Genetic versus phenotypic models of selection

Weissing, Franz

Published in:
Journal of Mathematical Biology

DOI:
10.1007/BF02409749

Publication date:
1996

J. Math. Biol. (1996) 34: 533–555

Journal of Mathematical Biology

Genetic versus phenotypic models of selection: can genetics be neglected in a long-term perspective?

Franz J. Weissing

Department of Genetics, University of Groningen, P.O. Box 14, 9750 AA Haren, The Netherlands.
e-mail: weissing@biol.rug.nl

Received 24 June 1994; received in revised form 15 May 1995

Abstract. Game theoretical concepts in evolutionary biology have been criticized by populations geneticists, because they neglect such crucial aspects as the mating system or the mode of inheritance. In fact, the dynamics of natural selection does not necessarily lead to a fitness maximum or an ESS if genetic constraints are taken into account. Yet, it may be premature to conclude that game theoretical concepts do not have a dynamical justification. The new paradigm of long-term evolution postulates that genetic constraints, which may be dominant in a short-term perspective, will in the long run disappear in the face of the ongoing influx of mutations. Two basic results (see Hammerstein; this issue) seem to reconcile the dynamical approach of long-term population genetics with the static approach of evolutionary game theory: (1) only populations at local fitness optima (Nash strategies) can be long-term stable; and (2) in monomorphic populations, evolutionary stability is necessary and sufficient to ensure long-term dynamic stability. The present paper has a double purpose. On the one hand, it is demonstrated by fairly general arguments that the scope of the results mentioned above extends to non-linear frequency dependent selection, to multiple loci, and to quite general mating systems. On the other hand, some limitations of the theory of long-term evolution will also be stressed: (1) there is little hope for a game theoretical characterization of stability in polymorphic populations; (2) many interesting systems do not admit long-term stable equilibria; and (3) even if a long-term stable equilibrium exists, it is not at all clear whether and how it is attainable by a series of gene substitution events.

Key words: Frequency dependent selection – ESS – Nash equilibrium – Game dynamics – Mendelian population – Multilocus model – External stability – Phenotypic stability – Optimization

1 Introduction

The theory of long-term evolution, developed by Eshel, Matessi, Hammerstein and others, has attracted considerable attention since it seems to reconcile two apparently incompatible approaches towards natural selection. The population genetical approach bases its evolutionary predictions on a dynamical model of the selection process. This approach is "genetic" since detailed assumptions on mating, reproduction, and inheritance have to be made in order to specify a dynamical selection model. In contrast, the predictions of evolutionary ecology are usually based on a notion of "adaptedness" and hence on one or the other (game theoretical) optimality principle. This second, static approach is "phenotypic" in that it focuses on fitness differences and neglects the details of mating and reproduction. The two complementary approaches are widely used but, unfortunately, they may lead to different and even contradictory predictions. It is now well established (e.g. Moran 1964; Akin 1979; Karlin and Lessard 1986) that natural selection is often not an optimizing process and that stable evolutionary equilibria do often not correspond to fitness maxima or evolutionarily stable strategies. These results may seem counterintuitive at first sight, but they merely reflect the fact that natural selection does not only depend on phenotypic fitness differences but also on the genetic transmission of these differences from one generation to the next. As a consequence, information which has been built up during selection at the phenotypic level may be destroyed by the reshuffling processes of Mendelian genetics.

The fact that selection at the pheotypic level may be dominated by constraints at the genetic level has led many population geneticists to conclude that a full genetic specification is required for a proper understanding of selection. Evolutionary ecologists have difficulties to accept this standpoint since virtually nothing is known about the genetic basis of most evolutionarily interesting traits. For them, the paradigm of *long-term evolution* provides new hope that a purely phenotypic characterization of selection is still feasible. In other words, the claim that the two approaches towards natural selection are inconsistent may be premature. Proponents of the new paradigm concede that genetic constraints are often decisive for the outcome of selection in a short-term perspective. They argue, however, that these constraints will be alleviated in the long run by newly arising mutations whose evolutionary success is less governed by genetic constraints than by their phenotypic properties. As a consequence, they expect that the discrepancy between dynamic stability and fitness optimality will vanish in a long-term evolutionary perspective.

The present paper aims at a generalisation and a critical evaluation of two results which may be viewed as the cornerstones of the theory of long-term evolution (see Hammerstein, this issue). For linear frequency dependent selection at two loci in a randomly mating population, Hammerstein and Selten (1994) showed that only local fitness optima ("Nash strategies") can be long-term stable (Theorem 1) and that, in monomorphic populations,

evolutionary stability is necessary and sufficient for long-term dynamic stability (Theorem 2). Here, I shall demonstrate that the scope of these results can be extended to non-linear frequency dependence, to multiple loci, and to some forms of nonrandom mating. By means of a simple example, I shall point to a gap in the proof of Hammerstein and Selten's Theorem 2. I have not been able to close this gap completely, and instead I derive a slightly weaker version of this result. The new proofs given here try to avoid technical detail and to make the arguments as transparent as possible. I hope that this will shed some new light on the scope and the limitations of the paradigm of long-term evolution.

2 The phenotypic approach to frequency-dependent selection

The phenotypic approach to selection is a (game theoretical) optimization approach which is based on three ingredients (e.g. Parker and Maynard Smith 1990). First, optimality theory requires a clearcut description of all traits that are phenotypically feasible. The resulting *phenotype set* (or *strategy set*) characterizes the phenotypic constraints on adaptive evolution. Second, a *fitness function* (or: *payoff function*) is needed which relates the "adaptedness" of a phenotypic trait to characteristics of the base population or the environment. Finally, an *optimization criterion* is required which singles out "well-adapted" phenotypic traits as primary candidates for the outcome of an adaptive selection process.

In the context of evolutionary game theory, phenotypic traits correspond to the (mixed) strategies of an evolutinary game. A *mixed strategy* (denoted by bold-faced letters such as $\boldsymbol{p}$, $\boldsymbol{q}$, or $\boldsymbol{Q}$) is a probability distribution over a finite set $\sigma_1, \ldots, \sigma_n$ of pure strategies. If selection is *frequency dependent*, the fitness of an individual, $F(\boldsymbol{q}, \bar{\boldsymbol{p}})$, will not only depend on its own strategy $\boldsymbol{q}$ but also on the *population strategy* $\bar{\boldsymbol{p}}$, i.e. on the average strategy of the population. We shall assume that the fitness function F is continuously differentiable. For most fitness concepts, it is plausible to assume that the fitness of a mixed strategy $\boldsymbol{q}$ is the expected value of the fitness values of the pure strategies used by $\boldsymbol{q}$:

$$F(\boldsymbol{q}, \bar{\boldsymbol{p}}) = \sum_{i=1}^{n} q_i \cdot F(\sigma_i, \bar{\boldsymbol{p}}) \; . \tag{1}$$

Let us therefore suppose that the fitness function F is linear in its first component. Many applications of evolutionary game theory (see, e.g., Maynard Smith 1982) consider pairwise contests with random matching of opponents. In such a case, F is given by a payoff matrix $\boldsymbol{A}$, $F(\boldsymbol{q}, \bar{\boldsymbol{p}}) = \boldsymbol{q} \cdot \boldsymbol{A}\bar{\boldsymbol{p}}$, and, as a consequence, F is also linear in its second component. We shall refer to *linear* frequency dependent selection if F is linear in both components (Technically speaking, the term "affine" is more adequate, but the term "linear" is generally used; see Taylor 1996).

The phenotypic approach founds its evolutionary predictions on a detailed analysis of the fitness function F. In classical game theory, a payoff function F is in the first place investigated with respect to Nash equilibrium strategies. A strategy $\boldsymbol{p}^*$ is a *Nash strategy* if no alternative strategy has a higher payoff against $\boldsymbol{p}^*$ then $\boldsymbol{p}^*$ itself:

$$F(\boldsymbol{p}^*, \boldsymbol{p}^*) \geqq F(\boldsymbol{q}, \boldsymbol{p}^*) \tag{2}$$

for all $\boldsymbol{q}$. In classical game theory, the Nash solution concept is based on strict rationality requirements which have of course no place in evolutionary game theory. The solution concept of evolutionary game theory is based on the idea that the stability of a wild-type population requires that the resident strategy has a fitness advantage over invading mutant strategies. Consider a mutant strategy $\boldsymbol{q}$ which enters a monomorphic population of $\boldsymbol{p}^*$-strategists with a small frequency ε. $\boldsymbol{p}^*$ will be *immune* against invasion by $\boldsymbol{q}$ if the $\boldsymbol{p}^*$-strategists have a fitness advantage in the resulting dimorphic population with population strategy $\boldsymbol{p}_\varepsilon = (1 - \varepsilon)\boldsymbol{p}^* + \varepsilon\boldsymbol{q}$. $\boldsymbol{p}^*$ is called an *evolutionarily stable strategy* (*ESS*) if it is immune against invasion by any alternative strategy $\boldsymbol{q}$. Formally, $\boldsymbol{p}^*$ is an ESS if

$$F(\boldsymbol{p}^*,(1 - \varepsilon)\boldsymbol{p}^* + \varepsilon\boldsymbol{q}) > F(\boldsymbol{q}, (1 - \varepsilon)\boldsymbol{p}^* + \varepsilon\boldsymbol{q}) \tag{3}$$

for all $\boldsymbol{q} \neq \boldsymbol{p}^*$ and all sufficiently small ε, say for all $\varepsilon < \boldsymbol{E}$. In the literature, some equivalent and some alternative definitions of an ESS can be found (see, e.g., Lessard 1990). I prefer the present form since it best reflects the basic intuition behind the ESS concept.

Let me close this section with a few remarks. First, every ESS is a Nash strategy (let ε approach zero in (3)). Hence, an ESS prescribes "quasi-rational" behaviour even though no rationality assumptions are made in evolutionary game theory (see, e.g., Parker and Hammerstein 1985). On the other hand, the ESS condition is much more stringent than the Nash condition. While every evolutionary game has at least one Nash strategy, many games do not have an ESS. Second, all considerations on frequency dependent selection also apply to frequency independent selection. In the latter case, the fitness of a strategy $\boldsymbol{q}$ does not depend on the population strategy $\bar{\boldsymbol{p}}$, i.e. $\bar{\boldsymbol{p}}$ appears as a dummy variable in the fitness function $F(\boldsymbol{q},\bar{\boldsymbol{p}})$. Now, a Nash strategy and an ESS correspond to a weak and a strict maximum of the fitness function, respectively. Third, the term evolutionary "stability" is quite misleading. The ESS concept is not a stability concept but a (game theoretical) optimality concept which is based on fitness comparisons between strategies. Simple example show that an ESS is not necessarily stable with respect to the dynamics of natural selection (e.g., Weissing 1991; Cressman 1992). The notion of "stability against invading mutants" may be intuitively appealing but it has no clearcut dynamical justification. Stability concepts always have to refer to an underlying dynamics, and dynamical considerations are lacking in the phenotypic approach to natural selection.

3 Viability selection in a multilocus context

Natural selection takes place whenever there are *phenotypic* differences in mortality and reproduction. But selection per se will have no evolutionary consequences unless the phenotypic differences have a *genetic* basis. Unfortunately, the genetic basis of most evolutionarily "interesting" traits is completely unknown. This ignorance forms the major justification for the phenotypic approach to natural selection which avoids all genetic assumptions and instead focuses on interactions at the phenotypic level. In contrast, genetic selection models directly address the dynamics of genotype frequency change. A genetic model is inherently more complex than a corresponding phenotypic model. In fact, genotype frequencies are not only affected by selection but also by other factors such as the mating structure or recombination patterns.

Consider a dipolid Mendelian population where the genetic trait in question is affected by a number of autosomal loci A, B, $C, \ldots$. In the present paper, we shall focus on one of these loci, say A. The alleles at this locus will be indicated by indices such as i and j. The genetic constitution at the other loci will be indicated by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. If in total m loci are involved, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are multi-indices involving $m-1$ components. A diploid genotype will be characterized by a pair $(ij, \boldsymbol{\alpha\beta})$ with an obvious interpretation. In the three-locus context, for example, the pair $(13, \boldsymbol{\alpha\beta})$ with $\boldsymbol{\alpha} = (2, 2)$ and $\boldsymbol{\beta} = (4, 3)$ corresponds to the genotype $A_1B_2C_2 \mid A_3B_4C_3$.

Let $X_{ij,\,\boldsymbol{\alpha\beta}}$ denote the frequency of genotype $(ij, \boldsymbol{\alpha\beta})$ at the zygote stage and let $w_{ij,\,\boldsymbol{\alpha\beta}}$ denote the fitness (= viability) of this genotype. As usual, it will be assumed that the fitness parameters are *position independent*, i.e. that $w_{ij,\,\boldsymbol{\alpha\beta}}$ does not depend on the order of the components in $(ij, \boldsymbol{\alpha\beta})$. In particular, this implies $w_{ij,\,\boldsymbol{\alpha\beta}} = w_{ji,\,\boldsymbol{\alpha\beta}}$. The $X_{ij,\,\boldsymbol{\alpha\beta}}$ should be understood as *ordered* frequencies, i.e. we distinguish between the frequencies of $(ij, \boldsymbol{\alpha\beta})$ and $(ji, \boldsymbol{\alpha\beta})$ and assume $X_{ij,\,\boldsymbol{\alpha\beta}} = X_{ji,\,\boldsymbol{\alpha\beta}}$. Viability selection leads to a shift of genotype frequencies from $X_{ij,\,\boldsymbol{\alpha\beta}}$ at the zygote stage to

$$\tilde{X}_{ij,\,\boldsymbol{\alpha\beta}} = X_{ij,\,\boldsymbol{\alpha\beta}} \cdot \frac{w_{ij,\,\boldsymbol{\alpha\beta}}}{\bar{w}} \tag{4}$$

at the adult stage. Here, $\bar{w}$ denotes the mean fitness (= mean viability) of the population:

$$\bar{w} = \sum_{ij} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} X_{ij,\,\boldsymbol{\alpha\beta}} w_{ij,\,\boldsymbol{\alpha\beta}} \,. \tag{5}$$

After viability selection, mating and reproduction takes place which involves the reshuffling of *gametic types* $(i, \boldsymbol{\alpha})$ due to recombination. Let

$$X_{i\boldsymbol{\alpha}} = \sum_{j} \sum_{\boldsymbol{\beta}} X_{ij,\,\boldsymbol{\alpha\beta}} \tag{6}$$

denote the frequency of gametic type (i, α) in the parent generation and let $X'_{i\boldsymbol{\alpha}}$ denote the corresponding frequency in the offspring generation (both

evaluated at zygote stage). In the two-locus case, $X'_{i\alpha}$ is given by

$$X'_{i\alpha} = \sum_j \sum_\beta [(1-r)\tilde{X}_{ij,\,\alpha\beta} + r\tilde{X}_{ji,\,\alpha\beta}] \,, \tag{7}$$

where we have assumed random mating, no fertility differences between matings, Mendelian segregation and recombination between A and B with frequency r (see, e.g., Eshel and Feldman 1984). In the multi-locus context, the gamete type dynamics looks more complicated since many recombination patterns are possible (see, e.g., Liberman 1988; Matessi and Di Pasquale 1996). Fortunately, we may, for our purposes, get rid of the details of the genetic reshuffling process. In fact, we shall concentrate on the *marginal allele frequencies* at the focus locus A:

$$x_i = \sum_\alpha X_{i\alpha} = \sum_j \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} \,. \tag{8}$$

In the two-locus context, we obtain

$$x'_i = \sum_\alpha X'_{i\alpha} = (1-r)\cdot \sum_j \sum_{\alpha,\beta} \tilde{X}_{ij,\,\alpha\beta} + r\cdot \sum_j \sum_{\alpha,\beta} \tilde{X}_{ji,\,\alpha\beta} = \sum_j \sum_{\alpha,\beta} \tilde{X}_{ij,\,\alpha\beta} = \tilde{x}_i \,.$$

In other words, the marginal allele frequency at the zygote stage of the offspring generation, x'_i, is identical with the marginal allele frequency $\tilde{x}_i$ of the parent generation after selection. This example illustrates a general principle. The genetic reshuffling processes of Mendelian reproduction lead to a reassociation of alleles within and between loci, but the allele frequencies are not affected by these processes. This principle holds under a wide range of circumstances and it forms the basis of the analysis to follow. In particular, it extends to the multilocus context if we assume random mating, Mendelian inheritance and the absence of fertility selection. Even most systems of nonrandom mating, like inbreeding, do not lead to deviations from this principle (see, e.g., Crow and Kimura 1970). Hence, we get quite in general:

$$x'_i = \tilde{x}_i = \sum_j \sum_{\alpha,\beta} \tilde{X}_{ij,\,\alpha\beta} = \frac{1}{\bar{w}} \cdot \sum_j \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} w_{ij,\,\alpha\beta} \,. \tag{9}$$

Summarizing, we have derived

Result 1. *Under random mating, Mendelian inheritance and selection acting only on viability differences, the dynamics of allele frequency change at a specific locus A is given by the recurrence equation*

$$x'_i = x_i \cdot \frac{w_i}{\bar{w}} \,. \tag{10}$$

In this equation, the parameters w_i and $\bar{w}$ denote the marginal fitness of allele i:

$$w_i = \frac{1}{x_i} \cdot \sum_j \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} w_{ij,\,\alpha\beta} \tag{11}$$

and the mean fitness of the population:

$$\bar{w} = \sum_{ij} \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} w_{ij,\,\alpha\beta} = \sum_{i} x_i w_i \,. \tag{12}$$

In particular, allele i will increase in frequency if and only if its marginal fitness is larger than the mean fitness of the population:

$$x_i' > x_i \Leftrightarrow w_i > \bar{w} \,. \tag{13}$$

Equation (10) closely resembles the discrete replicator dynamics for selection in an asexual population (e.g. Weissing 1991) and the Haldane-Fisher-Wright model for selection at a single autosomal locus (e.g. Crow and Kimura 1970). Notice, however, that in contrast to these other models the (marginal) allele frequencies per se do not fully specify the allele frequency dynamics. In fact, the genotype frequencies, $X_{ij,\,\alpha\beta}$ are required for a full specification of eqn (10). Nevertheless, relation (13) remains crucial for judging the invasion chances of a rare mutant.

4 Maynard Smith meets Mendel: the multilocus game dynamics

Let us now integrate the two approaches outlined above. Consider an evolutionary game with a fixed number of pure strategies and a fixed fitness function $F(\boldsymbol{q}, \bar{\boldsymbol{p}})$. To add genetics to this game, we assume that the mixed strategy of an individual is the expression of its genotype at m autosomal loci. In other words, the genotype $(ij, \boldsymbol{\alpha\beta})$ corresponds to a mixed strategy $\boldsymbol{Q}_{ij,\,\alpha\beta}$, where we assume that $\boldsymbol{Q}_{ij,\,\alpha\beta} = \boldsymbol{Q}_{ji,\,\alpha\beta}$. The population strategy is then given by

$$\bar{\boldsymbol{p}} = \sum_{ij} \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} \boldsymbol{Q}_{ij,\,\alpha\beta} \,. \tag{14}$$

If we assume that individual fitness (= viability) is determined by the outcome of the underlying game, the fitness of genotype $(ij, \boldsymbol{\alpha\beta})$ is given by

$$w_{ij,\,\alpha\beta} = F(\boldsymbol{Q}_{ij,\,\alpha\beta}, \bar{\boldsymbol{p}}) \,. \tag{15}$$

Some authors interpret payoffs $F(\boldsymbol{Q}_{ij,\,\alpha\beta}, \bar{\boldsymbol{p}})$ in terms of *changes* in fitness. According to this interpretation, total fitness is given by $w_{ij,\,\alpha\beta} = F(\boldsymbol{Q}_{ij,\,\alpha\beta}, \bar{\boldsymbol{p}}) + w_0$, where the "basis fitness" w_0 is a positive constant. All results in this paper remain true if we switch to this interpretation.

We can now define the *marginal strategy* $\boldsymbol{q}_i$ of allele i:

$$\boldsymbol{q}_i = \frac{1}{x_i} \cdot \sum_{j} \sum_{\alpha,\beta} X_{ij,\,\alpha\beta} \boldsymbol{Q}_{ij,\,\alpha\beta} \,. \tag{16}$$

This is the mean strategy of all individuals harbouring allele i at locus A. Since the fitness function is linear in its first component, the marginal fitness of allele i is given by

$$w_i = F(\boldsymbol{q}_i, \bar{\boldsymbol{p}}) \,, \tag{17}$$

while the mean fitness can be written as

$$\bar{w} = F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) \; . \tag{18}$$

Hence, in the present context, Result 1 can be formulated as

Result 2. *Allele i will increase in frequency if and only if its marginal strategy yields a higher payoff than the population strategy*:

$$x_i' > x_i \Leftrightarrow F(\boldsymbol{q}_i, \bar{\boldsymbol{p}}) > F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) \; . \tag{19}$$

This result shows that dynamic properties of the allele frequency dynamics can be derived from the fitness function of the underlying game. In this respect, Result 2 provides a crucial link between the phenotypic and the genetic approach to natural selection.

5 Complete external stability and the Nash property

Genetic selection models have the useful property that the stability of an equilibrium can be judged on the basis of two complementary stability concepts. *Internal stability* refers to stability with respect to perturbations of the equilibrium which only involve gametic types already present at equilibrium. In contrast, *external stability* considers the stability of an equilibrium with respect to "new" types that are not represented in equilibrium with positive frequency. In non-degenerate cases, an equilibrium is stable if and only if it is internally stable *and* externally stable.

Let us first focus on the external stability of an equilibrium. We shall only consider stability against *single* mutants, i.e. against gametic types that differ in one allele at one locus, say locus A, from the gametic types of the resident population (for stability against multiple mutants see Matessi and Di Pasquale 1996). Assume that, at equilibrium, $n-1$ alleles are present at locus A and that a new allele, n, enters the population with low frequency. The equilibrium is (neutrally) *externally stable against n* if, for each $\varepsilon > 0$ and a small enough starting frequency (depending on ε), the marginal frequency of n, x_n, remains smaller than ε. In view of Result 2, the question of external stability against allele n boils down to a fitness comparison between the marginal strategy of the mutant allele, $\boldsymbol{q}_n$, and the population strategy at equilibrium, $\bar{\boldsymbol{p}} = \boldsymbol{p}^*$. This is, however, not completely straightforward: the marginal strategy of the mutant allele is not constant but dependent on the genetic background of this allele which changes over time ($\boldsymbol{q}_n$ depends on the genotype frequencies $X_{nj,\,\alpha\beta}$; see (16)). Eshel and Feldman (1984) and Liberman (1988) showed that this problem can be overcome: The linearisation of the selection dynamics at equilibrium has a simple, dominant eigenvalue. The normalized eigenvector corresponding to this eigenvalue yields a mixed strategy $\hat{\boldsymbol{q}}_n$ which may be viewed as the *characteristic strategy* of the mutant allele near equilibrium. In fact, the invasion chances of a mutant allele can be judged on basis of its characteristic strategy:

Result 3 (Liberman 1988). *Consider an equilibrium of the multilocus game dynamics with population strategy* $\boldsymbol{p}^*$. *A newly arising allele n at locus A will successfully invade and thereby destabilize the resident population if* $F(\hat{\boldsymbol{q}}_n, \boldsymbol{p}^*) > F(\boldsymbol{p}^*, \boldsymbol{p}^*)$, *i.e. if the characteristic strategy* $\hat{\boldsymbol{q}}_n$ *of the mutant allele is superior in fitness at equilibrium. In contrast, the resident equilibrium is externally stable against n if* $\hat{\boldsymbol{q}}_n$ *yields a lower fitness than the resident strategy* $\boldsymbol{p}^*$, *i.e. if* $F(\hat{\boldsymbol{q}}_n, \boldsymbol{p}^*) < F(\boldsymbol{p}^*, \boldsymbol{p}^*)$.

In principle, Result 3 allows to assess the external stability of an equilibrium against a specific mutant allele. The practical applicability of this result is, however limited by the fact that the fitness comparison between $\hat{\boldsymbol{q}}_n$ and $\boldsymbol{p}^*$ is not a purely phenotypic criterion. In fact, the characteristic strategy $\hat{\boldsymbol{q}}_n$ of the mutant allele comprises all kinds of *genetic constraints* acting upon the mutant. Detailed genetic information is required to calculate $\hat{\boldsymbol{q}}_n$ since it strongly depends on the association of allele n with the genetic background provided by alleles at the same locus and at other loci. Accordingly, there is little hope to arrive at a purely phenotypic characterization of external stability if we follow the "standard" approach of population genetics and focus on the invasion chances of a *specific* mutant allele. However, in evolutionary time, an equilibrium is challenged repeatedly by a whole range of alternative alleles. Among others, IIan Eshel, Carlo Matessi, and Peter Hammerstein (see their contributions in this issue) have championed the view that evolutionary theory should focus less on the invasion chances of a specific mutation but rather on external stability against *any conceivable* mutant allele. Let us call an equilibrium *completely externally stable* if it is (neutrally) externally stable against any feasible single mutant n, i.e. against any conceivable combination of strategies $\boldsymbol{Q}_{nj,\,\alpha\beta}$. Complete external stability is a restrictive stability concept that does not leave room for equilibria which are only externally stable due to specific genetic constraints. As a consequence, we obtain at least a partial phenotypic characterization of complete external stability.

Result 4. *The population strategy* $\boldsymbol{p}^*$ *of a completely externally stable equilibrium is a Nash strategy.*

Proof. The proof of this important result is almost trivial. We have to show that an equilibrium can be destabilized by the invasion of a mutant allele if its population strategy $\boldsymbol{p}^*$ is not a Nash strategy. If $\boldsymbol{p}^*$ is not a Nash strategy, an alternative strategy $\boldsymbol{q}$ with a higher fitness can be found: $F(\boldsymbol{q}, \boldsymbol{p}^*) > F(\boldsymbol{p}^*, \boldsymbol{p}^*)$. Consider a mutant allele n at locus A which is "dominant" in the sense that it induces strategy $\boldsymbol{q}$ irrespective of its genetic background (i.e. $\boldsymbol{Q}_{nj,\,\alpha\beta} = \boldsymbol{q}$ for all j, α, and β). The marginal strategy of n is, of course, again $\boldsymbol{q}$: $\boldsymbol{q}_n = \boldsymbol{q}$. This implies $F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) > F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}})$ in a neighbourhood of $\boldsymbol{p}^*$. In view of Result 2, allele n will increase in frequency until this neighbourhood of $\boldsymbol{p}^*$ is left. Hence, the equilibrium is externally unstable against n. □

Let me close this section with a few remarks. First, Result 4 reflects the fact that complete external stability requires stability against a huge set of alternatives. All kinds of mutations (including dominant ones with an adequate

strategy) are considered feasible, and the role of genetic constraints is marginalized. It depends on the biological context whether this is indeed a reasonable assumption. Second, Result 4 provides a necessary condition for complete external stability which is by far not sufficient. It is easy to construct examples of a Nash strategy $\boldsymbol{p}^*$ where the underlying equilibrium can be invaded by *any* mutant with a different strategy (e.g. Weissing 1991). Third, consider an equilibrium whose population strategy $\boldsymbol{p}^*$ is not a Nash strategy. We have shown that this equilibrium can be invaded by any mutant coding for a "better response" strategy $\boldsymbol{q}$. The set of "better response" strategies is convex, it contains at least one pure strategy, and it contains strategies from each neighbourhood of $\boldsymbol{p}^*$ (e.g. Weissing 1991). As a consequence, the equilibrium can be destabilized by mutants whose strategy deviates only slightly from the population strategy. Fourth, the fact that a population can be successfully invaded if $\boldsymbol{p}^*$ is not a Nash strategy does *not* imply that a successful mutant will shift the population closer to a Nash strategy or an ESS. If $\boldsymbol{p}^*$ is already close to an ESS, only those mutants can invade whose marginal strategy is "in the approximate direction" of the ESS (Eshel and Feldman 1984). Unfortunately, this does not guarantee that the population strategy after a successful invasion is closer to the ESS than the original one. In fact, there is no clearcut relationship between the properties of successful mutants and the properties of a Nash strategy or an ESS. It is, for instance, easy to construct an evolutionary game with a unique ESS where a resident population can be destablized by a variety of mutants, but not by mutants representing the ESS strategy.

6 Phenotypic attractivity and long-term stability

According to the paradigm of long-term evolution, an evolutionary equilibrium can only be considered long-term stable if it combines the properties of internal stability (= short-term stability) and complete external stability (= stability against the ongoing influx of new mutant strategies). One might be inclined to use the term "stability" in the sense of "asymptotic stability" (e.g. Hofbauer and Sigmund 1988), i.e. to require that the selection dynamics drives the system back to equilibrium after any sufficiently small perturbation. However, this requirement is too strict in the context of evolutionary game theory. In fact, an equilibrium with a mixed population strategy $\bar{\boldsymbol{p}}$ can *never* be asymptotically stable against all kinds of perturbations. Consider such an equilibrium and an invasion attempt by a rare mutant allele n that induces the same marginal strategy ($\boldsymbol{q}_n = \bar{\boldsymbol{p}}$). In view of Result 2, this allele is not selected against and a new equilibrium is attained that incorporates the mutant allele. Hence, at the genetic level the population is not driven back to its original state. At the phenotypic level, however, nothing has changed since the population strategy has remained constant. More generally, each genetic equilibrium forms part of a manifold of equilibria which all induce the same population strategy. All one can hope for is that this manifold is attractive, i.e. that, after

any small (internal or external) perturbation, the population strategy of the perturbed population is driven back to the population strategy of the resident population. Let us call a genetic equilibrium *phenotypically attractive* (Weissing 1983) if it has this property. An equilibrium will be called *long-term stable* if it is (1) internally stable, (2) completely externally stable, and (3) phenotypically attractive.

We have already seen (Result 4) that the population strategy of a long-term stable equilibrium is a Nash strategy. For the special case of a single gene locus and two pure strategies, Eshel (1982) and Lessard (1984) have shown that the genetic concept of long-term stability is closely related to the phenotypic concept of an ESS. In our terminology, their main conclusion can be framed as

Result 5 (Lessard 1984). *In the context of a single locus and two pure strategies, a genetic equilibrium is long-term stable if and only if its population strategy is an ESS.*

Result 5 shows that at least for a restricted class of models a full phenotypic characterization of long-term stability can be achieved. Therefore the seminal papers of Eshel (1982) and Lessard (1984) may be viewed as the starting point of the theory of long-term evolution. One should realize, however, that the proof of Result 5 rests on the one-dimensionality of the strategy space and the fact that one-locus systems are "well-behaved" when compared to multi-locus systems (they are "locally adaptive", i.e. Fisher's Fundamental Theorem of Natural Selection holds for these systems in the context of frequency-independent selection). Accordingly, it would be premature to conclude that a full phenotypic characterization of long-term stability is within reach.

In fact, I have little hope that it will ever be possible to arrive at a phenotypic characterization of internal stability (the "short-term component" of long-term stability). This is exemplified by a simple observation: Phenotypic optimality concepts like the Nash equilibrium concept or the ESS concept are not affected if the fitness function is changed by adding a constant. On the other hand, adding a constant to the fitness function changes the strength of selection and – in the context of three or more pure strategies – such a change may affect the internal stability properties of an equilibrium (see Weissing 1991 for a class of examples). Hence, phenotypic concepts will probably never be fully congruent with internal stability and we have to be content with the characterization of long-term stability for those equilibrium points where internal stability does not present a problem. Such equilibria will be considered in the following section.

7 Long-term stability of monomorphic populations

The question of internal stability does not arise for a *genetic monomorphism*, i.e. for a genetic state where only a single allele is present at each locus. Internal stability is also automatically obtained for a *phenotypic*

monomorphism, i.e. for a (polymorphic) genetic state where all resident genotypes induce the same strategy ($\boldsymbol{Q}_{ij,\,\alpha\beta} = \bar{\boldsymbol{p}}$ for all $i, j, \boldsymbol{\alpha}$ and $\boldsymbol{\beta}$; Hammerstein and Selten 1994). At a phenotypic monomorphism, all frequency distributions of resident gametic types are *phenotypically equivalent* (Weissing 1983) since they induce the same population strategy. As a consequence, all "internal" perturbations of a phenotypic monomorphism are selectively neutral (see Result 2), and the monomorphism is (neutrally) internally stable.

In this section, we shall investigate to what extent the phenotypic concept of evolutionary stability is able to characterize those monomorphic states which are long-term stable, i.e. which are completely externally stable and phenotypically attractive. We have already seen (Result 4) that the population strategy of a completely externally stable equilibrium is a Nash strategy. For a (phenotypic) monomorphism, this result is complemented by

Result 6. *The population strategy* $\boldsymbol{p}^*$ *of a phenotypically attractive monomorphism is an ESS.*

Proof. Recall that an ESS is characterized by the family of inequalities (3). The linearity of the fitness function in its first component implies that $\boldsymbol{p}^*$ is an ESS if and only if

$$F(\boldsymbol{p}_\varepsilon, \boldsymbol{p}_\varepsilon) - F(\boldsymbol{q}_n, \boldsymbol{p}_\varepsilon) > 0 \tag{20}$$

for $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ and all $0 < \varepsilon < \boldsymbol{E}$, where the strategies $\boldsymbol{p}_\varepsilon = \boldsymbol{p}_\varepsilon(\boldsymbol{q}_n)$ are of the form

$$\boldsymbol{p}_\varepsilon = (1 - \varepsilon)\boldsymbol{p}^* + \varepsilon \boldsymbol{q}_n \,. \tag{21}$$

Consider now a phenotypically monomorphic equilibrium where all resident genotypes induce the strategy $\boldsymbol{p}^*$: $\boldsymbol{Q}_{ij,\,\alpha\beta} = \boldsymbol{p}^*$ for all $i, j, \boldsymbol{\alpha}$, and $\boldsymbol{\beta}$. Suppose that $\boldsymbol{p}^*$ is not an ESS, i.e. suppose that an alternative strategy $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ exists such that $F(\boldsymbol{q}_n, \boldsymbol{p}_\varepsilon) \geqq F(\boldsymbol{p}_\varepsilon, \boldsymbol{p}_\varepsilon)$ for small values of ε. Consider a dominant mutant allele n at locus A which induces this strategy irrespecitve of its genetic background ($\boldsymbol{Q}_{nj,\,\alpha\beta} = \boldsymbol{q}_n$ for all $j, \boldsymbol{\alpha}$, and $\boldsymbol{\beta}$). If this allele enters the population with frequency x_n, the resulting population strategy will be of the form $\bar{\boldsymbol{p}} = \boldsymbol{p}_\varepsilon$, where $\varepsilon < 2x_n$. Hence $F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) \geqq F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}})$. Result 2 implies that the mutant allele will not be selected against. Accordingly, the population strategy will not converge to $\boldsymbol{p}^*$, the population strategy of the resident population. In other words, the resident population is not phenotypically attractive. □

Let us now turn to the question whether evolutionary stability of the population strategy $\boldsymbol{p}^*$ is not only a necessary requirement but also sufficient to ensure the long-term stability of a monomorphic equilibrium. Consider therefore a phenotypic monomorphism with an evolutionarily stable population strategy $\boldsymbol{p}^*$: $\boldsymbol{Q}_{ij,\,\alpha\beta} = \boldsymbol{p}^*$ for all $i, j, \boldsymbol{\alpha}$, and $\boldsymbol{\beta}$. We want to show that any rare mutant allele n will be selected against. In view of Result 2, if is sufficient to show that

$$F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) - F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) > 0 \tag{22}$$

for all mutant strategies $\boldsymbol{q}_n \neq \boldsymbol{p}^*$. Here $\bar{\boldsymbol{p}}$ denotes the mean strategy of the population which results after introducing allele n with a small frequency x_n.

Notice that inequality (20) would immediately yield (22) if the population strategy of the perturbed population were of the form $\bar{\boldsymbol{p}} = \boldsymbol{p}_\varepsilon$. However, $\bar{\boldsymbol{p}}$ is typically not a convex combination of $\boldsymbol{p}^*$ and $\boldsymbol{q}_n$. In fact, $\bar{\boldsymbol{p}}$ is given by

$$\bar{\boldsymbol{p}} = \sum_{ij<n} \sum_{\alpha,\beta} X_{ij,\alpha\beta}\, \boldsymbol{p}^* + 2 \sum_{j \leqq n} \sum_{\alpha,\beta} X_{nj,\alpha\beta} \boldsymbol{Q}_{nj,\alpha\beta} - \sum_{\alpha,\beta} X_{nn,\alpha\beta} \boldsymbol{Q}_{nn,\alpha\beta} \,.$$

This can be rewritten as

$$\bar{\boldsymbol{p}} = (1 - 2x_n + x_{nn})\boldsymbol{p}^* + 2x_n \boldsymbol{q}_n - x_{nn} \boldsymbol{q}_{nn} \,,$$

where x_{nn} and $\boldsymbol{q}_{nn}$ denote the marginal frequency and marginal strategy of mutant homozygotes:

$$x_{nn} = \sum_{\alpha,\beta} X_{nn,\alpha\beta} \quad \text{and} \quad \boldsymbol{q}_{nn} = \frac{\boldsymbol{1}}{x_{nn}} \sum_{\alpha,\beta} X_{nn,\alpha\beta} \boldsymbol{Q}_{nn,\alpha\beta} \,.$$

Hence, for $\varepsilon = 2x_n$, the difference between $\bar{\boldsymbol{p}}$ and $\boldsymbol{p}_\varepsilon$ is of order x_{nn}:

$$\bar{\boldsymbol{p}} = \boldsymbol{p}_\varepsilon + x_{nn}(\boldsymbol{p}^* - \boldsymbol{q}_{nn}) \,. \tag{23}$$

We shall henceforth assume that the frequency of mutant homozygotes, x_{nn}, is of order $(x_n)^2$, an assumption that is certainly satisfied in randomly mating populations but which also applies to other mating systems (but not, for example, to populations with inbreedings). Under this assumption, the difference between the left-hand sides of (20) and (22) is of order $o(x_n)$ and hence becomes negligible for small mutant frequencies:

Lemma 1.

$$F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) - F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) = F(\boldsymbol{p}_\varepsilon, \boldsymbol{p}_\varepsilon) - F(\boldsymbol{q}_n, \boldsymbol{p}_\varepsilon) + o(x_n), \quad \text{where } \varepsilon = 2x_n \,. \tag{24}$$

Proof. The first-order Taylor approximations of the "pure-strategy payoff" functions $F_i(\boldsymbol{p}) := F(\sigma_i, \boldsymbol{p})$ are given by

$$F_i(\boldsymbol{p}) = F_i(\boldsymbol{p}^*) - \sum_j \frac{\partial F_i}{\partial p_j}(\boldsymbol{p}^*)(p_j^* - p_j) + o(\|\boldsymbol{p} - \boldsymbol{p}^*\|) \,.$$

As a consequence,

$$\begin{aligned} F(\boldsymbol{p}, \boldsymbol{p}) - F(\boldsymbol{q}_n, \boldsymbol{p}) &= F(\boldsymbol{p}, \boldsymbol{p}^*) - F(\boldsymbol{q}_n, \boldsymbol{p}^*) - (\boldsymbol{p} - \boldsymbol{q}_n) \cdot A(\boldsymbol{p}^* - \boldsymbol{p}) \\ &\quad + o(\|\boldsymbol{p} - \boldsymbol{p}^*\|) \,, \end{aligned}$$

where A denotes the Jacobian matrix

$$A = \left(\frac{\partial F_i}{\partial p_j}(\boldsymbol{p}^*) \right)_{i,j} \,. \tag{25}$$

Inserting $\bar{\boldsymbol{p}}$ and $\boldsymbol{p}_\varepsilon$ for $\boldsymbol{p}$ and neglecting higher order terms, we obtain

$$\begin{aligned} F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) - F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) &= (1 - 2x_n)[F(\boldsymbol{p}^*, \boldsymbol{p}^*) - F(\boldsymbol{q}_n, \boldsymbol{p}^*)] \\ &\quad - 2x_n(\boldsymbol{p}^* - \boldsymbol{q}_n) \cdot A(\boldsymbol{p}^* - \boldsymbol{q}_n) + o(x_n) \,, \end{aligned} \tag{26}$$

and

$$F(\boldsymbol{p}_\varepsilon, \boldsymbol{p}_\varepsilon) - F(\boldsymbol{q}_n, \boldsymbol{p}_\varepsilon) = (1-\varepsilon)[F(\boldsymbol{p}^*, \boldsymbol{p}^*) - F(\boldsymbol{q}_n, \boldsymbol{p}^*)] - \varepsilon(\boldsymbol{p}^* - \boldsymbol{q}_n) \cdot A(\boldsymbol{p}^* - \boldsymbol{q}_n) + o(\varepsilon) \; . \tag{27}$$

This immediately implies (24). □

At an ESS, the left-hand side of (27) is positive for all $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ and small enough ε. Let us call an ESS *regular*, if already the linear approximation is positive for all $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ and small enough ε:

$$(1-\varepsilon)[F(\boldsymbol{p}^*, \boldsymbol{q}^*) - F(\boldsymbol{q}_n, \boldsymbol{p}^*)] - \varepsilon(\boldsymbol{p}^* - \boldsymbol{q}_n) \cdot A(\boldsymbol{p}^* - \boldsymbol{q}_n) > 0 \; . \tag{28}$$

Regularity is a mild requirement and I am not aware of a biological example involving a non-regular ESS. In case of a *linear* fitness function, the Jacobian A corresponds to the payoff matrix, and it is easy to see that (28) is always satisfied for an ESS $\boldsymbol{p}^*$. Hence, all ESSs of the linear games considered by Hammerstein and Selten (1994) are regular. More generally, an ESS is regular whenever the quadratic form induced by the matrix A is non-degenerate, i.e. if $\det(A + A^T) \neq 0$ (see, e.g., Weissing 1983). Actually, regularity may hold under even weaker conditions (Bomze and Pötscher 1989), but the determinant criterion is easy to check and it shows that regularity is a generic condition.

At a regular ESS, either $F(\boldsymbol{p}^*, \boldsymbol{p}^*) - F(\boldsymbol{q}_n, \boldsymbol{p}^*)$ is positive or $(\boldsymbol{p}^* - \boldsymbol{q}_n) \cdot A(\boldsymbol{p}^* - \boldsymbol{q}_n)$ is negative. Since both terms do not depend on the mutant frequency, the left-hand side of (26) will become positive for small enough x_n. Hence, we obtain

Lemma 2. *Consider a phenotypic monomorphism with population strategy $\boldsymbol{p}^*$. If $\boldsymbol{p}^*$ is a regular ESS, for each alternative strategy $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ an "invasion barrier" $\Delta(\boldsymbol{q}_n) > 0$ can be found such that all mutant alleles with marginal strategy $\boldsymbol{q}_n$ and marginal frequency $0 < x_n < \Delta(\boldsymbol{q}_n)$ have a smaller than average fitness*:

$$F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) < F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) \; . \tag{29}$$

The invasion barrier can be chosen in such a way that $\Delta(\boldsymbol{q}_n)$ depends continuously on $\boldsymbol{q}_n$.

In the context of linear frequency dependence, Hammerstein and Selten (1994) claim that Lemma 2 directly implies the external stability and phenotypic attractivity of a phenotypically monomorphic ESS population. This claim may be premature. Hammerstein and Selten's conclusion would be correct if the marginal strategy $\boldsymbol{q}_n$ of a mutant allele would not change in time. For a mutant allele with a *fixed* strategy $\boldsymbol{q}_n \neq \boldsymbol{p}^*$ and a small initial frequency $x_n < \Delta(\boldsymbol{q}_n)$, Lemma 2 does indeed imply that the mutant is indefinitely selected against and that its frequency will monotonically decline to zero. In general, however, the marginal strategy of a mutant allele changes in the course of selection and there is a priori no guarantee that $x_n(t) < \Delta(\boldsymbol{q}_n(t))$ for all times.

Hammerstein and Selten's claim could still be justified if a *uniform* invasion barrier $\Delta_{\min} > 0$ would exist, i.e. if the set of invasion barriers $[\Delta(\boldsymbol{q}_n) \mid \boldsymbol{q}_n \neq \boldsymbol{p}^*]$ would be bounded below by a positive number. However, even in the most simple scenarios, the invasion barrier of mutant strategies may become arbitrarily small:

Consider the linear fitness function $F(\boldsymbol{q}, \bar{\boldsymbol{p}}) = \boldsymbol{q} \cdot A\bar{\boldsymbol{p}}$ that is given by the payoff matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It is easy to see that $\boldsymbol{p}^* = (0.5, 0.5)$ is the unique ESS of this evolutionary game. Suppose that $\boldsymbol{p}^*$ is the population strategy of a genetic or a phenotypic monomorphism. Consider a mutant allele n which induces the pure strategy $\boldsymbol{q}_{nn} = (0, 1)$ in homozygous condition and the mixed strategy $\boldsymbol{q}_{\text{het}} = (0.5 + \lambda, 0.5 - \lambda)$ in heterozyous condition ($0 \leqq \lambda \leqq 0.5$). If we assume that the genotypes are in Hardy-Weinberg proportions, the marginal mutant strategy and the population strategy of the perturbed population are of the form

$$\boldsymbol{q}_n = (0.5 + \mu, 0.5 - \mu) \quad \text{where} \quad \mu = (1 - x_n)\lambda - 0.5x_n ,$$

$$\bar{\boldsymbol{p}} = (0.5 + \nu, 0.5 - \nu) \quad \text{where} \quad \nu = 2x_n(1 - x_n)\lambda - 0.5(x_n)^2 .$$

The fitness difference between mutant strategy and population strategy is given by

$$F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) - F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) = 2\nu(\nu - \mu) = -2x_n(1 - x_n)[2(1 - x_n)\lambda - 0.5x_n] \times [(1 - 2x_n)\lambda - 0.5x_n] .$$

Now it is easy to derive when the mutant allele has a selective advantage (see Fig. 1):

$$F(\boldsymbol{q}_n, \bar{\boldsymbol{p}}) > F(\bar{\boldsymbol{p}}, \bar{\boldsymbol{p}}) \Leftrightarrow \frac{2\lambda}{4\lambda + 1} < x_n < \frac{4\lambda}{4\lambda + 1} .$$

Accordingly the invasion barrier of the marginal mutant strategy is (for $\lambda > 0$) given by:

$$\Delta(\boldsymbol{q}_n) = \frac{2\lambda}{4\lambda + 1} . \tag{30}$$

Notice that the invasion barrier of a mutant strategy is negatively related to the parameter λ. In fact, the invasion barrier converges to zero when λ tends to zero, i.e. when $\boldsymbol{q}_{\text{het}}$ approaches the ESS. As a consequence, for *each* value $x_n > 0$ a mutant allele can be found which is not yet selected against at this frequency. In other words, there does not exist a uniform invasion barrier $\Delta_{\min} > 0$ which could be applied to all mutant strategies.

In view of Result 5, only a complete multi-locus specification of the above example could turn it into a counter-example to Hammerstein and Selten's

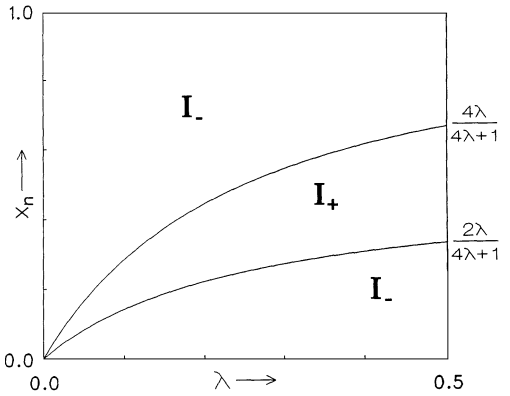

Fig. 1. Regions of selective advantage ($\boldsymbol{I}_+$) and selective disadvantage ($\boldsymbol{I}_-$) of a mutant allele that has reached the frequency x_n in a previously monomorphic ESS population. The mutant allele induces the strategies $\boldsymbol{q}_{\text{het}} = (0.5 + \lambda, 0.5 - \lambda)$ and $\boldsymbol{q}_{nn} = (0, 1)$ in heterozygous and homozygous condition, respectively

claim that a phenotypically monomorphic ESS population is long-term stable. However, the example illustrates that it may be difficult if not impossible to justify this claim. Consider a mutant allele whose heterozygote strategy $\boldsymbol{q}_{\text{het}}(\lambda)$ changes in the course of selection. Hence λ and the invansion barrier (30) change with time. As long as $x_n(t) < \Delta(t)$, the mutant frequency will decrease. But $\lambda(t)$ may also decrease in time, implying a decrease in $\Delta(t)$. If $\lambda(t)$ decreases at a faster rate than $x_n(t)$, the pair $(\lambda(t), x_n(t))$ might enter the region $\boldsymbol{I}_+$ in Fig. 1 where the mutant allele is no longer selected against. It is therefore conceivable that a monomorphic ESS population can be destabilized by the invasion of a single, rare mutant.

The problem outlined above is closely related to the nonexistence of a uniform invasion barrier. It only arises if the marginal strategy $\boldsymbol{q}_n(t)$ of the invading mutant approaches the resident ESS $\boldsymbol{p}^*$ in the course of selection. Suppose that the marginal strategies that a mutant allele may attain cannot approach the ESS $\boldsymbol{p}^*$. This is, for example, the case if $\boldsymbol{p}^*$ is not contained in the convex hull of the strategies $\boldsymbol{Q}_{nj,\alpha\beta}$. In such a case, a compact set C can be found that does not contain $\boldsymbol{p}^*$ but which contains all feasible mutant strategies $\boldsymbol{q}_n$. The continuous "invasion barrier function" of Lemma 2 attains its minimum

$$\Delta_C = \min\{\Delta(\boldsymbol{q}) \mid \boldsymbol{q} \in C\} > 0 \tag{31}$$

on C. Δ_C is a uniform invasion barrier for all strategies in C. Hence, for a sufficiently small starting frequency ($x_n(0) < \Delta_C$), the mutant allele will be selected against for all times (see (29)). As a consequence, the mutant frequency will monotonically decline to zero. In other words, a monomorphic ESS population is externally stable against all mutants whose marginal strategies are restricted to a compact set not containing the ESS $\boldsymbol{p}^*$.

A phenotypically monomorphic ESS population will therefore be externally stable against the majority of invasion attempts. A slight modification of the above argument shows that even more can be said for a *genetic* monomorphism:

Result 7. *A genetic monomorphism is long-term stable if its population strategy* $\boldsymbol{p}^*$ *is a regular ESS.*

Proof. At a genetic monomorphism, only one allele is present at each locus. Newly arising alleles at locus A do not affect the genetic background of this locus which remains constant. Hence all other loci may be neglected and, in essence, we are dealing with a one-locus situation. Assume that at the monomorphism allele A_1 is present at locus A (i.e. $\boldsymbol{Q}_{11} = \boldsymbol{p}^*$) and that the mutant allele A_2 enters the population with a small frequency x_2. The mutant allele induces the strategies $\boldsymbol{Q}_{12}$ and $\boldsymbol{Q}_{22}$ in heterozygote and in homozygote condition, respectively. We shall consider three cases:

(A) $\boldsymbol{Q}_{12} = \boldsymbol{Q}_{22} = \boldsymbol{p}^*$. In this case, the mutant allele induces the same strategy as the resident population and it will therefore neither increase nor decrease in frequency (Result 2). Hence the monomorphism is phenotypically attractive and (neutrally) externally stable against the mutant allele.

(B) $\boldsymbol{Q}_{12} = \boldsymbol{p}^*$, $\boldsymbol{Q}_{22} \neq \boldsymbol{p}^*$. In this case, $\bar{\boldsymbol{p}}$ can be put into the form (21), i.e. it is a convex combination of $\boldsymbol{p}^*$ and $\boldsymbol{q}_2$: $\bar{\boldsymbol{p}} = \boldsymbol{p}_\varepsilon$, where $\varepsilon = x_{22}/x_2$. Since x_{22} is of order $(x_2)^2$, the ESS property (20) immediately yields (22) for a sufficiently small mutant frequency x_2 (for $x_{22}/x_2 < E$). Hence the frequency of a rare mutant allele A_2 will monotonically decline to zero (Result 2). Accordingly the resident population is externally stable against A_2, and the population strategy converges to $\boldsymbol{p}^*$.

(C) $\boldsymbol{Q}_{12} \neq \boldsymbol{p}^*$. In this case, a compact neighbourhood C of $\boldsymbol{Q}_{12}$ can be found which does not contain $\boldsymbol{p}^*$. The marginal mutant strategy can be written as $\boldsymbol{q}_2 = \boldsymbol{Q}_{12} + (x_{22}/x_2)\cdot(\boldsymbol{Q}_{22} - \boldsymbol{Q}_{12})$ and it is therefore close to $\boldsymbol{Q}_{12}$ if x_2 is small. Hence, there exists a $\delta > 0$ such that $\boldsymbol{q}_2 \in C$ for $x_2 < \delta$. Suppose that the initial frequency of the mutant allele is smaller than δ and Δ_C, the uniform invasion barrier for the strategies in C. In view of (29), the mutant will have a smaller than average fitness, and the mutant frequency will monotonically decline to zero. Again, the monomorphism is externally stable, and the population strategy converges back to the ESS $\boldsymbol{p}^*$. □

Result 7 justifies Hammerstein and Selten's claim for the context of a genetic monomorphism. It is, however, important to realize the limitations of this result. Consider a genetically monomorphic ESS population that is challenged by the appearance of a rare mutant allele A_2. If $\boldsymbol{Q}_{12} \neq \boldsymbol{p}^*$ or $\boldsymbol{Q}_{22} \neq \boldsymbol{p}^*$, the mutant allele will be driven out of the population ($\lim_{t\to\infty} x_2(t) = 0$), provided that its initial frequency is small enough. But how small is "small enough"? We have seen that a uniform invasion barrier does not always exist. Accordingly, the degree of external stability may be "relative" in the sense that the invasion barrier depends on the marginal strategy of the mutant allele.

A second problem may be even more important. If $\boldsymbol{Q}_{12} = \boldsymbol{Q}_{22} = \boldsymbol{p}^*$, the mutant allele will not be selected against and stay in the population. Hence, the population is transformed from a *genetic* monomorphism into a *phenotypic* monomorphism, and it is not yet clear whether such a monomorphism is stable against further invasion attempts. Even if we suppose that a phenotypically monomorphic ESS population is always completely externally stable and phenotypically attractive, dynamic stability is still rather shaky. Suppose that $\boldsymbol{p}^*$ is not a pure strategy. For each value of x_n, we can construct a mutant allele n that is phenotypically equivalent to the resident population (i.e. $\boldsymbol{q}_n = \boldsymbol{p}^*$) but for which not all strategies $\boldsymbol{Q}_{nj,\alpha\beta}$ are identical to $\boldsymbol{p}^*$. Such a mutant will not be selected against, but it transforms the phenotypic monomorphism into a phenotypic polymorphism. It is well known (e.g. Weissing 1991) that *polymorphic* ESS populations may be dynamically unstable. Hence, the resulting polymorphic population might be destabilized by internal perturbation or further invasion attempts.

In conclusion, we have shown that the evolutionary stability of the population strategy guarantees the long-term stability of a genetic monomorphism. Whether the same holds true for a phenotypic monomorphism has still to be established. Long-term stability of a (genetic or phenotypic) monomorhism is, however, a rather weak property, since it is conceivable that a monomorphic ESS population can be destabilized by a series of successive invasion attempts.

Discussion

Evolution by natural selection is the result of an interplay between forces at the phenotypic and the genetic level. An integrated study of natural selection is, however, hampered by the fact that the genetical component of selection is typically unknown in practice. In reaction to this ignorance, phenotypic approaches towards natural selection tend to neglect the genetic constraints on adaptive evolution. Instead, these approaches focus on the complexity of interactions at the phenotypic level. The theory of long-term evolution (the "streetcar theory of natural selection", Hammerstein 1996) seems to provide a justification for such an approach. In contrast to population genetical models which concentrate on a fixed number of alleles at a fixed number of loci, the theory of long-term evolution considers a huge variety of mutants which are all viewed as potential invaders into a resident population. In this broader context, the analysis of external stability rests on the assumption that, in the long run, mutants should arise whose invasion chances are less affected by genetic factors than by their phenotypic effects. As a consequence, the *long-term* outcome of selection should rather be dominated by fitness properties than by genetic constraints. The two "streetcar theorems" of Hammerstein and Selten (1994) reflect this principle: only Nash strategies can be long-term stable, and the ESS property characterizes the long-term stability of monomorphic populations.

In the present paper, these theorems (Results 4, 6, and 7) are derived from rather general genetical considerations. The derivation is based on the idea that phenotypic and genetic processes are often fairly independent from one another. Within a generation, the allele frequency changes brought about by (viability) selection are governed by fitness differences (Li 1955; Turner 1970): $\tilde{x}_i = x_i w_i/\bar{w}$. In contrast, the transition from one generation to the next is governed by genetic factors such as Mendelian segregation or recombination. These factors lead to a reassociation between alleles but they usually do not affect the allele frequencies. Consequently, as long as we focus on the level of *allele* frequencies, the transition from the zygote stage of the parent generation to the zygote stage of the offspring generation is still governed by fitness differences (Results 1 and 2): $x_i' = \tilde{x}_i = x_i w_i/\bar{w}$. We may conclude that the two "streetcar theorems" hold true whenever (a) selection can be decoupled from reproduction, and (b) mating and reproduction do not lead to changes in allele frequencies. As a consequence, the theorems apply to a wide range of genetical contexts, including viability selection at multiple loci in a randomly mating population.

Hammerstein and Selten's proof makes use of the special properties of an ESS under *linear* frequency dependent selection. The present paper shows that linearity of the fitness function in its second component is not really required. This is an important generalisation since "playing the field" situations (Maynard Smith 1982) or games with non-random matching of opponents easily lead to this kind of non-linearity. Moreover, the more transparent arguments given here show that there is a gap in Hammerstein and Selten's proof of the second streetcar theorem. This gap has partially been closed by demonstrating that a *genetically* monomorphic ESS population is stable in a long-term perspective. Whether the same holds true for a *phenotypic* monomorphism remains to be shown.

If we neglect this technical problem, the two fundamental theorems of the "streetcar theory" hold under a broad range of circumstances. Nevertheless, it would be premature to conclude that the theory of long-term evolution does already rest on a firm foundation. Conceptually, the theory has opened new roads to the study of evolution. But I do not want to end this article without mentioning some serious drawbacks and limitations.

The monomorphism problem. The ESS concept is based on the heuristics that a monomorphic wildtype population should be immune against the invasion attempts of single mutants. It is therefore not too suprising that the most convincing result of the streetcar theory (Result 7) only applies to monomorphic populations. In my opinion, a similar characterization of long-term stability will hardly be possible for polymorphic equilibria, since phenotypic criteria are not well-suited to study internal stability (Weissing 1991). In case of asexual reproduction, one may argue that monomorphic equilibria have a slight but consistent advantage over phenotyically equivalent polymorhic equilibria (Weissing, in preparation). But I do not see a convincing reason why such a "trend towards monomorphism" should also be expected in Mendelian

populations. One might even argue that polymorphisms are inherently advantageous in a long-term perspective since they buffer a population against environmental fluctuations. In any case, polymorphisms abound in nature and one should not expect monomorphisms to play a dominant role in the face of a reproductive system whose main features are related to the generation and conservation of genetic variation.

The time scale problem. To apply the concept of long-term stability to the natural world, one has to assume that the situation found in nature does resemble an evolutionary equilibrium which has resisted invasion by a huge variety of mutations. This includes the assumption that the selective forces have been present and consistent long enough to allow all kinds of mutants to appear and to challenge the resident population. But how long is "long enough"? Consider the classical example of sickle cell anaemia, where human populations remain polymorphic at the β-globin locus due to overdominance (heterozygote advantage). This polymorphism is short-term stable due to genetic constraints but can hardly be considered long-term stable. In fact, the theory of long-term evolution would predict the occurrence of mutations which combine the advantages of the overdominant heterozygote in homozygous condition. To explain the discrepancy between theoretical prediction and empirical findings, one might argue that selection for malaria resistance is of recent origin, that the selective forces have been weak or inconsistent in the past, that the required "super-mutations" have not yet shown up, and so on. There is probably some truth in these arguments, and transient phenomena have certainly to be taken into account in a long-term perspective. Nevertheless, one should be sceptical towards these kinds of arguments since they provide an ad hoc explanation for *any* phenomenon which does not fit into the theory. The empirical relevance of the theory of long-term evolution will strongly depend on the development of criteria which make it possible to judge what "long-term" really means and to which phenomena the theory can be expected to apply.

The attainability problem. In the sickle cell anaemia example, the prevalence of the polymorphic overdominance equilibrium is often explained by the fact that no superior homozygotes are available. This is, however, not the whole story. Actually, homozygous individuals of a genotype CC have the highest fitness, where C is an alternative allele found in several human populations. Despite its fitness advantage in homozygous condition, this allele cannot invade the resident population since it provides a lower fitness in combination with the resident alleles (see e.g., Cavalli-Sforza and Bodmer 1971; Hartl and Clark 1989). This example illustrates the general principle that well adapted genotypes are not necessarily good invaders. Even more important is the opposite: successful invaders have a fitness advantage in the neighbourhood of a resident equilibrium, but there is no guarantee that they also provide a high fitness once this context is left. Consequently, one should not expect that the destabilization of an equilibrium by a superior invader will typically lead to

the establishment of a superior equilibrium. Even if we envisage long-term evolution as a succession of equilibria where each new equilibrium results from a successful invasion of the previous one, there is no reason to assume that the equilibria that are reached later in the series have a higher degree of external stability than their predecessors. In fact, the theory of long-term evolution says virtually nothing about how a long-term stable equilibrium can be reached and whether it will be reached at all.

The existence problem. The Nash property is a necessary condition for long-term stability, but it is far from being sufficient. In fact, many situations are conceivable where long-term stability is not feasible at all. In the terminology of the streetcar metaphor, a final stop does not exist in these cases. Consider, for instance, the situation where rare alleles have a systematic, structural advantage over common alleles. Examples for such a *minority advantage* are more common than one might expect. Apart from phenomena as (marginal) overdominance and segregation distortion (e.g. Feldman and Otto 1991), they include many of the "standard examples" for frequency dependent selection: negative assortative mating and rare male preferences (e.g. Partridge 1988); genetic incompatibility systems (e.g. Uyenoyama 1988); apostatic selection mediated by predators (e.g. Allen 1988); the arms race between hosts and infectious agents (e.g. May 1985). If rarity provides an advantage per se, the role of fitness differences is marginalized and the outcome of selection can be highly counterintuitive (Weissing et al., in preparation). The streetcar theory will hardly apply to these situations, since *all* equilibria are externally unstable, and a high degree of polymorphism should be expected. It is difficult to judge how common these situations are, but I have the impression that evolutionary biology underestimates their relevance.

Let me conclude with a personal evaluation of the streetcar theory which is necessarily highly subjective. The paradigm of long-term evolution has opened new ways to think about selection. It is therefore certainly an important contribution to evolutionary biology, if not a conceptual break-through. It is more difficult to judge the empirical relevance of the theory when it is applied to specific situations. I see no problem to apply the theory to phenotypic traits such as the morphology of locomotory or sensory organs, since these traits have probably been shaped by strong and consistent selection forces. Much more problematic is the application of the theory to traits which are of recent origin or subject to weak selection (e.g. many behavioral traits). And the theory of long-term evolution does certainly not apply to fluctuating selection (e.g. host-parasite coevolution) or to traits where minority per se provides a systematic advantage.

Acknowledgements. I thank Ilan Eshel and Carlo Matessi for introducing me to the problem, I am grateful to Peter Hammerstein and Reinhard Selten for many stimulating discussions, and above all I thank Odo Diekmann and Sabin Lessard for their constructive critique and helpful suggestions.

References

Akin, E. (1979): The Geometry of Population Genetics. Berlin: Springer-Verlag
Allen, J. A. (1988): Frequency-dependent selection by predators. Phil. Trans. R. Soc. Lond. B **319**: 485–503
Bomze, I. M. and B. M. Pötscher (1989): Game Theoretical Foundations of Evolutionary Stability. Lecture Notes in Economics and Mathematical Systems, Vol. 324. Berlin: Springer-Verlag
Cavalli-Sforza, L. L. and Bodmer, W. F. (1971): The Genetics of Human Populations. San Francisco: W. H. Freeman
Cressman, R. (1992): The Stability Concept of Evolutionary Game Theory. Berlin: Springer-Verlag
Crow, J. F. and Kimura, M. (1970): An Introduction to Population Genetics Theory. New York: Harper & Row
Eshel, I. (1982): Evolutionarily stable strategies and viability selection in Mendelian populations. Theor. Pop. Biol. **22**: 204–217
Eshel, I. (1996): On the changing concept of population stability as a reflection of a changing problematics in the quantitative theory of evolution. J. Math. Biol. **34**: 485–510
Eshel, I. and Feldman, M. W. (1984): Initial increase of new mutants and some continuity properties of ESS in two-locus systems. Amer. Nat. **124**: 631–640
Feldman, M. W. and Otto, S. P. (1991): A comparative approach to the population-genetics theory of segregation distortion. Amer. Nat. **137**: 443–456
Hammerstein, P. (1996): Darwinian adaptation, population genetics and the streetcar theory of evolution. J. Math. Biol. **34**: 511–532
Hammerstein, P. and Selten, R. (1994): Game theory and evolutionary biology. In: R. J. Aumann and S. Hart (Eds), Handbook of Game Theory with Economic Applications, Vol. II. Amsterdam: Elsevier, pp. 929–993
Hartl, D. L. and Clark, A. G. (1989): Principles of Population Genetics. Sunderland, Ms: Sinauer.
Hofbauer, J. and Sigmund, K. (1988): The Theory of Evolution and Dynamical Systems. Cambridge: Cambridge Univ. Press
Karlin, S. and Lessard, S. (1986): Theoretical Studies on Sex Ratio Evolution. Princeton, N. J.: Princeton Univ. Press
Lessard, S. (1984): Evolutionary dynamics in frequency-dependent two-phenotype models. Theor. Pop. Biol. **25**: 210–234
Lessard, S. (1990): Evolutionary stability: one concept, several meanings. Theor. Pop. Biol. **37**: 159–170
Li, C. C. (1955): Population Genetics. Chicago: Univ. of Chicago Press.
Liberman, U. (1988): External stability and ESS: criteria for initial increase of new mutant allele. J. Math. Biol. **26**: 477–485
Matessi, C. and Di Pasquale, C. (1996): Long-term evolution of multilocus traits. J. Math. Biol. **34**: 613–653
May, R. M. (1985): Host-parasite associations: their population biology and population genetics. In: D. Rollinson and R. M. Anderson (eds): Ecology and Genetics of Host-Parasite Interactions. London: Academic Press, pp. 243–262
Maynard Smith, J. (1982): Evolution and the Theory of Games. Cambridge: Cambridge Univ. Press
Moran, P. A. P. (1964): On the nonexistence of adaptive topographies. Ann. Hum. Genet. **27**: 383–393
Parker, G. A. and Hammerstein, P. (1985): Game theory and animal behavior. In: P. J. Greenwood, P. H. Harvey and M. Slatkin (eds): Evolution. Essays in Honour of John Maynard Smith. Cambridge: Cambridge Univ. Press, pp. 73–94
Parker, G. A. and Maynard Smith, J. (1990): Optimality theory in evolutionary biology. Nature, **348**: 27–33
Partridge, L. (1988): The rare-male effect: what is its evolutionary significance? Phil. Trans. R. Soc. Lond. B **319**: 525–539

Taylor, P. D. (1996): Inclusive fitness arguments in genetic models of behaviour. J. Math. Biol. **34**: 654–674

Turner, J. R. G. (1970): Changes in mean fitness under natural selection. In K. I. Kojima (ed.): Mathematical Topics in Population Genetics. Berlin: Springer-Verlag.

Uyenoyama, M. K. (1988): On the evolution of genetic incompatibility systems. In: R. E. Michod and B. R. Levin (eds): The Evolution of Sex. An Examination of Current Ideas. Sunderland, Ms: Sinauer, pp. 212–232

Weissing, F. (1983): Populationsgenetische Grundlagen der Evolutionären Spieltheorie. Bielefeld: Materialien zur Mathematisierung der Einzelwissenschaften, 41 and 42.

Weissing, F. (1991): Evolutionary stability and dynamic stability in a class of evolutionary normal form games. In: R. Selten (ed.): Game Equilibrium Models. I. Evolution and Game Dynamics. Berlin: Springer-Verlag, pp. 29–97.

.