SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
THEORY
COMPUTATIONAL BIOLOGY
ECOLOGY
BIOLOGICAL MODELS

Received
30 April 2012

Accepted
28 May 2012

Published
13 June 2012

Correspondence and requests for materials should be addressed to C.H. (hilbe@evolbio.mpg.de)

Emergence of responsible sanctions without second order free riders, antisocial punishment or spite

Christian Hilbe & Arne Traulsen

Evolutionary Theory Group, Max Planck Institute for Evolutionary Biology, D-24306 Plön, Germany.

While empirical evidence highlights the importance of punishment for cooperation in collective action, it remains disputed how responsible sanctions targeted predominantly at uncooperative subjects can evolve. Punishment is costly; in order to spread it typically requires local interactions, voluntary participation, or rewards. Moreover, theory and experiments indicate that some subjects abuse sanctioning opportunities by engaging in antisocial punishment (which harms cooperators), spiteful acts (harming everyone) or revenge (as a response to being punished). These arguments have led to the conclusion that punishment is maladaptive. Here, we use evolutionary game theory to show that this conclusion is premature: If interactions are non-anonymous, cooperation and punishment evolve even if initially rare, and sanctions are directed towards non-cooperators only. Thus, our willingness to punish free riders is ultimately a selfish decision rather than an altruistic act; punishment serves as a warning, showing that one is not willing to accept unfair treatments.

Numerous experiments demonstrate that human subjects are eager to punish others for unjust behaviour[1–6], thereby suggesting that we are equipped with an inclination for retaliation[7]. The evolutionary origin of this inclination, however, is puzzling because punishment is costly and therefore unlikely to evolve unless it results in direct or indirect benefits[8–13]. One avenue of research has argued that sanctions are more costly for the punished and that punishment thus gives a relative payoff advantage to the punisher[9,11,12,14]. In these models, punishment can evolve, sometimes even if punishers are initially rare, if there are accompanying mechanisms such as voluntary participation in the collective endeavor[11,12], local interactions on a lattice or on a network[9,15–19], or the option to reward cooperators[13]. Surprisingly, it was also demonstrated that defectors who punish other defectors help to pave the way for a cooperative society[14–16]. However, while a relative payoff advantage for the punisher may explain the emergence of punishment, it cannot account for the emergence of responsible punishment, targeted at defectors only. If the mere act of punishing others gives an edge to the punisher, then spite and antisocial punishment should eventually take over. Most previous studies presumed that only defectors are punished, which is clearly contradicting experimental evidence from numerous countries[20]. Two recent models that also allow cooperators to be punished have shown that anti-social punishment can fully prevent the evolution of cooperation and responsible sanctions in both, well-mixed[21] and lattice-structured[22] populations. Therefore, the question arises whether punishment can promote cooperation at all[6].

However, in most real interactions, the decision to punish others does not only affect the relative payoffs of the players, but also their reputation. If the punishment act can be observed by others, it can pay to sanction only defectors. A recent experiment suggests that emotions such as anger or moral disgust may have evolved as a commitment device; they lead people to disregard the immediate consequences of their behaviour in order to preserve integrity and to maintain their reputation[23]. If individuals are able to build up a strict reputation by displaying a low tolerance for unfair behaviour, then future interaction partners may act more cooperatively. Recently, *dos Santos et. al.* have presented an analytical model, combined with computer simulations, showing that reputation indeed facilitates the co-evolution of cooperation and punishment[24]. However, their analytical model does not allow antisocial punishment, and individuals can only resort to the last action of their peers. A responsible use of sanctions requires a long-run reputation advantage[25]. Here, we underpin this argument with an evolutionary model. We derive an exact condition for the evolution of responsible punishment in the presence of antisocial punishment. Our model shows that reputation allows the co-evolution of cooperation and responsible sanctions even if both are initially rare.

Results

We consider a pairwise game with two stages. Before the game starts, a coin toss determines which player is in the role of the donor and which one is in the role of the recipient. In the initial helping stage, donors may cooperate and transfer a benefit b to their recipients, at their own cost $c < b$, or they may refuse to do so. In the subsequent punishment stage, recipients decide whether or not to punish the donor at a cost γ, thereby reducing the payoff of the donor by β. Depending on the outcome of the helping stage, there are four possible reactions of the recipient: Punishing defectors only (denoted by R for responsible sanctions), punishing cooperators only (A for antisocial punishment), punishing everybody (S for spiteful punishment) or punishing nobody (N). Because sanctions are costly, immediate self-interest speaks against either form of punishment, leading to a destabilization of punishment in the absence of reputation[12]. In order to incorporate reputation, we assume that donors can anticipate their co-player's behaviour with probability λ, either from previous encounters, from observation, or from gossip. We can therefore distinguish four different types of donors. The first type are the C-players who always cooperate, whereas the second type, the D-players, never cooperate, regardless of λ and the opponent's reputation. The third type are the opportunistic cooperators, O_C, who optimally adapt their behaviour on the co-player's punishment reputation: They cooperate against social sanctioners, while saving the cooperation costs against all other recipients, N, A, and S. If no information on the co-player's reputation is available, O_C-donors cooperate by default. The last type of donors, opportunistic defectors O_D, also adjust their behaviour to the recipient's reputation (in the same way as O_C-donors), but play defect if the recipient's reputation is unknown.

Thus, if there is no information about the reputation of the other group members available, opportunistic cooperators O_C just behave as unconditional cooperators C, and opportunistic defectors O_D are indistinguishable from defectors D. However, once the others' reputation is known, opportunists can be swayed by the threat of punishment, whereas the unconditional strategies cannot. As players can be in both roles, donor and recipient, and since we consider four strategies for each role (C, O_C, O_D, D for donors and R,N,A,S for recipients), there are 16 strategies in total. Note that this is only a subset of the full strategy space; for example, donors might also apply the rather counter-intuitive rule to cooperate only against anti-social punishers. However, such a strategy is clearly dominated by O_C, and we show in the Supplementary Information (SI) that our results remain unchanged if we consider the full strategy space.

We study the transmission of strategies with a frequency-dependent birth-death process[26] in a finite population of size n. In each time step, two randomly chosen individuals compare their payoffs and one of them can switch to the other one's strategy. This process can be interpreted as a model for social learning, whereby successful strategies spread, and, occasionally, random strategy exploration introduces novel strategies (corresponding to mutations in biological models). In the limit of low exploration rates, we provide an analytical approximation, which is complemented with simulations for frequent exploration (SI Text).

When interactions are completely anonymous ($\lambda = 0$), then neither responsible punishment nor cooperation occurs at notable frequencies (Fig. 1). Instead, donors tend to defect either unconditionally, or because they are not swayed by responsible sanctions. Because of the absence of cooperators, antisocial punishment incurs no costs and can therefore increase to substantial levels through neutral drift, which is in line with previous studies[21,22]. These results, however, change drastically when the recipient's reputation is at stake: If the probability of knowing the others' type fulfills (see SI)

$$\lambda > \frac{(n-1)\gamma - \beta}{(n-1)(\gamma + b) + c - \beta}, \qquad (1)$$

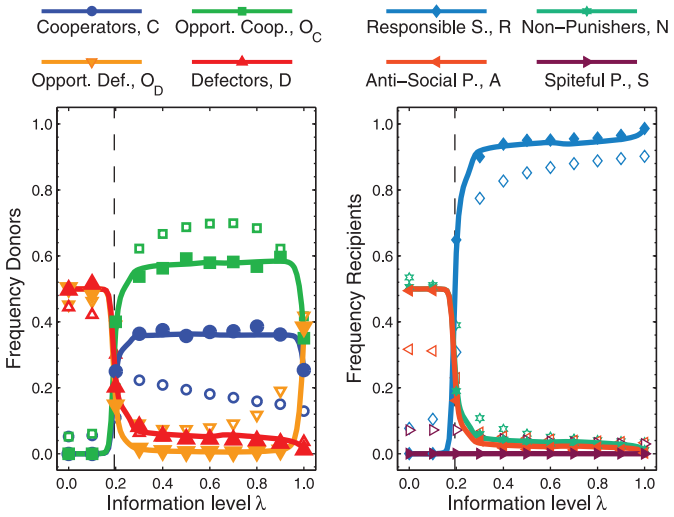

Figure 1 | Information promotes the co-evolution of cooperation and responsible punishment. Time-averaged frequencies for the strategies of donors (left graph) and recipients (right graph), respectively. Solid lines indicate exact results for the limiting case of rare exploration. Filled symbols represent simulation results for low exploration rates ($\mu = 0.0001$) and open symbols are simulations for high exploration rates ($\mu = 0.1$). The black dashed line represents the critical information level given by Eq. (1). Above this information level, individuals make use of responsible sanctions to deter opportunists from defection. Parameter values are $n = 80$, $b = 4$, $\beta = 3$, $c = \gamma = 1$, the strength of selection is set to $s = 0.5$. Simulations were run over a period of 10^{10} time steps (i.e., each individual was allowed to implement more than 10^8 strategy changes).

then it pays off for the recipient to engage in responsible sanctions to deter opportunists from defection. Notably, this expression simplifies to $\lambda > \gamma/(\gamma + b)$ for large populations, indicating that responsible punishment is the result of balancing the costs of punishment γ with the prospects of future benefits b, but does neither depend sensitively on cost of cooperation c nor on the magnitude of the punishment β. In fact, we find that above this threshold, recipients almost immediately switch to responsible punishment, which in turn promotes the evolution of cooperative strategies among the donors. Remarkably, this positive effect of information is largely independent of the exploration rate, although frequent exploration has a distinct impact on the abundance of opportunism.

To illustrate the emergence of responsible punishment, we have traced the evolutionary dynamics (Fig. 2). In the absence of reputation effects, both, spite and responsible sanctions soon go extinct, followed by a long period of neutral drift between unconditional and opportunistic defection, such that everyone defects, as well as between antisocial punishment and no punishment, such that no one punishes. On the other hand, if recipients have the opportunity to build a reputation, then they turn to responsible punishment, which promotes the evolution of opportunism and, eventually, establishes cooperation. This holds true even if responsible sanctions are absent in the initial population (Fig. 3): Indeed, starting from a population of antisocial defectors (DA), mutation and neutral drift can lead to a population of non-punishing opportunists (O_DN). This kind of opportunism paves the way for responsible sanctions (O_DR or O_CR).

Our results demonstrate that with and without information, spite is immediately driven to extinction (see Figs. 1–3). This is in contrast to a recent model considering the evolution of antisocial behaviour in locally subdivided populations[22]. However, we show in the Supplementary Information that spite requires a high degree of anonymity, small population sizes and low costs of punishment to

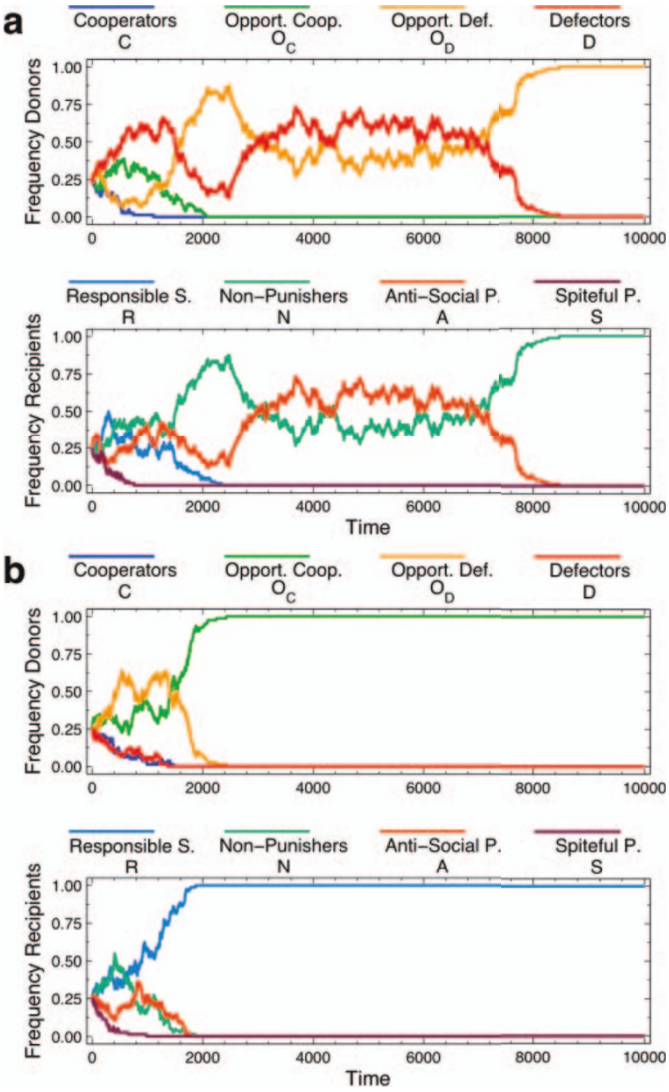

Figure 2 | Time evolution of responsible punishment. Two typical individual-based simulation runs, without (a) and with (b) reputation. In both cases, the upper graph depicts the dynamics among the donors' strategies, whereas the lower graph shows the evolution of strategies among recipients. While a low information regime results in neutral drift between different non-cooperative strategies, individuals almost immediately switch to social sanctions and cooperation if their reputation is at stake. Parameter values are $\mu = 0.0001$ and $\lambda = 0$ for (a) and $\lambda = 0.3$ for (b), respectively, the other parameter values being the same as in Figure 1.

gain a foothold in the population. These conditions are intuitive for they lead to local competition where relative payoff advantages matter. It is noteworthy that the three conditions of anonymity, small group sizes and cheap punishment are characteristic for many laboratory experiments, suggesting that such behavioural studies may overestimate the impact of spite on human decision making.

The positive effects of reputation are robust with respect to errors in the perception of the co-players' reputation, and to extensions of the strategy space (SI Text and Figs. S3 and S4). Moreover, these results are not restricted to pairwise interactions: Our results also carry over to public good games between more than two players (SI Text and Figs. S5 and S6). Also in that case, there is a critical threshold for the reputation parameter λ which needs to be met for cooperation and responsible sanctions to evolve. This critical threshold, however, increases with the number of group members. Thus, large group sizes threaten the emergence and the stability of responsible peer punishment, which may explain why most large societies rather rely on centralized punishment institutions than on self-governance[19,27,28].

Discussion

Previous evolutionary models could not explain why individuals learn to deal responsibly with sanctions. Instead, it was either presumed that punishment is targeted at defectors only[9–16,18], or it was predicted that evolution leads to non-punishing defectors or spite, respectively[22]. Here, we have shown how reputation can resolve these issues. Non-anonymity makes anti-social punishment and spite unappealing, and if punishment evolves, then it is systematically targeted at non-cooperators. Hence, we also question the conventional wisdom that any behaviour, even if abstruse, can become a common norm as long as deviations are punished[8]. Opportunistic individuals will stop to impose sanctions on pro-social activities, simply because it is in their own interest to let cooperative outcomes evolve. In particular, the emergence of anti-social punishment in some models[21,22], is likely to be a consequence of their assumption of anonymous interactions. Antisocial punishment has been observed experimentally in repeated games, but there it could be a component of retaliation[20,29].

In our model, individuals learn to make use of responsible punishment because these sanctions serve as a signal to bystanders. In this way, responsible sanctions are a form of weak reciprocity[30]: they are beneficial in the long run, despite being costly in the short run. If this individual long-run benefit of punishment is absent (e.g. if reputation effects are precluded), then responsible sanctions do not evolve. Strong reciprocators (i.e., individuals that are willing to punish others even if it reduces their absolute fitness in the long run[31]) do not emerge in our model. Thus, responsible

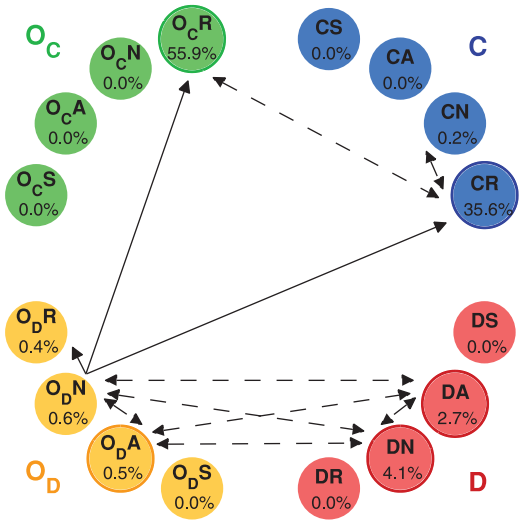

Figure 3 | Responsible punishment can invade when rare. Time-averaged frequencies of the 16 possible strategy combinations and typical transitions between homogeneous populations. Arrows with dashed lines indicate neutral drift between the two corresponding strategy combinations. Arrows with solid lines represent transitions where the target strategy has a fixation probability that exceeds the neutral probability $1/n$. Populations marked with a colored ring can only be invaded through neutral drift. The Figure illustrates that unconditional defectors can be subverted by opportunistic defectors, which in turn can be swayed by responsible sanctions. However, once established, responsible sanctions can be replaced by unconditional defectors via the (unlikely) path via non-punishing cooperators, which can be invaded by non-punishing defector strategies. Parameter values are $n = 80$, $b = 4$, $\beta = 3$, $c = \gamma = 1$, $s = 0.5$, $\lambda = 0.3$ and frequencies are calculated for the limit of rare exploration. For clarity, we have only plotted arrows starting from strategies that are played in more than 0.5% of all cases.

punishment is a selfish act, rather than an altruistic service to the community. Opportunism (that is, the propensity to be swayed by sanctions), on the other hand, emerges endogenously, once individuals are able to anticipate the punishment behaviour of their peers. Of course this implies some cognitive requirements on the subjects: They have to monitor their co-players and need to process and remember this information properly. Subjects in behavioural experiments show an enhanced memory for faces of defectors[32] and although not tested empirically, one may expect similar results for the faces of punishers. Humans highly regard reputation[33]; the mere picture of an eye, indicating that someone is watching[34] or the physical presence of an experimenter[35] can affect the subjects' behaviour, often making them more cooperative or increasing their willingness to punish non-cooperators. In fact, the capability to gather and transmit information might be a major cause for the the high levels of cooperation in humans[36].

Under non-anonymity, reputation becomes a strategic variable and experiments reveal that we make use of sophisticated strategies when it comes to publicising or concealing information about ourselves[37]. While explicit penalties serve as a warning to others, they also bear the risk of counter-punishment[38]. However, we show that responsible sanctions remain prevalent even if counter-punishment is a sure event (in which case the costs for the punisher, γ, are as high as the costs for being punished, β, SI Text and Fig. S2), implying that we are willing to pay a high price to uphold our reputation[39,40].

1. Ostrom, E., Walker, J. & Gardner, R. Covenants with and without a sword: Self-governance is possible. *Am. Polit. Sci. Rev.* **86**, 404–417 (1992).
2. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
3. Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723 (2006).
4. Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
5. Yamagishi, T. The provision of a sanctioning system as a public good. *J. Pers. and Soc. Psychology* **51**, 110–116 (1986).
6. Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. Winners don't punish. *Nature* **452**, 348–351 (2008).
7. de Quervain, D. J. F. *et al.* The neural basis of altruistic punishment. *Science* **305**, 1254–1258 (2004).
8. Boyd, R. & Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology* **13**, 171–195 (1992).
9. Nakamaru, M. & Iwasa, Y. The evolution of altruism by costly punishment in the lattice structured population: score-dependent viability versus score-dependent fertility. *Evol. Ecol. Research* **7**, 853–870 (2005).
10. Boyd, R., Gintis, H. & Bowles, S. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* **328**, 617–20 (2010).
11. Fowler, J. H. Altruistic punishment and the origin of cooperation. *Proc. Natl. Acad. Sci. USA* **102**, 7047–7049 (2005).
12. Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907 (2007).
13. Hilbe, C. & Sigmund, K. Incentives and opportunism: From the carrot to the stick. *Proc. R. Soc. B* **277**, 2427–2433 (2010).
14. Eldakar, O. T. & Wilson, D. S. Selfishness as second-order altruism. *Proc. Natl. Acad. Sc.* **105**, 6982–6986 (2008).
15. Nakamaru, M. & Iwasa, Y. The coevolution of altruism and punishment: role of the selfish punisher. *J. Theor. Biol.* **240**, 475–488 (2006).
16. Helbing, D., Szolnoki, A., Perc, M. & Szabó, G. Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput Biol* **6**, e1000758 (2010).
17. Helbing, D., Szolnoki, A., Perc, M. & Szabo, G. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New J. Physics* **12**, 083005 (2010).
18. Perc, M. & Szolnoki, A. Self-organization of punishment in structured populations. *New J. Physics* **14**, 043013 (2012).
19. Perc, M. Sustainable institutionalized punishment requires elimination of second-order free-riders. *Sci. Rep.* **2**, 344 (2012).
20. Herrmann, B., Thöni, C. & Gächter, S. Antisocial punishment across societies. *Science* **319**, 1362–1367 (2008).
21. Rand, D. G. & Nowak, M. A. The evolution of antisocial punishment in optional public goods games. *Nature Communications* **2** (2011).
22. Rand, D. G., Armao IV, J. J., Nakamaru, M. & Ohtsuki, H. Anti-social punishment can prevent the co-evolution of punishment and cooperation. *J Theor Biol* **265**, 624–632 (2010).
23. Yamagishi, T. *et al.* The private rejection of unfair offers and emotional commitment. *Proc. Natl. Acad. Sci.* **106**, 11520–11523 (2009).
24. Dos Santos, M., Rankin, D. J. & Wedekind, C. The evolution of punishment through reputation. *Proc. R. Soc. B* **278**, 371–377 (2011).
25. Gächter, S., Renner, E. & Sefton, M. The long-run benefits of punishment. *Science* **322**, 1510 (2008).
26. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).
27. Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863 (2010).
28. Szolnoki, A., Szabó, G. & Perc, M. Phase diagrams for the spatial public goods game with pool punishment. *Phys Rev E* **83**, 036101 (2011).
29. Nikiforakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* **92**, 91–112 (2008).
30. Guala, F. Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behav. Brain Sci.* **35**, 1–59 (2012).
31. Gintis, H. Strong reciprocity and human sociality. *J. Theo. Biol.* **206**, 169–179 (2000).
32. Mealy, L., Daood, C. & Krage, M. Enhanced memory for faces of cheaters. *Behav. Ecol. Sociobiol.* **17**, 119–128 (1996).
33. Semmann, D., Krambeck, H. J. & Milinski, M. Strategic investment in reputation. *Behav. Ecol. Sociobiol.* **56**, 248–252 (2004).
34. Haley, K. J. & Fessler, D. M. T. Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* **26**, 245–256 (2005).
35. Kurzban, R., DeScioli, P. & O'Brien, E. Audience effects on moralistic punishment. *Evol. Hum. Behav.* **28**, 75–84 (2007).
36. Brosnan, S. F., Salwiczek, L. & Bshary, R. The interplay of cognition and cooperation. *Phil. Trans. Roy. Soc. London B* **365**, 2699–2710 (2010).
37. Rockenbach, B. & Milinski, M. To qualify as a social partner, humans hide severe punishment, although their observed cooperativeness is decisive. *Proc. Natl. Acad. Sci.* doi: 10.1073/pnas.1108996108 (2011).
38. Jansen, M. A. & Bushman, C. Evolution of cooperation and altruistic punishment when retaliation is possible. *J Theor Biol* **254**, 541–545 (2008).
39. Fehr, E. Human behaviour: don't lose your reputation. *Nature* **432**, 449–450 (2004).
40. Barclay, P. Reputational benefits for altruistic punishment. *Evol. Hum. Behav.* **27**, 325–344 (2006).

Acknowledgements

We thank M. Abou Chakra, J. García, K. Sigmund and M. Milinski for helpful comments.

Author contributions

Both authors were involved in the design and analysis of the model and wrote the paper.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare that they have no competing financial interest.

How to cite this article: Hilbe, C. & Traulsen, A. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Sci. Rep.* **2**, 458; DOI:10.1038/srep00458 (2012).