

Mutation in evolutionary games can increase average fitness at equilibrium

Martin Willensdorfer^{a,b}, Martin A. Nowak^{a,c,*}

^aProgram for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA

^bDepartment of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA

^cDepartment of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Received 9 September 2004; received in revised form 20 April 2005; accepted 25 April 2005

Available online 24 June 2005

Abstract

We study game dynamical interactions between two strategies, A and B, and analyse whether the average fitness of the population at equilibrium can be increased by adding mutation from A to B. Classifying all two by two games with payoff matrix $[(a, b), (c, d)]$, we show that mutation from A to B enhances the average fitness of the whole population (i) if both a and d are less than $(b + c)/2$ and (ii) if c is less than b . Furthermore, we study conditions for maximizing the productivity of strategy A, and we analyse the effect of mutations in both directions. Depending on the biological system, a mutation in an evolutionary game can be interpreted as a genetic alteration, a cellular differentiation, a change in gene expression, an accidental or deliberate modification in cultural transmission, or a learning error. In a cultural context, our results indicate that the equilibrium payoff of the population can be increased if players sometimes choose the strategy with lower payoff. In a genetic context, we have shown that for frequency-dependent selection mutation can enhance the average fitness of the population at equilibrium.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Evolutionary game theory; Replicator dynamics; Frequency-dependent selection

1. Introduction

Normally, mutation is thought to reduce the average fitness of a population at equilibrium. A fraction of the population consists of the fittest genotype (or phenotype), but mutations constantly introduce less-fit mutants which lower the average fitness. This argument is based on the assumption of constant selection. In this paper, we will show that for frequency-dependent selection, which is the defining property of evolutionary game dynamics, mutation can increase the average fitness of a population at equilibrium.

The question what is the optimum mutation rate for a biological population has received much attention. Haldane (1937) has estimated the ‘mutational load’ in a population which is essentially the loss in productivity due to mutation. Kimura (1967) has argued that the optimum mutation rate is zero in a constant environment. The actual mutation rate will adapt to a minimum level which represents a balance between the effect of mutational load and the cost of further reducing the mutation rate. In a changing environment, however, a certain amount of mutations will be essential to stay at the same level of adaptation; the Red Queen has to run to remain in the same position. Other approaches for studying the optimum mutation rate start with the assumption that an organism is never perfectly adapted and there is always a fitter solution within reach. In this context, the question that needs to be solved is which

*Corresponding author. Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, MA 02138, USA.

E-mail address: martin_nowak@harvard.edu (M.A. Nowak).

mutation rate maximizes the probability of finding the fitter solution (Taddei et al., 1997; Sniegowski et al., 1997; Sasaki and Nowak, 2006).

In this paper, we will explore another approach of asking about optimum mutation rates by studying the effect of mutation in evolutionary games. Consider the interaction between two strategies, A and B, in an evolutionary game with payoff matrix

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{matrix} \quad (1)$$

The fitness of player A is given by $f_A = ax_A + bx_B$ and the fitness of player B is given by $f_B = cx_A + dx_B$. Denote by x_A and x_B the frequencies of players A and B in an (infinitely large) population. Let $x_A + x_B = 1$. The average fitness of the population is given by $\Phi = x_A f_A + x_B f_B$. Deterministic evolutionary dynamics can be described by the replicator equation (Taylor and Jonker, 1978; Hofbauer et al., 1979; Hofbauer and Sigmund, 1989, 1998; Fudenberg and Harris, 1992; Weibull, 1995; Fudenberg and Levine, 1998; Nowak and Sigmund, 2004)

$$\begin{aligned} \dot{x}_A &= (f_A - \Phi)x_A, \\ \dot{x}_B &= (f_B - \Phi)x_B. \end{aligned} \quad (2)$$

Note that $\dot{x}_A + \dot{x}_B = 0$. Thus the equation is one dimensional; dynamics are defined on the interval $x_A \in [0, 1]$.

If $a < c$ and $d < b$, there is a globally stable equilibrium in the interior given by $x_A^* = (d - b)/(a - b - c + d)$. If $a > c$ and $d > b$ the interior equilibrium, x_A^* , is unstable and two stable equilibria $x_A^* = 0$ and 1 exist. In this case both A and B are strict Nash equilibria (Nash, 1950, 1951) and evolutionary stable strategies (Maynard-Smith and Price, 1973; Maynard-Smith, 1982). If $a > c$ and $d < b$ then strategy A dominates strategy B. The globally stable equilibrium is given by $x_A^* = 1$. Strategy A is a strict Nash equilibrium. Conversely, if $a < c$ and $d > b$ then strategy B dominates strategy A. The globally stable equilibrium is given by $x_B^* = 1$, and strategy B is a strict Nash equilibrium. These are the four generic cases for interaction between two strategies A and B.

The replicator equation (with arbitrary fitness function) is a general description of frequency-dependent selection dynamics (assuming asexual reproduction). The replicator mutator equation (Hadeler, 1981; Stadler and Schuster, 1992; Bomze and Burger, 1995; Nowak et al., 2001, 2002) describes frequency-dependent selection and mutation and can be transformed into a Price equation (Page and Nowak, 2002) and adaptive dynamics (Nowak and Sigmund, 1990).

In this paper, we consider a replicator–mutator equation of two strategies A and B. We will study if mutation can increase the average fitness of the population. Intuitively, if the off-diagonal entries in the payoff matrix (1) are larger than the diagonal entries, then it might be possible for mutation from A to B or in both directions to increase the average fitness Φ at equilibrium.

Symbiosis, for example, is a widespread phenomenon in biology. In most cases it involves two organisms feeding from two different resources resulting in a mutual benefit. That is, individuals of type A increase the fitness of individuals of type B and vice versa. The fitness of both A and B is dependent on the composition of the symbiotic community and, therefore, frequency dependent. The replicator equation (2) can be used to describe such systems. It is known that symbiosis is only possible/stable if the cross-diagonal elements of the payoff matrix are larger than the diagonal ones ($a < c$ and $d < b$), that is, A benefits more from B than from A and vice versa.

In this work we have a similar scenario in mind. The major difference is that A is able to (accidentally or deliberately) produce B at rate u_A . That is, A produces cells that are specialized on performing a certain task that benefits A. Evolution has developed such systems at several occasions. Especially in multicellular structures, specialized cells are ubiquitous (e.g. cyanobacteria, sponges, fungi).

From a biological point of view the most interesting question is “when is such a cellular differentiation beneficial either for A alone or the whole population of A and B cells?” We will answer this question. The framework we present here can be used to discuss conditions that promote the evolution of cellular differentiation.

Replicator dynamics can also be interpreted to describe cultural (rather than genetic) transmission of strategies (see Hofbauer and Sigmund, 1998). In this context, strategy A is imitated or learned by other individuals proportional to its payoff. Surprisingly, for maximizing the average payoff of the population at equilibrium it may be best to include occasional or deliberate mistakes in this process of imitation or learning.

In general, we analyse the question of what is the optimum mutation rate for maximizing the performance of the whole population under frequency-dependent selection.

In Section 2, we provide a complete classification of all 2×2 games that allow a mutation from A to B to maximize the fitness $\Phi = x_A f_A + x_B f_B$ at equilibrium. In Section 3 we analyse mutations in both directions (from A to B and from B to A). In Section 4, we study mutations that maximize $x_A f_A$, the reproductive potential of A, at equilibrium. In Section 5 we summarize our results.

2. Mutation from A to B

Let $x = x_A = 1 - x_B$. The average fitness of Eq. (2) is given by

$$\Phi(x) = (a + d - b - c)x^2 + (b + c - 2d)x + d. \quad (3)$$

If $a + d - b - c \geq 0$, then $\Phi(x)$ is either linear or convex and assumes its maximum values for $x \in [0, 1]$ either at $x = 0$ or $x = 1$. We have $\Phi(0) = d$ and $\Phi(1) = a$. Let \hat{x} denote the frequency of A which maximizes $\Phi(x)$. If $a < d$, then $\hat{x} = 0$. If $a > d$, then $\hat{x} = 1$.

If $a + d - b - c < 0$, then $\Phi(x)$ has its maximum at $[(b + c)/2 - d]/(b + c - a - d)$. This value is in $(0, 1)$ if and only if $a, d < (b + c)/2$. If this condition is not satisfied, then $\hat{x} = 0$ for $a < d$ and $\hat{x} = 1$ for $a > d$, as before. We summarize

$$\hat{x} = \begin{cases} \frac{\frac{b+c}{2} - d}{b+c-a-d} & \text{if } \max(a, d) < \frac{b+c}{2}, \\ 0 & \text{if } a < d \text{ and } \max(a, d) \geq \frac{b+c}{2}, \\ 1 & \text{if } a > d \text{ and } \max(a, d) \geq \frac{b+c}{2}. \end{cases} \quad (4)$$

Note that \hat{x} maximizes the average fitness Φ , but is not necessarily an equilibrium of Eq. (2).

Now we assume that A mutates into B with probability u_A . Instead of Eq. (2), we now have

$$\begin{aligned} \dot{x}_A &= (1 - u_A)f_A x_A - \Phi x_A, \\ \dot{x}_B &= u_A f_A x_A + f_B x_B - \Phi x_B. \end{aligned} \quad (5)$$

The average fitness Φ and, therefore, the results summarized in Eq. (4) remain unchanged. In system (5), however, the inner equilibrium x^* depends on the actual payoff matrix as well as on u_A . Obviously, if $u_A = 0$ we get the replicator equation (2). If $u_A = 1$, then $x = 0$ is the only equilibrium. We note that if $b > c$ and $\hat{x} \in (0, 1)$, then a u_A can be found for which we have $x^* = \hat{x}$, i.e. the equilibrium of the system maximizes the average fitness. This mutation rate is given by

$$u_A = \frac{((b + c)/2 - a)((b - c)/2)}{((b + c)/2 - d)a + ((b + c)/2 - a)b} \quad (6)$$

and is always less than $\frac{1}{2}$.

In the following we will denote the optimum mutation rate, which maximizes the equilibrium fitness, by \hat{u}_A . Table 1 shows \hat{u}_A for the four different fitness landscapes (columns) and the four possible equilibrium structures of the replicator equation (rows). There are four possible fitness landscapes: (i) $\Phi(x)$ has its maximum in $(0, 1)$; (ii) $\Phi(x)$ is monotonically decreasing; (iii) $\Phi(x)$ is monotonically increasing; (iv) $\Phi(x)$ has its minimum in $(0, 1)$. There are four different equilibrium structures: (i) B dominates A, (ii) A dominates B, (iii) A and B

coexist, and (iv) A and B are bistable. In the following we will describe the results summarized in Table 1.

- **B dominates A:** Here $x = 0$ is the only stable equilibrium. Mutations from A to B have no effect.
- **A dominates B:** If $\hat{x} \in (0, 1)$ then the optimum u_A is given by Eq. (6). If $\hat{x} = 0$ then any u_A is optimum that exceeds $1 - d/b$. If $\hat{x} = 1$ then $u_A = 0$ is optimum.
- **A and B coexist:** If $\hat{x} \in (0, 1)$ and if $b > c$ then the optimum u_A is given by Eq. (6). If $\hat{x} \in (0, 1)$ and $b < c$ then $\hat{u}_A = 0$. If $\hat{x} = 0$ then any u_A is optimum that exceeds $1 - d/b$. If $\hat{x} = 1$, then $\hat{u}_A = 1$.
- **A and B are bistable:** The bistable equilibrium structure is incompatible with $\hat{x} \in (0, 1)$. For $\hat{x} = 1$ the optimum mutation rate is $\hat{u}_A = 0$. For $\hat{x} = 0$ any u_A is optimum that exceeds a critical value u_c . This critical value is given by the bifurcation point of the two inner equilibria. If $u_A = u_c$, only one inner equilibrium, x_c , exists; it is a saddle point. We have

$$u_c = \frac{a(a - b + d - c) + ad - bc - 2\sqrt{a(a - b + d - c)(ad - bc)}}{(b - a)^2} \quad (7)$$

and

$$x_c = \frac{(a - b + d - c)b - \sqrt{a(a - b + d - c)(ad - bc)}}{(a - b + d - c)(b - a)}. \quad (8)$$

If $u_A > u_c$, inner equilibria disappear and $x = 0$ is the only remaining equilibrium; it maximizes Φ . Hence any $u_A > u_c$ can be chosen as \hat{u}_A .

Thus, we have derived a complete classification of when cooperative mutation from A to B can increase the average fitness Φ .

3. Mutation in both directions

Let us now analyse mutation in both directions. Let u_A and u_B be the mutation rate of A and B. The differential equations for this system are given by Eq. (9).

$$\begin{aligned} \dot{x}_A &= u_B f_B x_B + ((1 - u_A)f_A - \Phi)x_A, \\ \dot{x}_B &= u_A f_A x_A + ((1 - u_B)f_B - \Phi)x_B. \end{aligned} \quad (9)$$

Again, we have $\Phi = x_A f_A + x_B f_B$ and can use the results from Eq. (4). In Appendix A, we show that if a particular pair $u_A, u_B > 0$ is able to maximize the average fitness, then a linear manifold of parameters exist that maximize the fitness as well. In particular, a pair of parameters exists with either $x_A = 0$ or $x_B = 0$, i.e. cooperative mutations occur only in one direction, either from A to B or from B to A.

Table 1
Can mutation from A to B increase the average fitness of the population?

	<p>$a, d < \frac{b+c}{2}$</p> <p>$\hat{x} \in (0, 1)$</p>	<p>$a < \frac{b+c}{2} < d$</p> <p>$\hat{x} = 0$</p>	<p>$d < \frac{b+c}{2} < a$</p> <p>$\hat{x} = 1$</p>	<p>$\frac{b+c}{2} < a, d$</p> <p>$\hat{x} = 0$ if $a < d$</p> <p>$\hat{x} = 1$ if $d < a$</p>	
<p>B dominates A</p> <p>$B \leftarrow A$</p> <p>$a < c, b < d$</p>	<p>cooperative mutations from A to B are not able to influence the outcome</p>				
<p>A dominates B</p> <p>$B \rightarrow A$</p> <p>$a > c, b > d$</p>	<p>u_A from Eq. (6)</p>	<p>Any $u_A > 1 - \frac{d}{b}$</p>	<p>$\hat{u}_A = 0$</p>	<p>Any $u_A > 1 - \frac{d}{b}$</p>	<p>$\hat{u}_A = 0$</p>
<p>A and B coexist</p> <p>$B \rightleftarrows A$</p> <p>$a < c, b > d$</p>	<p>If $b > c$ u_A from Eq. (6) else $\hat{u}_A = 0$</p>	<p>Any $u_A > 1 - \frac{d}{b}$</p>	<p>$\hat{u}_A = 0$</p>	<p>Not possible</p>	
<p>A and B are bistable</p> <p>$B \rightleftarrows A$</p> <p>$a > c, b < d$</p>	<p>Not possible</p>	<p>Any $u_A > u_c$</p>	<p>$\hat{u}_A = 0$</p>	<p>Any $u_A > u_c$</p>	<p>$\hat{u}_A = 0$</p>

Here, we answer this question by indicating the optimum amount of cooperative mutation. Rows distinguish the possible equilibrium structures and columns the fitness landscapes.

Even though, the average fitness can equally well be increased with mutations in only one direction, mutations in both directions are especially interesting for small but equal mutation rates, i.e. for $u = u_A = u_B$. In this case, we can think of u as an error rate of reproduction rather than of a deliberate mutation. Table 2 shows when small error rates are able to increase the average fitness.

- **B dominates A:** A small mutation rate, u , can increase the average fitness if $\Phi(x)$ has a maximum in the open interval $(0, 1)$ or $\Phi(x)$ is monotonically increasing on $(0, 1)$. If $d < (b + c)/2$, one of these two conditions holds.
- **A dominates B:** A small mutation rate, u , can increase the average fitness if $\Phi(x)$ has a maximum in $(0, 1)$ or $\Phi(x)$ is monotonically decreasing on $(0, 1)$. If $a < (b + c)/2$, one of these two conditions holds.
- **A and B coexist:** Let x^* denote the inner equilibrium for $u = 0$ and x_u^* the inner equilibrium for $u > 0$. We have $\Phi'(x^*) = c - b$. Therefore, $\Phi(x)$ at $x = x^*$ is increasing if $b < c$ and decreasing if $b > c$. For a small mutation rate, we have $x_u^* < x^*$ if and only if $a + b > c + d$. Hence, a small mutation rate, u , increases Φ if

(i) $b < c$ and $a + b < c + d$

or

(ii) $b > c$ and $a + b > c + d$.

- **A and B are bistable:** Since there are two stable equilibria, we have to consider x_0 , the initial value of x_A . We will distinguish two cases, $x_0 = 0$ and 1. If $x_0 = 0$, then a small mutation rate can increase Φ only if $\Phi(x)$ is increasing on $(0, 1)$. If $x_0 = 1$ then a small mutation rate can increase Φ only if $\Phi(x)$ is decreasing on $(0, 1)$.

4. Maximizing the reproductive potential of A

In Sections 2 and 3, we studied the effect of mutations on the average fitness of the whole population, i.e. A and B. Naturally, the question of whether mutations can be beneficial for just one of the players arises. Is, for example, player A able to increase his reproductive potential by (deliberately or undeliberately) mutating into B?

To address this question we analysed Eq. (5) with respect to A's reproductive potential, $F_A = x_A f_A$ (see Appendix B). As we can see in Table 3, mutations from A to B can indeed increase A's reproductive potential. A necessary condition for this is $2a < b$. Hence, only if B contributes more than twice as much to A's fitness than A does, can mutations from A to B increase the reproductive potential of A.

In Appendix B we also analyse whether F_A can be increased by replication errors, that is, low-frequency mutations in both directions (see Table 4). We found an

Table 2
Can replication errors increase the average fitness?

	<p>$a, d < \frac{b+c}{2}$ $\hat{x} \in (0, 1)$</p>	<p>$a < \frac{b+c}{2} < d$ $\hat{x} = 0$</p>	<p>$d < \frac{b+c}{2} < a$ $\hat{x} = 1$</p>	<p>$\frac{b+c}{2} < a, d$ $\hat{x} = 0$ if $a < d$ $\hat{x} = 1$ if $d < a$</p>	
B dominates A B \leftarrow \rightarrow A $a < c, b < d$	Yes	No	Yes	No	No
A dominates B B \rightarrow \leftarrow A $a > c, b > d$	Yes	Yes	No	No	No
A and B coexist B \rightleftarrows A $a < c, b > d$	Yes if $c+d < a+b \wedge c < b$ or $a+b < c+d \wedge b < c$	Yes if $c+d < a+b$	Yes if $a+b < c+d$	Not possible	
A and B are bistable B \leftarrow \rightarrow A $a > c, b < d$	Not possible	Yes if $x_0 = 1$	Yes if $x_0 = 0$	No	No

For every equilibrium structure (rows) and every fitness landscapes (columns) we indicate whether low-frequency replication errors $u = u_A = u_B$ increase the average fitness. For the bistable equilibrium structure (fourth row) the result depends on x_0 , the initial frequency of A.

Table 3
Is it possible for A to increase its reproductive potential $F_A = x_A f_A$ by producing B?

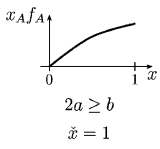
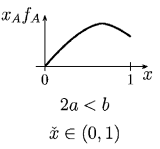
	<p>$2a \geq b$ $\tilde{x} = 1$</p>	<p>$2a < b$ $\tilde{x} \in (0, 1)$</p>
B dominates A B \leftarrow \rightarrow A $a < c, b < d$	Cooperative mutations from A to B are not able to influence the outcome	
A dominates B B \rightarrow \leftarrow A $a > c, b > d$	$\tilde{u}_A = 0$	u_A from Eq. (16)
A and B coexist B \rightleftarrows A $a < c, b > d$	$\tilde{u}_A = 0$	If $c-a < (1-\frac{d}{b})(b-2a)$ then u_A from Eq. (16) else $\tilde{u}_A = 0$
A and B are bistable B \leftarrow \rightarrow A $a > c, b < d$	$\tilde{u}_A = 0$	If $b^2(a-b+d-c) < 4a(ad-bc)$ then u_A from Eq. (16) else u_A from Eq. (7)

We answer this question by providing the optimal mutation rate \tilde{u}_A . We write $\tilde{u}_A = 0$ if mutations from A to B are not able to increase F_A .

increase of F_A in surprisingly many cases. The trivial cases are the ones with a dominating B or $x_B = 1$. But even if A dominates or $x_A = 1$ is an increase of F_A

possible. In particular, if $2a < b$ and $x_A = 1$, F_A can always be increased by low-frequency mutations in both directions.

Table 4
Can low-frequency replication errors $u = u_A = u_B$ increase A's reproductive potential $F_A = x_A f_A$?

	 <p style="text-align: center;">$2a \geq b$ $\hat{x} = 1$</p>	 <p style="text-align: center;">$2a < b$ $\hat{x} \in (0, 1)$</p>
B dominates A B \leftarrow \rightleftarrows A $a < c, b < d$	Yes	Yes
A dominates B B \rightleftarrows A $a > c, b > d$	No	Yes
A and B coexist B \rightleftarrows A $a < c, b > d$	Yes if $a + b < c + d$	Yes if $a + b > c + d \wedge c - a < (1 - \frac{d}{b})(b - 2a)$ or $a + b < c + d \wedge c - a > (1 - \frac{d}{b})(b - 2a)$
A and B are bistable B \leftarrow \rightleftarrows A $a > c, b < d$	Yes if $x_0 = 0$	Yes

5. Discussion

In this paper, we have analysed a replicator–mutator equation of two strategies A and B. The frequency of A and B in the infinitely large population are given by x_A and x_B with $x_A + x_B = 1$. The fitness of A and B is given by $ax_A + bx_B$ and $cx_A + dx_B$, respectively.

We have answered the question of when a mutation from A to B increases the average fitness Φ of the population (see Table 1). We have also calculated the condition for mutation from A to B to increase $F_A = x_A f_A$, the reproductive potential of A (see Table 3). In both cases, the maximum value (of Φ or F_A) is achieved for a mutation from A to B, if the fitness of A increases sufficiently as a function of the frequency of B, that is, if the coefficient b in the payoff matrix is large enough.

We have also studied the effect of low-frequency mutations in both directions. Such mutations can be interpreted as replication or transition errors during the evolutionary process described by the replicator equation. Again, we have calculated the conditions for these mutations to increase the equilibrium values of the average fitness Φ of the population (Table 2), and the reproductive potential of A, F_A (Table 4). We showed that such replication errors can increase Φ as well as F_A in many cases.

Normally, we think of mutation as something detrimental in a non-changing environment. For constant selection, the average fitness at equilibrium is a declining function of the mutation rate. The purpose of this paper is to point out that in the case of frequency-dependent selection (= evolutionary games) the effect of mutation can be subtle. It is even possible that mutation increases the equilibrium fitness of a population.

Appendix A. Mutations in both directions

The equilibria of Eq. (9) are given by the solutions of a third-order polynomial. In particular by the solutions of

$$[a - b + d - c]x^3 + \left[(1 - u_A)(b - a) + u_B(c - d) + 2\left(\frac{b + c}{2} - d\right) \right]x^2 - [u_B(c - 2d) + (1 - u_A)b - d]x - u_B d = 0. \tag{10}$$

We are interested in the relation between u_A and u_B for $x = \hat{x}$, i.e. for equilibria that maximize Φ . We define

$$k = \frac{(b - a)\hat{x}^2 - b\hat{x}}{(c - d)\hat{x}^2 - (c - 2d)\hat{x} - d},$$

$$h = \frac{(-a + b - d + c)\hat{x}^3 + (a - 2b - c + 2d)\hat{x}^2 + (b - d)\hat{x}}{(c - d)\hat{x}^2 - (c - 2d)\hat{x} - d} \tag{11}$$

and rewrite Eq. (10) for $x = \hat{x}$ as

$$u_B = ku_A + h. \tag{12}$$

The linear relation between u_A and u_B becomes obvious. To discuss the properties of the coefficients k and h we will use the parameters

$$\alpha = \frac{b + c}{2} - a, \quad \beta = \frac{b + c}{2},$$

$$\delta = \frac{b + c}{2} - d, \quad \gamma = \frac{b - c}{2} \tag{13}$$

instead of $a, b, c,$ and d . For these parameters we have $\hat{x} = \delta / (\delta + \alpha)$ and \hat{u}_A from Eq. (6) becomes

$$\frac{\alpha\gamma}{\delta(\beta - \alpha) + \alpha(\beta + \gamma)}.$$

Obviously, $\hat{x} \in (0, 1) \Leftrightarrow \alpha, \delta > 0$ and $\hat{u}_A > 0 \Leftrightarrow \gamma > 0$. We can write Eq. (11) as

$$k = \frac{\delta(\alpha(\beta - \delta) + \delta\beta + \gamma\alpha)}{\alpha(\alpha(\beta - \delta) + \delta(\beta - \gamma))},$$

$$h = \frac{-\delta\gamma}{\alpha(\beta - \delta) + \delta(\beta - \gamma)}. \quad (14)$$

Since $\beta > \alpha, \delta, \gamma$, we have $k > 0$ and $h > 0 \Leftrightarrow \gamma < 0$. Hence, if \hat{x} is a stable equilibrium for a system with $u_A, u_B > 0$, it is also a stable equilibrium of a system where $u_A = 0$ (if $\gamma < 0$) or $u_B = 0$ (if $\gamma > 0$).

From this we conclude that for every system in which cooperative mutations in both directions are able to maximize the average fitness, a system exist that maximizes the average fitness equally well but requires cooperative mutations only in one direction. In particular, cooperative mutations from A to B are sufficient if $c < b$, and mutations from B to A are sufficient if $b < c$.

Appendix B. Maximizing the reproductive potential of A

Let $F_A = x_A f_A$ denote the reproductive potential of A. We want to study if F_A can be increased by mutation from A to B. As before, let $x_A = x$ denote the frequency of A and $x_B = 1 - x$ the frequency of B. We have

$$F_A(x) = (a - b)x^2 + bx. \quad (15)$$

There are two different fitness landscapes. If $b < 2a$ then $F_A(x)$ increases monotonically on the interval $[0, 1]$. If $b > 2a$ then $F_A(x)$ has a maximum at $\check{x} = b/[2(b - a)]$. Table 3 gives the mutation rates \check{u}_A that maximize F_A . Obviously, $\check{u}_A = 0$ for a monotonically increasing $F_A(x)$. Therefore we will discuss Table 3 only for $\check{x} \in (0, 1)$.

- **B dominates A:** Mutations from A to B cannot influence the equilibrium structure; hence, u_A is unimportant.
- **A dominates B:** The optimum mutation rate is

$$\check{u}_A = \frac{(2a - b)(b(a - b + d - c) - 2(ad - bc))}{2b(b - a)^2}. \quad (16)$$

- **A and B coexist:** Let x^* denote the inner equilibrium for $u_A = 0$. Cooperative mutations can increase F_A only if $\check{x} < x^*$, which is equivalent to $c - a < (1 - \frac{d}{b})(b - 2a)$. In this case \check{u}_A is given by Eq. (16). If $c - a \geq (1 - \frac{d}{b})(b - 2a)$, then $\check{u}_A = 0$.
- **A and B are bistable:** Let x_0 denote the initial value of x_A . Since mutations from A to B are without consequence for $x_0 = 0$, we will only consider $x_0 = 1$. We recollect the critical mutation rate u_c from Eq. (7) and the corresponding equilibrium x_c from Eq. (8). If $x_c < \hat{x}$, then \check{u}_A is given by Eq. (16) else $\check{u}_A = u_c$. We note that $x_c < \hat{x}$ is equivalent to $b^2(a - b + d - c) < 4a(ad - bc)$.

After a complete classification of when cooperative mutations from A to B can increase the reproductive potential of A, we will analyse when small replication errors are able to do the same. The differential equations are given by Eq. (9). Let $u = u_A = u_B$ be the error rate. Table 4 shows the results.

- **B dominates A:** A small mutation rate always increases F_A .
- **A dominates B:** F_A can only be increased by a small mutation rate if $\check{x}_A \in (0, 1)$, which means $2a < b$.
- **A and B coexist:** Let x^* denote the inner equilibrium for $u = 0$ and x_u^* the inner equilibrium for $u > 0$. We know that $x^* < x_u^*$ is equivalent to $a + b < c + d$ (see Section 3). We have

$$F'_A(x^*) > 0 \Leftrightarrow c - a > \left(1 - \frac{d}{b}\right)(b - 2a).$$

Therefore, a small mutation rate increases F_A if

$$(i) \ a + b < c + d \quad \text{and} \quad c - a > \left(1 - \frac{d}{b}\right)(b - 2a)$$

or

$$(ii) \ a + b > c + d \quad \text{and} \quad c - a < \left(1 - \frac{d}{b}\right)(b - 2a).$$

- **A and B are bistable:** If $\check{x} \in (0, 1)$, replication errors increase F_A . If $\check{x} = 1$, errors are only able to increase F_A if x_0 , the initial value for x_A , equals 0.

References

Bomze, I., Burger, R., 1995. Stability by mutation in evolutionary games. *Games Econ. Behav.* 11 (2), 146–172.

Fudenberg, D., Harris, C., 1992. Evolutionary dynamics with aggregate shocks. *J. Econ. Theory* 57 (2), 420–441.

Fudenberg, D., Levine, D.K., 1998. *The Theory of Learning in Games*. MIT Press, Cambridge, MA.

Hadeler, K., 1981. Stable polymorphisms in a selection model with mutation. *SIAM J. Appl. Math.* 41 (1), 1–7.

Haldane, J.B., 1937. The effect of variation on fitness. *Am. Nat.* 71 (735), 337–349.

Hofbauer, J., Sigmund, K., 1989. On the stabilizing effect of predators and competitors on ecological communities. *J. Math. Biol.* 27 (5), 537–548.

Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.

Hofbauer, J., Schuster, P., Sigmund, K., 1979. A note on evolutionary stable strategies and game dynamics. *J. Theor. Biol.* 81 (3), 609–612.

Kimura, M., 1967. On evolutionary adjustment of spontaneous mutation rates. *Genet. Res.* 9 (1), 23.

Maynard-Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.

Maynard-Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18.

- Nash, J., 1950. Equilibrium points in n -person games. *Proc. Natl Acad. Sci. USA* 36 (1), 48–49.
- Nash, J., 1951. Non-cooperative games. *Ann. Math.* 54 (2), 286–295.
- Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoners-dilemma. *Acta Appl. Math.* 20 (3), 247–265.
- Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. *Science* 303 (5659), 793–799.
- Nowak, M.A., Komarova, N.L., Niyogi, P., 2001. Evolution of universal grammar. *Science* 291 (5501), 114–118.
- Nowak, M.A., Komarova, N.L., Niyogi, P., 2002. Computational and evolutionary aspects of language. *Nature* 417 (6889), 611–617.
- Page, K.M., Nowak, M.A., 2002. Unifying evolutionary dynamics. *J. Theor. Biol.* 219 (1), 93–98.
- Sasaki, A., Nowak, M.A., 2006. Mutation rate that maximizes evolvability. Manuscript in preparation.
- Sniegowski, P.D., Gerrish, P.J., Lenski, R.E., 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387 (6634), 703–705.
- Stadler, P., Schuster, P., 1992. Mutation in autocatalytic reaction networks—an analysis based on perturbation-theory. *J. Math. Biol.* 30 (6), 597–632.
- Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., Godelle, B., 1997. Role of mutator alleles in adaptive evolution. *Nature* 387 (6634), 700–702.
- Taylor, P., Jonker, L., 1978. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* 40 (1–2), 145–156.
- Weibull, J., 1995. *Evolutionary Game Theory*. MIT Press, Cambridge, MA.