Mutation enhances cooperation in direct reciprocity

Josef Tkadlec[a] (ID), Christian Hilbe[b,1] (ID), and Martin A. Nowak[a,c,1,2] (ID)

Direct reciprocity is a powerful mechanism for the evolution of cooperation based on repeated interactions between the same individuals. But high levels of cooperation evolve only if the benefit-to-cost ratio exceeds a certain threshold that depends on memory length. For the best-explored case of one-round memory, that threshold is two. Here, we report that intermediate mutation rates lead to high levels of cooperation, even if the benefit-to-cost ratio is only marginally above one, and even if individuals only use a minimum of past information. This surprising observation is caused by two effects. First, mutation generates diversity which undermines the evolutionary stability of defectors. Second, mutation leads to diverse communities of cooperators that are more resilient than homogeneous ones. This finding is relevant because many real-world opportunities for cooperation have small benefit-to-cost ratios, which are between one and two, and we describe how direct reciprocity can attain cooperation in such settings. Our result can be interpreted as showing that diversity, rather than uniformity, promotes evolution of cooperation.

evolution of cooperation | direct reciprocity | donation game | mutation rate

In evolutionary game theory, cooperation is an action in which an individual voluntarily incurs a cost to give a benefit to someone else. While socially beneficial, cooperation is opposed by natural selection unless a mechanism for evolution of cooperation is in place (1, 2). One such mechanism is direct reciprocity: When the same two individuals interact repeatedly, mutual cooperation becomes a viable option (3–16).

The phenomenon of cooperation can be described quantitatively by the donation game (17), which is a simplified prisoner's dilemma (18). In the donation game, each of two players chooses between cooperation and defection (Fig. 1A). Cooperation means paying a cost, $c > 0$, for the other player to receive a benefit, $b > c$. Defection incurs no cost and causes no benefit. When players only interact for one round of the donation game, evolutionary game theory predicts that both players learn to defect (19). This prediction, however, changes when the game is repeated. In that case, individuals can react to their coplayer's previous behavior. They can employ conditionally cooperative strategies, such as Grim-Trigger (17), Tit-for-Tat (18), Generous Tit-for-Tat (20, 21), or Win–Stay Lose–Shift (7, 8) to incentivize their coplayer to cooperate. With such conditional strategies, mutual cooperation can be sustained as a Nash equilibrium.

Full cooperation, however, is not the only possible equilibrium outcome of the repeated donation game. On the contrary, the so-called Folk theorem guarantees the existence of a multitude of equilibria with all possible levels of cooperation, provided that each player gets at least the payoff for mutual defection (22). For example, in addition to cooperating in every round, there are equilibria in which players defect unconditionally, or in which they alternate between cooperation and defection (23).

Because there are many equilibria, it becomes natural to ask which equilibrium emerges in populations of evolving players. This question can be explored with computer simulations of stochastic evolutionary dynamics (24–26). In these simulations, players can choose among many different strategies for the repeated interaction. Over time, they abandon strategies that yield inferior payoffs and instead adopt strategies that perform comparably well. By analyzing the resulting evolutionary trajectories, researchers explore how likely players learn to cooperate and which strategies they eventually use.

The results of these individual-based simulations depend on a number of parameters, which include the benefit-to-cost ratio, the population's size, the intensity of selection, and the mutation rate. The latter specifies how often players randomly explore new strategies. The values of these parameters not only affect whether or not cooperation evolves but also how long it takes for populations to converge and whether or not analytical approximations are feasible. One way to minimize computation time is to assume that mutations are exceedingly rare (27–29). In this limit, populations are homogeneous most of the time. Only occasionally a mutant strategy arises, and this mutant either fixes in the population or goes extinct before the next mutation occurs.

Significance

Evolution is not only fierce competition but also cooperation. The latter can be seen as the main architect of biological complexity. Cooperation means that one individual pays a cost for another individual to receive a benefit. Cooperation is opposed by natural selection unless a mechanism for evolution of cooperation is at work. One such mechanism is direct reciprocity which utilizes repeated interactions between the same individuals. Here, we report that intermediate mutation rates, which lead to diverse communities of cooperators, dramatically enhance the levels of cooperation in competitive settings. Therefore, mutation "cooperates" with cooperation to maximize its beneficial effect on life.

Author affiliations: [a]Department of Mathematics, Harvard University, Cambridge, MA 02138; [b]Max Planck Research Group 'Dynamics of Social Behavior', Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany; and [c]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

[1]C.H. and M.A.N. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: martin_nowak@harvard.edu.
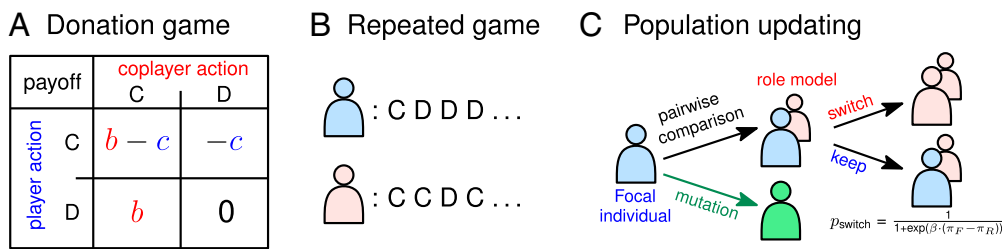
Fig. 1. Evolutionary dynamics of cooperation. (*A*) In the donation game, players choose between cooperation and defection. Cooperation incurs a cost c and provides a benefit b to the coplayer. Defection incurs no cost and provides no benefit. (*B*) We consider the repeated donation game, in which players can use conditional strategies that depend on the outcome of previous rounds. (*C*) In each evolutionary update step, a focal individual, F, either explores a new random strategy (with probability u) or compares his own payoff to that of a random role model, R. He is then more likely to switch to the role model's strategy if she performs better than him.

Since there is an explicit formula for the mutant's fixation probability (24), simulations that make use of the rare-mutation assumption tend to be many orders of magnitudes faster than conventional simulations (26). In addition, simulation results can be interpreted more easily when mutations are rare because the evolving population compositions are often closely connected to the Nash equilibria of the game (23). Due to these advantages, the rare-mutations assumption has become a standard approach to explore the evolution of direct reciprocity (30–41).

However, there is by now substantial evidence from experimental games among human subjects suggesting that empirical mutation rates are sizeable (42). In particular, estimated mutation rates are far beyond the threshold for which the rare-mutation approximation is valid (43). This observation raises the question how individuals learn to engage in direct reciprocity when mutations occur more frequently. This is the question that we explore in the present paper.

If mutations are frequent, evolutionary dynamics lead to communities in which many different strategies coexist. Surprisingly, we find that such diverse communities facilitate cooperation. The previous literature based on rare mutations emphasized that cooperation can only evolve if the benefit-to-cost ratio, b/c, is large. For example, the well-known strategy Win–Stay Lose–Shift can only maintain cooperation if $b/c > 2$. Among reactive strategies (which only take into account the coplayer's previous action), it takes an even larger benefit-to-cost ratio for cooperation to evolve (44). Here, we show that once mutation rates are nonnegligible, high cooperation levels can occur even as the benefit-to-cost ratio approaches one, the theoretical minimum. To this end, we present extensive simulations for various strategy sets, including the set of stochastic reactive strategies, stochastic memory-1 strategies, and deterministic memory-2 strategies. In each case, we find that intermediate mutation rates are able to facilitate cooperation in parameter regions in which cooperative populations are unlikely to emerge otherwise. The emerging diverse communities destabilize the equilibrium around all out defectors, while keeping the cooperative equilibria stable. Thus, certain levels of diversity—rather than uniformity—promote cooperation.

Results

Model Framework. We consider a population of size N. Individuals engage in repeated donation games (Fig. 1). For most of the main text, we assume that individuals make their decision whether or not to cooperate based on reactive strategies (21). Such strategies consist of two parameters: p is the probability to cooperate if the coplayer has cooperated in the previous round, whereas q is the probability to cooperate if the coplayer has defected in the previous round. A strategy is called deterministic if all cooperation probabilities are either zero or one. If at least one probability is in between, the strategy is called stochastic. The space of stochastic reactive strategies is the unit square, and the deterministic strategies correspond to the four corners of this square. Reactive strategies include: always defect ALLD $= (0, 0)$, always cooperate ALLC $= (1, 1)$, the random strategy $(0.5, 0.5)$, tit-for-tat TFT $= (1, 0)$, and generous tit-for-tat GTFT $= (1, q)$, where $q \in [0, 1]$ is the probability of forgiveness or the level of generosity. We note that it is straightforward to compute the cooperation rate $C(S, S')$ for a reactive strategy S when facing a reactive strategy S'. This in turn also allows us to compute the resulting payoff $\pi(S, S')$, that is, the average benefit minus cost derived from that interaction (45). For details, see *SI Appendix*, Note 2. We focus on reactive strategies for simplicity; analogous results hold when players use memory-1 (*SI Appendix*, Fig. S1) or memory-2 strategies (*SI Appendix*, Fig. S9).

At any point in time, the composition of the population is described by a list of the employed strategies $(S_1, S_2, ..., S_N)$. Each individual i derives an average payoff $\pi_i = \sum_{j \neq i} \pi(S_i, S_j)/(N - 1)$ from all pairwise interactions. Individuals learn to adopt more profitable strategies over time. To describe the resulting dynamics, several processes have been proposed, including processes based on stochastic best responses (46) or based on stochastic imitation (47). Here, we use a pairwise comparison process (48), which is a variant of stochastic imitation. When a focal individual F revises their strategy, they can do so in two ways (Fig. 1*C*). With probability u, the focal individual adopts a new strategy at random, which represents mutation. With probability $1 - u$, the focal individual considers imitating the strategy of another population member, which corresponds to reproduction and selection. To this end, the focal player randomly picks a role model R from the population and compares its payoff, π_F, to that of the role model, π_R. The focal player switches to the role model's strategy with a probability given by the Fermi function, $1/[1 + \exp(\beta \cdot (\pi_F - \pi_R))]$; otherwise, the focal individual keeps their strategy. The parameter $\beta \geq 0$ measures the intensity of selection. It reflects how clearly payoffs can be evaluated. If $\beta = 0$, the imitation probability simplifies to one-half, meaning that payoffs become irrelevant and we are in the realm of neutral evolution. In the other limit $\beta \to \infty$, the focal player adopts the role model's strategy only if $\pi_R \geq \pi_F$, and selection becomes very strong. For most our simulations, we choose $\beta = 10$, which represents an intermediate intensity of selection.

We consider two versions of this process. For our simulations, we consider a Wright–Fisher type model with nonoverlapping generations (49): in each time step, all players are given the opportunity to revise their strategy. For analytical calculations, we complement this framework with a Moran-type model (24), in which in each step only one randomly chosen player can revise
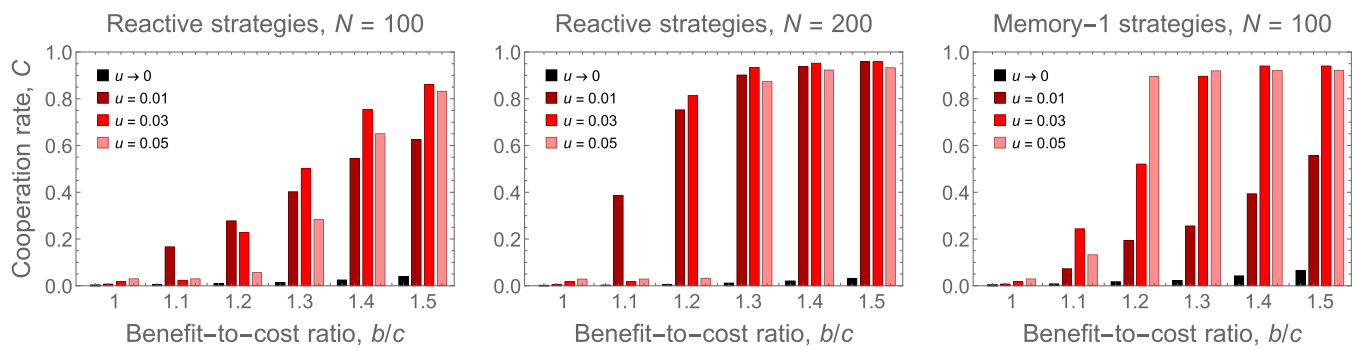
Fig. 2. The effect of diversity on cooperation. In the limit of rare mutations ($u \rightarrow 0$, black bars), the average cooperation rate increases very slowly with the benefit-to-cost ratio. Adding mutation (red bars) substantially enhances cooperation for small benefit-to-cost ratios. Parameters: $\beta = 10$; simulations are run for at least 10^9 updates to get reliable averages.

their strategy. In both cases, we obtain similar results, but the model with nonoverlapping generations is computationally more efficient. Our simulations start out with random populations. Over time, players adopt new strategies according to the above process. The results depend on the benefit-to-cost ratio b/c, the mutation rate u, the population size N, and the intensity of selection β. In the following, we explore how these parameters affect evolutionary outcomes. The quantity of interest is the average cooperation rate C, with the average being taken over all pairwise interactions over sufficiently long time.

The Diversity Effect. In Fig. 2, we show how the average cooperation rate C depends on the benefit-to-cost ratio b/c. In the limit of rare mutations, $u \rightarrow 0$, the process is relatively well understood (50). In that setting, the cooperation rate increases only very slowly with rising benefit-to-cost ratio. But if we add mutation we find a dramatic increase in the cooperation rate. For population size $N = 100$, we find that mutation rates between $u = 0.01$ and $u = 0.03$ are ideal for reactive strategies, while slightly higher mutations rates (from $u = 0.03$ to $u = 0.05$) are ideal for memory-1 strategies. The optimum mutation rate depends on the exact value of b/c.

In the limit of rare mutation, $u \rightarrow 0$, the population is mostly homogeneous at any one time and is exploring the strategy space by making transitions between states that are dominated by single strategies. For larger mutation rates, communities are more diverse and, unexpectedly, this enhances the average cooperation rate. Our goal is to understand this surprising effect. Why does diversity promote cooperation? In the limit of vanishing mutation, all players update their strategies based on performance, but in the presence of mutation, some players choose randomly. Why do random choices augment cooperation?

Characteristic Curve and Optimum Mutation Rate. In the donation game, the benefit-to-cost ratio varies between one and infinity. Consequently, the cost-to-benefit ratio varies between zero and one. For a given mutation rate u and a population size N, we can plot the average cooperation rate in the population versus the cost-to-benefit ratio, c/b. We call this graph the characteristic curve of the evolutionary process (Fig. 3A). Since increasing the cost-to-benefit ratio makes cooperation less rewarding, all characteristic curves are expected to decline monotonically. For $u \rightarrow 0$, we find very low levels of cooperation if $c/b > 1/2$. For $u = 0.01$, high levels of cooperation are observed for population sizes $N = 100$ and 200 even if $c/b > 1/2$.

To examine the role of diversity in more detail, another perspective is useful. In Fig. 3B, we show the average cooperation rate C as a function of the mutation rate u. Proceeding from high ($u = 1$) to low mutation rates ($u \approx 10^{-4}$), we observe three trends: First, the cooperation rate declines from 0.5 toward near zero; subsequently, it rises suddenly to very high values (near one); and then, it declines again. We refer to those trends as the *valley* and the *hump*. The exact locations of the minimum of the valley and the maximum of the hump depend on the cost-to-benefit ratio and the population size. However, both the valley and the hump occur consistently across many combinations of parameters considered.

Stationary Distribution. To understand the two effects of valley and hump, we show in Fig. 4 how often each strategy in the space of all reactive strategies is used at various mutation rates. The heat maps suggest that two regions of the strategy space are visited predominantly. The first region corresponds to a set of defective strategies $(p, 0)$, with $0 \leq p \leq c/b$, including the strategy ALLD $= (0, 0)$. The second region consists of generous tit-for-tat strategies GTFT $= (1, q)$ with q satisfying $0 \leq q \leq 1 - c/b$. Both regions are favored by selection for all mutation rates: They are visited more often than expected under neutrality. However, the relative abundance of each region changes with u. For small mutation rates such as $u = 10^{-4}$, the region of defectors is most abundant. For larger mutation rates up to $u = 10^{-1}$, individuals predominantly use GTFT.

Detailed Analysis of a Reduced Strategy Set. To gain intuition for these findings, it is useful to consider a reduced strategy set. This set ought to be large enough to reproduce the above findings, yet small enough to be tractable in detail. In *SI Appendix*, Note 2, we show that a set with two strategies is not sufficient to reproduce all qualitative findings. Specifically, if individuals can only choose between ALLD $= (0, 0)$ and GTFT $= (1, q)$, we obtain neither the valley nor the hump (*SI Appendix*, Fig. S2). Intuitively, this two-strategy system fails to capture one key feature of the full system, namely that all strategies $(1, q)$ with $q \in [0, 1]$ are neutral with respect to each other. This neutrality implies that the mass of the blue peak in Fig. 4 can move freely along the edge $p = 1$ of the unit square $[0, 1]^2$.

To capture this effect, we study a three-strategy system consisting of ALLD, GTFT, and ALLC. For this system, we observe both the valley and the hump, independent of whether mutations introduce all three strategies equally often or whether they are biased toward ALLD, *SI Appendix*, Fig. S3.

To explain the valley and the hump in this reduced strategy space, we plot the relative abundance of each possible population composition. Any compositions can be depicted as a point
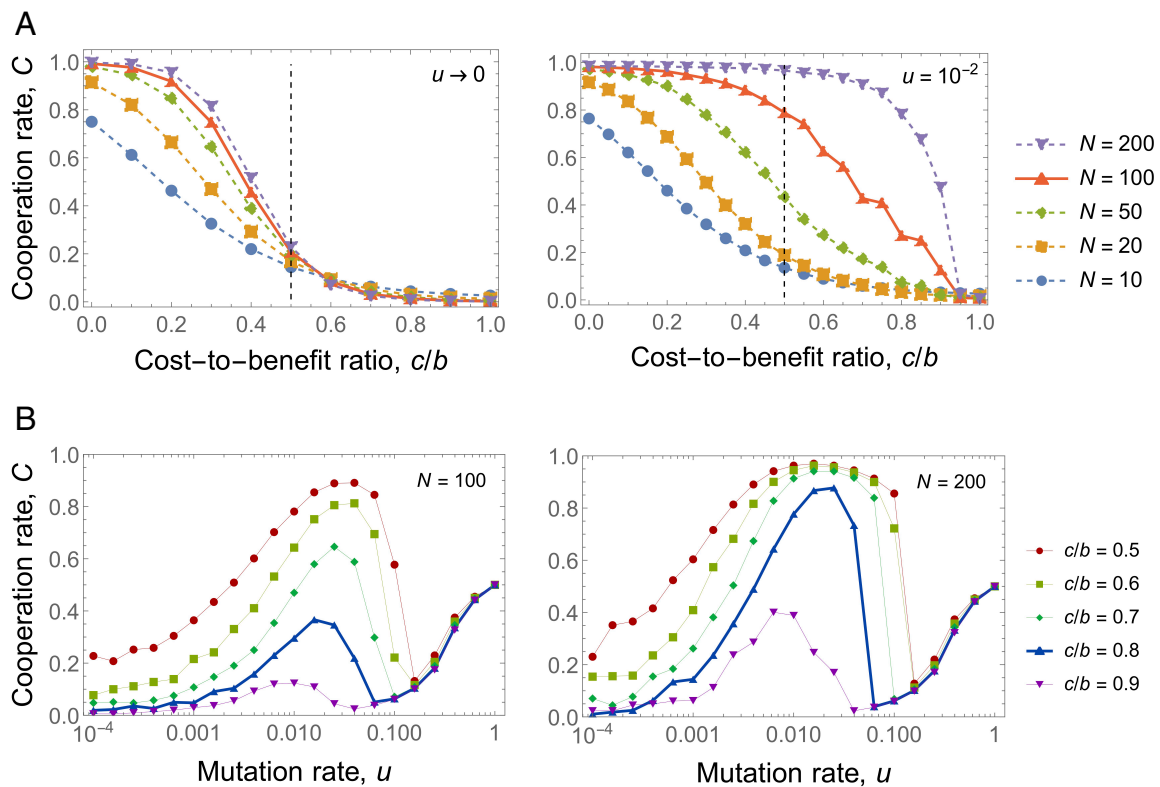
Fig. 3. Characteristic curve and the optimum diversity for reactive strategies. (A) The characteristic curve of an evolutionary process of cooperation and defection is the graph of average cooperation rate versus cost-to-benefit ratio. In the limit of rare mutations ($u \to 0$, left), no substantial cooperation occurs for $c/b > 0.5$. For mutation rate $u = 10^{-2}$ (right), we observe substantial levels of cooperation even if $c/b > 0.5$, especially when $N \geq 50$. (B) The average cooperation rate C as a function of the mutation rate u, for $N = 100$ (*Left*) and $N = 200$ (*Right*). When there are only mutations, $u = 1$, the cooperation rate approaches $C = 0.5$, regardless of the cost-to-benefit ratio c/b (Lemma 2 in *SI Appendix,* Note 2). As u goes down, the cooperation rate first drops toward zero, then jumps up toward one (here at around $u \approx 10^{-1}$), and finally, it drops toward zero again (here at around $u \approx 10^{-4}$). We call these phenomena the "valley" and the "hump". The optimum mutation rate is around $u \approx 10^{-2}$. We use $\beta = 10$.

in a simplex with corners ALLD, GTFT, ALLC. The relative frequency of each strategy is proportional to the distance of the point from the opposite side (Fig. 5A). Corners correspond to homogeneous populations, where all individuals play the same strategy.

The evolutionary dynamics on the simplex can be interpreted most easily in the limiting cases, when there are either only mutations ($u = 1$) or very rare mutations ($u \ll 1$). When there are only mutations, individuals use each of the three available strategies with equal probability. Since there is no selection, the law of large numbers implies that populations are concentrated around the center of the simplex (Fig. 5B). In such populations, the strategy receiving the highest payoff is ALLD. If u is slightly decreased, ALLD is favored by selection and becomes most frequent. Since ALLC is exploited by ALLD, the remaining population members are more likely to adopt GTFT rather than ALLC (Fig. 5C).

Looking at the other extreme, when u is very small (Fig. 5G–I), populations are typically homogeneous. That is, populations are at one of the three corners of the state space most of the time; they are less often on one of the edges; and they are least often in the interior. In this limit, selection makes the population oscillate around the simplex as follows (Fig. 5J). If the current population predominantly uses ALLD, then GTFT can invade with a certain probability (24). Once this happens, evolution leads toward a pure GTFT state. Because ALLD players are now absent, the strategies ALLC and GTFT are neutral with respect to each other. Thus, a pure GTFT population drifts toward a mix of GTFT and ALLC. Once the proportion of ALLC is high, the population

becomes susceptible to invasion by ALLD. This pattern is consistent with the behavior of the full process in the limit of rare mutations, $u \to 0$. In that limit, the population resolves to a pure state between every two consecutive mutations. As a consequence, the relative frequencies of the three pure states can be computed from the probabilities $p_{A \to B}$ that a single mutant with strategy B successfully invades and fixes in a resident population with strategy A, for $A, B \in \{\text{ALLD}, \text{GTFT}, \text{ALLC}\}$ (27). In particular, for large N (and any c and q), we show that the population spends virtually all the time at a pure ALLD state, Theorem 3 in the *SI Appendix,* Note 2.

For intermediate mutation rates, all three strategies are typically present in the population simultaneously, but selection events occur sufficiently often to drive inefficient strategies to low frequencies. As a result, we observe two peaks for intermediate u. One peak is near ALLD and the other one is near GTFT (Fig. 5D), which is consistent with the behavior observed for the full strategy space (Fig. 4). In this regime, mutations have two positive effects on the evolution of cooperation. On the one hand, they destabilize the peak around ALLD: because mutations recurrently introduce GTFT players, these GTFT players are more likely to reach a critical number after which cooperation is more profitable. On the other hand, mutations make the peak around GTFT more stable: because mutations recurrently introduce small minorities of defectors, GTFT yields a strictly better payoff than ALLC, which prevents ALLC from invading through neutral drift.

This situation is reminiscent of a counterintuitive effect in rock–paper–scissors games. In these games, providing a payoff
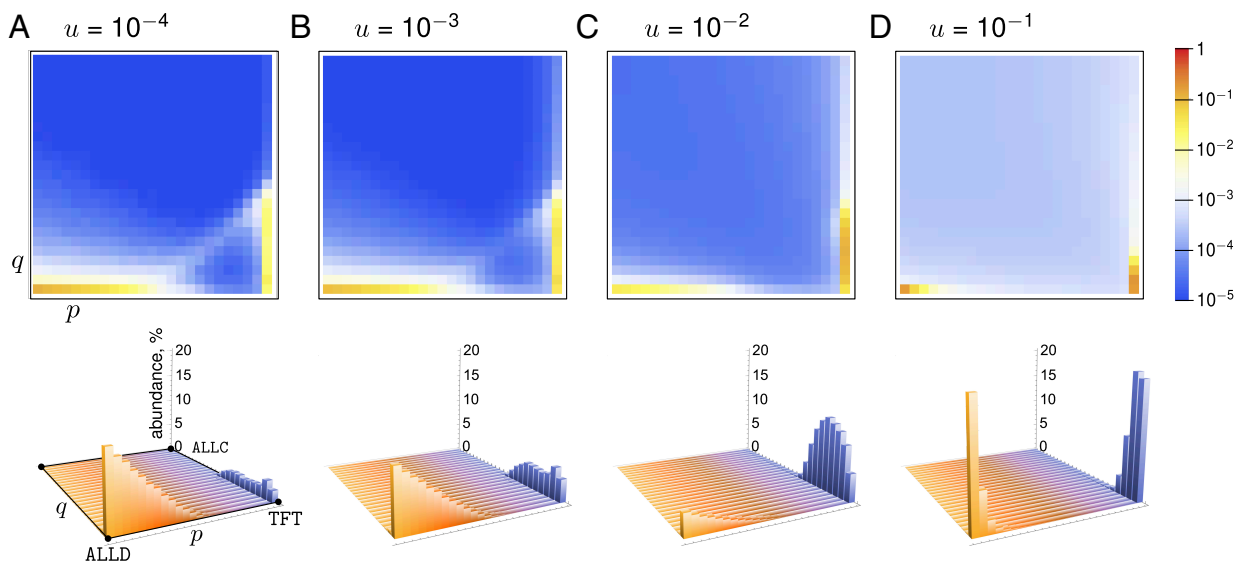
Fig. 4. Frequencies of reactive strategies. (*A–D*) For 4 distinct values of the mutation rate, $u = 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}$, we simulate evolutionary dynamics for at least 10^9 steps. We collect the appearing strategies into a 25×25 grid and plot their relative abundance as a heat map (*Top* row, log scale) and as a bar chart (*Bottom* row, linear scale). In all 4 cases, the strategies with nonnegligible frequency have either $q \approx 0$ (orange peak, *Bottom Left*) or $p \approx 1$ (blue peak, *Right*). For intermediate mutation rate, $u = 10^{-2}$ (third column), the blue peak has more mass than the orange one, and we observe high overall cooperation rates. Parameters: $N = 100$, $c/b = 0.5$, $\beta = 10$.

advantage to one strategy may eventually lower that strategy's frequency in equilibrium (51, 52). This counterintuitive effect could explain our findings, as higher mutation rates seem to give a payoff advantage to defectors. However, we believe that this explanation does not fully capture our results. First, in addition to (indirectly) affecting the payoffs of each strategy, mutations alter the evolutionary dynamics altogether. They provide the population with a larger pool of role models that can be imitated subsequently. Second, the dynamics between ALLD, GTFT, and ALLC does not follow a strict rock–scissors–paper cycle (Fig. 5*J*). Instead, the competition between ALLD and GTFT is bistable, whereas the competition between GTFT and ALLC is neutral. In each case, mutations favor cooperation: They help GTFT to overcome the invasion barrier (against ALLD) and to resist neutral drift (against ALLC).

We emphasize that unconditional ALLC players catalyze the transition from the cooperative equilibrium to the defective one

in the limit of rare mutation. If we instead consider the triplet {ALLD, GTFT, ALLD}, where ALLC is replaced by another copy of ALLD, then high levels of cooperation occur for all low enough mutation rates (*SI Appendix*, Fig. S4).

Beyond Reactive Strategies. To illustrate the robustness of our results, we also explored the space of stochastic memory-1 strategies (8). These strategies are given by four parameters p_{CC}, p_{CD}, p_{DC}, and p_{DD} that denote the probabilities to cooperate after each of the possible outcomes of the last round, CC, CD, DC, and DD, respectively. The space of memory-1 strategies is the hypercube, $[0, 1]^4$. The space includes all reactive strategies but also Win–Stay Lose–Shift and many other strategies. In *SI Appendix*, Fig. S1 and in Fig. 2, we show that our findings continue to hold in this broader space. In the limit of rare mutations, $u \to 0$, there is no substantial cooperation when the cost-to-benefit ratio exceeds 1/2. In contrast, for nonnegligible
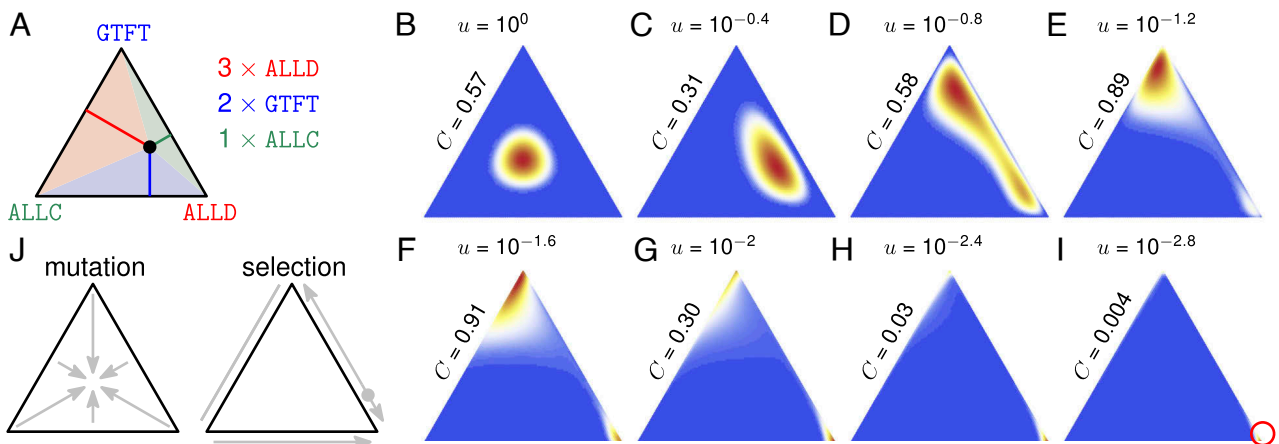

Fig. 5. Evolution of the reduced three strategy system $\mathcal{S}_3 = \{\text{ALLD}, \text{GTFT}, \text{ALLC}\}$. (*A*) Each point in the simplex represents a composition of the population. (*B–I*) Relative frequencies of the population composition (red is high) as u decreases from *B*, $u = 10^0 = 1$ to **i**, $u = 10^{-2.8}$. (*B*) For $u = 1$, the population typically consists of roughly 1/3 of each of ALLD, GTFT, and ALLC and the average cooperation rate is $5.1/9 \doteq 0.57$ (Theorem 1 in *SI Appendix*, Note 2). (*E–F*) For $u \in [10^{-1.2}, 10^{-1.6}]$, the population typically consists mostly of GTFT players and the cooperation rate is ≈ 0.9. (*H* and *I*) For $u \le 10^{-2.4}$, virtually all players play ALLD virtually all the time. (*J*) The effects of mutation and selection on the total mass. Parameters: $N = 100$, $c = 0.7$, $\beta = 10$. For GTFT we use $(1, 0.1)$.

mutation rates, cooperation does occur even if $c/b > 1/2$. Moreover, the average cooperation rate as a function of u again exhibits both the valley and the hump. In *SI Appendix*, Note 4, we report similar results for (deterministic) memory-two strategies (*SI Appendix*, Fig. S9). Overall, these findings suggest that the positive effects of mutations are not restricted to one particular strategy space.

Discussion

In repeated social dilemmas, individuals can sustain cooperation by reacting to their interaction partner's previous actions. This mechanism for cooperation is called direct reciprocity (3–16). For direct reciprocity, fully cooperative Nash equilibria exist if the benefit-to-cost ratio is greater than one, $b/c > 1$ (53). But there is still a range of other equilibria with lower levels of cooperation (23). In particular, full defection remains an equilibrium for all benefit-to-cost ratios. Therefore, it is a question of evolutionary dynamics which equilibrium is chosen and how much cooperation is achieved on average. When individuals react to their interaction partner's last decision, previous papers (33, 34, 44) suggest that high levels of cooperation only emerge if the benefit-to-cost ratio is greater than two, $b/c > 2$. Hence, the interval $2 > b/c > 1$ was void of cooperation. Here, we show how to attain high levels of cooperation for $b/c > 1$ in general, and thus for $2 > b/c > 1$ in particular.

Many previous studies of direct reciprocity explore selection dynamics either in the absence of mutation (54–61), or in the limit of very low levels of mutation (30–41) (for a more detailed discussion of the previous literature, *SI Appendix*, Note 1). In the context of evolutionary game theory, selection means to adopt strategies based on performance, while mutation means to adopt strategies randomly. We find that intermediate mutation rates allow high levels of cooperation even if benefit-to-cost ratios are only marginally above one. Mutations generate diverse communities which we find to have two beneficial effects for evolution of cooperation. First, diversity undermines defective equilibria by seeding clusters of potential invaders. Second, diversity stabilizes cooperative equilibria by preventing neutral drift toward strategies that can be exploited.

The importance of mutant strategies has also been highlighted in earlier studies of direct reciprocity (5, 12, 62). These studies stress that any resident population becomes unstable when interacting with an appropriate ensemble of mutants. Importantly, however, the reported effects of mutant strategies in these studies are symmetric. Just as there are mixtures of mutants that can invade an ALLD equilibrium, there are mutants that can invade WSLS (or any other cooperative resident strategy). In contrast, the effect of mutations that we report is asymmetric. We find that the variation introduced by mutations systematically favors cooperation. This effect is robust, and it arises independently of the strategy spaces we considered. Interestingly, the positive effect of variation does not even require this variation to be heritable. Indeed, further simulations reported in *SI Appendix*, Note 3 suggest that even phenotypic (nonheritable) variation in the player's behaviors can foster cooperation (*SI Appendix*, Fig. S8).

The essential question of evolution of cooperation is not only whether cooperation can evolve but for which benefit-to-cost ratio. The efficiency of a mechanism could be defined by the minimum benefit-to-cost ratio needed for the evolution of cooperation. We show that the efficiency of direct reciprocity is greatly enhanced by nonnegligible mutation rates. Remarkably, the path to cooperation identified in this paper does not require complex strategies with extended memory capacities, as proposed earlier (63–67). Instead, this path is already available to individuals with the most basic strategies of direct reciprocity. In our paper, cooperation does not evolve because individual strategies are sufficiently sophisticated. Instead, cooperation evolves because the evolutionary process is sufficiently erratic. Our observation is both surprising and relevant. It is surprising because one would not expect a-priori that replacing performance-based update events with random ones could promote cooperation. It is relevant because many opportunities for cooperation naturally arise in situations where the benefit-to-cost ratio is only slightly above one.

Materials and Methods

Here, we sketch the intuition behind the mathematical claims made in the main text. The full mathematical proofs appear in the *SI Appendix*, Note 2.

Reactive Strategies. A reactive strategy is given by a pair (p, q), where $p \in [0, 1]$ (resp. $q \in [0, 1]$) are the probabilities to cooperate in the next round, assuming that the coplayer has just cooperated (resp. defected). When two individuals employing reactive strategies s and s' play an infinitely repeated donation game, their payoffs $\pi(s, s')$, $\pi(s', s)$ derived from that interaction can be computed using a standard formula (45). The same applies to the pairwise cooperation rates $C(s, s')$, $C(s', s)$, Lemma 1 in the *SI Appendix*, Note 2.

Case $u = 1$ (No Selection). When $u = 1$ then selection does not play any role and at each point in time, each individual plays a strategy selected uniformly at random from the space $\mathcal{S}$ of available strategies. As a consequence, when the space $\mathcal{S}$ consists of finitely many strategies, by linearity of expectation, the overall cooperation rate can be computed by averaging the pairwise cooperation rates among the strategies, Theorem 1. Moreover, when the space $\mathcal{S}$ includes all reactive strategies, we show that the average cooperation rate of the population is precisely $1/2$. This is because of a certain symmetry between cooperation and defection, Lemma 2.

Case $u \to 0$ (Rare Mutations). In the regime of rare mutations ($u \to 0$), we obtain a simplified process studied by Fudenberg and Imhof (27), as well as Imhof and Nowak (50). In that regime, the population resolves to a pure (homogeneous) state between every two consecutive mutations. When the space $\mathcal{S}$ of available strategies is finite, the long-term fate of the population can be characterized by a frequency vector $\mathbf{f} = (f_S \mid S \in \mathcal{S})$ that records, for each strategy S, the relative proportion f_S of time for which everyone in the population employs strategy S. This frequency vector can be found in time that is polynomial in the population size N and in the size k of the strategy space, by computing all the pairwise fixation probabilities $\{\rho^N(S, S') \mid S, S' \in \mathcal{S}\}$ and solving for the stationary distribution of the underlying Markov chain (27). Given the frequency vector, it is straightforward to compute the overall average cooperation rate C as $C = \sum_{S \in \mathcal{S}} f_S \cdot C(S, S)$.

In the *SI Appendix*, Note 2, we show that for large population sizes N, each of the pairwise fixation probabilities $\rho^N(S, S')$ is either very small (namely exponentially small in N) or quite large (namely at least inversely proportional to N), see Lemma 3. This allows us to argue that for the reduced strategy space $\mathcal{S}_3$, the cooperation rate C tends to either 0 or to 1, as the population size N grows large. The intuition is that for large N, one of the entries f_S in the frequency vector $\mathbf{f}$ tends to 1, whereas all other entries tend to 0. Thus, the overall cooperation rate tends to the pairwise cooperation rate $C(S, S)$ of this most frequent strategy against itself. Theorem 2 and Theorem 3 for details.

1. M. A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
2. B. Skyrms, *Evolution of the Social Contract* (Cambridge University Press, 2014).
3. R. L. Trivers, The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971).
4. R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, NY, 1984).
5. R. Boyd, J. Lorberbaum, No pure strategy is evolutionary stable in the iterated prisoner's dilemma game. *Nature* **327**, 58–59 (1987).
6. R. Boyd, Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game. *J. Theor. Biol.* **136**, 47–56 (1989).
7. D. P. Kraines, V. Y. Kraines, Pavlov and the prisoner's dilemma. *Theory Decision* **26**, 47–79 (1989).
8. M. A. Nowak, K. Sigmund, A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* **364**, 56–58 (1993).
9. T. Killingback, M. Doebeli, N. Knowlton, Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proc. R. Soc. B* **266**, 1723–1728 (1999).
10. T. Killingback, M. Doebeli, The continuous Prisoner's Dilemma and the evolution of cooperation through reciprocal altruism with variable investment. *Am. Nat.* **160**, 421–438 (2002).
11. I. Fischer *et al.*, Fusing enacted and expected mimicry generates a winning strategy that promotes the evolution of cooperation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10229–10233 (2013).
12. J. García, M. van Veelen, In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *J. Econ. Theory* **161**, 161–189 (2016).
13. E. Akin, Good strategies for the iterated prisoner's dilemma: Smale vs. Markov. *J. Dyn. Games* **4**, 217–253 (2017).
14. C. Hilbe, K. Chatterjee, M. A. Nowak, Partners and rivals in direct reciprocity. *Nat. Hum. Behav.* **2**, 469–477 (2018).
15. N. Glynatsi, V. Knight, Using a theory of mind to find best responses to memory-one strategies. *Sci. Rep.* **10**, 1–9 (2020).
16. N. Glynatsi, V. Knight, A bibliometric study of research topics, collaboration and centrality in the field of the Iterated Prisoner's Dilemma. *Hum. Soc. Sci. Commun.* **8**, 45 (2021).
17. K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, NJ, 2010).
18. A. Rapoport, A. M. Chammah, *Prisoner's Dilemma* (University of Michigan Press, Ann Arbor, 1965).
19. M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
20. P. Molander, The optimal level of generosity in a selfish, uncertain environment. *J. Conf. Res.* **29**, 611–618 (1985).
21. M. A. Nowak, K. Sigmund, Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
22. J. Friedman, A non-cooperative equilibrium for supergames. *Rev. Econ. Stud.* **38**, 1–12 (1971).
23. A. J. Stewart, J. B. Plotkin, Collapse of cooperation in evolving games. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17558–17563 (2014).
24. M. A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).
25. J. García, M. van Veelen, No strategy can win in the repeated prisoner's dilemma: Linking game theory and computer simulations. *Front. Robot. AI* **5**, 102 (2018).
26. L. Hindersin, B. Wu, A. Traulsen, J. García, Computation and simulation of evolutionary game dynamics in finite populations. *Sci. Rep.* **9**, 6946 (2019).
27. D. Fudenberg, L. A. Imhof, Imitation processes with small mutations. *J. Econ. Theory* **131**, 251–262 (2006).
28. B. Wu, C. S. Gokhale, L. Wang, A. Traulsen, How small are small mutation rates? *J. Math. Biol.* **64**, 803–827 (2012).
29. A. McAvoy, Comment on "Imitation processes with small mutations". *J. Econ. Theory* **159**, 66–69 (2015).
30. M. A. Nowak, K. Sigmund, The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl. Math.* **20**, 247–265 (1990).
31. A. J. Stewart, J. B. Plotkin, The evolvability of cooperation under local and non-local mutations. *Games* **6**, 231–250 (2015).
32. J. G. Reiter, C. Hilbe, D. G. Rand, K. Chatterjee, M. A. Nowak, Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nat. Commun.* **9**, 555 (2018).
33. C. Hilbe, M. A. Nowak, K. Sigmund, The evolution of extortion in iterated prisoner's dilemma games. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6913–6918 (2013).
34. A. J. Stewart, J. B. Plotkin, From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15348–15353 (2013).
35. A. J. Stewart, T. L. Parsons, J. B. Plotkin, Evolutionary consequences of behavioral diversity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7003–E7009 (2016).
36. K. Donahue, O. Hauser, M. Nowak, C. Hilbe, Evolving cooperation in multichannel games. *Nat. Commun.* **11**, 3885 (2020).
37. L. Schmid, K. Chatterjee, C. Hilbe, M. Nowak, A unified framework of direct and indirect reciprocity. *Nat. Hum. Behav.* **5**, 1292–1302 (2021).
38. P. S. Park, M. A. Nowak, C. Hilbe, Cooperation in alternating interactions with memory constraints - source code and data. *Nat. Commun.* **13**, 737 (2022).
39. S. Kurokawa, Y. Ihara, Emergence of cooperation in public goods games. *Proc. R. Soc. B* **276**, 1379–1384 (2009).
40. S. van Segbroeck, J. M. Pacheco, T. Lenaerts, F. C. Santos, Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.* **108**, 158104 (2012).
41. F. L. Pinheiro, V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, Evolution of all-or-none strategies in repeated public goods dilemmas. *PLoS Comput. Biol.* **10**, e1003945 (2014).
42. A. Traulsen, D. Semmann, R. D. Sommerfeld, H. J. Krambeck, M. Milinski, Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2962–2966 (2010).
43. J. Grujic *et al.*, A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Sci. Rep.* **4**, 4615 (2014).
44. S. K. Baek, H. C. Jeong, C. Hilbe, M. A. Nowak, Comparing reactive and memory-one strategies of direct reciprocity. *Sci. Rep.* **6**, 25676 (2016).
45. W. H. Press, F. J. Dyson, Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10409–10413 (2012).
46. L. E. Blume, The statistical mechanics of strategic interaction. *Games Econ. Behav.* **5**, 387–424 (1993).
47. G. Szabó, C. Tőke, Evolutionary Prisoner's Dilemma game on a square lattice. *Phys. Rev. E* **58**, 69–73 (1998).
48. A. Traulsen, M. A. Nowak, J. M. Pacheco, Stochastic dynamics of invasion and fixation. *Phys. Rev. E* **74**, 011909 (2006).
49. L. A. Imhof, M. A. Nowak, Evolutionary game dynamics in a Wright-Fisher process. *J. Math. Biol.* **52**, 667–681 (2006).
50. L. A. Imhof, M. A. Nowak, Stochastic evolutionary dynamics of direct reciprocity. *Proc. R. Soc. B* **277**, 463–468 (2010).
51. K. Tainaka, Paradoxial effect in a three-candidate voter model. *Phys. Lett. A* **176**, 303–306 (1993).
52. M. Frean, E. R. Abraham, Rock-scissors-paper and the survival of the weakest. *Proc. R. Soc. B* **268**, 1323–1327 (2001).
53. E. Akin, "The iterated prisoner's dilemma: Good strategies and their dynamics" in *Ergodic Theory, Advances in Dynamics*, I. Assani, Ed. (de Gruyter, Berlin, 2016), pp. 77–107.
54. M. A. Nowak, K. Sigmund, Game-dyamical aspects of the prisoner's dilemma. *Appl. Math. Comput.* **30**, 191–213 (1989).
55. H. Brandt, K. Sigmund, The good, the bad and the discriminator - errors in direct and indirect reciprocity. *J. Theor. Biol.* **239**, 183–194 (2006).
56. C. Hilbe, Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory. *Bull. Math. Biol.* **73**, 2068–2087 (2011).
57. J. Grujic, J. A. Cuesta, A. Sanchez, On the coexistence of cooperators, defectors and conditional cooperators in the multiplayer iterated prisoner's dilemma. *J. Theor. Biol.* **300**, 299–308 (2012).
58. I. Núñez Rodríguez, A. G. M. Neves, Evolution of cooperation in a particular case of the infinitely repeated prisoner's dilemma with three strategies. *J. Math. Biol.* **73**, 1665–1690 (2016).
59. G. Szabó, T. Antal, P. Szabó, M. Droz, Spatial evolutionary prisoner's dilemma game with three strategies and external constraints. *Phys. Rev. E* **62**, 1095–1103 (2000).
60. A. Szolnoki, M. Perc, Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Sci. Rep.* **4**, 5496 (2014).
61. A. Szolnoki, M. Perc, Evolution of extortion in structured populations. *Phys. Rev. E* **89**, 022804. (2014).
62. M. van Veelen, J. García, D. G. Rand, M. A. Nowak, Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9929–9934 (2012).
63. C. Hauert, H. G. Schuster, Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: A numerical approach. *Proc. R. Soc. B* **264**, 513–519 (1997).
64. C. Hauert, O. Stenull, Simple adaptive strategy wins the prisoner's dilemma. *J. Theor. Biol.* **218**, 261–72 (2002).
65. A. J. Stewart, J. B. Plotkin, Small groups and long memories promote cooperation. *Sci. Rep.* **6**, 26889 (2016).
66. C. Hilbe, L. A. Martinez-Vaquero, K. Chatterjee, M. A. Nowak, Memory-*n* strategies of direct reciprocity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4715–4720 (2017).
67. J. Li *et al.*, Evolution of cooperation through cumulative reciprocity. *Nat. Comput. Sci.* **2**, 677–686 (2022).
68. J. Tkadlec, C. Hilbe, M. A. Nowak. Data and Code for "Mutation enhances cooperation in direct reciprocity." Figshare. https://doi.org/10.6084/m9.figshare.21583554.v1. Deposited 14 March 2023.