

Genetic Stability and Territorial Structure Facilitate the Evolution of Tag-Mediated Altruism

Lee Spector
Cognitive Science
Hampshire College
Amherst, MA 01002
lspector@hampshire.edu

Jon Klein
Cognitive Science
Hampshire College
Amherst, MA 01002
and
Physical Resource Theory
Chalmers University of Technology
and Göteborg University
SE-412 96 Göteborg, Sweden

Abstract Evolutionary theorists have long been interested in the conditions that permit the evolution of altruistic cooperation. Recent work has demonstrated that altruistic donation can evolve in surprisingly simple models, in which agents base their decisions to donate solely on the similarity of evolved “tags” relative to evolved tag-difference tolerances. There is disagreement, however, about the conditions under which tag-mediated altruism will in fact evolve. Here we vary two critical parameters in a standard model of tag-mediated altruism—genetic stability and territorial structure—and show that altruism evolves in a wide range of conditions. We demonstrate the evolution of significant levels of altruism even when the immediate costs to donors equal the benefits to recipients. We describe the mechanism that permits the emergence of altruism in the model as a form of kin selection that is facilitated by interactions between altruism, genetic drift, and fecundity.

Keywords

Altruism, tags, computational models, genetic stability, territorial structure, fecundity

I Tag-Mediated Altruism

Evolutionary theorists have long been interested in the conditions under which altruism can be produced by evolution. Prior work has explained the evolution of altruistic behavior toward kin [6, 7], reciprocating partners [1, 20], and partners with good reputations [14], as well as non-reciprocal altruism in Mendelian populations driven by population structure [22]. Recently attempts have been made to explain altruism that is based solely on partner similarity, as assessed relative to arbitrary *tags* [9] that can serve as models of either genetic or cultural markers [16–18]. These attempts, if successful, would provide extremely general explanations, as they make few assumptions regarding genetic structure or the knowledge and abilities of the evolving agents. In particular, they would help to explain how altruism can evolve among agents that have no explicit knowledge of kinship, past behavior, or reputation. The models would also be applicable to organisms that are far too simple to engage in activities that underlie some other models of the evolution of altruism, such as mate selection [11] and explicit repression of competition [4].

There is disagreement, however, about the conditions under which tag-mediated altruism will in fact evolve. Roberts and Sherratt [18] argue that altruism is actually “built in” to the model of Riolo et al. [16], and that altruism no longer evolves if an implicit bias favoring cooperation among agents with identical tags is removed. We show here, however, that tag-mediated altruism does in fact evolve in a range of conditions, even in the context of Roberts and Sherratt’s modifications. In particular, we show that it evolves when lower mutation rates are used and when territorial structure limits the interaction radius of each agent. We also show that this holds even when the immediate

costs and benefits of altruistic acts are equivalent. We discuss this surprising result and the mechanisms that produce it in the final section of the article.

2 The Model and Prior Work

We began with a computational model equivalent to those used in the prior work [16], in which a population of 100 agents is progressed through a number of generations (30,000 in prior work, 400,000 here). Each agent has a tag τ and a tolerance T , both of which are represented as floating point numbers and uniformly sampled from $[0, 1]$ in the initial generation. All agents begin each generation with a score of 0 and are then paired with $P=3$ randomly selected others for potential acts of altruism. If the difference between the tags of the two agents is less than or equal to the potential donor's tolerance, then the altruistic act (donation) is performed: a cost c is deducted from the donor's score, and a benefit b is added to the recipient's score. The resulting scores may be positive, negative, or zero. For most of our runs we used $c=0.1$, $b=1.0$, as in most of the prior work, but we also studied other values of c and b . After all donation attempts are completed, the population of the next generation is produced by pairing each agent a_i in a reproductive tournament with another randomly selected agent a_j and producing, at the location of a_i , a child from the parent with the higher score (or from a_j in case of a tie).¹ The child inherits its tag and tolerance from the parent, but each may each be mutated with probability m ($= 0.1$ in the prior work, though we vary it here). When a tag is mutated, it is replaced with a new tag randomly selected from $[0,1]$; when a tolerance is mutated, it is summed with mean 0, standard deviation 0.01 Gaussian noise.² This mutation scheme, like the donation, selection, and reproduction schemes described above, is somewhat arbitrary, but it is also standard in the literature.

The primary measure of interest is the average donation rate, which ranged from 2.1% to 79.2% in the data presented in the original report by Riolo et al. [16]. We also measured tag diversity by counting the number of unique tags in each generation and averaging this number over all generations in a run.³

Riolo et al. did not allow tolerances to mutate below 0, so agents with identical tags would always donate to one another. Roberts and Sherratt [18] reported that the removal of this restriction (implemented by changing the lower bound for T from 0 to -10^{-6}) prevents the emergence of altruism; they observed donation rates of 1.48%, as opposed to Riolo et al.'s 73.6%. Riolo et al. replied that the two models have opposite, strong biases regarding cooperation among agents with identical tags, and that cooperation continues to evolve in models with intermediate biases [17]. Here we investigated only the condition that prevented the evolution of cooperation for Roberts and Sherratt ($T \geq -10^{-6}$), but we varied other parameters and found that cooperation does in fact evolve in a range of conditions.

3 Variation of Genetic Stability and Territorial Structure

The mutation rate used in the prior work, 0.1, is exceptionally high compared to naturally occurring rates. We ran experiments with additional mutation rates m (0.001, 0.01, and 0.5) and found that in more stable genetic environments (with lower m) higher donation rates often evolve; for example, using Roberts and Sherratt's parameters and varying the mutation rate, we observed a donation rate of 37.61% for $m=0.001$. Tag diversity was lower in this condition than in Roberts and Sherratt's condition, but it was still considerable; we saw an average of 33.17 unique tags in each generation. Our results agreed well with those of Roberts and Sherratt for $m=0.1$ (1.30% donation, tag diversity

¹ The location of the child is not specified in the prior literature; it becomes relevant only under the modifications presented here.

² The lower bound for the resulting tolerance is discussed below; there is no upper bound.

³ A variety of other diversity measures might also be informative, including measures of "effective diversity" that are sensitive to the tolerance levels found in a population.

79.87), and we also found low donation rates with $m=0.5$ (2.59%) and $m=0.01$ (3.93%). These results are graphed in Figure 1, which also displays the results of additional experiments discussed below.

Natural populations also have territorial structure that limits both the potential recipients of altruistic acts and the reproductive competitors; in some cases such a structure has been shown to affect the evolution of altruism. For example, Nowak and May showed that a regular two-dimensional grid structure facilitates the evolution of cooperation in a prisoner's dilemma game [13]. Even when populations are large and irregularly structured, patterns of interaction tend to form "small world networks" that have been shown to affect game theoretic dynamics [21], and several researchers have characterized relations between population structure and evolutionary dynamics more generally [10, 22]. An earlier analytic model of Eshel [3] predicted, in agreement with Wright [24], that low mobility in a population with territorial structure would produce a "neighbor effect" that promotes the evolution of altruism. Others have subsequently studied the effects of low mobility (a.k.a. "viscosity"), which can also sometimes inhibit altruism by intensifying competition among neighboring altruists [12, 23]. For example, recent work by Hauert and Doebeli demonstrates that spatial structure can inhibit the evolution of cooperation in the "snowdrift game" [8]. Riolo et al. described possible connections between territorial structure and the evolution of altruism in their model, but without presenting data [17]. More recently Axelrod et al. [2] demonstrated the emergence of tag-mediated altruism in populations structured as two-dimensional lattices.

Here we have modeled a simple form of territorial structure by considering the population as a one-dimensional array wrapped around a ring. Each agent's interactions are limited to others within radius R ; for example, with $R=5$ each agent has 10 potential partners for altruistic acts and for reproductive competition (5 on each side). The model of Roberts and Sherratt results from setting $R=50$ (as does the model of Riolo et al. [16], with the additional setting of $T \geq 0$). Note that this scheme, while simple, has features more similar to those of natural populations than some of the alternatives that have been explored in the literature; for example, in contrast to the scheme used by Hauert and Doebeli [8], agents in our model may interact with agents that are not directly adjacent, and the amount of overlap between neighborhoods decreases relatively smoothly for pairs of agents at successively greater distances from one another. A more standard cellular-automaton-like model, in which agents can interact only when immediately adjacent, results from setting $R=1$.

We observed that as R decreases, the amount of altruism often increases. For example, in Roberts and Sherratt's condition (aside from this change) we observed donation rates of 12.48% for $R=10$ (tag diversity 58.74), 31.74% for $R=5$ (tag diversity 40.66), and 55.80% for $R=1$ (tag diversity 41.81). Notice that the condition $R=1$ produces both more altruism and more diversity than $R=5$; as discussed below, this demonstrates that the observed altruism is not a simple product of homogeneity.

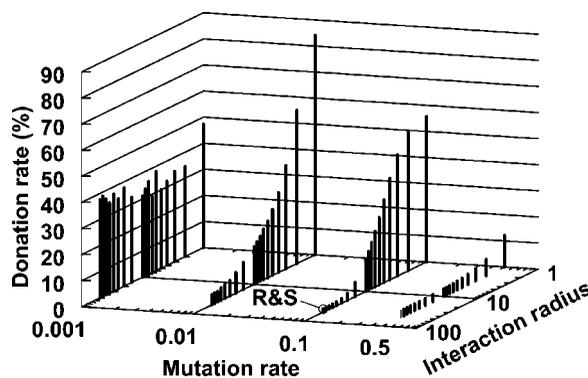


Figure 1. Donation rate as a function of mutation rate and interaction radius, with a cost/benefit ratio $c/b=0.1$ (as in prior work). The configuration described by Roberts and Sherratt [18] is marked with the circle and "R&S." Here 92 independent runs were conducted for each of 72 conditions. A log scale was used for the interaction radius axis to highlight the behavior of the system at small radii.

Territorial structure can be combined with lower mutation rates, producing cumulative effects. For example, $R=1$ with $m=0.01$ yields a donation rate of 84.11%, while the change to R alone yields only 55.80% and the change to m alone yields only 3.93%. By varying both R and m we produce the landscape of altruism shown in Figure 1. Donation rates of up to 84.11% (with tag diversity 9.52) were obtained by combining changes in mutation rate with changes in territorial structure. We conclude that while the conditions studied by Roberts and Sherratt produce minimal altruism, significant altruism evolves in many other reasonable conditions.

4 Variation of Cost/Benefit Ratio

Altruism can evolve even when the cost/benefit ratio of altruistic donation is much higher than the value 0.1 considered above. The value 0.1 was used in prior work without substantial justification, but it might be considered to model alarm calls, or the sharing of surplus food, or other altruistic acts that benefit the recipient greatly at little cost to the actor. These justifications notwithstanding, one might wonder about the extent to which the observed effects depend on the large net gain to the population, and about the ability of the model to explain the evolution of more costly forms of altruism.

Riolo et al. found that donation rates fell to 24.7% when c/b reached 0.5 (even though, in their model, $T \geq 0$). In contrast, we observed a donation rate of 75.82% with $c/b=0.5$, $m=0.01$, and $R=1$. At higher ratios c/b one would expect less altruism, and Hamilton's rule (discussed further below) might lead one to expect altruism to approach zero with $c/b=1.0$. Consistent with this expectation, Riolo et al. report that when c/b reaches 0.6 the donation rate drops to 2.2%. Surprisingly, however, we find that donation rates average 12.13% even with $c/b = 1.0$ (with $m=0.01$ and $R=1$). This effect is remarkably consistent; although the rate fluctuates over the course of a run, in our 92 runs in the condition $m=0.01$, $R=1$, $c/b = 1.0$ we found average donation rates ranging only from 11.19% to 13.01%. This range is an order of magnitude higher than rates found in many of our other conditions and much higher than control runs in which large penalties were assigned to all altruists (producing rates of less than 0.6%). This means that nontrivial levels of altruism can arise even when donation produces no net gain to the population, a situation analogous to the sharing of scarce food in a natural ecosystem. The levels of donation found in our simulations with $c/b = 1.0$ are graphed in Figure 2. The effect is the strongest for $R=1$, a condition that resembles the cellular automata upon which many other spatial models have been based; but there are also effects at other values of R . We discuss the mechanism responsible for this effect in the following section.

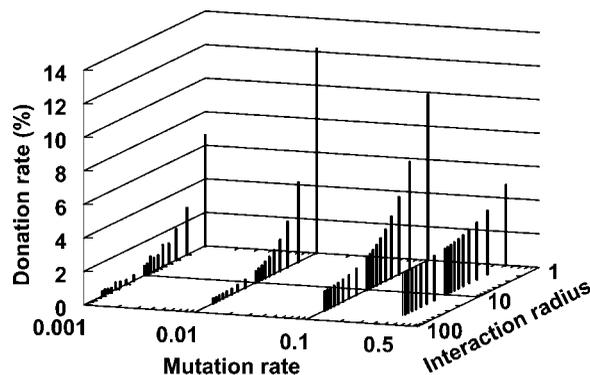


Figure 2. Donation rate as a function of mutation rate and interaction radius, with a cost/benefit ratio $c/b=1.0$. Here 92 independent runs were conducted for each of 72 conditions.

5 Discussion

The simulation results have several implications: they provide evidence supporting the viability of tag-mediated models of altruism in general, they demonstrate the effects of genetic stability and territorial structure on these models, and they demonstrate that altruism can arise even when the immediate cost to the donor is equivalent to the immediate benefit to the recipient. We address these implications in this order.

The data show that alternative but reasonable mutation rates and a simple form of territorial structure both facilitate the evolution of altruism in a standard tag-mediated model, potentially explaining prior non-altruistic outcomes as resulting from particular assumptions about these parameters. Consistent with discussions of these factors in the literature, our diversity data suggests that altruism in the model depends to some extent on homogeneity among interacting agents. It should be noted, however, that we observed significant levels of diversity in all conditions and that we observed no cases in which diversity collapsed completely. In all of our conditions the average number of unique tags in a generation was at least 4, and in many of our conditions with high levels of altruism the average was over 40; the diversity averaged over all of the runs graphed in Figures 1 and 2 was 67.09. Considering that the population size is just 100, these are remarkably diverse populations. The observed altruism is therefore not a simple consequence of complete homogeneity.

Neither is the observed altruism a simple consequence of mutation rates that are so low as to prevent any occurrence of negative-tolerance defectors. We observed substantial altruism with $m=0.01$, which produces about one mutation every generation, and a mutation to the agent at any particular location about once every 100 generations. The initial population is likely to include at least one agent with tolerance around 0.01, and there is a strong initial selective advantage to low tolerance. Since, in addition, tolerance mutation is performed by summation with Gaussian noise of mean 0 and standard deviation 0.01, one would expect to produce at least one individual with zero or negative tolerance within the first few generations. Since we ran our simulations for 400,000 generations, there was ample time for low or negative tolerances to arise through mutations and to spread through the population. One would expect the average tolerance to converge to values near zero quickly, and for most of the tolerance values to cluster around zero even when our low mutation rates are in effect. This is indeed what we observe; the average tolerance, averaged over all of our conditions, was just 0.014. Mutations to zero or negative tolerance were therefore quite common.

It is also important to note that the relatively high incidence of altruistic behavior in these simulations cannot be attributed to random variation alone. Mutation does constantly introduce small numbers of new altruists, an effect that increases with higher mutation rates. Figures 1 and 2 show, however, that higher mutation rates often lead to *lower* levels of altruism in the model discussed here. In these cases mutation is disrupting heredity and thereby dampening the effects of selection. Additional tests with the extreme mutation rate of $m=1.0$ in the condition $c=b=1$ produced an average donation rate of 4.88, which is less than half of that produced with $m=0.01$.

The altruism-promoting effects of lower mutation rates and territorial structure can be seen as arising from a dynamic similar to Eshel's "neighbor effect" [3]. Altruistic agents are always vulnerable, but they are less likely to meet defectors in environments with low mutation rates and/or small interaction radii. Although there is no direct reciprocation, no explicit knowledge of kinship, and no knowledge of past behavior or reputation in any of the models under consideration here, altruistic agents in sufficiently stable environments will generally find themselves in the vicinity of other altruistic agents, many of which will in fact be kin. Tag-mediated altruism in such models can therefore be seen as a form of *probabilistic kin recognition*, in which agents with similar tags are not explicitly recognized as kin but have a high probability of being kin nevertheless.

The fact that we observe significant altruism even when $c=b=1$ is surprising and might even be thought to violate Hamilton's rule, which states that altruism can evolve only in environments in which the cost to the altruistic donor is less than the product of a relatedness coefficient times the benefit to the recipient; that is, $c < rb$, where $r \leq 1$ [6, 7]. Whether or not the results reported here constitute a violation depends on the interpretation of c and b (which are variously construed as immediate costs

and benefits to individuals, as changes to expected numbers of offspring, and as contributions to “inclusive fitness”) and also on the interpretation of r (about which there is also debate; see [15]). The literature on the proper interpretation of Hamilton’s rule is rich (see, for example, [5]), and the significance of the results reported here in the context of this literature is an interesting, open question.

We investigated the mechanism responsible for the evolution of altruism in the condition $c=b=1$ by observing graphic depictions of simulations over evolutionary time, displaying individual tags, tolerances, and scores using grayscale and color values. From these observations we derived simple illustrations of some of the processes by which altruism can form and grow, even when $c/b=1$. The simplest illustration begins with a population with $R=1$ in which all agents have negative tolerances; that is, in which all agents are initially selfish (Figure 3a).⁴ Here there is no selection pressure on tag values, and it is possible for tag drift to produce high levels of tag diversity (Figure 3b), even if the original tag values were all identical. When tags are sufficiently diverse, then tolerances can themselves begin to drift, as small positive tolerances will have no effect (Figure 3c). This two-stage genetic drift process explains how the first potentially altruistic (positive tolerance) agents can survive to produce offspring.

Next consider two adjacent agents with positive tolerance, one of which does actually produce a net positive donation to the other (Figure 3d); it may do this, for example, because its tolerance has mutated to a larger value than its differently tagged neighbor, or because the two agents are clones of the same parent and their donations to one another are asymmetric by chance. After this happens, the true altruist will not reproduce, but the recipient, which is also a potential altruist (it has a positive tolerance), is likely to be unusually fecund. It will win reproductive tournaments against the loser, and it will also win most tournaments against others, since it has received donations that raise its score above that of agents in a selfish population. It is therefore likely to produce two or more adjacent children which, unless the mutation rate is unusually high, are likely to be clones of one another (Figure 3e). Because these children have identical tags and non-negative tolerances, they will also donate to one another. If they donate symmetrically, then they are neither advantaged nor disadvantaged by the donations, but if they donate asymmetrically, then the winners are again likely to be unusually fecund. When a winner is at the edge of the block of altruists, it will also be likely to win tournaments against differently tagged neighbors, the scores of which will be those of agents in a selfish population. This interaction may cause the block of altruists to grow, as the agents near the ends are sacrificed to allow the agents at the ends to extend the altruistic community (Figure 3f). As a result, altruism-supporting genes and actual acts of altruism may increase in the population until equilibrium is reached with the trend toward selfishness produced by defections. The fact that this equilibrium need not be near zero is a consequence of the way in which genetic drift, geography, altruism, and fecundity interact in the model. Ironically, it is the disparities that arise among the altruistic agents because of random donation asymmetries that allow some of them to gain the upper hand on their selfish neighbors, and thereby to propagate altruistic behavior to a wider community.

With reasonable changes to another critical parameter, reproductive tournament group size, this mechanism can support even more extreme forms of altruism in which the cost actually exceeds the benefit. While our preliminary experiments with $c/b > 1$ produced no significant altruism, we have observed consistently high levels of altruism when reproductive tournaments involve more than two agents. For example, with $c=1.1$, $b=1$, $R=5$, and a tournament size of 4 (meaning that the agent and 3 others in the neighborhood of 10 agents participate in each tournament), we observe donation rates consistently over 30%. The appropriate setting for the tournament size parameter is not obvious, and it seems reasonable to assume that the appropriate setting will vary for different ecological circumstances. For some settings of this parameter the mechanism described in this article can explain the evolution of costly forms of altruism.

In related work we have demonstrated the emergence of significant levels of altruism with $c=b=1$ in more complex models involving continuous 3D dynamics and agents controlled by evolved

⁴ In our actual simulations the tags and tolerances in the initial population are randomly assigned, as discussed above. We discuss a purely selfish population here only for the sake of illustration; this presents the most challenging case for the emergence of altruism.

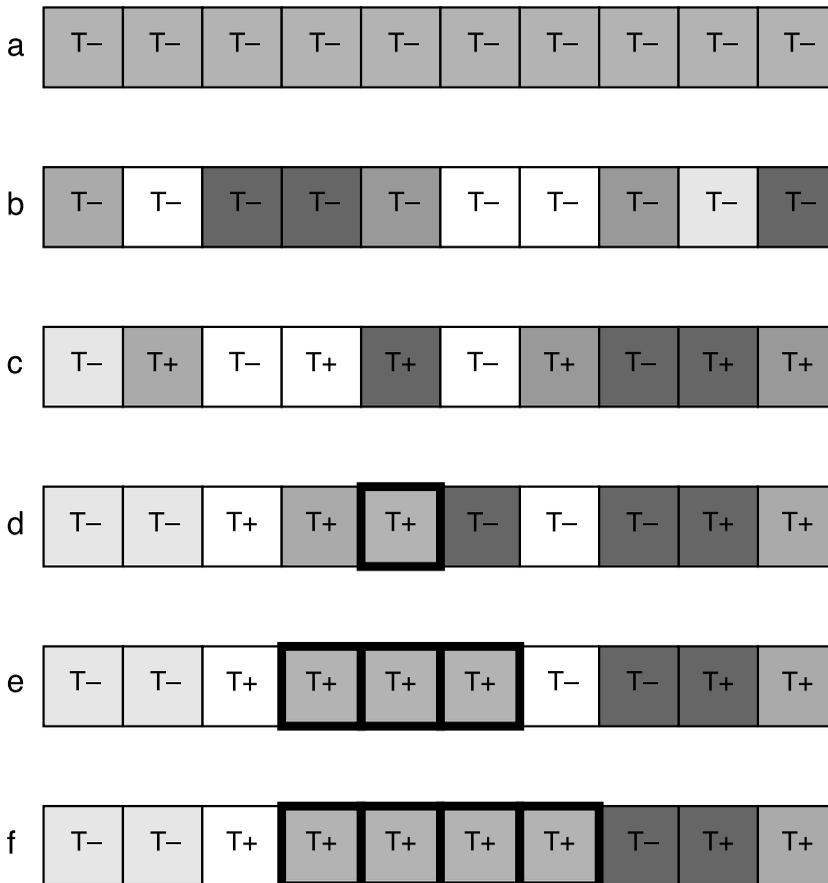


Figure 3. Illustration of the spread of altruism in a condition with $c/b=1.0$. Tags are displayed as grayscale values, $T-$ indicates negative tolerance, and $T+$ indicates positive tolerance. (a) To present the most challenging case, we begin with a uniform population of defectors. (b) Because all tolerances are negative, there will be no donations and tags will drift. (c) In neighborhoods with high tag diversity, small positive tolerances will have no effect and tolerances will drift. (d) When a positive-tolerance agent (shown with the heavy outline) receives net positive donations, either from a neighbor with higher tolerance or from a clone with which interactions were asymmetric by chance, the agent will be likely to sire adjacent clones in the subsequent generation. (e) If the adjacent positive-tolerance clones donate to one another symmetrically, then they will be neither advantaged nor disadvantaged by their potential altruism, but if they donate asymmetrically (which is likely), then some of them will fail to reproduce, but others may again be unusually fecund. (f) This process can repeat, causing the block of altruists to grow as agents at the ends use donations from agents within the block to outcompete other neighbors. Eventually a defector will arise (through a mutation to negative tolerance), and the block of altruists will degrade from within.

computer programs [19]. These models are difficult to analyze, and we cannot yet be certain that the effects described in this article are responsible for the results. Nonetheless, the evolution of tag-mediated altruism in complex models, even with $c=b=1$, suggests that tag-mediated altruism can evolve in a variety of interesting circumstances.

Acknowledgments

This material is based upon work supported by the United States National Science Foundation under Grant No. 0308540 and Grant No. 0216344. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors also thank the anonymous reviewers for their vigorous challenges and attention to detail.

References

1. Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.
2. Axelrod, R., Hammond, R. A., & Grafen, A. (2004). Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. *Evolution*, *58*(8), 1833–1838.
3. Eshel, I. (1972). On the neighbor effect and the evolution of altruistic traits. *Theoretical Population Biology*, *3*, 258–277.
4. Frank, S. A. (2003). Repression of competition and the evolution of cooperation. *Evolution*, *57*, 693–705.
5. Grafen, A. (1984). Natural selection, kin selection and group selection. In J. R. Krebs & N. B. Davies (Eds.), *Behavioral Ecology: An Evolutionary Approach* (3rd ed.) (pp. 62–84). Sunderland, MA: Sinauer Associates.
6. Hamilton, W. D. (1963). The evolution of altruistic behavior. *American Naturalist*, *97*, 354–356.
7. Hamilton, W. D. (1964). The genetical evolution of social behavior. I. *Journal of Theoretical Biology*, *7*, 1–16.
8. Hauert, C., & Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, *428*, 643–646.
9. Holland, J. H. (1996). *Hidden order: How adaptation builds complexity*. Cambridge, MA: Perseus Publishing.
10. Lieberman, E., Hauert, C., & Nowak, M. A. (2005). Evolutionary dynamics on graphs. *Nature*, *433*, 312–316.
11. Koeslag, J. H., & Terblanche, E. (2003). Evolution of cooperation: Cooperation defeats defection in the cornfield model. *Journal of Theoretical Biology*, *224*, 399–410.
12. Mitteldorf, J., & Wilson, D. S. (2000). Population viscosity and the evolution of altruism. *Journal of Theoretical Biology*, *204*, 481–496.
13. Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, *359*, 826–829.
14. Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.
15. Pepper, J. W. (2000). Relatedness in trait group models of social evolution. *Journal of Theoretical Biology*, *206*, 355–368.
16. Riolo, R. L., Cohen, M. D., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, *414*, 441–443.
17. Riolo, R. L., Cohen, M. D., & Axelrod, R. (2002). Riolo et al. reply. *Nature*, *418*, 500.
18. Roberts, G., & Sherratt, T. N. (2002). Does similarity breed cooperation? *Nature*, *418*, 499–500.
19. Spector, L., Klein, J., & Perry, C. (2004). Tags and the evolution of cooperation in complex environments. *Proceedings of the AAAI 2004 Symposium on Artificial Multiagent Learning* (pp. 111–117). Menlo Park, CA: AAAI Press.
20. Trivers, R. (1972). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.
21. Watts, T. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*, 440–442.
22. Wilson, D. S. (1987). Altruism in Mendelian populations derived from sibling groups: The haystack model revisited. *Evolution*, *41*, 1059–1070.
23. Wilson, D. S., Pollock, G. B., & Dugatkin, L. A. (1992). Can altruism evolve in purely viscous populations? *Evolutionary Ecology*, *6*, 331–341.
24. Wright, S. (1945). Tempo and mode in evolution: A critical review. *Ecology*, *26*, 415–419.