical Analyses of Trends in Densities (Fig. 9.1) and Ratio (Fig. 9.2) of Introduced and Indigenous Populations

	ln Density of introduced population	ln Density of indigenous population	ln Ratio of introduced to indigenous populations
Sample points	11	11	11
Correlation coefficient	0.068	0.310	0.785
Estimated y-intercept	10.391	13.154	−2.763
Estimated slope	−0.052	0.244	−0.296
Standard error of slope	0.254	0.249	0.078
t-statistic	−0.204	0.978	−3.803
Degrees of freedom	9	9	9
Significance level	$0.7 < p$	$0.3 < p$	$p < 0.01$
Estimated difference in slope		−0.296	
Standard error of difference		0.356	
t-statistic		−0.831	
Degrees of freedom		18	
Significance level		$0.4 < p$	

ence in the two slopes (Kleinbaum and Kupper, 1978), one can then perform a t-test to determine whether the two slopes are significantly different from one another; this is equivalent to testing whether the selection coefficient is significantly different from 0. Owing to the rather larger standard error of the difference in slopes, one cannot reject the null hypothesis that the selection coefficient is zero (see Table 9.2).

There is nothing overtly incorrect in the preceding analysis, but it lacks statistical power (i.e., ability to discriminate) for reasons that we shall now see. In Fig. 9.2, the same set of data has been plotted, except that now the two densities have been converted to a single ratio prior to the natural logarithmic transformation. The line indicates the least-squares linear regression fit to the 11 sample points; summary statistics are provided in Table 9.2. The slope of this line provides an estimate of the selection coefficient (Dykhuizen and Hartl, 1983); using these data, we obtain a slope of -0.296, which is equal to the difference between the two slopes obtained from the separate regressions for each population. Note, however, that the scatter of the data points around the regression line calculated from the ratios (see Fig. 9.2) is much less than the scatter around the regression lines calculated from the densities (see Fig. 9.1). This is reflected by a much higher correlation coefficient and by a much smaller standard error of the slope (see Table 9.2). Indeed, the standard error is such that, based upon a t-test, one can claim with a high degree of statistical confidence

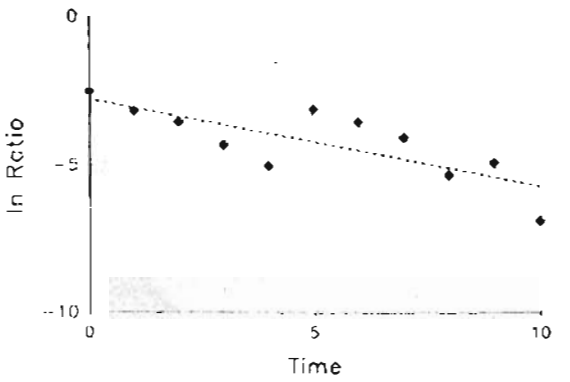

Figure 9.2 Natural logarithm of the ratio of the density of the introduced population to the density of the indigenous population versus time (Table 9.1). The line indicates the least-squares linear regression (Table 9.2). The slope is significantly less than zero ($p < 0.01$), which indicates that the introduced strain is less fit than the indigenous population under the prevailing environmental conditions.

($p < 0.01$) that the selection coefficient is less than zero. Thus, one can conclude that the introduced population is declining relative to the indigenous population at a significant rate.

Why is the selection coefficient based upon the difference in the two separately calculated rates of population change not significant, whereas the selection coefficient based upon the change in the ratio of the two population densities highly significant? The magnitudes of the two selection coefficients are comparable, but the standard error of the latter (0.078) is much smaller than the standard error of the former (0.356).

The smaller standard error arises because the fluctuations in the densities of the two populations are positively correlated in the hypothetical data set, and therefore the fluctuations in the two population densities tend to cancel one another. [By contrast, if the fluctuations in the two populations were independent (i.e., uncorrelated), then the standard error of the selection coefficient would not be reduced by calculating it from the rate of change in the ratio of the two densities.] There are two reasons to expect that, in real data sets, the fluctuations in the densities of two such populations would tend to covary in this manner. (1) Sampling variation: The densities of the two populations will often be estimated from the same physical sample, such as a plug of soil. Any uncontrolled variation in the volume of that sample, or in the efficiency of extraction of the microorganisms from it, will cause the estimated densities of the two populations to covary positively. (2) Environmental variation: Any temporal or spatial variation in environmental qualities, such as temperature or resource concentration, will cause the densities of the two populations to covary positively, provided that the two populations respond similarly to such variables.

These sources of variation are likely to be ubiquitous and will be especially important in natural or seminatural conditions, such as may exist in a microcosm. In effect, one has an internal control for these sources of variation when one analyzes the ratio of the densities of the introduced and indigenous populations, which is lacking when one analyzes separately the densities of the two populations. Therefore, in order to achieve greater statistical power, it is recommended that selection coefficients be estimated from the rate of change in the ratio of the two population densities.

The primary function of statistical inference is to formalize the degree of one's confidence in some conclusion. In the preceding analysis, a single estimate of the selection coefficient was computed by linear regression of the natural logarithm of the ratio of two population densities against time. The standard error of the slope was then used to determine whether that single estimate of the selection coefficient was significantly different from zero.

TABLE 9.3 Hypothetical Set of Five Replicate Estimates of the Selection Coefficient and a Summary of Their Statistical Analysis

Replicate estimates	−0.296
	−0.495
	−0.607
	−0.271
	−0.403
Sample mean	−0.414
Sample standard deviation	0.140
Standard error of the mean	0.063
t-statistic	−6.616
Degrees of freedom	4
Significance level	$p < 0.01$

An alternative approach is to obtain several estimates of the selection coefficient, each based upon an independent experimental replicate. Selection coefficients would be calculated by linear regression as before, but no significance level would be attached to any single estimate. Instead, the standard error of the selection coefficient would be calculated from the sample standard deviation based upon the several independent estimates. One could then use a *t*-test to determine whether the *mean* of the several estimates of the selection coefficient was significantly different from zero. Table 9.3 provides a hypothetical set of five replicate estimates of the selection coefficient and the corresponding statistical analysis. Statistical inferences based on somewhat less intensive sampling of many independent experimental units are generally more reliable than those based on exhaustive sampling of a single experimental unit, and so this second approach is usually preferable. Hurlbert (1984) provides a useful discussion of the importance of proper replication in ecological experiments.

Assumptions

The preceding methods of analysis, like any others, make several assumptions. It is important to be aware of the major assumptions and the circumstances under which the assumptions might be seriously violated. In some instances, it may be possible to modify the methods so that the results and conclusions are less sensitive to certain assumptions.

Assumption 1: The physiological states and ecological circumstances of the introduced and indigenous populations are comparable. If this assumption is met, then any significant difference in fitness must be due to genetic differences. For example, we inferred from our hypothetical data that the introduced strain was less fit than the indigenous population under the prevailing set of environmental conditions. Implicit in this conclusion is the assumption that there were no differences between the introduced and indigenous populations with respect to their physiological states and ecological circumstances. In fact, however, the indigenous population may be physiologically acclimated to the environmental conditions in a particular test system, whereas a recently introduced population may experience some death or delay in growth as it acclimates to those conditions. Also, some fraction of the introduced population may be inoculated into microenvironments that are unsuitable for survival or growth of that organism (e.g., soil interstices for a bacterium that lives in the film associated with soil particles), thereby producing a high rate of decline at the start of an experiment. Either of these artifacts may cause one to systematically underestimate the introduced strain's true fitness in that environment.

One might get around this problematic assumption in a couple of different ways. First, one could exclude from the calculation of the selection coefficient one or more sample ratios obtained at the beginning of an experiment, if visual inspection (or, better yet, an appropriate statistical test) indicates nonlinearity in a plot of the logarithmically transformed ratio against time. Alternatively, one might redesign the experiment so as to avoid this assumption. This could be accomplished by admixing and simultaneously introducing two populations into the test system: one the strain of interest, and the other a strain isolated from the indigenous population and then genetically marked such that it can be distinguished from its counterparts in situ. One would then calculate the selection coefficient from the rate of change in the ratio of the densities of the strain of interest and the marked indigenous strain, with greater confidence that the two populations being compared have experienced similar trauma during their inoculation into the test system. In essence, the marked indigenous strain provides a control for the effects of physiological acclimation and ecological circumstance, which is otherwise lacking. (Of course, one should also perform an additional control to determine whether the genetic marker used to identify the indigenous strain affects its relative fitness.)

Assumption 2: The change in the ratio of the two population densities is due to selection. The ratio of the two population densities may be affected by processes other than selection. Genetic drift (see previous section) is another process that may affect the ratio of two strains, although it is usually an important influence only when one or both populations are extremely small. Also, genetic drift should not generally bias the estimation of the selection coefficient; rather, it introduces another

source of statistical error but, by definition, does not cause systematic deviations in favor of one strain or another.

Potentially more serious violations of this assumption may arise when genetic processes that convert one genotype into another, such as plasmid segregation or conjugation, occur at a high rate. In these instances, the rate of change in the relative abundances of the two strains may not be adequately described by a single parameter, and more complicated analyses may be required. Lenski (1991) presents methods that can be used to distinguish the effects on population dynamics of selection from the effects of segregation. Stewart and Levin (1977) and Levin and Rice (1980) present mathematical models of bacterial conjugation that may be adapted to disentangle selection and gene transfer. In theory, either segregation or gene transfer can permit two strains to coexist, even if one is more fit, provided that the less fit type is regenerated de novo by the relevant genetic process. However, if selection coefficients are large relative to the rates governing these genetic processes, as may often be the case, then these genetic processes should have little effect on an estimate of the selection coefficient obtained from a relatively short-term experiment.

Assumption 3: The selection coefficient is constant. The utility of relative fitness as a criterion for predicting the fate of an introduced strain is affected only very little by certain kinds of variation in the selection coefficient, but it may be much more sensitive to other sources of variation. Of particular concern are those instances in which the sign of the selection coefficient, rather than simply its magnitude, may vary.

Figure 9.3 illustrates the effects of three different types of variation in the selection coefficient on the short-term persistence of an introduced population. In each case, the introduced strain is usually less fit than the indigenous population (i.e., the selection coefficient is less than zero), but occasionally the introduced strain is more fit (i.e., the selection coefficient is greater than zero). And in all three cases, the observed dynamics over the short term (e.g., 20 or so time units) would seem to indicate that the introduced strain will not persist.

In the first and simplest case, the variation in the selection coefficient occurs temporally over a relatively short scale. In particular, imagine that the introduced strain has a -0.1 selection coefficient relative to the indigenous population for 99 percent of the time. For the other 1 percent of the time, however, the introduced strain has a $+0.05$ selection coefficient relative to the indigenous population. Over the long term, the introduced population is expected to decline relative to the indigenous population at a rate that is simply the average of the variable rates in time, i.e., $(0.99)(-0.10) + (0.01)(+0.05) = -0.0985$. Therefore, over both the short term and the long term, the introduced population declines toward extinction.

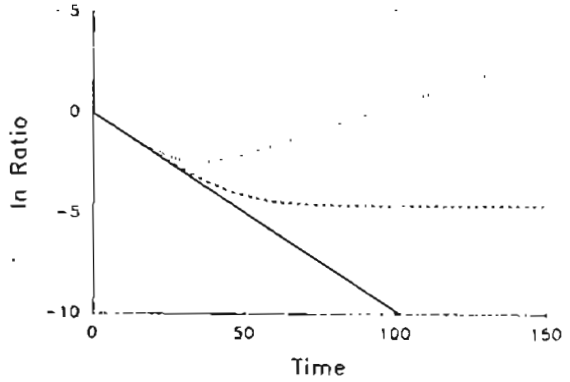

Figure 9.3 Effects of three different types of variation in the selection coefficient on the persistence of an introduced population. The natural logarithm of the ratio of the density of an introduced population to the density of an indigenous population is plotted against time. *Solid line:* Fine-scale temporal variation in the selection coefficient. *Dashed line:* Fixed spatial variation in the selection coefficient. *Dotted line:* Genetic variation in the selection coefficient. In each of the three cases that is illustrated, the introduced strain is usually less fit than the indigenous population (selection coefficient = -0.1), but occasionally the introduced strain is more fit (selection coefficient = $+0.05$). See text for further details.

In the second case, the variability in the selection coefficient exists spatially and remains fixed in time. This variation may arise from differences in the qualities of certain habitats or microenvironments with respect to their suitability for one population or the other. For example, imagine that the introduced strain has a -0.1 selection coefficient relative to the indigenous population in 99 percent of the habitats. In the other 1 percent of the habitats, the introduced strain has a $+0.05$ selection coefficient relative to the indigenous population. (The average densities supported per unit volume in the habitats are assumed to be equal in this example, and the two populations are initially distributed randomly across the habitats.) Over the short term, the introduced population declines relative to the indigenous population. Over the long term, however, the introduced and indigenous populations stably coexist by virtue of their differential utilization of the two types of habitats. Frequency-dependent selection (see previous section) may similarly give rise to stable coexistence after a period of initial change.

In the third case, there is genetic variation in the selection coeffi-

cient such that some subset of the introduced population is more fit than the indigenous population, even though the majority is much less fit. In particular, imagine that 99 percent of the individuals in the introduced population have a −0.1 selection coefficient relative to the indigenous population, while the other 1 percent of the individuals in the introduced population have a +0.05 selection coefficient relative to the indigenous population. These differences are assumed to be heritable (i.e., genetically determined). Over the short term, the introduced population once again declines relative to the indigenous population. Over the long term, however, the introduced population increases relative to the indigenous population and may even cause its extinction, owing to the fitness advantage that accrues to an initial minority of the introduced population.

A hypothetical example of how this situation might arise is as follows. The introduced strain has been modified genetically to provide it with the enzymatic functions necessary to utilize as a resource some compound in the environment not available to the indigenous population. At the same time, the introduced strain has been deliberately handicapped ecologically by the incorporation of some restrictive nutritional requirement. If a spontaneous mutant that lost the ecological handicap occurred in the introduced population, then the mutant subpopulation would have the advantage of the strain intended for introduction, without the disadvantage, and hence it could increase unexpectedly. Kim et al. (1991) have developed quantitative models that may be useful in predicting the likelihood of a reversal in the outcome of selection due to secondary genetic changes in an introduced population.

Summary and Conclusions

The ability to predict whether a microbial population that may be released into the environment will persist or disappear is an important consideration when evaluating the possible benefits and risks of an application, particularly one that is unfamiliar. An important factor in determining the likelihood of persistence of the introduced microorganism is its fitness in the target environment. Fitness is a broadly inclusive term that encompasses the combined effects of all biotic and abiotic interactions on an organism's capacity to survive and reproduce in a particular environment.

In many cases, a target environment may already contain an ecologically self-sustaining population of an indigenous microorganism that is closely related to the microorganism proposed for introduction. In such cases, the difference in fitness between the introduced and indigenous microorganisms will be especially important in determining the fate of the introduced population. The difference in fitness between two clonally reproducing genotypes, or strains, is termed a selection coefficient. Selection coefficients have units of inverse time and indicate the rate at which one strain replaces another. This chapter has illustrated the basic methods used to estimate selection coefficients.

The basic design of an experiment to estimate the selection coefficient, or difference in fitness, for two strains is quite simple. Populations of the two strains are mixed together in some initial ratio in a test environment. At various time points, samples are obtained from the test system and the ratio of the two population densities in each sample is determined. If one strain is more fit than the other in the test environment, then this ratio should increase or decrease with time in a systematic fashion; if the ratio of the population densities remains essentially constant, then one must conclude that the two strains are equally fit in the test environment, at least within the statistical limit of resolution.

It should be emphasized that these methods of estimating selection coefficients are widely applicable to a variety of different test environments, ranging from simple laboratory culture systems, such as shake flasks and chemostats, to more complex seminatural systems, such as microcosms. These methods may also be useful in monitoring field trials. In principle, all that is required to measure a selection coefficient is the ability to monitor the ratio of two strains sharing a common environment.

Selection coefficients that are estimated in this manner invariably have an internal control; whether a particular sample or experimental unit is nutrient rich or poor, for example, the same is true for both strains. Because of this internal control, differences in relative fitness that are calculated directly from rates of change in the ratio of two population densities will often be more accurate than differences in fitness calculated from separate measurements of the growth properties for each strain.

Of course, these methods are not without their assumptions. The most serious violations may occur when the introduced and indigenous populations are not in comparable physiological states or ecological circumstances; when segregation or transfer of genes, and not selection, is primarily responsible for the change in the relative densities of the strains; and when the sign of the selection coefficient varies as a consequence of spatial heterogeneity, frequency-dependent selection, or secondary genetic changes in the less fit population. In such cases, successful prediction of the fate of an introduced microbial population may require more complex analyses or changes in the experimental design.

Acknowledgments

This chapter is respectfully dedicated to the memory of Steve Cuskey. Preparation of this chapter was supported by a cooperative research grant from the Gulf Breeze Laboratory of the U.S. Environmental Protection Agency (CR-815380) and by an award from the National Science Foundation (BSR-8858820).

References

Brill, W. J. 1985. Safety concerns and genetic engineering in agriculture. *Science* 227:381–384.

Colwell, R. K., E. A. Norse, D. Pimentel, F. E. Sharples, and D. Simberloff. 1985. Genetic engineering in agriculture. *Science* 229:111–112.

Committee on Scientific Evaluation of the Introduction of Genetically Modified Microorganisms and Plants into the Environment. 1989. *Field Testing Genetically Modified Organisms: Framework for Decisions*. National Academy Press, Washington, D.C.

Davis, B. D. 1987. Bacterial domestication: Underlying assumptions. *Science* 235:1329–1335.

Dykhuizen, D. E., and D. L. Hartl. 1983. Selection in chemostats. *Microbiol. Rev.* 47: 150–168.

Futuyma, D. J. 1986. *Evolutionary Biology*, 2d ed. Sinauer Associates, Sunderland, Mass.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* 54:187–211.

Kim, J., L. R. Ginzburg, and D. E. Dykhuizen. 1991. Quantifying the risks of invasion by genetically engineered organisms. In L. R. Ginzburg (ed.), *Assessing Ecological Risks of Biotechnology*. Butterworth-Heineman, Boston, pp. 193–214.

Kleinbaum, D. G., and L. L. Kupper. 1978. *Applied Regression Analysis and Other Multivariable Methods*. Duxbury Press, North Scituate, Mass.

Lenski, R. E. 1991. Quantifying fitness and gene stability in microorganisms. In L. R. Ginzburg (ed.), *Assessing Ecological Risks of Biotechnology*. Butterworth-Heinemann, Boston, pp. 173–192.

Lenski, R. E., and S. E. Hattingh. 1986. Coexistence of two competitors on one resource and one inhibitor: A chemostat model based on bacteria and antibiotics. *J. Theor. Biol.* 122:83–93.

Lenski, R. E., and T. T. Nguyen. 1988. Stability of recombinant DNA and its effects on fitness. In J. Hodgson and A. M. Sugden (eds.), *Planned Release of Genetically Engineered Organisms (Trends in Biotechnology / Trends in Ecology and Evolution Special Issue)*. Elsevier, Cambridge, England, pp. 18–20.

Levin, B. R., and V. A. Rice. 1980. The kinetics of transfer of nonconjugative plasmids by mobilizing conjugative factors. *Genet. Res., Camb.* 35:241–259.

Pritchard, P. H., and A. W. Bourquin. 1984. The use of microcosms for evaluation of interactions between pollutants and microorganisms. *Adv. Microb. Ecol.* 7:133–215.

Regal, P. J. 1986. Models of genetically engineered organisms and their ecological impact. In H. A. Mooney and J. A. Drake (eds.), *Ecology of Biological Invasions of North America and Hawaii*. Springer-Verlag, New York, pp. 149–162.

Regal, P. J. 1988. The adaptive potential of genetically engineered organisms in nature. In J. Hodgson and A. M. Sugden (eds.), *Planned Release of Genetically Engineered Organisms (Trends in Biotechnology / Trends in Ecology and Evolution Special Issue)*. Elsevier, Cambridge, England, pp. 36–38.

Sharples, F. E. 1987. Regulation of products from biotechnology. *Science* 235:1329–1332.

Stewart, F. M., and B. R. Levin. 1977. The population biology of bacterial plasmids: A priori conditions for the existence of conjugationally transmitted factors. *Genetics* 87: 209–228.

Tiedje, J. M., R. K. Colwell, Y. L. Grossman, R. Hodson, R. E. Lenski, R. N. Mack, and P. J. Regal. 1989. The planned introduction of genetically engineered organisms: Ecological considerations and recommendations. *Ecology* 70:298–315.

Microbial Ecology

Principles, Methods, and Applications

Morris A. Levin
Maryland Biotechnology Institute
University of Maryland
Baltimore, Maryland

Ramon J. Seidler
U.S. Environmental Protection Agency
Corvallis, Oregon

Marvin Rogul
Maryland Biotechnology Institute
University of Maryland
Baltimore, Maryland

McGraw-Hill, Inc.
New York St. Louis San Francisco Auckland Bogotá
Caracas Lisbon London Madrid Mexico Milan
Montreal New Delhi Paris San Juan São Paulo
Singapore Sydney Tokyo Toronto