

Genetics of Mutualism: The Evolution of Altruism between Species

STEVEN A. FRANK

*Department of Ecology and Evolutionary Biology, University of California, Irvine,
CA 92717, U.S.A.*

(Received on 20 January 1994, Accepted on 6 June 1994)

Conditions are analyzed under which natural selection favors an individual to help another species at a cost to its own reproduction. Traditional models for the evolution of altruism between species focus on the genetic relatedness between the original donor and the recipients of return benefits from the mutualistic partner species. A more general model is analyzed here that focuses on the synergistic effects between partner species caused by genetic variability. The model shows that the spread of altruism is enhanced by spatial correlations between species in the genetic tendency to give aid to partners. These spatial correlations between species are similar to the kin selection coefficients of relatedness that determine the course of social evolution within species. The model also shows that natural selection and ecological dynamics can create genetic correlations between neighbors of different species, even when the initial spatial distributions of the species are uncorrelated. Genetic correlations between species may play an important role in the origin and maintenance of altruism between species.

1. Introduction

Why would an individual sacrifice its own direct reproduction to aid a member of another species? Williams (1966) suggested that mutualistic aid of this sort is unlikely to evolve by natural selection. The problem is that, even if the recipient returned benefits to the donor species, the original donor individual would rarely receive those benefits.

Hamilton (1972) and Wilson (1980, 1983a) agreed with Williams's logic but argued that, in many natural systems, limited dispersal keeps individuals of both species in close proximity. Thus aid given to another species is likely to be returned to the original donor (Trivers, 1971) or its immediate neighbors. Because, under limited dispersal, the neighborhood constitutes a kin group, returned benefits enhance the inclusive fitness of the original donor.

The Hamilton–Wilson argument focused on genetic variability in the donor, with kin or group selection as the mechanism by which mutualism can evolve. Although there is some disagreement about whether “group selection” or “kin selection” is the proper phrase for this scenario (Wilson,

1983b), the process itself is unambiguous (Hamilton, 1975).

This paper presents models in which both species of a mutualistic pair vary genetically in their tendency to aid their partners. Two interesting results emerge. First, the correlation between species in the tendency to aid partners plays an important role. The correlation between mutualistic partners has an effect similar to the genetic correlations within species that form the standard coefficients of relatedness in kin selection models. Positive correlations between species can greatly enhance the spread of mutualism.

The second result concerns how correlations between species may arise. Correlations between neighbors of the same species are easily understood: neighbors are often kin, and kin share common ancestors. By contrast, correlations between species must be created by special migration or selection processes. In the models presented here, the synergistic effect of mutual aid between species creates positive spatial correlations in the tendency to give aid to mutualistic partners. Wilson (1993) has recently suggested that a similar process may occur at the

TABLE 1

Definitions of notation used throughout this paper

Symbol	Definition
<i>Genetic variability only in the host</i>	
i	Genotypic value of a host individual. The amount of aid given by an individual increases linearly with its genotypic value.
i_s	Average genotypic value of hosts in the s -th subgroup of the population.
\bar{i}	The average genotypic value over all hosts in the population.
H_{si}	The abundance of the i -th host genotype in the s -th group.
h_{si}	The absolute fitness of the i -th host genotype in the s -th group.
\bar{h}	The average fitness of hosts over all genotypes and groups. Relative fitness of the i -th genotype in the s -th group is h_{si}/\bar{h} .
c	The cost per unit aid given to a partner species.
b	The fitness benefit per unit aid received from a partner species.
r_{isi}	Regression coefficient of the average group values for the hosts, i_s , on the individual values, i . This is the kin selection coefficient of relatedness among hosts within each group.
<i>Genetic variability in both host and mutualist</i>	
b_1	The fitness benefit to a host individual per unit of aid available in the local group from other hosts. These are indirect benefits received via the mutualist species, as shown in Fig. 1.
b_2	The fitness benefit to a host individual per unit of aid available in the local group from the mutualists.
j_s	Average genotypic value of mutualists in the s -th group of the population. Genotypic value determines the amount of aid given to hosts.
r_{jsi}	Regression coefficient of the mutualists' average genotypic value for groups s on individual host genotypes, i . Describes the genotypic association across space between the hosts' and mutualists' tendency to give aid to each other.
<i>The role of abundance</i>	
H_s	Total abundance of hosts in s -th group, $H_s = \sum_i H_{si}$.
M_{sj}	The abundance of the j -th mutualist genotype in the s -th group.
M_s	Total abundance of mutualists in the s -th group, $M_s = \sum_j M_{sj}$.
ϕ_s	Total aid provided by the hosts to the mutualists in the s -th group, $\phi_s = H_s i_s$.
μ_s	Total aid provided by the mutualists to the hosts in the s -th group, $\mu_s = M_s j_s$.
$r_{\phi si}$	Regression coefficient of the total host aid provided to mutualists in group s on individual host genotype in that group, i . Describes the association across space between host genotype and total aid provided by hosts to mutualists.
$r_{\mu si}$	Regression coefficient of the total mutualist aid provided to hosts in group s on individual host genotype in that group, i . Describes the association across space between host genotype and total aid provided by mutualists to hosts.
<i>The role of demography</i>	
b_1	The fitness benefit to a mutualist individual per unit of aid available in the local group from hosts.
c_1	The cost to hosts per unit aid given to the mutualist species.
c_2	The cost to mutualists per unit aid given to the host species.
V_a^m	The variance among groups in the average tendency of hosts to give aid, the variance of i_s .

V_a^m The variance among groups in the average tendency of mutualists to give aid, the variance of j_s .

Build-up of genetic correlation between species

$r_{\phi sj}$ Regression coefficient of the total host aid provided to mutualists in group s on individual mutualist genotype in that group, j . Describes the association across space between mutualist genotype and total aid provided by hosts to mutualists.

Symbols are listed according to the section of the paper in which they are first introduced. The symbol b_1 's meaning changes slightly in different sections and is therefore listed twice.

community level, in which ecological dynamics create a positive spatial correlation between mutualistic species. This community-level problem remains unsolved, however.

To analyze the evolution of ecological interactions between species, the population genetic tools that have been used to study social evolution are applied here. This approach puts ecological analysis in an evolutionary context, providing mechanisms for the coevolution of ecological traits in a way that emphasizes social interactions. Simultaneously, the ecological side of social evolution is made clearer, because genetic models traditionally analyze genotype frequency but ignore abundance. This expanded analysis provides a more fluid notion of sociality. In the Discussion I consider a wider interpretation of coefficients of relatedness, and the question of whether natural selection can favor the functional coherence of interacting species.

2. Models

2.1. GENETIC VARIABILITY ONLY IN THE HOST

Suppose that an individual of a host species provides aid to a potentially mutualistic species. The fitness advantage for providing this aid depends on the additional benefits that the mutualist returns to the host individual's relatives. The first model formalizes this idea and introduces the notation and approach for later extensions:

$$H'_{si} = H_{si}(1 + bi_s - ci), \quad (1)$$

where H_{si} is the abundance of the i -th host genotype in the s -th subgroup of the population, and the prime denotes abundance at one time step into the future. The genotype i specifies a quantitative character that determines the amount of aid given to mutualists. The amount of aid, i , is determined by additive genes, and i_s is the average amount of aid given by hosts in the s -th group. The cost to the host per unit aid given to the mutualists is c , and the benefit that the mutualists return to the s -th group per unit aid received from the

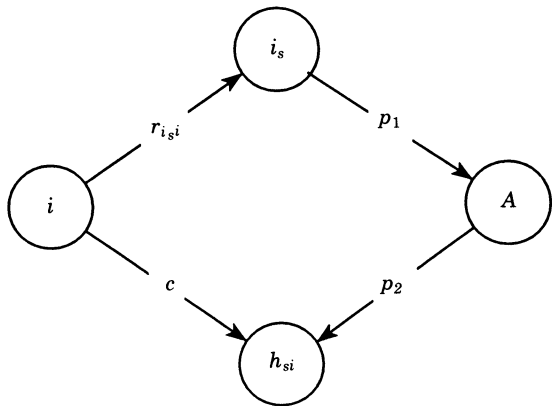


FIG. 1. Diagram of the model given by eqn (1).

i -th genotype is b . This is the standard linear additive model of kin-selected altruism (Hamilton, 1970, 1972; Uyenoyama *et al.*, 1981; Grafen, 1985; Wade, 1985; Queller, 1992). Notation for this and later sections is listed in Table 1.

A picture of this interaction is given in Fig. 1. The level of aid to the mutualist, i , directly affects fitness, h_{si} , by the rate $-c$. Fitness is defined by eqn (1), with $H'_{si} = H_{si} h_{si}$. The genotypic value of an individual, i , affects the average genotypic value of the group, i_s , by the rate $r_{i_s i}$; this coefficient is defined below. A host group with average genotype i_s provides aid to the mutualists at a rate p_1 . The aid the mutualists provide to hosts, A , is augmented at a rate p_2 by aid received from the hosts; this extra benefit is given to all members of the host subpopulation. The benefit to the group per unit cost to an individual is $b = p_1 p_2$, and the benefit to the original donor is $r_{i_s i} b$.

Behaviors are called altruistic if they cause individuals to give up some of their own direct reproduction in order to aid others (Hamilton, 1964). One way to study the evolution of altruistic behavior is to ask, for given values of cost to an individual and benefit to neighbors per level of altruism, whether natural selection favors an increase or decrease in the level of altruism. We can use the expression in eqn (1) for the fitness of a genotype with level of altruism i in an s type subgroup, $h_{si} = 1 + b_i s - ci$, to ask whether natural selection causes the average level of altruism to increase or decrease in the population. The simplest method for studying the evolution of altruism is the Price equation (Price, 1970, 1972; Hamilton, 1970). The technique as described by Queller (1992) will be used here.

According to the Price equation,

$$\bar{h} \Delta \bar{i} = \text{Cov}(h_{si}, i), \quad (2)$$

where \bar{h} and \bar{i} are the average values in the population for fitness and altruism, respectively, and Cov is the

covariance. The average level of altruism increases when $\text{Cov}(h_{si}, i) > 0$, which yields the condition known as Hamilton's rule (Grafen, 1985):

$$r_{i_s i} b - c > 0, \quad (3)$$

where $r_{i_s i}$ is the regression of average group values, i_s , on individual values, i .

This regression coefficient, $r_{i_s i}$, is the standard coefficient of relatedness r from kin selection theory (Hamilton, 1970, 1972). This kin selection coefficient measures the genetic similarity between a donor, who gives up c units of reproduction, and the recipients in group s , who gain b units of reproduction. The coefficient $r_{i_s i}$ can also be viewed as a measure of genetic variability among groups. In the case here, where all genetic variability is additive, $r_{i_s i} = V_a/V_i$, where V_a is the variance in genotypic values among groups—the variance in i_s , and V_i is the total variance in genotypic values in the population—the variance in i over all groups.

2.2. GENETIC VARIABILITY IN BOTH HOST AND MUTUALIST

The previous model showed that spatial patterns of genetic variability in the host determine the evolution of mutualism. This section demonstrates that spatial correlations in the genetic variability of hosts and mutualists can also influence the level of aid given to mutualistic partners. The model is

$$H'_{si} = H_{si}(1 + b_1 i_s + b_2 j_s - ci), \quad (4)$$

where j_s is the average tendency of the s -th group of the mutualists to provide benefit to the hosts. Using the Price equation as above, the condition for an increase in \bar{i} , the average level of aid given to mutualists, is

$$b_1 r_{i_s i} + b_2 r_{j_s i} - c > 0, \quad (5)$$

where $r_{j_s i}$ is the regression of the mutualists' average genotypic value for subgroups s on individual host genotypes, i .

The term $b_2 r_{j_s i}$ is the only change from the previous model. This term arises from the correlation between host genotype i and the tendency of the local mutualists to give aid. The role of this correlation can be seen in Fig. 2. If hosts provide no direct benefit to mutualists, $p_1 = 0$, the benefit to their neighbors is $b_1 = p_1 p_2 = 0$. The average level of altruism can, however, still increase by natural selection. This occurs when individuals with high values of i live in subgroups with mutualists that have a greater than average genetic tendency to give aid, high j_s , as measured by $r_{j_s i}$. In genetical language i is correlated with a high fitness trait, which in this case happens to

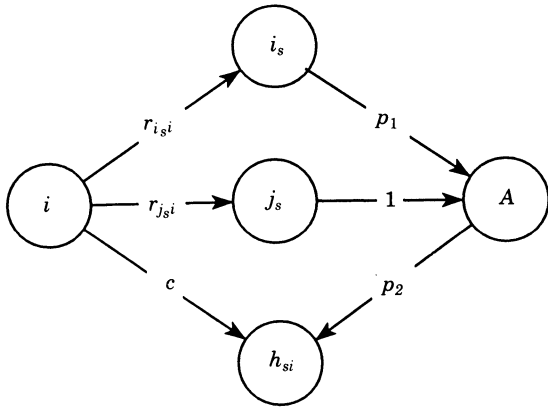


FIG. 2. Diagram of the model given by eqn (4). The aid provided by the local group of mutualists, A , is affected by the average genotype of mutualists j_s at a rate of one.

be the trait of another species. The benefit gained per unit measure of $r_{j_s i}$ is $b_2 = p_2$ (Fig. 2).

2.3. THE ROLE OF ABUNDANCE

The average tendency to give aid by donors, i_s and j_s for hosts and mutualists, influenced the fitness benefits of a host in the previous model. A more realistic assumption in many cases is that benefits depend on the total value of aid available in the group rather than the average value per donor. Specifically, the new model is

$$H'_{si} = H_{si}(1 + b_1 H_s i_s + b_2 M_s j_s - ci), \quad (6)$$

where H_s and M_s are the host and mutualist abundance in group s , so that $\phi_s = H_s i_s$ and $\mu_s = M_s j_s$ are total amounts of aid. The condition for an increase in the average aid given by hosts to mutualists, $\Delta \bar{i} > 0$, is

$$b_1 r_{\phi_s i} + b_2 r_{\mu_s i} - c > 0, \quad (7)$$

where $r_{\phi_s i}$ is the regression of total host aid to mutualists in group s on individual host genotype in that group, and $r_{\mu_s i}$ is the regression of total mutualist aid to hosts in group s on individual host genotype in that group.

2.4. THE ROLE OF DEMOGRAPHY

Spatial patterns of abundance and genetic variability on both sides of a mutualistic interaction determine how mutualistic traits evolve. The previous models set the spatial patterns as parameters and did not address how correlations within and between species might arise.

Genetic correlations within species typically occur because relatives live near each other and share common ancestors. How might correlations in genetic variability and abundance arise between species? To

address this question a model is examined in which, initially, there is no correlation between species.

Studying the course of selection within subgroups along with occasional migration episodes is tedious, so two limited goals are pursued. First, conditions are derived under which selection causes an increase in the average level of aid given from one species to its partner species. Second, conditions under which selection creates spatial associations between species, given initially uncorrelated distributions, are derived. The analysis could be extended to an explicit migration scheme, but selection-migration dynamics are not developed here.

The model of population dynamics and selection within each group s of the population is given by:

$$H'_{si} = H_{si}(1 + b_2 M_s j_s - c_1 i) \quad (8)$$

$$M'_{sj} = M_{sj}(1 + b_1 H_s i_s - c_2 j), \quad (9)$$

where M_{sj} is the abundance of the j -th mutualist genotype in the s -th group, and H_s and M_s are the total abundances of hosts and mutualists in the s -th group. In this model hosts are repaid for aid given to mutualists only after a time delay: the hosts aid the mutualists, which increases the abundance of mutualists, M_s , which then feeds back to provide a benefit in the term $b_2 M_s j_s$. A similar delay occurs from the mutualists' point of view. Thus a minimal model for the creation of spatial associations has two time periods of selection in each group s , before processes such as migration are allowed to mix the population and destroy the developing associations. Specifically, the above equations must be expanded to give the dynamics for H''_{si} and M''_{sj} . The analysis is given in the Appendix.

The condition for the increase in \bar{i} , the average level of aid given to mutualists by hosts, is

$$r_{i_s i} b^2 [\bar{j} + b(V_a^m + \bar{j}^2)] - 2c[1 + \bar{j}b + \bar{i}\bar{j}b^2/2] > 0, \quad (10)$$

where V_a^m is the variance among subpopulations in the average tendency of mutualists to give aid, j_s . A number of simplifying assumptions were made to obtain the condition (see Appendix). The most important assumptions are: host and mutualist abundances and genotypes are spatially uncorrelated before the first round of selection, so that $r_{j_s i} = 0$; $b_1 = b_2 = b$ and $c_1 = c_2 = c$; and c and $r_{i_s i}$ are both much smaller than b and \bar{j} . Also, any negative density-dependent effects occur globally rather than within local subpopulations. Note that, because of the symmetry of the dynamical system in eqns (8) and (9), conditions for change in \bar{j} are easily obtained from eqn (10).

Increasing genetic variance of the mutualist among subpopulations, V_a^m , favors the spread of helping behavior in the host. Likewise, increasing genetic variance in the host among subpopulations, V_a^h , favors the spread of helping behavior in the mutualist. Genetic variability among groups creates a spatial correlation between hosts and mutualists in the total amount of aid per subpopulation. This may be one reason why genetic variability favors mutualism. The association can be seen by the covariance after one generation of selection between the total amount of aid by hosts, $\phi'_s = H'_s i'_s$, and the total amount of aid by mutualists, $\mu'_s = M'_s j'_s$:

$$\text{Cov}(\phi'_s, \mu'_s) = \bar{j}b_1 V_a^h (1 - 2c_1 \bar{i}) + \bar{i}b_2 V_a^m (1 - 2c_2 \bar{j}) + b_1 b_2 V_a^h V_a^m, \quad (11)$$

where the key assumptions are given below eqn (10), and initial abundances are equal, $H_s = M_s$, before the first round of selection. (See the Appendix for assumptions about the distribution of host and mutualist genotypes.) Some assumptions make it easier to see what is required for a positive covariance. Let $\bar{j} = \bar{i} = \bar{k}$ and the values for b , c , and V_a be the same for the two species, thus

$$\text{Cov}(\phi'_s, \mu'_s) = 2b\bar{k}V_a(1 - 2c\bar{k}) + b^2V_a^2.$$

Small or moderate values of $c\bar{k}$, cost multiplied by current average level of aid, yield a positive covariance.

2.5. BUILD-UP OF GENETIC CORRELATION BETWEEN SPECIES

One key component of mutual aid between species is an association between the genotype of the host and the total value of aid by the local mutualists, r_{μ_i} [eqn (7)]. The previous section showed that selection often causes a build-up in the association between total aid of hosts and mutualists, $\text{Cov}(\phi'_s, \mu'_s) > 0$. Similar analyses show that selection also tends to create positive associations between the genotype of one species and the total value of aid given by the partner species, $r'_{\mu_i} > 0$ and $r'_{\phi_j} > 0$.

In this ecological model the total value of aid—abundance multiplied by average genotypic value—determines the benefit returned for aid provided to a mutualistic partner. However, given the traditional interest in purely genetic correlations in studies of social behavior, a brief look at the build-up in genetic correlations between species is in order.

The genetic correlation between species in average genotypic value per subpopulation can build up after one generation of selection, given initially uncorre-

lated distributions (see Appendix). When the effects of selection are weak, the correlation is approximately

$$\text{Corr}(i'_s, j'_s) = \frac{2bcV_w}{(1 + b\bar{j} - c\bar{i})^2} \left(1 - \frac{c^2 V_w}{(1 + b\bar{j} - c\bar{i})^2} \right),$$

where V_w is the average within group genetic variance, and for simplicity the parameters b , c and V_w are assumed to be equal for both species.

The correlation of neutral loci between species will remain zero after selection unless these neutral alleles are linked with the altruism loci.

3. Discussion

3.1. SUMMARY OF MAIN RESULTS

Natural selection favors helping another species only when benefits are returned to relatives of the original donor (Hamilton, 1972; Wilson, 1980). The first model formalizes this idea by an expression of Hamilton's rule for kin selection that is appropriate for interspecific interactions [eqn (3)]. In this kin selection model, benefits returned by a mutualistic partner species are weighted by the coefficient of relatedness between the original donor and the recipients, r_{i_j} . This coefficient is a type of genetic correlation between members of the same species.

The new models show that genetic associations between species create an additional force favoring mutual aid between species [eqn (5)]. In this case a strong tendency to give aid may be spatially correlated with a strong tendency to receive aid. Reciprocity is genetically determined, and natural selection favors an increase in mutual aid.

The models also show that selection can create positive correlations between species. For example, a subpopulation of one species that is relatively generous in giving aid to its partner species will have more partners from which to receive aid in the future. The same is true from the partner's point of view. Thus subpopulations with mutually generous individuals grow disproportionately large, creating a positive correlation between species in tendency to give aid. The only requirement of the model is that spatial associations last long enough that aid given by an individual to a partner species feeds back to provide additional benefits to the donor's relatives. This may occur within or between generations.

3.2. INTERPRETATION

In what sense can aid given to another species be considered "altruistic"? In the biological context,

“altruism” is usually taken to mean sacrifice of an individual’s own direct reproduction in order to aid the reproductive effort of another individual. It is customary to distinguish two types of altruism: kin selection (Hamilton, 1964), in which aid is provided to genetic relatives, and reciprocal altruism, in which aid is provided to non-relatives of the same or different species (Trivers, 1971). According to this tradition, aid given to another species is reciprocal altruism if benefits are returned, or maladaptive altruism if no benefits are returned.

The altruism caused by the regression coefficients between species, $r_{j,i}$ or $r_{\mu_s,i}$, does not fit easily into a traditional category. On the one hand, these coefficients arise in the equations in exactly the same way as the coefficients of relatedness within species: both types of coefficient measure the fitness benefits caused by an association between the genotype of an individual and the phenotypic properties of a group of neighbors. On the other hand, the coefficients between species have a relatively indirect effect: high altruism is favored when it is correlated with another high fitness trait, in this case the amount of altruism given by neighbors of another species. Coefficients of relatedness within species are more direct: high altruism is favored when it confers benefits directly to correlated genotypes.

The coefficients between species measure reciprocity. However, the reciprocity in this particular case is determined by a diffuse genetic and ecological correlation rather than the complex behavioral mechanisms between individuals typically associated with reciprocal altruism (Trivers, 1971). The uncertain status for the coefficients between species suggests either the need for a broader view, or a widening of old definitions and more elegant phrasing. (See Brown *et al.*, 1982, for a definition of a coefficient of reciprocity in a Trivers-type model of intraspecific reciprocal altruism.)

Both kin selection and associations between species create functional coherence for a group, where “functional coherence” means shared reproductive interests. This group-level coherence can be explained by stressing the fact that the selfish interests of the component individuals are aligned, or by stressing the fact that, as a unit, the group has beneficial traits that the components lack when alone. There is a long a rather painful literature about this distinction (Brandon & Burian, 1984). The point here is that selection can create associations between species and the co-transmission of mutually favorable traits. This causes either the functional coherence of the species pair, or the appearance of functional coherence, depending on one’s tastes.

3.3. EVOLUTIONARY DYNAMICS

Genetic correlation between species is determined by two opposing forces. Selection changes genetic variability within species into correlation between species, whereas migration of one species relative to its partner breaks down correlation. Thus a dynamical analysis requires knowledge about how fast new genetic variability within species is created relative to the rate at which migration destroys the correlation between species.

The dynamic tension between selection and migration determines the maintenance of a spatial correlation in mutual aid between species. It is possible, however, that spatial correlations played an important role only in the origin of mutual aid between species, and that genetic variability does not influence the maintenance of these traits [see eqn (10)]. Specifically, genetic variability and correlation between species may be required for the initial increase in the average levels of aid provided by partner species. As these average levels increase, the condition for the maintenance or further increase of mutual aid becomes more easily satisfied. Correlations between species may no longer be needed to satisfy the condition at this later stage.

The natural histories of symbiosis span the complete range of co-dispersal patterns, from vertically transmitted endosymbionts to widely dispersed propagules of infectious diseases. The shared interests of hosts and their vertically transmitted endosymbionts are well understood. The models presented here show that this extreme form of co-dispersal is just one end of a continuum. Natural selection builds correlations in genotype and abundance between more mobile partners. This creates varying degrees of shared reproductive interests between symbiotic species across a wide range of natural communities.

I thank F. Breden and Y. Michalakis for helpful comments on the manuscript. My research is supported by National Science Foundation grant DEB-9057331 and National Institutes of Health grants GM42403 and BRSG-S07-RR07008.

REFERENCES

- BARTON, N. H. & TURELLI, M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* **49**, 157–173.
- BRANDON, R. N. & BURIAN, R. M. (1984). *Genes, Organisms, Populations: Controversies over the Units of Selection*. Cambridge, MA: MIT Press.
- BROWN, J. S., SANDERSON, M. J. & MICHOD, R. E. (1982). Evolution of social behavior by reciprocation. *J. theor. Biol.* **99**, 319–339.
- FRANK, S. A. (1986). Hierarchical selection theory and sex ratios I. General solutions for structured populations. *Theor. Popul. Biol.* **29**, 312–342.

- FRANK, S. A. (1987). Demography and sex ratio in social spiders. *Evolution* **41**, 1267–1281.
- GRAFEN, A. (1985). A geometric view of relatedness. *Oxf. Surv. evol. Biol.* **2**, 28–89.
- HAMILTON, W. D. (1964). The genetical evolution of social behavior. I, II. *J. theor. Biol.* **7**, 1–52.
- HAMILTON, W. D. (1970). Selfish and spiteful behaviour in an evolutionary model. *Nature, Lond.* **228**, 1218–1220.
- HAMILTON, W. D. (1972). Altruism and related phenomena, mainly in social insects. *A. Rev. ecol. Syst.* **3**, 193–232.
- HAMILTON, W. D. (1975). Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (Fox, R., ed.) pp. 133–155. New York: Wiley.
- KIMURA, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. natn Acad. Sci. U.S.A.* **54**, 731–736.
- LANDE, R. (1975). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235.
- PRICE, G. R. (1970). Selection and covariance. *Nature, Lond.* **227**, 520–521.
- PRICE, G. R. (1972). Extension of covariance selection mathematics. *Ann. hum. Genet.* **35**, 485–490.
- QUELLER, D. C. (1992). A general model for kin selection. *Evolution* **46**, 376–380.
- SLATKIN, M. (1987). Heritable variation and heterozygosity under a balance between mutations and stabilizing selection. *Genet. Res.* **50**, 53–62.
- TRIVERS, R. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57.
- UYENOYAMA, M. K., FELDMAN, M. W. & MUELLER, L. D. (1981). Population genetic theory of kin selection: multiple alleles at one locus. *Proc. natn Acad. Sci. U.S.A.* **78**, 5036–5040.
- WADE, M. J. (1985). Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**, 61–73.
- WILLIAMS, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.
- WILSON, D. S. (1980). The natural selection of populations and communities. Menlo Park, CA: Benjamin/Cummings.
- WILSON, D. S. (1983a). The effect of population structure on the evolution of mutualism: a field test involving burying beetles and their phoretic mites. *Am. Nat.* **121**, 851–870.
- WILSON, D. S. (1983b). The group selection controversy: history and current status. *A. Rev. ecol. Syst.* **14**, 159–189.
- WILSON, D. S. (1993). Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. *Ecology* **73**, 1984–2000.

APPENDIX

Role of Demography

The details are given here for the derivation of the condition in eqn (10) in the main text. The model uses the recursions in eqns (8) and (9) in the text for two time periods of selection in each subpopulation.

The analysis here focuses on the host. Results for the mutualist follow directly by the symmetry of the model. To obtain information on selection both within and among subpopulations an expanded form of the Price equation is used (Frank, 1986, 1987):

$$\bar{h}\Delta\bar{i} = \text{Cov}(h_s^{(2)}, i_s) + E_s[\text{Cov}(h_{st}^{(2)}, i)], \quad (\text{A.1})$$

where $h_{st}^{(2)}$ and $h_s^{(2)}$ are the fitnesses after two time periods of selection for, respectively, host genotype i in subpopulation s , and the average over host geno-

types in subpopulation s . The operator E_s takes the average over subpopulations s .

The first step is to obtain an expression for the total change in each subpopulation after two time periods:

$$\begin{aligned} H_{st}'' &= H_{st}'(1 + b_2 M_s' j_s' - c_1 i) \\ &= H_{st}(1 + b_2 M_s j_s - c_1 i) \\ &\quad \times (1 + b_2 M_s' j_s' - c_1 i) = H_{st} h_{st}^{(2)}, \end{aligned}$$

where definitions for the terms can be found after eqn (8) in the main text.

Next we need expressions for M_s' and j_s' that are calculated from initial values in the subpopulation before the first round of selection:

$$\begin{aligned} M_s' &= M_s(1 + b_1 H_s i_s - c_2 j_s) \\ &= M_s m_s^{(1)} \\ j_s' &= j_s + \Delta j_s \\ &= j_s - c_2 V_s^m / m_s^{(1)}, \end{aligned}$$

where $m_s^{(1)}$ is the total fitness of the mutualists in subpopulation s after one round of selection, and V_s^m is the genetic variance of the mutualists in subpopulation s before the first round of selection.

Now we need an expression for the group fitness of subpopulation s :

$$\begin{aligned} H_s'' &= H_s'(1 + b_2 M_s' j_s' - c_1 i_s) \\ &= H_s(1 + b_2 M_s j_s - c_1 i_s)(1 + b_2 M_s' j_s' - c_1 i_s') \\ &= H_s h_s^{(2)} \\ i_s' &= i_s + \Delta i_s \\ &= i_s - c_1 V_s^h / h_s^{(1)} \\ h_s^{(1)} &= (1 + b_2 M_s j_s - c_1 i_s), \end{aligned}$$

where $h_s^{(1)}$ and $h_s^{(2)}$ are the total fitness of the hosts in subpopulation s after one or two rounds of selection, respectively, and V_s^h is the genetic variance of the hosts in subpopulation s before the first round of selection.

The condition for the increase in aid to the mutualist, $\Delta\bar{i} > 0$, is obtained by applying the expanded Price equation, eqn (A.1), and the following assumptions: the distributions of hosts and mutualists are independent before selection; the initial abundance of hosts and mutualists in each subpopulation before selection is $H_s = M_s = 1$; the distribution of genotypic values within subpopulations and among the averages of subpopulations are assumed to follow a normal distribution so that the fourth central moment is equal to three multiplied by the square of the variance, the third central moment is zero, and subpopulation means are uncorrelated with subpopulation

variances. The assumption of a normal distribution is typical in quantitative genetic analyses, and the moment relations follow as properties of normality. See, for example, Kimura (1965), Lande (1975), Barton & Turelli (1987), and Slatkin (1987).

Following through, the condition for an increase in \bar{i} is

$$r_{i,i}[b_1 b_2 \bar{j} + b_2^2 (b_1 - c_1)(V_a^m + \bar{j}^2) - c_1 b_2 \bar{j} (1 + b_1 \bar{i}) + 2\bar{i}c_1^2 (1 + 2b_2 \bar{j}) - c_1^3 (4V_a^h + 3\bar{i}^2)] - c_1 [2(1 - c_1 \bar{i} + b_2 \bar{j}) + b_1 b_2 \bar{i} \bar{j} - c_2 b_2 (V_i^m + \bar{j}^2)] > 0,$$

where V_a and V_i are, respectively, the genetic variance among subpopulations, s , and the total variance in the population, and superscripts on V denote whether the variance is for the mutualist, m , or the host, h .

The main features of interest in this condition can be seen more easily when $b_1 = b_2 = b$, $c_1 = c_2 = c$, and the order of magnitudes of c and r are approximately the same, and much smaller than the magnitudes of b , \bar{i} , and \bar{j} . With these assumptions the condition reduces to the expression given in eqn (10) in the main text.

Build-up of Genetic Correlations

This section shows how selection creates spatial correlations in genotypic value between species. Changes by selection are studied given that initial genotypic distributions of host and mutualists are independent, $\text{Cov}(i_s, j_s) = 0$. The selection equations are eqns (8) and (9) in the main text. Assumptions about the distributional properties of genotypes before selection are given in the previous section.

The condition of interest is $\text{Cov}(i'_s, j'_s) > 0$. The first step is to break the covariance into manageable pieces:

$$\begin{aligned} \text{Cov}(i'_s, j'_s) &= \text{Cov}(i_s + \Delta i_s, j_s + \Delta j_s) \\ &= \text{Cov}(i_s, \Delta j_s) + \text{Cov}(j_s, \Delta i_s) \\ &\quad + \text{Cov}(\Delta i_s, \Delta j_s) \end{aligned}$$

$$\begin{aligned} &= -c_2 V_w^m \text{Cov}(i_s, 1/m_s) \\ &\quad - c_1 V_w^h \text{Cov}(j_s, 1/h_s) \\ &\quad + c_1 c_2 V_w^h V_w^m \text{Cov}(1/h_s, 1/m_s), \end{aligned}$$

where V_w^h and V_w^m are the average within group variances for hosts and mutualists, respectively, and

$$\begin{aligned} h_s &= 1 + b_2 \bar{j} - c_1 \bar{i} + b_2 \delta_{j_s} - c_1 \delta_{i_s} \\ m_s &= 1 + b_1 \bar{i} - c_2 \bar{j} + b_1 \delta_{i_s} - c_2 \delta_{j_s} \end{aligned}$$

where δ_{i_s} and δ_{j_s} are independent normally distributed random variables with mean zero and variances V_a^h and V_a^m , respectively. The partitioned covariance has three terms. The first term, $\text{Cov}(i_s, \Delta j_s) = -c_2 V_w^m \text{Cov}(i_s, 1/m_s)$, is positive, because an increase in i_s causes an increase in m_s and thus a decrease in $1/m_s$. Similarly, the second term $\text{Cov}(j_s, \Delta i_s) = -c_1 V_w^h \text{Cov}(j_s, 1/h_s)$ is also positive. The third term, $\text{Cov}(\Delta i_s, \Delta j_s) = c_1 c_2 V_w^h V_w^m \text{Cov}(1/h_s, 1/m_s)$, can be negative, because an increase in i_s or j_s within a group changes the fitnesses, h_s and m_s in opposite directions. Thus $\text{Cov}(i'_s, j'_s)$ will be positive when $\text{Cov}(i_s, \Delta j_s) + \text{Cov}(j_s, \Delta i_s) > \text{Cov}(\Delta i_s, \Delta j_s)$, and negative when the inequality is reversed.

I have not been able to derive an exact condition that determines whether the genetic covariance is positive or negative. If we assume that the effects of the interactions are weak, that is, $h_s \approx 1$ and $m_s \approx 1$ and the genetic variances are small because $\delta_{i_s} \ll 1$ and $\delta_{j_s} \ll 1$, then an approximation for the genetic covariance can be obtained by taking the Taylor series expansions of $1/h_s$ and $1/m_s$ in δ_{i_s} and δ_{j_s} and keeping only the linear and second-order terms, yielding

$$\text{Cov}(i'_s, j'_s) \approx \frac{2bc V_w V_a}{(1 + b\bar{j} - c\bar{i})^2} \left(1 - \frac{c^2 V_w}{(1 + b\bar{j} - c\bar{i})^2} \right),$$

where, for simplicity, the parameters b , c , V_w and V_a are assumed to be equal for both species. The variance among groups after selection, V_a' , will typically be smaller than before selection, V_a , therefore the correlation between species is at least $\text{Cov}(i'_s, j'_s)/V_a'$.