6-1-2004

Strong Altruism Can Evolve in Randomly Formed Groups

Jeffrey Fletcher
Portland State University, jeff@pdx.edu

Martin Zwick
Portland State University

Citation Details

Fletcher, J. A., & Zwick, M. (2004). [Post-print]. Strong altruism can evolve in randomly formed groups. Journal of theoretical biology, 228(3), 303-313.

Strong altruism can evolve in randomly formed groups

Jeffrey A. Fletcher[*], Martin Zwick
Systems Science Ph.D. Program
Portland State University, Portland, Oregon 97207, U.S.A.

[*] Corresponding author. Tel.: +1-503-725-4995; fax: +1-503-725-8489
Email addresses: jeff@pdx.edu (J.A. Fletcher), zwickm@pdx.edu (M. Zwick).

Abstract

Although the conditions under which altruistic behaviors evolve continue to be vigorously debated, there is general agreement that altruistic traits involving an absolute cost to altruists (strong altruism) cannot evolve when populations are structured with randomly formed groups. This conclusion implies that the evolution of such traits depends upon special environmental conditions or additional organismic capabilities that enable altruists to interact with each other more than would be expected with random grouping. Here we show, using both analytic and simulation results, that the positive assortment necessary for strong altruism to evolve does not require these additional mechanisms, but merely that randomly formed groups exist for more than one generation. Conditions favoring the selection of altruists, which are absent when random groups initially form, can naturally arise even after a single generation within groups—and even as the proportion of altruists simultaneously decreases. The gains made by altruists in a second generation within groups can more than compensate for the losses suffered in the first and in this way altruism can ratchet up to high levels. This is true even if altruism is initially rare, migration between groups allowed, homogeneous altruist groups prohibited, population growth restricted, or kin selection precluded. Until now random group formation models have neglected the significance of multigenerational groups—even though such groups are a central feature of classic "haystack" models of the evolution of altruism. We also explore the important role that stochasticity (absent in the original random group models) plays in the evolution of altruism. The fact that strong altruism can increase when groups are periodically and randomly formed suggests that altruism may evolve more readily and in simpler organisms than is generally appreciated.

Keywords: Altruism; Haystack model; Multilevel selection; Positive Assortment; Randomly formed groups

1. Introduction

Nearly three decades ago Hamilton (1975) and Wilson (1975) independently developed models which showed that strong altruism (involving an absolute cost to altruists) cannot evolve in randomly formed groups. This conclusion is still generally accepted even among those who debate how best to define altruism and the mechanisms by which it evolves (Hamilton 1975, Maynard Smith 1998, Nunney 1985, 2000, Sober and Wilson 2000, Wilson 1975, 1990). Here we challenge this conclusion by exploring what happens when groups exist for more than one generation. Multigenerational groups are a central feature of Maynard Smith's classic "haystack" model (1964, Wilson 1987), but the role of multiple generations within groups was not explored in Hamilton's (1975) and Wilson's (1975) models. Although the initial conditions after random group formation favor non-altruists over altruists, paradoxically these conditions can switch to favor altruists after even one generation of selection. Thus even though the overall proportion of altruists decreases after one generation, it can increase *even more* after a second generation spent within groups.

Besides single-generation groups, these original analytic models rely on other simplifying assumptions such as an infinite population and no migration between groups. We begin by showing how strong altruism can evolve under the assumptions of the original models, with the only modification being delayed reformation of random groups. Multigenerational groups introduce additional issues such as interactions among related offspring, persistent homogeneous groups of altruists, and exponential growth of population size. We explore model modifications—preventing altruists from benefiting kin, precluding homogeneous groups from forming, and adding a population-level carrying capacity—that mitigate each of these factors. We find that under *all* these modifications (imposed both separately and concurrently) strong altruism, although dampened, can still evolve in randomly formed multigenerational groups. We then transform the basic analytic model into an evolutionary simulation in which population size is finite and stable and migration between groups is allowed. In the simulation model migration between groups dampens selection for altruism, but the stochasticity inherent in finite random group formation, benefit distribution, and carrying capacity appears to enhance the likelihood that altruism will evolve compared to the average non-stochastic characteristics of an infinite population.

2. Classifications of Altruism

Both the analytic and simulation models discussed here involve what Pepper (2000) has termed an *other-only* altruistic trait because none of the altruist's benefits come back to itself, as opposed to *whole-group* traits (also called group-beneficial traits) where the benefit is divided among *all* group members including the altruist. Wilson (1979, 1990) previously classified altruistic traits in a related but different way as either *strong* (involving an absolute cost to altruists) or *weak* (involving only a relative cost to altruists). Other-only altruistic traits are always strong while whole-group traits are strong if the cost to an altruist is greater than its share of the benefit it provides. Note that the same whole-group behavior involving the same sacrifice and provided benefit may be strong or weak depending on group size (Pepper 2000).

In contrast to strong altruism, Wilson showed (1979, 1990) that weakly altruistic traits *can* increase when groups are randomly formed every generation. That is, for an infinite population where a binomial trait is randomly redistributed every generation, the resulting between-group component of total variance can be enough for weak, but not for strong, altruism to evolve. Nevertheless in finite populations where fitness is relative the distinction between strong and weak altruism may be less important as both types are selected against within groups and require selection (or differential productivity) among groups in order to increase (Wilson 1979, 1990). In this paper we focus on other-only, strong altruism (the most restrictive situation) to address the random group models of Hamilton and Wilson directly, but the consequences of multigenerational groups and stochasticity also apply to weak, whole-group traits and therefore these traits can even more readily increase via randomly formed groups than was previously shown (Wilson 1979, 1990).

3. Analytic Model

We focus on Hamilton's model (1975) because he developed a formal proof that altruism cannot evolve in single-generation randomly formed groups (Wilson's model (1975) is similar in all important aspects). In this model a haploid infinite population is randomly subdivided into groups of equal size, n. Group members interact for one generation, affecting each other's fitness (offspring count), before the population is pooled and then again randomly assigned to new groups. In every generation each altruist behaves in a way that costs itself c offspring and provides a *total* benefit of b offspring divided evenly among the other n-1 group members. Each non-altruist receives its share of benefits, but does not provide any benefit to others. Therefore, within every group non-altruists have more offspring than altruists, but groups with more altruists have more offspring per capita than groups with less. This is an example of multilevel selection where here selection within groups opposes selection between groups.

Hamilton (1975) using Price's covariance equation (1970) showed that under his model's assumptions, between-group selection (due to the variance between groups in altruist frequency, p) must always be weaker than average within-group selection (due to the expected variance in the altruistic trait within groups) and therefore the overall frequency of altruists, P, must decrease in every generation. (Capital letters indicate whole population values; small letters indicate group values.) To illustrate this, we calculate $\Delta_1 P$ for an infinite binomial distribution, where Δ_1 indicates that the change occurs over one generation within groups, $g = 1$. The variable g is the number of generations spent within groups before each reformation event. (See Appendix A for model details.) Fig. 1(a) shows how $\Delta_1 P$ depends on the level of benefit, b, provided by altruists for different starting P values. (For convenience, all results reported in this paper use $c = 1$ such that benefit b is also the benefit to cost ratio.) The results shown in Fig. 1(a) are the same for any group size n. Note that as benefit increases, zero is an upper limit on $\Delta_1 P$—hence the conclusion that strong altruism cannot increase under the assumptions of this model for all values of P and n (Hamilton 1975, Wilson 1975).

Yet quite different results are obtained if groups persist for even one additional generation ($g = 2$) before random mixing and the formation of new groups. Fig. 1(b) shows how the change in P after two generations within groups, $\Delta_2 P$, depends on benefit values for different starting P values. The only difference between Figs. 1(a) and

(b) is that the latter measures the change in altruist frequency after an additional generation spent within groups. In this case strong altruism can clearly increase ($\Delta_2 P > 0$) for sufficient values of benefit. Fig. 1(c) shows that smaller groups give a larger increase in altruist frequency which is consistent with previous findings on the relationship between group size and the evolution of altruistic traits (Avilés 1993, Boyd and Richerson 1988). Additionally Fig. 1(c), for which $P = 0.001$, shows that strong altruism can increase due to multigenerational groups even when the altruistic trait is rare, although higher benefit levels are needed for $\Delta_2 P > 0$ when P is low.

4. Applying Hamilton's Rule

We can also understand these results in terms of Hamilton's rule (1964) which states that the condition for an altruistic trait to increase in the next generation is:

$$rb > c\,, \tag{1}$$

where r is the "coefficient of relatedness" or more generally the regression coefficient between the frequency of the trait in recipients and actors (Hamilton 1972). Thus r is a measure of positive assortment—the degree to which the benefits of altruists fall to other altruists. The value of r differs for other-only and whole-group traits because in the latter case, but not the former, altruists are recipients of their own actions (Pepper 2000). We use superscripts w and o respectively for whole-group and other-only measures of r. For whole-group traits r^w is the between-group variance in p over the total variance in the altruistic trait (Breden 1990, Frank 1995a). For an infinite binomial population of randomly formed groups of size n, the variance ratio $r^w = 1 / n$. Thus according to Hamilton's rule (Eq. (1)) the trait increases after one generation if $b / n > c$, but for whole-group traits this means that an altruist's share of its benefit must be greater than its cost—this is the definition of weak altruism so as Wilson (1979, Wilson 1990) noted only weak traits can increase after one generation.

For groups of uniform size the r values are related by the following expression (Pepper 2000):

$$r^o = \frac{nr^w - 1}{n - 1}\,. \tag{2}$$

Therefore $r^o = 0$ for an initial random distribution where $r^w = 1 / n$. Obviously there are no positive values of b and c that can satisfy Hamilton's rule (Eq. (1)) for an other-only (strong) altruistic trait when $r = 0$ and such traits must decrease. Note however, that any modifications to the model that make $r^o > 0$ can yield an increase in P, given a sufficient value of b. Hamilton noted that any positive assortment of altruists beyond that produced at random could allow altruism to increase (1975). Surprisingly, for many parameter settings r^o increases above zero after one generation of selection—even as the proportion of altruists decreases. That is, this transient one-generation-long "population viscosity" of the original models is enough (without any other mechanisms for creating positive assortment) to create conditions that favor altruism in the following generation. If groups are randomly reformed after this single generation then this gain in positive assortment is destroyed before being used by selection; r^o returns to zero and P subsequently declines. On the other hand, additional generations within groups can take advantage of this increased positive assortment so that strong altruism increases, as shown in Fig. 1. (Whether altruism actually increases or not depends on parameters including P, b, and n.)

Note that although Hamilton emphasized a ratio of variances in his proof, in this other-only model the regression coefficient between actors and recipients, r^o, is an easier to interpret measure of the changing conditions affecting altruism as selection occurs. For instance, after one generation of selection (starting with randomly formed groups), the between over total variance, r^w, can decrease while r^o increases. It is the increase in r^o that accurately reflects whether altruism can increase in the next generation. For most parameter settings both measures increase after one generation, but the range of parameters where r^w decreases is greater than the range in which r^o decreases. In general the r values can decrease when there is a combination of low P, low b, and high n.

According to Hamilton's rule whether altruism increases in the second generation within groups depends on whether r^o after the first generation is greater than c / b. Of course it is not enough for altruism to increase in the second generation for altruism to increase overall. The increases in subsequent generations within groups must make up for any losses in the initial generation(s). Fig. 2(a) shows the expected dynamic change in P values over successive generations when groups persist for one and two generations before random reformation and Fig. 2(b) shows the concurrent changes in r^o. Altruist frequency P decreases monotonically when groups are reformed every generation and $r^o = 0$ before each round of selection. On the other hand when groups exist for two generations, P oscillates (and can ratchet upward). The every-other-generation saw-toothed peaks in P correspond to similar (but offset) oscillating peaks in r^o (Fig. 2(b)). Here r^o increases after a generation within groups and we indicate the critical c / b value with a solid horizontal line. Troughs on the other hand correspond to global mixing, new group formation, a decrease of r^o back to zero and a subsequent decrease in P. In Fig. 2 we also show a case with the same parameters except bigger group size ($g = 2$; $n = 10$). Here, although P can increases during the second generations within groups, it is not enough to make up for losses in the first generations. Note that when peaks in r^o fail to reach the c / b value (after generation 21 in Fig. 2(b) for $g = 2$; $n = 10$), as predicted by Hamilton's rule, P can no longer increase and instead falls during both generations within groups (Fig. 2(a)).

5. Analytic Model Modifications

Multiple generations within groups complicate the simple single-generation model in several ways: 1) kin interactions within groups become possible; 2) the contribution of homogeneous groups of altruists increases—these groups uniquely retain their initial (maximal) level of altruism; 3) the additive frequency-dependent fitness functions can now lead to exponential growth of the population. Yet as we demonstrate below, while not inconsequential, none of these factors are essential to explain why strong altruism increases in randomly formed multigenerational groups—especially when altruism is initially rare. The following three paragraphs elaborate on each issue and describe modifications to the basic model to address them. We follow this with a summary of the results produced by each modification.

5.1. No Kin Selection

In the original model groups are formed by randomly selecting individuals from an infinite population and therefore groups contain unrelated individuals. In a second generation within groups, when the benefits provided by an altruist are divided among other group members, some of this benefit (in the form of additional offspring) will fall to those with the same parent as the altruist. In general the proportion of benefit falling to relatives (defined by common ancestry) in subsequent generations will depend on parameters n, b, and P, but this proportion is bounded by $1 / n$ (Appendix B). This limit is approached for high b and P, but when altruism is initially rare this proportion is much less significant. For instance, for groups of size 4, the maximum possible proportion of benefit that could fall to relatives is $1 / n = 25\%$, but for $P = 0.1$ and $b = 10$ the proportion is actually 0.61% during the second generation within groups. For the same P and b the proportion decreases with larger group size and more generations within groups. To eliminate kin selection we modify our model so that altruists only divide their benefit among non-relatives (Appendix B).

5.2. No Homogeneous Groups

In the infinite population of this model, homogeneous groups of altruists will be randomly created whenever $P > 0$. These groups are unique in being the only group composition for which p does not decrease with successive generations within groups. They are also the fastest growing groups as they contain no free-riding non-altruists. One might suspect that such homogeneous groups account for altruism being able to increase after multiple generations within groups. To check this we modify our model (Appendix B) such that immediately after group formation all homogeneous groups of altruists have one altruist switched to a non-altruist. Note that this artificially decreases P, making it even harder for altruism to evolve.

5.3. No Population Growth

Even with additive (linear) fitness functions, multiple generations within groups can cause a population to grow exponentially (Wilson 1987). To study the effect of stable population size we implement a global carrying capacity by scaling the offspring count of all population members each generation by the inverse of the expected overall growth rate during that generation (Appendix B). This holds the population size constant (albeit infinite) at every generation, but allows groups with more altruists to have relatively more offspring each generation than groups with less.

5.4. Modification Results

Fig. 3 compares the results for each of these modifications with the unmodified model for two generations within groups, $g = 2$. We also include results for the original model where $g = 1$. For each of the three modifications $\Delta_2 P$ is dampened, but still positive given sufficient benefit. This is true even when all of the modifications are imposed simultaneously. That is, in a model where no benefit is given to kin, homogeneous groups are always corrupted, and population size is held constant, strong altruism can still increase after two generations within groups.

In contrast to the *unmodified* dynamic model shown in Fig. 2 where altruism tends to evolve to $P = 1.0$ or $P = 0.0$ given enough generations, it does not *necessarily* evolve to saturation under all these modifications. Corrupting homogeneous groups for example necessarily keeps $P < 1.0$. In the case of a population-level carrying capacity, for $n = 4$, $P = 0.1$, $g = 2$, and $b = 15$, a stable limit cycle is reached in which P oscillates every other generation between 0.616 and 0.636. (Yet as shown in the next section, when stochasticity is introduced populations tend to evolve to one extreme or the other in these models.)

6. Simulation Model

So far, like Hamilton, we have used the assumption of an infinite population in order to calculate the expected distribution of group compositions when individuals are randomly distributed. But infinity here has the special consequence of converting a seemingly stochastic model (random group formation) to a deterministic one—the *expected value* of r^o is produced by *every* group reformation event. For any *finite* population, group reformation events will produce r^o values that fluctuate both above and below zero. Depending on the value of other parameters, randomly occurring above zero r^o values constitute a *second* mechanism by which strong altruism can increase (albeit rarely), even if groups are reformed every generation (data not shown).

We now transform the analytic model above into a computer simulation of a finite evolving population and the following stochastic features (see Appendix C for further details):

1. When reforming groups each individual is assigned at random to an unfilled group (rather than by using the expected distribution).
2. Population size is held constant by a global carrying capacity—each generation excess offspring are removed at random without regard to altruistic trait or group membership.
3. Altruists' benefits are distributed evenly (in units of whole offspring) to other group members with any remainder distributed among other members randomly.

Fig. 4 shows results obtained with this computer simulation where each data point represents the results of a set of 1,000 runs done with the given parameter values but different random number seeds. Because here there is no mutation, empirically we observe that $P = 1.0$ and $P = 0.0$ act as stable equilibrium points and intermediate values do not persist indefinitely. All runs were done until one of these equilibrium points was reached and we use the percentage of runs reaching altruist saturation, $P = 1.0$, as a measure of how readily altruism evolves under the given conditions. As was the case for the analytic model, Fig. 4(a) shows that both higher starting P and higher benefit values favor selection for altruism and Fig. 4(b) shows that less altruistic benefit is required to evolve altruism for smaller group sizes.

Note that the stochasticity of the simulation model appears to enhance selection for altruism compared to the analytic model. The dynamic example mentioned earlier for the scaled analytic model resulted in a stable limit cycle where P oscillated in a small range slightly above $P = 0.6$. Here for the same parameters ($P = 0.1$, $n = 4$, $b = 15$, $g = 2$) 99.9% of 1,000 runs resulted in altruism saturation (Fig. 4(a)). In the scaled analytic model if benefit is lowered to $b = 9$ while keeping other parameters the same, then P goes to zero, yet in the simulation model for this benefit level 47.4% of runs evolved to $P = 1.0$ (Fig. 4(a)).

We now investigate the effect of migration in our simulation model where the migration rate, m, specifies the probability that an individual will leave its group during each generation, moving to a randomly selected group (weighted proportionately by group size). The idea here is that larger, thriving groups are proportionately more attractive to migrants, but similar results obtain when migrants join groups at random, independent of group size. Fig. 5 shows how the interaction between the number of generations spent within groups and the migration rate influence selection for strong altruism. Predictably, migration lessens selection for altruism (Fig. 5) by working to dampen the positive assortment, r^o, between actors and recipients each generation, but for intermediate numbers of generations spent within groups, even at relatively high migration rates (i.e. 30%), strong altruism evolves to saturation in some runs.

Fig. 5 also shows that even without migration ($m = 0.0$) intermediate numbers of generations within groups are most favorable to the evolution of altruism. The advantage of an intermediate number of generations is consistent with similar findings in haystack models (Wilson 1987) and models of biased sex ratios (Wilson and Colwell 1981). Note that in the simulations of Fig. 5 it is initially unlikely that any homogeneous groups of altruists will form. With initial $P = 0.1$ and $n = 4$ the probability of forming homogeneous altruist groups is one in 10,000 and only 250 groups are formed ($N = 1{,}000$, $n = 4$) at each group reformation. Yet, in the absence of homogeneous groups strong altruism can still initially increase overall even as p declines in every group. This is because groups with a higher frequency of altruists grow faster—population P increasing while every group p decreases is an example of Simpson's paradox (Simpson 1951, Sober and Wilson 1998). With more generations within groups P must eventually decrease as the altruists are eliminated from every group. Altruism evolves most readily when the number of generations spent within groups takes full advantage of the increase in P due to Simpson's paradox, but avoids the inevitable decline in P.

Fig. 6 illustrates this tension. Here representative individual runs are shown for 2, 4, and 10 generations within groups using the same parameters as Fig. 5 without migration ($m = 0.0$). To aid in comparison the same random number seed (same initial group distribution) is used in all three runs. For 10 generations within groups ($g = 10$), reformation clearly takes place well after peak P values are reached and altruism eventually goes extinct. For $g = 4$, reformation takes place near peak P values and altruism rapidly evolves towards saturation—even though altruism always decreases in the first generation after reformation. (This is true for $g = 10$ also, but harder to see as the rate of decline after reformation matches the rate before.) On the other hand when groups are reformed every other generation ($g = 2$), the potential additional increase in P that would result from staying within groups longer is lost and altruism increases more gradually. Note that the initial increase in P in these three runs takes place in the absence of homogeneous groups. For $g = 2$ and $g = 4$ no such groups are formed until P reaches about 0.3 (initial $P = 0.1$) and in the case of $g = 10$ homogeneous groups never formed. In contrast, in the infinite analytic model homogeneous groups are always initially present and more generations within groups can allow these fastest-growing groups to become more and more dominant, even if initially rare.

7. Conclusion

The main purpose of this paper is to demonstrate that strong altruism can evolve in randomly formed groups and thereby challenge a presumed theoretic limitation on the evolution of altruism. Although allowing groups to last more than one generation introduces new complications, we have demonstrated that kin selection, homogeneous groups, and population expansion are not essential to account for this phenomenon. The fundamental explanation is that, for many initial conditions, after even just one generation of selection in randomly formed groups, the positive assortment between altruists and their potential recipients increases as measured by the regression coefficient, r^o. The groups that are by chance initially dominated by altruists grow larger compared to other groups and even though the fraction of altruists declines in these groups, the absolute number of altruists poised to benefit other altruists in a subsequent generation increases. On the other hand, the groups that are by chance dominated by non-altruists do not grow as large *and* the relatively few altruists in these groups are eliminated or greatly diminished after one to several generations within groups. This also increases positive assortment as these non-altruists are stuck with each other and will receive less benefit from altruists than they did in the first generation. Of course the few non-altruists lucky enough to end up in altruist-dominated groups are the fittest individuals, but overall the conditions that favored non-altruists in the initial random distribution can switch to favor altruists in subsequent generations.

We emphasize again that even when groups are multigenerational, the vast majority of the benefit provided by altruists will fall to non-relatives—especially when altruism is initially rare (Appendix B). Altruism evolves due to the positive assortment among heritable helping behaviors regardless of whether there is a positive assortment among relatives by descent. The regression coefficient used here, r^o, measures the former. This positive assortment can be viewed equivalently (Frank 1998, Queller 1985, 1992, Sober and Wilson 1998, Wade 1980) as causing selection on the altruistic trait (allele) via inclusive fitness or as causing selection among groups that vary in their trait composition. While interactions among kin in nature no doubt often contribute to the positive assortment of altruistic traits, kin interactions are not in themselves a requirement for altruism to evolve.

Whether strong altruism evolves in nature via mechanisms similar to those illustrated here will depend on the degree to which the assumptions of these models are representative of natural conditions. For instance, in both the analytic and simulation models we demonstrated that strong altruism can evolve even when population size is held constant by a global carrying capacity. In nature, in addition to population-level limits on growth there are often limits on group size. While not explored here, group-level limits will dampen between-group selection for altruism, so further investigation is needed to elucidate the relative import of global vs. local levels of population control in the evolution of altruism. A lack of mutation is also unrealistic. We experimented with mutation in our models (data not shown), but in the simple binary genetics used here a mutation that switches behavioral types exerts pressure towards $P = 0.5$ and thus favors altruism when P is initially low. This is because the more common type experiences more mutations. Even if this bias could be compensated for, low mutation rates are unlikely to alter our basic results, which are robust under fairly high levels of migration among groups and even when homogeneous groups of altruists are "mutated" to contain at least one non-altruist.

This model started with the original assumptions of the random group models (Hamilton 1975, Wilson 1975) and added the idea of multigenerational groups from haystack models (Maynard Smith 1964, Wilson 1987). Just as Wilson (1987) created a simulation model to study a more realistic version of Maynard Smith's (1964) original haystack model, we have created a simulation model that adds finite population size, stochasticity, and mutigenerational groups to the original analytic random group models. Whereas Wilson's (1987) haystack simulation corrected the "worst case" assumption made by Maynard Smith (1964) that groups would persist until altruism was eliminated in all mixed groups; here we correct an opposite "worst case" assumption made in random group models that groups only exist for a single generation. As demonstrated here and in the haystack simulations (Wilson 1987), an intermediate number of generations within groups is most favorable to the evolution of altruism.

Maynard Smith (1998) in discussing different views on the evolution of altruism recently echoed the original findings of Hamilton (1975) and Wilson (1975) and the current consensus opinion when he wrote: "If costs and benefits combine additively, and groups are formed randomly, then altruism cannot evolve. But if altruists tend to associate with altruists, and non-altruists with non-altruists, then altruism can evolve. This conclusion is agreed." Many mechanisms which result in a positive assortment among self-sacrificing behaviors have been proposed including passive methods such as foraging in non-uniform resource distributions which can be depleted (Pepper and Smuts 2002), continuous population viscosity with periodic environmental disturbances (Mitteldorf and Wilson 2000), the coevolution of group joining and cooperative behaviors (Avilés 2002), and the presence of non-participants (Hauert et al. 2002), as well as active methods such as kin recognition (Gamboa et al. 1991), conditional strategies based on past actions (Axelrod and Hamilton 1981, Trivers 1971) or reputation (Nowak and Sigmund 1998, Panchanathan and Boyd 2003), policing (Frank 1995b, 2003), punishment of non-altruists (Boyd et al. 2003, Boyd and Richerson 1992, Fehr and Gächter 2002), and even recognition of arbitrary tags (Riolo et al. 2001). Here we have shown that when groups exist for more than one generation such specific or more complex mechanisms for creating positive assortment, although certainly important if present, are not needed—the positive assortment that develops between randomly created multigenerational groups can suffice for between-group selection to dominate within-group selection and thus for strong altruism to evolve.

Acknowledgements

We wish to thank Leticia Avilés and Mark Bedau for helpful comments on an earlier draft of this manuscript and Len Nunney for helpful discussion.

Appendix A. Analytic Model

Here we describe Hamilton's original model with recursion added to accommodate multiple generations within groups. If a_g, s_g, and n_g are respectively the number of altruists, non-altruists (selfish individuals), and total individuals in a group after g generations spent within groups, then:

$$a_g = a_{g-1}\left(1 + b\frac{a_{g-1} - 1}{n_{g-1} - 1} - c\right), \tag{A.1}$$

$$s_g = s_{g-1}\left(1 + b\frac{a_{g-1}}{n_{g-1} - 1}\right), \text{ and} \tag{A.2}$$

$$n_g = n_{g-1}\left(1 + (b - c)\frac{a_{g-1}}{n_{g-1}}\right). \tag{A.3}$$

In Hamilton's model g is always one, but in our model we vary g by using these equations recursively—inputting the results from one generation into the calculations for the next. Note that when first formed all groups are size n, but after reproduction group sizes vary. (Terms without g subscripts indicate initial values, i.e. n is n_0.)

The overall number of altruists, A_g, and individuals, N_g, in the population after g generation within groups is then the number contributed (after g generations) by groups of every possible original composition ($a = 0$ to n) times the number of such groups expected in a random binomial distribution. If G is the total number of groups, then the expected count of groups with a initial altruists out of n group members is:

$$h(a) = G\binom{n}{a}P^a(1-P)^{(n-a)}. \tag{A.4}$$

The total population values after g generations spent within groups are then given by:

$$A_g = \sum_{i=0}^{n} h(i)\ a_g(i) \quad \text{and} \tag{A.5}$$

$$N_g = \sum_{i=0}^{n} h(i)\ n_g(i), \tag{A.6}$$

where $a_g(i)$ is the a_g value from Eq. (A.1) given the particular starting a value specified by the index i and similarly for $n_g(i)$. Although G is infinite, it cancels in the calculation of $P_g = A_g / N_g$ and $\Delta_g P = P_g - P$.

Appendix B. Analytic Model Modifications

Here we describe modification to the analytic model such that altruists do not give benefit to kin, homogeneous groups are disallowed, and the population size is kept constant.

B.1. No Benefit to Kin

The modified fitness functions for when altruists only give to non-relatives are implemented by substituting the size of altruist kin groups, *k*, for the minus-one term in Eqs. (A.1) and (A.2). The minus-one term subtracted the altruist from the number of its beneficiaries; here we subtract the altruist's kin (those having a common ancestor) as well. A preceding superscript *k* is used to designate fitness calculations that subtract *k* instead of one from a group's altruist count and group size:

$$^{k}a_{g} = {^{k}a_{g-1}}\left(1 + b\frac{^{k}a_{g-1} - k_{g-1}}{n_{g-1} - k_{g-1}} - c\right), \tag{B.1}$$

$$^{k}s_{g} = {^{k}s_{g-1}}\left(1 + b\frac{^{k}a_{g-1}}{n_{g-1} - k_{g-1}}\right), \tag{B.2}$$

where the size of a kin group of altruists in generation *g* is given by:

$$k_{g} = k_{g-1}\frac{a_{g}}{a_{g-1}}. \tag{B.3}$$

This is the size of the kin group in the last generation times an altruist's clutch size for this generation. The initial *k* value $k_0 = 1$ (altruists are only related to themselves). Shifting benefit from kin to non-kin in this way does not affect the total group size and Eq. (A.3) works for calculating n_g.

Note that in the unmodified model the average proportion of a group that is related to an altruist, k_g / n_g, can never be above $1 / n$ and therefore the proportion of an altruist's benefit that falls to kin $(k_g - 1)/(n_g - 1)$ is also bounded by $1 / n$. To see this note that k_g / n_g will be largest within homogeneous groups of altruists compared to mixed groups. In such groups (given our convention that $c = 1$) k_g is multiplied by *b* each generation and total group size also increases with *b*. Therefore the proportion k_g / n_g remains at its original value of $1 / n$. In all other groups this proportion falls with successive generations. Only when *P* is high (so that homogeneous altruist groups are common) or when *b* is high (so homogeneous altruist groups grow proportionally bigger than other groups) is this limit approached.

B.2. No Homogeneous Groups

In an infinite population there will always be some homogeneous groups whenever $P > 0$. For this modification, each time groups are randomly formed we eliminate homogeneous groups by simply converting all groups where the altruist number $a = n$ to groups where $a = n - 1$. In this way all homogeneous groups of altruists have one member switched to a non-altruist. Because non-altruists always increase faster than altruists, groups become more non-homogeneous with each successive generation. Note that this modification causes *P* to decrease, but this effect is small when *P* is small.

B.3. No Population Growth

The global carrying capacity is implemented by scaling back all offspring numbers each generation by $N_{g-1} \,/\, N_g$ where N_g is first calculated without scaling. We use a preceding asterisk to denote values calculated with scaling. For instance, the number of altruists in a group after g generations with scaling is,

$$^{*}a_g = a_g \frac{N_{g-1}}{N_g} \qquad \text{(B.4)}$$

and similarly for group size $^{*}n_{\mathrm{g}}$, where scaling is imposed at each recursion (generation). Whole population values with scaling $^{*}A_{\mathrm{g}}$ and $^{*}N_{\mathrm{g}}$ then sum over $^{*}a_{\mathrm{g}}$ and $^{*}n_{\mathrm{g}}$ instead of a_{g} and n_{g} respectively in Eqs. (A.5) and (A.6) and $^{*}P_{\mathrm{g}} = {}^{*}A_{\mathrm{g}} \,/\, {}^{*}N_{\mathrm{g}}$.

Appendix C. Computer Simulation Model

For each run of the model, individuals (N = 1,000) are initially randomly distributed into groups of size n using a random number generator to assign individuals to unfilled groups. The proportion of altruists and non-altruists is determined by the starting P value. The sequential steps of the simulation are then:

1. In each group the new number of altruists and non-altruists (to the closest whole individual) are determined as in the analytic model above with each remainder benefit offspring allocated at random to other group members.
2. Individuals are chosen at random across the whole population and eliminated until the original population size N is reached.
3. If g generations have passed within groups since the last group reformation, all individuals are randomly assigned to new groups of size n; otherwise if the migration rate, m, is greater than zero, mN individuals are chosen at random from the whole population and moved to new random locations in the population array which is ordered by groups—consequently larger groups are proportionately more likely to send out and receive migrants.

These steps are repeated until an equilibrium at P = 0.0 or P = 1.0 is reached. For runs where n = 6, N was 1,002 instead of 1,000 and initial P = 0.0998 instead of 0.1000 to allow an even distribution into groups.

References

Avilés, L., 1993. Interdemic selection and the sex ratio: A social spider perspective. Am. Nat. 142, 320-345.

Avilés, L., 2002. Solving the freeloaders paradox: Genetic associations and frequency-dependent selection in the evolution of cooperation among nonrelatives. Proc. Natl. Acad. Sci. 99, 14268-14273.

Axelrod, R., Hamilton, W. D., 1981. The evolution of cooperation. Science 211, 1390-1396.

Boyd, R., Gintis, H., Bowles, S., Richerson, P. J., 2003. The evolution of altruistic punishment. Proc. Natl. Acad. Sci. 100, 3531-3535.

Boyd, R., Richerson, P. J., 1988. The evolution of reciprocity in sizeable groups. J. Theor. Biol. 132, 337-356.

Boyd, R., Richerson, P. J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology 13, 171-195.

Breden, F., 1990. Partitioning of covariance as a method for studying kin selection. Trends Ecol. Evol. 5, 224-228.

Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137-140.

Frank, S. A., 1995a. George price's contributions to evolutionary genetics. J. Theor. Biol. 175, 373-388.

Frank, S. A., 1995b. Mutual policing and repression of competition in the evolution of cooperative groups. Nature 377, 520-522.

Frank, S. A. 1998. Foundations of social evolution. Princeton University Press, Princeton.

Frank, S. A., 2003. Perspective: Repression of competition and the evolution of cooperation. Evolution 57, 693-705.

Gamboa, G. J., Reeve, H. K., Holmes, W. G., 1991. Conceptual issues and methodology in kin-recognition research: A critical discussion. Ethology 88, 109-127.

Hamilton, W. D., 1964. The genetical evolution of social behavior i and ii. J. Theor. Biol. 7, 1-52.

Hamilton, W. D., 1972. Altruism and related phenomena, mainly in social insects. Annual Review of Ecology and Systematics 3, 193-232.

Hamilton, W. D., 1975. Innate social aptitudes of man: An approach from evolutionary genetics. In: Fox, R. (Ed.), Biosocial anthropology. John Wiley and Sons, New York, pp. 133-155.

Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002. Volunteering as red queen mechanism for cooperators in public goods games. Science 296, 1129-1132.

Maynard Smith, J., 1964. Groups selection and kin selection. Nature 201, 1145-47.

Maynard Smith, J., 1998. The origin of altruism. Nature 393, 639-640.

Mitteldorf, J., Wilson, D. S., 2000. Population viscosity and the evolution of altruism. J. Theor. Biol.

Nowak, M. A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. Nature 393, 573-577.

Nunney, L., 1985. Group selection, altruism, and structured-deme models. Am. Nat. 126, 212-230.

Nunney, L., 2000. Altruism, benevolence and culture: Commentary discussion of Sober and Wilson's 'Unto Others'. J. Consciousness Stud. 7, 231-236.

Panchanathan, K., Boyd, R., 2003. A tale of two defectors: The importance of standing for evolution of indirect reciprocity. J. Theor. Biol. 224, 115-126.

Pepper, J. W., 2000. Relatedness in trait group models of social evolution. J. Theor. Biol., 355-368.

Pepper, J. W., Smuts, B. B., 2002. Assortment through environmental feedback. Am. Nat. 160, 205-213.

Price, G. R., 1970. Selection and covariance. Nature 277, 520-521.

Queller, D. C., 1985. Kinship, reciprocity and synergism in the evolution of social behavior. Nature 318, 366-367.

Queller, D. C., 1992. Quantitative genetics, inclusive fitness, and group selection. Am. Nat. 139, 540-558.

Riolo, R. L., Cohen, M. D., Axelrod, R., 2001. Evolution of cooperation without reciprocity. Nature 414, 441-443.

Simpson, E. H., 1951. The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society B 13, 238-241.

Sober, E., Wilson, D. S. 1998. Unto Others, the evolution and psychology of unselfish behavior. Harvard University Press, Cambridge, MA.

Sober, E., Wilson, D. S., 2000. Morality and 'Unto Others': Response to commentary discussions. J. Consciousness Stud. 7, 257-268.

Trivers, R. L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35-57.

Wade, M. J., 1980. Kin selection: Its components. Science 210, 665-667.

Wilson, D. S., 1975. A theory of group selection. Proc. Natl. Acad. Sci. 72, 143-146.

Wilson, D. S., 1979. Structured demes and trait-group variation. Am. Nat. 113, 606-610.

Wilson, D. S., 1987. Altruism in Mendelian populations derived from sibling groups: The haystack model revisited. Evolution 41, 1059-1070.

Wilson, D. S., 1990. Weak altruism, strong group selection. Oikos 59, 135-140.

Wilson, D. S., Colwell, R. K., 1981. Evolution of sex ratio in structured demes. Evolution 35, 882-897.

Figure Legends

Fig. 1 Change in altruist frequency ($\Delta_g P$) as a function of altruist benefit, b. (a) and (b) compare the effect of different starting P values after one ($g = 1$) and two ($g = 2$) generations spent within groups, respectively, where founding group size $n = 4$ (although in (a) the results are the same for all n). (c) Compares the effect of different n for multigenerational groups ($g = 2$) when altruism is rare—here $P = 0.001$. The cost $c = 1$ in all calculations.

Fig. 2 Calculated dynamics in the analytic model for one and two generations spent within groups ($g = 1$, $g = 2$) for groups of two different sizes ($n = 3$; $n = 10$). (a) Shows the dynamics in overall altruist frequency P given different starting P values and different sized groups. (b) Shows the concurrent change in the regression coefficient between actors and recipients, r^o. The critical r^o value of c / b is also shown with a solid horizontal line. Here $P = 0.1$, $b = 10$, and $c = 1$. Both r^o and P are calculated at the end of the indicated generation and after group reformation if it occurs. (a) and (b) use the same legend.

Fig. 3 Change in altruist frequency ($\Delta_g P$) after two ($g = 2$) generations within groups as a function of altruist benefit, b, for several modifications of the original binomial model including preventing homogeneous groups from forming, scaling the population size to its original size each generation, and distributing altruist benefit only to the non-relatives of an altruist. The results for the original model after one ($g =1$) and two ($g = 2$) generations are also shown for comparison. The original $P = 0.25$, $n = 4$, and $c = 1$ in all calculations.

Fig. 4 The percentage of simulation runs reaching altruist saturation as a function of altruist benefit, b. (a) Compares the effect of different starting P values, where the number of generations spent within groups $g = 2$ and initial group size $n = 4$. (b) Compares the effect of different n where $P = 0.1$ and $g = 2$. The total population size $N = 1{,}000$ and cost $c = 1$ in all runs. 1,000 runs were done for each unique set of parameters with different random number seeds. All runs were done until $P = 1.0$ or $P = 0.0$.

Fig. 5 The percentage of simulation runs reaching altruist saturation as a function of the number of generations spent within groups g. Compares the effect of different migration rates, m, where for all runs $P = 0.1$, $g = 2$, $b = 10$, $n = 4$, $N = 1{,}000$, and $c = 1$. 1,000 runs were done for each unique set of parameters with different random number seeds. All runs were done until $P = 1.0$ or $P = 0$. For $m = 0.4$ all runs resulted in $P = 0.0$ (data not shown).

Fig. 6 Dynamics in altruist frequency P for individual simulation runs of $g = 2$, $g = 4$, and $g = 10$ generations spent within groups. For all runs initial $P = 0.1$, $b = 10$, $n = 4$, $N = 1{,}000$, and $c = 1$. To aid in comparison, all three runs were initiated with the same random number seed (same initial distribution into groups).

Fig. 1

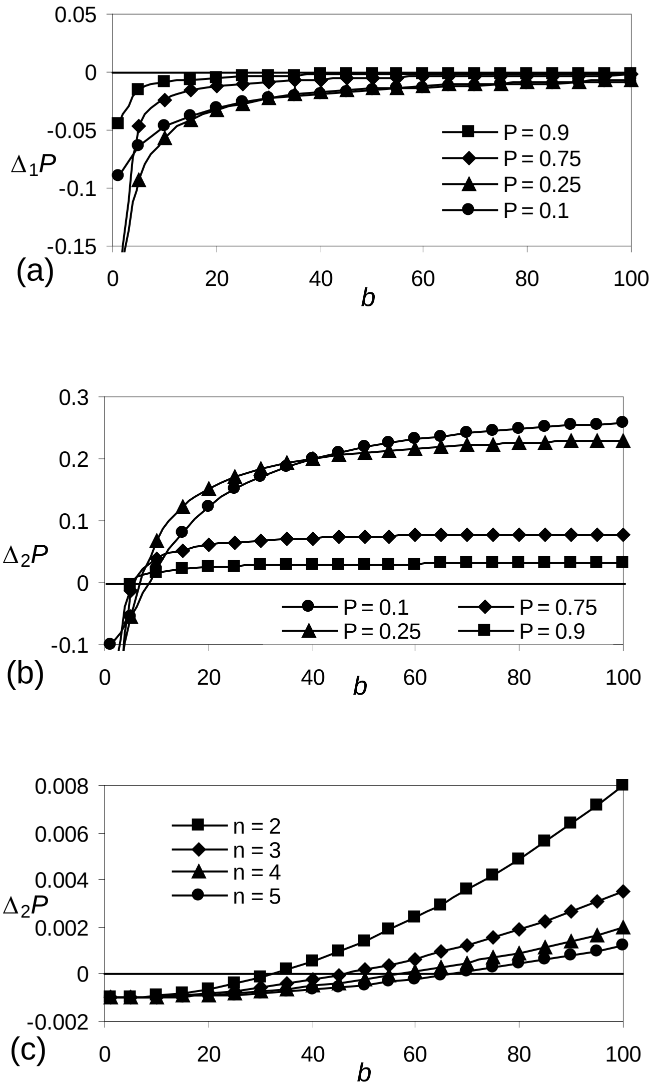

0.05
0
-0.05
$\Delta_1 P$
-0.1
-0.15
(a)
0
20
40
60
80
100
b
P = 0.9
P = 0.75
P = 0.25
P = 0.1
0.3
0.2
0.1
$\Delta_2 P$
0
-0.1
(b)
0
20
40
60
80
100
b
P = 0.1
P = 0.75
P = 0.25
P = 0.9
0.008
0.006
0.004
$\Delta_2 P$
0.002
0
-0.002
(c)
0
20
40
60
80
100
b
n = 2
n = 3
n = 4
n = 5

Fig. 2

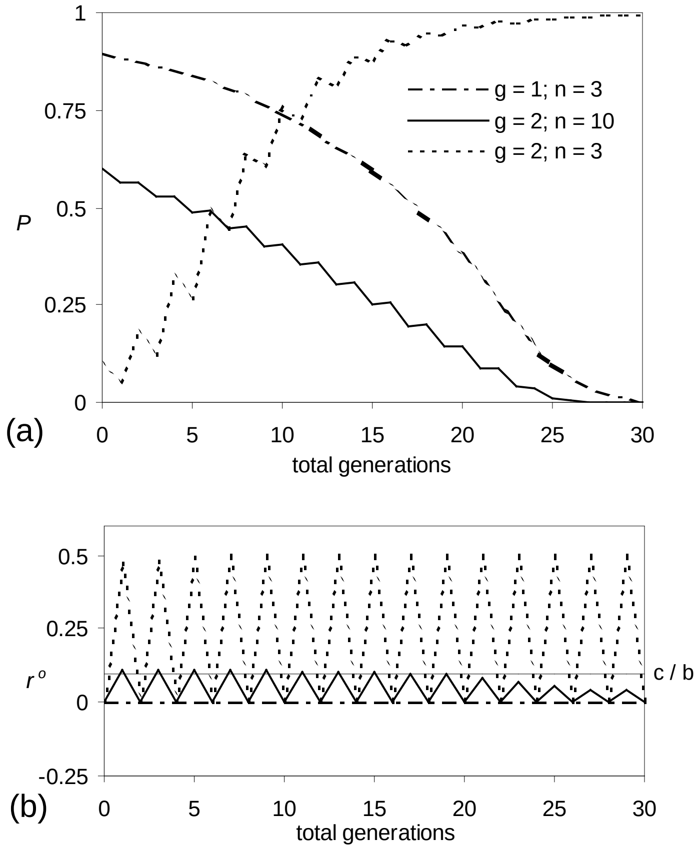

1
0.75
0.5
0.25
0
P
g = 1; n = 3
g = 2; n = 10
g = 2; n = 3
(a)
0
5
10
15
20
25
30
total generations
0.5
0.25
0
-0.25
r o
c / b
(b)
0
5
10
15
20
25
30
total generations

Fig. 3

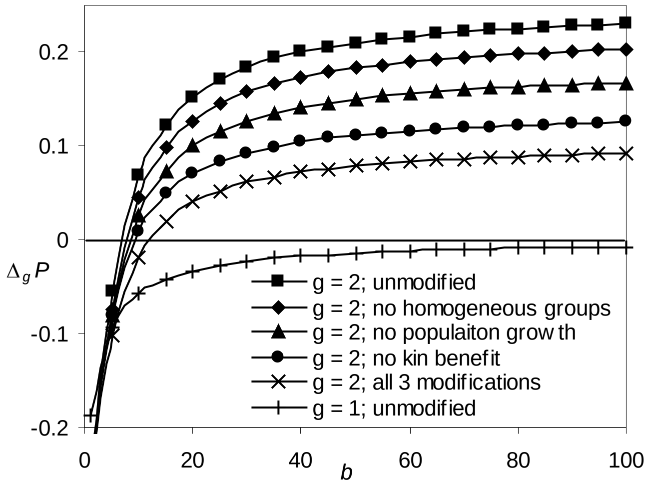
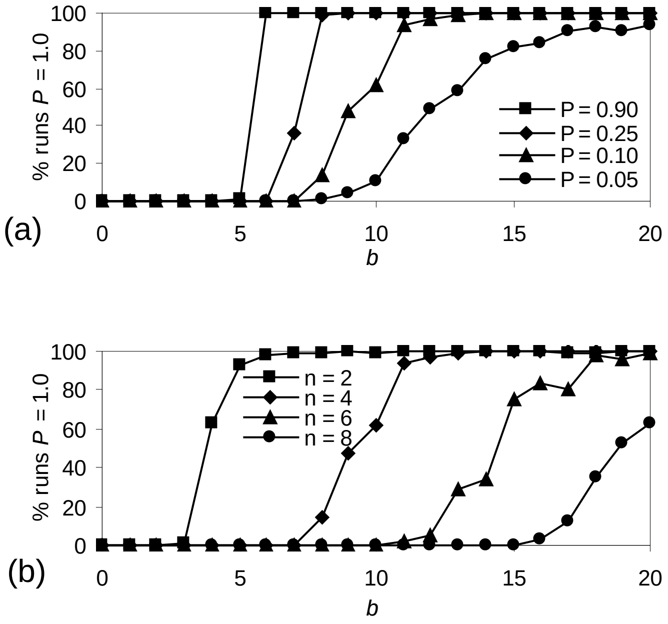

0.2
0.1
0
-0.1
-0.2
$\Delta_g P$
0
20
40
60
80
100
b
g = 2; unmodified
g = 2; no homogeneous groups
g = 2; no populaiton grow th
g = 2; no kin benefit
g = 2; all 3 modifications
g = 1; unmodified

Fig. 4



(a)
100
80
60
40
20
0
% runs P = 1.0
0
5
10
15
20
b
P = 0.90
P = 0.25
P = 0.10
P = 0.05
(b)
100
80
60
40
20
0
% runs P = 1.0
0
5
10
15
20
b
n = 2
n = 4
n = 6
n = 8

Fig. 5

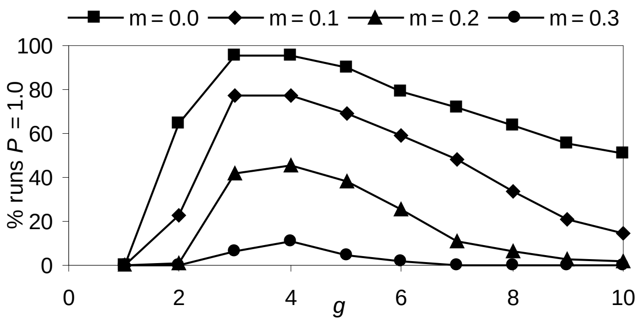


Fig. 6