Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 8676–8680, July 1998
Evolution

A 38

The evolution of interspecific mutualisms

MICHAEL DOEBELI*† AND ‡NANCY KNOWLTON

*Zoology Institute, University of Basel, Rheinsprung 9, CH-4051 Basel, Switzerland; and ‡Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Republic of Panama

Edited by John Maynard Smith, University of Sussex, Brighton, United Kingdom, and approved May 12, 1998 (received for review February 17, 1998)

ABSTRACT **Interspecific mutualisms are widespread, but how they evolve is not clear. The Iterated Prisoner's Dilemma is the main theoretical tool to study cooperation, but this model ignores ecological differences between partners and assumes that amounts exchanged cannot themselves evolve. A more realistic model incorporating these features shows that strategies that succeed with fixed exchanges (e.g., Tit-for-Tat) cannot explain mutualism when exchanges vary because the amount exchanged evolves to 0. For mutualism to evolve, increased investments in a partner must yield increased returns, and spatial structure in competitive interactions is required. Under these biologically plausible assumptions, mutualism evolves with surprising ease. This suggests that, contrary to the basic premise of past theoretical analyses, overcoming a potential host's initial defenses may be a bigger obstacle for mutualism than the subsequent recurrence and spread of noncooperative mutants.**

Mutually beneficial interactions between members of different species play a fundamental role in all ecosystems (1, 2), but their evolution has challenged theoreticians for decades. In 1971, Trivers (3) pointed out the relevance of the Prisoner's Dilemma, in which two unrelated players benefit by cooperating but do even better by cheating their partners. Ten years later, Axelrod and Hamilton (4) proposed the Iterated Prisoner's Dilemma, in which the same partners interact repeatedly, as the basic framework for finding cooperative solutions to this problem. Although other approaches have been developed (1, 5–11), analyses of reciprocation have dominated the theoretical literature on cooperation without kinship ever since (11–18).

The general conclusions from this large body of work are that strategies that reward cooperation and punish defection or that persist in previously successful behaviors (Tit-for-Tat, Pavlov, and their relatives) often do well. However, two assumptions underlying these analyses are inconsistent with the biology of most interspecific mutualisms (19): that the players compete directly with their partners and that the payoffs associated with cooperation, defection, cheating, and being cheated are constants. The former assumption ignores the reality that partners in interspecific mutualisms (e.g., corals and dinoflagellates, higher plants and mycorrhizal fungi) usually have different ecological requirements. The latter assumption makes it impossible to analyze how initially neutral relationships evolve into mutually beneficial ones and is probably also invalid for many cases of intraspecific cooperation (11).

We therefore reframed the problem as a game in which payoffs vary according to investments made while still satisfying the inequalities of the Prisoner's Dilemma (4). If a player invests I, it incurs a cost $C(I)$ to itself and provides a benefit $B(I)$ to its partner, and failure to invest has neither costs nor benefits (Fig. 1 *A* and *B*). As long as players interact only once, no investment is the best strategy, even though both players would be better off if both invested. In the classical Iterated Prisoner's Dilemma, this paradox is solved (4, 12–16) by strategies that lead to repeated mutual investments. However, if the amount of investment can itself evolve, investments of each player should decrease to 0 whenever the partner's decision to invest depends only qualitatively on whether a player cooperates at all, rather than quantitatively on how much the player invests. For in that case the partner's response does not depend on the amount of one's investments, and because investments are costly, they should be minimized. It follows that Tit-for-Tat and Pavlov (and similar, successful strategies in the Iterated Prisoner's Dilemma) have the same pitfalls as simple cooperation in the noniterated Prisoner's Dilemma once the assumption of fixed costs and benefits is relaxed.

To solve this new dilemma, we assumed that the investment decision in one round of the iterated game depends quantitatively on the payoff received in the previous round. This assumption is biologically equivalent to the plausible notion that healthy organisms have more to offer their partners. In our evolutionary game, hosts and symbionts are described by their investment strategies, which are given by two parameters: a, the initial offer, and b, the reward rate or rate of increase of investment as a function of past payoff (Fig. 1*C*). In each round of the iterated game, investments lead to benefits and costs that are given by two functions $B(I)$ and $C(I)$ (Fig. 1*B*). For these functions, we used the formulas $B(I) = B_0(1 - \exp[-B_1 I])$ and $C(I) = C_0 I$, where B_0, B_1, and C_0 are parameters determining the cost–benefit relation (thus we assumed that costs increase linearly with investment I, but benefits level off, Fig. 1*B*).

The interaction between a host with strategy (a,b) and a symbiont with strategy (a',b') is defined as follows. The partners start by investing a and a', respectively. Therefore, in the first iteration, the host gets the payoff $B(a') - C(a)$, i.e., it gets the benefit $B(a')$ from the symbiont's investment and it pays a cost $C(a)$ because of its own investment. Similarly, the symbiont gets the payoff $B(a) - C(a')$. In every further iteration, the host invests $a + bp$, where p is the payoff to the host from the previous iteration, except if that payoff was negative, in which case the investment is 0. Similarly, the symbiont invests $a' + b'p'$ in every further iteration, where p' is the payoff to the symbiont from the previous round, except if that payoff was negative, in which case the investment is 0. For example, in the second iteration, the host invests $a + b[B(a') - C(a)]$ and the symbiont invests $a' + b'[B(a) - C(a')]$ unless $B(a') - C(a)$ or $B(a) - C(a')$ are negative, in which case the corresponding investment is 0. The total payoff of the mutualistic interaction within one generation is the sum of the payoffs from a fixed number of iterations.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/958676-5$2.00/0
PNAS is available online at http://www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
†To whom reprint requests should be addressed. e-mail: doebeli@ubaclu.unibas.ch.

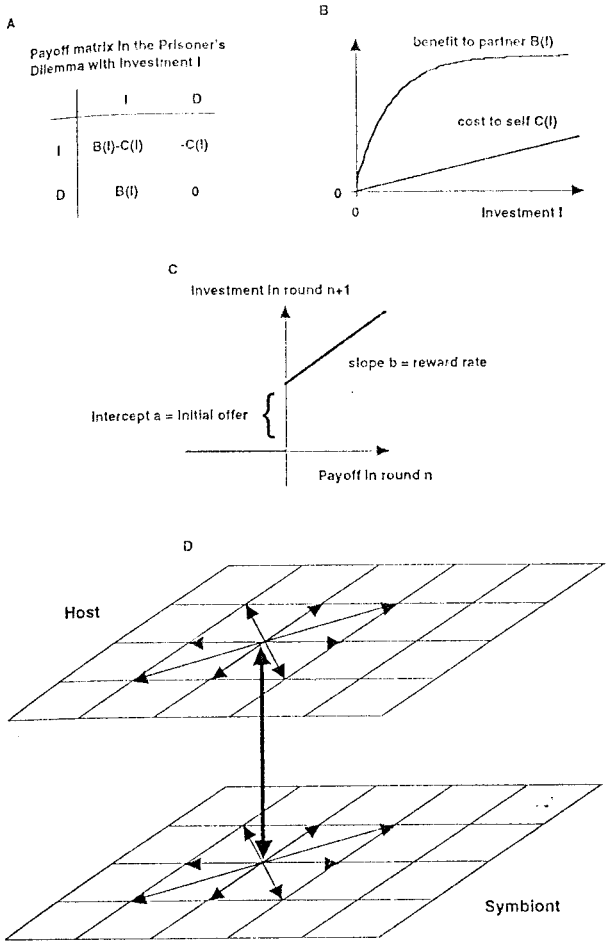

FIG. 1. Model structure. (*A*) Payoff matrix for partners that invest *I* or do nothing *D*. As long as $B(I) > C(I)$, the payoffs satisfy the assumptions of the Prisoner's Dilemma. (*B*) Benefit $B(I)$ and cost $C(I)$ functions used in simulations. (*C*) Strategy for repeated interactions, defined by the initial offer (intercept *a*) and by the reward rate (slope *b*). (*D*) Spatial lattices for host and symbiont; example illustrated shows 30 host-symbiont pairs; heavy vertical arrow symbolizes mutualistic (between lattice) interaction; horizontal arrows symbolize competitive (within lattice) interactions.

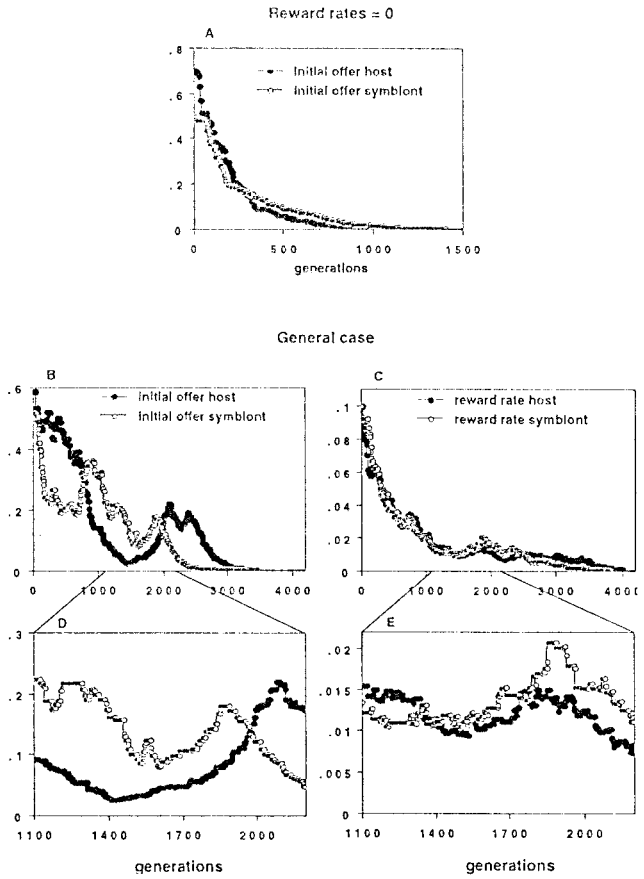

FIG. 2. Representative examples of simulations without spatial structure. (*A*) With 0 reward rates, the initial offer evolves monotonically to zero. (*B* and *C*) More complex dynamics when both initial offer and reward rate evolve. (*D* and *E*) Detailed view of dynamics, showing transient period of increasing mutualism (generations 1400–1900). For simulations shown here, cost and benefit parameters were $B_0 = 4.7$, $B_1 = 1.8$, and $C_0 = 0.6$ (Fig. 1*B*), with 15 iterations per mutualistic interaction.

To simulate evolution, we assigned hosts and symbionts a starting phenotype for *a* and *b* and tested them against mutants in each generation. If in any generation the mutant host does better than the existing host in the iterated game against the existing symbiont, then the mutant phenotype replaces that of the existing host. Replacement in the symbiont occurs analogously. This evolutionary scenario assumes that "invasion implies fixation," i.e., that initial increase of a rare mutant will lead to replacement of the former resident by the mutant, and that ecological, that is, replacement, dynamics occur on a much faster time scale than evolutionary dynamics, so that replacement has occurred before the next mutation arises. The evolution of mutualism from a neutral relationship occurs when the parameters *a* and *b* evolve away from 0 in both partners.

Simulations confirmed that when the reward rate *b* is fixed at 0 (as in classical Tit-for-Tat or Pavlov), then the initial offer *a* evolves monotonically to 0 (Fig. 2*A*). When both *a* and *b* were allowed to vary, the reward rate *b*, and consequently the initial offer *a*, again invariably evolved to 0 in our simulations (Fig. 2 *B* and *C*), but the evolutionary dynamics were much more complex. If the host has a higher reward rate than the symbiont, it is better for the host to have a low initial offer to keep costs down, whereas it is better for the symbiont to have a high initial offer because otherwise returns would be too low (and *vice versa*, e.g., generations 1,100–1,300 and 2,000–2,200 in Fig. 2 *D* and *E*). Mutations can lead to switches in the ranking of the reward rates in the host and the symbiont. Until the corresponding switch in the initial offers evolves, there may be transient periods during which selection favors higher parameter values in both partners (e.g., generations 1,400–1,900 in Fig. 2 *D* and *E*), causing the average payoff to increase temporarily. Thus, even though all parameters and payoffs eventually evolve to 0, the evolutionary dynamics are not monotonic. This complexity, together with earlier results from spatial game theory (17, 18, 20, 21), suggested that considering spatially structured populations in which competitive interactions occur locally could yield different results.

Therefore, we placed the hosts and the symbionts on separate spatial lattices (Fig. 1*D*). The iterated game was played between hosts and symbionts at corresponding positions on these lattices, after which competition occurred between nearest neighbors on each lattice separately. Thus, hosts and symbionts at corresponding positions in the two lattices interact according to their phenotypes as described above. The payoffs from these interactions are then used to determine the winners in intraspecific competition within hosts and within symbionts. At each position in both lattices, we compared the payoffs of the occupant with the payoffs of the eight nearest neighbors (Fig. 1*D*). The individual with the highest payoff among those compared then was placed at this focal position. To complete the evolutionary scenario, we assumed that, after

competition, mutations occur at each lattice position with a certain probability. Mutants had a phenotype drawn from a bivariate normal distribution with mean the previous occupant's phenotype and variance a certain percentage of the mean.

The dual lattice structure with two separate lattices for hosts and symbionts corresponds to the idea of biological markets (9) and reflects the fact that, in nature, ecological differences between hosts and symbionts normally limit direct competition between them. Nevertheless, interspecific mutualisms contribute indirectly to the outcome of intraspecific competition because successful mutualists are potentially better competitors. Our use of spatial lattices implies that competition for reproductive success is stronger between neighboring individuals than between individuals far apart. Thus, the bigger the size of the lattices, the larger the total number of hosts and symbionts relative to the size of the local neighborhood within which competition occurs. One rather restrictive assumption in our model is equal lattice size for hosts and symbionts and hence equal population size of the two partners. Future work should allow for different lattice sizes and for empty lattice sites, thus allowing the species abundances to vary between the partners and over time.

In our model, numerical simulations showed that, with the addition of spatial structure, the initial offer a and the reward rate b can increase from very low levels, and long term persistence of mutualism is possible (Fig. 3*A*). Maintenance of mutualism is characterized by large fluctuations in the average payoffs in the host and in the symbiont (Fig. 3*A*). These fluctuations are due to fluctuations in the average reward rates and initial offers in the host and the symbiont (Fig. 2). At any one point in time, the partner with the higher average reward rate and lower average initial offer has the higher average payoff. However, the ranking of average reward rates and initial offers in the two partners changes frequently (Fig. 2), leading to the observed fluctuations in average payoffs. A large number of numerical simulations showed that spatial structure can maintain mutualistic relationships for a wide range of model parameters (see below).

Our results are related to those of Mar and St. Denis (22), who also modeled continuously varying strategies in the Iterated Prisoner's Dilemma. They similarly concluded that a continuous version of Tit-For-Tat does well against a number

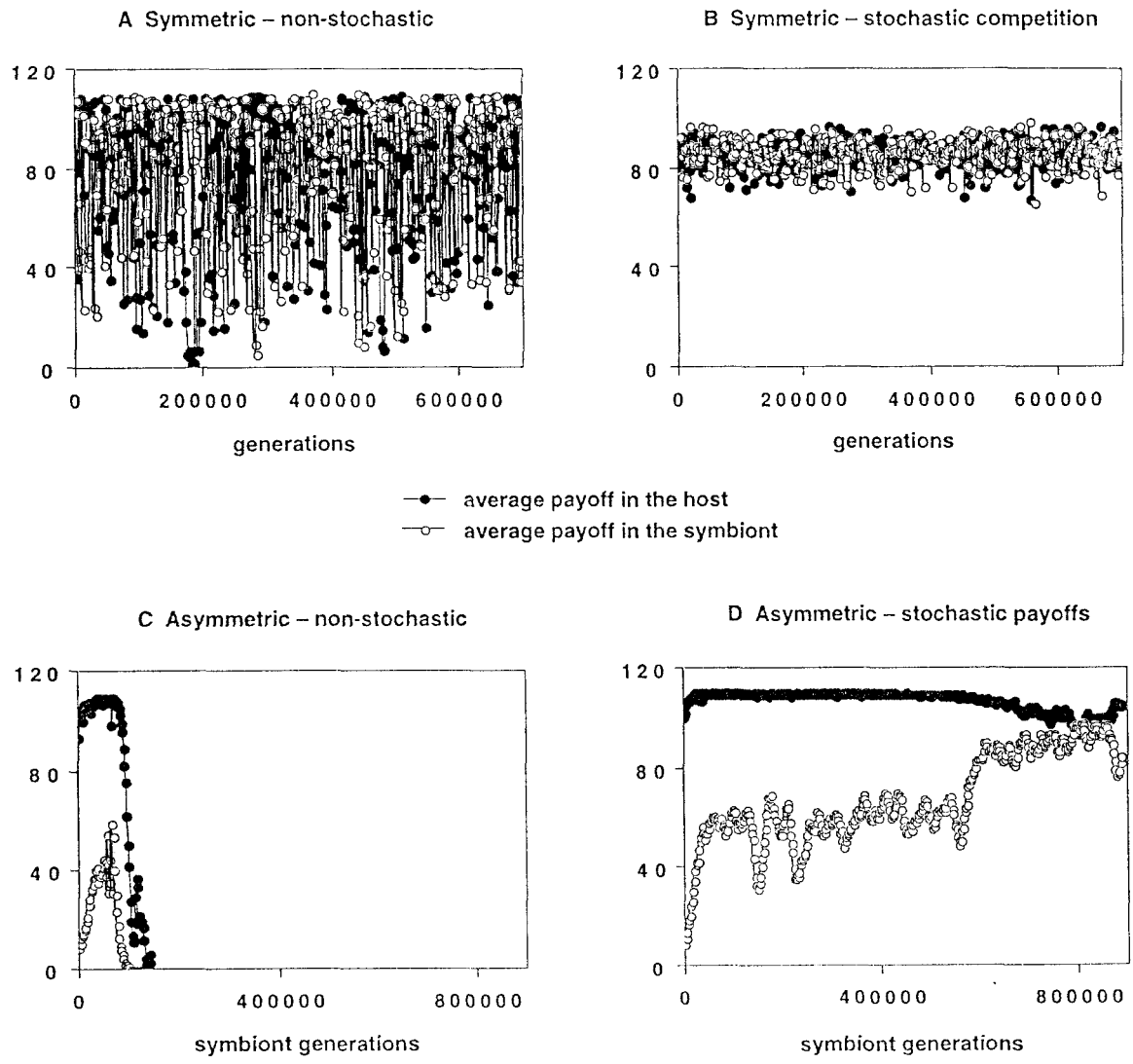

FIG. 3. Representative examples of simulations with spatial structure. (*A*) Mutualism persisting with large fluctuations in the average payoffs. (*B*) Same as *A* but with stochastic competition and much smaller fluctuations. (*C*) Asymmetry in generation time (100 symbiont generations per host generation) resulting in asymmetry in the payoffs and loss of mutualism. (*D*) Same as *C* but with stochastic payoffs, leading to persistence with payoff asymmetry. For these simulations, $B_0 = 7$, $B_1 = 2.5$, and $C_0 = 0.1$ (Fig. 1*B*). There were 15 iterations per mutualistic interaction, mutation probability per lattice site was 0.02, mutational variance was 5% of the mean, lattice size was 25×25, and a and b were set initially to 0.005 across both lattices.

Table 1. Summary of the effects of different parameters on the evolution of mutualism

Increase of parameter	Effect on initial spread and persistence
Population (lattice) size	positive (*cf.* Fig. 4)
Number of mutualistic interactions (iterations) per generation	positive
Maximum benefit B_0	positive (*cf.* Fig. 4)
Unit cost C_0	negative (*cf.* Fig. 4)
Vertical transmission	positive
Nonlocal dispersal	negative
Stochasticity in contest	positive (*cf.* Figs. 3*B* and 4)
Stochasticity in payoff	positive (*cf.* Figs. 3*D* and 4)
Asymmetry in generation time	negative (*cf.* Fig. 3*C*)
Asymmetry in mutation rate or magnitude	negative

The robustness of these patterns was confirmed by hundreds of simulations by using a variety of baseline conditions.

of other strategies in spatial games, but there are three important differences between their approach and ours. First, their strategies are not defined by investment decisions. Second, they do not consider the dual lattices that account for ecological differences between potentially mutualistic species. Most importantly, they do not simulate evolutionary dynamics by introducing a constant stream of randomly generated mutant strategies but by competition between a limited and fixed number of strategies. This *a priori* prevents a gradual evolutionary decay of cooperation. Thus our results provide more robust support for the importance of spatial structure when strategies vary continuously.

Unfortunately, analytical results for the conditions that favor the evolution of mutualism are not feasible. Even in the nonspatial case (Fig. 2), the iterative procedure for the mutualistic interaction renders the dependence of the total payoff on phenotypic values analytically intractable. To appreciate the complications, recall that the evolutionary dynamics toward loss of mutualism in the nonspatial case are not monotonic (Fig. 2 *D* and *E*). In particular, the direction of selection on mutualism in one of the partners depends on the present phenotype of the other partner, leading to complicated interactions which nevertheless always result in the eventual loss of mutualism. In the spatial model with dual lattices, there is even less hope for analytical results. An intuitive explanation of why mutualism can persist with spatial heterogeneity is as follows. Across the lattices, various dynamically homogeneous pockets of decaying mutualism (Fig. 2) develop. Interactions at the boundaries of such pockets lead to transient local selection for more mutualistic phenotypes with higher average payoffs (similar to the transient effects shown in Fig. 2 *D* and *E*). Therefore, new and more mutualistic pockets form that again start to decay and at the boundaries of which new transients are generated. This leads to continual recurrence of mutualistic types (think of a boiling sea of mutualistic bubbles) with overall maintenance of mutualism. However, persistence does not imply an evolutionary steady state. Instead, initial offers and reward rates vary in space and time, inducing fluctuations in average payoffs (Fig. 3*A*). This variation suggests that real world mutualisms should be characterized by considerable genetic heterogeneity in the amount that partners invest in each other, a result similar to those obtained by Ikegami and Kaneko (23).

We varied specific features of the model to see how biologically important factors should influence the evolution of mutualism. Larger populations were modeled by increasing the size of the spatial arrays. Increasing the duration of a relationship was modeled by increasing the number of iterations in the mutualistic interaction per generation. The effects of vertical transmission were simulated by assuming that hosts

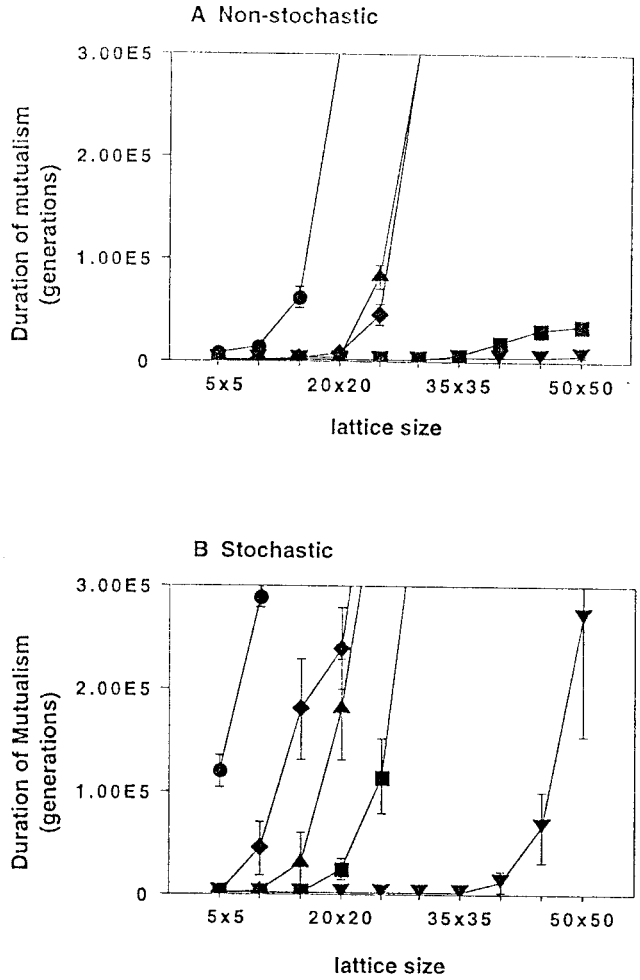

FIG. 4. Effects of population size and stochasticity on persistence of mutualism. Mean and SD of the times until the average payoff fell below 0.001 is shown for 10 runs starting from randomized low initial phenotypic values (a, $b < 0.005$) for a range of population (lattice) sizes; lines leaving the upper horizontal indicate persistence beyond 300000 generations. (*A*) No stochasiticity. (*B*) Same as *A* but with stochastic competition and stochastic payoffs. ●, baseline parameters $B_0 = 8$, $B_1 = 4.2$, and $C_0 = 0.4$ in the benefit and cost functions; other scenarios were obtained by varying one of these parameters: ▲, $B_0 = 1.3$; ◆, $B_1 = 0.7$; ■, $C_0 = 12$; different parameters with a high cost-benefit ratio were used for ▼: $B_0 = 20$, $B_1 = 4.2$, and $C_0 = 50$. The number of iterations per mutualistic interaction was 10.

that win the competition bring their symbionts with them with a certain probability, i.e., the success of these symbionts was determined by the success of their hosts rather than by competition among symbionts. Nonlocal dispersal was modeled by competition occurring not with nearest neighbors but with eight individuals randomly chosen from the lattice. The sensitivity of the results to different costs and benefits was assessed by varying the parameters B_0, B_1, and C_0 in the cost and benefit functions. We also tested two types of stochasticity. For stochastic competition (Fig. 3*B*), the occupant at each lattice site was left unchanged with a certain probability. For stochasticity in the payoff (Fig. 3*D*), we drew the actual benefit from investment I in any round of the iterated game from a normal distribution with mean $B(I)$ and variance a certain percentage of the mean.

The results of these modifications are summarized in Table 1. Larger population size, greater number of interactions per generation, high benefit-to-cost ratios, and vertical transmission favored mutualism, whereas dispersal across the lattice rather than to neighboring sites was disadvantageous. These

results parallel those summarized by others (10, 11, 16). Contrary to previous results (11, 16, 24), however, were the effects of stochasticity in our model (Fig. 3 *B* and *D*). Both types of stochasticity were highly beneficial for the evolution of mutualism (Fig. 3 *B* and *D*). Our interpretation of this result is that stochasticity generates many more of the instances of transient local selection for more mutualistic phenotypes that lead to the overall maintenance of mutualism (see above). Also unexpected was the effect of asymmetries in evolutionary rate, which were modeled by giving hosts and symbionts different values for generation time, mutation rate, or mutation magnitude (Fig. 3*C*). Without exception, the partner with the higher evolutionary rate [probably the smaller symbiont in many natural systems (25) (but see also ref. 26)] had a higher investment and a lower payoff, a result consistent with otherwise puzzling patterns of unequal relative benefits in invertebrate–algal and lichen mutualisms (2).

Table 1 summarizes the effects of these parameters in isolation, but the likelihood of mutualism in natural systems will be determined by the combined influence of various factors. Of particular importance are the positive effects of stochasticity and large population size because the real world is noisy and vastly larger than the spatial arrays we were able to use. Fig. 4 reflects the results of numerous simulations that demonstrate that large spatial arrays and stochasticity can rescue otherwise doomed mutualisms for a large range of parameter combinations in the cost and benefit functions.

Our analyses do not apply to indirect mutualisms or mutualisms without repeated interactions, for which other approaches are more appropriate (1, 5–11). We also made a number of simplifying assumptions. For example, we assumed equal lattice sizes for hosts and symbionts and hence equal population size of the two partners, which is unrealistic for all intracellular symbioses. In addition, we assumed one-on-one interactions between hosts and symbionts, that is, we assumed that one host interacts with only one symbiont in every generation and *vice versa*. More realistically, one could assume that a host interacts with several symbionts in one generation (and *vice versa*), which would increase the potential for conflict among cooperators and cheaters. Finally, we made the simplifying assumption that all reproduction is asexual. The effects of relaxing these and other unrealistic assumptions will be explored in future studies.

Nevertheless, our model represents the first specific attempt to combine the Prisoner's Dilemma approach with the basic features of interspecific mutualisms. Overall, our results suggest that the transition from neutral to mutually beneficial interactions should often occur. Moreover, selection for higher benefits and lower costs, an option we did not model, should further stabilize mutualism once it evolves. Thus, the real evolutionary challenge for mutualism may not be the spread of cheaters in established associations, as has been assumed traditionally by theoreticians, but rather the ability to survive in intimate association in the first place. This may explain why symbionts, although much more diverse than previously realized, often have relatively few independent origins (26, 27). Highly conserved molecular mechanisms for the formation of associations (28) and the origin of new mutualisms through lateral genetic transfer (29) also support this perspective.

We thank M. Ackermann, D. Haig, J. B. C. Jackson, T. Killingback, R. Law, J. Maynard Smith, M. Mesterton-Gibbons, R. Rowan, L. Sanders, S. C. Stearns, and an anonymous referee for their comments; F. Rösel and P. Zimak for computing advice; and the Swiss National Science Foundation and the Zoology Institute of the University of Basel for financial support.

1. Boucher, D. H., ed. (1985) *The Biology of Mutualism: Ecology and Evolution* (Oxford Univ. Press, New York).
2. Smith, D. C. & Douglas, A. E. (1987) *The Biology of Symbiosis* (Edward Arnold, London).
3. Trivers, R. L. (1971) *Q. Rev. Biol.* **46**, 35–57.
4. Axelrod, R. & Hamilton, W. D. (1981) *Science* **211**, 1390–1396.
5. Roughgarden, J. (1975) *Ecology* **56**, 1201–1208.
6. Bull, J. J. & Rice, W. R. (1991) *J. Theor. Biol.* **149**, 63–[illegible].
7. Law, R. (1991) in *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*, eds. Margulis, L. & Fester, R. (MIT Press, Cambridge, MA), pp. 57–71.
8. Renaud, F. & De Meeus, T. (1991) *J. Theor. Biol.* **152**, 319–327.
9. Noë, R. & Hammerstein, P. (1995) *Trends Ecol. Evol.* **10**, 336–339.
10. Maynard Smith, J. & Szathmáry, E. (1995) *The Major Transitions in Evolution* (Freeman, Oxford).
11. Dugatkin, L. A. (1997) *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, New York).
12. Axelrod, R. & Dion, D. (1988) *Science* **242**, 1385–1390.
13. Nowak, M. A. & Sigmund, K. (1992) *Nature (London)* **355**, 250–253.
14. Nowak, M. A. & Sigmund, K. (1993) *Nature (London)* **364**, 56–58.
15. Boerlijst, M. C., Nowak, M. A. & Sigmund, K. (1997) *J. Theor. Biol.* **185**, 281–293.
16. Brembs, B. (1996) *Oikos* **76**, 14–24.
17. Nowak, M. A. & May, R. M. (1992) *Nature (London)* **359**, 826–829.
18. Grim, P. (1995) *J. Theor. Biol.* **173**, 353–359.
19. Leimar, O. & Axén, A. H. (1993) *Anim. Behav.* **46**, 1177–1182.
20. Axelrod, R. (1984) *The Evolution of Cooperation* (Basic Books, New York).
21. Killingback, T. & Doebeli, M. (1996) *Proc. R. Soc. Lond. B* **263**, 1135–1144.
22. Mar, G. & St. Denis, P. (1994) *Int. J. Bif. Chaos* **4**, 943–958.
23. Ikegami, T. & Kaneko, K. (1990) *Physica D* **42**, 235–243.
24. Wu, J. & Axelrod, R. (1995) *J. Confl. Resol.* **39**, 183–189.
25. Moran, N., von Dohlen, C. D. & Baumann, P. (1995) *J. Mol. Evol.* **41**, 727–731.
26. Law, R. & Lewis, D. H. (1983) *Biol. J. Linn. Soc.* **20**, 249–276.
27. Douglas, A. E. (1995) *Adv. Ecol. Res.* **26**, 69–103.
28. van Rhijn, P., Fang, Y., Galili, S., Shaul, O., Atzmon, N., Winiger, S., Eshed, Y., Lum, M., Li, Y., To, V., *et al.* (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5467–5472.
29. Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A. & Perret, X. (1997) *Nature (London)* **387**, 394–401.