

## Hamilton's inclusive fitness in finite-structured populations

Peter D. Taylor and Wes Maciejewski

*Phil. Trans. R. Soc. B* 2014 **369**, 20130360, published 31 March 2014

---

### References

**This article cites 40 articles, 6 of which can be accessed free**

<http://rstb.royalsocietypublishing.org/content/369/1642/20130360.full.html#ref-list-1>

### Subject collections

Articles on similar topics can be found in the following collections

[behaviour](#) (519 articles)

[evolution](#) (671 articles)

[theoretical biology](#) (61 articles)

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)



## Review

**Cite this article:** Taylor PD, Maciejewski W. 2014 Hamilton's inclusive fitness in finite-structured populations. *Phil. Trans. R. Soc. B* **369**: 20130360. <http://dx.doi.org/10.1098/rstb.2013.0360>

One contribution of 14 to a Theme Issue 'Inclusive fitness: 50 years on'.

### Subject Areas:

behaviour, evolution, theoretical biology

### Keywords:

allele frequency, relatedness, heterogeneous populations, class structure, reproductive value

### Author for correspondence:

Peter D. Taylor  
e-mail: [peter.taylor@queensu.ca](mailto:peter.taylor@queensu.ca)

# Hamilton's inclusive fitness in finite-structured populations

Peter D. Taylor<sup>1</sup> and Wes Maciejewski<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada K7L 3N6

<sup>2</sup>Department of Mathematics, The University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z2

Hamilton's formulation of inclusive fitness has been with us for 50 years. During the first 20 of those years attention was largely focused on the evolutionary trajectories of different behaviours, but over the past 20 years interest has been growing in the effect of population structure on the evolution of behaviour and that is our focus here. We discuss the evolutionary journey of the inclusive-fitness effect over this epoch, nurtured as it was in an essentially homogeneous environment (that of 'transitive' structures) having to adapt in different ways to meet the expectations of heterogeneous structures. We pay particular attention to the way in which the theory has managed to adapt the original constructs of relatedness and reproductive value to provide a formulation of inclusive fitness that captures a precise measure of allele-frequency change in finite-structured populations.

## 1. Introduction

'Hamilton's Rule' [1] states that a trait will be favoured by selection when the overall fitness benefit to the recipient multiplied by its genetic relatedness to the actor is greater than the overall fitness cost to the actor [2]. In symbols,

$$\mathbf{b}R > \mathbf{c}, \quad (1.1)$$

and here we use boldface notation for  $\mathbf{b}$  and  $\mathbf{c}$  to emphasize the principle that they must represent the total fitness effect on the individual [3,4]. This will distinguish it from the fecundity effects  $b$  and  $c$ , which will generally represent partial fitness effects. This rule is a mantra for many evolutionary biologists and there is good reason for that. Although the formulation covers only the fitness effects on two individuals (actor and recipient), it is simple and compelling and does capture Hamilton's fundamental insight that the effect on the recipient should carry a weight representing relative probability that affects the gene responsible for the behaviour. More generally of course, particularly in evolutionary graphs, an item of behaviour will typically affect the fitness of many individuals, who can all be regarded as 'recipients', and what is called *the inclusive-fitness effect* is the weighted sum of these fitness effects  $\mathbf{b}_i$  where the weights are the relatedness coefficients  $R_i$  of the actor to the  $i$ th recipient (one of whom will be the actor itself)

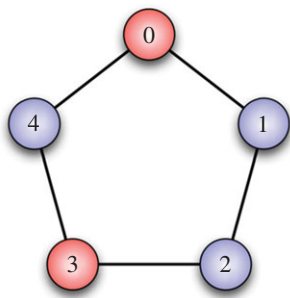
$$W_{\text{IF}} = \sum_i \mathbf{b}_i R_i. \quad (1.2)$$

In his 1964 paper, Hamilton argued that under certain general assumptions, an allele causing such behaviour will increase in frequency precisely when  $W_{\text{IF}}$  is positive. In rare cases, there might be only two affected recipients, the actor 0 and its 'partner' 1, and in that case

$$W_{\text{IF}} = \mathbf{b}_0 R_0 + \mathbf{b}_1 R_1 \quad (1.3)$$

and with  $\mathbf{b}_0 = -\mathbf{c}$  and the relatedness of an actor to itself equal to  $R_0 = 1$ , we recover the simple version.

Our purpose here is to trace the development of the theory of inclusive fitness with particular attention paid to the effect on the evolutionary outcome of population structure, all the way from the implicit homogeneity of the early models to a study of the explicit effects of heterogeneity and class structure. Our setting is finite evolutionary graphs and our particular focus is on recent



**Figure 1.** A 5-cycle with an *A* actor at position 0. This actor gives a primary benefit  $b$  to both neighbours 1 and 4, at a total cost of  $2c$ . These primary effects translate into secondary competitive effects that act on survival under BD and on fecundity under DB. Details are given in table 1. (Online version in colour.)

developments around coefficients of relatedness and reproductive values (RVs). We illustrate our remarks with two examples, one homogeneous (the 5-cycle) and the other heterogeneous (the 3-star).

## 2. Evolutionary graphs

We work with a finite-structured population of constant size. We represent it as a graph, that is, a finite set of nodes  $i$  with edges  $(i, j)$  between certain pairs of nodes [5]—see, for example, figure 1. Nodes that are joined by an edge are called neighbours. Each node is occupied by a haploid asexual individual and at the end of each time interval, one node is chosen to produce a single offspring, which displaces one of its neighbours. Each individual carries one of two possible alleles *A* and *B* assorting at a fixed locus. Offspring carry the parental allele except for mutation, which occurs at birth with a small probability  $u$ . In that case, both *A* and *B* mutate to an *A* form with probability  $p_N$  and to a *B* form with probability  $1 - p_N$ . We choose this notation as  $p_N$  turns out to be the long-term allele frequency (frequency of *A*) in the neutral population.

We suppose that in each time interval each *A*-individual, as actor, gives pay-off  $b$  to each of its neighbours at cost  $c$  per neighbour. We assume that these pay-offs represent small increments in fecundity, small enough that we can ignore second-order effects (w-weak selection as in [6]). These pay-offs are added to a baseline fecundity of 1 to give each individual a *relative fecundity*. To be clear, this assumption of additivity applies to pay-offs received from multiple *A*-nodes. For example, an *A*-node with three neighbours, one of which is *A*, will have relative fecundity  $1 + b - 3c$ .

At the end of each time interval, the fecundity (probability of giving birth) of each node is determined by an ‘update rule’ and here we will work with a Moran process with either a birth–death (BD) or death–birth (DB) updating [7,8]. Under a BD updating, the birth node is determined at random using the relative fecundities as weights. Under a DB updating, a random (unweighted) node is chosen and the occupying individual is removed. The replacement offspring is chosen at random from the neighbours of the chosen node using the relative fecundities as weights. In other treatments, these pay-offs might also represent increments in survival but we do not consider that here. It turns out that in the class of models we work with here there is a symmetry between the results for fecundity and survival pay-offs [9].

**Table 1.** Calculation of the inclusive-fitness effect on the 5-cycle with fecundity pay-offs. Relatedness coefficients are calculated in appendix A.

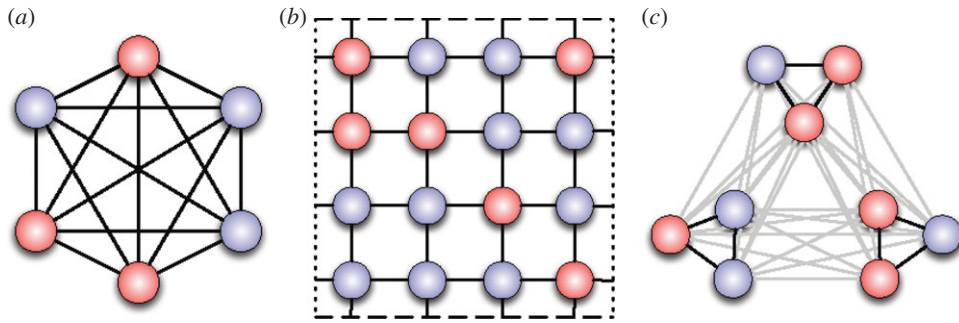
pay-off	primary		secondary	
	recipient	relatedness	recipient	relatedness
BD focal actor at node 0				
$-2c$	0	$R_0$	1,4	$R_1$
$b$	1	$R_1$	0,2	$(R_0 + R_2)/2$
$b$	4	$R_1$	0,3	$(R_0 + R_2)/2$
$W_{if} = 2b \left( R_1 - \frac{R_0 + R_2}{2} \right) - 2c(R_0 - R_1) = -b/2 - 2c$				
DB focal actor at node 0				
$-2c$	0	$R_0$	2,3	$R_2$
$b$	1	$R_1$	3,5	$(R_2 + R_1)/2$
$b$	4	$R_1$	1,3	$(R_1 + R_2)/2$
$W_{if} = 2b \left( R_1 - \frac{R_1 + R_2}{2} \right) - 2c(R_0 - R_2) = b/2 - 3c$				

It is clear that as a result of these interactions the fitness effect on each individual will in general be the result of the actions of a number of *A*-nodes. However, Hamilton’s inclusive-fitness effect, briefly described above, considers the effects of only one *A*-actor. For this to capture the overall picture, this ‘focal’ actor must somehow represent all actors and for that to happen we need some assumptions of symmetry in the population structure. This question of structure, and the different forms of internal symmetry it might exhibit, is our main objective of study in this paper.

## 3. An example: the 5-cycle

As an example, we consider the cycle graph (figure 1) [10] with both BD and DB updating and fecundity pay-offs. We fasten attention on an *A*-individual at node 0 and examine the inclusive-fitness effect of its behaviour [11]. The primary effects are fecundity gifts of  $b$  to each of nodes 1 and 4 and a consequent fecundity loss of  $2c$  for node 0.

From this point on, the story is different for the two update rules (table 1). Under BD, the gifts represent changes in the probability that the nodes will be selected to reproduce. These fecundity effects are called *primary*, as they are directly connected to the pay-offs. They will, however, lead to additional effects on survival classed as *secondary*. For example, the gain of  $b$  offspring to node 1 will give nodes 0 and 2 a survival loss of  $b/2$  each, and the loss of  $2c$  offspring from node 0 will give nodes 1 and 4 a survival gain of  $c$  each. However, under DB, no reproduction is possible until we have a death. For example, the fecundity gift of  $b$  to node 1 must await a death from either node 0 or node 2 to be realized, and in that case the competitive effect of that gift will be to reduce the effective fecundity of the node ‘on the other side’. For example, if node 0 dies, the competition to colonize the vacant node is between nodes 1 and 4, so that



**Figure 2.** Some examples of homogeneous graph-structured populations. (a) A well-mixed population, drawn as a complete graph, where all individuals are free to interact and replace any member of the population; (b) a lattice with periodic boundary, a toroidal lattice and (c) a structure where individuals interact more strongly in a triad than between triads, a type of island structure. (Online version in colour.)

node 1's gain of  $b$  will create an effective loss to node 4's fecundity. Actually, that only affects node 4's reproduction in half of its opportunities; the other half being the death of node 3. Thus, the overall effect on node 4 is  $-b/2$ . Similar stories apply to the fecundity gift of  $b$  to node 4 and the fecundity loss of  $2c$  to node 0. The take-away message is that the secondary effects of primary effects on fecundity act on survival under BD and on fecundity under DB.

The inclusive-fitness effect of all this is obtained by adding these effects up, each weighted by the relatedness of node 0 to the recipient node. The details are laid out in table 1 and the coefficients of relatedness are calculated in appendix A. The calculations give us the inclusive-fitness effects  $W_{IF} = -b/2 - 2c$  for BD updating and  $W_{IF} = b/2 - 3c$  for DB updating.

As we have mentioned,  $W_{IF}$  gives us the sign of the rate of increase of the allele  $A$  and, for the 5-cycle, the conditions for selection to act to increase the frequency of  $A$  are

$$-b > 4c \quad \text{BD updating} \quad (3.1)$$

and

$$b > 6c \quad \text{DB updating.} \quad (3.2)$$

Under DB updating, altruism ( $b, c > 0$ ) will be selected if the benefit inflicted on a neighbour is at least six times the cost. However, under BD updating, altruism can never be selected, but spite ( $-b, c > 0$ ) will be selected if the harm inflicted on a neighbour is at least four times the cost. It turns out that these are examples of a general result for 'homogeneous' populations, those with a special type of internal symmetry called transitivity.

## 4. Transitive graphs

A graph is called *transitive* [12,13] if given any pair  $i, j$  of nodes there is a bijection  $T$  of the node set mapping  $i$  to  $j$ , which preserves all edges (that is  $T(i)$  is a neighbour of  $T(j)$  if and only if  $i$  is a neighbour of  $j$ —such bijections are called *isomorphisms*). Clearly our 5-cycle is transitive (the isomorphisms are the rotations) and some additional examples are provided in figure 2. With this assumption, the graph 'looks the same' from every node and that allows us to work with a single 'focal' node, as we have done in the example above. But a striking result is that in this case of a transitive graph, the condition  $W_{IF} > 0$  is independent of the structure of the graph and depends only on the number  $N$  of nodes and the update rule. For example, for the Moran process with

fecundity pay-offs, the inclusive-fitness effect is positive exactly when [9,13–16]

$$-b > (N - 1)c \quad \text{BD updating} \quad (4.1)$$

and

$$b\left(\frac{N}{k} - 2\right) > (N - 2)c \quad \text{DB updating,} \quad (4.2)$$

where  $k$  is the degree of each node (the number of neighbours) so that for a cycle graph,  $k = 2$ . For the case  $N = 5, k = 2$ , we get the conditions above for the 5-cycle. A transparent explanation in terms of 'circles of compensation' of why the DB protocol is friendlier to altruism than BD is found in Grafen & Archetti [13]. A further discussion of the relationship between these two protocols is found in Taylor [9]. For example, if pay-offs affect survival rather than fecundity, the transitive population conditions above are reversed: equation (4.1) belongs to DB and equation (4.2) belongs to BD.

It turns out, not surprisingly, that the transitive graphs are precisely those for which the simple (one might say 'initial') form of Hamilton's inclusive-fitness analysis works. For these graphs, the fitness effects on the population of the actions of a single focal  $A$ -individual are able to capture the selective advantage of an  $A$ -allele. Our objective is to see how inclusive fitness can be generalized to apply to graphs with more complex symmetry patterns. The short answer is that the analysis works fine but that we need to incorporate RV and consider a focal individual for each reproductive class [17].

## 5. Allele-frequency change: the Price equation

Price's [18] covariance formula for the selective allele-frequency change (ignoring mutation) over a single time step can be written as

$$\Delta \bar{x} = \text{cov}(W, x), \quad (5.1)$$

where  $W$  is individual fitness effect (the overall fitness effect of all pay-offs),  $x$  is individual genotype ( $=1$  for an  $A$ -node and  $0$  for a  $B$ -node) and  $\bar{x}$  is the population-wide allele frequency (the average genotype). To calculate the overall fitness effect  $W_i$  on each node  $i$ , we need to know the *state* of the population, that is, which nodes are  $A$  and which are  $B$  and then the fitness effects for each node are calculated as the difference between the fecundity and the mortality effects. For example, for the 5-cycle (figure 1) with BD updating suppose that nodes 0 and 2 are  $A$  and the other four nodes are all  $B$ . As a result of the assumption of additivity among effects, we can calculate

$W_i$  by adding the separate effects of the behaviour of nodes 0 and 2. The calculation is given in table 2 and the resulting selective change of allele frequency is

$$\begin{aligned}\Delta\bar{x} &= \text{cov}(W, x) = E(Wx) - E(W)E(x) = E(Wx) \\ &= \frac{1}{6} \sum_i W_i x_i = \frac{1}{6} (W_0 + W_2) = \frac{-3b - 4c}{6},\end{aligned}\quad (5.2)$$

where we have used the fact that  $E(W) = 0$ , as the population size is constant.

The calculation in equation (5.2) gives the allele-frequency change in a particular state. However, over time, the population will wander among many possible states each with its own long-term frequency. One might well ask, in this case, how we might construct a reasonable measure of the overall selective advantage of the allele  $A$ . The generally accepted answer to this was provided in 2000 by Rousset & Billiard: take the average state-specific selective allele-frequency change  $\Delta\bar{x}$  and then average these over all states, each state weighted by its long-term frequency of occurring. That is not an easy calculation, as the state frequencies are not generally analytically accessible, but under certain general conditions this average can be calculated with the inclusive-fitness effect.

The transition from this average allele-frequency change to the inclusive-fitness effect is technically rather interesting. We summarize the process here and provide the missing calculation in appendix B. For simplicity, we work with a transitive population but we show later how the analysis generalizes to heterogeneous populations. What we need to calculate is

$$\sum_s \pi_s \Delta\bar{x}_s = \sum_s \pi_s \text{cov}_s(W, x), \quad (5.3)$$

where  $\pi_s$  is the frequency of state  $s$  and  $\text{cov}_s$  is the covariance over all nodes of the population in state  $s$ . The notation can get a bit hard to keep track of, so we adopt a notational device found in Taylor *et al.* [16]. We use round brackets to represent an average or a covariance over all nodes in the population, and square brackets to represent an average or a covariance over all population states. With this notation, equation (5.3) is written

$$E[\Delta\bar{x}] = E[\text{cov}(W, x)]. \quad (5.4)$$

The expression on the right-hand side asks us to calculate the covariance of fitness with genotype over the population in a fixed state and then average these over all states. In this form, the calculation is intractable. The trick, and the essence of Hamilton's 1964 approach, is to argue that under certain assumptions it is permissible to interchange the order of the operations and to calculate instead

$$E[\Delta\bar{x}] = E(\text{cov}[W, x]). \quad (5.5)$$

The equivalence of (5.4) and (5.5) is established in appendix B. In equation (5.5), we take a fixed node and calculate the covariance between its genotype and fitness over all possible states, and then average this over all nodes.

In the case of a transitive population, the result of this first covariance will be the same for all nodes and the final average is unnecessary. In this case, the formulation for average allele-frequency change becomes

$$E[\Delta\bar{x}] = \text{cov}[W_0, x_0], \quad (5.6)$$

where  $i = 0$  is a randomly chosen 'focal' node. We see later how to extend this formulation to heterogeneous graphs.

**Table 2.** Inclusive-fitness effects on the 5-cycle at each node  $i$  for  $A$  on nodes 0 and 2.

node $i$	effect of node 0	effect of node 2	total effect $W_i$
0	$-2c - b/2$	$-b/2$	$-2c - 3b/2$
1	$c + b$	$c + b$	$2c + 2b$
2	$-b/2$	$-2c - b/2$	$-2c - 3b/2$
3	$-b/2$	$c + b$	$c + b/2$
4	$c + b$	$-b/2$	$c + b/2$

## 6. The inclusive-fitness effect: relatedness

In a graph-structured population, the overall fitness  $W_i$  of each node  $i$  will be, to first order in the pay-offs, a linear function of the node genotypes

$$W_i = 1 + \sum_j b_{ij} x_j, \quad (6.1)$$

where  $b_{ij}$  is the fitness effect on  $i$  of an  $A$ -allele at node  $j$ . It is important to note that  $b_{ij}$  are independent of state, and in that case they can be pulled out of the square brackets in equation (5.5):

$$E[\Delta\bar{x}] = E(\text{cov}[W_i, x_i]) = E\left(\sum_j b_{ij} \text{cov}[x_j, x_i]\right). \quad (6.2)$$

Here, for clarity, we have replaced the random variables  $W$  and  $x$  in equation (5.5) with  $W_i$  and  $x_i$  and the expectation  $E$  is over all nodes  $i$ . The covariance in equation (6.2) takes a fixed pair of nodes and asks how their genotypes covary over the long-term life of the population. And the point is that this can be calculated with a simple argument. As a result of our assumption on the mutation process, the nodes  $i$  and  $j$  are either identical by descent (IBD) or independent [16]. The probability in the first case is the coefficient of consanguinity (CC)  $G_{ij}$  (appendix A), and in this case the covariance will be the genic variance  $\text{var}[x]$ , and in the second case the covariance is zero. Thus,

$$\text{cov}[x_j, x_i] = G_{ij} \text{var}[x] \quad (6.3)$$

and

$$E[\Delta\bar{x}] = E\left(\sum_j b_{ij} G_{ij}\right) \text{var}[x] = E\left(\sum_i b_{ij} G_{ij}\right) \text{var}[x]. \quad (6.4)$$

In the middle term of equation (6.4), the  $j$ -summation ranges over the actors and the final expectation is over all recipients  $i$ . The final term moves towards an inclusive-fitness formulation by reversing them so that the summation ranges over the recipients and the expectation is taken over the actors [17,19].

The  $G_{ij}$  are sometimes used as coefficients of relatedness, and indeed in infinite population models they do play that role, though under diploidy they should be normalized by dividing by  $G_{jj}$  the CC of the actor with itself [20]. However in finite population models, they should not, strictly speaking, be considered as relatedness coefficients. For one thing, they are always greater than or equal to 0, and in finite populations altruism and spite are often studied [21] and negative relatedness plays an important conceptual role. In the literature, there seems to be a diversity of candidates for relatedness in finite population models [4,22–24]. Such formulations subtract from

$G$  an average value calculated over some reference population, so that the relatedness of actor  $j$  to recipient  $i$  has the form

$$R_{ij} = \frac{G_{ij} - \bar{G}}{1 - \bar{G}}, \quad (6.5)$$

where the denominator ensures that the focal relatedness to itself is 1. This  $G$ -average might be calculated over the whole population or over an ‘economic neighbourhood’ [25]—the set of nodes that experiences the negative effects of the primary pay-offs. With a graph structure, such a node set is not so easy to get hold of, as the secondary effects are not uniform. In a transitive population, a standard normalization takes  $\bar{G}$  to be the average CC of the focal actor to the whole population; an advantage of that is that the focal node has average relatedness zero to the population.

As we have assumed the population size is constant, the sum of all fitness effects on any actor  $j$  must be zero:  $\sum_i b_{ij} = 0$ , and we can (except for the denominator in equation (6.5)) replace  $G_{ij}$  in equation (6.4) by  $R_{ij}$ . We define the inclusive-fitness effect to be

$$W_{\text{IF}} = E\left(\sum_i b_{ij} R_{ij}\right), \quad (6.6)$$

and then from equation (6.4)

$$W_{\text{IF}} = \frac{E[\Delta\bar{x}]}{(1 - \bar{G})\text{var}[x]} \quad (6.7)$$

so that the inclusive-fitness effect has the same sign as  $E[\Delta\bar{x}]$ , the long-term average one-step change in allele frequency.

As above, in a transitive population, the summation in the round brackets of equation (6.6) will be the same for each node  $j$ , and we can take a focal actor and eliminate the expectation:

$$W_{\text{IF}} = \sum_i b_i R_i. \quad (6.8)$$

Here  $b_i$  is the fitness effect of the focal actor on recipient  $i$  and  $R_i$  is the probability the focal actor is IBD to recipient  $i$ . We now look for an analogue of equation (6.8) in a heterogeneous population.

## 7. Heterogeneous populations: reproductive value

The RV of a node is roughly defined as its expected contribution to the future gene pool of the population. Fisher [26] was perhaps the first to understand that individuals with different RV need to be treated as belonging to different ‘reproductive classes.’ As an illustration, he produced a plot of RV against age in humans (p. 28) and pointed out that ‘The value shown is probably correct . . . for assessing how far it is worthwhile to give assistance to immigrants in respect of infants . . . for such infants will usually emigrate with their parents’ (p. 29).

Possibly the first appearance of RV in a calculation of inclusive fitness is found in Hamilton’s [27] work with altruistic behaviour in eusocial species, in which, in the case of the haplodiploid genetic system, the diploid females have twice the RV of the haploid males. In this paper, Hamilton corrects (p. 204) an error he made in his 1967 paper on extreme sex ratios in which a neglect of this RV ratio led him to an incorrect value for the evolutionarily stable strategy sex allocation under haplodiploidy (p. 485). To our knowledge, this is the first example of an inclusive-fitness calculation in a heterogeneous population in which the wrong result is obtained if RVs are not used.

In the post-Hamiltonian world, it was assumed, without much in the way of formality, that inclusive-fitness

arguments need to take account of variations in RV—after all, it is surely the future gene pool that counts [28]. As an ‘obvious’ example, consider a mother feeding her nestlings giving total benefit  $b$  at cost  $c$ . Suppose that half her nestlings will die. Suppose that we are at the point where the mother can tell which half will die. Then, we might imagine two alternative strategies: feed only those who will live and feed only those who will die. An inclusive-fitness calculation that did not take account of RV would provide the same condition for both cases, but clearly they have very different implications for gene-frequency increase. Later (tables 3 and 4), we provide specific results for the  $N = 3$  star graph.

Another early example is found in West Eberhard’s [29] suggestion that altruistic acts by individuals of low RV towards related individuals of high RV are more likely to be selected for than is altruism in the reverse direction. This is taken from Charlesworth & Charnov [30] who provided the first rigorous connection between RV-weighted fitness effects and gene-frequency increase with an inclusive-fitness model in an age-structured population. A nice example of the deployment of RVs in a class-structured population (in which there are breeders, helpers and waiters) is found in Pen & Weissing [31].

In the more recent graph-structured evolutionary models, the class structure derives not so much from different roles (male–female, parent–offspring, breeder–worker) as different edge configurations, but the same general considerations apply. To define these classes, we return to our concept of graph isomorphism, and define the reproductive classes to be the orbits under the set of all graph isomorphisms. That is, two nodes are in the same class if there is an isomorphism mapping one to the other. Thus, a transitive graph has only one reproductive class; otherwise, there is more than one class and the graph is called heterogeneous (figure 3).

Our rough definition above of RV seems to neglect mutation. In our finite graph model, the long-term future of a population without mutation is simply fixation of one of the current genes, and thus our definition of RV is equivalent to fixation probability. Place an allele  $A$  at node  $i$  with  $B$  at all other nodes. Let the population run with neutral selection (all pay-offs zero) and no mutation. Then, the RV  $v_i$  of node  $i$  is the probability that the population will become fixed in a pure  $A$  state. With this definition,  $v_i$  are normalized to have sum 1.

It should be clear that nodes in the same reproductive class will have the same RV, but the converse is not true. In fact, it has been shown [34] that the RV of a node is proportional under DB updating to its degree and under BD updating to the reciprocal of its degree. Note that this makes qualitative sense. Under DB, it is good for a node to have high degree, as this provides many opportunities for reproduction; conversely, under BD, it is good for a node to have low degree, as this reduces the probability of death. But note that degree does not determine class, as seen, for example, in the Frucht graph [33] (figure 3c). Here, all nodes have the same degree, and hence the same RV but they are all in different reproductive classes. The same behaviour at two such nodes will have the same effect on the population allele frequency, but in other ways can affect the population differently.

## 8. The heterogeneous price equation

Now let us return to the Price equation  $\Delta\bar{x} = \text{cov}(W, x)$  for the one-step selective allele-frequency change. We must first

**Table 3.** Calculation of the inclusive-fitness effect on the 3-star with BD updating and fecundity payoffs. Reproductive values are proportional to the reciprocal of the degree of the node so that  $v_H = 1/5$ ,  $v_L = 2/5$ . Relatedness coefficients are calculated in appendix A.

pay-off	primary		secondary		RV
	recipient	relatedness	recipient	relatedness	
hub actor (node 0)					
$-2c$	0	$R_0$	1	$R_1$	$v_L$
$b$	1	$R_1$	0	$R_0$	$v_H$
$b$	2	$R_1$	0	$R_0$	$v_H$
$W_{IF}(H) = 2b(R_1 - R_0)v_H - 2c(R_0 - R_1)v_L = (-21b - 42c)/40$					
leaf actor (node 1)					
$-c$	1	$R_0$	0	$R_1$	$v_H$
$b$	0	$R_1$	0,2	$(R_0 + R_2)/2$	$v_L$
$W_{IF}(L) = b\left(R_1 - \frac{R_0 + R_2}{2}\right)v_L - c(R_0 - R_1)v_H = (-12b - 21c)/80$					
$W_{IF} = W_{IF}(H) + 2W_{IF}(L) = -\frac{3}{40}u(11b + 21c)$					

**Table 4.** Calculation of the inclusive-fitness effect on the 3-star with DB updating and fecundity payoffs. Reproductive values are proportional to the degree of the node so that  $v_H = 2/4$ ,  $v_L = 1/4$ . Relatedness coefficients are calculated in appendix A.

pay-off	primary		secondary		RV
	recipient	relatedness	recipient	relatedness	
hub actor (node 0)					
$-2c$	0	$R_0$	0	$R_0$	$v_L$
$b$	1	$R_1$	1,2	$R_1$	$v_H$
$b$	2	$R_1$	1,2	$R_1$	$v_H$
$W_{IF}(H) = 2b(R_1 - R_1)v_H - 2c(R_0 - R_0)v_L = 0$					
leaf actor (node 1)					
$-c$	1	$R_0$	0,2	$(R_0 + R_2)/2$	$v_H$
$b$	0	$R_1$	1	$R_1$	$v_L$
$W_{IF}(L) = b(R_1 - R_1)v_L - c\left(R_0 - \frac{R_0 + R_2}{2}\right)v_H = -c\left(\frac{R_0 - R_2}{2}\right)v_H = -\frac{1}{2}c$					
$W_{IF} = W_{IF}(H) + 2W_{IF}(L) = -c$					

ask, in a heterogeneous graph, what we should even mean by allele frequency. A gene on a more valuable node will contribute many more copies of itself to the future than a gene on a less valuable node. It would thus appear that any measure of the 'effective' frequency of an allele  $A$  should take account of the value of the nodes it occupies. That has prompted the definition of *RV-weighted allele frequency* in any state  $S$  as  $\hat{x} = \hat{x}_S = \sum_i v_i x_i$  [35–37]. A closely related way to realize the vector  $\langle \hat{x}_S \rangle$  is as a left eigenvector. Regard  $\hat{x}_S$  as the total RV of all  $A$ -genes in state  $S$ . Under neutrality, any copy of  $A$  must hold its RV over each update step, and that gives us a system of forward recursive equations for  $\hat{x}_S$ . Now those equations, when framed in vector-matrix form, tell us that  $\langle \hat{x}_S \rangle$  is the left eigenvector of the state transition matrix  $A$  at neutrality [35].

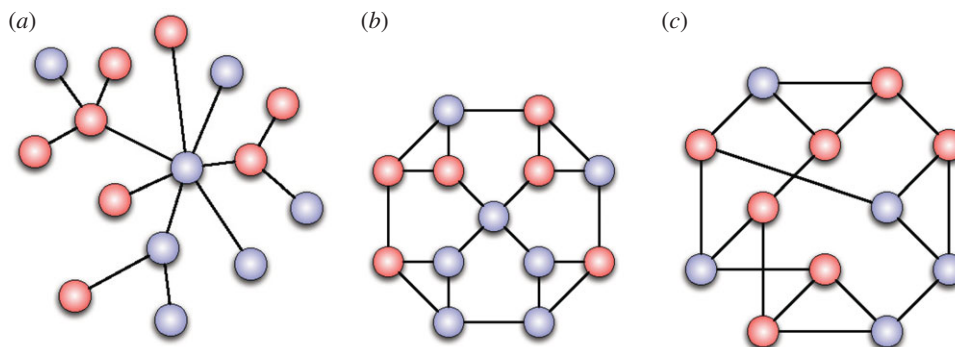
To calculate the one-step selective change in  $\hat{x}$ , fitness effects must be RV-weighted, that is, when an individual dies the RV of its node is lost and when an offspring is born the RV of the node it occupies is gained. Using  $\hat{W}$  for this RV-weighted form of fitness, Price's formula for average RV-weighted allele-frequency change is

$$\Delta \hat{x} = \text{cov}(\hat{W}, x). \quad (8.1)$$

This formula calculates the one-step selective change in RV-weighted allele frequency of the population in any particular state. Its long-term average is

$$E[\Delta \hat{x}] = E[\text{cov}(\hat{W}, x)] \quad (8.2)$$

and this is the general version of equation (5.4).



**Figure 3.** Three examples of heterogeneous graph-structured populations. (a) In contrast to the homogeneous structures considered previously, many authentic social structures are known to have a few highly connected individuals while the majority of individuals have far fewer connections. Such models can be generated by a preferential attachment algorithm [32]. Nodes with different numbers of connections will certainly be in different reproductive classes. However this is far from being the only distinguishing characteristic—for two nodes  $i$  and  $j$  to be in the same class, there must be an isomorphism of the graph mapping  $i$  to  $j$ . (b) The central node has degree 4 and the 12 others have degree 3. There are three reproductive classes, eight nodes in the outer octagon, four nodes forming the middle square and the central node. The symmetry in the drawing of the graph provides the isomorphisms. Thus, nodes may have the same number of connections but may relate to the population as a whole differently. (c) An extreme example of this is represented by the Frucht graph [33] with 12 nodes. All nodes have degree 3, yet this structure has no isomorphism (other than the identity); thus, each node is its own class. (Online version in colour.)

## 9. Heterogeneous inclusive fitness

Our transformation above of Price's formula into an inclusive-fitness effect will also work with RV-weighted allele frequency. In particular, the argument in appendix B which leads to equation (5.5) remains valid with  $W$  replaced by  $\hat{W}$ . The analogue of equation (6.4) is

$$E[\Delta \hat{x}] = E\left(\sum_i \hat{b}_{ij} G_{ij}\right) \text{var}[x], \quad (9.1)$$

where  $\hat{b}_{ij}$  is the effect of an  $A$ -allele at node  $j$  on the RV-weighted fitness of individual  $i$ , and, following equation (6.6), the inclusive-fitness effect is more generally defined as

$$W_{\text{IF}} = E\left(\sum_i \hat{b}_{ij} R_{ij}\right), \quad (9.2)$$

where  $R_{ij}$  is again given by equation (6.6) with  $\bar{G}$  some suitable state-independent average CC value. We return to this in the Discussion.

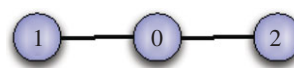
Now we generalize equation (6.8). In a heterogeneous population, the expectation in equation (9.2) will be the same for nodes  $j$  in the same reproductive class so that we can in fact take the expectation over each class of actor using the relative class size as a weight. We get

$$W_{\text{IF}} = \sum_J \frac{N_J}{N} \sum_i \hat{b}_{ij} R_{ij}, \quad (9.3)$$

where  $N_J$  is the size of class  $J$ ,  $\hat{b}_{ij}$  is effect of an  $A$ -allele at a class  $J$  focal node on the RV-weighted fitness of individual  $i$ , and  $R_{ij}$  is the relatedness of the class  $J$  focal actor to recipient  $i$ . One further comment: the  $J$ -sum in equation (9.3) needs only be taken over these classes, which exhibit the behaviour being studied. We return to this in the Discussion.

## 10. Example. The star graph with $N = 3$ nodes

The graph (figure 4) has a hub and two leaves; the hub is connected to both leaves such that in each time step the hub has two interactions and each leaf has one. Offspring dispersal from a leaf is only to the hub and from the hub is to each leaf with neutral probability 1/2. There are clearly two reproductive classes, class  $H$  containing the hub and class  $L$



**Figure 4.** The star graph on three vertices. The single hub node 0 interacts with the two leaf nodes 1 and 2, and vice versa; there is no primary interaction between the leaf nodes. Offspring from either leaf node disperse to the hub, while hub offspring disperse to either leaf with probability 1/2 (tables 3 and 4). (Online version in colour.)

containing the two leaves. In tables 3 and 4, we present the inclusive-fitness analyses for both BD and DB updating. Both follow the pattern found in the homogeneous case of table 1 with a few notable differences.

- *Focal actors.* As we have two reproductive classes, we need two focal actors, a hub actor and a leaf actor. We do the analysis for each actor class separately and then add the results with each class weighted by its size (the number of actors in the class).
- *RV.* Fitness effects are weighted with RV—a death weighted by the RV of its node and a birth weighted by the RV of the node the offspring occupies. The way we have set up the calculations facilitates this weighting: each fitness effect occupies its own line and displays both the primary and secondary relatedness coefficients. The RVs differ for BD and DB updating, being proportional to either the degree (DB) or the reciprocal of the degree (BD) of the node [34].
- *Relatedness.* Perhaps the first thing to note is that the relatedness coefficients change between BD and DB updating. This is not the case in a homogeneous graph. The difference can be seen in the structure of the recursive equations. To form a recursive equation for  $G_{ij}$ , the CC between nodes  $i$  and  $j$ , we look at where the most recent offspring replacement came from. In a homogeneous graph, the symmetry between nodes  $i$  and  $j$  means that we need only look at one of them. But in a heterogeneous graph, not only will the new CC depend on which node is replaced (dies) but the death probability can also be different between the



nodes. Take for example, the CC  $G_1$  between hub and leaf. Under DB (table 4), the two nodes have equal probability of death, but under BD (table 3) the hub is replaced four times as often as is the leaf, giving us probabilities of 4/5 and 1/5 in the  $G_1$  recursion.

- *Secondary effects.* As is the case for homogeneous graphs, when the primary effects are on fecundity, the secondary effects act on survival under BD and on fecundity under DB.

## 11. Discussion

### (a) Class-specific behaviour

In early models of heterogeneous populations, the classes were role based and actors typically belonged to only one class. More precisely, even when studying interaction between classes the behaviours belonged to different sets of genes. For example, in parent–offspring conflict the genes for offspring begging were different from the genes for parental resistance. In studying the coevolution of these traits, we would work with two inclusive-fitness effects,  $W_{IF}(P)$  and  $W_{IF}(O)$ , each depending upon the level of behaviour of both, and we would set them both to zero and solve this system for the equilibrium for both behaviours. Thus, the problem of how to combine the two effects into one did not arise. Indeed, both the general class-structured analysis in Taylor & Frank [17] and the model of helping at the nest of Pen & Weissing [31] assumed that there was only one actor class, although recipients belonged to different classes. Even in more general studies of altruism, one might generally suppose that the altruism of a parent was different from that of an offspring.

In evolutionary graph studies, it has been often the case that all nodes had the same behavioural repertoire and this was determined at the same locus or loci. In that case, the inclusive-fitness effects belonging to different actor classes would be added together, weighted by class size (equation (9.3)) and that is the assumption presented in tables 3 and 4. One can of course suppose that facultative behaviour is possible and a hub-dweller might want to use a different level of altruism (or spite) from a leaf-dweller. In that case, we would have a system of two interacting equations as in the parent–offspring example above.

### (b) The calculation of $W_{IF}$

The calculations we have presented here are ‘heuristic’ and a careful analysis is needed to justify them. For example, our analysis presented in table 1 of the 5-cycle (figure 1) worked directly with the fecundity effects  $b$  and  $c$  rather than using the realized fecundity (the probability of reproducing). Is that valid? Suppose that with BD updating node 1 gets a fecundity gift of  $b$  and all others have only their neutral fecundity. In that case, the realized fecundity (probability of reproducing) of node 1 is  $(1+b)/(5+b) \approx (1+4b)/5$  to first order in  $b$ , while the other four nodes have realized fecundity  $(1-b)/5$ , again to first order. Thus, the effective fecundity increase to node 1 is only  $4b/5$  while the others have an effective loss of  $b/5$ . The adjustment can be regarded as a secondary *fecundity* effect of the original gift, but in our analysis of table 1 we do not bother to account for that. In fact it turns out that, in the analysis, one does not need to take account of effects which apply equally to all individuals in the population, but one needs to ‘know’ that this is the case

or else at some point to do a calculation to verify it. It turns out that in a heterogeneous population this principle, that one does not need to account for effects that apply to all individuals, need *not* hold if fitness effects are not weighted by RV and it is the RV-weighting that allows us to ignore such effects [35]. The take-home message is that there is a ‘lore’ surrounding the formulation of inclusive fitness and one must think carefully. Hamilton himself (PDT Hamilton 1987, personal communication) said that he is not absolutely sure an inclusive-fitness calculation is correct until he can check his calculations by making them another way.

### (c) Relatedness

What should we use for coefficients of relatedness in a heterogeneous graph? The problem is what to choose for  $\bar{G}$  in equation (6.5). In a homogeneous population, a standard choice is the average CC of a focal node to the whole population, but in a heterogeneous population this will generally be different for focal nodes in different classes. If the behaviour is expressed by only one class  $K$ , then the specification:

$$R_{ij} = \frac{G_{ij} - \bar{G}_K}{1 - \bar{G}_K}, \quad (11.1)$$

for  $j$ , a focal class  $K$  node [37], will serve well where  $\bar{G}_K$  is the average CC of a focal node of class  $K$  to the whole population. In this case, the average relatedness of a class  $K$  actor to the whole population will be zero. But if there are several classes acting from the same allele, we cannot use these class-specific coefficients and still combine the contributions of different classes of actors in the simple additive way given in tables 3 and 4. A reasonable solution is to use the average of the  $\bar{G}_K$  and define

$$\bar{G} = \sum_K \frac{N_K \bar{G}_K}{N},$$

where  $N_K$  is the size of class  $K$ . This is what we have done for the calculations presented in tables 3 and 4 (appendix A).

### (d) Inclusive-fitness effect and fixation probability

The fixation probability  $\rho_A$  (or  $\rho_B$ ) of the allele  $A$  (or  $B$ ) is the probability that an  $A$ -mutation (or  $B$ -mutation) arising in a pure  $B$ -state (or  $A$ -state) will colonize the next pure state. In a homogeneous (transitive) population, it has been shown [12,22] that the inclusive-fitness effect of the allele  $A$  has the same sign as the difference  $\rho_A - \rho_B$  in the fixation probabilities, but as yet no parallel result has been obtained for heterogeneous populations. Tarnita and Taylor have a forthcoming paper that investigates this interesting question.

### (e) Inclusive fitness at age 50

The theoretical development of inclusive fitness over the past 50 years has enhanced and sharpened many of its positive features and has clarified its limitations. The essence of the method is to turn the standard (direct) way of calculating allele-frequency change inside–out. For example, Price’s equation (equations (5.1) and (8.1)) adds up the fitness effects on a focal allele of the behaviour of many actors, whereas inclusive fitness adds up the behavioural effects of a focal allele on the fitness of many recipients, each effect weighted by a measure of genetic closeness to the actor. A powerful consequence of

this point of view is its construction of a ‘maximizing agent’ [38], whereby an actor is able to attach to its own fitness various fractions of the effects of a proposed action, and thereby obtain a quantity that its behaviour should maximize.

For us, the most interesting challenge that inclusive fitness has faced in its 50-year journey has been the incorporation of population structure, and this is the main theme of our paper. Already in 1964 Hamilton speculated that population structure could well have a moderating effect on the fitness benefits and costs of a trait, particularly in what he called ‘viscous’ populations [1,39] but it was only many years later [25,40–42] that a more formal study of the ‘secondary’ effects of population structure began. Later still, a formal distinction between homogeneous and heterogeneous structures was made and it was realized that general results on evolutionary stability could be formulated for all homogeneous populations [13,16]. These results also suggested how a class-structure could be formally defined in abstract populations (graphs) that mirrored in some sense the traditional structure based on ‘roles’ (male–female, parent–offspring, breeder–helper) and that Fisher’s original concept of RV can be incorporated into the formulation of inclusive fitness such that it is able to predict changes in allele frequency under heterogeneity [17]. Currently, the effect of heterogeneity on the formulation and nature of inclusive fitness is one of its most active areas of investigation and it is giving us new insights into the evolutionary process.

**Acknowledgements.** The authors thank Corina Tarnita for insightful comments that helped in shaping this work. Two anonymous reviewers provided hard-hitting and well-deserved valuable comments.

**Funding statement.** P.D.T acknowledges support from the Natural Sciences and Engineering Research Council of Canada.

## Appendix A. Relatedness

The coefficient of consanguinity (CC) between two nodes is the probability that they are identical by descent (IBD), that is, that they are descended from a common ancestor without an intervening mutation. To first order in the mutation rate  $u$ , these are expected to have the form  $G = 1 - gu$ , where  $g$  is the coefficient of the first-order term in  $u$ , as when  $u = 0$ , the population will drift to a state in which all nodes are IBD.

A significant property of  $G$  is that it can be readily calculated recursively and examples are found in tables 1, 3 and 4. (But see Maciejewski [43] for an interesting alternative approach that uses established results from the theory of random walks.) The equations are obtained by asking for the  $G$ -coefficient just before the most recent replacement affecting the pair. For example, for  $G_1$  in the 5-cycle (below), the replacement offspring came from one of the pair or from a neighbour on the other side, each with probability 1/2 such that the average previous coefficient was  $(G_0 + G_2)/2$ . Of course, that applies only if the replacement did not mutate, and in that case the coefficient is zero.

An important remark is that these recursive equations do not take account of small selective variations in fecundity or survival, that is, they assume a neutral distribution of alleles. That introduces an error in the coefficients, which is first order in the pay-offs  $b$  and  $c$ , but it is clear from the form of the equation for the inclusive-fitness effect  $W_{IF}$  that the resulting error will then be of *second* order and can be ignored.

### (a) The 5-cycle (figure 1 and table 1)

Let  $G_k$  (and  $R_k$ ) be the CC (and relatedness) between nodes at distance  $k$ . We have the recursive equations

$$G_1 = (1 - u) \left( \frac{G_0 + G_2}{2} \right) \quad \text{and} \quad G_2 = (1 - u) \left( \frac{G_1 + G_2}{2} \right).$$

These solve to give  $G_1 = 1 - 4u$  and  $G_2 = 1 - 6u$ .

Then,

$$\bar{G} = \left( \frac{G_0 + 2G_1 + 2G_2}{5} \right) = 1 - 4u.$$

Finally,  $R_i = (G_i - \bar{G}) / (G_0 - \bar{G})$ , giving us  $R_0 = 1$ ,  $R_1 = 0$ ,  $R_2 = -1/2$ .

### (b) The $N = 3$ star (figure 4)

Let  $G_1$  (and  $R_1$ ) be the hub–leaf CC (and relatedness) and  $G_2$  (and  $R_2$ ) be the CC (and relatedness) between leaves.

#### (i) Birth–death updating (table 3)

The recursive equations are

$$G_1 = (1 - u) \left( \frac{4G_0 + G_2}{5} + \frac{1}{5}G_0 \right) \quad \text{and} \quad G_2 = (1 - u)G_1.$$

These solve to give  $G_1 = 1 - (7/3)u$  and  $G_2 = 1 - (10/3)u$ .

Then, the average CCs are

$$\text{for the hub: } \bar{G}_H = \left( \frac{G_0 + 2G_1}{3} \right) = 1 - \frac{14}{9}u,$$

$$\text{for the leaf: } \bar{G}_L = \left( \frac{G_0 + G_1 + G_2}{3} \right) = 1 - \frac{17}{9}u,$$

$$\text{giving an overall average: } \bar{G} = \left( \frac{\bar{G}_H + 2\bar{G}_L}{3} \right) = 1 - \frac{16}{9}u.$$

Finally,  $R_i = (G_i - \bar{G}) / (G_0 - \bar{G})$ , giving us  $R_0 = 1$ ,  $R_1 = -5/16$ ,  $R_2 = -7/8$ .

#### (ii) Death–birth updating (table 4)

The recursive equations are

$$G_1 = (1 - u) \left( \frac{1G_0 + G_2}{2} + \frac{1}{2}G_0 \right) \quad \text{and} \quad G_2 = (1 - u)G_1.$$

These solve to give  $G_1 = 1 - (5/3)u$  and  $G_2 = 1 - (8/3)u$ .

Then, the average CCs are

$$\text{for the hub: } \bar{G}_H = \left( \frac{G_0 + 2G_1}{3} \right) = 1 - \frac{10}{9}u,$$

$$\text{for the leaf: } \bar{G}_L = \left( \frac{G_0 + G_1 + G_2}{3} \right) = 1 - \frac{13}{9}u,$$

$$\text{giving an overall average: } \bar{G} = \left( \frac{\bar{G}_H + 2\bar{G}_L}{3} \right) = 1 - \frac{4}{3}u.$$

Finally,  $R_i = (G_i - \bar{G}) / (G_0 - \bar{G})$ , giving us  $R_0 = 1$ ,  $R_1 = -1/4$ ,  $R_2 = -1$ .

## Appendix B

Let fitness  $W$  and genotype  $x$  depend on both node  $i$  and population state  $S$ . We use round brackets to represent an average or a covariance over all nodes in the population, and square brackets to represent an average or a covariance over all population states *endowed with the neutral distribution*.

**Theorem B.1.** Suppose that for each node the genotype value averaged over all states with the neutral distribution is  $p_N$ :  $E[x] = p_N$ , and in each state  $S$ , the average fitness effect is zero:  $E(W) = 0$ . The latter will hold in a population of constant size. Then,

$$E[\text{cov}(W, x)] = E(\text{cov}[W, x]).$$

Proof:

$$\begin{aligned} E[\text{cov}(W, x)] &= E[E(Wx) - E(W)E(x)] \\ &= E[E(Wx)] \quad \text{as in each state } E(W) = 0 \end{aligned}$$

$$\begin{aligned} &= E(E(Wx)) \\ &= E(\text{cov}[W, x] + E[W]E[x]) \\ &= E(\text{cov}[W, x] + E[W]p_N) \quad \text{since } E[x] = p_N \text{ for every node} \\ &= E(\text{cov}[W, x]) + p_N E(E[W]) \\ &= E(\text{cov}[W, x]) + p_N E(E[W]) \\ &= E(\text{cov}[W, x]) \quad \text{as again, } E(W) = 0. \quad \blacksquare \end{aligned}$$

## References

- Hamilton WD. 1964 The genetical evolution of social behaviour I. *J. Theor. Biol.* **7**, 1–16. (doi:10.1016/0022-5193(64)90038-4)
- West S, Gardner A. 2010 Altruism, spite, and greenbeards. *Science* **327**, 1341–1344. (doi:10.1126/science.1178332)
- Grafen A. 2006 A theory of Fisher's reproductive value. *J. Math. Biol.* **53**, 15–60. (doi:10.1007/s00285-006-0376-4)
- Lehmann L, Rousset F. 2010 How life history and demography promote or inhibit the evolution of helping behaviours. *Phil. Trans. R. Soc. B* **365**, 2599–2617. (doi:10.1098/rstb.2010.0138)
- Lieberman E, Hauert C, Nowak MA. 2005 Evolutionary dynamics on graphs. *Nature* **433**, 312–316. (doi:10.1038/nature03204)
- Wild G, Traulsen A. 2007 The different limits of weak selection and the evolutionary dynamics of finite populations. *J. Theor. Biol.* **247**, 382–390. (doi:10.1016/j.jtbi.2007.03.015)
- Moran PAP. 1958 Random processes in genetics. *Proc. Cambridge Phil. Soc.* **54**, 60–71. (doi:10.1017/S0305004100033193)
- Moran PAP. 1962 *The statistical processes of evolutionary theory*. Oxford, UK: Oxford University Press.
- Taylor PD. 2010 Birth–death symmetry in the evolution of a social trait. *J. Evol. Biol.* **23**, 2569–2578. (doi:10.1111/j.1420-9101.2010.02122.x)
- Ohtsuki H, Nowak MA. 2006 Evolutionary games on cycles. *Proc. R. Soc. B* **273**, 2249–2256. (doi:10.1098/rspb.2006.3576)
- Grafen A. 2007 An inclusive fitness analysis of altruism on a cyclical network. *J. Evol. Biol.* **20**, 2278–2283. (doi:10.1111/j.1420-9101.2007.01413.x)
- Taylor PD, Day T, Wild G. 2007 Evolution of cooperation in a finite homogeneous graph. *Nature* **447**, 312–316. (doi:10.1038/nature05770)
- Grafen A, Archetti M. 2008 Natural selection of altruism in inelastic viscous homogeneous populations. *J. Theor. Biol.* **252**, 649–710. (doi:10.1016/j.jtbi.2008.01.021)
- Ohtsuki H, Hauert C, Lieberman E, Nowak MA. 2006 A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505. (doi:10.1038/nature04605)
- Lehmann L, Keller L, Sumpter DJT. 2007 The evolution of helping and harming on graphs: the return of the inclusive fitness effect. *J. Evol. Biol.* **20**, 2284–2295. (doi:10.1111/j.1420-9101.2007.01414.x)
- Taylor PD, Day T, Wild G. 2007 From inclusive fitness to fixation probability in homogeneous structured populations. *J. Theor. Biol.* **249**, 101–110. (doi:10.1016/j.jtbi.2007.07.006)
- Taylor PD, Frank S. 1996 How to make a kin-selection argument. *J. Theor. Biol.* **180**, 27–37. (doi:10.1006/jtbi.1996.0075)
- Price GR. 1970 Selection and covariance. *Nature* **227**, 520–521. (doi:10.1038/227520a0)
- Frank SA. 2013 Natural selection. VII. History and interpretation of kin selection theory. *J. Evol. Biol.* **26**, 1151–1184. (doi:10.1111/jeb.12131)
- Michod R, Hamilton WD. 1980 Coefficients of relatedness in sociobiology. *Nature* **288**, 694–697. (doi:10.1038/288694a0)
- Hamilton WD. 1970 Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**, 1218–1220. (doi:10.1038/2281218a0)
- Rousset F, Billiard S. 2000 A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *J. Evol. Biol.* **13**, 814–825. (doi:10.1046/j.1420-9101.2000.00219.x)
- Rousset F. 2002 Inbreeding and relatedness coefficients: what do they measure? *Heredity* **88**, 371–380. (doi:10.1038/sj.hdy.6800065)
- Lehmann L, Bargum K, Reuter M. 2006 An evolutionary analysis of the relationship between spite and altruism. *J. Evol. Biol.* **19**, 1507–1516. (doi:10.1111/j.1420-9101.2006.01128.x)
- Queller D. 1994 Genetic relatedness in viscous populations. *Evol. Ecol.* **8**, 70–73. (doi:10.1007/BF01237667)
- Fisher RA. 1930 *The genetical theory of natural selection*. Oxford, UK: Oxford University Press.
- Hamilton WD. 1972 Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.* **3**, 193–232. (doi:10.1146/annurev.es.03.110172.001205)
- Grafen A. 2006 Various remarks on Lehmann and Keller's article. *J. Evol. Biol.* **19**, 1397–1399. (doi:10.1111/j.1420-9101.2006.01161.x)
- West-Eberhard MJ. 1975 The evolution of social behaviour by kin selection. *Q. Rev. Biol.* **50**, 1–33. (doi:10.1086/408298)
- Charlesworth B, Charnov EL. 1981 Kin selection in age-structured populations. *J. Theor. Biol.* **88**, 103–119. (doi:10.1016/0022-5193(81)90330-1)
- Pen I, Wessing FJ. 2000 Sex ratio optimization with helpers at the nest. *Proc. R. Soc. Lond. B* **267**, 539–544. (doi:10.1098/rspb.2000.1034)
- Barabasi L, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)
- Frucht R. 1949 Graphs of degree three with a given abstract group. *Can. J. Math.* **1**, 365–378. (doi:10.4153/CJM-1949-033-6)
- Maciejewski W. 2014 Reproductive value in graph-structured populations. *J. Theor. Biol.* **340**, 285–293. (doi:10.1016/j.jtbi.2013.09.032)
- Leturque H, Rousset F. 2002 Dispersal, kin competition, and the ideal free distribution in a spatially heterogeneous population. *Theor. Popul. Biol.* **62**, 169–180. (doi:10.1006/tpbi.2002.1600)
- Taylor PD. 1990 Allele-frequency change in a class-structured population. *Am. Nat.* **135**, 95–106. (doi:10.1086/285034)
- Taylor PD. 2009 Decompositions of Price's formula in an inhomogeneous population structure. *J. Evol. Biol.* **22**, 201–213. (doi:10.1111/j.1420-9101.2008.01640.x)
- Grafen A. 2009 Formalizing Darwinism and inclusive fitness theory. *Phil. Trans. R. Soc. B* **364**, 3135–3141. (doi:10.1098/rstb.2009.0056)
- Hamilton WD. 1971 Selection of selfish and altruistic behaviour in some extreme models. In *Man and beast: comparative social behaviour* (eds JF Eisenberg, WS Dillon), pp. 57–91. Washington, DC: Smithsonian Press.
- Wilson DS, Pollock GB, Dugatkin LA. 1992 Can altruism evolve in purely viscous populations? *Evol. Ecol.* **6**, 331–341. (doi:10.1007/BF02270969)
- Taylor PD. 1992 Altruism in viscous populations: an inclusive fitness model. *Evolutionary Ecology* **6**, 352–356. (doi:10.1007/BF02270971)
- Taylor PD. 1992 Inclusive fitness in a homogeneous environment. *Proc. R. Soc. Lond. B* **249**, 299–302. (doi:10.1098/rspb.1992.0118)
- Maciejewski W. 2012 Resistance and relatedness on an evolutionary graph. *J. R. Soc. Interface* **9**, 511–517. (doi:10.1098/rsif.2011.0429)