doi:10.1111/j.1558-5646.2007.00191.x

CONSTRAINTS ON THE ORIGIN AND MAINTENANCE OF GENETIC KIN RECOGNITION

François Rousset[1,2] and Denis Roze[3,4]

[1]*Laboratoire Génétique et Environnement, Institut des Sciences de l'Évolution, Montpellier, France*

[2]*E-mail: Rousset@isem.univ-montp2.fr*

[3]*Station Biologique de Roscoff, Pl. G. Teissier, B.P. 74, 29680 Roscoff, France*

[4]*E-mail: Roze@sb-roscoff.fr*

Received January 5, 2007
Accepted May 28, 2007

Kin-recognition mechanisms allow helping behaviors to be directed preferentially toward related individuals, and could be expected to evolve in many cases. However, genetic kin recognition requires a genetic polymorphism on which recognition is based, and kin discriminating behaviors will affect the evolution of such polymorphism. It is unclear whether genetic polymorphisms used in kin recognition should be maintained by extrinsic selection pressures or not, as opposite conclusions have been reached by analytical one-locus models and simulations exploring different population structures. We analyze a two-locus model in a spatially subdivided population following the island model of dispersal between demes of finite size. We find that in the absence of mutation, selection eliminates polymorphism in most cases, except with extreme spatial structure and low recombination. With mutation, the population may reach a stable limit cycle over which both loci are polymorphic; however, the average frequency of conditional helping can be high only under strong structure and low recombination. Finally, we review evidence for extrinsic selection maintaining polymorphism on which kin recognition is based.

KEY WORDS: **Cooperation, inclusive fitness, multilocus evolution, spatial structure**

Kin recognition based on learned phenotypic cues of other individuals during development (familiarity) is "expected, common, overlooked" (Hamilton 1987). By contrast, the plausibility and actual occurrence of assortative kin recognition based on matching with one's own phenotypic cues have been more debated. Social behaviors dependent on such self-referent kin recognition are now reported in various animal taxa and perhaps in fungi (see Tsutsui 2004 for a recent review and the Discussion for additional references) but conditions favoring their evolution have remained obscure. The maintenance of conditional helping behaviors requires balanced polymorphism at some locus on which phenotypic cues allowing kin recognition are based (henceforth, different recognition alleles at one or more matching locus). Theoretical analyses employing different modeling techniques have reached divergent conclusions about the maintenance of such polymorphism. It is sometimes assumed that balancing selection is an automatic consequence of conditional helping behavior (e.g., Grafen 1990; Crampton and Hurst 1994). Crozier (1986) came to the opposite conclusion based on single-locus models (see also Grosberg and Quinn 1989; Ratnieks 1991; Ratnieks et al. 2007). In these models, each individual helps other individuals carrying the same recognition allele at a given locus. Therefore, the most common recognition allele is favored (because it receives more helping), causing the loss of polymorphism at the matching locus. Based on this conclusion, extrinsic factors that could maintain polymorphism at the matching loci have been discussed (e.g., Crozier 1987; Grosberg and Quinn 1989). On the other hand, Grafen (1990) argued that Crozier's conclusion held only because he did not

consider "cheating" (alleles that do not reciprocate), and that when cheating occurs, gametic disequilibria between the helping allele and rare recognition alleles would build up, which would favor rare alleles. This conclusion has been reiterated on the basis of simulations. In particular, Axelrod et al. (2004) considered a model in which individuals occupy sites on a lattice, and interact with their neighbors. Different alleles code for different strategies: always defect, always help, or help only individuals carrying the same allele at a matching locus. The model assumes asexual reproduction, and a mutation rate of 10^{-3} on each trait. Axelrod et al.'s simulation results show that discriminating helpers can reach high frequencies in the population. More recently, Jansen and van Baalen (2006) considered a similar model, but allowing for recombination between the locus coding for conditional helping and the matching locus (and assuming a mutation rate of 10^{-4} at each locus). They also find that conditional helping can be maintained at an intermediate frequency in the population, this frequency being higher under low recombination than without recombination.

In this work, we analyze the joint evolution of helping versus cheating alleles and of polymorphism at a matching locus, using extensions of previous methods (Barton and Turelli 1991; Kirkpatrick et al. 2002; Roze and Rousset 2005) to analyze multilocus models in spatially subdivided populations. Our analysis quantifies separately the roles of linkage disequilibrium and of other genetic associations in this process, and this allows us to demonstrate a continuum of situations, with kin recognition maintaining balanced polymorphism only in limiting cases with extreme structure and low recombination between the helping locus and the matching locus. In most cases, selection leads to the loss of polymorphism at the matching locus, and to the loss of the helping allele. In between these two cases, there is a range of population structures in whcih the purging effect of selection is weak enough that, when mutation is added to the model, the population may reach a limit cycle over which both loci are polymorphic. This last result is in agreement with Axelrod et al.'s and Jansen and van Baalen's simulation results, but still holds under restrictive conditions. Evidence for extrinsic balancing selection at the matching locus is reviewed.

Biological Scenario

The population follows the infinite island model of dispersal, with N haploid adults settled in each deme. We assume that there are interactions between all pairs of individuals in each deme. In each such pair each individual helps its partner if it bears a helping allele (say allele H_1) at a helping locus and if the two individuals bear the same allele at some "matching" locus. The interaction is reciprocal in that each individual in a pair takes the actor's role. It is not reciprocal in the sense that one individual will not reciprocate the fecundity gift it may have received when it does not bear the helping allele. Thus the alternative allele to the helping allele (say allele H_0) may be seen as a "cheating" allele. In our general model (presented in the online Supplementary Appendix S1), the number of alleles segregating in the population at the matching locus can be arbitrary. Here, for simplicity, we will introduce the model and results for the case in which only two alleles (say R_0 and R_1) segregate at this locus; however, we will see that some results extend readily to the case of an arbitrary number of alleles.

Let $X_{\mathrm{R}(ij)}$ be an indicator variable that equals 1 if individual j in deme i bears allele R_1 at the matching locus, and 0 if it bears allele R_0. Similarly, let $X_{\mathrm{H}(ij)}$ be an indicator variable that equals 1 if individual j in deme i bears allele H_1 at the helping locus, and 0 if it bears allele H_0. We assume that helping increases the fecundity of the partner by an amount $B/(N-1)$, but decreases the fecundity of the actor by an amount $C/(N-1)$. Therefore, interacting with individual k in deme i increases (or decreases) the fecundity of individual j in deme i by an amount $(-CX_{\mathrm{H}(ij)} + BX_{\mathrm{H}(ik)})/(N-1)$, if both individuals bear the same allele at the matching locus, which is the case if the quantity $X_{\mathrm{R}(ij)}X_{\mathrm{R}(ik)} + (1-X_{\mathrm{R}(ij)})(1-X_{\mathrm{R}(ik)})$ equals 1 (otherwise, this quantity equals 0). Because we assume that the effects of different interactions are additive, B and C thus measure the maximal benefit and cost in fecundity to an individual, after having interacted with all other individuals from its deme. Taking all possible interactions into account, the fecundity of individual j in deme i (relative to the fecundity of an individual that did not give or receive helping) is given by

$$1 + \tfrac{1}{N-1}\sum_{k\neq j}\Big[X_{\mathrm{R}(ij)}X_{\mathrm{R}(ik)} + \big(1-X_{\mathrm{R}(ij)}\big)\big(1-X_{\mathrm{R}(ik)}\big)\Big]\Big(-CX_{\mathrm{H}(ij)} + BX_{\mathrm{H}(ik)}\Big). \quad (1)$$

We assume a haploid life cycle in which individuals reproduce asexually according to such fecundities, gamete dispersal occurs, and is followed by syngamy and immediate meiosis during which recombination occurs with probability r between the two loci considered. Population regulation then brings back the number of settled haploid individuals to N.

In the following, we first present weak-selection approximations for the change in allele frequencies at the helping and matching loci, where the derivation is detailed in the Online Appendix. These approximations are first-order results in B and C, assuming B and C are weak relative to r and m. There is no constraint on deme size relative to the other parameters.

The approximations show that conditions for the evolution of helping are less stringent when helping can be directed toward kin, using polymorphism at the matching locus to discriminate kin from nonkin. However, the results also show that when helping does evolve, it eliminates polymorphism at the matching locus.

Thus, helping conditional on kin recognition can be maintained only if another source of selection stabilizes polymorphism at the matching locus. We then consider stronger forms of selection, by deriving second-order results. This shows that in some cases, the population evolves toward an equilibrium where both the helping and matching locus are polymorphic. However, this only occurs under rather restrictive conditions. Finally, we consider the stabilizing effect of mutation on polymorphism at the two loci.

WEAK SELECTION RESULTS

To the first order in B and C, the change in allele frequency at the helping locus is of the form

$$\Delta p_{\mathrm{H}} = p_{\mathrm{H}} q_{\mathrm{H}} \Big(-C\,[1 - F - 2(1-2F) p_{\mathrm{R}} q_{\mathrm{R}}] + 2B\phi\, p_{\mathrm{R}} q_{\mathrm{R}} - 2(B-C)(1-m)^2 \tfrac{F+\phi+(N-2)\gamma}{N} p_{\mathrm{R}} q_{\mathrm{R}} \Big) \quad (2)$$

where p_{H} and q_{H}(p_{R} and q_{R}) are the frequencies of alleles H_1 and H_0 (R_1 and R_0) in the whole population, m is the immigration rate, and F, ϕ, and γ measure various probabilities of identity by descent, computed under neutrality (see Fig. 1). These identity probabilities are readily obtained by solving recurrence equations (see online Supplementary Appendix S1); F is a function of N and m, whereas ϕ and γ are functions of N, m, and r.

To understand this result, first note that in the absence of polymorphism at the matching locus, the whole expression reduces to

$$\Delta p_{\mathrm{H}} = -C(1-F) p_{\mathrm{H}} q_{\mathrm{H}} \quad (3)$$

for any N, showing that helping is favored if and only if a focal individual's helping allele increases its fecundity. This result is consistent with the previous analyses of the same life cycle with dispersal, population regulation, and reproduction in this order (e.g., Taylor 1992; see Taylor and Irwin 2000; Roze and Rousset 2003; Lehmann et al. 2006; Lehmann 2007 for results under other life cycles).

The $2B\phi p_{\mathrm{R}} q_{\mathrm{R}}$ term arises because individuals who are identical at the matching locus also tend to share the same allele at the helping locus; therefore, polymorphism at the matching locus increases the proportion of helpers among the recipients of helping acts. However, these benefits are partially lost through kin-competition effects (last term of eq. 2, which vanishes when $m = 1$). The complete expression can be positive only if $B > C$, as intuition would suggest and as can be verified by setting $B = C$ and using the fact that $\phi \leq F$ and $\gamma \leq F$.

The change in allele frequency is maximal when the polymorphism (as measured by $p_{\mathrm{R}} q_{\mathrm{R}}$) is maximum at the matching locus (i.e., 1/4 in the two-alleles model), and intuition suggests that identity at the matching locus provides "better information" about identity at the helping locus when several recognition alleles segregate. For multiple allele at frequencies p_k at the matching locus,

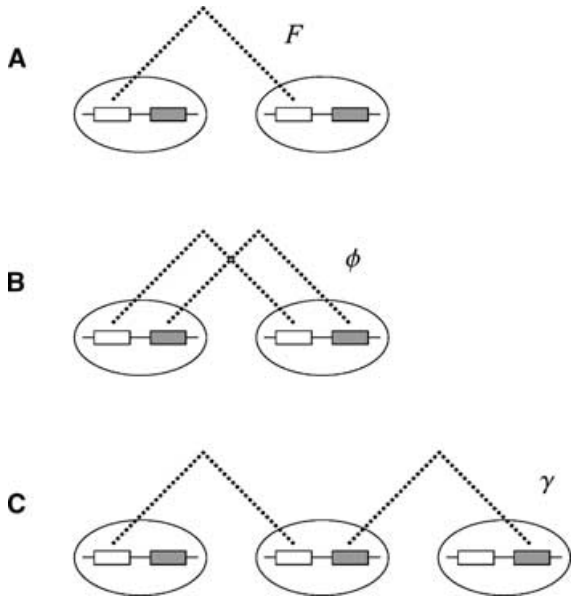

Figure 1. Association between two loci in a deme. Each individual is represented by an oval box with two loci (empty and gray rectangles). Dotted lines connect gene copies that probability of joint coalescence within the deme ("identity by descent") is measured by the coefficients F, ϕ, and γ. F measures the probability that the ancestral lineages of two genes sampled at the same locus from two different individuals from the same deme stay in the same deme and coalesce (in a neutral model), whereas ϕ (γ) is the probability that two pairs of genes sampled at two different loci from two (three) different individuals from the same deme stay in the same deme and coalesce (still in the neutral case).

selection on the helping allele actually depends on $\sum_i \sum_{j>i} p_i p_j$ in place of $p_{\mathrm{R}}\, q_{\mathrm{R}}$ in equation (2). The conclusion that selection can be positive only if $B > C$ is therefore unaffected (see online Supplementary Appendix S1 for further details about multiallelic models).

One can also check that Δp_{H} has the sign of $-C$ when $\phi = \gamma = F^2$, that is when the two-locus probabilities of identity are equal to the product of independent one-locus probabilities of identity. Helping alleles are favored when ϕ is increased above this level. The quantity $\phi - F^2$ is known as a measure of identity disequilibrium (e.g., Vitalis and Couvet 2001). Positive identity disequilibrium occurs because population structure generates a correlation between probabilities of coalescence at different loci: if the ancestral lineages of two genes sampled at the same locus from two individuals from the same deme stay in the same deme and coalesce (rather than migrating to different demes), then the probability of coalescence at any other locus from the same two individuals is increased. There is no gametic ("linkage") disequilibrium at the total population level: the altruistic allele does not need to be associated to any given recognition allele on average in the total population. Rather, what matters is that two individuals

from the same deme that share the same allele at one locus also tend to share the same allele at the other locus. Identity disequilibrium can thus be seen as a measure of local gametic disequilibrium. γ measures a similar phenomenon but between three random individuals within a deme: individuals "1" and "3" are more likely to share the same allele at a locus when individuals "1" and "2" share an allele at the other locus, again as a consequence of correlations in coalescence events within demes. This genetic association appears in the term measuring fitness effects of local competition because the fitness of an individual is affected by pairwise interactions between its neighbors in the deme (as highlighted by the derivation of eq. 2 in the online Supplementary Appendix S1).

With free dispersal, some degree of identity disequilibrium may occur in finite populations, but will remain rather small (of order $1/N_T$, where N_T is the total population size, i.e., deme size times number of demes). Spatial structure is thus important because it is efficient in generating identity disequilibrium. On the other hand, local competition for resources decreases the strength of selection for conditional helping. Overall, equation (2) indicates that selection for costly conditional helping is stronger when deme size is small, dispersal is low, and recombination is low. Nevertheless, the process operates even with free recombination ($r = 1/2$) between the two loci. Exactly the same mechanisms would operate in family-structured populations (e.g., sib-recognition in social insect colonies without localized dispersal).

Whether costly helping is selected for obviously depends on the efficiency of kin recognition, which should be greater when the C/B ratio is higher. When based on several matching loci, kin recognition may practically restrict positive interactions to members of a single clone or, in fungi, to mycelial filaments from the same individual. Nevertheless, such self-recognition is not fundamentally distinct from kin recognition.

In the conditions in which conditional costly helping may be selected for, polymorphism is not maintained at the matching locus. This problem was first pointed out by Crozier (1986). In the present model, the change in frequency is

$$\Delta p_R = p_H p_R q_R (2p_R - 1)(B - C) Z \qquad (4)$$

where Z is a function of N and m, which is given in the online Supplementary Appendix S1 and which is always positive. The $B - C$ factor in equation (4) is expected: helping occurs only when recognition alleles are identical, so that one copy of any given recognition allele pays the cost if and only if another copy of the same allele receives the benefits. Note that this is true whatever the magnitude of B and C—not only for weak selection. Equation (4) shows that Δp_R has the sign of $(2p_R - 1)(B - C)$: therefore, the polymorphism is stable if and only if $C > B$, a conclusion that holds for any number of recognition alleles (see online Supplementary Appendix S1).

However, we have seen previously that helping is not selected when $C > B$. Thus, recognition-based helping could evolve only if the polymorphism is maintained by some extraneous force. Various forms of balancing selection could be considered, and they could favor either locally (at the level of the deme) or globally (at the level of the total population) rare alleles. The latter case can be described phenomenologically by an additional factor $1 + D(1 - 2X_{R(ij)})(1/2 - p_R)$ affecting the fecundity of individual j in deme i, where the parameter D measures the strength of balancing selection. In that case, one obtains that an additional term $(1 - F)Dp_R q_R(1 - 2p_R)$ appears in the expression for change at the matching locus, whereas the change in helping allele frequency is unaffected. Hence, with D large enough relative to $B - C$, the polymorphism at the matching locus can be stable and helping can be selected. The most efficient forms of helping (large $B - C$) would evolve only if conditional on recognition systems with high intrinsic "cost" D, the minimal D being approximately $(B - C)p_H Nm/(1 + Nm)$, so it decreases with decreasing number of immigrants Nm. Helping with stronger effect may invade but will eventually destabilize the polymorphism when the helping allele approaches fixation. If extrinsic balancing selection favors locally, rather than globally, rare alleles, the minimal D is $(B - C)p_H N/(N - 1)$. These results hold unchanged in the multiallelic case (see online Supplementary Appendix S1).

STRONGER SELECTION

When strong fecundity effects are considered, the above approximations may fail. We have investigated such complications and in particular we have sought cases in which extrinsic forces are no longer needed to maintain the polymorphism at the matching locus. Although the second-order expressions for selection differentials are too complicated to be reproduced here, they allow an efficient numerical exploration of the parameter space. Importantly, our method (which is outlined in the online Supplementary Appendix S1) assumes that selection is weaker than migration and recombination, so that genetic associations between genes at different loci and/or in different individuals equilibrate fast relative to changes in allele frequencies (e.g., Kirkpatrick et al. 2002; Roze and Rousset 2005). This allows us to express genetic associations in terms of allele frequencies and of the different parameters of the model, and to describe the evolution of the system by recursions on allele frequencies only (we thus have two recursions when the number of alleles K at the matching locus equals two).

As a result of the selective interactions, rare recognition alleles become preferentially associated with the helping allele. This association, which was negligible in the first-order approximation, creates a positive frequency-dependent selective force at the matching locus, which may overcome the negative frequency-dependent effect previously described. Then, rare recognition alleles may be favored, while at the same time the helping

behavior is also selected. However, partly because this association decreases when the recognition allele frequency approaches 0.5, partly because of changes in other associations (see online Supplementary Appendix S1), the recognition polymorphism is eventually destabilized. As this polymorphism decreases, selection for helping decreases, and may eventually become negative. This may bring back the population to an initial state of low helping frequency and low recognition polymorphism. Thus, stable equilibria or limit cycles might occur.

Both dispersal and recombination prevent the occurrence of stable polymorphisms by reducing the two-locus associations generated by selection. Indeed, we found that a joint stable polymorphism is possible only under low recombination and strong population structure. The upper left plot of Figure 2 shows the range of values of B and C where stable polymorphic equilibria were predicted when $N = 6$, $m = 0.1$, and $r = 0.05$ (strong spatial structure and strong linkage), and with two recognition alleles. It shows that B has to be greater than a minimum value, but less than a maximal value, above which the equilibrium becomes unstable. For $m = 0.1$, no internal equilibrium was predicted for $r > 0.2$ except for $N = 2$, and for $m = 0.25$ no internal equilibrium was predicted for $N > 2$ and any r.

Because the two recognition alleles are functionally equivalent, the evolutionary dynamics is completely symmetric around the $p_R = 0.5$ axis, and only the dynamics for $p_R < 0.5$ will be shown in Figures. Stable equilibria were not found on the $p_R = 0.5$ axis (i.e., equal frequencies at the recognition allele are never stable), nor on the $p_H = 1$ edge (i.e., recognition polymorphism is never stable when helping is fixed in the population). Even when a stable polymorphic equilibrium does exist, its basin of attraction may be rather small (see Fig. 2). Further, the deterministic paths toward these points are slowly converging spirals (Figs. 2 and 3). As simulation shows, genetic drift then sends the population on random orbits around such equilibrium points (Fig. 3). Therefore,

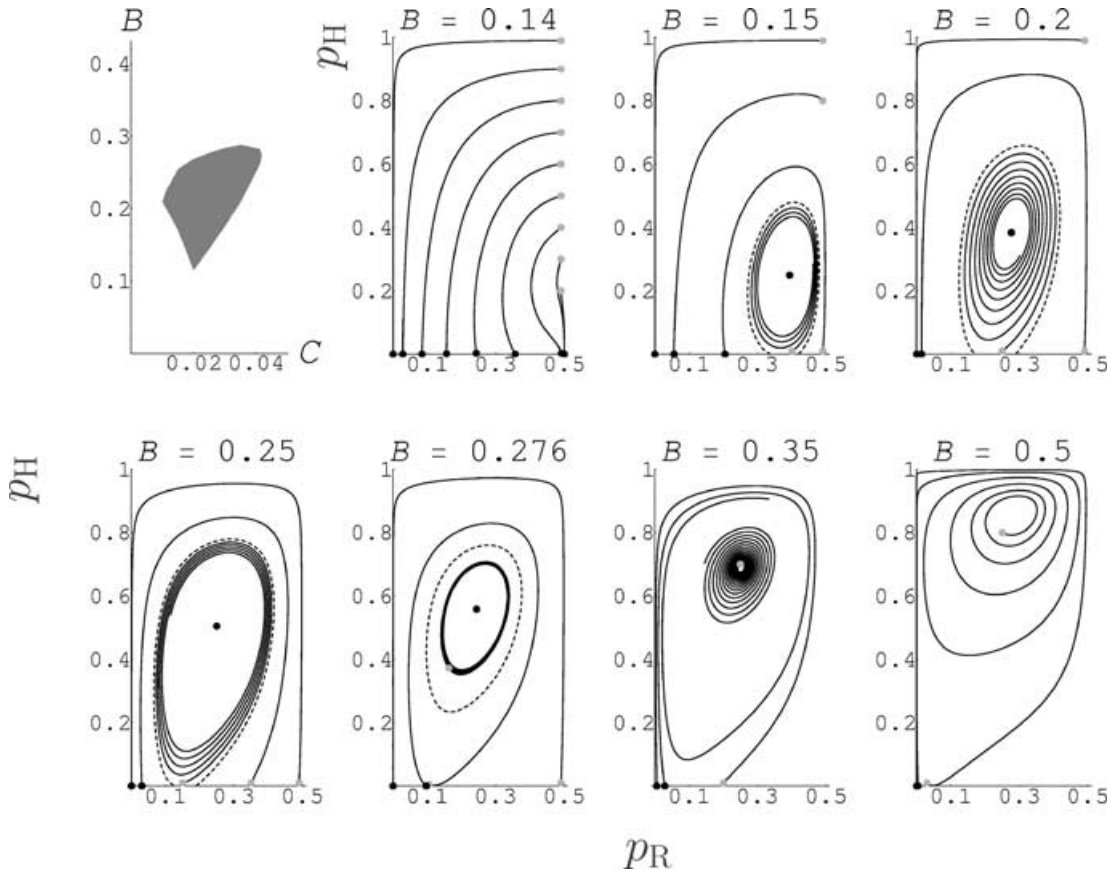

Figure 2. Predicted dynamics under strong structure. These plots are drawn for $N = 6$, $r = 0.05$, and $m = 0.1$. Stable jointly polymorphic equilibria are predicted for B and C values in the range shown on the upper left plot. The stability of these internal equilibrium points was determined by local stability analysis, using the second-order analytical approximations. The other plots show the predicted dynamics for $C = 0.025$ and variable B. The completely symmetric results for $p_R > 0.5$ are not shown. Basins of attraction are delimited by dashed lines. Initial points of shown orbits are marked as gray dots. Stable polymorphic equilibria and endpoints of orbits resulting in loss of polymorphism are marked as black dots. The plot for $B = 0.276$ shows the shrinking of the basin of attraction as B approaches the upper threshold $B = 0.2768$ for stable polymorphic equilibrium.

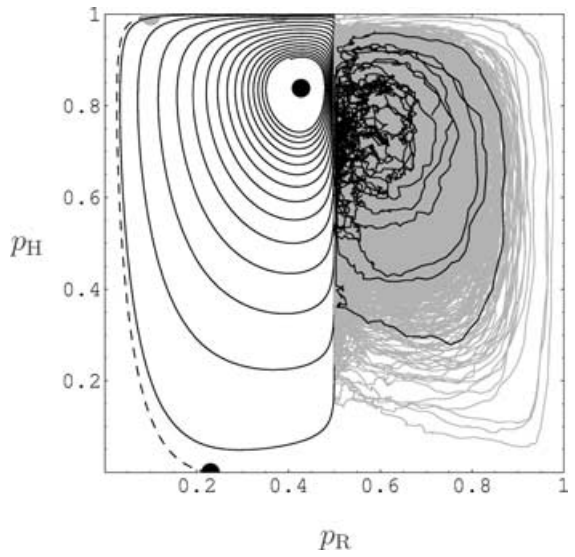

Figure 3. Predicted and observed dynamics under strong structure. Simulation parameters were $B = 0.205$, $C = 0.05$, $r = 0.05$, $m = 0.1$, and $N = 2$. The second-order prediction (shown for $p_R < 0.5$; the right-hand side would be symmetrical) is compared to the observed dynamics from a simulation of a population of 20,000 demes (shown for $p_R > 0.5$; a few paths are enhanced for better visibility). The predicted paths and basin of attraction are shown as in Figure 2. In the simulation, genetic drift sends the population on random orbits around a putative stable equilibrium point. Only one fixation event occurred in 20×10^6 generations. Estimation of phase portraits from additional simulations showed that the putative equilibrium is near $(p_R, p_H) = (0.6, 0.72)$.

a stable polymorphism seems possible only under a restricted set of parameters and initial conditions. In most biological settings, nonhelper mutants at additional unlinked ($r = 0.5$) loci would be expected to destabilize any preexisting polymorphism not maintained by extrinsic forces.

Although we obtained second-order recursions for an arbitrary number K of recognition alleles, the dynamics are difficult to analyze when K is greater than 2 or 3, due to the high number of dimensions involved. With three alleles, and for the parameter values used in Figure 2 ($N = 6$, $m = 0.1$, $r = 0.05$, $C = 0.025$), we found that internal equilibria where all three recognition alleles are present in the population may exist for some values of B, but these equilibria are never stable, and no stable limit cycle is observed. Accordingly, stable polymorphisms (as measured by comparison of fixation times at the recognition locus to those for a neutral model) were not observed in simulations. The same conclusion holds for simulations conducted in the same conditions but with more recognition alleles. Finally, we will see in the following that K has generally little effect on the average frequency of helping under recurrent mutation. These results thus suggest that increasing the number of recognition alleles does not help to maintain polymorphism and conditional helping.

Although selection alone appears generally insufficient to maintain polymorphism at the recognition locus, a low mutation rate may allow the population to reach a limit cycle, during which the frequency of helping and the polymorphism may periodically reach high values. This is easily seen in Figure 2, for the case $B = 0.5$: all trajectories converge to the origin (black dot at $p_H = 0$, $p_R = 0$), but if the population can move slightly away from the origin (from the black dot to the gray dot close to the origin), helping and polymorphism increase, reach high values, and decrease again to zero. This is confirmed by adding mutation to our second-order recursions on allele frequencies: for example, Figure 4 shows limit cycles reached under $B = 0.35$ and the same other parameter values as in Figure 2, but with two-way mutation at both loci. When the mutation rate is $u = 10^{-5}$ at both loci, the population reaches the large limit cycle, when $u = 10^{-4}$ (as in Jansen and van Baalen 2006) it reaches the small cycle, whereas when $u = 10^{-3}$ (as in Axelrod et al. 2004) it reaches a stable equilibrium, shown by the black dot.

Finally, we used simulations to investigate the level of polymorphism and of helping maintained in the population under recurrent mutation, assuming mutation rates $u = 10^{-5}$ at both loci. Parameters values were $N = 6$, $r = 0.05$, $m = 0.1$, $K = 5$, $C = 0.025$ unless specified otherwise, B was varied from 0.1 to 0.5 by steps of 0.05, and either r, m, N, or K were varied. Results are shown in Figure 5. A general trend is that levels of helping

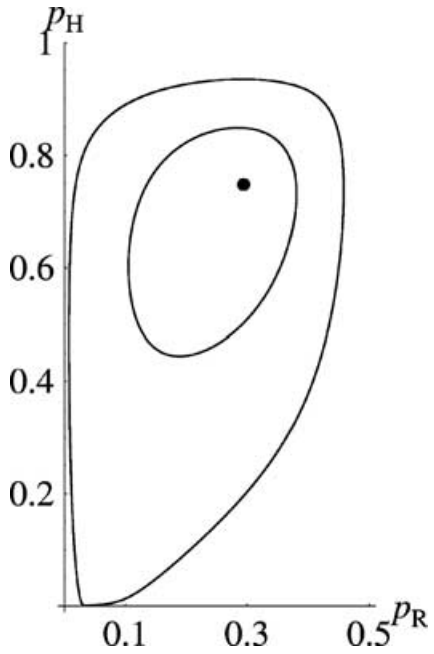

Figure 4. Effect of mutation on predicted dynamics. This figure shows predicted deterministic dynamics for $u = 10^{-5}$ (outer stable cycle), $u = 10^{-4}$ (inner stable cycle), and $u = 10^{-3}$ (single stable equilibrium point), when $C = 0.025$ $N = 6$, $r = 0.05$, $m = 0.1$ (as in Fig. 2), and $B = 0.35$.

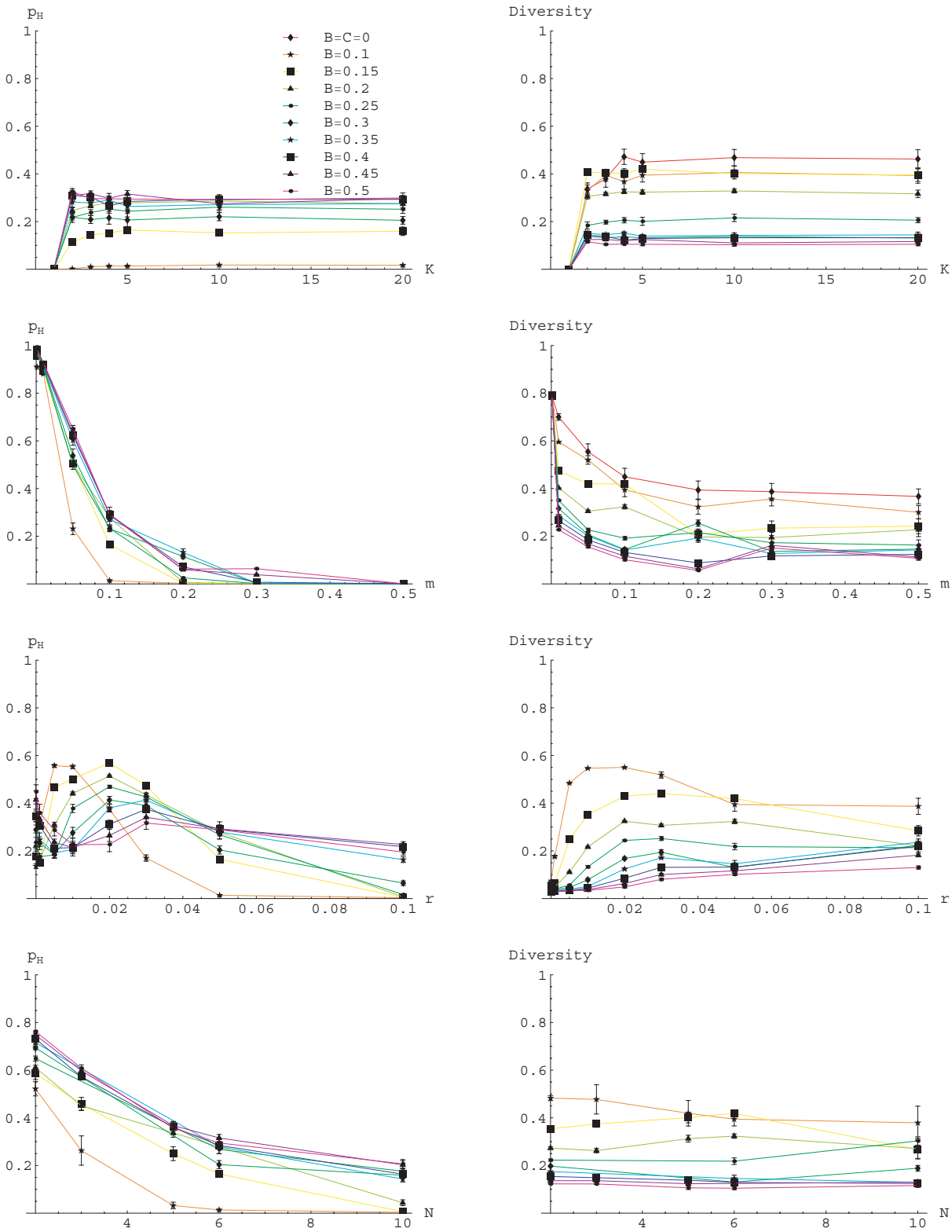

Figure 5. Frequency of helping and gene diversity at the matching locus under recurrent mutation. Simulation parameters were, $C = 0.025$, $r = 0.05$, $m = 0.1$, $N = 6$, $u = 10^{-5}$, and $K = 5$ unless specified. Note that neutral expectations (in red) are shown only in the two upper right-hand plots. The left-hand column shows estimates of the helping allele frequency and the right-hand column the gene diversity ($1 - \sum_{k=1}^{K} p_k^2$) at the matching locus. For each simulation conditions, simulations lasted at least 10^6 generations, and confidence intervals were constructed by the batching method for Markov chains (Hastings 1970). Rows show successively the effects of varied K, m, r, and N. The number of demes n was such that total population size was kept constant ($Nn = 30,000$).

increase when spatial genetic structure increases (lower m, lower N), which is expected. The effect of recombination is more complex, with often a maximum level of helping at $r = 0$ and another maximum at some intermediate r. Noticeably, increasing K above 2 has only a slight effect on the average frequency of helping. Only when population structure is strong does the frequency of helping reach high values; otherwise, it usually remains lower than the expectation for a neutral allele in a mutation-drift model with the same parameters (here 0.5). Moderately high p_H frequencies (≈ 0.55) are also observed for low recombination ($r = 0.01$ or 0.02) and $B < 0.25$. The second-order model predicts stable equilibria at even higher p_H values in these cases, except for $B = 0.1$. Polymorphism at the recognition locus is always lowest in the low-recombination limit; indeed, under low recombination, the helping allele remains associated with a single recognition allele when it increases in frequency, and diversity at the matching locus is eroded by these recurrent hitchhiking events. This may partly explain the nonmonotonous effect of recombination on helping frequency: the evolution of conditional helping needs an association between the two loci, but when this association is too strong, diversity is greatly reduced by hitchhiking at the matching locus, which in turn affects selection on helping. This effect was already noted by Jansen and van Baalen (2006).

Discussion

Different population structures allow the evolution of more or less costly helping; kin recognition allows for the evolution of more costly helping (Hamilton 1987). Previous theoretical studies of this process have been based on simulation or on models that did not account for genetic associations between loci among different individuals. We have analyzed a two-locus model to delineate conditions for the evolution of conditional costly helping. According to the first-order approximation, strongly costly altruism requires independent strongly balanced polymorphism at the matching locus for its evolution. According to the second-order approximation, some exceptions are found to this conclusion, where complex patterns of frequency dependence can be observed. In particular, jointly balanced polymorphisms at both loci may be maintained by selection, but only when there is both extreme population structure and low recombination. This result relies in part not only on the building of positive gametic disequilibria between rare recognition alleles and the helping allele (Grafen 1990; Axelrod et al. 2004) but also on the effect of selection on genetic associations between genes in different individuals within a deme. When recurrent mutation is introduced, the population reaches a stable limit cycle, over which both loci are polymorphic, over a wider region of the parameter space. Still, population structure has to be strong, and linkage sufficiently tight, so that the average frequency of the helping allele can reach high values. These results are qualitatively consistent with the simulation results of Axelrod et al. (2004). In particular, Axelrod et al.'s model involved spatial structure (with isolation by distance), complete linkage (asexual reproduction), and high mutation rates (10^{-3} per loci), which allows conditional helping to be maintained at high frequencies. Apart from the case of strong structure and tight linkage, we find that the first-order results are a good guide to the evolution of helping conditional on kin recognition, and in all cases helping is more likely to evolve when there is independent balancing selection at the matching locus.

Our simulation results, including the maximum level of helping for an intermediate level of recombination, are also qualitatively consistent with those of Jansen and van Baalen (2006). Although their model was described as a "green beard" model rather than as kin recognition, it describes the evolution of helping between individuals that share any allele at a recognition locus rather than between individuals that share a specific allele at the recognition locus. It is thus a model of kin recognition rather than of the green beard process as usually understood (e.g., Dawkins 1982, chapter 8). The two processes may be difficult to distinguish when recombination vanishes, because each recognition allele may then become associated to a distinct lineage of helping alleles.

In the simulations, intermediate level of helping and of recognition polymorphism are maintained at mutation–selection balance. In most cases the net effect of selection is to reduce polymorphism below the neutral mutation-drift expectation, and the average frequency of helping is low (< 0.5). This is a much less efficient process for the evolution of conditional helping than one depending on extrinsic balancing selection, because in the latter case the helping allele can go to fixation. The main exceptions were found for low dispersal and low recombination ($m < 0.1$ and $r < 0.1$, Fig. 5). Less-stringent conditions on r were found for high values of B and moderately large values of N. In particular, high frequencies of helping can be observed for $B = 2$ or 5, $N = 10B$ to $20B$, and $r = 0.1$ to 0.2 when $m < 0.05$. In many of these cases, high frequencies of helping occur despite selection being purifying, as evidenced by fixation times (simulations not shown).

In light of our results, it is worth reconsidering evidence in favor of extrinsic factors in the few cases in which a genetic basis of kin recognition is known. Choice of MHC-similar social partners has been reported in mice (Manning et al. 1992) and may operate also in humans (Jacob et al. 2002). The MHC is a classic example of balanced polymorphism. Although observations of social behaviors conditional on MHC identity support the idea that polymorphisms balanced by extrinsic selection are preferential targets for kin recognition, focusing on this polymorphism may prevent the detection of other polymorphisms that might also be involved but not be maintained by extrinsic selection. Asking

whether loci under extrinsic balancing selection, such as the MHC, are involved in kin recognition is thus distinct from asking whether loci involved in kin recognition are under extrinsic balancing selection. There are two cases in which there is information specifically on the latter issue, and they both support the above conclusions. Extrinsic selection has been demonstrated recently at the fusion/histocompatibility locus, which controls fusion between colonies of the ascidian *Botryllus*. In this classic case of kin recognition, it has been found that heterozygotes have a fitness advantage over homozygous siblings in laboratory crosses (de Tomaso and Weissman 2004), although the mechanism is unknown. Previous suggestions that this locus controls gamete compatibility have not been confirmed (Grosberg and Hart 2000; de Tomaso and Weissman 2004). The other, almost forgotten example is from the cellular slime mold *Dictyostelium discoideum*, where multicellular reproductive stages (fruiting bodies) are formed by the aggregation of numerous individual amoebae. Although assortative kin recognition is not always reported to control aggregation in *D. discoideum* (Strassmann et al. 2000), it was once observed and then found to be closely linked if not identical to the mating type locus (Robson and Williams 1979), which by its function is under strong extrinsic balancing selection. Kin recognition occurs in other *Dictyostelium* species (Bonner and Adams 1958; Kaushik et al. 2006; Mehdiabadi et al. 2006) but its basis is unknown.

Other observations in animals are less detailed. Genetic kin recognition is assumed to occur in social insects, but its chemical and physiological basis is insufficiently known although the involvement of cuticular hydrocarbons has long been suspected (e.g., Boomsma et al. 2003; Tsutsui 2004; Howard and Blomquist 2005). In a parasitoid wasp, kin recognition is based on cues derived from an extraembryonic membrane involved in protection against host defenses, suggesting that host–parasite interactions maintain the polymorphism (Giron and Strand 2004). A similar suggestion has been made for cuticular hydrocarbons (Ratnieks et al. 2007). Nevertheless, there is relatively little evidence for genetic kin recognition in cases in which nepotism could pay, which can in part be explained by the constraints on the maintenance of a matching polymorphism (Ratnieks et al. 2006).

In *Botryllus*, colony fusion allows invasion of the gonads by the other colony's reproductive cells (Sabbadin and Zaniolo 1979; Laird et al. 2005). It has been suggested that compatibility systems, including vertebrate adaptive immunity, have originally evolved to prevent invasion of the germ line (or stem cells) by selfish mutants (e.g., Buss 1982; de Tomaso 2006). According to the present results, this scenario can only work in very restricted conditions. Our model describes conditional resistance or susceptibility to colony fusion in the case of identity at the matching locus, where fusion results in some total benefit ($B - C$) for the pair of colonies, but may be costly for a colony whose germ line is invaded. Thus, conditional rather than unconditional resistance can be selected only if $B - C > 0$ and again, to first order, no polymorphism is maintained at the matching locus by this mechanism.

Vegetative incompatibility loci in fungi could also be involved in kin recognition. Here matching-based kin recognition is known to control whether cytoplasmic fusion between different mycelia (plasmogamy) is stable or turns necrotic. This vegetative compatibility is distinguished from sexual compatibility, which controls nucleus migration or meiosis (for reviews see Worrall 1997; Saupe 2000; Brown and Casselton 2001); vegetative incompatibility may or may not prevent mating. The significance of vegetative incompatibility in natural populations remains poorly understood. Here too, resistance to germ line parasitism has been considered as an explanation for the evolution of compatibility loci. However, consistent with our results, earlier models assuming no spatial structure found that the polymorphism of vegetative compatibility loci could not be explained if kin recognition served as a means to avoid germ line parasitism (Nauta and Hoekstra 1994).

To explain the maintenance of the polymorphism at vegetative incompatibility loci, costs of fusion between individuals identical at all loci have been considered (Muirhead et al. 2002), but it would then be unclear why vegetative fusion persist between such individuals. Indeed, in several cases in which kin recognition exists, the costs and benefits are not well known. In our model, there are cases in which there is selection for an allele with effects $B < 0$ and $C < 0$, and where the recognition polymorphism is stable. However, strategies with such fitness effects are not stable in that the actor exploits the recipient ($B < 0$) only when the recipient matches at the matching locus. The unconditional behavior with the same fecundity effects would be favored, so this case should not be taken as an explanation for kin-conditional behavior (Grosberg and Quinn 1989). Rather, fusion might be beneficial on average (over the two partner mycelia) and the polymorphism may be maintained by extrinsic forces. Although the causes of balancing selection at vegetative incompatibility loci are generally unknown, one of the 11 known vegetative incompatibility loci in *Neurospora crassa* is the mating-type locus (reviewed in Xiang and Glass 2004), again a locus subject to extrinsic balancing selection.

Our results are thus consistent with current, although fragmentary knowledge of the genetic basis of assortative kin recognition in diverse organisms. They are congruent with the insight from some of the earliest models, according to which extrinsic selection is required to maintain polymorphism at other loci. Exceptions could be expected only under extreme spatial structure and low recombination. These exceptional cases cannot be completely excluded, in particular in poorly known systems such as some soil microorganisms and fungi. The hypothesis that defense against germ line parasites is the original function of

polymorphism at the MHC and similar compatibility systems could work only under the same conditions.

The conclusion that the first-order analysis is sufficient over most of the parameter space could be expected, and it should be independent of the details of the life cycle assumed in this work, as multilocus selective forces generally arise from associations between loci that are canceled by recombination, and thus one-locus forces tend to dominate except in condition of low recombination (e.g., Nagylaki 1993; Kirkpatrick et al. 2002). In spatially structured populations, a similar result holds, except that the decay of multilocus associations also depends on dispersal rates (see also Roze and Rousset 2005; Lehmann et al. 2007; Martin et al. 2006). On the other hand, some multilocus associations, such as identity disequilibria, are created by spatial structure in the absence of selection, and these should be taken into account to generate accurate first-order predictions.

ACKNOWLEDGMENTS

We thank S. Gandon, L. Lehmann, M.-A. Selosse, and S. West for discussions, and A. Grafen for very constructive comments on the ms. This is publication ISEM 07-067. Part of the computations for this work were performed on the ISEM cluster.

LITERATURE CITED

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.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229–255.

Bonner, J. T., and M. S. Adams. 1958. Cell mixtures of different species and strains of cellular slime moulds. J. Embryol. Exp. Morphol. 6:346–356.

Boomsma, J. J., J. Nielsen, L. Sundström, N. J. Oldham, J. Tentschert, H. C. Petersen, and E. D. Morgan. 2003. Informational constraints on optimal sex allocation in ants. Proc. Natl. Acad. Sci. U. S. A. 100:8799–8804.

Brown, A. J., and L. A. Casselton. 2001. Mating in mushrooms: increasing the chances but prolonging the affair. Trends Genet. 17:393–400.

Buss, L. W. 1982. Somatic cell parasitism and the evolution of somatic tissue compatibility. Proc. Natl. Acad. Sci. U.S.A. 79:5337–5341.

Crampton, W. G. R., and L. D. Hurst. 1994. True kin recognition, in the form of somatic incompatibility, has multiple independent origins. Anim. Behav. 47:230–234.

Crozier, R. H. 1986. Genetic clonal recognition abilities in marine invertebrates must be maintained by selection for something else. Evolution 40:1100–1101.

———. 1987. Genetic aspects of kin recognition: concepts, models, and synthesis. Pp. 55–73 *in* D. J. C. Fletcher and C. D. Michener, eds. Kin recognition in animals. Wiley, New York.

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

de Tomaso, A. W. 2006. Allorecognition polymorphism versus parasitic stem cells. Trends Genet. 22:485–490.

de Tomaso, A. W., and I. L. Weissman. 2004. Evolution of a protochordate allorecognition locus. Science 303:977.

Giron, D., and M. R. Strand. 2004. Host resistance and the evolution of kin recognition in polyembryonic wasps. Proc. Roy. Soc. (Lond.) B 271:S395–S398.

Grafen, A. 1990. Do animals really recognize kin? Anim. Behav. 39:42–54.

Grosberg, R. K., and M. W. Hart. 2000. Mate selection and the evolution of highly polymorphic self/nonself recognition genes. Science 289:2111–2114.

Grosberg, R. K., and J. F. Quinn. 1989. The evolution of selective aggression conditioned on allorecognition specificity. Evolution 43:504–515.

Hamilton, W. D. 1987. Discriminating nepotism: expectable, common, overlooked. Pp. 417–437 *in* D. Fletcher and C. D. Michener, eds. Kin recognition in animals. Wiley, New York.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Howard, R. W., and G. J. Blomquist. 2005. Ecological, behavioral, and biochemical aspects of insect hydrocarbons. Ann. Rev. Entomol. 50:371–393.

Jacob, S., M. K. McClintock, B. Zelano, and C. Ober. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. Nat. Genet. 30:175–179.

Jansen, V. A. A., and M. van Baalen. 2006. Altruism through beard chromodynamics. Nature 440:663–666.

Kaushik, S., B. Katich, and V. Nanjundiah. 2006. Social behaviour in genetically heterogeneous groups of *Dictyostelium giganteum*. Behav. Ecol. Sociobiol. 59:521–530.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Laird, D. J., A. W. de Tomaso, and I. L. Weissman. 2005. Stem cells are units of natural selection in a colonial ascidian. Cell 123:1351–1360.

Lehmann, L. 2007. The evolution of trans-generational altruism: kin selection meets niche construction. J. Evol. Biol. 20:181–189.

Lehmann, L., N. Perrin, and F. Rousset. 2006. Demographic sensitivity and the evolution of helping. Evolution 60:1137–1151.

Lehmann, L., F. Rousset, D. Roze, and L. Keller. 2007. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. Am. Nat.170:21–36.

Manning, C. J., E. K. Wakeland, and W. K. Potts. 1992. Communal nesting patterns in mice implicate MHC genes in kin recognition. Nature 360:581–583.

Martin, G., S. P. Otto, and T. Lenormand. 2006. Selection for recombination in structured populations. Genetics 172:593–609.

Mehdiabadi, N. J., C. N. Jack, T. T. Farnham, T. G. Platt, S. E. Kalla, G. Shaulsky, D. C. Queller, and J. E. Strassmann. 2006. Kin preference in a social microbe. Nature 442:881–882.

Muirhead, C. A., N. L. Glass, and M. Slatkin. 2002. Multilocus self-recognition systems in fungi as a cause of trans-species polymorphism. Genetics 161:633–641.

Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. Genetics 134:627–647.

Nauta, M. J., and R. F. Hoekstra. 1994. Evolution of vegetative incompatibility in filamentous ascomycetes. I. Deterministic models. Evolution 48:979–995.

Ratnieks, F. L., K. R. Foster, and T. Wenseleers. 2006. Conflict resolution in insect societies. Ann. Rev. Entomol. 51:581–608.

Ratnieks, F. L. W. 1991. The evolution of genetic odor-cue diversity in social Hymenoptera. Am. Nat. 137:202–226.

Ratnieks, F. L. W., H. Helanterä, and K. R. Foster. 2007. Are mistakes inevitable? Sex allocation specialization by workers can reduce the genetic information needed to assess queen mating frequency. J. Theor. Biol. 244:470–477.

Robson, G. E., and K. L. Williams. 1979. Vegetative incompatibilty and the mating-type locus in the cellular slime mold *Dictyostelium discoideum*. Genetics 93:861–875.

Roze, D., and F. Rousset. 2003. Diffusion approximations for selection and drift in subdivided populations: a straightforward method and examples

involving dominance, selfing and local extinctions. Genetics 165:2153–2166.

———. 2005. Inbreeding depression and the evolution of dispersal rates: a multilocus model. Am. Nat. 166:708–721.

Sabbadin, A., and G. Zaniolo. 1979. Sexual differentiation and germ cell transfer in the colonial ascidian *Botryllus schlosseri*. J. Exp. Zool. 207:289–304.

Saupe, S. J. 2000. Molecular genetics of heterokaryon incompatibility in filamentous Ascomycetes. Microbiol. Mol. Biol. Rev. 64:489–502.

Strassmann, J. E., Y. Zhu, and D. C. Queller. 2000. Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. Nature 408:965–967.

Taylor, P. D. 1992. Altruism in viscous populations—an inclusive fitness model. Evol. Ecol. 6:352–356.

Taylor, P. D., and A. J. Irwin. 2000. Overlapping generations can promote altruistic behaviour. Evolution 54:1135–1141.

Tsutsui, N. D. 2004. Scents of self: the expression component of self/non-self recognition systems. Ann. Zool. Fennici 41:713–727.

Vitalis, R., and D. Couvet. 2001. Estimation of effective population size and migration rate from one- and two-locus identity measures. Genetics 157:911–925.

Worrall, J. J. 1997. Somatic incompatibility in basidiomycetes. Mycologia 89:24–36.

Xiang, Q., and N. L. Glass. 2004. The control of the mating type heterokaryon incompatibility by *vib-1*, a locus involved in *het-c* heterokaryon incompatibility in *Neurospora crassa*. Fungal Genet. Biol. 41:1063–1076.

Associate Editor: S. Otto

Supplementary Material

The following supplementary material is available for this article:

Appendix S1

Figure S1. Components of second-order effects on p_R.

Figure S2. Fixation times under strong structure.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1558-5646.2007.00191.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.