

The evolution of punishment through reputation

Miguel dos Santos^{1,2,*}, Daniel J. Rankin^{2,3,†} and Claus Wedekind¹

¹*Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland*

²*Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*

³*Swiss Institute of Bioinformatics, Quartier Sorge Bâtiment Génopode, 1015 Lausanne, Switzerland*

Punishment of non-cooperators has been observed to promote cooperation. Such punishment is an evolutionary puzzle because it is costly to the punisher while beneficial to others, for example, through increased social cohesion. Recent studies have concluded that punishing strategies usually pay less than some non-punishing strategies. These findings suggest that punishment could not have directly evolved to promote cooperation. However, while it is well established that reputation plays a key role in human cooperation, the simple threat from a reputation of being a punisher may not have been sufficiently explored yet in order to explain the evolution of costly punishment. Here, we first show analytically that punishment can lead to long-term benefits if it influences one's reputation and thereby makes the punisher more likely to receive help in future interactions. Then, in computer simulations, we incorporate up to 40 more complex strategies that use different kinds of reputations (e.g. from generous actions), or strategies that not only include punitive behaviours directed towards defectors but also towards cooperators for example. Our findings demonstrate that punishment can directly evolve through a simple reputation system. We conclude that reputation is crucial for the evolution of punishment by making a punisher more likely to receive help in future interactions, and that experiments investigating the beneficial effects of punishment in humans should include reputation as an explicit feature.

Keywords: game theory; punishment; cooperation; humans; experimental game

1. INTRODUCTION

Cooperation is often enhanced if non-cooperators can be punished [1–5], but this simple fact cannot yet explain the evolution of punishment, especially not if punishment inflicts immediate costs to the punisher. Indeed, in settings where individuals interact repeatedly with the same partner [6–8], or when third-party punishment is possible (i.e. punishing players for being unkind to others [9,10]), punishers usually finish with lower pay-offs than non-punishers (but see [5]). It has therefore been concluded that punishment is mostly maladaptive within the respective games, and that it may have evolved for other reasons than for promoting cooperation [6–10]. However, the maladaptive argument is not very satisfying [11] and, given the widespread prevalence of punishment, a significant direct evolutionary advantage of punishment is still likely, for example, in the context of reputation games [4].

Although it is well established that reputation can play a key role in social interactions (e.g. in the evolution of human cooperation [12–21]), the possible advantage from a reputation of being a punisher has not been sufficiently explored as an explanation for the evolution of punishment (if I punish you because you have defected against me, others may later not defect against me). Some theoretical studies have suggested that, under certain circumstances, natural selection favours strategies

that take the likelihood of being punished into account [22–27], for example, when information about a neighbour's behaviour is available [25], or in situations where individuals cooperate according to the average punishment strategy played by all social partners [26]. Thus, punishment could act as an indirect threat to observers. However, the majority of empirical studies so far have investigated punishment in anonymous settings where players did not have information on the others' punitive behaviours, or in settings where this information was confounded with information about the others' cooperative behaviours [2,3,6,8,10,28–30].

Here, we explore the evolution of punishment in a helping game. In this game, a donor can either help, or refuse to help a receiver who then can punish in return [26,27,31]. Such punitive actions influence an individual's 'punishment score', which reflects how much an individual punished previously and can then be used by others to discriminate between punishers and non-punishers. Such a reputation system is analogous to the image scoring proposed for the evolution of indirect reciprocity [14–16,18,21,32,33]. In a simple analytical model, we investigate the evolutionary stability of a strategy that discriminates between punishers and non-punishers. We then use computer simulations to extend our model and to test the robustness of the major outcome of the model.

2. METHODS AND RESULTS

(a) Analytical model

We first build an analytical model. Individuals interact randomly and can choose to either help or not help their social partners. *Defectors*, who never help, are denoted

* Author for correspondence (miguel.dossantos@unil.ch).

† Both authors contributed equally to this work.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2010.1275> or via <http://rsob.royalsocietypublishing.org>.

by frequency x_j , *cooperators*, who always help, are denoted by frequency y_j and *discriminators*, who only help individuals that punished in their last encounter, are denoted by frequency z_j . The subscript $j \in \{N, P\}$ denotes whether individuals are *punishers* ($j = P$), who punish upon not receiving help, or *non-punishers* ($j = N$), who never punish. We assume a simple form of scoring, where players have either a positive punishment score, or a negative punishment score. The parameter q_{ij} (where $i \in \{x, y, z\}$ and $j \in \{N, P\}$) is the probability that an individual punished in their last encounter (and therefore has a positive punishment score) while $1 - q_{ij}$ is the probability that an individual did not punish in their last encounter (and therefore has a negative punishment score). For non-punishers, this probability is always zero. For punishers, it is contingent on them having experienced defection in their last encounter. Following replicator dynamics, and assuming an infinite population, the probability that a given individual punished at time $\tau + 1$ can be calculated by the following recursion relations:

$$q_{xP}(\tau + 1) = x_N + x_P + (1 - q_{xP}(\tau))(z_N + z_P),$$

$$q_{yP}(\tau + 1) = x_N + x_P + (1 - q_{yP}(\tau))(z_N + z_P)$$

$$\text{and } q_{zP}(\tau + 1) = x_N + x_P + (1 - q_{zP}(\tau))(z_N + z_P).$$

We assume that behavioural dynamics occur on a very fast time scale relative to evolutionary dynamics, and therefore the above recurrence relations will equilibrate very quickly to give the equilibrium punishment scores for each state, which are

$$q_{xP}^* = \frac{1 - y_N - y_P}{1 + z_N + z_P},$$

$$q_{yP}^* = \frac{1 - y_N - y_P}{1 + z_N + z_P}$$

$$\text{and } q_{zP}^* = \frac{1 - y_N - y_P}{1 + z_N + z_P}.$$

Since each punishment score is identical (as it merely depends on the relative frequency of defectors and discriminators in the population), we will write $q_{xP}^* = q_{yP}^* = q_{zP}^* = q^*$. We assume that cooperation imposes a cost c on an actor, and confers a benefit b on a receiver. Punishment imposes a cost s on the punisher, while inflicting a cost e on the individual being punished. The pay-off of each of the six strategies is denoted by g_{ij} (given in appendix A). The fitness of a given strategy is given by $w_{ij} = g_{ij}/\bar{g}$, where \bar{g} is the average pay-off in the population, such that

$$\bar{g} = x_N g_{xN} + x_P g_{xP} + y_N g_{yN} + y_P g_{yP} + z_N g_{zP} + z_P g_{zP}.$$

The dynamics of a given strategy are therefore given by $k_{ij}(t + 1) = k_{ij}(t)w_{ij}$, where k_{ij} is the frequency of the strategy in question. Table 1 shows a list of the symbols used in our model.

We analyse the condition under which *punishing discriminators* (with a frequency z_P) cannot be invaded by any other strategy, and is therefore an evolutionarily stable strategy (ESS). The condition under which punishing discriminators, z_P , will be an ESS with respect to *non-punishing defectors* (i.e. $x_N \rightarrow 0$ and $z_P \rightarrow 1$) is if z_P have a greater pay-off than x_N (i.e. $g_{zP} > g_{xN}$), which is fulfilled if

$$b - c > s - e,$$

under which they are also an ESS with respect to *non-punishing cooperators* (i.e. when $y_N \rightarrow 0$ and $z_P \rightarrow 1$). This

Table 1. List of symbols.

symbol	definition
x_j	defectors (with $j \in \{N, P\}$)
y_j	cooperators (with $j \in \{N, P\}$)
z_j	discriminators (with $j \in \{N, P\}$)
i_N	non-punishing i (with $i \in \{x, y, z\}$)
i_P	punishing i (with $i \in \{x, y, z\}$)
g_{ij}	pay-off of strategy i_j
\bar{g}	average pay-off in the population
w_{ij}	fitness of strategy i_j
c	cost of helping
b	benefit of receiving help
s	cost of punishing
e	cost of being punished
q^*	probability that an individual punished a defection in its last encounter
n	population size ^a
m	number of interactions per individual ^a
μ	mutation rate ^a
ε	error rate ^a
k_a	punishment score ^a (with $a \in \{1, 2, 3\}$)
I_s	image score ^a

^aOnly used in the simulations.

condition is always respected since $b > c$ and $e > s$. Punishing discriminators will also be an ESS with respect to *punishing defectors* (i.e. $x_P \rightarrow 0$ and $z_P \rightarrow 1$) if being punished is less costly than helping ($e > c$), and to *non-punishing discriminators* (i.e. $z_N \rightarrow 0$ and $z_P \rightarrow 1$) if the benefit of helping is greater than the cost of punishing ($b > s$). If the cost of helping is greater than the cost of being punished ($c > e$), *punishing cooperators* will be able to invade, and if this condition is held, punishing cooperators are an ESS with respect to all strategies (except non-punishing cooperators, to which they are neutral). This allows for non-punishing cooperators to invade through drift. If this occurs, any of the other four strategies (i.e. x_N , x_P , z_N and z_P) will be able to invade. However, punishment is frequency dependent, and, during an invasion of a population of non-punishing cooperators, punishing discriminators will be able to outcompete non-punishing defectors (i.e. $g_{zP} > g_{xN}$), which will occur if

$$z_P > \frac{s}{b - c + e},$$

which is independent of the frequency of non-punishing cooperators (y_N). Thus, as punishing discriminators become more common, they are increasingly favoured over *defectors*. But how can punishment invade a population of non-punishing defectors (x_N)? When common, x_N will be an ESS with respect to all strategies except z_N , who can invade through drift. Then, as soon as they are common in the population and that the benefit of help is greater than the cost of punishing ($b > s$), a single punisher can outweigh the costs of punishing by receiving help and z_P can invade.

(b) Computer simulations

Our analytical model considers a simplified case, with large populations and a limited number of strategies. To test the robustness of our results and to put them into the wider context of cooperative strategies, we built an individual-based model. We modelled a population of

finite size n . In each generation, pairs of players are randomly chosen to interact in the following manner: one player (the donor) has to decide whether to help the other one (the receiver) or not. Helping incurs a cost c to the donor and a benefit b to the receiver (where $b > c$). No help results in zero pay-off for both individuals. After the donor's decision, the receiver has the possibility to pay a cost s in order to punish the donor. Punishment reduces the donor's total pay-off by e (where $e > s$). In each generation, a player has on average m interactions, each of which could be in the donor or receiver role. The fitness of a player is given by its total number of points at the end of its m interactions. Individuals then leave offspring in proportion to their fitness. Mutations occur during reproduction and both the helping and the punishment strategies mutate independently with a probability μ , in which case they are replaced at random by another helping/punishment strategy, respectively. This potentially creates new combinations of punishment and helping strategies into our simulations. We used $n = 500$, $m = 30$ and $\mu = 0.02$ in all our simulations. It is well established that for a reputation system to be efficient, a relatively high number of interactions per individual is required [14,32,33], hence our choice of $m = 30$ interactions. As a consequence, we chose a sufficiently large number of individuals in the population to avoid direct reciprocity effects, that is, a given player meeting the same partner in reversed roles (i.e. probability < 0.03). In the electronic supplementary material, we show that decreasing the mutation probability μ can often hinder the emergence of punishment and cooperation for low benefits b and low punishment ratios (s/e ; electronic supplementary material, figure S1).

Our mechanism of punishment scoring is analogous to the image scoring of [14] but applied to punitive actions: individuals have a punishment score that starts at 0 and can reach -5 or $+5$. We implemented four different punishment/image scores, which we denote k_1 , k_2 , k_3 and I_s .

- k_1 —*punishment of defection*. Each time an individual punishes defection, his k_1 increases by 1 unit, whereas each time he does not punish defection, his k_1 decreases by 1 unit.
- k_2 —*punishment of either defection/cooperation*. Each time an individual punishes (defection or cooperation), his k_2 increases by 1 unit, whereas each time he does not punish, his k_2 decreases by 1 unit.
- k_3 —*punishment of cooperation*. Each time an individual punishes cooperation, his k_3 increases by 1 unit, whereas each time he does not punish cooperation, his k_3 decreases by 1 unit.
- I_s —*cooperation*. Each time an individual cooperates, his I_s increases by 1 unit, whereas each time he does not cooperate, his I_s decreases by 1 unit (i.e. image score).

These punishment scores are used to define the different strategies. Table 2 shows our 10 different helping strategies specifying how a player acts in the donor role, and table 3 shows the four different punishment strategies specifying how a player acts in the receiver role. In summary, we have 10 different helping strategies that could be combined with four different punishing strategies; i.e. there were in total 40 behavioural strategies possible.

Table 2. The 10 strategies in the simulations that specify how a player acts in the donor role.

strategy name	strategy notation	help if receiver ...	helping rule
defectors	x_j	—	never help
cooperators	y_j	—	always help
discriminators	z_j	punished defections	$k_1 \geq 0$
discriminators'	z'_j	did not punish defections	$k_1 < 0$
helpers to punishers	u_j	punished	$k_2 \geq 0$
helpers to punishers'	u'_j	did not punish	$k_2 < 0$
helpers to C-punishers	v_j	punished cooperative actions	$k_3 \geq 0$
helpers to C-punishers'	v'_j	did not punish cooperative actions	$k_3 < 0$
image scorers	a_j	cooperated	$I_s \geq 0$
image scorers'	a'_j	defected	$I_s < 0$

Table 3. The four strategies in the simulations that specify how a player acts in the receiver role. (Non-punishers never punish.)

strategy name	strategy notation	punish if donor ...
unconditional punishers	i_A	cooperated or defected
non-punishers	i_N	—
punishers of defection	i_P	defected
punishers of cooperation	i_C	cooperated

Figure 1 gives a typical example of how punishing strategies emerged in our simulations and how populations arrived at a mix of cooperative strategies. The mean frequencies across 20 replicates and for generations 18 000–20 000 were: punishing discriminators = 0.58; punishing cooperators = 0.12, non-punishing discriminators = 0.06, mean rate of cooperation = 0.98, mean rate of punishment = 0.01 (standard errors always < 0.01). Note that low levels of punishment are sufficient to prevent defectors from invading (figure 1b). Punishing discriminators (z_P) consistently evolved as the dominant strategy in a wide range of different parameters (figure 2), provided that the cost of being punished was greater than the cost of cooperating ($e > c$), and that the cost of punishing was smaller than the benefit of cooperation ($s < b$; figure 2). Increasing the benefit b of cooperation also had the consequence of increasing the frequency of punishing cooperators (y_P) and *punishing image scorers* (a_P ; figure 2). Although some drift effects were possible (e.g. non-punishing defectors can be invaded by other non-cooperative strategies or punishing cooperators can invade punishing discriminators; figure 2), variation brought by mutations maintained a certain selective pressure for discrimination, such that

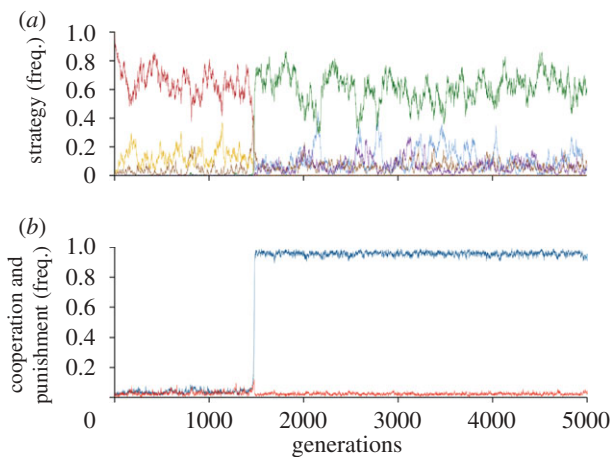


Figure 1. Punishment scoring leads to cooperation in computer simulations. (a) The six most frequent strategies (out of 40 possible ones) in a typical simulation for a finite population initially constituted of non-punishing defectors (red). Non-punishing discriminators (brown) appear and pave the way for punishing discriminators (green) to invade and dominate all other strategies. The other frequent strategies are *defectors punishing cooperation* (yellow), punishing cooperators (blue) and punishing image scorers (purple). (b) Frequency of cooperative moves (blue line) and punitive moves (red line). Parameters values are: $c = 1$, $b = 2$, $s = 1$, and $e = 4$.

unconditional cooperators could never dominate, i.e. the cycling between strategies seen in the analytical model was avoided, and cooperation was more stable (figure 1b). Decreasing the mutation rate μ made it harder for punishers to invade for low benefits b of cooperation (electronic supplementary material, figure S1a–f), and also decreased the selection pressure for discrimination (electronic supplementary material, figure S1d–i).

As in other indirect reciprocity models, our punishment scoring system depends on the ability of a player to correctly assess the punitive reputation of others. Thus, we went on to test whether the inclusion of errors had an impact on our results. We introduced errors in the perception of an individual's score. After each interaction, the donor's cooperation score I_s was replaced randomly with a probability ε by either $I_s - 1$, $I_s + 0$ or $I_s + 1$, and the receiver's punishment score k_i was replaced randomly with a probability ε by either $k_i - 1$, $k_i + 0$ or $k_i + 1$. As a consequence, the number of wrong actions caused by an incorrect score perception was greater at the beginning of each generation and decreased as the game went along (e.g. if my image score is 4 and I cooperate, replacing it by 3 instead of 5 will not result in a wrong action). This realistic assumption reflected the fact that more mistakes are made when players do not know their partners at the beginning of the interactions. As shown in the electronic supplementary material, the inclusion of errors often hinders the emergence of punishment and cooperation when the benefit b of cooperation is low (electronic supplementary material, figure S2a–c). With greater b , however, punishment still emerged, but increasing the error rate reduced the efficiency of punishing discriminators (z_p) and punishing image scorers (a_p) relative to punishing cooperators (y_p) who do not use reputation for their actions (electronic supplementary material, figure S2d–i).

3. DISCUSSION

Punishment that serves to prevent an individual from repeating a damaging action towards the punisher or that serves to prevent future defection towards the punisher seems to be very common in humans and some animals [1]. Recent studies have concluded, however, that punishment may have evolved for something else other than for promoting cooperation, because significant benefits to punishers could typically not be found in the context of pure cooperation games [4,6–11,34]. Third-party punishment in an indirect reciprocity setting, for example, rarely favours punishment [9,10]. Indeed, it has been suggested that benefits to groups rather than to individuals could explain the evolution of punishment [35], even though punishment is expected to provide either direct or indirect fitness benefits to the individual in order to evolve [26,31,36]. For example, punishment can be favoured by indirect reciprocity when it discourages future aggression by observers [37]. Our models show that even if punishing defectors is immediately costly, it acts to discourage future refusals to help from observers to such a degree that the immediate costs of punishment are outweighed by the additional donations it evokes over the long run. Hence, punishment evolved in our simulations entirely through a punitive reputation, i.e. without punishment directed towards non-punishers [4,22] or the need of spatial constraints [25,35,38].

The cognitive abilities of humans may allow reputation to be used not only for assessing the cooperativeness of individuals within a social group, but also with regard to the readiness of group members to punish defectors. In our simulations, both types of reputation could be used, and individuals merely using the other's reputation of being generous (i.e. image scoring [14]) do not fare better than punishing discriminators. The latter strategy has a higher pay-off than punishing cooperators, a strategy that could be interpreted as strong reciprocity [35,38]. Strong reciprocity can be evolutionarily stable when common, but when punitive actions are observable by others, punishing cooperators cannot invade punishing discriminators unless the cost of helping is greater than the cost of being punished ($c > e$, as shown in our analytical model above). It is likely that situations where individuals actions were observable, and thus formed public information, occurred during much of human evolutionary history [39].

In experimental public goods games, punishment is typically perceived as conferring benefits on a social group and being an act of cooperation in itself [3,40]. Indeed, humans seem to be more likely to punish if they are observed by others [41], suggesting that they care about the reputation effects that arise from punishing. It has been shown that refusing low offers (a sort of punishment) in the ultimatum game when there are observers made the 'punishers' more likely to receive higher offers in later interactions [42]. Thus, it is still unknown whether the proximate incentives to cooperate more often with punishers would come from the fear of being punished (and thus defecting when there is no threat of punishment would be opportunism [27]) or from the wish to reward punishers for their pro-social behaviour [3,40]. However, the latter hypothesis raises a second-order dilemma as rewarding punishers is also costly.

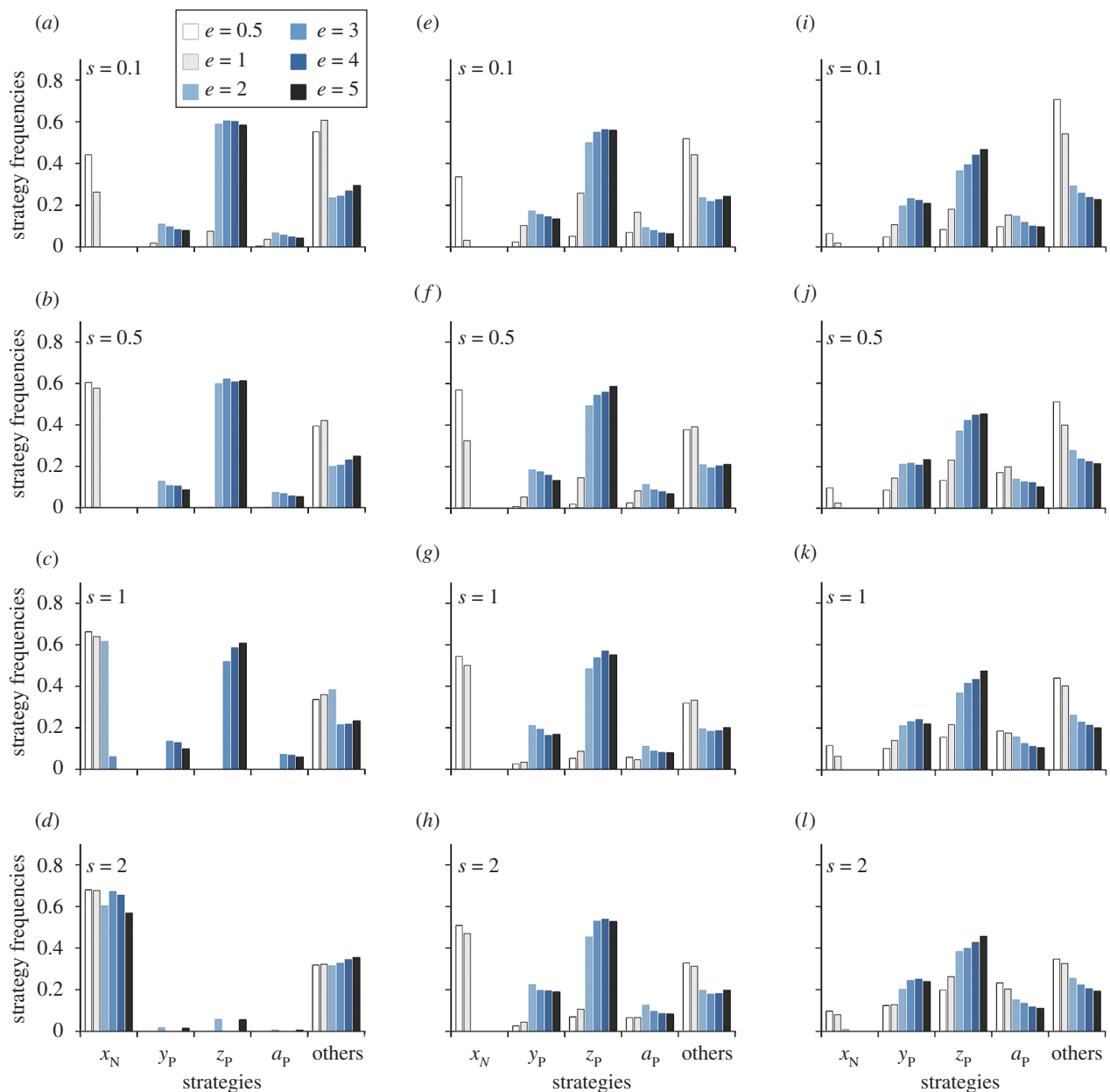


Figure 2. Evolution of punishment at cost of helping $c = 1$ and benefit $b = 2$ ($a-d$), $b = 4$ ($e-h$), and $b = 10$ ($i-l$). The average frequencies of the most successful strategies non-punishing defectors (x_N), punishing cooperators (y_P), punishing discriminators (z_P) and punishing image scorers (a_P) are calculated across 20 replicates for generations 18 000–20 000 for different costs s to the punisher and costs e of being punished (standard errors always < 0.1). Punishing discriminators (z_P) predominate in the simulations for a large set of parameters. The category ‘others’ is the sum of the remaining 36 strategies, some of which mainly arise through drift, for example, defectors punishing cooperation if non-punishing defectors (x_N) are most frequent (e.g. (d)).

We focused on strategies using indirect information on the punishment reputation of others. Strategies using their own experience of direct punishment (i.e. whether you punished me) compared with those using indirect information (i.e. whether you punished others) may perform differently depending on the number of interactions as well as the group size, as it is the case with direct and indirect reciprocity [33]. Similarly, the weight of information attached to reputation gained from punishment of defectors and cooperators was the same in our model, but both types of information could be weighted differently [21]. We also assumed no retaliation from punished individuals in our models. While more realistic, the option of retaliation seems to lower the cooperation level in public goods games [43], but it is still unknown what

would be the influence on cooperation if retaliative actions also impact one’s reputation. Another assumption of our model was that all individuals have the same capacity to punish. This is probably unrealistic for many animals when punishment could be used to establish and maintain dominance relationships, in which case dominant individuals would receive more cooperation [1,44,45].

Our model bears similarity to the model of Hilbe & Sigmund [27] (hereafter H&S), where a game was played in which reputation was implemented as a probability to know about the others’ behaviour and individuals had the option to reward cooperators. For instance, they found that punishing discriminators (denoted $[O_C, P]$ in H&S) can invade a population initially constituted of non-punishing defectors (denoted

[ALLD, N] in H&S) if the probability to know whether the co-player punishes or not is greater than $s(b+s)$, denoted $(\mu > \gamma)$ $(b + \gamma)$ in H&S. Moreover, the possibility for receivers to reward cooperation seems to foster the evolution of cooperation (and then punishment) when the probability to know the co-player's reputation is small [27]. We included 40 different strategies in our simulations including the possibility to always punish, or to punish cooperation, an often missing feature of many models [22,25–27,31], and still find that punishing discriminators can bring cooperation and prevent defectors from invading.

Our study highlights the importance of reputation in driving the evolution of punishment. By allowing reputation to be based on either the punishment of defectors, cooperators or both, we have shown that punishing defectors and always cooperating with punishers emerges as a dominant strategy. Our results are also robust to the other strategies, such as image scoring. We conclude that reputation is the key to the evolution of punishment, and that simple reputation games can explain the high prevalence of punishment in humans. The combination of reputation and punishment acts as a strong mechanism promoting the evolution of cooperation.

We thank R. Bergmüller, R. Bshary, M. Chapuisat, C. Clavien, C. El Mouden, A. Gardner, C. Hauert, M. Hochberg, L. Keller, C. Metzger, S. Nusslé, A. Ross-Gillespie and the anonymous referee for discussion or comments on a previous version of the manuscript, and the Swiss National Science Foundation (grants to C.W. and D.J.R.) for funding.

APPENDIX A

Our analytical model above consists of six strategies, namely x_N , x_P , y_N , y_P , z_N and z_P . The pay-offs for these respective strategies can be written as:

$$\begin{aligned} g_{xN} &= -e(x_P + y_P + z_P) + b(y_N + y_P), \\ g_{xP} &= -s x_N + x_P(-s - e) + b z_N q_{xP} + z_N(1 - q_{xP})(-s) \\ &\quad + b z_P q_{xP} + z_P(1 - q_{xP})(-s) + (y_P + z_P)(-e) \\ &\quad + b(y_N + y_P), \\ g_{yN} &= -c + b(y_N + y_P), \\ g_{yP} &= -c - s(x_N + x_P) + b z_N q_{yP} + z_N(1 - q_{yP})(-s) \\ &\quad + b z_P q_{yP} + z_P(1 - q_{yP})(-s) + b(y_N + y_P), \\ g_{zN} &= x_P q_{zP}(-c) + x_P(1 - q_{zP})(-e) + z_P q_{zP}(-c) \\ &\quad + z_P(1 - q_{zP})(-e) + y_P q_{zP}(-c) + y_P(1 - q_{zP})(-e) \\ &\quad + b(y_N + y_P) \end{aligned}$$

and
$$\begin{aligned} g_{zP} &= -s(x_N + x_P) + x_P q_{zP}(-c) + x_P(1 - q_{zP})(-e) \\ &\quad + b z_N q_{zP} + z_N q_{zP}(-s) + z_P q_{zP}(-c) \\ &\quad + z_P(1 - q_{zP})(-e) + y_P q_{zP}(-c) \\ &\quad + y_P(1 - q_{zP})(-e) + b z_P q_{zP} \\ &\quad + z_P(1 - q_{zP})(-s) + b(y_N + y_P). \end{aligned}$$

REFERENCES

- Clutton-Brock, T. H. & Parker, G. A. 1995 Punishment in animal societies. *Nature* **373**, 209–216. (doi:10.1038/373209a0)
- Fehr, E. & Gächter, S. 2000 Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994.
- Fehr, E. & Gächter, S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)
- Sigmund, K. 2007 Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22**, 593–600. (doi:10.1016/j.tree.2007.06.012)
- Raihani, N. J., Grutter, A. S. & Bshary, R. 2010 Punishers benefit from third-party punishment in fish. *Science* **327**, 171–171. (doi:10.1126/science.1183068)
- Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. 2008 Winners don't punish. *Nature* **452**, 348–351. (doi:10.1038/nature06723)
- Rand, D. G., Ohtsuki, H. & Nowak, M. A. 2009 Direct reciprocity with costly punishment: generous tit-for-tat prevails. *J. Theor. Biol.* **256**, 45–57. (doi:10.1016/j.jtbi.2008.09.015)
- Wu, J. J., Zhang, B. Y., Zhou, Z. X., He, Q. Q., Zheng, X. D., Cressman, R. & Tao, Y. 2009 Costly punishment does not always increase cooperation. *Proc. Natl Acad. Sci. USA* **106**, 17 448–17 451. (doi:10.1073/pnas.0905918106)
- Ohtsuki, H., Iwasa, Y. & Nowak, M. A. 2009 Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature* **457**, 79–82. (doi:10.1038/Nature07601)
- Ule, A., Schram, A., Riedl, A. & Cason, T. N. 2009 Indirect punishment and generosity toward strangers. *Science* **326**, 1701–1704. (doi:10.1126/science.1178883)
- Rankin, D. J., dos Santos, M. & Wedekind, C. 2009 The evolutionary significance of costly punishment is still to be demonstrated. *Proc. Natl Acad. Sci. USA* **106**, E135. (doi:10.1073/pnas.0911990107)
- Alexander, R. D. 1987 *The biology of moral systems*. New York, NY: Aldine de Gruyter.
- Zahavi, A. 1991 Arabian babblers: the quest for social status in a cooperative breeder. In *Cooperative breeding in birds: long term studies in behaviour and ecology* (eds P. B. Stacey & W. D. Koenig), pp. 105–130. Cambridge, UK: Cambridge University Press.
- Nowak, M. A. & Sigmund, K. 1998 Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577. (doi:10.1038/31225)
- Wedekind, C. & Milinski, M. 2000 Cooperation through image scoring in humans. *Science* **288**, 850–852. (doi:10.1126/science.290.5491.454)
- Milinski, M., Semmann, D., Bakker, T. C. M. & Krambeck, H. J. 2001 Cooperation through indirect reciprocity: image scoring or standing strategy? *Proc. R. Soc. Lond. B* **268**, 2495–2501. (doi:10.1098/rspb.2001.1809)
- Milinski, M., Semmann, D. & Krambeck, H. J. 2002 Reputation helps solve the 'tragedy of the commons'. *Nature* **415**, 424–426. (doi:10.1038/415424a)
- Wedekind, C. & Braithwaite, V. A. 2002 The long-term benefits of human generosity in indirect reciprocity. *Curr. Biol.* **12**, 1012–1015. (doi:10.1016/S0960-9822(02)00890-4)
- Nowak, M. A. & Sigmund, K. 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298. (doi:10.1038/nature04131)
- Brandt, H. & Sigmund, K. 2006 The good, the bad and the discriminator: errors in direct and indirect reciprocity. *J. Theor. Biol.* **239**, 183–194. (doi:10.1016/j.jtbi.2005.08.045)
- Rankin, D. J. & Eggmann, F. 2009 The evolution of judgment-bias in indirect reciprocity. *Proc. R. Soc. B* **276**, 1339–1345. (doi:10.1098/rspb.2008.1715)
- Boyd, R. & Richerson, P. J. 1992 Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* **13**, 171–195. (doi:10.1016/0162-3095(92)90032-Y)

- 23 Gintis, H., Smith, E. A. & Bowles, S. 2001 Costly signaling and cooperation. *J. Theor. Biol.* **213**, 103–119. (doi:10.1006/jtbi.2001.2406)
- 24 Sigmund, K., Hauert, C. & Nowak, M. A. 2001 Reward and punishment. *Proc. Natl Acad. Sci. USA* **98**, 10 757–10 762. (doi:10.1073/pnas.161155698)
- 25 Brandt, H., Hauert, C. & Sigmund, K. 2003 Punishment and reputation in spatial public goods games. *Proc. R. Soc. Lond. B* **270**, 1099–1104. (doi:10.1098/rspb.2003.2336)
- 26 Gardner, A. & West, S. A. 2004 Cooperation and punishment, especially in humans. *Am. Nat.* **164**, 753–764.
- 27 Hilbe, C. & Sigmund, K. Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B* **277**, 2427–2433. (doi:10.1098/rspb.2010.0065)
- 28 Rockenbach, B. & Milinski, M. 2006 The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723. (doi:10.1038/nature05229)
- 29 Gächter, S., Renner, E. & Sefton, M. 2008 The long-run benefits of punishment. *Science* **322**, 1510. (doi:10.1126/science.1164744)
- 30 Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M. A. 2009 Positive interactions promote public cooperation. *Science* **325**, 1272–1275. (doi:10.1126/science.1177418)
- 31 Lehmann, L., Rousset, F., Roze, D. & Keller, L. 2007 Strong reciprocity or strong ferocity? a population genetic view of the evolution of altruistic punishment. *Am. Nat.* **170**, 661.
- 32 Leimar, O. & Hammerstein, P. 2001 Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. Lond. B* **268**, 745–753. (doi:10.1098/rspb.2000.1573)
- 33 Roberts, G. 2008 Evolution of direct and indirect reciprocity. *Proc. R. Soc. B* **275**, 173–179. (doi:10.1098/rspb.2007.1134)
- 34 Tao, Y., Li, C., Wu, J. J. & Cressman, R. 2009 Reply to Rankin *et al.*: the efficiency ratio of costly punishment. *Proc. Natl Acad. Sci. USA* **106**, E136. (doi:10.1073/pnas.0912928107)
- 35 Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. 2003 The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100**, 3531–3535. (doi:10.1073/pnas.0630443100)
- 36 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. (doi:10.1111/j.1420-9101.2006.01119.x)
- 37 Johnstone, R. A. & Bshary, R. 2004 Evolution of spite through indirect reciprocity. *Proc. R. Soc. Lond. B* **271**, 1917–1922. (doi:10.1098/rspb.2003.2581)
- 38 Bowles, S. & Gintis, H. 2004 The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theor. Popul. Biol.* **65**, 17–28. (doi:10.1016/j.tpb.2003.07.001)
- 39 Dunbar, R. (ed.) 1996 *Grooming, gossip and the evolution of language*. Cambridge, MA: Harvard University Press.
- 40 Barclay, P. 2006 Reputational benefits for altruistic punishment. *Evol. Hum. Behav.* **27**, 325–344. (doi:10.1016/j.evolhumbehav.2006.01.003)
- 41 Kurzban, R., DeScioli, P. & O'Brien, E. 2007 Audience effects on moralistic punishment. *Evol. Hum. Behav.* **28**, 75–84. (doi:10.1016/j.evolhumbehav.2006.06.001)
- 42 Fehr, E. & Fischbacher, U. 2003 The nature of human altruism. *Nature* **425**, 785–791. (doi:10.1038/nature02043)
- 43 Janssen, M. A. & Bushman, C. 2008 Evolution of cooperation and altruistic punishment when retaliation is possible. *J. Theor. Biol.* **254**, 541–545. (doi:10.1016/j.jtbi.2008.06.017)
- 44 Bshary, R., Grutter, A. S., Willener, A. S. T. & Leimar, O. 2008 Pairs of cooperating cleaner fish provide better service quality than singletons. *Nature* **455**, 964–966. (doi:10.1038/nature07184)
- 45 Stamp Dawkins, M. 2010 Do asymmetries destabilize the Prisoner's Dilemma and make reciprocal altruism unlikely? *Anim. Behav.* **80**, 339–341. (doi:10.1016/j.anbehav.2010.05.005)