



The Unit of Selection in Viscous Populations and the Evolution of Altruism

MINUS VAN BAALEN*†‡ AND DAVID A. RAND*

**Nonlinear Dynamics Laboratory, Mathematics Institute, University of Warwick,
Coventry CV4 7AL, U.K.*

(Received on 5 February 1998, Accepted in revised form on 23 April 1998)

Group selection can overcome individual selection for selfishness and favour altruism if there is variation among the founders of spatially distinct groups, and groups with many altruists become substantially larger (or exist longer) than groups with few. Whether altruism can evolve in populations that do not have an alternation of local population growth and global dispersal (“viscous populations”) has been disputed for some time. Limited dispersal protects the altruists from the non-altruists, but also hinders the export of altruism. In this article, we use the Pair Approximation technique (tracking the dynamics of pairs of neighbours instead of single individuals) to derive explicit invasion conditions for rare mutants in populations with limited dispersal. In such viscous populations, invading mutants form clusters, and ultimately, invasion conditions depend on the properties of such clusters. Thus there is selection on a higher level than that of the individual; in fact, invasion conditions define the unit of selection in viscous populations. We treat the evolution of altruism as a specific example, but the method is of more general interest. In particular, an important advantage is that spatial aspects can be incorporated into game theory in a straightforward fashion; we will specify the ESS for a more general model.

The invasion conditions can be interpreted in terms of inclusive fitness. In contrast with Hamilton’s model, the coefficient of relatedness is not merely a given genetical constant but depends on local population dynamical processes (birth, dispersal and death of individuals). With a simple birth rate function, Hamilton’s rule is recovered: the cost to the donor should be less than the benefit to the recipient weighted with the coefficient of relatedness. As the coefficient of relatedness is roughly inversely proportional to an individual’s number of neighbours, benefits to the recipient must be substantial to outweigh the costs, confirming earlier studies. We discuss the consequences for the evolution of dispersal and outline how the method may be extended to study evolution in interacting populations.

© 1998 Academic Press

1. Introduction

When one individual behaves altruistically towards another, it increases the beneficiary’s fitness at the expense of its own. To explain how such behavior can evolve has long been one of the benchmark problems of evolutionary biology. How can altruists increase in numbers if non-altruists, who do not pay the cost associated with altruism, always seem to have the advantage?

Kin selection was the first mechanism proposed (Hamilton, 1963, 1964): a gene promoting altruism can increase in frequency when there is a chance that the beneficiary also carries a copy, which may happen if the individuals are related. [Incidentally, the term “kin selection” was coined by Maynard Smith (1964), Hamilton used the term “inclusive fitness” to describe the mechanism.] This is the basis of “Hamilton’s Rule”: the cost of an altruistic act must be less than the benefit to the recipient multiplied by the “coefficient of relatedness”. Kin selection models have contributed greatly to the understanding of altruism among relatives; sterile workers in social

†Author to whom correspondence should be addressed.

‡Present address: Université P. et M. Curie, Institut d’Ecologie, Bât A 7^{me} Etage CC237, 7 Quai St.-Bernard, 75252 Paris cedex 05, France.

insects and birds helping their parents are textbook examples.

Group selection, the second mechanism that may promote altruism, seems rather different. It focuses not so much on interactions between pairs of individuals, but on processes in and between spatially separate subpopulations. If subpopulations with many altruists fare better than subpopulations with few (i.e. have a greater contribution to future generations), the global proportion of altruists may increase, even if locally the altruists do less well than the non-altruists (Maynard Smith, 1964; Wilson, 1977; Nunney, 1985). Originally, the term group selection referred to selection of traits that benefit an entire species, but this mechanism has been discredited because such populations do not resist invasion of cheaters. We use the term group selection in the more limited sense of selection of traits that favour the members of a subpopulation in a spatially structured system. Group selection in this more limited sense is more robust and has been invoked to explain phenomena like reduced virulence in pathogens. Within a host individual fast reproducers (virulent pathogens) will replace slow reproducers (avirulent pathogens), but hosts infected with more avirulent pathogens may live so long that per-host transmission of the avirulent pathogens is larger than that of the virulent pathogens (Eshel, 1977; Levin & Pimentel, 1981; van Baalen & Sabelis, 1995). Many models for the origin of life are based on the principle of group selection (Maynard Smith & Szathmáry, 1995).

An important limitation in our understanding of the scope of group selection (in the modern sense of the term) is that it has been modeled mainly for populations that are subdivided into discrete subpopulations. To what extent the mechanism also operates in systems that have a more continuous spatial structure has been debated. Whether, and if so, under what conditions, altruism can evolve in such more continuous spatial settings is the main question of the present article.

1.1. POPULATION VISCOSITY

Most models for group selection assume a population that exhibits a synchronized cycle of colonization and dispersal because such models are relatively easy to analyse. However, many species do not have a clear alternation of local population growth and global mixing. Hamilton (1964) conjectured that the principle might also work in what he called “viscous” populations, populations without imposed subdivision but with limited dispersal, because offspring tend to remain close to their

relatives, any individual is likely to have relatives in its neighbourhood. An altruistic individual may therefore convey the benefit (at least partly) to its relatives. The problem is that population viscosity at the same time increases the intensity of competition (for space and/or resources) among relatives, a factor which is known to impede the evolution of altruism (Hamilton, 1964).

On the basis of simulations using a cellular automaton model, Wilson *et al.* (1992) found that altruism is favoured only in what they considered to be a very limited (“unrealistic”) set of the parameter domain. An altruistic individual will still benefit from its altruistic neighbours, but when dispersal is limited, it will also compete for space with its altruistic neighbours. In other words, the altruists sit in each other’s way, making it more difficult for them to “export” their strategy. Taylor (1992a, b) has shown that if the “spatial scale of competition” is equal to the “spatial scale of dispersal” the benefit of altruism and the cost of local competition cancel out exactly. Taylor claims that this is always the case for viscous populations, effectively inhibiting the evolution of altruism.

This conclusion would imply that in viscous populations the effect of relatedness caused by population viscosity can be ignored, and that the quantity that is maximized by natural selection is therefore “individual fitness” (in which case individuals should maximize their own life-time reproduction ignoring any effect on neighbours). Our main aim in this article is to show that this is not true in general, and that in viscous populations it is not individual fitness that is maximized by natural selection, but rather the rate of growth of a cluster of relatives, and that the effect, though subtle, may be significant. These results can be interpreted in terms of “inclusive fitness” (Hamilton, 1963, 1964) where individuals weigh the effect they have on their neighbours’ reproduction according to the probability that these are related (i.e. belong to the same cluster of relatives).

1.2. APPROXIMATION OF SPATIAL MODELS

Invasion in viscous populations is a process that is hard to analyse. Reaction–diffusion models can be used to address some questions of spatial population dynamics, but in the case of the evolution of altruism they are inappropriate. Because they allow infinitely small densities, even a very small population of non-altruists will diffuse into the entire spatial domain, allowing them to exploit the altruists everywhere. To give the altruists a chance, they must be protected against exploitation by such

“nano-individuals”. It has already been shown by Goodnight (1992) that for altruists to invade a system of so-called “budding” populations, the composition of daughter populations must be a small random sample from their parent population. If the daughter populations are of exactly the same composition as the parent population, the proportion of non-altruists will ceaselessly increase until all populations are dominated by non-altruists. However, if the number of individuals that buds off is small, there will be variation due to sampling error so that some daughter populations have a higher proportion of altruists. If such populations produce more daughter populations, the global proportion of altruists will eventually increase.

Thus, for the evolution of altruism there must be discreteness and associated stochasticity (Goodnight, 1992). This means that we should analyse models that are individual-based as well as spatial. Probabilistic cellular automaton (PCA) models (which we will describe in more detail) satisfy these criteria. However, even when we assume haploid reproduction (and thus ignore genetics) such PCA models are easy to simulate but very hard to analyse.

1.3. THE PAIR APPROXIMATION

Although simulations of PCA models are excellent for developing intuition and formulating conjectures, as models they are hard to analyse. Therefore, one should consider more controllable models for which there is more mathematical understanding and which can be more directly connected with biological data. The most promising of such models follow from a correlation equation formalism (Matsuda *et al.*, 1992; Keeling & Rand, 1995; Keeling, 1995). Such models are more robust to the assumptions underlying their derivation and these assumptions are more open to experimental verification. The approach has been applied to a range of systems, such as host-parasite models (Satō *et al.*, 1994; Keeling & Rand, 1995; Keeling, 1995), vegetation dynamics (Harada & Iwasa, 1994) and spatial games (Morris, 1997).

Matsuda *et al.*'s (1992) formalism, the so-called pair approximation technique, models space implicitly, by focusing on the interaction between nearest neighbours. Matsuda *et al.* (1992) and recently Harada *et al.* (1995) and Nakamaru *et al.* (1997) have applied this technique to demonstrate that altruists and non-altruists may coexist in a viscous population, thus contradicting Taylor's (1992a, b) conclusion. In this article, we will elaborate on these studies by focusing on the more general problem of invasion of a rare population in viscous

system, and use the results to derive the invasion conditions for altruists.

1.4. FITNESS IN VISCOUS POPULATIONS

Following Metz *et al.* (1992) and Rand *et al.* (1994) we define the fitness of a rare mutant to be simply its per capita rate of growth when rare. Therefore if the mutant's fitness is positive, the mutant increases in number and can invade, if it is negative the mutant will disappear.

Closely associated with the concept of fitness is that of the “unit of selection”. The definition of the unit of selection as that entity “whose fitness is maximized” (see, *e.g.* Dawkins, 1982) is dangerously circular. However, our method allows to link the two concepts closely together. In fact, fitness and unit of selection have to be calculated simultaneously.

In viscous systems, multiple differential equations are necessary to describe the invasion dynamics of a rare population, in contrast to well-mixed systems where a single equation suffices. These differential equations can be concisely represented in matrix form. Thus, the invasion exponent (fitness) will be given by the dominant eigenvalue of a matrix. The corresponding eigenvector describes the spatial structure of the clusters that form when the rare population invades (as in Fig. 1). Mathematically, fitness and unit of selection follow from a set of simultaneous equations.

We will use these results to determine under what conditions altruists can invade a system dominated by non-altruists (assuming haploid inheritance). It will turn out that the invasion condition is very similar to Hamilton's Rule, but the “coefficient of relatedness”

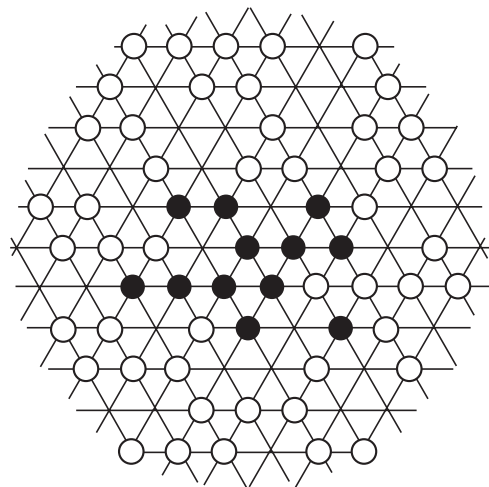


FIG. 1. An example of a cluster of altruists (black) invading a population of non-altruists (white) living on a triangular lattice ($n = 6$).

TABLE 1
Demographic events

	Event*	Rate†
Birth	$S_o \rightarrow SS$	ϕb_S
	$A_o \rightarrow AA$	ϕb_A
Death	$Sj \rightarrow oj$	ϕd_S
	$Aj \rightarrow oj$	ϕd_A
Migration	$So \rightarrow oS$	ϕm_S
	$Ao \rightarrow oA$	ϕm_A

*"Mirror image" events have the same rate.

†The factor $\phi = 1/n$ scales the *per capita* rates b , d and m to per-pair rates.

is not a genetical quantity (since we are assuming haploid inheritance, the genetic coefficient of relatedness can only assume the values zero or one). Instead, the coefficient of relatedness gives the probability that a site neighbouring an altruist is occupied by another altruist. This probability can be calculated from the characteristic cluster structure (unit of selection).

Finally, we will derive ESS conditions for the more general case in which there is a continuous range of strategies, varying in cost and altruistic benefit to neighbours. As an example, we will discuss the evolution of competition for light in a population of plants.

1.5. ERRORS

Approximations are never perfect, and for the simplest correlation equations, the associated errors can be shown to be significant. However, although we will discuss the errors introduced in the various approximation steps, we will focus on the basics of the technique. A more rigorous error analysis will be published elsewhere (Morris, 1997).

2. The Model Framework

2.1. THE LATTICE

In this framework space is represented by a network of sites, in which every site is connected to n neighbouring sites. Every site is either occupied by an altruist individual (A), occupied by a non-altruist individual (S) or it is empty (o).

2.2. EVENTS

The state of the lattice will change over time as a consequence of three types of events, birth, death and migration. Any of these events may occur with a given probability per unit of time (i.e. the lattice is asynchronously updated). Table 1 lists the events as "pair events", events that are not so much associated with the sites as well as with the connections between sites. The advantage of this theoretical framework is

that it allows the inclusion of migration, because such events change the state of two neighbouring sites at the same time. This framework is called an "artificial ecology" (Rand *et al.*, 1995) which is a more general framework than the classical "probabilistic cellular automaton" which allows only single site changes at a time (van Baalen, 1998).

We assume that the per capita rates of mortality (d_i , with $i = A, S$) and migration (m_i) are constant, and that an individual's rate of reproduction depends on its neighbourhood. Specifically, we assume that the rate of reproduction of any individual is increased by an amount B/n by every altruist in its neighbourhood. If the individual in question is an altruist itself, its rate of reproduction is decreased by an amount C , representing the cost of altruism. (This is actually a limited definition of altruism. Later on, we will also analyse the case where the cost of altruism increases the altruist's mortality rate instead of decreasing its birth rate.) Thus, the birth rate of a non-altruist at a site x with $n_A(x)$ altruist neighbours equals

$$b_S(x) = b_0 + B \frac{n_A(x)}{n} \tag{1}$$

(b_0 is the base-line rate of reproduction), whereas the birth rate of an altruist at the same site would equal

$$b_A(x) = b_0 + B \frac{n_A(x)}{n} - C \tag{2}$$

In the same environment, a non-altruist will therefore always have a higher probability to reproduce than an altruist. Incidentally, note that though we will call B the rate of reproduction or birth rate, in fact it should be called the rate of reproductive effort: whether or not an individual *actually* reproduces depends on whether it has neighbouring sites that are empty.

Whenever a birth, death or migration event occurs, the lattice jumps from one discrete state to another, and as the events are stochastic, the actual state of the lattice will quickly become unpredictable. One way of studying the properties of such systems is by computer simulation. For example, computer simulation could be used to determine whether a small number of altruists are able to invade a lattice that is dominated by non-altruists. A disadvantage of this approach is that large lattices must be simulated for a long time, to overcome demographic stochasticity and to allow them to settle at their attractor behaviour. More importantly, the approach yields no explicit relationships between parameter values and the dynamics of the system.

Such analytical insight is possible, however, through application of techniques developed in statistical mechanics. The basic idea behind these

so-called “correlation dynamics” models is that although the precise state of the lattice is unpredictable, it is nevertheless possible to work out the expected rates of change of certain average quantities such as the proportion of sites in state i or the proportion of pairs in state ij (see Durrett, 1988 and Durrett & Levin, 1994 for an outline of the underlying theory).

The resulting differential equations completely bypass the need of keeping track of the entire lattice. Space is thus modeled implicitly. Since the dynamics of a system with very many dimensions (that is, equal to the number of sites in the lattice) is reduced to a system of much lower dimensionality, inevitably information is lost. Where appropriate we will briefly discuss the errors that are associated with the approximation, but for a more in-depth discussion we refer to Morris (1997).

3. Pair Dynamics

3.1. GLOBAL AND LOCAL DENSITIES

The proportion of sites in state i , denoted by p_i , corresponds to the classical concept of the “density” of i . Correlation dynamics models, however, extend the density concept to larger configurations than single sites. The simplest of these is the pair of neighbouring sites, and therefore these models are called “pair approximation” models. Just as with the single sites, the “pair density” p_{ij} denotes the proportion of all pairs that happens to be in state ij . (Notice that since every site in the lattice is connected to n neighbours it forms part of n pairs.) Thus if a pair of neighbouring sites is picked, the probability that they are in states i and j is p_{ij} .

The main advantage of knowing the proportions of pairs is that the conditional probabilities

$$q_{j|i} = \frac{p_{ij}}{p_i} \tag{3}$$

can be calculated, which gives the probability that a given neighbour of a site in state i is in state j . Because $q_{j|i}$ specifies the density of species j as experienced by the average i individual (which may be different from the global density p_j) one may therefore speak of $q_{j|i}$ as a *local density* (Matsuda *et al.*, 1992 employ the term “environs density”).

Incidentally, because we assume that the number of neighbours is constant, the “singlet” density p_i follows from the pair densities,

$$p_i = \sum_j p_{ij} \tag{4}$$

Would the number of neighbours vary from site to site this will not hold, and the p_i would have to be tracked separately (Morris, 1997).

3.2. BOOKKEEPING

The technique boils down to tracking changes in the proportions of pairs, in much the same way as one would track the occupancy of single sites. This requires bookkeeping of how events change the proportions of all pair combinations. This bookkeeping is complicated because members of a pair form part of other pairs and therefore the rates of change in the proportion of a particular pair combination is affected by events in neighbouring pairs. For example, a given So pair may become an SS pair because the S individual reproduces into the o -site, but it can also become an SS pair because of a migration or reproduction event in a pair formed by the empty site and its other neighbours.

Averaging over all possible pairs on the lattice, and averaging over all possible events that may occur at these pairs leads to a set of differential equations that give the expected rate of change all possible pair densities p_{ij} . With three states, there are nine different pair combinations, but symmetry relations ($p_{ij} = p_{ji}$) and the fact that the pair densities sum to one, leaves us with a set of five differential equations, for $p_{oS} (= p_{So})$, p_{SS} , $p_{Ao} (= p_{oA})$, $p_{AS} (= p_{SA})$ and p_{AA} . These equations take into account all transitions shown schematically in Fig. 2; the full equations are given in Appendix A.

3.3. CLOSURE

An elementary aspect of these differential equations is that the rates of change in the pair frequencies depend on frequencies of configurations larger than pairs. Take for example the conditional probabilities

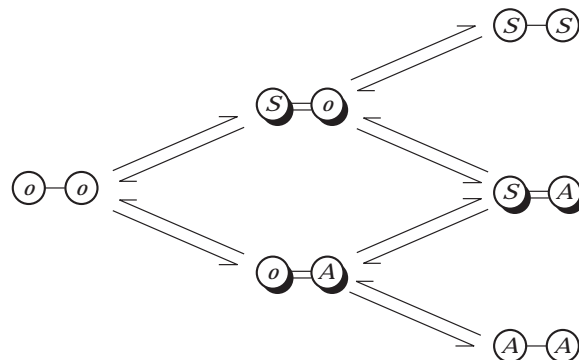


FIG. 2. The possible transitions between the state of doublets (pairs of neighbouring sites). Pairs that have a symmetric counterpart are shaded.

of the type q_{hij} , which give the probability that a neighbour of the i in an ij pair is occupied by an h . (For example, the probability that an oS becomes an AS pair because an A -neighbour of the pair reproduces will be proportional to $b_A q_A |oS$). From elementary probability theory we have,

$$q_{hij} = \frac{p_{hij}}{p_{ij}} \quad (5)$$

which implies that q_{hij} depends on the frequency of hij triplets. In fact, the differential equations will depend on the frequencies of more complex configurations, as birth rates depend on the entire configuration surrounding reproducing individuals.

Hence if we are to describe the dynamics of pairs in terms of pair frequencies (i.e. “close” the system) we have to estimate or approximate the distribution of these larger configurations in terms of pair frequencies. For conditional probabilities of the type q_{hij} the most straightforward strategy is to adopt the so-called pair approximation, i.e. to assume that

$$q_{h|ij} \approx q_{hi} \quad (6)$$

i.e. the probability to find an h next to the i is assumed not to be affected by i 's other neighbour j (Matsuda *et al.*, 1992).

This assumption may introduce a significant error. Consider, for example, $q_{A|SA}$. Under the pair approximation assumption, this would be approximated by $q_{A|S}$, a quantity that is very small when the altruists are rare (on average, the non-altruists do not “see” altruists). However, $q_{A|SA}$ is the probability that the non-altruist has a *second* altruistic neighbour. This implies that the S in question is likely to be in a region where A is locally abundant, and that $q_{A|SA}$ therefore does not approximate zero. The “standard” pair approximation thus ignores an important aspect of spatial structure. In the discussion we will present preliminary results using an improved approximation.

4. Mean-field Dynamics

Before analysing the spatial dynamics, it is instructive to consider the equivalent non-spatial (“mean-field”) model. When the migration rates m_A and m_S become very large, the populations become “well-mixed” and the dynamics are governed by the following differential equations:

$$\begin{aligned} \frac{dp_S}{dt} &= [(b_0 + Bp_A)p_o - d]p_S \\ \frac{dp_A}{dt} &= [(b_0 + Bp_A - C)p_o - d]p_A \end{aligned} \quad (7)$$

where $p_o = 1 - p_S - p_A$.

From this system it can be deduced immediately that (1) the non-altruists always have a higher birth rate as they do not pay the cost of altruism and therefore (2), the altruists can never invade the equilibrium population of the non-altruist. This, of course, is nothing but a restating of the classical dilemma of the evolution of altruism. Thus, if in the pair approximation model the altruists *can* invade, we know that it is a consequence of spatial structure.

5. Invasion

5.1. INVASION DYNAMICS

In order to determine under what conditions altruists can invade a system dominated by the non-altruists, we proceed exactly as we would for well-mixed populations. First, we work out the dynamics of the system in the absence of altruists, and then we derive the “invasion exponent” (Metz *et al.*, 1992; Rand *et al.*, 1994) for a small population of altruists.

In the absence of altruists, the non-altruists will settle at a stable equilibrium $\bar{p}_S = \bar{p}_{oS} + \bar{p}_{SS}$ (see Matsuda *et al.*, 1992 for its derivation). At this point, it is sufficient to verify that the resident has a positive equilibrium which is the case if b_0 is sufficiently larger than d_S (Matsuda *et al.*, 1992).

If the altruists are (globally) rare (i.e. p_{Ao} , p_{AS} and p_{AA} are all very small) they do not affect global dynamics of the resident non-altruists. As a consequence the invasion dynamics of the altruists are governed by three differential equations, for p_{Ao} , p_{AS} and p_{AA} . Matsuda *et al.* (1992) express these in the form

$$\frac{dp_{ij}}{dt} = M_{ij}p_{ij} \quad (8)$$

in which an expression M_{ij} is called the “Malthusian” of p_{ij} . Here, however, we represent the system in a different way (though formally equivalent), that is, in matrix form:

$$\frac{d\mathbf{p}_A}{dt} = \mathbf{M}(\mathbf{q}_A)\mathbf{p}_A \quad (9)$$

where

$$\mathbf{p}_A = \begin{bmatrix} p_{Ao} \\ p_{AS} \\ p_{AA} \end{bmatrix} \quad \text{and} \quad \mathbf{q}_A = \begin{bmatrix} q_{o|A} \\ q_{S|A} \\ q_{A|A} \end{bmatrix} \quad (10)$$

and $\mathbf{M}(\mathbf{q}_A)$ is a 3×3 matrix that is fully given in Appendix B.

The important observation about $\mathbf{M}(\mathbf{q}_A)$ is that it depends on the local densities $q_{i|A}$ and thus the pair frequencies p_{iA} . As a consequence, the system of differential equations is not linear. However, as Matsuda *et al.* (1992) have shown, these local densities equilibrate quickly even when the global density of a population (i.e. p_A) still changes over time. Denoting the equilibrated local densities by $\tilde{\mathbf{q}}_A$, the ‘‘invasion matrix’’ $\mathbf{M}(\mathbf{q}_A)$ will converge to a matrix with constant elements $\tilde{\mathbf{M}} = \mathbf{M}(\tilde{\mathbf{q}}_A)$, and once that happens, the invasion dynamics of the altruists is linear. The long-term dynamics of this linear system is easy to solve:

$$\mathbf{p}_A(t) = c\tilde{\mathbf{q}}_A e^{\lambda t} \quad (11)$$

where λ is the dominant eigenvalue of $\tilde{\mathbf{M}}$, $\tilde{\mathbf{q}}_A$ the normalized eigenvector and c a constant that depends on initial conditions. With this expression the invasion condition becomes obvious: if the dominant eigenvalue of $\tilde{\mathbf{M}}$ is positive, the altruists will invade; if it is negative they will die out. The dominant eigenvalue, or ‘‘invasion exponent’’ (Metz *et al.*, 1992; Rand *et al.*, 1994), is therefore the relevant fitness measure.

5.2. THE UNIT OF SELECTION

If the dominant eigenvalue of the invasion matrix is the relevant fitness measure, the eigenvector $\tilde{\mathbf{q}}_A$ can be interpreted as the associated ‘‘unit of selection’’: the altruists do not increase randomly distributed over the lattice, but as clusters that grow. It is precisely this cluster structure that is described by the eigenvector $\tilde{\mathbf{q}}_A$. Perhaps the best way to understand the relation between $\tilde{\mathbf{q}}_A$ and cluster structure is to reverse the argument, and calculate the neighbourhood vector from a given spatial distribution. Consider, for example the cluster of altruists in Fig. 1. It consists of 12 individuals and as every individual has six neighbours there are 62 pairs with at least one A ; of these 23 are oA pairs, 21 are SA pairs and 28 are AA pairs. Thus, this particular cluster corresponds to

$$\mathbf{q}_A = \begin{pmatrix} q_{o|A} \\ q_{S|A} \\ q_{A|A} \end{pmatrix} = \frac{1}{72} \begin{pmatrix} 23 \\ 21 \\ 28 \end{pmatrix} \approx \begin{pmatrix} 0.32 \\ 0.29 \\ 0.39 \end{pmatrix} \quad (12)$$

In an inverse, slightly more roundabout way, the vector \mathbf{q}_A can be used to ‘‘reconstruct’’ the spatial cluster structure of the mutant.

In a later section we will give an approximation for the invasion exponent of a rare mutant, but first we will focus on invasion conditions, which are considerably more simple.

5.3. INVASION CONDITIONS

The simplest case to analyse is that in which the non-altruists and altruists are identical in every respect except for the fact that altruists help their neighbours at their own cost. Then, the basic question is whether given common (base-line) birth, death and migration rates (b_0 , d and m), an altruistic mutant (with $B > 0$ and $C > 0$) can invade a non-altruistic population (with $B = C = 0$). Since birth rate is the only demographic aspect that can differ between altruists and non-altruists, it should come as no surprise that the invasion condition is

$$b_A > b_S \quad (13)$$

This does *not* imply that we have recovered individual selection here: birth rate of the altruists depends on their local density $\tilde{q}_{A|A}$ and thus on the altruist’s cluster structure. (The birth rate of the non-altruists is b_0 because the altruists are globally rare, so that $q_{A|S} \approx 0$). This implies that the altruists can invade if

$$B(1 - \phi)\tilde{q}_{A|A} - C > 0 \quad (14)$$

that is, we have recovered a variant of ‘‘Hamilton’s Rule’’ in which the coefficient of relatedness estimates how much of an altruist’s environment consists of other altruists [given by the factor $(1 - \phi)\tilde{q}_{A|A}$]. This has to be calculated from the invasion matrix; the full expression (given in Appendix B) is rather messy. However, if we assume that B and C are small compared with b_0 , we obtain

$$\tilde{q}_{A|A} \approx \phi \frac{b_0}{b_0 + (1 - \phi)m} \quad (15)$$

If the rate of migration (m) is low, $q_{A|A}$ will be approximately equal to $\phi = 1/n$: an altruist will have, on average, one other altruist in its neighbourhood. If m increases, $q_{A|A}$ decreases, as expected. Thus, we can conclude that for the altruists to invade, the benefit B should be at least n times as large as the cost C . This, already, suggests that the conditions for altruism to evolve are rather restrictive, which is in agreement with Wilson *et al.*’s (1992) simulations

For the more general case where the other demographic rates are allowed to differ as well, the invasion condition becomes

$$\frac{b_0 + B(1 - \phi)\tilde{q}_{A|A} - C + m_A}{d_A + (1 - \phi)m_A\tilde{q}_{o|A}} > \frac{b_0 + m_S}{d_S + (1 - \phi)m_S\tilde{q}_{o|S}} \quad (16)$$

Notice that this is basically a comparison of a modified “birth/death” ratio of altruists and non-altruists, that is, the invasion condition can be written as

$$\frac{\alpha_A}{\delta_A} > \frac{\alpha_S}{\delta_S} \tag{17}$$

where

$$\alpha_i = b_i + m_i \tag{18}$$

($\alpha_i q_{i0}$ gives the probability that an i -individual will arrive at a given empty site), and

$$\delta_i = d_i + (1 - \phi)m_i q_{0i} \tag{19}$$

gives the probability that an i -individual will disappear from a site. This suggests that in order to maximize its fitness, the altruist cluster should maximize this birth/death ratio, and therefore it could stand as a proxy for fitness (see Appendix B for a more formal derivation). However, the condition can already be used to infer the invasion conditions for the case where the cost of altruism is incurred as an increased mortality rate instead of a decreased birth rate (which is more in line with the popular image of an altruist sacrificing itself for its relatives!). If migration is zero, we obtain

$$\frac{b_0 + B(1 - \phi)\tilde{q}_{A|A}}{d + C} > \frac{b_0}{d} \tag{20}$$

which leads to

$$B(1 - \phi)\tilde{q}_{A|A} > \frac{b_0}{d} C \tag{21}$$

This is a variation of Hamilton’s Rule where costs have to be corrected for background birth and mortality rates. As background birth rate exceeds background mortality rate (a necessary assumption because otherwise the non-altruists would go extinct) the benefits of altruism should be even larger to allow the altruists to invade.

6. ESS Conditions

If altruists can invade, the end result may well be coexistence because non-altruists can easily invade a system dominated by altruists. However, this does not imply that local interactions in viscous populations will inevitably lead to polymorphism. A third type, intermediate to the pure non-altruist and the full-blown altruist might be able to invade and eventually displace both original types.

To determine which strategy is the true ESS (evolutionarily stable strategy, Maynard Smith & Price, 1973), a continuous range of options (strategy set) must be considered. This means that instead of two parameters, an entire relationship between B and C must be specified, which can become quite arbitrary if there is no underlying idea about the sort of interactions that are involved. A better option is to construct a more realistic example, as we will do for a simple model for competition for light among plants. To outline the approach, however, we will assume that individuals are characterized by a strategy s that affects both the benefits they confer to their neighbours and the costs they incur themselves.

$$\begin{aligned} B &= B(s) \\ C &= C(s) \end{aligned} \tag{22}$$

The level of altruism is then a single parameter strategy (which we assume can be chosen from a continuous strategy set), and the question becomes which strategy (or strategies) will be favored by natural selection. A first step in answering this question is to determine which of these strategies is evolutionarily stable. A strategy s^* is an evolutionarily stable strategy (ESS) if no rare mutant with a deviating strategy $s \neq s^*$ can invade (Maynard Smith & Price, 1973), or

$$\lambda_{s^*}(s) < \lambda_{s^*}(s^*) \tag{23}$$

for all $s \neq s^*$, where $\lambda_{s^*}(s)$ denotes the invasion exponent (i.e. fitness, Metz *et al.*, 1992; Rand *et al.*, 1994) of strategy s when s^* is the resident strategy.

The only difference with the standard ESS definition is that the invasion exponent is derived from the pair equation, and that “fitness” therefore may refer to a higher level of selection. The basic approach is the same as that of invasion of altruists in a non-altruist population: assume that the resident (now denoted R) is at equilibrium ($p_{R0} > 0$, $p_{RR} > 0$), and trace the dynamics of the mutant (M) when it is rare. Mutant dynamics is derived in the same way as in the previous section, the only difference is that it will depend on the level of altruism of the residents. (We have relabeled the types R and M , because there is no longer a qualitative distinction between altruists and non-altruists.)

The invasion exponent of a mutant close to the resident is approximated by

$$\lambda_{s^*}(s^* + \Delta s) = \frac{\alpha_R + \delta_R}{\alpha_R + \beta_R + \delta_R} \left(\frac{\delta_R}{\alpha_R} \Delta \alpha - \Delta \delta \right) \tag{24}$$

where $\Delta\alpha = \alpha_M - \alpha_R$ and $\Delta\delta = \delta_M - \delta_R$ (as shown in Appendix D). We can immediately see that the mutant's invasion exponent is zero if

$$\frac{\Delta\alpha}{\Delta\delta} = \frac{\alpha_R}{\delta_R} \quad (25)$$

Taking the limit $\Delta s \rightarrow 0$, we conclude that the ESS should satisfy

$$\frac{\frac{d\alpha}{ds}}{\frac{d\delta}{ds}} = \frac{\alpha_R}{\delta_R} \quad (26)$$

for $s = s^*$. This result is an example of the so-called "marginal value" principle (Charnov, 1976) and implies that at the ESS the ratio α/δ is maximized, as was already conjectured in the previous section.

If it is assumed that costs decrease birth rate, the per-capita demographic rates of type i ($i = R, M$) become

$$b_i = b_0 + B(s^*)(1 - \phi)q_{R|i} + B(s)(1 - \phi)q_{M|i} - C(s_i) \quad (27)$$

$$d_i = d$$

(where $s_M = s, s_R = s^*$). As in this case mortality does not depend on s , evolutionary stability implies

$$\frac{d\alpha_M}{ds} = 0 \quad (28)$$

which in turn leads to

$$\frac{dB}{ds}(1 - \phi)\tilde{q}_{M|M} = \frac{dC}{ds} \quad (29)$$

This is a marginal value analog of Hamilton's Rule: ESS conditions, a change in benefits multiplied by the coefficient of relatedness $\tilde{q}_{M|M}$ counterbalances the associated change in costs. Using results from the previous section, for a mutant at the invasion boundary the coefficient of relatedness is given by

$$\tilde{q}_{M|M} = \phi \frac{b_M}{b_M + (1 - \phi)m_M} \quad (30)$$

where, because we assume the mutant to be close to the resident,

$$b_M \approx b_0 + B(s^*)(1 - \phi)\tilde{q}_{R|R} - C(s^*) \quad (31)$$

Thus, except when $m = 0$ (no migration), the coefficient of relatedness of the mutant has to be calculated from the "subjective" resident density $\tilde{q}_{R|R}$, which is the positive solution of the resident equilibrium condition

$$(b_0 + B(s^*)(1 - \phi)\tilde{q}_{R|R} - C(s^*))(1 - \tilde{q}_{R|R}) - d_R = 0 \quad (32)$$

If we assume that costs increase death rate instead of decreasing birth rate, the ESS condition becomes more complex. The per-capita demographic rates of type i then become

$$b_i = b_0 + B(s^*)(1 - \phi)q_{R|i} + B(s)(1 - \phi)q_{M|i} \quad (33)$$

$$d_i = d + C(s_i)$$

The resident equilibrium then is the solution of

$$(b_0 + B(s^*)(1 - \phi)\tilde{q}_{R|R})(1 - \tilde{q}_{R|R})(d_R + C(s^*)) = 0 \quad (34)$$

and the ESS should satisfy

$$\frac{dB}{ds}(1 - \phi)\tilde{q}_{M|M} = \frac{b_R}{d_R} \frac{dC}{ds} \quad (35)$$

This is yet another variant of Hamilton's Rule, one in which the costs have to be corrected for birth, death and migration. Again other variants would arise if the benefits of altruism affected mortality rate instead of birth rate, which would be the case if individuals help each other to survive, for example by contributing to a communal defense against predators. Nonetheless, the basic approach remains the same. The example serves to warn that we should be careful, and specify as precisely as possible how "costs" and "benefits" affect demographic rates. It will make a difference if costs affect the rate of mortality instead of the rate of reproduction, even if measured in the same unit.

7. Competition for Light

Discussions of the evolution of altruism will remain rather academic if "costs" and "benefits" remain unspecified. We will work out a very simple example for the competition for light among plants to demonstrate the principle, and to show how to derive expressions for costs and benefits in a concrete example. Consider a plant species that reproduces entirely by vegetative reproduction, or a species that has a very narrow seed distribution. The only way for such plants to increase in frequency is to produce offspring into neighbouring sites as the individuals themselves cannot move (i.e. $m = 0$). Because all demographic processes are local, the dynamics of such plants have been studied using cellular automaton models (Crawley & May, 1987; Hendry & McGlade, 1995), but lend themselves also very well to a correlation dynamics approach (Harada & Iwasa, 1994).

The basic question we will address here is how much the plants should invest in vertical growth. The advantage of being tall is an increased amount of light capture, at the expense of neighbouring plants. The

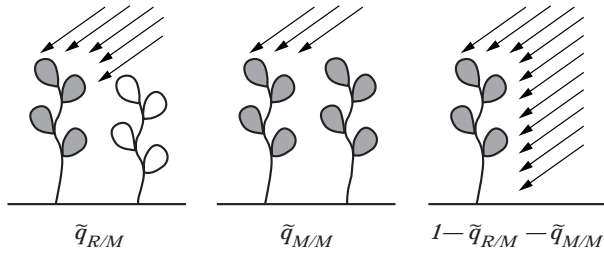


FIG. 3. Schematic representation of light capture by a mutant (shaded) plant. A neighbouring site can either be occupied by another mutant plant (shaded), by a resident plant (white) or be empty. Assuming the average amount of light captured is the average for the three configurations, it will depend on the height of mutant and resident plants but also on the average angle of the sun and the horizontal surface of the plants (assumed constant).

direct disadvantage is that the resources invested in growth cannot be used for seed production or ramet growth, but there is also the indirect disadvantage in that neighbouring relatives may be overshadowed.

Assume that after germination plants quickly grow to a (genetically determined) height h^* and then start producing seeds (or ramets) that disperse to neighbouring sites. Now consider a mutant that grows to a different size h . For simplicity we will assume the amount of light that is captured by a plant to be a constant that is offset by an amount proportional to the size difference with each of its neighbours (an empty site is counted as a plant of size 0, see Fig. 3). One may assume that the costs of being tall reduce seed production rate because resources must be allocated to structural growth and maintenance. Then, the seed production rate of a plant belonging to a cluster of mutants that grow to a height h in an environment dominated by plants of height h^* can be modeled as

$$b_M = B_0 + Lh(1 - \phi)\tilde{q}_{o|M} + L(h - h^*)(1 - \phi)\tilde{q}_{R|M} \quad (36)$$

where b_0 is base-line seed production, L is a proportionality constant that measures competition for light between two neighbouring plants (L will depend on the number of neighbouring sites, average angle of sun rays and so forth), and $K(h)$ measures the cost of maintaining size h .

Using $q_{o|M} = 1 - q_{R|M} - q_{M|M}$ we recover our basic model for continuous strategies, where

$$\begin{aligned} B(h) &= -Lh \\ C(h) &= K(h) - Lh \end{aligned} \quad (37)$$

and all of our results from the section on ESSs pertain to this plant model. The reason for the odd sign reversal is that “being altruistic” in this setting is the same as “growing less tall to benefit your neigh-

bours”. The only conceptual difference with the basic model is that individual costs incorporate a benefit because an altruistic plant saves on structural growth, which implies that there is an optimum height h_o , that minimizes the individual costs $C(h)$, given by

$$\frac{dC}{dh} = 0 \Leftrightarrow \frac{dK}{dh} = L \quad (38)$$

If the plant is surrounded by relatives, however, a plant can confer benefits to its neighbours—and thus to its relatives among them—by growing less tall. the optimum height h for a cluster of relatives should satisfy

$$\frac{dB}{dh}(1 - \phi)q_{M|M} - \frac{dC}{dh} = 0 \quad (39)$$

which, translated in terms of marginal values becomes

$$\frac{dK}{dh} = L(1 - (1 - \phi)\tilde{q}_{M|M}) \quad (40)$$

As the optimum is associated with a smaller marginal cost, the mutant plants should grow less tall. How much depends on the shape of the cost function, and the effect may be small.

The ESS depends on the cost function $K(h)$, the number of neighbouring sites n and the intensity of competition for light. Of these, the cost function will be most difficult to assess. A more serious flaw of this simple model is that it lacks phenotypic plasticity. A full model should incorporate this because in plants growth is strongly influenced by competition for light; seeds may not even be able to germinate on sites surrounded by tall neighbours. The model would then be more complex, but the same basic approach could be used to assess ESS allocation to growth and reproduction.

8. Discussion

We started analysing a model to study group selection in viscous populations, and we ended up with a number of variations on Hamilton’s Rule. If anything, this strengthens the notion that group selection and kin selection are two sides of the same coin (Grafen, 1984; Queller, 1994).

Traditionally, models for group selection assume a sharply subdivided environment. In many cases this is a reasonable assumption, for example when resources are patchily distributed. However, in many other cases subdivisions are not sharp, or even absent altogether. Just think of a continuous vegetation: although spatial structure may not be apparent, it is definitely not a “well-mixed” system: individuals do not move around, many plants reproduce (at least

partly) vegetatively and even seeds often do not disperse far. Hamilton already surmised that kin selection would occur in such systems as well. Limited dispersal leads to “viscous populations”, which means that individuals are likely to have relatives in their neighbourhood. Kin selection would then predict that altruism can evolve.

However, it turns out that it is not so simple. Limited dispersal indeed leads to clustering of relatives, but these relatives not only interact, they compete with each other as well. Altruists, for example, increase each other’s birth rate, but only to fill the same empty sites in the cluster’s neighbourhood. Thus, clustering of relatives in itself is not sufficient for kin selection to favor altruistic traits. For a rare mutant to invade successfully, it is essential that it is able to “export” its traits, i.e. a cluster of relatives should be able to grow and displace individuals belonging to the resident population (Wilson *et al.*, 1992). Here limited dispersal becomes a counteracting force, as mutants tend to sit in each other’s way. This is summarized in the statement that it is only when the “scale of dispersal” is larger than the “scale of regulation” (i.e. local competition) that altruists can invade (Kelly, 1992, 1994).

Taylor (1992a, b) argued that these scales are identical in purely viscous populations, and that altruism therefore cannot evolve in this setting. However, the picture is not as bleak as that. Using a correlation equation approach, Matsuda *et al.* (1992) and Harada *et al.* (1995) have shown that altruists may invade. Thus, in a viscous population, the scales of dispersal and interaction are not identical, even if they are very close. If the scales are close, spatial spread is a slow process, and on a finite lattice (as used in computer simulations) cluster growth may easily fail for stochastic reasons. Nonetheless, in the long run the altruists will invade, even if it requires many false starts.

In this article, we studied invasion dynamics in more detail, to show how the fitness of a rare mutant (its invasion exponent) is closely linked to the characteristic cluster structure of the mutant, given by the eigenvector of the mutant’s invasion matrix. Invasion in a viscous system implies growth of such clusters and, eventually, it is the properties of such clusters as coherent whole that determines whether or not a particular mutant will invade. Invasion analysis of the correlation equations leads to a natural definition of the “unit of selection” in viscous populations.

This is not to say that individuals are not important. Indeed, without individuality the whole process would not work, as selfish nano-individuals

would spread everywhere. Thus, the evolution of altruism requires that there is a “unit of population dynamics” as well, i.e. the individual. Recall that standard group selection also requires a unit of individuality: founders of populations have to be discrete individuals, because otherwise variation among groups will be lost (Goodnight, 1992). A condition for the evolution of altruism is then that the unit of selection should be larger than the unit of population dynamics.

From the unit of selection, the coefficient of relatedness can be calculated, and this turns out to be roughly proportional to the inverse of the number of neighbours (n). Thus for the altruists to invade, the total benefit to neighbours should be greater than n times the cost (Hamilton’s rule). This suggests that altruism does not easily evolve in viscous populations. If the altruists invade, it is because their higher net rate of reproduction rate leads to a higher local density than the non-altruists, which allows them to diffuse outward, into the domain dominated by the non-altruists. Incidentally, this points to the reason why altruists fail in Taylor’s (1992a) model: for mathematical convenience he set a parameter called “population elasticity” to unity. This implies that the lattice is completely filled, and that the altruists therefore cannot benefit from a higher local density (see also Kelly, 1994). Perhaps the most simple way to understand how this density effect works is to consider a harsh world where background birth rate is so low that non-altruists simply cannot maintain themselves on their own. Then *only* altruists can invade the empty world, due to the local Allee-effect that is generated by their mutual help.

8.1. LATTICE STRUCTURE

The results presented in this article are based on the assumption that the standard pair approximation holds. However, this approximation ignores some aspects that may be important. For example, one can see immediately that the cluster shown in Fig. 1 has $q_{A|oA} > 0$ (there are many oA -pairs where the o is neighbored by another A) whereas the standard pair approximation assumes $q_{A|oA} = 0$. This implies that there will be within-cluster competition for space, a phenomenon that the standard pair approximation ignores. At the same time, however, the average altruist has more altruists in its neighbourhood (and thus receives more help) than the standard pair approximation predicts (that is, the spatial segregation of altruists and non-altruists is more pronounced). Which of the two mechanisms will predominate is not a priori clear.

It should be realized, however, that the intensity of these mechanisms depends on the geometrical structure of the lattice. Most simulation studies assume regular, square two-dimensional lattices, with neighbourhood structures of either four or eight neighbours. However, it is also possible to construct random lattices, with the same number of neighbours per site but with a random structure. Then, one ends up with a probabilistic cellular automaton or artificial ecology in which the local rules are the same, but which will show different dynamics. Compare for example, the two $n = 3$ lattices in Fig. 4. In the regular lattice [Fig. 4(a)], neighbour's neighbours start overlapping only three links away, which will introduce correlations not accounted for in the standard pair approximation. In contrast, such overlap is not present in a random lattice, so there the standard pair approximation seems justified.

To a certain extent, lattice structure is incorporated in the following closure assumption:

$$q_{hij} \approx q_{hi}(1 - \theta + \theta\gamma_{hj})\tau_{hij} \quad (41)$$

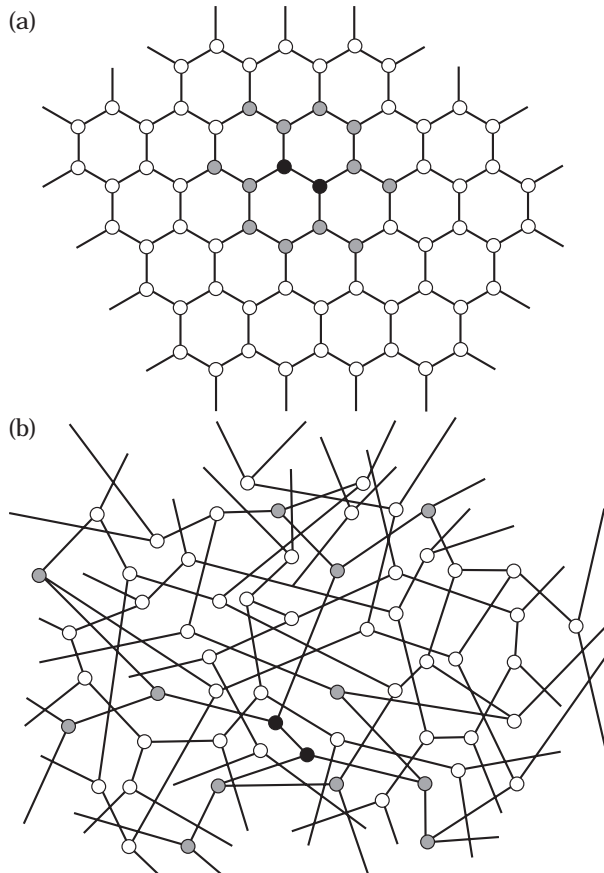


FIG. 4. Examples (a) of a regular $n = 3$ lattice, and (b) of a random $n = 3$ lattice. In both graphs, a central pair is indicated (black) with their neighbours up to two links away (grey).

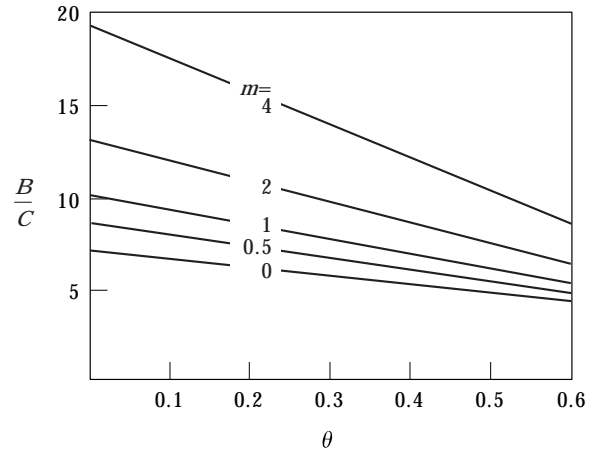


FIG. 5. Critical benefit/cost ratio (B/C) for the evolution of altruism as a function of lattice regularity θ (θ gives the proportion of triplets in triangular, as opposed to open, configuration). The critical ratio is shown for different migration rates $m (= m_s = m_A)$. Background birth rate $b_0 = 2$, mortality rate $d_s = d_A = 1$.

where $\gamma_{hj} = p_{hj}/(p_h p_j)$ denotes the correlation between i and j sites, and τ_{hij} is a correction factor to ensure that $\sum_h q_{hij} = 1$. This approximation basically assumes that a proportion θ of all triplets are in a closed, triangular configuration. Then, the j and the h may be direct neighbours, which introduces the extra correlation γ_{hj} . See Van Baalen (1998) for a more detailed discussion, and see Appendix C for an outline of how to calculate τ_{hij} . The parameter θ can be interpreted as measuring the regularity of the lattice: it gives an estimate of the overlap among i and j 's neighbours and this depends on the geometrical structure of the lattice. Notice that setting $\theta = 0$ leads to the uncorrected, standard pair approximation. Thus, the classical pair approximation is expected to work best for random lattices. This is borne out by comparing simulations with the improved approximation (see van Baalen, 1998).

Closure condition, eqn (41), makes the dynamical system more complex. We will not present a full analysis, but we will show some preliminary results. Given the equilibrium of the non-altruists (see Appendix C) the "relatedness coefficient" $(1 - \phi)q_{A|A_0}$ of an altruist close to the non-altruists (i.e. B and C are both small in absolute terms) can be quickly calculated). From this relatedness coefficient then follows the critical benefit/cost ratio for the evolution of altruism. As can be seen in Fig. 5, the critical benefit/cost ratio decreases when the lattice becomes more regular. Thus, the evolution of altruism is more easy on regular lattices than on random lattice. This suggests that the benefits of stronger clustering (more help) outweigh the disadvantages (within-cluster competition for space).

8.2. MULTITROPHIC SYSTEMS

In viscous populations altruism can evolve, but only if the benefits considerably outweigh the costs. This suggests that altruistic behavior would evolve only rarely, or, in more general terms, kin selection is not very intense in viscous populations.

The reason for this is that single populations do not exhibit very interesting spatial dynamics: they basically saturate the lattice (roughly analogous to the growth towards carrying capacity in the logistic population growth model) and do not produce large scale structures. Adding other trophic levels will change that. For example, spatial predator–prey models will give rise to more complex spatial patterns, ranging from chaotic structures to highly structured spiral patterns (Hassell *et al.*, 1991; Boerlijst *et al.*, 1993). Such large-scale structures may form higher-level units of selection, because it is the properties of such structures that may determine the outcome of natural selection. In the case of spiral hypercycles, for example, faster rotating spirals displace slower rotating ones (Boerlijst *et al.*, 1993). In host–parasite systems, spatial structure may limit the evolution of transmissibility: when a cluster of hosts succumbs before it mingles with neighbouring clusters, the parasites that “exploit” the cluster go extinct too (Rand *et al.*, 1995). The spatial dynamics of host–parasite systems may also provide an explanation for the evolution of sex: asexual hosts have an initial advantage because they do not pay the “two-fold cost”, but succumb in the longer term to parasites when their cluster grows too big (Keeling & Rand, 1995).

Using correlation dynamics models to derive the units of selection is a promising avenue of research to obtain more analytical insight into kin selection in multitrophic systems. For example, the unit of selection in a parasite population is affected by host traits like the host’s rate of reproduction, a parameter that will not affect parasite evolution in a well-mixed system (in absence of vertical transmission). Parasites that reduce their transmissibility in order to increase their host’s rate of reproduction may then invade a parasite population that sterilizes its hosts (M. van Baalen, unpublished results).

8.3. DISPERSAL

In this article, we assumed that dispersal rate was a given constant, the same for both altruists and non-altruists. Dispersal, however, is subject to natural selection as any other character. There is an important connection with the evolution of altruism: natural selection will favor non-altruists that disperse

as fast as possible, but for invading altruists there may be an optimal dispersal rate, as altruists depend on each other and must therefore “keep in touch”.

However, such an optimum dispersal rate is not an ESS. Would the altruists be the resident population there is no penalty associated with losing contact with relatives, and increased dispersal is selected for. But then, as the population becomes more and more well-mixed, less altruistic strategies are favored again. Therefore if altruism and dispersal are evolving simultaneously, the end result (ESS) may be a rapidly mixing non-altruistic strategy. However, then, there is scope for the altruists again: if the benefits of altruism are large enough, that is, if there exists a strategy s that satisfies

$$C(s) < B(s)(1 - \phi)\tilde{q}_{M|M} \quad (42)$$

a cluster of altruists that stay together may invade (recall that the coefficient of relatedness $\tilde{q}_{M|M}$ approximate $\phi = 1/n$ if $m_A = 0$).

Thus there may be also a second type of outcome, coexistence of rapidly moving non-altruists (m_S large), and sedentary altruists ($m_A = 0$). This is an interesting result, because strategy sets of simple continuous shapes usually lead to monomorphic ESSs. Here divergence into discrete types is made possible by the association of different expansion and interaction strategies. This suggests an interesting hypothesis for the evolution of multicellularity: multicellularity is not an inevitable consequence of “the quest for progress” but actually the result of a breakup of an original “slimy” ancestral population of unicellular organisms. In the end some unicellular lines specialized to divide and migrate as fast as they can, whereas others clung together to benefit from close cooperation and eventually evolved into multicellular organisms. The present analysis suggests that these benefits should be substantial.

Our research was greatly aided by our interaction with Andrew Morris. W. J. Edmunds, M. E. Hochberg, V. A. A. Jansen, C. de Mazancourt, M. W. Sabelis and P. Taylor are thanked for their valuable comments on an earlier version of this paper. This research was supported by the Applied Nonlinear Mathematics Initiative of the U.K. Engineering and Physical Science Research Council (EPSRC) and by the U.K. National Environmental Research Council (NERC)

REFERENCES

- BOERLIJST, M. C., LAMERS, M. & HOGEWEG, P. (1993). Evolutionary consequences of spiral patterns in a host-parasitoid system. *Proc. R. Soc. B* 253, 15–18.
- CHARNOV, E. L. (1976). Optimal foraging: the marginal value theorem. *Theor. Pop. Biol.* 9, 129–136.

- CRAWLEY, M. J. & MAY, R. M. (1987). Population dynamics and plant community structure: competition between annuals and perennials. *J. theor. Biol.* **125**, 475–489.
- DAWKINS, R. (1982). *The Extended Phenotype*. Oxford: Freeman.
- DURRETT, R. & LEVIN, S. (1994). The importance of being discrete (and spatial). *Theor. Pop. Biol.* **46**, 363–394.
- DURRETT, R. (1988). Ten Lectures on Particle Systems. Lecture notes.
- ESHEL, I. (1977). On the founder effect and the evolution of altruistic traits: an ecogenetical approach. *Theor. Pop. Biol.* **11**, 410–424.
- GOODNIGHT, K. F. (1992). The effect of stochastic variation on kin selection in a budding-viscous population. *Am. Nat.* **140**, 1028–1040.
- GRAFEN, A. (1984). Natural selection, group selection and kin selection. In: *Behavioural Ecology. An Evolutionary Approach* 2nd Edn (Krebs, J. R. & Davies, N. B., eds), pp. 62–84. Oxford: Blackwell.
- HAMILTON, W. D. (1963). The evolution of altruistic behaviour. *Am. Nat.* **97**, 354–356.
- HAMILTON, W. D. (1964). The genetical evolution of social behaviour. i and ii. *J. theor. Biol.* **7**, 1–16. 17–52.
- HARADA, Y. & IWASA, Y. (1964). Lattice population-dynamics for plants with dispersing seeds and vegetative propagation. *Res. Popul. Ecol.* **36**, 237–249.
- HARADA, Y., EZOE, H., IWASA, Y., MATSUDA, H. & SATŌ, K. (1995). Population persistence and spatially limited social interaction. *Theor. Pop. Biol.* **48**, 65–91.
- HASSELL, M. P., MAY, R. M., PACALA, S. W. & CHESSON, P. L. (1991). The persistence of host-parasitoid associations in patchy environments. I. A general criterion. *Am. Nat.* **138**, 568–583.
- HENDRY, R. J. & MCGLADE, J. M. (1995). The role of memory in ecological systems. *Proc. R. Soc. Lond. B* **259**, 153–159.
- KEELING, M. J. (1996). *The ecology and evolution of spatial host-parasite systems*. Ph.D. Thesis, University of Warwick, Coventry.
- KEELING, M. J. & RAND, D. A. (1995). A spatial mechanism for the evolution and maintenance of sexual reproduction. *Oikos* **74**, 414–424.
- KELLY, J. K. (1992). Restricted migration and the evolution of altruism. *Evolution* **46**, 1492–1495.
- KELLY, J. K. (1994). The effect of scale-dependent processes on kin selection-mating and density regulation. *Theor. Pop. Biol.* **46**, 32–57.
- LEVIN, S. & PIMENTEL, D. (1981). Selection of intermediate rates of increase in parasite–host systems. *Am. Nat.* **117**, 308–315.
- MATSUDA, H., OGITA, N., SASAKI, A. & SATŌ, K. (1992). Statistical mechanics of population: the lattice Lotka–Volterra model. *Prog. Theor. Phys.* **88**, 1035–1049.
- MAYNARD SMITH, J. (1964). Group selection and kin selection. *Nature* **201**, 1145–1147.
- MAYNARD SMITH, J. & PRICE, G. R. (1983). The logic of animal conflict. *Nature* **246**, 15–18.
- MAYNARD SMITH, J. & SZATHMÁRY, E. (1995). *The Major Transitions in Evolution*. Oxford, Freeman.
- METZ, J. A. J., NISBET, R. M. & GERITZ, S. A. H. (1992). How should we define fitness for general ecological scenarios. *TREE* **7**, 198–202.
- MORRIS, A. (1997). *Representing spatial interactions in simple ecological models*. Ph.D. Thesis, University of Warwick, Coventry.
- NAKAMARU, M., MATSUDA, H. & IWASA, Y. (1997). The evolution of cooperation in a lattice-structured population. *J. Theor. Biol.* **184**, 65–81.
- NUNNEY, L. (1985). Group selection, altruism, and structured-deme models. *Am. Nat.* **126**, 212–230.
- QUELLER, D. C. (1994). Genetic relatedness in viscous populations. *Evol. Ecol.* **8**, 70–73.
- RAND, D. A., KEELING, M. & WILSON, H. B. (1995). Invasion, stability and evolution to criticality in spatially extended, artificial host–pathogen ecologies. *Proc. R. Soc. Lond. B* **259**, 55–63.
- RAND, D. A., WILSON, H. B. & MCGLADE, J. M. (1994). Dynamics and evolution: evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Phil. Trans. R. Soc. Lond. B* **343**, 261–283.
- SATŌ, K., MATSUDA, H. & SASAKI, A. (1994). Pathogen invasion and host extinction in lattice structured populations. *J. Math. Biol.* **32**, 251–268.
- TAYLOR, P. D. (1992a). Altruism in viscous populations an inclusive fitness model. *Evol. Ecol.* **6**, 352–356.
- TAYLOR, P. D. (1992b). Inclusive fitness in a homogeneous environment. *Proc. R. Soc. Lond. B* **249**, 299–302.
- VAN BAALEN, M. (1998). Pair dynamics and configuration approximations. In: *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. (Metz, J. A. J. & Dieckman, U., eds) Cambridge, Cambridge University Press.
- VAN BAALEN, M. & SABELIS, M. N. (1995). The dynamics of multiple infection and the evolution of virulence. *Am. Nat.* **146**, 881–910.
- WILSON, D. S., POLLOCK, G. B. & DUGATKIN, L. A. (1992). Can altruism evolve in a purely viscous population? *Evol. Ecol.* **6**, 331–341.
- WILSON, D. S. (1977). Structured demes and the evolution of group advantageous traits. *Am. Nat.* **111**, 157–185.

APPENDIX A

The Full System

The full system of equations reads

$$\begin{aligned}
 \frac{dp_{S_o}}{dt} &= (b_S + m_S)\phi q_{S|o}p_{o_o} \\
 &- [\phi b_S + \bar{\phi}(b_S + m_S)q_{S|o_S} + \bar{\phi}(b_A + m_A)q_{A|o_S} + d_S \\
 &\quad - \bar{\phi}m_S q_{o|S_o}]p_{S_o} \\
 &\quad + [d_S + \bar{\phi}m_S q_{o|SS}]p_{SS} \\
 &\quad + [d_A + \bar{\phi}m_A q_{o|AS}]p_{SA} \\
 \frac{dp_{SS}}{dt} &= 2[\phi b_S + \bar{\phi}(b_S + m_S)q_{S|o_S}]p_{S_o} \\
 &\quad - 2[d_S + m_S \bar{\phi} q_{o|SS}]p_{SS} \\
 \frac{dp_{A_o}}{dt} &= (b_A + m_A)\bar{\phi} q_{A|o}p_{o_o} \quad (\text{A.1}) \\
 &- [\phi b_A + \bar{\phi}(b_A + m_A)q_{A|o_A} + \bar{\phi}(b_S + m_S)q_{S|o_A} + d_A \\
 &\quad + \bar{\phi}m_A q_{o|A_o}]p_{A_o} \\
 &\quad + [d_A + \bar{\phi}m_A q_{o|AA}]p_{AA} \\
 &\quad + [d_S + \bar{\phi}m_S q_{o|SA}]p_{SA} \\
 \frac{dp_{AA}}{dt} &= 2[\phi b_A + \bar{\phi}(b_A + m_A)q_{A|o_A}]p_{A_o} \\
 &\quad - 2[d_A + \bar{\phi}m_A q_{o|AA}]p_{AA} \\
 \frac{dp_{AS}}{dt} &= (b_S + m_S)\bar{\phi} q_{S|o}p_{A_o} \\
 &\quad + (b_A + m_A)\bar{\phi} q_{A|o_S}p_{S_o} \\
 &\quad - [d_S + \bar{\phi}m_S q_{o|SA} + d_A + \bar{\phi}m_A q_{o|AS}]p_{SA}
 \end{aligned}$$

where $\bar{\phi} = 1 - \phi = (n - 1)/n$. Symmetry relations imply $p_{ij} = p_{ji}$, and the fact that all frequencies sum to one implies

$$p_{oo} = 1 - 2(p_{So} + p_{Ao} + p_{AS}) - p_{SS} - p_{AA} \quad (\text{A.2})$$

It should be kept in mind that the events of birth, death and migration may be functions that depend on the individuals' environment. The vital rates b_i, d_i and m_i must therefore be averaged over the environments of all the individuals of type i . Since we assume that death and migration rates are environment-independent, these average rates are equal to the event rates d_i and m_i . Birth events do depend on the individual's environment, however, and care must be taken that these averages are taken properly (Morris, 1997),

$$b_i = \langle b_S(x) \rangle_i = b_0 + B \langle \frac{n_A(x)}{n} \rangle_i - C_i \quad (\text{A.3})$$

(where $\langle f(x) \rangle_i$ denotes the mean of $f(x)$ over all sites x occupied by type i), which thus depends on the mean proportion of altruists in the neighbourhood of type i . This procedure leads to the following result. An individual of type i has on average $nq_{A|i}$ altruist neighbours, and therefore receives an expected amount of benefit $(B/n)(nq_{A|i}) = Bq_{A|i}$.

However, eqn (A.3) is not yet the correct expression to substitute in the differential equations. These rates depend on the mean environment of individuals that actually reproduce (that is, on the environment of the i where $io \rightarrow ii$ events occur); such individuals therefore have at least one empty site among their neighbours, which reduces the potential benefit they receive. Averaging over all io pairs (instead of over all i sites, see van Baalen, 1998) leads to the better estimation

$$\begin{aligned} b_S &= b_0 + B(1 - \phi)q_{A|So} \\ b_A &= b_0 + B(1 - \phi)q_{A|Ao} - C \end{aligned} \quad (\text{A.4})$$

where the factor $1 - \phi = (n - 1)/n$ takes into account the fact that at least one of its neighbouring sites is empty. These values are to be substituted into the differential equations for the p_{ij} .

The structure of the system of equations becomes more clear if one defines for convenience

$$\begin{aligned} \alpha_i &= \bar{\phi}(b_i + m_i) \\ \beta_i &= \phi b_i + \bar{\phi}(b_i + m_i)q_{i|oi} \\ \delta_{ij} &= d_i + \bar{\phi}m_o q_{o|ij} \end{aligned} \quad (\text{A.5})$$

We then can write the system of equations as

$$\begin{aligned} \frac{dp_{So}}{dt} &= \alpha_S q_{S|oo} p_{oo} \\ &- [\beta_S + \alpha_A q_{A|oS} + \delta_{So}] p_{So} \\ &+ \delta_{SS} p_{SS} \\ &+ \delta_{AS} p_{SA} \\ \frac{dp_{SS}}{dt} &= 2\beta_S p_{So} - 2\delta_{SS} p_{SS} \\ \frac{dp_{Ao}}{dt} &= \alpha_A q_{A|oo} p_{oo} \\ &- [\beta_A + \alpha_S q_{S|oA} + \delta_{Ao}] p_{Ao} \\ &+ \delta_{AA} p_{AA} \\ &+ \delta_{SA} p_{SA} \\ \frac{dp_{AA}}{dt} &= 2\beta_A p_{oA} - 2\delta_{AA} p_{AA} \\ \frac{dp_{AS}}{dt} &= \alpha_S q_{S|oA} p_{Ao} + \alpha_A q_{A|oS} p_{So} \\ &- [\delta_{SA} + \delta_{AS}] p_{AS} \end{aligned} \quad (\text{A.6})$$

The system of equations depends on conditional probabilities involving triplets ($q_{i|ab}$). As we do not want to extend the analysis to the dynamics of triplets, we have to "close" the system, by expressing the $q_{i|ab}$ in terms of pair frequencies. In the main text of the article, we assume the standard pair approximation

$$q_{i|ab} \approx q_{i|a} \quad (\text{A.7})$$

In Appendix C we briefly discuss a more elaborate closure assumption.

APPENDIX B

Invasion

When resident S is at equilibrium, \bar{p}_{So} and \bar{p}_{SS} (and hence \bar{p}_{oo}) are given and constant in time, as are all $q_{h|i}$ and $q_{h|ij}$ with $h, i, j \in \{o, S\}$.

A (globally) rare population of altruists ($p_{Ao}, p_{AS}, p_{AA} \ll p_{So}, p_{SS}$) will not affect the resident dynamics. Effectively, this allows us to decouple the mutant's invasion dynamics from the resident's dynamics. In contrast to a well-mixed system, in a viscous system a mutant's invasion dynamics is governed by more than a single equation. Since we limit ourselves to derive equations for pairs, the mutant's invasion

dynamics is governed by three equations, for p_{Ao} , p_{AS} , p_{AA} (by definition, $p_{oA} = p_{Ao}$).

Using $q_{A|oA}p_{oo} = q_{o|oA}p_{oA}$, $q_{A|oS}p_{oS} = q_{S|oA}p_{oA}$, and $q_{o|oA} = 1 - q_{S|oA} - q_{A|oA}$, the differential equations for p_{Ao} , p_{AS} and p_{AA} can be written in matrix form,

$$\frac{d}{dt} \begin{pmatrix} p_{Ao} \\ p_{AS} \\ p_{AA} \end{pmatrix} = \mathbf{M} \begin{pmatrix} p_{Ao} \\ p_{AS} \\ p_{AA} \end{pmatrix} \quad (\text{B.1})$$

with

$$\mathbf{M} = \begin{pmatrix} \alpha_A(1 - q_{A|oA}) - \beta_A - (\alpha_S + \alpha_A)q_{S|oA} - \delta_{Ao} & & \\ (\alpha_S + \alpha_A)q_{S|oA} & & \\ 2\beta_A & & \\ & \delta_{SA} & \delta_{AA} \\ -\delta_{SA} - \delta_{AS} & & 0 \\ 0 & & -2\delta_{AA} \end{pmatrix} \quad (\text{B.2})$$

This matrix can be entirely expressed in terms of local densities $q_{i|A}$, even with the improved approximation proposed in Appendix C. Here, however, we will derive the invasion condition assuming the standard approximation is valid. The same approach can also be applied to improved approximations, but then it becomes more laborious.

The invasion exponent and the unit of selection have to be calculated simultaneously. As explained in the text, the most compact way to represent this set of simultaneous equations is

$$\mathbf{M}(\mathbf{q}_A)\mathbf{q}_A = \lambda\mathbf{q}_A \quad (\text{B.3})$$

From the invasion matrix, the normalized eigenvector that is associated with an eigenvector $\lambda = 0$ should satisfy

$$\begin{pmatrix} \bar{q}_{o|A} \\ \bar{q}_{S|A} \\ \bar{q}_{A|A} \end{pmatrix} = \frac{1}{K} \begin{pmatrix} \delta_{AA} \\ (\alpha_S + \alpha_A) \frac{\delta_{AA}}{\delta_{AS} + \delta_{SA}} q_{S|oA} \\ \beta_A \end{pmatrix} \quad (\text{B.4})$$

with

$$K = \delta_{AA} + (\alpha_S + \alpha_A) \frac{\delta_{AA}}{\delta_{AS} + \delta_{SA}} q_{S|oA} + \beta_A \quad (\text{B.5})$$

In the general case, this is a complicated set of equations, because α_i , β_i and δ_{ij} may all depend on local densities $q_{h|ij}$, which all have to be expressed in terms of the q_{hi} .

In principle, we can work everything out using improved pair approximations, such as proposed in Appendix C. However, here we will work out the

invasion condition assuming that the standard pair approximation holds. Then, $\delta_{ij} = \delta_i$, which simplifies the analysis considerably. With improved approximations, the analysis is similar but more laborious.

Though it should be possible to calculate fitness (λ) and unit of selection (\mathbf{q}_A) for any arbitrary mutant, explicit solutions, if obtainable at all, tend to become rather messy. If the demographic rates are density-independent constants, the characteristic equation that must be solved is cubic, and more complex cases (such as the one considered here, where birth rates depend on $q_{A|A}$) may require solution of 4-th order equations or higher. Invasion boundaries, manifolds in parameter space defined by $\lambda = 0$ are easier to analyse.

The invasion condition $\lambda = 0$ implies that the invasion matrix is singular ($|\mathbf{M}(\mathbf{q}_A)| = 0$), which leads to the following expression describing the ‘‘invasion boundary’’

$$(\alpha_A - \delta_A)(\delta_A + \delta_S) - \delta_A(\alpha_A + \alpha_S)\bar{q}_{S|o} = 0 \quad (\text{B.6})$$

Because we can define a ‘‘mutant’’ that is identical to the resident (*i.e.* a degenerate ‘‘altruist’’ characterized by $B = 0$ and $C = 0$), we can use this expression to solve the equilibrium of the resident. For such a mutant A^* we know that

$$\alpha_{A^*} = \bar{\phi}(b_{A^*} + m_{A^*}) = \bar{\phi}(b_S + m_S) = \alpha_S \quad (\text{B.7})$$

and (since $q_{o|A^*} = q_{o|S}$ as we will show below),

$$\delta_{A^*} = d_{A^*} + \bar{\phi}m_{A^*}\bar{q}_{o|A^*} = d_S + \bar{\phi}m_S\bar{q}_{o|S} = \delta_S \quad (\text{B.8})$$

This mutant will have an invasion exponent of exactly zero and should therefore be on the invasion boundary determined by eqn (B.6). This can be used to solve $\bar{q}_{S|o}$, yielding

$$\bar{q}_{S|o} = 1 - \frac{\delta_S}{\alpha_S} \quad (\text{B.9})$$

which can then be substituted back into eqn (B.6). (Notice that this equation gives us also $\bar{q}_{o|o} = 1 - \bar{q}_{S|o}$.) This gives the following characterization of the invasion boundary

$$(\alpha_A - \delta_A)(\delta_A + \delta_S) - \delta_A(\alpha_A + \alpha_S)\left(1 - \frac{\delta_S}{\alpha_S}\right) = 0 \quad (\text{B.10})$$

which can be simplified into

$$(\alpha_A + \alpha_S)(\alpha_S\delta_S - \delta_A\delta_S) = 0 \quad (\text{B.11})$$

This leads to the ‘‘birth–death’’ ratio condition eqn (17) that is discussed in the main text.

To calculate α_A and δ_A , we have to know its local densities $q_{i|A}$. Under the pair approximation, eqn (B.4) becomes

$$\begin{pmatrix} \tilde{q}_{o|A} \\ \tilde{q}_{s|A} \\ \tilde{q}_{A|A} \end{pmatrix} = \frac{1}{K} \begin{pmatrix} \delta_A \\ (\alpha_A + \alpha_S) \frac{\delta_A}{\delta_A + \delta_S} \tilde{q}_{s|o} \\ \beta_A \end{pmatrix} \quad (\text{B.12})$$

where K again is a normalizing factor. For an altruist that is close to the resident (B and C small in absolute terms) we obtain

$$\begin{aligned} \tilde{q}_{o|A} &= \frac{\delta_S}{\alpha_S + \beta_S} \\ \tilde{q}_{s|A} &= \frac{\alpha_S - \delta_S}{\alpha_S + \beta_S} \\ \tilde{q}_{A|A} &= \frac{\beta_S}{\alpha_S + \beta_S} \end{aligned} \quad (\text{B.13})$$

Note that this is in agreement with the assertion that for the degenerate altruist $\bar{q}_{o|A^*} = \bar{q}_{o|S}$, because

$$\begin{aligned} \tilde{q}_{o|S} &= \tilde{q}_{o|A^*} \Leftrightarrow \\ \tilde{q}_{o|S} &= \frac{\delta_S}{\alpha_S + \beta_S} = \frac{d_S + \bar{\phi} m_S \tilde{q}_{o|S}}{\bar{\phi}(b_S + m_S) + \phi b_S} \Leftrightarrow \\ \tilde{q}_{o|S} &= \frac{d_S}{b_S} \end{aligned} \quad (\text{B.14})$$

which is the equilibrium condition derived earlier.

APPENDIX C

Correction for Regular Lattices

In this appendix, a very short outline of the improved pair approximation [eqn (41)] will be given. A more extensive discussion can be found in van Baalen (1998). If, a fraction θ of the triplets is in a closed, triangular configuration, we can write

$$q_{hij} = q_{hi}(1 - \theta)T_{\triangle hij} + \theta\gamma_{hj}T_{\triangle hij} \quad (\text{C.1})$$

where $\gamma_{hj} = p_{hj}/p_h p_j$ stands for the pair correlation, and $T_{\triangle hij}$ and $T_{\triangle hij}$ stand for the triple correlations of open and closed triangles, respectively. The triple correlations are defined by $p_{\triangle hij} = p_h p_i p_j \gamma_{hi} \gamma_{ij} T_{\triangle hij}$ and $p_{\triangle hij} = p_h p_i p_j \gamma_{hi} \gamma_{ij} \gamma_{ih} T_{\triangle hij}$ (in other words triple correlations are the corrections one needs to apply to estimates based to pairs to obtain the ‘‘right’’ values.)

The pair correlations can be calculated from the pair and singlet frequencies, but the triple correlations are essentially unknown. Closing the system thus boils down to assuming a value for these triple correlations. Assuming no triangular triplets (corresponding to a random lattice) and the absence of triple correlations

(*i.e.* $T_{\triangle hij} = 1$) leads to the standard pair approximation. A more elaborate approximation is obtained if it is assumed that both types of triple correlation are equal (*i.e.* $T_{\triangle hij} = T_{\triangle hij}$) and estimated by τ_{hij} . Then we obtain a series of expressions

$$q_{hij} \approx q_{hi}(1 - \theta + \theta\gamma_{hj})\tau_{hij} \quad (\text{C.2})$$

Here, τ_{hij} cannot be simply set to unity, because this causes the q_{hij} to violate the consistency condition

$$\sum_{h=o,A,S} q_{hij} = 1 \quad (\text{C.3})$$

A possible solution to this consistency condition is

$$\tau_{hij} = \begin{cases} 1 & \text{if } h \neq j \\ \frac{1}{q_{jii}} (1 - \sum_{h \neq j} q_{hij} ((1 - \theta) + \theta\gamma_{hj})) & \text{if } h = j \end{cases} \quad (\text{C.4})$$

After substitution of these conditional probabilities into the differential equations for pairs, the equilibrium of the resident non-altruists and the invasion conditions can be solved using a symbolic mathematics package like Mathematica. This is how we produced Fig. 5.

APPENDIX D

ESS Conditions

In this appendix we will derive the invasion exponent $\lambda_{s^*}(s)$ of a rare mutant (M) with strategy s that does not differ much from the resident strategy s^* . [In our case, $s = (B, C)$.] The resident population (R) with strategy s^* is assumed to be at equilibrium.

The invasion exponent is the dominant eigenvalue of the matrix

$$\mathbf{M} = \begin{pmatrix} \alpha_M - \beta_M - (\alpha_R + \alpha_M)q_{R|o} - \delta_M & & \\ (\alpha_R + \alpha_M)q_{R|o} & \delta_R & \delta_M \\ 2\beta_M & -\delta_R - \delta_M & 0 \\ & 0 & -2\delta_M \end{pmatrix} \quad (\text{D.1})$$

Right and left eigenvalues of this matrix are

$$\tilde{\mathbf{v}} = \begin{pmatrix} \delta_M + \frac{1}{2}\lambda \\ \frac{\alpha_R + \alpha_M}{\alpha_R} \frac{\delta_M + \frac{1}{2}\lambda}{\delta_R + \delta_M + \lambda} (\alpha_R - \delta_R) \\ \beta_M \end{pmatrix} \quad (\text{D.2})$$

and

$$\tilde{\mathbf{u}} = \begin{pmatrix} 1 & \frac{\delta_R}{\delta_R + \delta_M + \lambda} & \frac{\delta_M}{2\delta_M + \lambda} \end{pmatrix} \quad (\text{D.3})$$

where $\lambda = \lambda_{s^*}(s)$ so that

$$\lambda_{s^*}(s) = \frac{\tilde{\mathbf{v}}\tilde{\mathbf{M}}\tilde{\mathbf{u}}}{\tilde{\mathbf{v}}\tilde{\mathbf{u}}} \quad (\text{D.4})$$

Now consider a mutant whose strategy differs little from the residents' strategy; the resulting changes in pair-related rates, $\Delta\alpha$, $\Delta\beta$ and $\Delta\delta$, will therefore be small. We can decompose the matrix and its eigenvectors in a constant part (depending only on the resident strategy), a part that depends only linearly on the changes caused by the mutant's strategy change, and higher order terms (which we will subsequently ignore):

$$\begin{aligned} \tilde{\mathbf{M}} &= \mathbf{M}^* + \Delta\mathbf{M} \\ \tilde{\mathbf{u}} &= \mathbf{u}^* + \Delta\mathbf{u} + \text{h.o.t} \\ \tilde{\mathbf{v}} &= \mathbf{v}^* + \Delta\mathbf{v} + \text{h.o.t} \end{aligned} \quad (\text{D.5})$$

Some algebra shows that

$$\mathbf{u}^* = \begin{pmatrix} \delta_R \\ \alpha_R - \delta_R \\ \beta_R \end{pmatrix}$$

and

$$\Delta\mathbf{u} = \begin{pmatrix} \Delta\delta \\ \left(\frac{\Delta\alpha}{\alpha_R} + \frac{\Delta\delta}{\delta_R}\right)(\alpha_R - \delta_R) \\ \Delta\beta \end{pmatrix} \quad (\text{D.6})$$

and

$$\mathbf{v}^* = (2 \ 1 \ 1) \text{ and } \Delta\mathbf{v} = (0 \ -\frac{\Delta\delta}{2\delta_R} \ 0) \quad (\text{D.7})$$

so that $\Delta\mathbf{u}$ and $\Delta\mathbf{v}$ vanish if $\Delta\alpha$, $\Delta\beta$ and $\Delta\delta$ go to zero. If we expand the expression for the invasion exponent while keeping only the linear terms, we obtain

$$\begin{aligned} \lambda_{s^*}(s^* + \Delta s) &\approx \frac{\mathbf{v}^*\mathbf{M}^*\mathbf{u}^*}{\mathbf{v}^*\mathbf{u}^*} + \frac{\mathbf{v}^*\Delta\mathbf{M}\mathbf{u}^*}{\mathbf{v}^*\mathbf{u}^*} \\ &\quad + \frac{\Delta\mathbf{v}\mathbf{M}^*\mathbf{u}^*}{\mathbf{v}^*\mathbf{u}^*} + \frac{\mathbf{v}^*\mathbf{M}^*\Delta\mathbf{u}}{\mathbf{v}^*\mathbf{u}^*} \end{aligned} \quad (\text{D.8})$$

which simplifies to

$$\lambda_{s^*}(s^* + \Delta s) \approx \frac{\mathbf{v}^*\Delta\mathbf{M}\mathbf{u}^*}{\mathbf{v}^*\mathbf{u}^*} \quad (\text{D.9})$$

because $\mathbf{v}^*\mathbf{M}^* = \mathbf{0}^T$ and $\mathbf{M}^*\mathbf{u}^* = \mathbf{0}$. Some more algebra then leads to

$$\lambda_{s^*}(s^* + \Delta s) \approx \frac{\alpha_R + \delta_R}{\delta_R + \beta_R + \delta_R} \left(\frac{\delta_R}{\alpha_R} \Delta\alpha - \Delta\delta \right) \quad (\text{D.10})$$

This gives us the actual rate of invasion (or extinction of the mutant, where factor in front of the "selection differential" represents the slowdown caused by spatial expansion.