Economic contract theory tests models of mutualism

E. Glen Weyl[a], Megan E. Frederickson[a,b], Douglas W. Yu[c,d,1], and Naomi E. Pierce[a,e]

[a]Society of Fellows, Harvard University, Cambridge, MA 02138; [b]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3G5; [c]State Key Laboratory of Genetic Resources and Evolution, Ecology, Conservation, and Environment Center, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; [d]School of Biological Sciences, University of East Anglia, Norwich, Norfolk NR47TJ, United Kingdom; and [e]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

Although mutualisms are common in all ecological communities and have played key roles in the diversification of life, our current understanding of the evolution of cooperation applies mostly to social behavior within a species. A central question is whether mutualisms persist because hosts have evolved costly punishment of cheaters. Here, we use the economic theory of employment contracts to formulate and distinguish between two mechanisms that have been proposed to prevent cheating in host–symbiont mutualisms, partner fidelity feedback (PFF) and host sanctions (HS). Under PFF, positive feedback between host fitness and symbiont fitness is sufficient to prevent cheating; in contrast, HS posits the necessity of costly punishment to maintain mutualism. A coevolutionary model of mutualism finds that HS are unlikely to evolve de novo, and published data on legume–rhizobia and yucca–moth mutualisms are consistent with PFF and not with HS. Thus, in systems considered to be textbook cases of HS, we find poor support for the theory that hosts have evolved to punish cheating symbionts; instead, we show that even horizontally transmitted mutualisms can be stabilized via PFF. PFF theory may place previously underappreciated constraints on the evolution of mutualism and explain why punishment is far from ubiquitous in nature.

evolution of cooperation | punishment | symbiosis | partner fidelity feedback | host sanctions

In contrast to the decades-long development of social co-operation theory (reviews in refs. 1–5), mutualism theory has only recently begun to be formalized (6–15), and major questions remain open. One of these questions concerns whether mutualists evolve to punish cheaters (1, 7–10, 16, 17). The host sanctions (HS) hypothesis suggests that the costly, selective punishment of cheating symbionts can evolve de novo in host species (i.e., in response to symbiont behaviors). As an alternative, partner fidelity feedback (PFF) suggests that punishment is only apparent and is instead an epiphenomenon of the fact that concordant life histories cause two species to be bound together for an "extended series of exchanges" (1), thereby linking their fitnesses. There has been, to date, no sharp way to distinguish these very different explanations for the maintenance of mutualism, and this has obscured our understanding of the factors that promote or constrain the evolution of cooperation between species.

PFF occurs when the benefits provided by a donor individual to a recipient individual automatically feed back to the donor (1, 9, 18, 19). The harder the donor works to assist the recipient, the better off the recipient is and the more benefits it, in turn, provides back to the donor. For example, a symbiont that provides a nutrient to a host improves the host's vigor, which can indirectly, but automatically, benefit the symbiont by decreasing the risk of host mortality. Analogously, if a donor harms the recipient, such as by failing to provide a valuable nutrient, the harm to the recipient also feeds back automatically to the donor. Note that PFF is possible only when partners associate long enough that the short-term costs of helping can be recuperated by the helper. The most straightforward example of PFF thus involves vertical trans-mission of symbionts, because partners are associated for multiple generations (20, 21).

Under PFF, natural selection favors mutualists rather than cheaters because an individual that fails to cooperate reduces its own fitness (or loses an opportunity to increase its own fitness); no further punishment is necessary. By contrast, HS posits that PFF is not sufficient to negate the incentive to cheat, and, thus, mutualism will persist only if hosts evolve to detect and punish cheaters (1, 13) (Box 1).

> **Box 1.   Partner fidelity feedback vs. host sanctions.** One way to contrast PFF and HS is by analogy to the criminal justice system. PFF is akin to the theory that incarceration should bring about incapacitation and/or rehabilitation, thus pre-venting any prospective (short-term future) harms the criminal may also cause society and only incidentally punishing mis-behavior. Each case is justified individually. By contrast, HS are analogous to theories that emphasize the value of pun-ishment as a public deterrent to criminal activity. According to Becker's economic theory of crime and punishment (45), the least expensive means of deterrence are punishments that are infrequently administered, widely publicized, and severe, such as corporal punishment or torture. The prospect of suffering similar punishment is transmitted culturally to the wider popu-lation and deters misbehavior by others. In a biological context, HS is the theory that host punishment selects for "evolutionary improvement" of the symbiont population. See also analysis in SI Text 1.
>
> An example is recent reaction to illegal downloading of copyrighted media, such as films and music. As in HS, copyright owners have tried to deter such violations by prosecuting indi-viduals who can be shown in court to have shared files illegally and widely publicizing these cases. However, recent years have shown that socially acceptable levels of punishment (typically, small fines) are too low to be an effective deterrent. Alterna-tively, simple nonretributive policies may help reduce violations at low social cost. Illegal downloads consume substantial band-width, so some have argued that establishing an automatic sys-tem of graduated charges for bandwidth consumption would significantly reduce such violations. Such policies are efficient even if illegal downloading is not taking place, as they help ensure efficient use of scarce bandwidth. Thus, as with PFF, natural incentives without any retributive motive may be sufficient to deter antisocial behavior.

Although both concepts (PFF and HS) have been discussed in some form in the literature at least since Trivers' seminal paper

in 1971 (22), it was Bull and Rice (23) who coined the term "partner fidelity" in 1991, which they distinguished from partner choice. Today, partner choice is often used interchangeably with HS, but Bull and Rice originally defined partner choice to mean interactions in which individuals "differentially reward cooperative vs. un-cooperative partners *in advance of any possible exploitation*" (italics added) (ref. 23, p 63), whereas both PFF and HS are differential rewards or punishments implemented *after* exploitation is possible. Here, we argue that a failure to clearly define the differences between PFF and HS has led to their conflation, with the result that experiments demonstrating what appears to be the punishment of cheating in a wide range of mutualisms, including those between yucca plants and yucca moths (24), legumes and nitrogen-fixing bacteria (25, 26), ants and plants (27), plants and mycorrhizal fungi (28), and figs and fig wasps (29), have been generally accepted as evidence for HS (1, 9, 13, 26–29), whereas PFF is the more likely explanation.

Model

To formally differentiate PFF from HS, we use a general principal–agent (host–symbiont) model that includes the possibility of costly punishment. For simplicity, we phrase our argument in terms of cheating actions and punishments. The same argument could be framed in terms of cooperative actions and rewards simply by changing signs [in other words, our value of π (described below) can take positive or negative values], in which case we would be describing the phenomenon of "pseudoreciprocity," which refers to investments in the partner that have been selected for because they trigger automatic (by-product) benefits in return (18, 30). Pseudoreciprocity is therefore a specific application of PFF and, like PFF, has been defined only verbally.

Consider a symbiotic *agent* (A) that may either take a cooperative *action* a^* that is in its host's interest or cheat by taking one of two alternative actions, a_1 or a_2. Something happens to the host as a result of the agent's action. Formally, the host, the *principal* (P), receives some (imperfect) *signal* $s \in S = \{s_1, \ldots, s_N\}$ of the action taken by A. The chance that the principal receives each of these signals depends on the agent's action. Formally, the probability of each signal is $p(s_i|a)$, with the property that $\sum_{i=1}^{N} p(s_i|a) = 1$ for any $a = a^*, a_1, a_2$.

On the basis of the signal, P carries out a response policy $\pi \in [0, \infty)$ that contains the possibility of reducing the fitness of A. The fitness of the principal U^P therefore depends on what happens to the principal (including any signals) and on her response policy π, which is assumed to be costly to carry out. These two may be related to one another, in the sense that after certain things happen to the principal, it may be more or less costly/beneficial to respond. We write this as $U^P(s, \pi)$, and if the policy is based on the signal that the principal receives, the fitness of the principal becomes only a function of the signal, $U^P(s, \pi(s))$. A natural strategy for P is to evolve a *subgame-perfect policy* $\pi^*(s)$ that maximizes its prospective fitness, $U^P(s, \pi(s))$, for each signal s. In other words, a policy $\pi^*(s)$ is a set of prospective-payoff-maximizing (or "best") responses to the signals received, executed without regard to the agent's actions per se. ("Prospective fitness" corresponds to the game-theoretic notion of "subgame-perfection" under which agents are unable to commit to their actions. For example, a threat to detonate a mutually destructive bomb if someone does not turn over their wallet may be useful for extortion, but it is not credible as it is not in the threat-maker's *prospective* interests once the threat has been disregarded.)

A's fitness following any particular action increases with the short-term benefits of any cheating $u^S(a)$, plus the feedback benefits that A generally derives from P's fitness, scaled by f, and decreases with P's response π that A receives. Recall that P's response also can affect P's own fitness, altering the feedback benefits that A receives. The expected fitness of any action $U^A(a; \pi)$ thus depends in two ways on P's response policy. If, for simplicity, we let these have independent influences on A, then

$$U^A(a; \pi) = u^S(a) + fU^P[s; \pi] - \pi \qquad [1.1]$$

By the definition of actions a_1 and a_2 being short-term profitable actions, we assume that $u^S(a_1), u^S(a_2) > u^S(a^*)$; that is, cheating improves A's short-term fitness. Importantly, because nature is stochastic, any single action a could result in any one of N signals (outcomes) s, albeit with different probabilities for different actions. For instance, a symbiont might not protect its host from parasites, which increases the probability of but does not doom the host to parasitism, or a symbiont might engage in protective behaviors but still fail to prevent parasitism. Eq. **1.1** is therefore expressed generally as an expected fitness across a distribution of possible signals, given a particular action:

$$U^A(a; \pi[\cdot]) \equiv u^S(a) + \sum_{i}^{N} (fU^P(s_i, \pi[s_i]) - \pi[s_i])p(s_i|a). \qquad [1.2]$$

This model allows us to differentiate PFF and HS once we have made an important, if common, assumption: Subject to the constraints imposed by the theories, both A and P maximize their fitness. If PFF is sufficient to maintain mutualism between A and P, then to explain the behavior observed we have no need to posit any response other than $\pi*$. Therefore, PFF is the theory that P uses π^* and that

$$U^A(a^*; \pi^*) \geq U^A(a_1; \pi^*), U^A(a_2; \pi^*). \qquad [2]$$

That is, under PFF, P responds only to signals that affect her own, prospective fitness, not to the actions of A that can be inferred from the signals, and this "natural" feedback is enough to sustain observed cooperation. For example, in vertically transmitted symbioses, a PFF response would simply be the reduction in offspring in a host lineage that contains a virulent parasite (i.e., no evolved punishment), which would favor other host lineages containing less virulent or even mutualistic symbionts and producing more host offspring (21). [As a counterexample, if the prospectively optimal host response is to tolerate any damage caused by a parasite (31), natural feedback would be weak, by definition, and PFF would not be sufficient to maintain mutualism (21).]

On the other hand, to say that a_1 *is disciplined by HS* is to insist that P has evolved to infer actions from signals and carry out appropriate responses, which in the case of a_1 would be a punishment. This is because, despite any natural feedback,

$$U^A(a^*; \pi^*) < U^A(a_1; \pi^*). \qquad [3]$$

Evolutionary stability under HS then requires that the punishment necessary to maintain cooperation be administered in the least costly manner, a scheme derived by Holmström (32) in his theory of moral hazard.

Under what conditions would we expect HS to evolve? In *SI Text 1*, we extend a one-sided HS model (12, 13) to a two-sided, coevolutionary model and find two evolutionarily stable strategies (ESS): (*a*) HS are absent and (*b*) HS are maintained, but only when the symbiont population is a mix of cheaters and cooperators and only for a limited set of parameter values. For the system to reach (*b*), it would need somehow to escape the first ESS of no sanctions and no cooperative agents. Alternatively, HS could evolve if the host induces evolution in the symbiont, something that is possible but is more likely to occur in highly specific situations with strong spatial structure and/or hosts with generation times much longer than those of the symbiont (*SI Text 1*). If the most likely starting point for a symbiosis is a largely homogeneous population of initially nonmutualistic symbionts and noninvesting, nonsanctioning hosts, then costly punishment by hosts is unlikely to evolve de novo. An exogenous input of cheaters, such as a stably coexisting parasite

species, or biased mutation could maintain HS, but cannot explain the origin of HS (10).

In the absence of HS, PFF-enforced mutualism [Eq. **2**] is more likely to evolve and be maintained when f takes an exogenously high value because of the life cycle concordance of the partners [Eq. **1**], as is well known to occur under spatial structuring (11) or vertical transmission (33).

Horizontally transmitted mutualisms can also have high values of f and therefore evolve via PFF. For example, plants may limit, withdraw, or even abscise investments in shoots, flowers, and roots (or subsets of any of these) after physical damage or deficits of pollen or other resources (e.g., refs. 27, 29, 34–36, and included references). If a sufficient amount of symbiont fitness is also reduced after reducing investment in a plant part, the condition for PFF is met [Eq. **2**], and symbionts are selected to protect, pollinate, or fertilize to maintain plant investment.

An example can clarify. In ant–plant protection mutualisms, ants patrol vulnerable plant parts, protecting them from herbivores. Edwards et al. (27) experimentally showed in the ant–plant *Cordia nodosa* that if a new shoot suffers heavy leaf loss, the shoot is abscised. They inferred that if the ant fails to patrol effectively (a cheating action) and herbivores exploit the opportunity to consume emerging new leaves (a signal correlated with cheating), host plants will respond by abscising the damaged shoot. Because each new shoot also contains a new unit of ant housing, new shoot abscission reduces ant fitness. Furthermore, because the fitness cost of losing a unit of housing is greater than the gain from not patrolling (27), the optimal action for the ant is to patrol. Edwards et al. concluded that this ant–plant had evolved to use the amount of leaf-area loss to infer and respond to ant cheating behavior—in short, a host sanction.

However, under a PFF interpretation, shoot abscission after herbivory is not taken to be a de novo adaptation but as a kind of preadaptation, in the sense that plants generally are selected to abort parts that lose their value after environmental damage. What has (or at least could have) evolved is the response policy π, which in this system are the various probabilities of shoot abscission after different amounts of leaf loss. The prospectively optimal response π* balances the cost of maintaining a damaged plant part against the cost of abscission per se and any opportunity cost of the lost contribution by the remaining shoot to future growth, all without any consideration of the effect of abscission on ant fitness. In *C. nodosa*, f takes a high value because housing units are modular and located "in harm's way" on new shoots. Interestingly, in most ant–plant protection mutualisms, housing is distributed in this way, and it has been hypothesized that this pattern of placement reflects the plant locations where ants preferentially tend Hemiptera and might therefore be most likely to be encountered (37). As a result, the high value of f in ant–plants is likely exogenous.

In summary, the evolution of PFF is constrained by the requirement for a high value of f, and the evolution of HS is constrained because of the requirement for heterogeneity in symbionts or strong spatial structure (*SI Text 1*). Despite these different evolutionary pathways, differentiating PFF from HS in practice is difficult. In both cases, a real (or experimentally simulated) cheating behavior is followed by a host response that outweighs the short-term benefit of cheating to the symbiont. This conflation arises from the fact that, even were it possible, it would be costly and time-consuming to measure the long-term value of f and the mutualists' fitness functions with sufficient precision to directly test the theories. Thus, we use very general properties of Holmström's (32) solution for optimal punishment to design two experimental tests to distinguish PFF and HS with much lower and more realistic informational requirements.

Theorem 1: Do Principals Respond to Inferred Actions? Suppose there are two signals s and s' with the property that $U^P(s, \pi) + k = U^P(s', \pi)$, where $k \in \mathbb{R}$ is a constant for all π. In other words, two

signals might lead to different fitnesses but not in ways that depend on the response chosen. Then, under PFF we should observe the same response following these two signals: $\pi^*(s) = \pi^*(s')$. Furthermore, suppose that $p(s|a^*) = p(s|a_1) = p(s|a_2)$ but $p(s'|a^*) < p(s'|a_1)$. Then, under HS, we should observe strictly higher punishment following s' than s: $\pi_{HS}(s) > \pi_{HS}(s')$.

Intuitively, if the principal punishes signals that differ only in terms of the indication they give of cheating, this result provides support for HS; however, if the principal does not punish such actions differently, it supports PFF. Thus, if we can find two events ("actions") that happen to the principal that we are confident have the same effect on its prospective fitness after any given response, PFF predicts that these should lead to the same response. If additionally one signal indicates cheating whereas the other does not, HS is inconsistent with this prediction, instead calling for more punishment conditional on the signal indicating cheating.

Theorem 2: Are Agents Punished Minimally to Deter Cheating? Suppose that $u^s(a_2) \geq u^s(a_1)$ and

$$\sum_{i=1}^{N} u^P(s_i, \pi_{HS}[s_i]) p(s_i|a_2) \geq \sum_{i=1}^{N} u^P(s_i, \pi_{HS}[s_i]) p(s_i|a_1).$$

That is, a_2 gives at least as much short-term fitness advantage to A as does a_1 and, under observed punishment rules, at least as high principal fitness. Then HS predicts that $\sum_i \pi_{HS}(s_i) p(s_i|a_2) \geq \sum_i \pi_{HS}(s_i) p(s_i|a_1)$: a_2 is punished, on average, at least as much as a_1. If either of the initial inequalities is strict, so is the punishment inequality.

Furthermore, suppose that $\sum_i \pi^*(s_i) p(s_i|a_2) < \sum_i \pi^*(s_i) p(s_i|a_1)$ so that, in terms of prospective fitness, punishment is less attractive following a_2 on average than it is following a_1. Then PFF makes the opposite prediction from HS: less punishment on average following a_2 than following a_1.

Intuitively, if a_2 brings as great cheating benefits and leaves the principal as healthy, then it will be at least as attractive a cheating option as a_1 in the absence of necessary punishment. Because HS disciplines a_1, a_1 is punished no more than is necessary to prevent cheating. If a_2 is more tempting than a_1, it must be punished more than a_1, as otherwise a_2 will be punished less than necessary to avoid cheating.

But if a_2 on average reduces principal fitness less than does a_1, then the severity of natural feedback (PFF) will be automatically less following a_2.

The proofs of these theorems are given in *SI Text 2*.

By identifying appropriate pairs of actions, *without any further knowledge of the fitness function or the value of f*, we can use these tests to infer whether PFF or HS maintains cooperation in two well-known mutualisms in which hosts are thought to punish cheater symbionts. The first is the mutualism between legume plants and rhizobia bacteria, which nodulate plant roots and expend energy to "fix" atmospheric nitrogen (N_2) into a form (NH_4^+) that the plant can use to build proteins (38). In return, the plant provides carbohydrates to the rhizobia. Some rhizobia genotypes cheat by fixing little or no N_2 and accumulating energy reserves for reproduction instead (38). Some authors have argued (1, 9, 12, 13, 25, 26) that HS disciplines cheating in rhizobia. When soybean plant nodules were exposed to an atmosphere in which N_2 had been replaced by Ar, thereby preventing nitrogen fixation, nodules grew to smaller sizes and supported fewer rhizobia than did controls exposed to the normal N_2-rich atmosphere. Kiers et al. (25) assumed that plants cannot, on a nodule-by-nodule basis, measure the concentration of N_2 and thus cannot distinguish cheating by rhizobia from a lack of N_2. If this assumption is true, then *Theorem 1* can be used to distinguish between PFF and HS.

Let s be the signal of low N_2 fixation in the argon experiment, indicating cheating under the above assumption. Let s' be the signal

from a second experiment that reduces the value of rhizobia to the host plant by the same amount as the argon experiment but could not possibly indicate cheating. Such a signal is provided by experiments in which the nitrogen requirements of host plants are fully met via fertilizers. PFF predicts equal reduction in rhizobia populations, whereas HS predicts greater punishment in the argon treatment. If we compare what happens when soybean root systems are (*i*) exposed to an $Ar:O_2$ atmosphere that permits little N_2 fixation or (*ii*) grown in high levels of nitrates, almost the same, reduced rhizobial densities are reported (39), which supports PFF. We emphasize two caveats. In the published experiment (39), we assumed that the level of added nitrates fully met the soybean's short-term nitrogen requirements. This might not be strictly true. Second, the fact that legumes are often infected by multiple strains of rhizobia remains a problem, because our model assumes only one type of agent.

The second mutualism is that between yucca plants and yucca moths, which actively pollinate yucca flowers but also lay eggs so that their larvae can consume the developing seeds. Selection should favor moths that lay more eggs, at the cost of seeds, which results in a conflict of interest between plant and moth. Pellmyr and Huth (24) documented that *Yucca filamentosa* selectively abscises flowers with high egg loads (and low pollen loads). Again, using *Theorem 1*, let *s* be the signal of high egg load. Let *s'* be the signal from a second experiment (35), in which pins are used to mimic the puncture damage to ovules caused by moth oviposition, while omitting the eggs themselves. PFF predicts equal levels of floral abscission across the two experiments, but HS predicts greater punishment when eggs are present. Marr and Pellmyr (35) found that mechanical damage alone was sufficient to trigger levels of floral abscission consistent with naturally observed levels (24) and that the plant did not appear to react to the experimental application of moth eggs. Ovule damage is also correlated with the probability of selective abortion in an independently evolved pollination mutualism between *Glochidion* trees and *Epicephala* moths (40). Note that *Theorem 1* requires that the yucca plant be able to distinguish between egg oviposition and mechanical damage; otherwise *s'* would indicate cheating to the host, and the test would not be valid.

However, the yucca system also provides an example of *Theorem 2*. *Yucca kanabensis* is associated with two moth species, one of which oviposits shallowly in the carpel wall and does not damage ovules (41). "Shallow" species execute a_2 because they achieve equal or greater cheating benefits (on average, slightly more surviving larvae per locule) relative to "deep" species, which execute a_1 by also damaging ovules during oviposition. Furthermore a_2 results in an equally fit principal ("equivalent numbers of mature, viable seeds") (ref. 41, p 200). Consistent with PFF, *Y. kanabensis* selectively abscises only deep species (41) whose damage to ovules directly degrades flower value. Under HS, the host plant would abscise flowers oviposited on by both moth species because both actions provide similar benefits to the respective moth species. Consistent with PFF, selective abortion of flowers is thought to be a "symplesiomorphy within the Agavaceae, and thus a preadaptation within the yuccas" (ref. 34, p 558).

Discussion

Our analysis suggests that well-known examples of mutualism, previously used as examples of HS, are probably more consistent with PFF. We conclude that PFF is the more general mechanism for maintaining costly mutualistic behavior in symbionts, at least in plant hosts, and can be selected for even under horizontal transmission. However, PFF places important but underappreciated constraints on the evolution of mutualism, in that PFF requires the appropriate preadaptations or transmission mechanisms. HS, on the other hand, seems more likely to evolve when symbionts have higher cognitive functions (42), so that the influence of sanctions on symbiont behavior can take place in ecological, rather than evolutionary, time.

Concepts such as PFF and pseudoreciprocity have been criticized as being "empty" and "redundant" because of the lack of a formal definition and ways to discriminate them from other mechanisms that maintain cooperation (43, 44). In this paper, we have started to address these criticisms. More generally, biological mutualisms and economic contract theory form a surprisingly natural marriage of data and theory that provides additional directions for empirical studies and the possibility of further insights into the evolution of cooperation.

1. Sachs JL, Mueller UG, Wilcox TP, Bull JJ (2004) The evolution of cooperation. *Q Rev Biol* 79:135–160.
2. Lehmann L, Keller L (2006) The evolution of cooperation and altruism—a general framework and a classification of models. *J Evol Biol* 19:1365–1376.
3. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563.
4. Bergmüller R, Johnstone RA, Russell AF, Bshary R (2007) Integrating cooperative breeding into theoretical concepts of cooperation. *Behav Processes* 76:61–72.
5. West SA, Griffin AS, Gardner A (2007) Evolutionary explanations for cooperation. *Curr Biol* 17:R661–R672.
6. Doebeli M, Knowlton N (1998) The evolution of interspecific mutualisms. *Proc Natl Acad Sci USA* 95:8676–8680.
7. Ferrière R, Bronstein JL, Rinaldi S, Law R, Gauduchon M (2002) Cheating and the evolutionary stability of mutualisms. *Proc Biol Sci* 269:773–780.
8. Ferrière R, Gauduchon M, Bronstein JL (2007) Evolution and persistence of obligate mutualists and exploiters: Competition for partners and evolutionary immunization. *Ecol Lett* 10:115–126.
9. Foster KR, Wenseleers T (2006) A general model for the evolution of mutualisms. *J Evol Biol* 19:1283–1293.
10. Foster KR, Kokko H (2006) Cheating can stabilize cooperation in mutualisms. *Proc Biol Sci* 273:2233–2239.
11. Szilágyi A, Scheuring I, Edwards DP, Orivel J, Yu DW (2009) The evolution of intermediate castration virulence and ant coexistence in a spatially structured environment. *Ecol Lett* 12:1306–1316.
12. West SA, Kiers ET, Simms EL, Denison RF (2002) Sanctions and mutualism stability: Why do rhizobia fix nitrogen? *Proc Biol Sci* 269:685–694.
13. West SA, Kiers ET, Pen I, Denison RF (2002) Sanctions and mutualism stability: When should less beneficial mutualists be tolerated? *J Evol Biol* 15:830–837.
14. Ferdy JB, Després L, Godelle B (2002) Evolution of mutualism between globeflowers and their pollinating flies. *J Theor Biol* 217:219–234.
15. Yu DW, et al. (2004) Oviposition strategies, host coercion and the stable exploitation of figs by wasps. *Proc Biol Sci* 271:1185–1195.
16. Bronstein JL, Alarcón R, Geber M (2006) The evolution of plant-insect mutualisms. *New Phytol* 172:412–428.
17. Douglas AE (2008) Conflict, cheats and the persistence of symbioses. *New Phytol* 177:849–858.
18. Connor RC (1986) Pseudoreciprocity: Investing in mutualism. *Anim Behav* 34:1562–1566.
19. Frederickson ME, Gordon DM (2009) The intertwined population biology of two Amazonian myrmecophytes and their symbiotic ants. *Ecology* 90:1595–1607.
20. Jeon KW (1972) Development of cellular dependence on infective organisms: Micrurgical studies in amoebas. *Science* 176:1122–1123.
21. Yamamura N (1993) Vertical transmission and evolution of mutualism from parasitism. *Theor Popul Biol* 44:95–109.
22. Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35–57.
23. Bull JJ, Rice WR (1991) Distinguishing mechanisms for the evolution of co-operation. *J Theor Biol* 149:63–74.
24. Pellmyr O, Huth CJ (1994) Evolutionary stability of mutualism between yuccas and yucca moths. *Nature* 372:257–260.
25. Kiers ET, Rousseau RA, West SA, Denison RF (2003) Host sanctions and the legume-rhizobium mutualism. *Nature* 425:78–81.
26. Simms EL, et al. (2006) An empirical test of partner choice mechanisms in a wild legume-rhizobium interaction. *Proc Biol Sci* 273:77–81.
27. Edwards DP, Hassall M, Sutherland WJ, Yu DW (2006) Selection for protection in an ant-plant mutualism: host sanctions, host modularity, and the principal-agent game. *Proc Biol Sci* 273:595–602.

ECONOMIC SCIENCES

EVOLUTION

28. Bever JD, Richardson SC, Lawrence BM, Holmes J, Watson M (2009) Preferential allocation to beneficial symbiont with spatial structure maintains mycorrhizal mutualism. *Ecol Lett* 12:13–21.
29. Jandér KC, Herre EA (2010) Host sanctions and pollinator cheating in the fig tree-fig wasp mutualism. *Proc Biol Sci* 277:1481–1488.
30. Leimar O, Connor RC (2003) By-product benefits, reciprocity and pseudoreciprocity in mutualism. *Genetic and Cultural Evolution of Cooperation*, ed Hammerstein P (MIT Press, Cambridge, MA), pp 203–222.
31. Edwards DP, Yu DW (2008) Tolerating castration by hiding flowers in plain sight. *Behav Ecol Sociobiol* 63:95–102.
32. Holmström B (1979) Moral hazard and observability. *Bell J Econ* 10:74–91.
33. Sachs JL, Wilcox TP (2006) A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc Biol Sci* 273:425–429.
34. Pellmyr O, Huth CJ (1995) Differential abortion in the yucca. *Nature* 376:558–559.
35. Marr DL, Pellmyr O (2003) Effect of pollinator-inflicted ovule damage on floral abscission in the yucca-yucca moth mutualism: The role of mechanical and chemical factors. *Oecologia* 136:236–243.
36. Hodge A (2004) The plastic plant: Root responses to heterogeneous supplies of nutrients. *New Phytol* 162:9–24.
37. Benson WW (1985) Amazon ant-plants. *Amazonia*, eds Prance GT, Lovejoy TE (Pergamon, Oxford), pp 239–266.
38. Kiers ET, Denison RF (2008) Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Annu Rev Ecol Evol Syst* 39:215–236.
39. Kiers ET, Denison RF (2006) Measured sanctions: Legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. *Evol Ecol Res* 8:1077–1086.
40. Goto R, Okamoto T, Kiers ET, Kawakita A, Kato M (2010) Selective flower abortion maintains moth cooperation in a newly discovered pollination mutualism. *Ecol Lett* 13:321–329.
41. Addicott JF, Bao T (1999) Limiting the costs of mutualism: Multiple modes of interaction between yuccas and yucca moths. *Proc Biol Sci* 266:197–202.
42. Bshary R, Grutter AS (2002) Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. *Anim Behav* 63:547–555.
43. Buston PM, Balshine S (2007) Cooperating in the face of uncertainty: A consistent framework for understanding the evolution of cooperation. *Behav Processes* 76: 152–159.
44. Mesterton-Gibbons M, Dugatkin LA (1997) Cooperation and the Prisoner's dilemma: Towards testable models of mutualism versus reciprocity. *Anim Behav* 54:551–557.
45. Becker GS (1968) Crime and punishment: An economic approach. *J Polit Econ* 76:169–217.