doi:10.1111/j.1558-5646.2010.01162.x

INCLUSIVE FITNESS ANALYSIS ON MATHEMATICAL GROUPS

Peter Taylor,[1,2] Timothy Lillicrap,[3,4] and Daniel Cownden[1,5]

[1]*Dept Math and Stats*

[2]*E-mail: taylor@queensu.ca*

[3]*Centre for Neuroscience, Queen's University, Kingston, ON, K7L 3N6, Canada*

[4]*E-mail: tim@biomed.queensu.ca*

[5]*E-mail: dcownden@gmail.com*

Received July 2, 2010
Accepted September 17, 2010

Recent work on the evolution of behaviour is set in a structured population, providing a systematic way to describe gene flow and behavioural interactions. To obtain analytical results one needs a structure with considerable regularity. Our results apply to such "homogeneous" structures (e.g., lattices, cycles, and island models). This regularity has been formally described by a "node-transitivity" condition but in mathematics, such internal symmetry is powerfully described by the theory of mathematical groups. Here, this theory provides elegant direct arguments for a more general version of a number of existing results. Our main result is that in large "group-structured" populations, primary fitness effects on others play no role in the evolution of the behaviour. The competitive effects of such a trait cancel the primary effects, and the inclusive fitness effect is given by the direct effect of the actor on its own fitness. This result is conditional on a number of assumptions such as (1) whether generations overlap, (2) whether offspring dispersal is symmetric, (3) whether the trait affects fecundity or survival, and (4) whether the underlying group is abelian. We formulate a number of results of this type in finite and infinite populations for both Moran and Wright–Fisher demographies.

KEY WORDS: **Allele frequency, group theory, homogeneous, population structure, relatedness, selection.**

A standard approach for the study of the selective effects of a social trait is provided by Hamilton's (1964) inclusive fitness effect, measured as the sum of the fitness effects of a behavioural deviation, each effect weighted by the relatedness of the actor to the recipient. It is well understood that this sum must include "all" effects, the "primary" effects that follow a direct interaction, for example on fecundity, as well as the resulting "secondary" competitive effects, for example, on mortality arising from changes in fecundity. At some level, this was well understood right from the beginning, but we suspect that the significance of the secondary effects was underestimated until more systematic studies of structured populations were undertaken. Wilson et al. (1992) in a Monte Carlo study of a large two-dimensional lattice population with limited dispersal, made the surprising discovery that an allele for altruistic behaviour to a neighbor declined in frequency no matter how great was the benefit b. As there was significant relatedness R between neighbors, this seemed to contradict Hamilton's (1964) rule that this allele should increase in frequency whenever $Rb > c$ where c is the cost of the altruistic act. This observation led to the an analysis of Taylor (1992a, 1992b) showing that in both an infinite island model and a one-dimensional lattice model the conferred benefit b of the altruistic act would be exactly cancelled by secondary competitive effects removing b completely from the inclusive fitness effect of the action. This effect was generalized by Queller's (1994) concept of "economic neighborhood," which distinguished the "scale of competition" from the "scale of interaction." The idea is that the former must be larger than the latter for altruism to be favored, and in Taylor's (1992a,b) examples, they were precisely the same.

From that time forward, a large body of literature grew up around the study of cooperation and altruism in structured populations, both theoretical and experimental. On the theoretical side, there was much work on evolutionary games on graphs, for example extensive work with prisoner's dilemma (Nowak and May 1992; Nakamaru et al. 1997, 1998; Ohtsuki and Nowak 2006b and Ohtsuki et al. 2007, the latter providing a good summary of the area). On the experimental side, we point to a pair of studies of cooperation in a structured population of the iron-scavenging bacteria *Pseudomonas aeruginosa.* Cooperators in this population incur a cost to produce siderophores, agents that release iron held in the environment, making it available for general use. Griffin et al. (2004) and Kummerli et al. (2009) manipulated the level of clumping and of dispersal to provide a range of values of both local relatedness and scale of competition, and found that cooperation was most favored when the latter was relatively small and the former was relatively large. The latter paper explicitly discussed "budding," a form of clumped dispersal, which is of interest to those who study homogeneous populations in that the given structure (rules for dispersal and interaction) appear to be the same for every individual, but no "group" structure can be found and altruism can in fact be favored (Gardner and West 2006).

However, it was some time until it was realized (Taylor et al. 2007b; Grafen and Archetti 2008) that the critical general assumption that made these scales the same was that of a homogeneous population structure. Our purpose here is to reformulate and extend these results. Our main result (Theorem 1) gives precise technical conditions on a homogeneous population so that the primary and secondary fitness effects of a social behaviour all sum to zero leaving us with only the fitness effects of the actor on herself. This result applies in both a finite and infinite population, and in continuous overlapping (Moran model) and discrete nonoverlapping (Wright–Fisher) generations models, although in both cases there are technical conditions which apply.

DEMOGRAPHY

We structure the population as a set of breeding sites, each occupied by a single asexual breeder. Diagrammatically, we represent the sites as the nodes of a graph and use the arc from node j to node k to represent the relationship of j to k. We let the breeder at site j have fecundity F_j and survival S_j. We work with two standard demographics, Wright–Fisher and Moran. The Wright–Fisher model has discrete nonoverlapping generations. The breeder at site j produces a large number F_j of offspring and then dies. Thus all survivals S_j are zero. Offspring born at site j disperse to compete at site k with probability $d(j, k)$, so that $\sum_k d(j, k) = 1$. At each site, the offspring then compete fairly for the breeding spot, and a new generation begins. In the Moran model, generations overlap and for any breeder, reproduction and death are events distributed in continuous time. We work with two standard protocols for this model BD (birth–death) and DB (death–birth) (Ohtsuki and Nowak 2006a). Under BD, a birth occurs at site j at rate F_j and the offspring replaces the breeder at site k with relative probability $d(j, k)(1 - S_k)$, the product of the dispersal probability and the node k mortality rate. Under DB, a death occurs at node k at rate given by breeder mortality $1 - S_k$ and the colonizing offspring comes from node j with relative probability $F_j\, d(j, k)$. Our main result will assume that dispersal is symmetric $[d(j, k) = d(k, j)]$ for the Moran process but not for Wright–Fisher.

INCLUSIVE FITNESS

We assume that the behaviour we are studying is genetically determined; thus the "effects" we are ultimately interested in capturing are those on allele frequency. A standard approach is to take a "resident" allele determining the behaviour we are interested in studying and look at the fitness of a "deviant" alternative allele defined for this purpose as the initial direction of frequency change of this allele—where "initial" here refers to the activation of the deviant allele from its neutral (resident) state. Under certain standard assumptions (of additive gene action and small effects of individual genes on fitness), it is known (Rousset and Billiard 2000; Taylor et al. 2007a) that this measure of fitness is captured (in sign) by Hamilton's (1964) inclusive fitness effect that can be written as

$$W_I = \sum_i w_i R_i. \tag{1}$$

Here, we assume that a single breeder (the actor) expresses the deviant behaviour and we take w_i to be the overall fitness effect of this deviation on the breeder at site i. Finally, R_i is the relatedness of the actor to breeder i (Michod and Hamilton 1980).

These overall fitness effects w_i incorporate both the primary and the secondary effects, but in a structured population, it makes little sense to specify the latter up front. It is more reasonable (and more useful) to specify the primary effects of the behavioural interactions (e.g., on fecundity or on survival), and let the secondary competitive effects be determined by the population structure. Thus our main result (Theorem 1) will be formulated in terms of the primary effects.

As an example and preview of Theorem 1, suppose that an actor in a population structured as a group incurs personal cost c to provide a total primary fitness increment b to one or more others (different from the actor), where b and c might measure effects on fecundity or on survival. Then, under certain conditions, the inclusive fitness effect W_I has the same sign as $-c$ in an infinite population and as $-c - b/(N - 1)$ in a finite population of size N. When the primary fitness effects are on fecundity (resp survival), this result applies in the Moran model with the BD

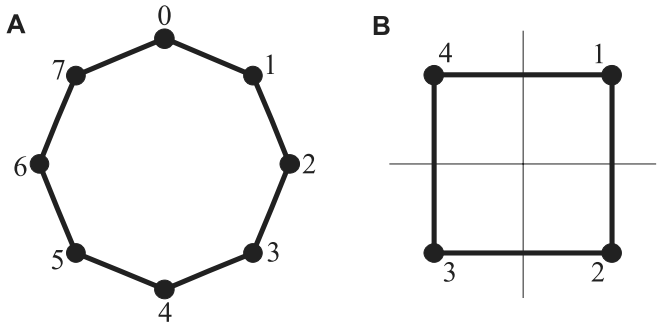

Figure 1. Two groups of order 8. The diagram of (A) illustrates the group of all rotations that preserve the octagon under composition. There are eight of these. They can also be represented as the set of integers n with $0 \le n \le 7$ under addition "mod 8," that is, "casting out 8." For example, $3 + 7 = 2$. This group is abelian. The diagram of (B) illustrates the group of all symmetries of the square—the permutations of the vertices which preserve the square, again under composition. One way to see that there are eight of these is to note that 1 can be mapped to any corner, k (4 choices), then 2 must be mapped to a corner adjacent to k (2 choices), and then the images of 3 and 4 are determined. This group is nonabelian. For example, a rotation about 90° followed by reflection in the vertical axis maps 1 into itself, whereas in the reverse order, 1 is mapped into 3.

(resp. DB) protocol with the proviso that dispersal is symmetric and the group is abelian (multiplication is commutative—discussed later). It also applies in the Wright–Fisher model with fecundity selection, and in this case dispersal is allowed to be asymmetric. We will comment later on some of the interesting exceptions, such as the Moran process with asymmetric dispersal, and the Wright–Fisher process with partial survival.

A Group-Structured Population

Our homogeneity assumption, roughly speaking, is that the population "look the same" in terms of dispersal rates and interaction effects when viewed from any breeding site. Here, we introduce a natural and elegant reformulation of this condition in terms of mathematical group theory. A *group* G is a set of elements i with a binary operation (which we represent multiplicatively) that satisfies the following three axioms:

1. There is an identity element e with the property
$$ei = ie = i \quad \text{for all } i.$$
2. Every element i has a unique inverse i^{-1} such that
$$ii^{-1} = i^{-1}i = e.$$
3. The operation is *associative*:
$$i(jk) = (ij)k \quad \text{for all } i, j, \text{ and } k.$$

The group is called *abelian* if $ij = ji$ for all i and j. Examples of two groups of order 8 are given in Figure 1.

To apply this to our population, we let the breeding sites (the nodes) be the elements of the group. To specify the group multiplication, we use the homogeneity of the structure. Take a random node and label it e. Now take another node i. The homogeneity property is that the population should look the same from i as it does from e. In particular, for any node j, there should be a node that looks the same from i as j looks from e. We call that node ij and this in fact defines the multiplication operation on the node set (Fig. 2). For all of the homogeneous population structures that we have encountered in the theoretical literature, in particular, stepping-stone structures, cycles, lattices, tori, and island structures (Fig. 3), this operation satisfies the axioms of a group.

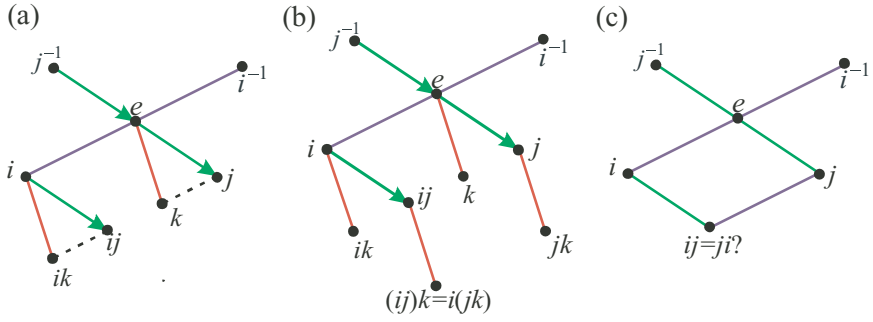

Figure 2. The group structure on the set of breeding sites. The group multiplication is illustrated in (A). For example, j^{-1} is the node from which e looks the same as j looks from e. In (B), the associativity is illustrated—that "extending" from ij by k is the same as "extending" from i by jk. The diagram is not meant to include all of the nodes nor all of the arcs, but enough connections are shown to capture the properties of node multiplication. Sometimes the offspring dispersal between sites i and j will be symmetric [$d(i, j) = d(j, i)$] and in this case we draw an edge between the nodes (no arrow). When the relationship is asymmetric, we use an arc signaled by an arrow. A group is called *abelian* if $ij = ji$ for all i and j. Panel (C) suggests that this might always be expected, but there are nonabelian groups (Fig. 4). The problem with (C) is that it uses a vector addition representation of group multiplication and vector addition is commutative.

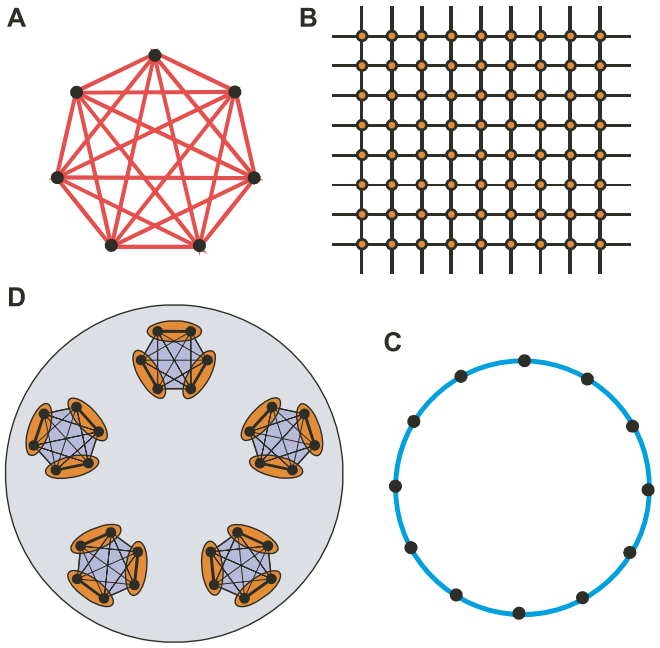

Figure 3. **A collection of homogeneous population structures. (A) a regular graph; (B) an infinite lattice—can be made finite (a torus) by identifying opposite boundaries, (C) a cycle, (D) an island model with 30 sites organized into five demes each with three dyadic patches. In example (D), the light and heavy edges correspond to two different dispersal rates. There will also be dispersal between demes but these edges are not shown. We assume however that offspring leaving the deme settle at random on the other demes. This population provides a good example of how a group structure can be imposed on a homogeneous population. Effectively we use the geometric symmetry obtained by arranging the demes and the sites within a deme in a cycle. If we index the demes in order from 0 to 4 and the nodes within each deme in order from 0 to 5, then every node can be specified by its angular displacement (θ, φ) from (0, 0) where the site angle θ is a multiple of 60° and the deme angle φ is a multiple of 72°. Then group multiplication is simply addition of the component angles. This gives us an abelian group structure. An isolated island of this type can be regarded as a finite population with $N = 30$ sites. To get an infinite population, we could take an infinite unstructured collection of such islands with some interisland dispersal.**

Now we make precise the homogeneity condition that the population "looks the same" when viewed from any breeding site. This is an assumption that the dispersal probabilities and the primary fitness effects are both invariant under the group structure. For the fitness effects we suppose that breeder fecundity F_j and survival S_j depend on the trait level z_k at a number of sites k. That is, for any i, j, and k

$$d(j,k) = d(ij, ik) \tag{2}$$

$$\frac{\partial F_j}{\partial z_k} = \frac{\partial F_{ij}}{\partial z_{ik}} \quad \text{(fecundity invariance)} \tag{3}$$

$$\frac{\partial S_j}{\partial z_k} = \frac{\partial S_{ij}}{\partial z_{ik}} \quad \text{(survival invariance).} \tag{4}$$

We emphasize that the product subscripts ij and ik in (2)–(4) are not double subscripts but are single subscripts that are products.

Now suppose that a breeder can observe when each breeder dies, where each breeder sends her offspring and with whom each breeder interacts. Then equations (2)–(4) tell us that if we took the breeder on one site, and without her knowledge moved her to another, she would be unable to tell that she had been moved. This is the homogeneity condition.

The Inclusive Fitness Effect

The relatedness coefficients $R(j, k)$ are determined as the solutions of a set of recursive equations (Appendix and Appendix S1) involving only the dispersal coefficients $d(j, k)$ and it follows from equation (2) that these are also invariant under left multiplication:

$$R(j,k) = R(ij, ik). \tag{5}$$

We calculate the inclusive fitness effect using the neighbor-modulated approach (Taylor and Frank 1996). We calculate the fitness of a focal breeder measured as the expected number of offspring who win breeding sites in the next generation (Wright–Fisher) or the difference between the birth rate and the death rate (Moran BD) as

$$w_e = \sum_j \frac{F_e d(e,j)}{\sum_i F_i d(i,j)} \quad \text{(Wright–Fisher)} \tag{6}$$

$$w_e = F_e - \sum_i F_i \frac{d(i,e)(1-S_e)}{\sum_j d(i,j)(1-S_j)} \quad \text{(Moran BD)} \tag{7}$$

$$w_e = \sum_i (1-S_i)\frac{F_e d(e,i)}{\sum_j F_j d(j,i)} - (1-S_e) \quad \text{(Moran DB).} \tag{8}$$

In equation (6) the numerator is the relative number of focal offspring competing at site j and the denominator is the total relative number at j, so that w_e is the probability a focal offspring will win the site. In equation (7) the summation term is focal mortality. Here F_i is the rate at which site i produces offspring and the quotient is the probability that a site i offspring will displace the focal breeder. Note that this depends on a product of the tendency of such offspring to disperse to the focal site and the focal vulnerability $1 - S_e$. In equation (8), offspring are produced only when there is a local vacancy to fill and the ith summand gives the rate at which focal offspring will colonize site i, this being determined by the mortality rate of breeder i, and the relative competitive pressure of the focal breeder at site i.

The inclusive fitness effect is then

$$W_I = \sum_k R(e,k)\frac{\partial w_e}{\partial z_k}, \tag{9}$$

where the derivatives are evaluated at the resident value $\hat{z}$.

Results

Our purpose is to establish Theorem 1. To have a specific version for the Moran model, we work with a BD demographic with selection only on fecundity F_i. We will comment later on other versions.

THEOREM 1

Suppose we have a population of breeding sites structured as an abelian group, each site inhabited by an asexual haploid breeder. Suppose that generations are nonoverlapping (Wright–Fisher process) or continuous (Moran process) with a BD protocol. In the Moran model, but not in the Wright–Fisher model, we require that dispersal be symmetric. Consider a behavioural trait which has a primary effect on the fecundity of the actor at node e and of others. Suppose the dispersal probabilities $d(j, k)$ and the fecundity effects $\partial F_j/\partial z_k$ are both invariant under left multiplication (eqs. 2 and 3). Then in an infinite population the inclusive fitness effect of the behaviour has the same sign as the effect of the focal behaviour on her own fecundity. In a finite population we have to normalize by subtracting the average effect of the focal behaviour on the fecundity of all other breeders in the population. [In an infinite population, this correction is zero.]

$$\text{Infinite population:}\quad W_I \sim \frac{\partial F_e}{\partial z_e} \tag{10}$$

Finite population:

$$W_I \sim \begin{cases} \dfrac{\partial F_e}{\partial z_e} - E_{i\neq e}\left(\dfrac{\partial F_i}{\partial z_e}\right) & \text{(inclusive fitness)} \\ \dfrac{\partial F_e}{\partial z_e} - E_{i\neq e}\left(\dfrac{\partial F_e}{\partial z_i}\right) & \text{(neighbour modulated)} \end{cases} \tag{11}$$

where the derivatives are evaluated at the resident value $\hat{z}$.

In the Appendix we provide a proof for one of the four cases of Theorem 1, the Moran model in an infinite population. The other three cases are variations on this and are proved in the Appendix S1.

FUNDAMENTAL ASSUMPTION BEHIND THE INCLUSIVE FITNESS METHOD

Theorem 1 requires the standard genetic assumptions of additive gene action and small effects of individual genes on fitness (Rousset and Billiard 2000; Taylor et al. 2007a). For example, there can be no synergistic effects between actor and recipient or among recipients (Queller 1985).

TWO COMPUTATIONAL APPROACHES

The equivalence of the two forms in equation (11) derives from the observation that $\frac{\partial F_i}{\partial z_e} = \frac{\partial F_e}{\partial z_{i-1}}$ (eq. 3) and the average of the latter over all $i \neq e$ is the same as the average found in the neighbor-modulated form of (11). These two forms correspond to the two standard ways of calculating the inclusive fitness effect, the classical inclusive fitness approach in which we consider the effect of the focal deviant behaviour on the fitness of others, and the neighbor-modulated approach in which we consider the effect of the behaviour of others on focal fitness. It is important, conceptually as well as computationally, to have both approaches at hand. Our proof of the Theorem works with the neighbor-modulated approach.

THE ASSUMPTION THAT DISPERSAL IS SYMMETRIC

Previous results have assumed the symmetry of the dispersal rates $[d(i,j) = d(j,i)]$ but we have shown that this is not always needed. Asymmetric dispersal is expected to be found at a significant level whenever the population sits in a "flow"-environment such as at the bottom of a stream or under prevailing winds. Of interest is our finding that Theorem 1 requires dispersal to be symmetric under a Moran process but not under Wright–Fisher. Example 3 below describes a cycle population with one dispersal rate clockwise and another counterclockwise. Under a Moran demography, this produces a selection pressure on altruism in one direction and on spite in the other, but the same asymmetric dispersal structure with a Wright–Fisher demography will satisfy the assumptions of Theorem 1 and cannot support altruism.

THE MORAN PROCESS—FECUNDITY VERSUS SURVIVAL

Behavioural interactions will generally affect one or both of the fecundity F_j and the survival S_j. Parallel to this, we have the two standard protocols for the Moran model BD and DB described above. Theorem 1 has assumed a BD protocol with primary effects on fecundity and uniform survival, but it also holds under a DB protocol with primary effects on survival and uniform fecundity—the proof is virtually identical with survival S_i in place of fecundity F_i and we outline it here. Under the DB protocol with survival selection and constant fecundity, focal fitness (8) can be written $w_e = S_e - \sum_i S_i d(e,i)$. This is the same as the fecundity equation (Appendix A1) for BD with $d(i, e)$ replaced by $d(e, i)$, and, of course, F replaced by S. Given symmetric dispersal, the argument goes through exactly as in the fecundity effect case.

PARTIAL SURVIVAL IN WRIGHT–FISHER

In the standard Wright–Fisher model generations do not overlap so that there is no variation in survival (all breeders die). However an intermediate demographic has discrete generations that

Table 1. Summary of results. Theorem 1 fails to hold in general for a Wright–Fisher demography with partial survival (Taylor and Irwin 2000) and for a Moran demography with survival selection under BD or fecundity selection under DB. Here, we summarize the situation for the remaining cases. Some of these that are not covered by Theorem 1 are discussed in the Examples or in the Supporting information. We consider the case of a nonabelian group to be of limited biological interest and the proofs for this case are not presented. They are available from PDT.

	When does Theorem 1 hold?			
	Group Abelian		Group Non-Abelian	
	Dispersal symmetric	Dispersal asymmetric	Dispersal symmetric	Dispersal asymmetric
Wright–Fisher (fecundity selection)	YES Theorem 1	YES Theorem 1	YES (proof not given)	NO Fig. 4 (Appendix S3)
Moran BD (fec. sel.) DB (surv. sel.)	YES Theorem 1	NO Example 3 Fig. 5 (Appendix S2)	YES (proof not given)	NO Fig. 4 (Appendix S3) not presented

overlap—rather than being replaced by an offspring, breeder j survives to the next generation with probability S_j. In this case, Theorem 1 does not hold and altruism and spite can both be selected. For example, Taylor and Irwin (2000) have investigated this demography in an infinite island model (with demes of constant size) and found that, for interactions at random with dememates, altruism can be favored (with a large enough b/c ratio) under fecundity selection (with constant breeder survival) and spite can be favored (with a large enough h/c ratio) under survival selection (with constant breeder fecundity). Here, h is the harm inflicted on the deme mate (increased mortality) at cost c. Similar results are found in Irwin and Taylor (2001) in a stepping-stone population.

THE ABELIAN ASSUMPTION

Theorem 1 assumes that the underlying group is abelian. This assumption is often not needed (see Table 1) but having it makes the proof of Theorem 1 simpler. In practice this is not a significant restriction as all standard homogenous structures found in the literature are abelian. For this reason, we have not in this treatment paid much attention to nonabelian groups. The smallest such group has 6 nodes and can be realized as the group of all permutations of 3 objects under composition (Fig. 4). Using this group we can construct an example of a Wright–Fisher process for which Theorem 1 fails and altruism can be selected (Appendix S3). But as we have suggested this is of more mathematical than biological interest.

EXAMPLE 1. "HAMILTON'S RULE"

Consider an altruistic trait in which individuals give primary fecundity benefits to various other individuals at total cost c. Then, focal fecundity might have the form $F_e = 1 - cz_e + \sum_i b_i z_i$ where b_i is the focal benefit received from site i when there is an altruist at that site. To move from there to a calculation of W_I, we need the offspring dispersal patterns, both to calculate the competitive effects (which are needed for focal fitness w_e) and to get the focal relatedness coefficients. However the Theorem tells us that none of that is actually needed (assuming symmetric dispersal in an abelian group):

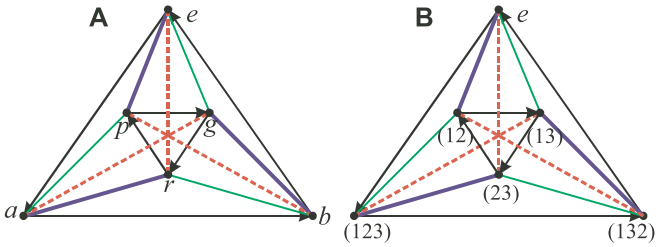

Figure 4. (A) A nonabelian group with six nodes. We have given the elements abstract names corresponding to the colors (or weights) of the edges from the identity *e*: *a*, arrow, *b*, back arrow, *p*, purple (heavy), *g*, green (light) and *r*, red (dotted). To multiply by any of these just follow the appropriate edge. Thus $ap = r$ because if we follow the purple edge from *a* we get to *r*. On the other hand, $pa = g$ because the arrow leaving *p* leads to *g*. Thus the group is nonabelian. To specify the dispersal probabilities, we are free to choose the six dispersal rates $d(e, h)$ for $h = e, a, b, p, g, r$. And then the rates $d(j, k)$ will be obtained from the invariance equation (2): $d(j, k) = d(e, j^{-1}k)$. Check that a breeder who can see only the dispersal rates will see the same structure from every node. In the Appendix S3, we use this population structure to construct an example of an infinite nonabelian group in which equation (8) fails under a Wright–Fisher process. (B) Demonstration that the group is isomorphic to the group S_3 of all permutations of three objects (which is the smallest nonabelian group). The six elements of this group are written: {e, (123), (132), (12), (13), (23)} where we use cyclic notation, so that (123) represents the permutation that maps 1 to 2, 2 to 3, and 3 to 1, and (12) is the "transposition" that interchanges 1 and 2 and leaves 3 fixed. Group multiplication is by composition (left to right), so that (123)(12) = (23) because if we apply (123) and then interchange 1 and 2, the net result is the interchange of 2 and 3. Observe that (12)(123) = (13).

$$\text{Infinite population:}\quad W_I \sim \frac{\partial F_e}{\partial z_e} = -c$$

$$\text{Finite population:}\quad W_I \sim \frac{\partial F_e}{\partial z_e} - E_{i\neq e}\left(\frac{\partial F_e}{\partial z_i}\right) = -c - \bar{b},$$

where $\bar{b}$ is the average value of the b_i over all nonfocal individuals. The finite population equation was obtained for the Moran process by Taylor et al. (2007b) and Grafen and Archetti (2008). In these versions, the actor gives b to a single other individual so that $\bar{b} = b/(N-1)$ where N is population size. These equations make it clear that altruism can never be selected. In an infinite population, spite ($b_i < 0$—Hamilton 1971; Gardner and West 2004) can also never be selected, but it "can" be selected in a finite population if the average harm done by a focal actor to other individuals in the population exceeds the focal cost.

What became of Hamilton's rule $br > c$? Well first of all there might be many b's so the rule would have to be written: $\sum_i b_i R_i > c$. But second, these are all primary fitness effects, and once the secondary effects are added in, the b_i evidently all cancel out. Indeed that is the thrust of Theorem 1. A significant question, for both theory and experiment, is to identify the structural variations on homogeneity, which will bring the b_i back and allow cooperation to evolve. We understand this to be a question of expanding the "scale of competition" in a way that will not reduce the relatedness among interactants (Queller 1994; Griffin et al. 2004 and Kummerli et al. 2009). Griffin et al. (2004) include a nice graph that shows cooperation approaching zero as the scale of competition becomes local. That is Theorem 1 kicking in.

EXAMPLE 2. COMPETITION

Take a finite population of annual plants with five demes of six breeders, each deme structured as three patches of size 2 (Fig. 3D). Consider a competitive trait such as height so that focal fecundity is a function F of the difference between focal height and the average height of the others in the deme, in which this average is weighted, with the focal patchmate given four times the weight of each of the others. In addition, we suppose that there is a quadratic cost to increased height. An expression with these properties has the form:

$$F_e = F\left(z_e - \frac{4z_1 + z_2 + z_3 + z_4 + z_5}{8}\right) - cz_e^2,$$

where z_e is focal height, z_1 is its patchmate's height, and the remaining z_i are the heights of the other four deme-mates.

Again to move from here to W_I, we need the offspring dispersal patterns within and between demes, both to calculate the competitive effects and to get the various focal relatedness coefficients. However, given that seed dispersal has the same geometric symmetry pattern as the population structure itself, equations (2) and (3) will hold, and equation (11) gives us the inclusive fitness effect of increased height to be

$$\begin{aligned} W_I &\sim \frac{\partial F_e}{\partial z_e} - E_{i\neq e}\left(\frac{\partial F_e}{\partial z_i}\right) \\ &= F'(0) - 2c\hat{z} - F'(0)\frac{1}{5}\left(-\frac{4+1+1+1+1}{8}\right) \\ &= \tfrac{6}{5}F'(0) - 2c\hat{z}. \end{aligned}$$

This gives us an equilibrium height of $\hat{z} = \frac{3F'(0)}{5c}$ and this can be shown to be stable.

EXAMPLE 3. A CYCLE WITH ASYMMETRIC DISPERSAL

The population is depicted in Figure 5. We use a Moran process with a BD protocol. We study an altruistic or spiteful trait whereby a breeder gives fecundity increment b to her clockwise neighbor at fecundity cost c. In the Appendix S2, we show that the condition $W_I > 0$ for the trait to be selected is that $b(\frac{\alpha-2\beta}{2(\alpha+\beta)}) > c$. For symmetric dispersal ($\alpha = \beta$) we get the Theorem 1 result: $-c - b/4 > 0$, as expected, but for $\alpha > 2\beta$ altruism is selected for sufficiently large b. For example, for $\alpha = 4\beta$, the condition is $b > 5c$—a gift to the "clockwise" neighbor is selected if it is at least five times the cost. At the same time, taking $\beta = 4\alpha$, harm h done to my "counter-clockwise" neighbor at cost c will be selected when $7h > 10c$. We have altruism one way and spite the other. In the Appendix S2, we also verify directly that the conclusion of Theorem 1 holds for this population under a Wright–Fisher demographic.

As a remark we report that for a cycle of size N the condition $W_I > 0$ is

$$b\left(\frac{\alpha(N-3) - \beta(N-1)}{(\alpha+\beta)(N-1)}\right) > c.$$

For large N, α needs only be a bit greater than β for altruism to be selected.

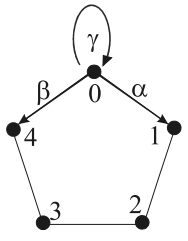

Figure 5. A five-cycle asymmetric dispersal. We show the dispersal pattern for offspring at node 0, but the same pattern applies to the other nodes. Offspring disperse with probability α to the immediate clockwise neighbor, β to the counter-clockwise neighbor. With probability $\gamma = 1 - \alpha - \beta$ they remain at home and displace their parent. To obtain the group structure, number the nodes from 0 to 4, clockwise and take the group multiplication to be addition mod 5, e.g., $2 + 4 = 6 = 1$ (cast out 5). Essentially, this gives us the rotation group generated by the clockwise rotation through one-fifth of a revolution. This group is clearly abelian.

EXAMPLE 4. FRANK'S MODEL OF WITHIN-DEME COMPETITION

We consider an infinite asexual haploid island population consisting of demes of size n. [For the model we are discussing here the word "group" is usually used but to avoid an obvious confusion we will use the term "deme."] We assume a Moran model of continuous reproduction in which an offspring displaces a breeder chosen at random from the natal deme with probability h and otherwise disperses to a breeding site on a distant deme. With this structure, a standard recursive argument shows that within-deme relatedness (focal to entire deme including self) will depend on both n and h:

$$R = \frac{1}{n - h(n-1)}. \tag{12}$$

Frank's (1994) model for within-deme competitiveness necessarily sits in a structured population, but it is nevertheless usually presented with an upfront assumption on the mathematical form of "fitness":

$$W(z_e, \bar{z}) = \frac{z_e}{\bar{z}}(1 - \bar{z}) \tag{13}$$

with z_e the focal phenotype and $\bar{z}$ the average phenotype of the focal deme. The inclusive fitness effect is then:

$$W_I = \frac{\partial W}{\partial z_e} + \frac{\partial W}{\partial \bar{z}} R$$

where R is the average relatedness to the deme. This gives us

$$W_I = \frac{1 - z - R}{z} \tag{14}$$

with an equilibrium at

$$z^* = 1 - R = \frac{(n-1)(1-h)}{n - h(n-1)} \tag{15}$$

using the formula for R in equation (12).

However suppose, as seems more reasonable, that the competitive behaviour directly affects only breeder "fecundity" F, leaving the indirect mortality effects of offspring recruitment to be determined by the population structure. In this case we might assume that fecundity itself has the form we assumed in equation (13) for fitness W:

$$F_e = f(z_e, \bar{z}) = \frac{z_e}{\bar{z}}(1 - \bar{z}). \tag{16}$$

Because the island structure is homogeneous, the inclusive fitness effect will be given by equation (10):

$$W_I = \frac{\partial F_e}{\partial z_e} = \frac{\partial f}{\partial z_e} + \frac{\partial f}{\partial \bar{z}} \frac{\partial \bar{z}}{\partial z_e} = \frac{\partial f}{\partial z_e} + \frac{\partial f}{\partial \bar{z}} \frac{1}{n} \tag{17}$$

This is the same expression as above (using f instead of W) with R replaced by $1/n$ (although the actual R is given by the formula at the top of this page). The equilibrium has the simple form:

$$z^* = 1 - \frac{1}{n} = \frac{n-1}{n}. \tag{18}$$

The two formulae, (15) and (18), for z^* are close but not quite the same.

Discussion

GROUP THEORY

Long before the theory of groups was properly formulated, mathematicians and physicists were using its ideas to try to understand the geometry of objects, such as polyhedra, that have a significant amount of internal symmetry. We have argued that such internal symmetry is effectively what is assumed in most "unstructured" inclusive fitness models, so that in a homogeneous structured population, group theory seems an ideal vehicle to formalize the analysis—and it gives us simpler and stronger analytical tools. Finally it provides a conjunction of two elegant theoretical domains, mathematical group theory and inclusive fitness analysis.

PREVIOUS RESULTS

Previous discussions of inclusive fitness in homogeneous populations (Taylor 1992a, 1992b; Queller 1994, West et al. 2002; Taylor et al. 2007b; Grafen and Archetti 2008) worked with altruism, but our formulation applies to any social trait that affects fecundity. Previous results work with assumptions of transitivity, which at first appear to be weaker than our assumption of group structure. [A structure is transitive if for every pair of nodes i and j, there is a node bijection mapping i to j and preserving dispersal and interaction probabilities.] Certainly any group-structured population is transitive (group multiplication provides the bijections), but it is not known whether there are transitive structures, which are not groups. This question is closely related to an open problem in group theory (Appendix S4).

RELAXING THE ASSUMPTIONS

In the Supporting information, we discuss a number of examples in which the assumptions of Theorem 1 do not hold. Section S2 looks at asymmetric dispersal and S3 presents an analysis of a nonabelian group.

TYPES OF INHOMOGENEITY

Theorem 1 has significant limitations. Although it can be made to work for a sexual diploid population if males and females are treated the same, it does not generally apply in a class- or age-structured population, nor to ploidies other than 1. In particular, it does not apply to sex-ratio traits, or to sex-specific behaviour. The recent study of Cornwallis et al. (2010) on the relationship between helping at the nest and local relatedness (affected by

levels of promiscuity) is a good example of a population structure exhibiting many of the complexities that Theorem 1 is unable to deal with. Second, although the trait is supposed to affect the fecundities F_i, it cannot affect the offspring dispersal probabilities $d(i, j)$. In particular, it does not apply in models of optimal dispersal. It seems to apply most readily in models of cooperation and competition.

ACKNOWLEDGMENTS

This work has been funded by a grant from the Natural Sciences and Engineering Research Council of Canada. We are grateful to A. Hurford, T. Day, and A. Grafen for useful comments. The project was initiated by PT who was the main contributor to the results on inclusive fitness. The mathematical results around graphs and group theory were mainly worked out and assembled by TL and DC. All participated in the writing of the article.

LITERATURE CITED

Cornwallis, C. K., S. A. West, K. E. Davis, and A. S. Griffin. 2010. Promiscuity and the evolutionary transition to complex societies. Nature 466:969–972.

Frank, S. A. 1994. Kin selection and virulence in the evolution of protocells and parasites. Proc. R. Soc. Lond. B 258:153–161.

Gardner, A., and S. A. West. 2004. Spite and the scale of competition. J. Evol. Biol. 17:1195–1203.

Gardner, A., and S. A. West. 2006. Demography, altruism, and the benefits of budding. J. Evol. Biol. 19:1707–1716.

Grafen, A., and M. Archetti. 2008. Natural selection of altruism in inelastic homogeneous viscous populations. J. Theor. Biol. 252:694–710.

Griffin, A. S., S. A. West, and A. Buckling. 2004. Cooperation and competition in pathogenic bacteria. Nature 430:1024–1027.

Hamilton, W. D. 1964. The genetical evolution of social behaviour I. J. Theor. Biol. 7:1–16.

Hamilton, W. D. 1971. Selection of selfish and altruistic behaviour in some extreme models. Pp. 57–91 *in* J. F. Eisenberg and W. S. Dillon, eds. Man and beast: comparative social behaviour. Smithsonian Press, Washington, D.C.

Irwin, A. J., and P. D. Taylor. 2001. Evolution of altruism in a stepping-stone population with overlapping generations. Theor. Pop. Biol. 60:315–325.

Kummerli, R., A. Gardner, S. A. West, and A. S. Griffin. 2009. Limited dispersal, budding dispersal, and cooperation: an experimental study. Evolution 63:939–949.

Michod, R. E., and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. Nature 288:694–697.

Nakamaru, M., H. Matsuda, and Y. Iwasa. 1997. The evolution of cooperation in a lattice structured population. J. Theor. Biol. 184:65–81.

Nakamura, M., H. Nogami, and Y. Iwasa. 1998. Score-dependent fertility model for the evolution of cooperation in a lattice. J. Theor. Biol. 194:101–124.

Nowak, M. A., and R. M. May, 1992. Evolutionary games and spatial chaos. Nature 359:826–829.

Ohtsuki, H., and M. A. Nowak. 2006a. The replicator equation on graphs. J. Theor. Biol. 243:86–97.

———. 2006b. Evolutionary games on cycles. Proc. R. Soc. Lond. B 273:2249–2256.

Ohtsuki, H., J. Pacheco, and M. A. Nowak. 2007. Evolutionary graph theory: breaking the symmetry between interaction and replacement. J. Theor. Biol. 246:681–694.

Queller, D. C. 1985. Kinship, reciprocity and synergism in the evolution of social behaviour: a synthetic model. Nature 318:366–367.

———. 1994. Genetic relatedness in viscous populations. Evol. Ecol. 8:70–73.

Rousset, F., and S. Billiard. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J. Evol. Biol. 13:814–825.

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

———. 1992b. Inclusive fitness in a homogeneous environment. Proc. R. Soc. Lond. B 249:299–302.

———. 1996. Inclusive fitness arguments in genetic models of behaviour. J. Math. Biol. 34:654–674.

Taylor, P. D., and S. Frank. 1996. How to make a kin selection argument. J. Theor. Biol. 180:27–37.

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

Taylor P. D., T. Day, and G. Wild. 2007a. From inclusive fitness to fixation probability in homogeneous structured populations. J. Theor. Biol. 249:101–110. doi:10.1016/j.jtbi.2007.07.006.

———. 2007b. Evolution of cooperation in a finite homogeneous graph. Nature 447:469–472.

West, S. A., I. Pen, and A. S. Griffin. 2002. Cooperation and competition between relatives. Science 296:72–75.

Wilson, D. S., G. B. Pollock, and L. A. Dugatkin. 1992. Can altruism evolve in purely viscous populations? Evol. Ecol. 6:331–341.

Associate Editor: S. West

Appendix

Here, we provide a proof for one of the four cases of Theorem 1, the Moran process in an infinite population. The other three cases have similar arguments and are treated in the Appendix S1. Our calculations assume that the group is abelian and that dispersal is symmetric. One observation we make is that is that the relatedness recursions (eq. A3 below) are linear in $d(k, h)$. For the Wright–Fisher process, these same recursions are quadratic (S1.6). This is an illustration of the difference between Grafen and Archetti's (2008) 1-circle and 2-circle (Table A2).

CALCULATION OF W_I

These calculations are the same for both infinite and finite populations. Under the BD protocol with fecundity selection and constant survival, focal fitness (7) can be written

$$w_e = F_e - \sum_i F_i d(i, e). \tag{A1}$$

Hence

$$\frac{\partial w_e}{\partial z_k} = \frac{\partial F_e}{\partial z_k} - \sum_i \frac{\partial F_i}{\partial z_k} d(i, e)$$

$$= \frac{\partial F_e}{\partial z_k} - \sum_i \frac{\partial F_e}{\partial z_{i^{-1}k}} d(i, e)$$

Table A2. "Circles of compensation" (Grafen and Archetti 2008) The "competitive distance" r_C: Suppose a focal breeder is given an extra unit of fecundity. Then r_C is the distance at which the competitive effects are felt. The "gene-flow distance" r_G is the number of dispersal steps the two genes in a pair of breeders had to travel in the most recent gene replacement involving the pair. These distances are recorded in (A) for a number of standard models. In (B), we tabulate the results of Grafen and Archetti's (2008) heuristic analysis, which assumes a primary fecundity gift b. Our results conform to this analysis except in the case of a nonabelian group structure under Wright–Fisher demography with asymmetric offspring dispersal. In this case, $r_C=r_G$ but the theorem fails to hold and altruism is favored.

(A)

Demography	r_C	r_G
Moran BD	1	1
Moran DB	2	1
Wright–Fisher	2	2
Wright–Fisher with partial survival s	2	$s1+(1-s)2$

(B)

	Benefit-cost interaction $c>0$
$r_C=r_G$.	theorem holds—neither altruism nor spite selected
$r_C>r_G$.	altruism favored ($b>0$) if b/c sufficiently large
$r_C<r_G$.	spite favored ($b<0$) if $-b/c$ sufficiently large

[Change variable $h = i^{-1}k$ so that $d(i, e) = d(e, i^{-1}) = d(e, i^{-1}kk^{-1}) = d(e, hk^{-1})$ using eq. 2]

$$= \frac{\partial F_e}{\partial z_k} - \sum_h \frac{\partial F_e}{\partial z_h} d(e, hk^{-1})$$

Now put this expression into equation (9)

$$W_I = \sum_k R(e,k)\left(\frac{\partial F_e}{\partial z_k} - \sum_h \frac{\partial F_e}{\partial z_h} d(e, hk^{-1})\right).$$

For all h, the coefficient of $\partial F_e/\partial z_h$ in W_I is

$$R(e,h) - \sum_k R(e,k) d(e, hk^{-1}) \qquad \text{(A2)}$$

RELATEDNESS RECURSIONS

The equilibrium values of relatedness can be found directly from the recursive equations for R. Recall that in this case we ignore mutation, giving long-distance dispersal the task of maintaining genetic variance.

Take h to be a site distinct from e. Because the alleles are neutral the two sites, e and h, will have equal probability ($^1/_2$) to be the one most recently replaced. In the expression for $R(e, h)$ below, the first sum assumes that this is e and the second assumes that this is h. At equilibrium:

$$R(e,h) = \frac{1}{2}\left[\sum_j R(j,h)d(j,e) + \sum_k R(e,k)d(k,h)\right] (h \neq e) \qquad \text{(A3)}$$

We begin by showing that the two terms in the square brackets are the same. The left-hand term is

$$\sum_j R(j,h)d(j,e)$$

$$= \sum_j R(e, j^{-1}h)d(j,e) \quad \text{(invariance under left multiplication)}$$

$$= \sum_k R(e,k)d(hk^{-1}, e) \quad (k = j^{-1}h, \text{ so that } j = hk^{-1})$$

$$= \sum_k R(e,k)d(k^{-1}h, e) \quad \text{(commutativity of multiplication)}$$

$$= \sum_k R(e,k)d(h,k) \quad \text{(invariance under left multiplication)}$$

and the symmetry of d gives us the right-hand term.

It follow from (A3) that $R(e, h)$ is equal to the right-hand term:

$$R(e,h) = \sum_k R(e,k)d(k,h) \qquad \text{(A4)}$$

Using invariance and commutativity:

$$d(k,h) = d(h^{-1}k, e) = d(kh^{-1}, e) = d(h^{-1}, k^{-1}) = d(e, hk^{-1})$$

and we deduce

$$R(e,h) = \sum_k R(e,k)d(e, hk^{-1})$$

this tells us that (A2) is zero so that the coefficient of $\partial F_e/\partial z_h$ is zero for $h \neq e$ and we are left with only the $h = e$ term in W_I. Setting $h = e$ in (A1) we get:

$$W_I = \left[1 - \sum_k R(e,k)d(e,k^{-1})\right]\frac{\partial F_e}{\partial z_e}.$$

Note that $\sum_k R(e,k)d(e,k^{-1})$ is an average of relatedness coefficients and will be less than 1. Hence the term in the square brackets is positive, and we get equation (10).

Supporting Information

The following supporting information is available for this article:

Appendix S1. Demonstration of Theorem 1.
Appendix S2. Investigating the assumption of symmetric dispersal—an example of an abelian group structure with asymmetric dispersal for which the conclusions of Theorem 1 fail for the Moran process (but not of course for Wright–Fisher).
Appendix S3. An example of a nonabelian group structure with asymmetric dispersal for which the conclusions of Theorem 1 fail for a Wright–Fisher process.
Appendix S4. Supplementary material for inclusive fitness on groups: Does every node-transitive weighted digraph have an equivalent group structure?

Supporting Information may be found in the online version of this article.

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