doi:10.1111/j.1558-5646.2009.00778.x

ON THE EVOLUTION OF HARMING AND RECOGNITION IN FINITE PANMICTIC AND INFINITE STRUCTURED POPULATIONS

Laurent Lehmann,[1,2] Marcus W. Feldman,[1] and François Rousset[3]

[1]*Department of Biology, Stanford University, California 94305*

[2]*E-mail: lehmann@stanford.edu*

[3]*Laboratoire Génétique et Environnement, Université de Montpellier II, France*

Received December 15, 2008
Accepted June 24, 2009

Natural selection may favor two very different types of social behaviors that have costs in vital rates (fecundity and/or survival) to the actor: helping behaviors, which increase the vital rates of recipients, and harming behaviors, which reduce the vital rates of recipients. Although social evolutionary theory has mainly dealt with helping behaviors, competition for limited resources creates ecological conditions in which an actor may benefit from expressing behaviors that reduce the vital rates of neighbors. This may occur if the reduction in vital rates decreases the intensity of competition experienced by the actor or that experienced by its offspring. Here, we explore the joint evolution of neutral recognition markers and marker-based costly conditional harming whereby actors express harming, conditional on actor and recipient bearing different conspicuous markers. We do so for two complementary demographic scenarios: finite panmictic and infinite structured populations. We find that marker-based conditional harming can evolve under a large range of recombination rates and group sizes under both finite panmictic and infinite structured populations. A direct comparison with results for the evolution of marker-based conditional helping reveals that, if everything else is equal, marker-based conditional harming is often more likely to evolve than marker-based conditional helping.

KEY WORDS: Competition, fixation probabilities, genetic diversity, genetic relationship, harming, helping, marker recognition, spite, two-locus models.

Helping behaviors, by which individuals in a population provide fitness benefits to others, are more likely to evolve in the presence of mechanisms that allow discrimination against defectors (Hamilton 1964; Hamilton 1971; Axelrod and Hamilton 1981; Eshel and Cavalli-Sforza 1982). Discrimination may occur if helpers can recognize each other by using conspicuous phenotypic cues, tags, or markers, and provide benefits to other individuals carrying the gene(s) underlying helping, instead of providing benefits to defectors. Both genetic kin-recognition and the so-called green-beard mechanism may involve the expression of helping conditional on actor and recipient bearing identical recognition markers (Hamilton 1964; Dawkins 1982; Grafen 1990). Such conditional helping relies on a tight linkage between the genes underlying helping and those producing the conspicuous markers. If genes for helping and markers were loosely coupled, defectors could acquire the markers expressed by helpers and then receive the benefits of helping without paying the cost. This would ultimately prevent the evolution of marker-based conditional helping.

Genetic kin-recognition based on actor and recipient bearing identical marker alleles might evolve in spatially structured populations (Axelrod et al. 2004; Jansen and van Baalen 2006; Rousset and Roze 2007). Two individuals from the same group (or spatial location) are more likely to have inherited both helping and recognition marker alleles from the same recent common ancestor than are two individuals from different groups. Common ancestry can then lead to the buildup of genetic associations between helping and recognition alleles between individuals that descend from the same group, even in the presence of recombination provided migration is limited and group size is not too

large. Because individuals within groups might interact with resident and immigrant individuals, recognition markers that are identical-by-descent may allow actors to discriminate among categories of recipients defined by markers within groups. Such marker-based discrimination then sustains the evolution of conditional helping under strong population structure (Axelrod et al. 2004; Jansen and van Baalen 2006; Rousset and Roze 2007).

An increase in group size or migration rate erodes population structure and weakens the genetic associations between individuals from the same group, within and across loci. With frequent migration, population structure is likely to vanish. In this case, marker-based helping is no longer expected to evolve. Nevertheless, if the population becomes panmictic but is of finite size, some variation may remain in the propensity of interacting individuals to share alleles identical-by-descent at many loci. Indeed, two offspring of the same parent are always more likely to have inherited identical helping and recognition marker alleles than are two individuals sampled at random from the population. The variation in the ancestry of pairs of interacting individuals within a panmictic population might then still allow actors to discriminate among categories of recipients. This variance supports the evolution of marker-based conditional helping when population size is very small (Traulsen and Nowak 2007) because the probability that individuals descending from the same parent interact in a panmictic population is likely to be small, approximately equal to the inverse of population size in the absence of searching. Thus, whether marker-based conditional helping evolves in panmictic or structured populations, finite population size is a crucial demographic requirement for the evolution of the behavior.

With finite local group (or total population) size individuals that help each other are also more likely to compete for the same local (or global) resources. Competition between interacting individuals has been shown to partially offset the benefits of helping in both finite panmictic (Hamilton 1971) and infinite structured populations (Taylor 1992a,b). Finite population size thus creates ecological conditions in which actors may actually benefit from expressing behaviors that reduce the fecundity of neighbors by harming them, instead of increasing it by helping them, even when the actor suffers a fecundity cost. This follows from the fact that in a finite population a single individual may have a marked effect on population productivity and decrease the intensity of competition experienced by its offspring.

From the green-beard in the red fire ant *Solenopsis invicta*, which kills individuals that do not have it, to bacteria releasing antagonistic compounds in their environment, to maternally transmitted symbionts generating cytoplasmic incompatibility, several examples have been documented where genotypes spread through natural populations by hampering the reproduction of those that do not carry them, thereby reducing the intensity of competition experienced by their carriers (e.g., Werren 1997; Keller and Ross 1998; Riley and Gordon 1999; Brown et al. 2006). It is thus useful to understand not only what are the ecological and demographic conditions leading to the evolution of helping behaviors (cooperation and altruism), but also those conducive to the evolution of harming (exploitation and spite). This might lead to a better understanding of the type of social interactions expected to occur in natural populations.

In this article, we try to understand the conditions under which marker-based conditional harming, whereby an actor decreases the fecundity of recipients conditional on them bearing a different phenotypic cue than the actor, is selected for. To that aim, we analyze the joint evolution of neutral recognition markers and marker-based conditional harming behaviors in a two-locus population genetic framework. The first locus controls the expression of neutral conspicuous markers. The second locus determines the expression of harming, conditional on actor and recipient bearing different conspicuous markers at the first locus. We show that under the Wright–Fisher scheme of reproduction, marker-based conditional harming can evolve for a large range of recombination rates and group sizes, in both finite panmictic and infinite structured populations. A direct comparison with results for the evolution of marker-based conditional helping reveals conditions under which, everything else being equal, the selective pressure favoring marker-based conditional harming is stronger than that on conditional helping. In particular, this is the case when only two conspicuous marker alleles at the recognition locus segregate in the population.

Model

LIFE CYCLE

We consider a population with an infinite number of groups each of finite size N (a list of symbols is given in Table 1). Individuals within groups interact with each of their $N-1$ neighbors. We assume that individuals are haploid and carry two different loci, with two alleles segregating at each locus. The first locus determines the expression of a conspicuous phenotypic feature, which is the basis of recognition: individuals bearing a mutant recognition allele, R, at this locus express a different phenotype from those bearing the wild-type, resident allele, r. Both mutant and resident alleles have the same effect on reproduction and can therefore be regarded as neutral alleles. The second locus controls the expression of a harming behavior. A focal actor bearing a mutant harming allele H at this locus decreases the fecundity of a single recipient by $D/(N-1)$ at a direct fecundity cost $C/(N-1)$ to itself, conditional on the recipient bearing a different allele at the recognition locus than the actor. Those individuals that carry the resident allele, h, at the harming locus do not express any phenotype, but their fecundity is decreased whenever they

Table 1. List of symbols.

Symbol	Definition
N	Deme size
m	Migration rate
r	Recombination rate between gametes
C	Fecundity cost to the actor of expressing the harming (or helping) allele
D	Reduction in the fecundity of recipient resulting from expressing the harming allele
B	Fecundity benefit generated by expressing the helping allele
w_{ij}	Fitness of individual j breeding in deme i
$1+f_{ij}$	Fecundity of individual j breeding in deme i
$1+f_i$	Average fecundity in deme i
$1+f$	Average fecundity in the population
$p_{\mathrm{A}(ij)}$	Frequency (0 or 1) of allele A in individual j breeding in deme i
p_{A}	Average frequency of allele A in the population
$\mathrm{E}_{i,j}[\cdot]$	Average over all i and j ($p_{\mathrm{A}}=\mathrm{E}_{i,j}[p_{\mathrm{A}(ij)}]$)
$\mathrm{E}_{i,j,k\neq j}[\cdot]$	Average over all i, j and $k\neq j$
$\mathbf{p}(t)$	Vector of genotype frequencies in the population
$\Pr(\mathbf{p}(t))$	Probability distribution of the vector $\mathbf{p}(t)$ of genotype frequencies in a finite population at time t
$\mathrm{E}[\cdot]$	Expectation over the distribution $\Pr(\mathbf{p}(t))$ of genotype frequencies
H_{R}	Genetic diversity at the recognition locus in the population
F	Probability of identity-by-descent between two genes randomly sampled at the same locus from two different individuals from the same group
ϕ, γ, and δ	Probability of identity-by-descent between two pairs of genes sampled at two different loci from, respectively, two, three, and four different individuals

interact with individuals carrying the harming allele and a different neutral conspicuous marker.

These phenotypic effects on fitness entail that if harming were expressed unconditionally, regardless of the recipient's genotype, then the act of harming would result in a net fecundity cost C to the actor [$(N-1)$ acts times the cost $C/(N-1)$ per act] and a total effect D on group productivity [$(N-1)$ acts times the cost $D/(N-1)$ to each recipient]. This corresponds to the usual parameterization of the evolution of unconditional helping in patch-structured populations, with the difference that the effect D on group productivity is a benefit in that case (e.g., Eshel 1972; Aoki 1982; Taylor 1992a; Gardner and West 2006). The conditional expression of the behavior will result in a situation in which the cost to the actor (or the total effect on group productivity) varies between zero (never expressing the act) and C (or D).

Events in the life cycle of the population occur in the following order. (1) Each adult individual produces a large number of juveniles with its fecundity being determined by the interactions with its $N-1$ neighbors. (2) Juveniles disperse independently from each other with probability m to another random group, and all adults die. (3) Juveniles fuse randomly to produce diploid zygotes (syngamy), which is immediately followed by meiosis with a recombination rate r between the harming and recognition loci to produce a new generation of haploid individuals. (4) Regulation occurs so that only N individuals are allowed to settle in each group.

GENE FREQUENCY CHANGE AND FIXATION PROBABILITY

We denote by $p_{\mathrm{A}(ij)}$ the frequency (0 or 1) of a mutant allele A (A stands for H or R) in individual j from group i. Individuals j and k from group i then bear different marker alleles with probability $p_{\mathrm{R}(ij)}(1-p_{\mathrm{R}(ik)})+(1-p_{\mathrm{R}(ij)})p_{\mathrm{R}(ik)}$. With our assumptions that individuals interact with each of their $N-1$ group neighbors and that actors harm only recipients bearing different marker alleles (and are harmed by others bearing the harming allele and different maker alleles), the fecundity of individual j from group i (relative to that of an individual not expressing the harming allele) can be written as

$$1+f_{ij}=1+\frac{1}{N-1}\sum_{k,k\neq j}(p_{\mathrm{R}(ij)}(1-p_{\mathrm{R}(ik)}) + (1-p_{\mathrm{R}(ij)})p_{\mathrm{R}(ik)})(-Cp_{\mathrm{H}(ij)}-Dp_{\mathrm{H}(ik)}). \quad (1)$$

The expected change in the frequency p_{A} of allele A over the life cycle, conditional on the current distribution of genotypes in the population (i.e., the current values of $p_{\mathrm{H}(ij)}$ and $p_{\mathrm{R}(ij)}$ for each individual in the population), is then given to leading order in C and D (weak selection) by

$$\Delta p_{\mathrm{A}}=\mathrm{E}_{i,j}[(f_{ij}-f)p_{\mathrm{A}(ij)}]-(1-m)^2\mathrm{E}_{i,j}[(f_i-f)p_{\mathrm{A}(ij)}], \quad (2)$$

where $\mathrm{E}_{i,j}[\cdot]$ denotes the average over all groups in the population and all individuals within groups ($p_{\mathrm{A}}=\mathrm{E}_{i,j}[p_{\mathrm{A}(ij)}]$); $1+f_i$ is the average fecundity in group i (average over $1+f_{ij}$ within group i); and $1+f$ is the average fecundity in the total population (eqs. A1 and A2 of the Appendix).

On substitution of equation (1) into equation (2), one obtains an expression for the expected change in allele frequency Δp_{A}, which is a sum of selection coefficients (e.g., C, D, $(1-m)^2C$, etc.), each weighted by moments of allele frequencies involving genes from the same or different loci, sampled from the same or different individuals (e.g., $\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}]$, $\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{R}(ik)}]$, where $\mathrm{E}_{i,j,k\neq j}$ denotes the average over all i, j, and $k\neq j$, see eq. A5). The moments of alleles frequencies must be evaluated explicitly to close equation (2). Under weak selection (i.e., retaining only first-order

effects of selection), it is sufficient to evaluate such moments under neutrality only. Any effect of selection on the moments of alleles frequencies will involve at least first-order effects, which will result in second-order or higher order effects of selection on allele frequency change. These second-order effects are neglected when the change in gene frequency is evaluated to the first order only (e.g., Kirkpatrick et al. 2002; Otto and Day 2007). We will evaluate the moments of gene frequencies in two different ways because we analyze the coevolution of recognition markers and harming alleles under two different but complementary demographic situations:

(1) Infinite structured population: we assume that migration is positive ($m > 0$) so that the evolving population is the collection of an infinite number of groups. Because the average in equation (2) is over all groups in the population, the change of allele frequency Δp_{A} at the level of the total population is deterministic. There is no genetic drift occurring at the level of the total population but genetic drift occurs locally, at the level of the deme, which generates fluctuations in gene frequencies between demes (i.e., relatedness). We then evaluate the moments of allele frequencies as is usually carried out for the infinite island model of dispersal (see the Appendix), by assuming that genetic associations (linkage- disequilibrium, relatedness) are fast variables that equilibrate before any significant change in allele frequencies (slow variables) has occurred at the level of the total population (e.g., separation of time-scale argument, Otto and Day 2007). This approximation is valid when the strength of selection is much weaker than both the recombination and the migration rates (e.g., Nagylaki 1993; Kirkpatrick et al. 2002; Roze and Rousset 2008).
(2) Finite panmictic population: we assume no migration ($m = 0$), which is the same as considering a single panmictic population of fixed finite size N, where the fate of any allele in the absence of mutation is either fixation or loss from the population. Our aim is then to evaluate the probability of fixation π_{A} of allele A in this population. For this case, we simplify the previous notation by dropping the group subscript i and denote by $p_{\mathrm{A}(j)}$ the frequency (0 or 1) of allele A in individual j of the population, $1 + f_j$ the fecundity of that individual (equivalent to eq. 1 with subscript i dropped), and by $1 + f$ the average fecundity in the total population (average over N individuals only). With these assumptions, equation (2) boils down to

$$\Delta p_{\mathrm{A}} = \mathrm{E}_j[(f_j - f)p_{\mathrm{A}(j)}], \tag{3}$$

where $\mathrm{E}_j[\cdot]$ denotes the average over all individuals in the population. Because the number of individuals is finite, Δp_{A} is now a random variable. That is, equation (3) is interpreted as the expected change in allele frequency conditional on the current realization of genotype frequencies in the population, which is a random variable because genetic drift causes the population to take different sample paths. By taking the expectation of Δp_{A} over all possible sample paths one obtains the unconditional expected change in allele frequency (see eqs. A20–A22), which allows us to calculate directly the fixation probability π_{A} by evaluating exactly under neutrality the moments of allele frequencies that will appear in equation (3) (no assumption of separation of time scale here).

HARMING VERSUS HELPING

We will compare our results with those obtained previously by Rousset and Roze (2007) for the evolution of marker-based conditional helping under the same infinite-island life cycle as that described above. For conditional helping, it was assumed that individuals bearing allele H provide a benefit $B/(N-1)$ to a neighbor [instead of harming it with intensity $D/(N-1)$], whenever actor and recipient carry the same conspicuous feature. The model for conditional helping can then be obtained directly from equation (1) by substituting $-D$ with B, and $p_{\mathrm{R}(ij)}(1 - p_{\mathrm{R}(ik)}) + (1 - p_{\mathrm{R}(ij)})p_{\mathrm{R}(ik)}$ with its complement $p_{\mathrm{R}(ij)}\, p_{\mathrm{R}(ik)} + (1 - p_{\mathrm{R}(ij)})(1 - p_{\mathrm{R}(ik)})$, which is the probability that individual j and k from group i bear identical conspicuous marker alleles. With these changes our equation (1) is equivalent to equation (1) of Rousset and Roze (2007).

Results

INFINITE STRUCTURED POPULATION

Invasion condition for harming

We find that the change in the expected frequency of the conditional harming allele H under weak selection is given by

$$\Delta p_{\mathrm{H}} = p_{\mathrm{H}}(1-p_{\mathrm{H}})H_{\mathrm{R}}\Big(-C(1-F) - D(F-\phi) + \big(D+C\big)(1-m)^2 \times \left[\frac{1}{N}(1-F) + \frac{1}{N}(F-\phi) + \left(\frac{N-2}{N}\right)(F-\gamma)\right]\Big), \tag{4}$$

where $H_{\mathrm{R}} = 2p_{\mathrm{R}}(1-p_{\mathrm{R}})$ is the genetic diversity in the population at the recognition locus; F measures the probability that two genes randomly sampled at the same locus from two different individuals from the same group are identical-by-descent (i.e., the two alleles stayed in the same deme and coalesced in that deme); ϕ (γ) is the probability that two pairs of genes sampled at two different loci from two (three) different individuals are identical-by-descent (see eqs. A5–A12). The probabilities ϕ and γ are measures of association between two pairs of homologous genes and are complicated functions of the demographic and genetic parameters N, m, and r (eqs. A15–A19).

The term $1 - F$ in equation (4) is the probability that a pair of genes does not coalesce and $H_{\mathrm{R}}(1 - F)$ then gives the probability that two recognition genes sampled from two different individuals are different. The term $H_{\mathrm{R}}(1 - F)$ can thus be interpreted as the

probability of interaction between a focal individual carrying the mutant harming allele and a recipient carrying a different marker allele from the focal individual, in which case the focal individual harms the recipient at a fecundity cost C to itself. The term $H_R(F - \phi)$ can be interpreted as the probability of interaction between the focal individual and an individual carrying a harming allele identical-by-descent to that of the focal individual and a different marker allele from the focal individual ($F - \phi$ is the probability that, among two pairs of homologous genes sampled from two distinct individuals, one coalesces within the deme, but the other, here at the recognition locus, does not), in which case the focal individual is harmed and loses D fitness units. Finally, $H_R(F - \gamma)$ can be interpreted as the probability of interaction in the focal deme between an individual, which is different from the focal individual but carries a harming allele identical-by-descent to that of the focal individual, with a third individual that carries a different marker allele from the former individual ($F - \gamma$ is the probability that, among two pairs of homologous genes sampled from three distinct individuals, one pair coalesces, but the other does not), in which case the former harms the latter at a cost to the former.

With these three interaction probabilities, the term in brackets in the second line in equation (4) can be thought of as the average frequency of interactions in the focal deme in which an actor bearing a harming allele identical-by-descent to that of the focal individual (including the focal individual himself) harms a recipient bearing a different phenotypic cue from the actor, and which causes a decrease in average patch productivity by $D + C$. This decrease in focal deme productivity decreases the intensity of competition faced by the focal individual's offspring, which then have a higher chance of reaching adulthood, provided they stay in the focal deme and compete against other offspring from that deme (hence the term $(1 - m)^2$ in eq. 4). The last term in equation (4) is thus the fitness benefit of harming accruing to the focal individual.

If pairs of alleles sampled at the harming and recognition loci in different individuals from the same group were to coalesce independently of each other (no associations between alleles from different loci in different individuals), we would have $\phi = \gamma = F^2$, in which case equation (4) reduces to $\Delta p_H = -Cp_H(1 - p_H)H_R(1 - F)$ (obtained by using eq. A15 at equilibrium). The direction of selection on the mutant allele is then given by $-C > 0$, which is the classical result for the selective pressure on an unconditional social behavior under our life-cycle assumptions (Taylor 1992a; Taylor 1992b). With limited dispersal ($m < 1$) and finite group size, alleles sampled at different loci from two different individuals are unlikely to be independent because the two individuals may descend from a common ancestor and may thus carry replica chromosomes. Individuals bearing recognition markers identical-by-descent are then also likely to bear harming alleles identical-by-descent; therefore $\phi > F^2$ and $\gamma > F^2$, in which case harming might eventually be selected for.

Harming versus helping

Using the explicit expressions for F, ϕ, and γ in equation (4) (eqs. A15–A19), we find that the mutant harming allele H might be selected for when group size is large (neglecting terms of higher order than $1/N$) if

$$\frac{C}{D} < \frac{(1-m)^2(1-r)^2}{N\{m(1-r)+r\}\{2-m(1-r)-r\}}, \tag{5}$$

which is a monotonically decreasing function of N, m, and r. We now compare this invasion condition with that obtained for conditional helping. Substituting the explicit expressions for F, ϕ, and γ into the condition for invasion of conditional helping (eq. 2 of Rousset and Roze 2007 or eq. A13 of the Appendix of this article), we find that if allele H results in helping, conditional on the recipient bearing an identical recognition marker allele to the actor, it is selected for under large group size if

$$\frac{C}{B} < \frac{H_R}{1-H_R} \times \frac{(1-m)^2(1-r)^2}{N\{m(1-r)+r\}\{2-m(1-r)-r\}}. \tag{6}$$

Because the ratio $H_R/(1 - H_R)$ is equal to one only when the two recognition alleles are equally frequent, is smaller otherwise, and tends to zero when the frequency of one of the two recognition markers is close to fixation, the selective pressure on marker-based conditional helping is generally weaker than that on marker-based conditional harming when there are only two recognition alleles in the population. It can be shown that this result holds regardless of population size, and when population size is small the difference between the two selection pressures becomes even more pronounced. However, if there are more than two recognition alleles segregating in the population, the diversity H_R can exceed one half (with K neutral marker alleles, the maximal diversity is $(K - 1)/K$). In this case, the selective pressure on conditional helping may be stronger than that on conditional harming, but this will depend on the parameter values.

Invasion condition for recognition markers

Following similar calculations to those carried out in the Appendix for the harming allele (e.g., eqs. A5–A12), we find that the change in the frequency of the recognition allele R is given by

$$\Delta p_R = p_H p_R(1 - p_R)(2p_R - 1)\big(C + D\big)Z, \tag{7}$$

where Z is a function of N, m, r, which is always positive and is exactly the same as that obtained for the change in the recognition allele under conditional helping (Rousset and Roze 2007, eq. 19 of their appendix). Hence, unless an extrinsic force maintains polymorphism at the recognition locus, selection will completely eliminate the diversity at the recognition locus, ultimately

preventing the evolution of conditional harming. This is exactly the same result as that found for marker-based conditional helping and whose implications were discussed at length in Rousset and Roze (2007).

Finally, we mention that we have also investigated the joint evolution of conditional harming and recognition markers including second-order effects of selection in exactly the same way as has been carried out previously for conditional helping, and which resulted in the identification of stable internal polymorphism under a restricted set of parameter values (Rousset and Roze 2007). However, for conditional harming we have not found a stable internal polymorphism for a biologically relevant set of parameter values. This suggests that polymorphism is unlikely to be maintained by frequency-dependent selection in the case of conditional harming unless introduced by mutation.

FINITE PANMICTIC POPULATION

Invasion condition for harming

We find that the fixation probability of a single mutant harming allele H, averaged over the two marker allele backgrounds in which the mutant allele might eventually appear, is given under weak selection by

$$\bar{\pi}_{\mathrm{H}} = \frac{1}{N} + H_{\mathrm{R}}(0)\Bigg(-CP_1 - DP_2 + (D+C)\left[\frac{1}{N}P_1 + \frac{1}{N}P_2 + \left(\frac{N-2}{N}\right)P_3\right]\Bigg), \quad (8)$$

where $1/N$ is the fixation probability under neutrality; $H_{\mathrm{R}}(0) = 2p_{\mathrm{R}}(0)\{1 - p_{\mathrm{R}}(0)\}$ is the genetic diversity at the recognition locus when the mutant is initially introduced into the population; P_1, P_2, and P_3 are functions of N and r (eqs. A27–A35).

The term $H_{\mathrm{R}}(0)P_1$ in equation (8) gives the sum over the whole invasion time of the mutant allele of the probability that a focal individual carries the mutant harming allele H and interacts with another individual bearing a different marker allele from the focal individual. Therefore, $H_{\mathrm{R}}(0)P_1$ can be thought off as the counterpart of $p_{\mathrm{H}}(1 - p_{\mathrm{H}})H_{\mathrm{R}}(1 - F)$ in the deterministic model given by equation (4). The term $H_{\mathrm{R}}(0)P_2$ gives the sum over the whole invasion time of the probability that a focal individual bears allele H and interacts with another individual bearing allele H but a different marker allele (counterpart of $p_{\mathrm{H}}(1 - p_{\mathrm{H}})\, H_{\mathrm{R}}(F - \phi)$ in the deterministic model); that is, the alleles sampled at the harming locus in the two individuals have coalesced but those sampled at the recognition locus have not. Finally, $H_{\mathrm{R}}(0)P_3$ gives the sum over the whole invasion time of the probability that a focal individual carries allele H and a second individual carries the same allele and interacts with a third individual that carries a different marker allele from the second (counterpart of $p_{\mathrm{H}}(1 - p_{\mathrm{H}})H_{\mathrm{R}}(F - \gamma)$ in the deterministic model).

Under weak selection in a finite population, selection favors a mutant allele H whenever it results in a fixation probability higher than that expected under neutrality, and this might occur when the second term in equation (8) is positive, which can be thought of as the gradient of selection on the mutant allele (force of directional selection). Hence, the direction of selection on the harming allele is independent of the initial genetic diversity $H_{\mathrm{R}}(0)$ at the marker locus. The invasion condition for the finite panmictic population model is, therefore, qualitatively similar in form to that given by the infinite structured population model (eq. 4). Further, the interpretation of the probabilities weighting the cost C and damage D depends on qualitatively similar coalescence probabilities. In both cases an actor may interact with two categories of recipients: those that are more and those that are less likely to have inherited recognition markers and behavioral alleles from one common ancestor. This may allow an actor to discriminate between the classes of recipients if recombination is not too frequent.

In Figure 1, we graph the threshold cost-to-benefit value C/D obtained from equation (8) above which the harming allele is selected against as a function of the recombination rate and population size (eq. A46). This figure shows that the harming allele may be selected for under a wide range of parameter values, even if the recombination rate is large. In the absence of recombination, the threshold cost-to-benefit is approximately given by $C/D < 1/2$ (eq. A48). If recognition is perfect and the population remains of

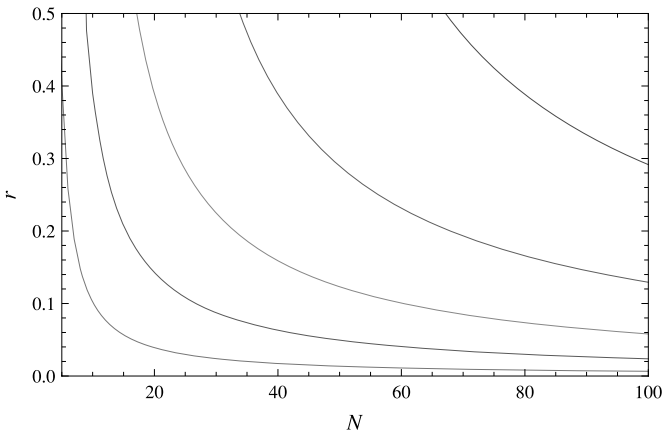

Figure 1. Threshold values of population size, *N*, and recombination rate, *r*, separating positive and negative selective pressure on a single harming allele for a given cost-to-benefit ratio, *C*/*D*, in a panmictic population. The lines were obtained by solving equation (A46) for *N* for given values of *r* and *C*/*D*. For each line, the harming allele is selected against for the combination of *N*, *r* values above it and selected for below it. From the right-most curve to the left-most curve, the *C*/*D* ratio was set to 0.02, 0.04, 0.08, 0.16, and 0.32; that is, an increasing cost-to-benefit ratio, which entails a lower combination of *N*, *r* values under which harming is favored.

finite size, it becomes very likely that a harming allele is favored by selection.

Harming versus helping

If population size is large (neglecting terms of higher order than $1/N$), conditional harming may be selected for in a finite panmictic population when

$$\frac{C}{D} < \frac{1}{Nr(2-r)} \tag{9}$$

(eqs. A46 and A47). We can again compare this invasion condition for conditional harming to that obtained for conditional helping. Results for marker-based conditional helping in a finite panmictic population with recombination have not been derived previously, and we also derived this case here (see eqs. A49 and A50). We find that a single mutant conditional helping allele may be selected for in a finite panmictic population when population size is large if

$$\frac{C}{B} < \frac{H_R(0) - 2r(2-r)}{Nr(2-r)(2-H_R(0))} \tag{10}$$

(eq. A51). This invasion condition depends on the initial genetic diversity $H_R(0)$ at the recognition locus, and is qualitatively the same result as that obtained for the deterministic model for the evolution of conditional helping in an infinite structured population (Rousset and Roze 2007).

As was also the case for the infinite structured population model, the selective pressure on conditional helping is generally weaker than that on conditional harming, but the difference is greater in the finite panmictic population case. Indeed, equation (10) suggests that if the genetic diversity $H_R(0)$ at the marker locus is not large enough relative to the recombination rate r, conditional helping can never invade, whatever the cost-to-benefit ratio. In Figure 2, we have plotted the exact threshold invasion condition as a function of the recombination rate and population size. Comparing Figure 1 and Figure 2, we see that the parameter space that favors the evolution of conditional harming is much larger than that for conditional helping. Finally, we have also evaluated the fixation probability of a mutant recognition marker, but as the resulting expression is complicated and not more informative than that for the deterministic model, we do not discuss this further.

Monte Carlo simulations

Because our results are accurate only to first-order effects of selection (theoretically when C, D, and B are close to zero), we compared predicted changes in fixation probabilities due to the effect of selection with those observed using Monte Carlo simulations. Figure 3 shows a general trend for the present model: the direction of selection is very well approximated by the first-order conditions (even for D as large as 0.5), although the predicted fixation probabilities themselves are not accurate for strong selection.

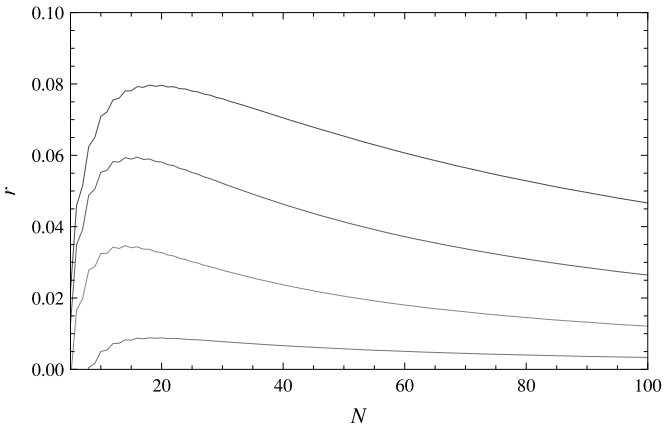

Figure 2. Same as Figure 1 but for a single mutant conditional helping allele introduced in a population. The lines were obtained by solving equation (A50) for N for given values of r and C/B. Helping is selected for below the lines and counter-selected above. In this figure, we assumed that the genetic diversity $H_R(0) = 2p_R(0)(1 - p_R(0))$ at the recognition locus was maximal; we used $p_R(0) = N/2$ for N even and $p_R(0) = (N + 1)/2$ for N odd. These assumptions entail that the selective pressure on helping will be strongest. With lower values of $H_R(0)$ the selective pressure on conditional helping will be reduced, even to the point where actors might actually benefit from harming individuals bearing identical marker alleles (see eq. 10). From the top curve to bottom curve, the C/B ratio was set to 0.02, 0.04, 0.08, 0.16, and 0.32.

THE RELATIONSHIP BETWEEN CONDITIONAL HARMING AND HELPING

Under our life-cycle assumptions, we find that marker-based conditional harming is often more likely to evolve than marker-based conditional helping (in both finite panmictic and infinite structured populations), and that the direction of selection on conditional harming is independent of the genetic diversity H_R at the recognition locus whereas the direction of selection on conditional helping depends on H_R (or initial genetic diversity for the finite panmictic population model). Why is this so? To understand these features it is useful to express the selective pressure on conditional helping in terms of that on conditional harming.

Selection on helping in terms of that on harming

To compare the direction of selection on helping and harming, we now consider that the magnitude of the phenotypic effects on fitness under harming and helping are the same in absolute terms; namely, $|D| = |B|$ and we use the symbol B to denote these effects under both cases (alternatively, we could have chosen D). The gradient of selection S_{harm} on conditional harming, which is the weight for $p_H(1 - p_H)$ in equation (4) and the second term in

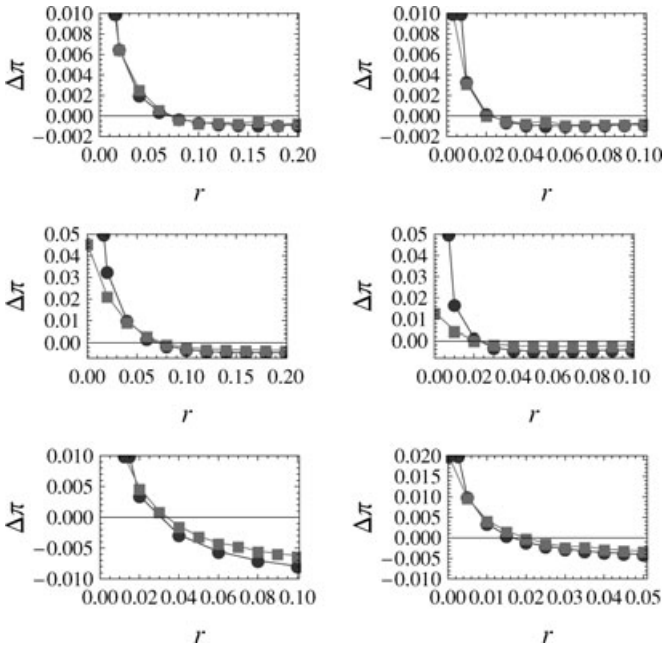

Figure 3. Comparison of the change in the fixation probability $\Delta\pi_H = \pi_H - 1/N$ of a harming allele relative to neutrality as a function of r predicted by the analytical model (circles) with that observed from Monte Carlo simulations (squares) for the case in which a single mutant harming (helping) allele appears simultaneously with a single mutant marker allele in the same individual (i.e., $N - 1$ individuals in the population bear the resident alleles at both loci). The match between the predicted and observed values for the direction of selection is usually good for the direction of selection, regardless of population size and selection strength. The two top panels are for conditional harming, and we have $C = 0.01$ and $D = 0.1$ for $N = 50$ (panel on the left) and $N = 150$ (panel on the right). The middle row of panels is again for conditional harming, but with stronger selection: $C = 0.05$ and $D = 0.5$ for $N = 50$ (panel on the left) and $N = 150$ (panel on the right). The bottom two panels are for conditional helping, and we have $C = 0.01$ and $B = 0.1$ for $N = 50$ (panel on the left) and $N = 150$ (panel on the right).

equation (8), can then be expressed generically as

$$S_{\text{harm}} = -C\kappa_1 + B\kappa_2. \tag{11}$$

The first component, $C\kappa_1$, can be thought of as the net inclusive fitness cost of expressing conditional harming. Under our two models, the weight κ_1 (varying between zero and one) can be interpreted as the probability that the focal individual expresses harming (weighted probability over the whole invasion time for the finite panmictic population model), standardized by the average probability that an individual expresses harming in the population (including the focal individual himself) times the probability that the resulting change in patch productivity affects the competition experienced by the focal individual's offspring. The second component, $B\kappa_2$, in equation (11) can be thought of as the net inclusive fitness benefit of expressing harming. In this case, the weight κ_2 (varying between zero and one) can be interpreted as the average probability that an individual expresses harming in the population times the probability that the resulting change in patch productivity affects the competition experienced by the focal individual's offspring, standardized by the probability that the focal individual gets harmed.

Because any neighbor is either identical to or different from the focal individual at the recognition locus, the probabilities that any neighbor is the recipient of actions by either a harming or a helping focal individual add up to one. It follows that the selective pressure S on unconditional helping (where the mutant allele codes for an act of helping of intensity B, which is shared equally among all neighbors, and which results in a cost C to the actor) can be expressed as

$$S = S_{\text{help}} + S_{\text{harm}}(C, -B), \tag{12}$$

where S_{help} is the selective pressure on conditional helping and $S_{\text{harm}}(C, -B)$ is the selective pressure on conditional harming (eq. 11), where the argument B is replaced by $-B$ to account for the fact that all categories of recipients are helped. The equation for S simply splits the interaction frequencies of a focal individual helping patch members unconditionally into two classes: those that involve individuals carrying identical marker alleles to it and those that involve individuals carrying different marker alleles.

Suppose that the selective pressure on unconditional helping is equal to zero ($S = 0$), but that individuals would benefit from helping recipients carrying identical marker alleles ($S_{\text{help}} > 0$). Then, the net inclusive fitness benefit for providing help to one category of individuals is exactly compensated by an inclusive fitness loss for providing help to the other category. Hence, harming the latter category of individuals would actually increase inclusive fitness. In this special case, the inclusive fitness benefit obtained by helping one category of individuals is exactly equal to the inclusive fitness benefit that would be obtained by harming the other category. This does not imply that the interaction frequencies with the two categories of individuals are equal. On the contrary, when there are only two recognition marker alleles, we have $H_R \leq 1/2$, so that individuals are much more likely to interact with others bearing identical marker alleles than with those bearing different marker alleles. By contrast, when there more than two recognition alleles, say K, then $H_R \leq (K - 1)/K$, in which case it becomes possible that individuals interact more often with others bearing different marker alleles.

In general, the selective pressure on unconditional helping is unlikely to be equal to zero ($S \neq 0$), in which case it can be expressed as

$$S_{\text{help}} = S + C\kappa_1 + B\kappa_2, \tag{13}$$

which, by comparing with equation (11), shows that whether the inclusive fitness benefit of conditional helping will be larger than that under conditional harming depends on whether the inclusive fitness benefit of unconditional helping is greater than zero. From equations (11) and (13), the selective pressure on conditional helping is stronger than that on conditional harming when

$$S > -2C\kappa_1. \tag{14}$$

Comparing the intensities of selection

For the infinite structured population model, we have $S = -(1 - F)C$ (eq. A14 of the Appendix), so that the inclusive fitness benefit under both unconditional helping and harming are exactly the same. From equation (4), one can see that $0 \leq \kappa_1 \leq (1 - F)H_R$; the inclusive fitness cost will never exceed $(1 - F)H_R$ for this model. When there are only two recognitions markers ($H_R \leq 1/2$), the inclusive fitness cost under conditional helping is equal to or greater than that under harming. Then, equation (14) can never be satisfied and the total selective pressure on conditional harming will either be equal to or stronger than that on conditional helping (see also eqs. 5 and 6). But the situation can be reversed when there are more than two recognition alleles. In this case, the inclusive fitness cost under conditional helping may be weaker than that under harming because individuals may interact more frequently with recipients carrying different marker alleles ($H_R \leq (K - 1)/K$ with K marker alleles), and equation (14) can now be satisfied (see also eqs. 5 and 6).

For the finite panmictic population model, we have $S = -C - (B - C)/N$ (see eq. A49), so that the inclusive fitness benefit under both unconditional helping and harming no longer cancel each other out. This stems from the fact that when interactions are unconditional, harming is selected for in a finite panmictic population (Hamilton 1971) whereas the sign of the behavioral effect on neighbors does not affect the direction of selection in the infinite structured population model. Everything else being equal, the dice are thus more loaded in favor of conditional harming than in favor of conditional helping in finite panmictic populations (e.g., compare eqs. 5 and 9).

More generally, the above arguments allowing us to contrast the selective pressure on helping and harming should apply more widely. They should apply whenever a focal individual may express conditionally helping or harming with complementary probabilities over all the events it might face throughout the life cycle and when the selective pressure on conditional harming can be expressed as equation (11). For instance, this is the case when a focal individual helps patch members only when it is a native and harms patch members only when it is an immigrant. It was then found in models with similar baseline life cycle as ours (namely where $S = -(1 - F)C$) that the selective pressure on conditional harming is often, but not always, stronger than that on conditional helping (Lehmann 2003; El Mouden and Gardner 2008). In this case, one has $\kappa_1 = m\{1 - (1 - m)^2/N\}$ and $\kappa_2 = m(1 - m)^2/N$ because an individual expresses harming only when it disperses, in which case no other individuals carries a harming allele identical-by-descent to that of the focal individual, which is then not harmed more often than average (substituting these two expressions into eq. 11 and eq. 13 allows one to recover, respectively, eq. 19 and eq. 12 of El Mouden and Gardner 2008). These expressions allows us to show that equation (14) can be satisfied only if migration is strong because in that case kin competition becomes negligible, while there may still be benefits from helping relatives as long as relatedness remains positive.

In the same vein as just discussed, one may assume that a focal individual perfectly recognizes and helps only individuals born in its natal deme and harms all others, in which case it can again be shown that for a large set of parameter values the selective pressure for conditional harming is stronger than that for conditional helping (Lehmann 2003; $\kappa_1 = \{(1 - m)m + m\} - (1 - m)^2[\{(1 - m)m + m\}/N + m(1 - m)^2F^R(N - 1)/N]$ and $\kappa_2 = (1 - m)^2[\{(1 - m)m + m\}/N + m(1 - m)^2F^R(N - 1)/N]$, where $(1 - m)m + m$ is the probability that a focal individual interacts with another individual that is not born in its natal deme, $m(1 - m)^2$ is the probability that the focal individual is a native and that another native individual interacts with a third individuals that is a nonnative, and $F^R = 1/N + (N - 1)F/N$ is the probability of identity between homologous genes sampled in two native individuals). Finally, we mention that Johnstone and Cant (2008) analyzed a patch-structured model with overlapping generations and sex-biased dispersal of juveniles, where selection on harming was also found to be stronger than that on helping for a large set of parameter values, but in their case the behavior was expressed unconditionally so that the above considerations do not apply to their model.

The effect of H_R on helping and harming

Equations (4) and (8) show that the force of selection S_{harm} on marker-based conditional harming is proportional to the amount of genetic diversity H_R in the population at the recognition locus (or initial genetic diversity for the panmictic population model). This stems from the fact that an interaction between an actor–recipient pair is conditional on the probability that the two individuals carry different marker alleles. This probability is proportional to the genetic diversity H_R because genetic drift reduces diversity within demes (or total population diversity for the finite panmictic population model) by a fractional amount in each generation, until equilibrium is eventually reached (Crow and Kimura 1970; Hartl and Clark 2007). Consequently, the interaction probabilities between actor and recipient determining κ_1 and κ_2 are proportional to H_R (e.g., eq. 4). It follows that the magnitude of population diversity (or initial diversity for the panmictic population model)

should not affect the direction of selection on conditional harming, and equation (4) also holds in the case in which there is an arbitrary number of recognition alleles segregating in the population (in which case $H_R = 1 - \sum_h p_{R,h}^2$, where $p_{R,h}$ is the frequency of the hth neutral marker allele, see eq. A5; notice that we did not demonstrate that this result also holds for the panmictic population case).

In contrast to harming, the direction of selection on marker-based conditional helping was found to depend on the magnitude of genetic diversity H_R at the recognition locus (Rousset and Roze 2007, eq. 2 or eq. A13 in the Appendix, and see eq. A49 for the panmictic population model). In this case, the interaction between an actor and a recipient within a deme is conditional on the probability that the pair carry identical marker alleles, which is the complement of the probability that they carry different alleles. This probability is typically larger but not proportional to the total population probability of identity, given by $1 - H_R$ (see eq. A13), because in each generation genetic drift adds new genetic identity to existing one through the action of coalescence (Crow and Kimura 1970; Hartl and Clark 2007). Thus, the direction of selection on conditional helping will not be proportional to $1 - H_R$ but will rather be affine in H_R (say $\alpha + \beta H_R$). This can be seen from equations (11) and (13): if the direction of selection on conditional harming, S_{harm}, is proportional to H_R, the direction of selection on S_{help} will depend on the value of H_R.

Discussion

Natural selection favors those genotypes that confer on their carriers the highest lifetime reproductive success (fitness defined here as the expected number of offspring that reach the stage of reproduction) because these genotypes are more likely to introduce replicate copies of themselves into the next generation than alternative genotypes. There are two basic and very different means by which a mutant allele can cause its carriers to have a higher fitness than those individuals bearing an alternative, resident allele. Either the mutant confers higher vital rates on its carriers (higher fecundity or survival) or the mutant confers lower vital rates to noncarriers. The latter case can be defined as harming, and when it occurs it may decrease the intensity of competition experienced by a carrier of the harming allele, or that experienced by its offspring.

EVOLUTION OF MARKER-BASED CONDITIONAL HARMING

We have analyzed the joint evolution of neutral recognition markers and marker-based conditional harming under two different but complementary demographic scenarios: finite panmictic and infinite structured populations. Our results show that for a mutant harming allele to be selected for under these two scenarios, the costs to an actor of expressing harming and being harmed must be offset by the benefits obtained from the reduction in competition faced by the actor's offspring, which is due to all actors in the population expressing conditional harming (eqs. 4 and 8). The first cost (cost of harming) depends on the probability that an actor interacts with another individual bearing a different marker allele from itself; the second cost (cost of being harmed) depends on the probability that an actor interacts with another individual bearing the harming allele and a different marker allele from the actor; finally, the benefit of harming depends on the probability that in the population of a focal actor, actors (including the focal actor) interact with other individuals bearing different marker alleles from those of the actor.

The interaction probabilities that weight the costs and benefits of harming depend on population size, N, recombination rate, r, and migration rate m (for the structured population case). Total finite population size (or finite local group size with limited dispersal) results in genetic drift, which entails that an actor may interact with other individuals that have the same common ancestor as the actor (i.e., coalescence of alleles sampled in different individuals occurs). These individuals are then likely to carry the mutant harming allele and the same marker allele as the actor. Interaction with such individuals will reduce the costs of being harmed, but also the benefits, because fewer individuals are likely to be harmed in the population (or local group). In the absence of recombination, the benefits might exceed the costs because an individual bearing the harming allele will never be harmed (perfect recognition), but as long as there is genetic variation in the population, individuals bearing the resident allele are harmed, which results in a decrease in competition felt by the actor or its offspring. Recombination increases the cost of being harmed, because the descendants of an ancestor bearing the harming allele may carry different marker alleles, but it also increases the benefits of harming for the same reason, as more individuals are harmed in the population.

Our results show that the selective pressure on marker-based conditional harming is a decreasing function of the three parameters N, m, and r (eq. 6). This is qualitatively exactly what is usually found for the selective pressure on unconditional helping in a spatially subdivided population with the island model of dispersal (e.g., Eshel 1972; Aoki 1982; Rogers 1990; Taylor and Irwin 2000; Gardner and West 2006), and is also what was found for the invasion of marker-based conditional helping (Rousset and Roze 2007). But our results also suggest that the selective pressure on conditional harming can be stronger than that on conditional helping under otherwise similar life-cycle assumptions. This result holds for both our finite panmictic and infinite structured population scenarios (compare eqs. 5 and 6, and eqs. 9 and 10). The difference between the selective pressure on conditional harming and helping is actually expected to be greatest

in finite panmictic populations (see section "The relationship between conditional harming and helping"). We then observe that a single mutant harming allele can be selected for under a wide range of parameter values, even in the presence of recombination (Fig. 1).

Under our life-cycle assumptions, the selective pressure on conditional harming is often stronger than that on conditional helping because the inclusive fitness benefits obtained through conditional harming and helping (total effect through B and D, assuming they are of similar magnitude) are identical in the infinite structured population case, and larger under harming for the panmictic population case (see section "The relationship between conditional harming and helping"). But at the same time, the inclusive fitness cost of expressing harming (total effect through C, eq. 11) can be lower than that of expressing helping. In particular, when there are only two recognition markers segregating in the population, individuals tend to interact more often with others having identical marker alleles ($H_R \leq 1/2$) and thus pay the direct cost of expressing conditionally helping more often than they would if they expressed harming conditionally. When there are more marker alleles segregating at the recognition locus ($H_R \leq (K-1)/K$ with K marker alleles), the selective pressure on conditional helping can become stronger than that on conditional harming as individuals tend to pay the direct cost of harming more often than that of helping (e.g., eqs. 5 and 6).

Different life-cycle assumptions might also lead to different selective regimes on conditional harming. For instance, in patch-structured populations with maternally transmitted symbionts spreading through host populations by hampering the reproduction of uninfected females (cytoplasmic incompatibility), the condition for invasion of harming was found to be a nonmonotonic, dome-shaped function of group size in a model with similar basic structure to ours (Reuter et al. 2008). It is also well known that introducing overlapping generations with only juvenile dispersal into the type of models considered here can increase the selective pressure on unconditional helping ($S > 0$ in eq. 13, Taylor and Irwin 2000; Irwin and Taylor 2001), which may tip the balance in favor of conditional helping instead of conditional harming, although this is likely to depend on the life-cycle parameter values (Johnstone and Cant 2008).

HARMING VERSUS SPITE

The selective pressure on a mutant harming allele depends on how its expression benefits its carrier and the carrier's relatives by reducing the fecundity of individuals bearing the alternative allele. Hamilton (1970) called "spiteful" a behavior decreasing the fitness of the actor and that of the recipient of the act of harming. We mention that harming might qualify as spiteful (sensu Hamilton (1970)) in both our finite panmictic and infinite structured population models. This can be seen by noting that the net change in the fitness of a carrier due to it expressing the harming allele (and holding everything else constant) can be obtained for the infinite structured population model by setting $H_R(F - \phi)$ and $H_R(F - \gamma)$ equal to zero in equation (4) and for the finite panmictic population model by setting P_2 and P_3 equal to zero equation (8). In both cases, there is a wide range of a parameter values (D, C, m, N, and r) where the resulting change in fitness can be negative (hence expressing the mutant allele results in a net fitness cost to the carrier when everything else is held constant) but the harming allele will still be favored by selection because these direct costs are offset by the reduction in competition felt by the offspring of relatives of the focal individual. However, the conditions under which harming qualifies as spiteful in the sense of Hamilton (1970) are complicated and do not help us here to further understand the ecological and demographic conditions under which conditional harming is selected for, so that we did not present such computations.

MARKER-BASED RECOGNITION, GREEN-BEARDS, AND MATE CHOICE

Marker-based recognition in panmictic populations is sometimes referred to as the green-beard mechanism: a conspicuous phenotypic effect of a gene is recognized by other individuals bearing that gene, and where that gene also has a pleiotropic effect on the behavior of individuals expressing the conspicuous phenotype (Dawkins 1982). Strictly speaking our marker-based conditional harming and helping models for both finite panmictic and infinite structured populations correspond to kin-recognition mechanisms. This is so because in both cases the marker alleles are exchangeable (whether the harming or helping alleles arise on an R or r marker allele background does not affect discrimination). By contrast, under the green-beard mechanism, a particular marker allele is postulated to be associated with a specific behavioral phenotype so that conspicuous markers are not exchangeable (see below for an empirical example).

Independently of the exact nature of the marker-based recognition mechanism (e.g., kin recognition or green beard), our results suggest that the evolution of marker-based helping in panmictic populations may be selected for only under small population size whereas marker-based harming might evolve under a much larger set of parameter values. In the light of this observation it is interesting that the compelling documented examples of green-beards in natural populations are of the harming type. Indeed, the green-beard found in the red fire ant *S. invicta* is of this type (Keller and Ross 1998), where workers homozygous for allele b at the Gp-9 locus kill those individuals that do not contain it (BB queens) while not inducing killing of individuals that do (Bb queens). Note that here allele B cannot be exchanged with allele b without affecting discrimination. Other examples of marker-based conditional harming may be found among bacterial

strains. Some bacteria release into their environment intraspecific antagonistic compounds such as bacteriocins and bacteriophages, which allows them to suppress the growth of competing strains (Riley and Gordon 1999; Gardner et al. 2004). Recognition in this case is molecular with the bacteriocin gene tightly linked to specific immunity genes that block the effect of the bacteriocins; and molecular discrimination may even occur between carriers and noncarriers of isogenic phages (Brown et al. 2006). For bacteria with recurrent cycles of colonization and population growth, the number of founding clones will probably be more relevant than the stationary population size to describe the change in genotype frequency in the population; in that case our parameter N can be thought of as the number of founding clones.

Another situation in which marker-based recognition can be used to discriminate between categories of recipients is in the context of assortative mating (Crow and Kimura 1970; Kirkpatrick 1982; Seger 1985). Females (or males) could prefer to mate with those individuals of the opposite sex that carry identical marker alleles to them at an arbitrary recognition locus that has no direct effects on fitness (Castro and Toro 2006). There are several similarities between our models and the mate choice model of Castro and Toro (2006). These authors also consider that individuals carry two loci: one where arbitrary recognition alleles segregate and another that codes for mating expressed conditionally on pairs of individuals bearing identical marker alleles at the recognition locus; the result is that individuals are more likely to interact with relatives. Castro and Toro then show by simulations that the spread of a choice allele resulting in females mating only with males carrying identical markers is enhanced by finite population size effects (whether the population is panmictic or structured), which corroborates our own results. However, our formalization does not apply directly as it stands to mate choice. By contrast to the model of Castro and Toro, we do not consider a process by which individuals search for others carrying identical (or different) recognition markers. Such a search process could be included in our models by introducing different acceptance probabilities for individuals bearing identical or different marker alleles, so that individuals would stop searching once they have found a partner they accept to interact with (or mate with in the context of mate choice). This deserves further formalization, especially because mate choice is also likely to depend on inbreeding, an inevitable consequence of finite patch or population size.

CONCLUSIONS

Harming behaviors may not be uncommon in nature. For instance, segregation distorter alleles may produce toxins during meiosis to which they but not their alternatives are resistant; the distorter thus increases in frequency by reducing competition for fertilization (e.g., Lyttle 1991; Ridley 2003; Burt and Trivers 2006). Maternally transmitted symbionts can spread through host populations by hampering the reproduction of uninfected females, thereby reducing competition for symbiont carriers (e.g., Werren 1997; Ridley 2003; Burt and Trivers 2006). In all these cases a mutant allele spreads by harming others and this functions because the interaction neighborhood is small enough that the reduction of vital rates of others due to the behavior of the actor, or that of its relatives, decreases the intensity of competition experienced by the actor or its offspring (the interaction neighborhood is actually very small for segregation distorters).

In addition to the results reported here, several models have already identified ecological and demographic conditions for the evolution of harming behaviors in structured populations, where localized migration generates small interaction neighborhoods (Gardner et al. 2004; Lehmann et al. 2006; Gardner et al. 2007; Lehmann et al. 2007a; Johnstone and Cant 2008; El Mouden and Gardner 2008). All these results broaden the scope of biological situations where harming may occur. They show not only that harming might evolve in both finite panmictic and structured populations, but suggest that, under certain situations, harming is actually more likely to evolve than helping. This should encourage behavioral ecologists to seek evidence for conditional harming rather than conditional helping.

ACKNOWLEDGMENTS

We thank A. Gardner for making helpful comments on the manuscript, and are grateful to N. Barton for unleashing the brakes of a congested review and for its helpful suggestions on the manuscript. This work was supported by a grant from the Swiss NSF to LL and by NIH grant GM28016 to MWF. This is publication ISEM 09-088.

LITERATURE CITED

Aoki, K. 1982. A condition for group selection to prevail over counteracting individual selection. Evolution 36:832–842.

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

Axelrod, R., R. A. Hammond, and A. Grafen. 2004. Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. Evolution 58:1833–1838.

Brown, S. P., L. Le Chat, M. De Paepe, and F. Taddei. 2006. Ecology of microbial invasions: amplification allows virus carriers to invade more rapidly when rare. Curr. Biol. 16:2048–2052.

Burt, A., and R. Trivers. 2006. Genes in conflict. Harvard Univ. Press, Harvard.

Castro, L., and M. A. Toro. 2006. Assortative mating through a mechanism of sexual selection. J. Theor. Biol. 243:386–392.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper and Row, New York.

Dawkins, R. 1982. The extended phenotype. Oxford Univ. Press, Oxford.

El Mouden, C., and A. Gardner. 2008. Nice natives and mean migrants: the evolution of dispersal-dependent social behaviour in viscous populations. J. Evol. Biol. 21:1480–1491.

Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. Theor. Popul. Biol. 11:258–277.

Eshel, I., and L. L. Cavalli-Sforza. 1982. Assortment of encounters and evolution of cooperativeness. Proc. Natl. Acad. Sci. USA 79:1331–1335.

Gardner, A., and S. A. West. 2006. Demography, altruism, and the benefits of budding. J. Evol. Biol. 19:1707–1716.
Gardner, A., S. A. West, and A. Buckling. 2004. Bacteriocins, spite and virulence. Proc. R. Soc. Lond. B 271:1529–1535.
Gardner, A., I. C. W. Hardy, P. D. Taylor, and S. A. West. 2007. Spiteful soldiers and sex ratio conflict in polyembryonic parasitoid wasps. Am. Nat. 169:519–533.
Grafen, A. 1990. Do animals really recognize kin? Anim. Behav. 39:42–54.
Hamilton, W. D. 1964. The genetical evolution of social behaviour, II. J. Theor. Biol. 7:17–52.
———. 1970. Selfish and spiteful behavior in an evolutionary model. Nature 228:1218–1220.
———. 1971. Selection of selfish and altruistic behaviour in some extreme models. Pp. 59–91 *in* J. Eisenberg and W. Dillon, eds. Man and beast: Comparative social behavior. Smithsonian Institutions Press, Washington, DC.
Hartl, D., and A. G. Clark. 2007. Principles of population genetics. 4th ed. Sinauer Associates Inc, Sunderland MA.
Irwin, A. J., and P. D. Taylor. 2001. Evolution of altruism in stepping-stone populations with overlapping generations. Theor. Popul. Biol. 60:315–325.
Jansen, V. A. A., and M. van Baalen. 2006. Altruism through beard chromodynamics. Nature 440:663–666.
Johnstone, R. A., and M. A. Cant. 2008. Sex differences in dispersal and the evolution of helping and harming. Am. Nat. 172:318–330.
Karlin, S. 1968. Equilibrium behavior of population genetic models with nonrandom mating: part ii: pedigrees, homozygosity and stochastic models. J. Appl. Prob. 5:487–566.
Keller, L., and K. G. Ross. 1998. Selfish genes: a green beard in the red fire ant. Nature 394:573–575.
Kimura, M. 1963. A probability method for treating inbreeding systems, especially with linked genes. Biometrics 19:1–17.
Kirkpatrick, M. 1982. Sexual selection and the evolution of female choice. Evolution 36:1–12.
Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.
Lehmann, L. 2003. Altruism, Spite, and Choosy Females. Ph.D. thesis, University of Lausanne. Available at http://www2.unil.ch/cyberdocuments/pratique/acces/sciences/these/Lehmann.
Lehmann, L., and F. Rousset. 2009. Perturbation expansions of multilocus fixation probabilities for frequency-dependent selection with applications to the Hill-Robertson effect and to the joint evolution of helping and punishment. Theor. Pop. Biol. 76:35–51.
Lehmann, L., K. Bargum, and M. Reuter. 2006. An evolutionary analysis of the relationship between spite and altruism. J. Evol. Biol. 19:1507–1523.
Lehmann, L., L. Keller, and D. Sumpter. 2007a. The evolution of helping and harming on graphs: the return of the inclusive fitness effect. J. Evol. Biol. 20:2284–229.
Lehmann, L., F. Rousset, D. Roze, and L. Keller. 2007b. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. Am. Nat. 170:21–36.
Lessard, S., and V. Ladret. 2007. The probability of fixation of a single mutant in an exchangeable selection model. J. Math. Biol. 54:721–744.
Lyttle, T. W. 1991. Segregation distorters. Annu. Rev. Genet. 25:511–517.
Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. Genetics 134:627–647.
Otto, S. P., and T. Day. 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton Univ. Press, Princeton, NJ.
Price, G. R. 1970. Selection and covariance. Nature 227:520–521.
Reuter, M., L. Lehmann, and F. Guillaume. 2008. The spread of incompatibility-inducing parasites in subdivided host populations. BMC Evol. Biol. 8:1–11.
Ridley, M. 2003. Evolution. 3th ed. Blackwell Publishers, Oxford.
Riley, M. A., and D. M. Gordon. 1999. The ecological role of bacteriocins in bacterial competition. Trends Microbiol. 7:129–133.
Rogers, A. R. 1990. Group selection by selective emigration: the effects of migration and kin structure. Am. Nat. 135:398–413.
Rousset, F. 2003. A minimal derivation of convergence stability measures. J. Theor. Biol. 221:665–668.
———. 2004. Genetic structure and selection in subdivided populations. Princeton Univ. Press, Princeton, NJ.
Rousset, F., and D. Roze. 2007. Constraints on the origin and maintenance of genetic kin recognition. Evolution 61:2320–2330.
Roze, D., and F. Rousset. 2005. Inbreeding depression and the evolution of dispersal rates: a multilocus model. Am. Nat. 166:708–721.
———. 2008. Multilocus models in the infinite island model of population structure. Theor. Pop. Biol. 73:529–542.
Seger, J. 1985. Unifying genetic models for the evolution of female choice. Evolution 39:1185–1193.
Taylor, P. D. 1992a. Altruism in viscous populations - an inclusive fitness model. Evol. Ecol. 6:352–356.
———. 1992b. Inclusive fitness in a homogeneous environment. Proc. R. Soc. Lond. B 240:299–302.
Taylor, P. D., and A. J. Irwin. 2000. Overlapping generations can promote altruistic behavior. Evolution 54:1135–1141.
Traulsen, A., and M. Nowak. 2007. Chromodynamics of cooperation in finite populations. Plos One 2:270.
Werren, J. H. 1997. Biology of wolbachia. Annu. Rev. Entomol. 42:587–609.
Wolfram, S. 2003. Mathematica, 5th ed. Cambridge Univ. Press, Cambridge.

Associate Editor: M. Van Baalen

Appendix

INFINITE STRUCTURED POPULATION

Gene frequency change

The expected change in the average frequency p_A of allele A over one generation can be written under our life cycle as

$$\Delta p_A = E_{i,j}[w_{ij} p_{A(ij)}] - p_A, \quad \text{(A1)}$$

where w_{ij} is the expected number of offspring of individual j from group i that will reach the next adult generation. This is the Price equation (Price 1970; Hamilton 1970), but as the population is assumed to be of constant size, the mean fitness is equal to one ($E_{i,j}[w_{ij}] = 1$). With our notations introduced in the main text, the fitness of individual j from group i can be written as

$$w_{ij} = \frac{(1-m)(1+f_{ij})}{(1-m)(1+f_i)+m(1+f)} + \frac{m(1+f_{ij})}{(1+f)}. \quad \text{(A2)}$$

(Roze and Rousset 2005; Lehmann et al. 2007b; Rousset and Roze 2007). Assuming that C and D are of small order δ, a Taylor expansion of w_{ij} substituted into equation (A1) produces equation (2) of the main text.

Mutant harming allele

We derive here the change in the expected allele frequency Δp_H of the harming allele when evolution occurs in the infinite island

model of dispersal. The derivation of the change in the expected frequency of a recognition allele is carried out in exactly the same way.

Our analysis closely follows that of Rousset and Roze (2007) for marker-based conditional helping. We also assume that K alleles, denoted $R_1, R_2, \ldots R_K$, may segregate at the recognition locus ($K = 2$ in the main text). We denote by $\mathbf{x}_{\mathrm{R}(ij)} \equiv (p_{\mathrm{R},1(ij)}, p_{\mathrm{R},2(ij)}, \ldots, p_{\mathrm{R},K(ij)})$ the vector with $p_{\mathrm{R},l(ij)}$ being the frequency (0 or 1) of recognition marker l in individual j from group i. Hence, the element l of this vector is equal to one if individual j from group i bears allele R_l, zero otherwise. With this, the effect of social interactions on the relative fecundity of individual j from group i (i.e., relative fecundity without the baseline fecundity unit) can be written as

$$f_{ij} = \frac{1}{N-1}\sum_{k,k\neq j}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})(-Cp_{\mathrm{H}(ij)} - Dp_{\mathrm{H}(ik)}), \quad \text{(A3)}$$

where $\cdot$ denotes the dot product, and with only two alleles segregating in the population this equation reduces to equation (1) of the main text. The average of this equation over all individuals in group i is then given by

$$f_i = \frac{1}{N(N-1)}\sum_{j}\sum_{k,k\neq j}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})(-Cp_{\mathrm{H}(ij)} - Dp_{\mathrm{H}(ik)}), \quad \text{(A4)}$$

whose average over all i gives f ($f = \mathrm{E}_i[f_i]$).

Inserting equations (A3) and (A4) into equation (2) allows us to write

$$\begin{aligned}\Delta p_{\mathrm{H}} &= -C(\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})] - p_{\mathrm{H}}M)\\ &\quad - D(\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})] - p_{\mathrm{H}}M)\\ &\quad + (1-m)^2C(\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(i)}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})] - p_{\mathrm{H}}M)\\ &\quad + (1-m)^2D(\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ik)}p_{\mathrm{H}(i)}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})] - p_{\mathrm{H}}M),\end{aligned} \quad \text{(A5)}$$

where $p_{\mathrm{H}(i)}$ is the average of $p_{\mathrm{H}(ij)}$ over all individuals within group i, and

$$M = \mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})], \quad \text{(A6)}$$

where $\mathrm{E}_{i,j,k\neq j}[\cdot]$ denotes the average over all i, j, and $k \neq j$. By substituting B for $-D$ and $\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)}$ for $(1-\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})$ in equations (A5) and (A6), one obtains the equation for conditional helping (Rousset and Roze 2007, eqs. 6 and 7 of their appendix).

To close the equation for Δp_{H}, we need to evaluate under neutrality all the moments of the form $\mathrm{E}_{i,j,k\neq j}[\cdot]$ appearing in equation (A5). This derivation has already been detailed in the earlier work on conditional helping (Rousset and Roze 2007, p. 3–6 of their appendix). For this reason, we present only the results we need here, without providing the derivations. To evaluate equation (A5), we first need

$$\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}] = Fp_{\mathrm{H}} + (1-F)p_{\mathrm{H}}^2, \quad \text{(A7)}$$

where F is the probability that two genes randomly sampled at the same locus from two different individuals are identical-by-descent (e.g., they stayed in the same deme and coalesced in that deme). We also need

$$\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(i)}] = \frac{1}{N}p_{\mathrm{H}} + \left(\frac{N-1}{N}\right)\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}] \quad \text{(A8)}$$

and

$$\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}(\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})] = p_{\mathrm{H}}[F + (1-F)H_{\mathrm{R}}], \quad \text{(A9)}$$

where $H_{\mathrm{R}} = 1 - \sum_{h=1}^{K} p_{\mathrm{R},h}^2$ is total genetic diversity in the population. We further need

$$\begin{aligned}&\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}(\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})]\\ &\quad = p_{\mathrm{H}}[\phi + (F-\phi)(p_{\mathrm{H}} + H_{\mathrm{R}}) + (1-2F+\phi)p_{\mathrm{H}}H_{\mathrm{R}}],\end{aligned} \quad \text{(A10)}$$

where ϕ is the probability that two pairs of genes sampled at two different loci from two different individuals are identical-by-descent. Finally, we need

$$\begin{aligned}&\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(i)}(\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})]\\ &\quad = \frac{1}{N}\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}(\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})]\\ &\qquad + \frac{1}{N}\mathrm{E}_{i,j,k\neq j}[p_{\mathrm{H}(ij)}p_{\mathrm{H}(ik)}(\mathbf{x}_{\mathrm{R}(ij)}\cdot\mathbf{x}_{\mathrm{R}(ik)})]\\ &\qquad + \left(1-\frac{2}{N}\right)p_{\mathrm{H}}[\gamma + (F-\gamma)(p_{\mathrm{H}} + H_{\mathrm{R}})\\ &\qquad + (1-2F+\gamma)p_{\mathrm{H}}H_{\mathrm{R}}],\end{aligned} \quad \text{(A11)}$$

where γ is the probability that two pairs of genes sampled at two different loci from three different individuals are identical-by-descent.

On substitution of equations (A7)–(A11) into equation (A5), we find after simplification that the change in frequency of the mutant harming allele can be written as

$$\begin{aligned}\Delta p_{\mathrm{H}} = p_{\mathrm{H}}(1-p_{\mathrm{H}})H_{\mathrm{R}}\Bigg(&-C(1-F) - D(F-\phi)\\ &+ (D+C)(1-m)^2\left[\frac{1}{N}(1-F)\right.\\ &\left.+\frac{1}{N}(F-\phi) + \left(\frac{N-2}{N}\right)(F-\gamma)\right]\Bigg),\end{aligned} \quad \text{(A12)}$$

which is equation (4) of the main text.

Mutant helping allele

For comparative analysis, we recall the results obtained for the change in frequency of the mutant allele H when it results in conditional helping (Rousset and Roze 2007, p. 11 of the appendix).

Substituting B for $-D$ and $\mathbf{x}_{R(ij)} \cdot \mathbf{x}_{R(ik)}$ for $(1 - \mathbf{x}_{R(ij)} \cdot \mathbf{x}_{R(ik)})$ in equation (A5) and equation (A6), and using again equations (A7)–(A11), the change in the frequency of allele H for conditional helping can be written as

$$\begin{aligned}\Delta p_H = p_H(1-p_H)\Big(&-C\{1-H_R(1-F)\}\\&+B\{F-H_R(F-\phi)\}-\big(B-C\big)(1-m)^2\\&\times\left[\frac{1}{N}\{1-H_R(1-F)\}+\frac{1}{N}\{F-H_R(F-\phi)\}\right.\\&\left.+\left(\frac{N-2}{N}\right)\{F-H_R(F-\gamma)\}\right]\Big)\end{aligned} \tag{A13}$$

which is equation (1) of Rousset and Roze (2007) written a bit differently. The term $1 - H_R(1 - F)$ in this equation can be interpreted as the probability of interactions between a focal individual carrying the mutant helping allele and a recipient carrying an identical marker allele to that of the focal individual. The term $F - H_R(F - \phi)$ can be interpreted as the probability of interaction between the focal individual and an individual carrying a helping allele identical-by-descent to that of the focal individual and an identical marker allele. Finally, $F - H_R(F - \gamma)$ can be interpreted as the probability of interaction in the focal deme between an actor, which is different from the focal individual but carries a helping allele identical-by-descent to that of the focal individual, with a third individual who carries an identical marker allele to that of the actor.

When the diversity at the recognition locus in a deme is zero ($H_R = 0$), equation (A13) reduces to

$$\begin{aligned}\Delta p_H = p_H(1-p_H)\Big(&-C+BF\\&-(B-C)(1-m)^2\left[\frac{1}{N}+\left(\frac{N-1}{N}\right)F\right]\Big),\end{aligned} \tag{A14}$$

where the term in parentheses is the classical selective pressure on unconditional helping derived by Taylor (1992a), and after simplification (using eq. A15 at equilibrium) further reduces to $\Delta p_H = -p_H(1 - p_H)C(1 - F)$. Comparing equations (A12)–(A14) illustrates that the gradient of selection on conditional helping is equal to the gradient of selection on unconditional helping minus the gradient of selection on conditional harming when D is replaced with $-B$ in equation (A12). This result holds regardless of the specificities of the life cycle and follows from equation (1) (or eq. A5) because the frequency of interaction, $\mathbf{x}_{R(ij)} \cdot \mathbf{x}_{R(ik)}$, with individuals bearing similar recognition alleles (conditional helping) is exactly one minus the frequency of interaction, $1 - \mathbf{x}_{R(ij)} \cdot \mathbf{x}_{R(ik)}$, with individuals bearing dissimilar recognition alleles (conditional harming), so that the sum of interactions over the two cases is equal to that occurring when the trait is expressed unconditionally.

Probabilities of identity-by-descent

To evaluate the change of allele frequency Δp_H explicitly (either the harming our the helping allele), it remains to evaluate F, ϕ, and γ, which can be obtained by writing down recursion equations for these variables (e.g., Kimura 1963; Karlin 1968; Rousset and Roze 2007, p. 10 of their appendix). The recursion for F is given by

$$F' = (1-m)^2\left[\frac{1}{N}+\left(1-\frac{1}{N}\right)F\right]. \tag{A15}$$

The recursions for ϕ and γ depend on δ, which is the probability that two pairs of genes sampled at two different loci from four different individuals are identical-by-descent, and they satisfy

$$\begin{aligned}\phi' &= (1-r)^2K_1 + 2r(1-r)K_2 + r^2K_3\\ \gamma' &= (1-r)K_2 + rK_3\\ \delta' &= K_3,\end{aligned} \tag{A16}$$

where

$$K_1 = (1-m)^2\left[\frac{1}{N}+\left(1-\frac{1}{N}\right)\phi\right] \tag{A17}$$

$$\begin{aligned}K_2 = (1-m)^3\Bigg[&\frac{1}{N^2}+\frac{1}{N}\left(1-\frac{1}{N}\right)(2F+\phi)\\&+\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\gamma\Bigg]\end{aligned} \tag{A18}$$

$$\begin{aligned}K_3 = (1-m)^4\Bigg[&\frac{1}{N^3}+\frac{1}{N^2}\left(1-\frac{1}{N}\right)(1+4F+2\phi)\\&+\frac{2}{N}\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)(F+2\gamma)\\&+\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\left(1-\frac{3}{N}\right)\delta\Bigg].\end{aligned} \tag{A19}$$

Solving these recursions gives complicated expressions for ϕ, γ, and δ, in terms of N, m, and r. Substitution of the resulting expressions into equations (A12) and (A13), and assuming large population size produces equations (5) and (6) of the main text.

FINITE PANMICTIC POPULATION

Probability of fixation

In this section, we give a very brief summary of the argument developed in earlier work to compute the probability of fixation of a mutant allele under weak selection in a one-locus setting (Rousset 2003; Lessard and Ladret 2007), and which can directly be applied to a multilocus setting (Lehmann and Rousset 2009).

From equation (3) the change of allele frequency at time t is given under weak selection by

$$\Delta p_A(t) = E_j[(f_j - f)p_{A(j)}(t) \mid \mathbf{p}(t)], \tag{A20}$$

where we have now made explicit that the change in allele frequency is conditional on the vector $\mathbf{p}(t)$ of genotypes frequencies

in the population at time t. Call $\Pr(\mathbf{p}(t))$ the distribution of $\mathbf{p}(t)$ at time t, conditional on the initial state $\mathbf{p}(0)$ of the population. We will use the expectation operator $\mathrm{E}[\cdot]$ without subscripts to denote an expectation over the distribution $\Pr(\mathbf{p}(t))$ in the population, e.g., $\mathrm{E}[p_{\mathrm{A}}(t)] = \sum_{\mathbf{p}(t)} \Pr(\mathbf{p}(t)) p_{\mathrm{A}}(t)$. With this and equation (A20), the expected unconditional change of allele frequency at time t can be written as

$$\Delta \mathrm{E}[p_{\mathrm{A}}(t)] = \sum_{\mathbf{p}(t)} \Pr(\mathbf{p}(t)) \mathrm{E}_j[(f_j - f) p_{\mathrm{A}(j)}(t) \,|\, \mathbf{p}(t)], \quad \text{(A21)}$$

from which the fixation probability π_{A} of allele A can be obtained as

$$\pi_{\mathrm{A}} = \mathrm{E}p_{\mathrm{A}}(\infty) = p_{\mathrm{A}}(0) + \sum_{t=0}^{\infty} \Delta \mathrm{E}[p_{\mathrm{A}}(t)], \quad \text{(A22)}$$

where $p_{\mathrm{A}}(0)$ is the initial frequency of the mutant in the population (Rousset 2003; Lessard and Ladret 2007; Lehmann and Rousset 2009).

The problem is thus to obtain a closed-form solution to equation (A21). This can be obtained by noting that the term $\mathrm{E}_j[(f_j - f) p_{\mathrm{A}(j)}(t) \,|\, \mathbf{p}(t)]$ can be expressed, as in the previous infinite structured population model, as a sum of selection coefficients (e.g., C, D, etc.), each weighted by averages of products of allele frequencies sampled from the same or different loci, from the same or different individuals (e.g., $\mathrm{E}_{j,k\neq j}[p_{\mathrm{H}(j)} p_{\mathrm{H}(k)}]$, $\mathrm{E}_{j,k\neq j}[p_{\mathrm{H}(j)} p_{\mathrm{R}(k)}]$, see eq. A5). These moments must then be integrated over the distribution $\Pr(\mathbf{p}(t))$. But as any effect of selection on the distribution $\Pr(\mathbf{p}(t))$ will be at least of first order, any effect of selection on this distribution will result in second-order or higher order effects of selection on allele frequency change. Hence, to evaluate the first-order effect of selection on allele frequency change, it is sufficient to consider this distribution under neutrality only. Practically, this consists of evaluating expectations of products of allele frequencies under neutrality in the same way as was carried out in the last section, and more generally in population genetics (e.g., Kimura 1963; Karlin 1968; Crow and Kimura 1970).

To evaluate these moments we use the notation for multilocus models in infinite populations from earlier work (Kirkpatrick et al. 2002; Roze and Rousset 2008) with some specificities to account for total finite population size (Lehmann and Rousset 2009). We will use the expectation operator notation without brackets (e.g., $\mathrm{E}p^{\circ}_{\mathrm{A}}(t) \equiv \mathrm{E}^{\circ}[p_{\mathrm{A}}(t)]$), where the superscript $\circ$ signifies that the expectation is evaluated in the neutral process. Because alleles can be sampled from different individuals, we denote the expectation of an average of products of sets of allele frequencies sampled in different individuals by $\mathrm{E}p^{\circ}_{S}$ with $S \equiv S_1/S_2/\ldots/S_{|S|}$, where each S_j is a set of alleles sampled from the same individual, the "/" symbol separates sets of alleles sampled from distinct individuals, and $|S|$ is the total number of different individuals from which sets of loci have been sampled. For instance,

$$\mathrm{E}p^{\circ}_{\mathrm{HR}} = \mathrm{E}^{\circ}[\mathrm{E}_j[p_{\mathrm{H}(j)} p_{\mathrm{R}(j)}]] \quad \text{(A23)}$$

is the probability that a randomly sampled individual from the population carries allele H at the helping locus and allele R at the recognition locus (expectation of the frequency of gamete HR); and

$$\mathrm{E}p^{\circ}_{\mathrm{HR/R}} = \mathrm{E}^{\circ}[\mathrm{E}_{j,k\neq j}[p_{\mathrm{H}(j)} p_{\mathrm{R}(j)} p_{\mathrm{R}(k)}]] \quad \text{(A24)}$$

is the probability that, among two distinct randomly sampled individuals in the population, one has chromosome HR whereas the other individual carries allele R.

Mutant harming allele

In this section, we derive the average fixation probability of the harming allele H. Because the analysis for finite populations turns out to be more complicated than for infinite populations we assume, for simplicity, that only two alleles segregate at the recognition locus. Dropping subscript i in equation (1) gives

$$f_j = \frac{1}{N-1} \sum_{k,k\neq j} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2 (-C p_{\mathrm{H}(j)} - D p_{\mathrm{H}(k)}) \quad \text{(A25)}$$

and

$$f = \frac{1}{N(N-1)} \sum_{j} \sum_{k,k\neq j} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2 (-C p_{\mathrm{H}(j)} - D p_{\mathrm{H}(k)}). \quad \text{(A26)}$$

On substitution of these equations into equation (A20) we have

$$\begin{aligned} \Delta p_{\mathrm{H}} = & -C \mathrm{E}_{j,k\neq j}\left[p_{\mathrm{H}(j)} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2\right] \\ & - D \mathrm{E}_{j,k\neq j}\left[p_{\mathrm{H}(j)} p_{\mathrm{H}(k)} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2\right] \\ & + C \mathrm{E}_{j,k\neq j}\left[p_{\mathrm{H}(j)} p_{\mathrm{H}} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2\right] \\ & + D \mathrm{E}_{j,k\neq j}\left[p_{\mathrm{H}(k)} p_{\mathrm{H}} (p_{\mathrm{R}(j)} - p_{\mathrm{R}(k)})^2\right]. \end{aligned} \quad \text{(A27)}$$

By taking the expectation of this unconditional change in allele frequency and using the notations introduced above, one obtains the conditional change of allele frequency as

$$\begin{aligned} \Delta \mathrm{E}p_{\mathrm{H}} = & -C\left(\mathrm{E}p^{\circ}_{\mathrm{HR}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/R}} + \mathrm{E}p^{\circ}_{\mathrm{H/R}}\right) \\ & - D\left(2\mathrm{E}p^{\circ}_{\mathrm{HR/H}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/HR}}\right) \\ & + (D + C)\left(\mathrm{E}p^{\circ}_{\mathrm{H}\widehat{/}\mathrm{HR}} - 2\mathrm{E}p^{\circ}_{\mathrm{H}\widehat{/}(\mathrm{HR/R})} + \mathrm{E}p^{\circ}_{\mathrm{H}\widehat{/}(\mathrm{H/R})}\right), \end{aligned} \quad \text{(A28)}$$

where

$$\mathrm{E}p^{\circ}_{U\widehat{/}V} = \frac{1}{N}\mathrm{E}p^{\circ}_{UV} + \left(\frac{N-1}{N}\right)\mathrm{E}p^{\circ}_{U/V}, \quad \text{(A29)}$$

and

$$\mathrm{E}p^{\circ}_{U\widehat{/}(V/W)} = \frac{1}{N}\mathrm{E}p^{\circ}_{UV/W} + \frac{1}{N}\mathrm{E}p^{\circ}_{UW/V} + \left(\frac{N-2}{N}\right)\mathrm{E}p^{\circ}_{U/V/W}. \quad \text{(A30)}$$

Inserting the last two equations into equation (A28) produces

$$\begin{aligned}\Delta \mathrm{E}p_{\mathrm{H}} = &-C\left(\mathrm{E}p^{\circ}_{\mathrm{HR}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/R}} + \mathrm{E}p^{\circ}_{\mathrm{H/R}}\right)\\ &- D\left(2\mathrm{E}p^{\circ}_{\mathrm{HR/H}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/HR}}\right)\\ &+ (D+C)\left[\frac{1}{N}\left(\mathrm{E}p^{\circ}_{\mathrm{HR}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/R}} + \mathrm{E}p^{\circ}_{\mathrm{H/R}}\right)\right.\\ &+ \frac{1}{N}\left(2\mathrm{E}p^{\circ}_{\mathrm{HR/H}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/HR}}\right)\\ &+ \left.\left(\frac{N-2}{N}\right)\left(\mathrm{E}p^{\circ}_{\mathrm{HR/H}} - 2\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}} + \mathrm{E}p^{\circ}_{\mathrm{H/H/R}}\right)\right].\end{aligned} \tag{A31}$$

On substitution of this equation into equation (A22), one can then evaluate the first-order Taylor polynomial for the fixation probability of a mutant harming allele conditional on some initial genotype distribution given by $\mathbf{p}(0) = (p_{\mathrm{HR}}(0), p_{\mathrm{Hr}}(0), p_{\mathrm{hR}}(0), p_{\mathrm{hr}}(0))$, where $p_V(0)$ is the initial frequency of gamete V.

A single mutant harming allele ($p_{\mathrm{H}}(0) = 1/N$) can initially arise on either of the two marker-allele backgrounds at the recognition locus, and our aim is to evaluate the average $\bar{\pi}_{\mathrm{A}}$ fixation probability of a single mutant harming allele H, averaged over the two marker allele backgrounds in which the mutant allele could appear. A single initial copy of the harming allele appears on an R background with probability $p_{\mathrm{R}}(0)$ in which case the initial gamete frequencies in the population are given by $\mathbf{p}_{\mathrm{R}}(0) = (p_{\mathrm{HR}}(0) = 1/N, p_{\mathrm{Hr}}(0) = 0, p_{\mathrm{hR}}(0) = p_{\mathrm{R}}(0) - 1/N, p_{\mathrm{hr}}(0) = 1 - p_{\mathrm{R}}(0))$. The same copy of the mutant appears on the alternative background with probability $1 - p_{\mathrm{R}}(0)$ in which case the initial gamete frequencies in the population is given by $\mathbf{p}_{\mathrm{r}}(0) = (p_{\mathrm{HR}}(0) = 0, p_{\mathrm{Hr}}(0) = 1/N, p_{\mathrm{hR}}(0) = p_{\mathrm{R}}(0), p_{\mathrm{hr}}(0) = 1 - p_{\mathrm{R}}(0) - 1/N)$. With this, the first-order Taylor polynomial of the average $\bar{\pi}_{\mathrm{A}}$ fixation probability of a single mutant harming allele can be written as

$$\begin{aligned}\bar{\pi}_{\mathrm{A}} = \frac{1}{N} + H_{\mathrm{R}}(0)\Bigg(&-CP_1 - DP_2\\ &+ (D+C)\left[\frac{1}{N}P_1 + \frac{1}{N}P_2 + \left(\frac{N-2}{N}\right)P_3\right]\Bigg),\end{aligned} \tag{A32}$$

where

$$\begin{aligned}P_1 = \frac{1}{H_{\mathrm{R}}(0)}\sum_{t=0}^{\infty}\big(&p_{\mathrm{R}}(0)\left[\mathrm{E}p^{\circ}_{\mathrm{HR}}(t) - 2\mathrm{E}p^{\circ}_{\mathrm{HR/R}}(t)\right.\\ &\left.+ \mathrm{E}p^{\circ}_{\mathrm{H/R}}(t) \mid \mathbf{p}_{\mathrm{R}}(0)\right] + (1 - p_{\mathrm{R}}(0))\left[\mathrm{E}p^{\circ}_{\mathrm{HR}}(t)\right.\\ &\left.- 2\mathrm{E}p^{\circ}_{\mathrm{HR/R}}(t) + \mathrm{E}p^{\circ}_{\mathrm{H/R}}(t) \mid \mathbf{p}_{\mathrm{r}}(0)\right]\big)\end{aligned} \tag{A33}$$

$$\begin{aligned}P_2 = \frac{1}{H_{\mathrm{R}}(0)}\sum_{t=0}^{\infty}2\big(&p_{\mathrm{R}}(0)\left[\mathrm{E}p^{\circ}_{\mathrm{HR/H}}(t) - \mathrm{E}p^{\circ}_{\mathrm{HR/HR}}(t) \mid \mathbf{p}_{\mathrm{R}}(0)\right]\\ &+ (1 - p_{\mathrm{R}}(0))\left[\mathrm{E}p^{\circ}_{\mathrm{HR/H}}(t) - \mathrm{E}p^{\circ}_{\mathrm{HR/HR}}(t) \mid \mathbf{p}_{\mathrm{r}}(0)\right]\big)\end{aligned} \tag{A34}$$

$$\begin{aligned}P_3 = \frac{1}{H_{\mathrm{R}}(0)}\sum_{t=0}^{\infty}\big(&p_{\mathrm{R}}(0)\left[\mathrm{E}p^{\circ}_{\mathrm{HR/H}}(t) - 2\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}(t)\right.\\ &\left.+ \mathrm{E}p^{\circ}_{\mathrm{H/H/R}}(t) \mid \mathbf{p}_{\mathrm{R}}(0)\right] + (1 - p_{\mathrm{R}}(0))\left[\mathrm{E}p^{\circ}_{\mathrm{HR/H}}(t)\right.\\ &\left.- 2\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}(t) + \mathrm{E}p^{\circ}_{\mathrm{H/H/R}}(t) \mid \mathbf{p}_{\mathrm{r}}(0)\right]\big).\end{aligned} \tag{A35}$$

Probabilities of identity-by-descent

To evaluate the average fixation probability $\bar{\pi}_{\mathrm{A}}$ explicitly, it now remains to evaluate the neutral moments $\mathrm{E}p^{\circ}_{\mathrm{HR}}$, $\mathrm{E}p^{\circ}_{\mathrm{H/R}}$, $\mathrm{E}p^{\circ}_{\mathrm{HR/H}}$, $\mathrm{E}p^{\circ}_{\mathrm{HR/R}}$, $\mathrm{E}p^{\circ}_{\mathrm{H/H/R}}$, $\mathrm{E}p^{\circ}_{\mathrm{HR/HR}}$, $\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}$, and $\mathrm{E}p^{\circ}_{\mathrm{H/H/R/R}}$. These moments will be affected by reproduction and recombination, and we evaluate them again by using standard methods (e.g., Kimura 1963; Karlin 1968; Crow and Kimura 1970). Only moments involving alleles sampled from the same individual at the two different loci will be affected by recombination because these alleles may descend from different individuals before recombination, and only moments involving genes sampled from different individuals may be affected by reproduction because coalescence of these genes may occur. Over the recombination phase $\mathrm{E}p^{\circ}_{\mathrm{H/R}}$, $\mathrm{E}p^{\circ}_{\mathrm{H/H/R}}$, and $\mathrm{E}p^{\circ}_{\mathrm{H/H/R/R}}$ remain constant, and the remaining expectations change according to the recursions

$$\begin{aligned}\mathrm{E}p^{\circ}_{\mathrm{HR}}{}'' &= (1-r)\mathrm{E}p^{\circ}_{\mathrm{HR}}{}' + r\mathrm{E}p^{\circ}_{\mathrm{H/R}}{}'\\ \mathrm{E}p^{\circ}_{\mathrm{HR/H}}{}'' &= (1-r)\mathrm{E}p^{\circ}_{\mathrm{HR/H}}{}' + r\mathrm{E}p^{\circ}_{\mathrm{H/R/H}}{}'\\ \mathrm{E}p^{\circ}_{\mathrm{HR/R}}{}'' &= (1-r)\mathrm{E}p^{\circ}_{\mathrm{HR/R}}{}' + r\mathrm{E}p^{\circ}_{\mathrm{H/R/R}}{}'\\ \mathrm{E}p^{\circ}_{\mathrm{HR/HR}}{}'' &= (1-r)^2\mathrm{E}p^{\circ}_{\mathrm{HR/HR}}{}' + 2(1-r)r\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}{}'\\ &\quad + r^2\mathrm{E}p^{\circ}_{\mathrm{H/R/H/R}}{}'\\ \mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}{}'' &= (1-r)\mathrm{E}p^{\circ}_{\mathrm{HR/H/R}}{}' + r\mathrm{E}p^{\circ}_{\mathrm{H/R/H/R}}{}',\end{aligned} \tag{A36}$$

whereas over the reproduction phase, one has

$$\mathrm{E}p^{\circ}_{U/V}{}' = \frac{1}{N}\mathrm{E}p^{\circ}_{UV} + \left(\frac{N-1}{N}\right)\mathrm{E}p^{\circ}_{U/V}, \tag{A37}$$

$$\begin{aligned}\mathrm{E}p^{\circ}_{U/V/W}{}' = &\frac{1}{N^2}\mathrm{E}p^{\circ}_{UVW}\\ &+ \frac{1}{N}\left(\frac{N-1}{N}\right)\left(\mathrm{E}p^{\circ}_{UV/W} + \mathrm{E}p^{\circ}_{U/VW} + \mathrm{E}p^{\circ}_{UW/V}\right)\\ &+ \left(\frac{N-1}{N}\right)\left(\frac{N-2}{N}\right)\mathrm{E}p^{\circ}_{U/V/W}.\end{aligned} \tag{A38}$$

$$\begin{aligned}\mathrm{E}p^{\circ}_{U/V/U/V}{}' = &\frac{1}{N^3}\mathrm{E}p^{\circ}_{UVUV} + \frac{1}{N}\left(\frac{N-1}{N}\right)\\ &\times\left(\mathrm{E}p^{\circ}_{V/U} + 2\mathrm{E}p^{\circ}_{UV/V} + 2\mathrm{E}p^{\circ}_{UV/U} + 2\mathrm{E}p^{\circ}_{UV/UV}\right)\\ &+ \frac{1}{N}\left(\frac{N-1}{N}\right)\left(\frac{N-2}{N}\right)\\ &\times\left(\mathrm{E}p^{\circ}_{U/U/V} + \mathrm{E}p^{\circ}_{U/V/V} + 4\mathrm{E}p^{\circ}_{UV/U/V}\right)\\ &+ \left(\frac{N-1}{N}\right)\left(\frac{N-2}{N}\right)\left(\frac{N-3}{N}\right)\mathrm{E}p^{\circ}_{U/V/U/V}.\end{aligned} \tag{A39}$$

By solving the above equations with the initial genotype distribution given by $\mathbf{p}(0)$, we can then evaluate P_1, P_2, P_3. We find that $P_1 = N/(N-1)$, whereas P_2, and P_3 are complicated functions of N and r but they are independent of the frequency of the marker alleles; that is $H_R(0)$ factors out of the numerator in equations (A33)–(A35) and is thus cancelled by the denominator. To the leading order in $1/N$ we have $P_1 = 1$,

$$P_2 = \frac{1}{2} - \frac{(2 - r(2-r))}{4Nr(2-r)} \tag{A40}$$

and

$$P_3 = \frac{1}{2} + \frac{1}{4N} > P_2. \tag{A41}$$

Invasion condition for the mutant harming allele

The expressions for P_1, P_2, P_3 allow us to evaluate equation (A32) explicitly. After simplification carried out with Mathematica (Wolfram 2003), we find that the average fixation probability of the mutant harming allele can be expressed as

$$\bar{\pi}_A = \frac{1}{N} + H_R(0)\left(\frac{DX_1 - CX_2}{X_3}\right), \tag{A42}$$

where

$$\begin{aligned} X_1 = {} & 2N^3[6 - 11N + 6N^2 + (N-4)(N-2)(N-1)r \\ & + (N-2)(N-1)r^2], \end{aligned} \tag{A43}$$

$$\begin{aligned} X_2 = {} & 2(N-1)N^2(1-r)\left[\frac{(2-r)r^2N^4}{1-r} - r(8r-13)N^3\right. \\ & + (12 - (43 - 23r)r)N^2 \\ & \left. - 2(r(14r-27)+11)N + 12(1-r)^2\right], \end{aligned} \tag{A44}$$

and

$$\begin{aligned} X_3 = {} & 2(N-1)N[(3N-2)(6+N(6N-11)) \\ & + (N-1)(N(N(26N-83)+96)-36)r \\ & + (N-1)(36 - 104N + 103N^2 + 4(N-10)N^3)r^2 \\ & - (N-2)(N-1)^2(6+N(2N-9))r^3]. \end{aligned} \tag{A45}$$

From equation (A42), the threshold cost-to-benefit ratio above which a mutant harming allele is selected against is then given by

$$\frac{C}{D} = \frac{X_1}{X_2}. \tag{A46}$$

When group size N is large, equation (A46) simplifies to

$$\frac{C}{D} = \frac{1}{Nr(2-r)}, \tag{A47}$$

which gives equation (9) of the main text. In the absence of recombination ($r = 0$), the threshold cost-to-benefit ratio above which the harming allele is selected against is given by

$$\frac{C}{D} = \frac{N}{2(N-1)}. \tag{A48}$$

Invasion condition for the mutant helping allele

Substituting $-D$ with B and $(p_{R(ij)} - p_{R(ik)})^2$ with $1 - (p_{R(ij)} - p_{R(ik)})^2$ in equation (A27), and using an analogous argument to that above, one can evaluate the average fixation probability of a single mutant allele expressing helping conditionally on both individuals bearing identical phenotypic markers. But the change of fixation probability can also be obtained by using equation (A42) and equation (12) of the main text. Replacing D with B in equation (A42), one has $S_{harm} = H_R(0)(BX_1 - CX_2)/X_3$. For this model, one also has $S = -C - (B - C)/N$, which is the perturbation of the fixation probability of a mutant allele expressing unconditional helping toward neighbors in a finite panmictic population (Rousset 2004; Lehmann et al. 2007a), and was anticipated by Hamilton (1971). With this and using $S_{help} = S - S_{harm}(C, -B)$, the average fixation probability of a single mutant allele expressing helping conditionally can then be expressed as

$$\begin{aligned} \bar{\pi}_A &= \frac{1}{N} + S_{help} \\ &= \frac{1}{N} + \left(-C - \frac{B-C}{N}\right) + H_R(0)\left(\frac{BX_1 + CX_2}{X_3}\right). \end{aligned} \tag{A49}$$

The average fixation probability given by equation (A49) then allows us to evaluate the threshold cost-to-benefit ratio above which the helping allele is selected against as

$$\frac{C}{B} = \frac{NH_R(0)X_1}{N(X_3 - H_R(0)X_2) - X_3}, \tag{A50}$$

which is more complicated than the invasion condition on conditional harming because it involves the function X_3. When group size N is large, the threshold reduces to

$$\frac{C}{B} = \frac{H_R(0) - 2r(2-r)}{Nr(2-r)(2 - H_R(0))}, \tag{A51}$$

which is equation (10) of the main text. In the absence of recombination ($r = 0$), the threshold cost-to-benefit ratio above which the helping allele is selected against is given by

$$\frac{C}{B} = \frac{N\{N^2H_R(0) - 3N + 5\} - 2}{(N-1)\{2 + (3N-5)N - 2N^2H_R(0)\}}. \tag{A52}$$