

# Supporting appendix for the paper “Group selection and kin selection. Two concepts but one process”

## Inclusive fitness effect

### Probability of fixation

We first give a brief summary of the argument developed by Rousset (1; 2) to compute the first derivative of the probability of fixation of a new allele (say  $A$ ) with respect to its phenotypic effect. Call  $\pi$  the probability of fixation of allele  $A$ , and  $E[p(t)]$  the expected frequency of  $A$  at time  $t$ , where  $t = 0$  is the time at which the allele enters the population. We have:

$$\pi = E[p(\infty)] = p(0) + \sum_{t=0}^{\infty} E[p(t+1) - p(t)], \quad (1)$$

which can be expressed as:

$$\pi = p(0) + \sum_{t=0}^{\infty} \sum_{\mathbf{x}} E[\Delta p(t) | \mathbf{p}(t) = \mathbf{x}] \Pr[\mathbf{p}(t) = \mathbf{x}] \quad (2)$$

where  $\mathbf{p}(t)$  is a vector giving the frequency of allele  $A$  in all interaction groups (or demes, or families) at time  $t$ , the second sum is over all possible values  $\mathbf{x}$  that this vector may take,  $\Pr[\mathbf{p}(t) = \mathbf{x}]$  is the probability that  $\mathbf{p}(t)$  equals  $\mathbf{x}$ , and  $E[\Delta p(t) | \mathbf{p}(t) = \mathbf{x}]$  is the expected change in frequency of  $A$  between times  $t$  and  $t+1$  given that  $\mathbf{p}(t) = \mathbf{x}$ . Call  $\delta$  the phenotypic deviation of allele  $A$  relative to the phenotypic effect of the resident allele. When  $\delta = 0$ , the allele has the same phenotypic effect as the resident allele and is neutral, in which case  $E[\Delta p(t) | \mathbf{p}(t) = \mathbf{x}] = 0$ . Therefore, we have:

$$\left. \frac{d\pi}{d\delta} \right|_{\delta=0} = \sum_{t=0}^{\infty} \sum_{\mathbf{x}} \left. \frac{dE[\Delta p(t) | \mathbf{p}(t) = \mathbf{x}]}{d\delta} \right|_{\delta=0} \Pr^{\circ}[\mathbf{p}(t) = \mathbf{x}] \quad (3)$$

where  $\Pr^{\circ}[\mathbf{p}(t) = \mathbf{x}]$  is the probability that  $\mathbf{p}(t)$  equals  $\mathbf{x}$  in a neutral model ( $\delta = 0$ ). Allele  $A$  is selected for when its probability of fixation is greater than the probability of fixation of a neutral allele, namely, when  $d\pi/d\delta > 0$ .

### Example

We now consider a specific application of the result of the last section by assuming that the population is subdivided into  $n_g$  groups, each comprising  $N$  adult individuals so that

the population is of total constant size  $n_g N$ . We assume that migration can occur between groups but, for simplicity, there is no isolation by distance. Call  $z_{ij}$  the phenotype of individual  $j$  in group  $i$ ,  $z_i$  the average phenotype in group  $i$ , and  $z$  the average phenotype in the whole population. Without loss of generality, we can assume that individuals bearing allele  $A$  have phenotype  $\delta$ , while the carriers of the resident allele (say allele  $a$ ) have phenotype 0. Therefore, we have  $z_{ij} = \delta p_{ij}$ ,  $z_i = \delta p_i$ , and  $z = \delta p$ , where  $p_{ij}$ ,  $p_i$  and  $p$  are the frequencies of  $A$  in individual  $ij$ , in group  $i$  and in the whole population, respectively ( $p_{ij}$  equals 0 or 1 if individuals are haploid). The “fitness”  $w_{ij}$  of individual  $ij$  is defined as its expected number of offspring who will be part of the next adult generation;  $w_{ij}$  may depend on  $z_{ij}$ ,  $z_i$ , and  $z$ . Note that with such a definition of fitness, the average fitness equals one ( $E[w_{ij}] = 1$ ) and the expected change in frequency of  $A$  over one generation is given by (getting rid of  $t$  for simplicity):

$$E[\Delta p] = \frac{1}{n_g N} \sum_{i=1}^{n_g} \sum_{j=1}^N w_{ij} p_{ij} - p. \quad (4)$$

As  $w_{ij}$  depends on  $z_{ij}$ ,  $z_i$  and  $z$ , which themselves depend on  $\delta$ , one obtains (using the fact that  $p_{ij}^2 = p_{ij}$ ):

$$\begin{aligned} \left. \frac{dE[\Delta p]}{d\delta} \right|_{\delta=0} &= \frac{1}{n_g N} \sum_{i=1}^{n_g} \sum_{j=1}^N \left( \frac{\partial w_{ij}}{\partial z_{ij}} \frac{dz_{ij}}{d\delta} + \frac{\partial w_{ij}}{\partial z_i} \frac{dz_i}{d\delta} + \frac{\partial w_{ij}}{\partial z} \frac{dz}{d\delta} \right) p_{ij} \\ &= \frac{\partial w_{ij}}{\partial z_{ij}} p + \frac{\partial w_{ij}}{\partial z_i} \overline{p_i^2} + \frac{\partial w_{ij}}{\partial z} p^2, \end{aligned} \quad (5)$$

where  $\overline{p_i^2}$  is the average of  $p_i^2$  over all groups  $i$ , and where the partial derivatives are evaluated at  $z_{ij} = z_i = z = 0$ . Finally, using the fact that the sum of the partial derivatives of the fitness function sum up to zero (3), one obtains:

$$\left. \frac{dE[\Delta p]}{d\delta} \right|_{\delta=0} = \frac{\partial w_{ij}}{\partial z_{ij}} (p - p^2) + \frac{\partial w_{ij}}{\partial z_i} (\overline{p_i^2} - p^2). \quad (6)$$

From equations 3 and 6, one arrives at:

$$\left. \frac{d\pi}{d\delta} \right|_{\delta=0} = \frac{\partial w_{ij}}{\partial z_{ij}} (S - S_0) + \frac{\partial w_{ij}}{\partial z_i} (S_0 - S_1) \quad (7)$$

with

$$S = \sum_{t=0}^{\infty} E^\circ [p(t)], \quad S_0 = \sum_{t=0}^{\infty} E^\circ [\overline{p_i^2}(t)], \quad S_1 = \sum_{t=0}^{\infty} E^\circ [p^2(t)], \quad (8)$$

where  $E^\circ [x(t)]$  is the expected value of  $x$  at time  $t$  in the neutral case ( $\delta = 0$ ). In the case where allele  $A$  is present as a single copy in the whole population at time zero, these

expectations are given by:

$$\mathbb{E}^\circ [p(t)] = \frac{1}{n_g N}, \quad \mathbb{E}^\circ [\overline{p_i^2}(t)] = \frac{Q_0^R(t)}{n_g N}, \quad \mathbb{E}^\circ [p^2(t)] = \frac{Q_1^R(t)}{n_g N} \quad (9)$$

where  $Q_0^R(t)$  and  $Q_1^R(t)$  are the probabilities that the ancestral lineages of two homologous genes sampled with replacement from the same group, and from the whole population (respectively) at time  $t$  coalesce between time 0 and time  $t$ . These probabilities of identity between pairs of homologous genes are functions of the parameters describing population structure (e.g., number of groups, group size, migration rates). This finally gives:

$$\left. \frac{d\pi}{d\delta} \right|_{\delta=0} = \left( \frac{\partial w_{ij}}{\partial z_{ij}} + \frac{\partial w_{ij}}{\partial z_i} R \right) \frac{Z}{n_g N}, \quad (10)$$

where the first term in the parentheses measures the effect of the phenotype of an individual on its fitness, the second term measures the effect of the group phenotype on that individual's fitness and

$$R = \frac{\sum_{t=0}^{\infty} [Q_0^R(t) - Q_1^R(t)]}{\sum_{t=0}^{\infty} [1 - Q_1^R(t)]} \quad (11)$$

is the coefficient of relatedness measuring by how much two individuals sampled with replacement from the same group are more related than two individuals sampled randomly from the whole population. Lastly,

$$Z = \sum_{t=0}^{\infty} [1 - Q_1^R(t)] \quad (12)$$

is a constant of proportionality that is always positive and depends on the demographic assumption on the population (e.g., number of groups, group size, migration rates). The effect of the mutant on its probability of fixation  $d\pi/d\delta$  fits within Hamilton's definition of the inclusive fitness effect (4), which is a weighted sum of the effects of the phenotypes of different actors on the fitness of a focal individual. The probability of fixation of allele  $A$  is greater than the probability of fixation of a neutral allele if the inclusive fitness effect is positive, that is, when

$$\frac{\partial w_{ij}}{\partial z_{ij}} + \frac{\partial w_{ij}}{\partial z_i} R > 0. \quad (13)$$

## T&N's model

### Inclusive fitness effect and T&N's model

In order to apply the previous formalism to T&N's model, it proves convenient to measure a "generation" by the time between two migration or group division events because fixation

within groups occurs on a faster time scale compared to migration and group division. More precisely, time 0 corresponds to the time at which allele  $A$  enters the population (as a single copy), while times 1, 2, 3, ... are the times just before reproduction events that will cause either a migration or a group division. Therefore, at any time unit all groups are genetically homogeneous, and all lineages from the same group coalesce during time  $t$  and time  $t+1$ . The method described previously can be applied here, the only difference being that we have to treat the first time interval separately: indeed, during time 0 and time 1, coalescence occurs within all groups (so that the mutant allele either fixes within its group or disappears), but no migration or group division occurs, while during all other time intervals, one migration or group division event occurs. Therefore, equation 10 becomes:

$$\left. \frac{d\pi}{d\delta} \right|_{\delta=0} = \left( \frac{\partial w_{ij}^p}{\partial z_{ij}} + \frac{\partial w_{ij}^p}{\partial z_i} R \right) \frac{Z}{n_g N} + \left. \frac{dE[\Delta_0 p]}{d\delta} \right|_{\delta=0} \quad (14)$$

where  $w_{ij}^p$  is a fitness function giving the expected number of individuals descending from individual  $ij$  (possibly including himself) after a reproductive event causing either a migration or a group division. Here, we have  $R = \sum_{t=1}^{\infty} [Q_0^R(t) - Q_1^R(t)] / \sum_{t=1}^{\infty} [1 - Q_1^R(t)]$ ,  $Z = \sum_{t=1}^{\infty} [1 - Q_1^R(t)]$ , and  $E[\Delta_0 p]$  is the expected change in frequency of allele  $A$  between time 0 and time 1 given by:

$$\begin{aligned} E[\Delta_0 p] &= \pi_g \left( \frac{1}{n_g} - \frac{1}{n_g N} \right) + (1 - \pi_g) \left( -\frac{1}{n_g N} \right) \\ &= \frac{1}{n_g} \pi_g - \frac{1}{n_g N} \end{aligned} \quad (15)$$

where  $\pi_g$  is the probability of fixation of a single allele at the level of a single group. The derivative of the probability of fixation of allele  $A$  with respect to its phenotypic effect  $\delta$  can be obtained by an application of the method exposed in the last section and by following T&N's assumption that the mutation appears in a group of size  $N$ . We then have:

$$\left. \frac{d\pi_g}{d\delta} \right|_{\delta=0} = \sum_{h=0}^{\infty} \sum_x \frac{dE[\Delta p(h) | p(h) = x]}{d\delta} \Big|_{\delta=0} \Pr^{\circ}[p(h) = x] \quad (16)$$

where  $p(h)$  is the frequency of allele  $A$  in the group at time  $h$  after its appearance, where  $h$  is the "fast" time scale of reproductive events within the group (i.e.  $h = 0, 1, \dots$  are the times just before the first, second... reproductive events within the group), which is small enough so that allele  $A$  fixes within the group before the first migration or group division. The second sum in eq. 16 is over all possible values  $x$  that the frequency of allele  $A$  may take in the group it appears in,  $\Pr[p(h) = x]$  is the probability that  $p(h)$  equals  $x$ , and

$E[\Delta p(h)|p(h) = x]$  is the expected change in frequency of  $A$  between times  $h$  and  $h + 1$  given that  $p(h) = x$ . Without loss of generality, call the group where the mutation arises group  $i$  and call  $w_{ij}^g$  the expected number of successful offspring of individual  $ij$  (possibly including himself) after a reproductive event that causes neither a migration nor a group division. Noting that the fitness of individual  $ij$  is affected only by its own phenotype  $z_{ij}$  and the average phenotype  $z_i$  in group  $i$ , the expected change in frequency of  $A$  over one generation is given by (getting rid of  $h$  for simplicity)

$$E[\Delta p] = \frac{1}{N} \sum_{j=1}^N w_{ij}^g p_{ij} - p_i. \quad (17)$$

As  $w_{ij}$  depends on  $z_{ij}$  and  $z_i$  only, which themselves depend on  $\delta$ , one obtains (using the fact that  $p_{ij}^2 = p_{ij}$ ):

$$\begin{aligned} \left. \frac{dE[\Delta p]}{d\delta} \right|_{\delta=0} &= \frac{1}{N} \sum_{j=1}^N \left( \frac{\partial w_{ij}^g}{\partial z_{ij}} \frac{dz_{ij}}{d\delta} + \frac{\partial w_{ij}^g}{\partial z_i} \frac{dz_i}{d\delta} \right) p_{ij} \\ &= \frac{\partial w_{ij}^g}{\partial z_{ij}} p_i + \frac{\partial w_{ij}^g}{\partial z_i} p_i^2, \end{aligned} \quad (18)$$

where the partial derivatives are evaluated at  $z_{ij} = z_i = 0$ . Finally, using the fact that the sum of the partial derivatives of the fitness function sum up to zero ( $\partial w_{ij}^g / \partial z_i = -\partial w_{ij}^g / \partial z_{ij}$ ) and following the same stream of calculations as in the previous section shows that

$$\left. \frac{d\pi_g}{d\delta} \right|_{\delta=0} = \frac{\partial w_{ij}^g}{\partial z_{ij}} \frac{Z_g}{N}, \quad (19)$$

with  $Z_g = \sum_{h=0}^{\infty} [1 - Q_0^R(h)]$ , where  $Q_0^R(h) = NE^\circ [p_i^2(h)]$  and the sum being over the “fast” time scale of reproductive events within the group.

With this result, the inclusive fitness effect (eq. 14) can be written as

$$\begin{aligned} \left. \frac{d\pi}{d\delta} \right|_{\delta=0} &= \frac{\partial w_{ij}^g}{\partial z_{ij}} \frac{Z_g}{n_g N} + \left( \frac{\partial w_{ij}^p}{\partial z_{ij}} + \frac{\partial w_{ij}^p}{\partial z_i} R \right) \frac{Z}{n_g N} \\ &= \frac{d\pi_g}{d\delta} \frac{1}{n_g} + \frac{d\pi_p}{d\delta} \frac{1}{N}, \end{aligned} \quad (20)$$

where

$$\left. \frac{d\pi_p}{d\delta} \right|_{\delta=0} = \left( \frac{\partial w_{ij}^p}{\partial z_{ij}} + \frac{\partial w_{ij}^p}{\partial z_i} R \right) \frac{Z}{n_g} \quad (21)$$

and  $\pi_p$  is the probability of fixation of the mutant at the level of the population conditional on having fixed at the level of a single group.

In order to recover T&N's result that  $d\pi/d\delta$  approximately equals  $1+(N/n_g)+(\lambda/q)$  when  $N$  and  $n_g$  are large, we now calculate explicitly the terms  $Z_g$ ,  $w_{ij}^g$ ,  $R$ ,  $Z$  and  $w_{ij}^p$  appearing in the inclusive fitness effect (eq. 20). In so doing, we follow exactly the assumptions that are implicit in eqs. 2-3 and eqs. 26-27 of T&N's and that are the following: 1) whenever reproduction occurs within a group without a migration or a group division, the new offspring replaces a randomly chosen individual from the group, 2) whenever a group division occurs, one of the two offspring group survives with probability one, while a group among the  $n_g$  remaining in the population is removed at random, 3) whenever a migration event occurs, migration is random to the  $n_g$  groups of the population, so that a migrant returns to its natal group with probability  $1/n_g$  and, finally, 4) a migrant replaces a randomly chosen individual from the group where it immigrates.

### Explicit functionals for fixation within the group

The term  $Z_g = \sum_{h=0}^{\infty} [1 - Q_0^R(h)]$ , where  $h$  is the “fast” time scale of reproductive events, can be obtained as follows:  $1 - Q_0^R(h)$  is the probability that the ancestral lineages of two homologous genes sampled with replacement from the same group at time  $h$  do not coalesce within the group between time 0 and time  $h$ . If the two genes are sampled from the same individual (probability  $1/N$ ), they coalesce immediately, and so:

$$[1 - Q_0^R(h)] = \left(1 - \frac{1}{N}\right) [1 - Q_0^D(h)], \quad (22)$$

where  $1 - Q_0^D(h)$  is the probability that the ancestral lineages of two homologous genes sampled from two different individuals in the same group at time  $h$  do not coalesce. A recursion for  $[1 - Q_0^D(h)]$  between pairs of genes is obtained as follows: the probability that two individuals sampled without replacement come from the same parent just before the last reproductive event is  $2(1/N)(1 - 1/N)1/(N - 1) = 2/N^2$ , where  $1/N$  is the probability of sampling the new individual produced by the reproductive event,  $1 - 1/N$  is the probability that the parent is still alive, and  $1/(N - 1)$  is the probability of sampling the parent. Hence:

$$[1 - Q_0^D(h)] = \left(1 - \frac{2}{N^2}\right) [1 - Q_0^D(h - 1)]. \quad (23)$$

Using the fact that  $1 - Q_0^D(0) = 1$ , one obtains after simplification:

$$Z_g = \frac{N(N - 1)}{2}. \quad (24)$$

It now remains to evaluate the fitness function  $w_{ij}^g$ , which gives the expected number of individuals descending from individual  $j$  in group  $i$  (possibly including  $ij$  himself) after a reproductive event that does not lead to a migration or group division. Remember that  $z_{ij} = \delta$  if individual  $ij$  carries allele  $A$ , while  $z_{ij} = 0$  if he carries allele  $a$ . Because we only need the first derivative of  $w_{ij}^g$  with respect to  $z_{ij}$  (taken in  $z_{ij} = z_i = 0$ ), it is sufficient to express  $w_{ij}^g$  to the first order in  $z_{ij}$  and  $z_i$ . From the assumptions on the costs and benefits of T&N's model, the probability that individual  $ij$  is chosen for reproduction in its group is:

$$\frac{1}{N} \left[ 1 - \left( c + \frac{b}{N-1} \right) (z_{ij} - z_i) \right] \quad (25)$$

and the probability that individual  $ij$  is not killed after the reproductive event is  $(N-1)/N$ . Combining these two terms, we obtain:

$$w_{ij}^g = \frac{N-1}{N} + \frac{1}{N} \left[ 1 - \left( c + \frac{b}{N-1} \right) (z_{ij} - z_i) \right]. \quad (26)$$

Inserting this equation and eq. 24 into eq. 19, yields:

$$\left. \frac{d\pi_g}{d\delta} \right|_{\delta=0} = -\frac{1}{2} \left( c + \frac{(b-c)}{N} \right). \quad (27)$$

### Explicit functionals for fixation in the whole population

The coefficient of relatedness  $R = \sum_{t=1}^{\infty} [Q_0^R(t) - Q_1^R(t)] / \sum_{t=1}^{\infty} [1 - Q_1^R(t)]$  equals one here, since at any time unit  $t$  between a migration or a group division (i.e., “large” time scale), the ancestral lineages of two genes sampled from the same group coalesce within groups with probability one (i.e.,  $Q_0^R(t) = 1$  for all  $t \geq 1$ ). The term  $Z = \sum_{t=1}^{\infty} [1 - Q_1^R(t)]$  can be obtained as follows:  $1 - Q_1^R(t)$  is the probability that the ancestral lineages of two homologous genes sampled with replacement from the whole population at time  $t$  do not coalesce between time 0 and time  $t$  (i.e., probability of non-identity). If the genes are sampled from the same group (probability  $1/n_g$ ), they will coalesce, and so:

$$[1 - Q_1^R(t)] = \left( 1 - \frac{1}{n_g} \right) [1 - Q_1^D(t)] \quad (28)$$

where  $1 - Q_1^D(t)$  is the probability that the ancestral lineages of two genes sampled from two different groups at time  $t$  do not coalesce. A recursion for the probability of non-identity  $[1 - Q_1^D(t)]$  between pairs of genes is obtained as follows: with probability  $q/(\lambda + q)$ , a group division occurred just after time  $t-1$ . With probability  $2/n_g^2$ , the two groups where the

genes are sampled at time  $t$  are the offspring groups of the group that has divided; in this case, the two lineages will coalesce. In the other case, the lineages coalesce with probability  $Q_1^D(t)(t-1)$ . Then, with probability  $\lambda/(\lambda+q)$ , a migration occurred just after time  $t-1$ . Since migration is random to any group (including the parental group), the probability that two genes sampled at random from the population after migration descend from the same group before migration is given by  $2/n_g^2$ . But now, we have to multiply  $2/n_g^2$  by the probability that a migrant fixes within a group (in the neutral case), which is  $1/N$ . We thus have:

$$[1 - Q_1^D(t)] = \left[ \frac{q}{\lambda+q} \left( 1 - \frac{2}{n_g^2} \right) + \frac{\lambda}{\lambda+q} \left( 1 - \frac{2}{n_g^2 N} \right) \right] [1 - Q_1^D(t-1)] \quad (29)$$

Using the fact that  $1 - Q_1^D(1) = 1$ , one obtains after simplification:

$$Z = \frac{n_g(n_g-1)(q+\lambda)N}{2(qN+\lambda)}. \quad (30)$$

We now evaluate the fitness function  $w_{ij}^P$ , which is the expected number of offspring present in the population at time  $t+1$  just before a reproductive event causing either a migration or a group division, and that descend from individual  $j$  in group  $i$  who is present in the population at time  $t$ , just before a reproductive event causing either a migration or a group division. Since all groups become genetically homogeneous between two such reproductive events, we have  $z_{ij} = z_i$ , and we can thus express  $w_{ij}^P$  as a function of  $z_{ij}$  and  $z$  only. Taking into account both migration and group division,  $w_{ij}^P$  can be expressed as:

$$w_{ij}^P = \frac{q}{\lambda+q} w_{ij}^Q + \frac{\lambda}{\lambda+q} w_{ij}^A \quad (31)$$

where  $w_{ij}^Q$  is the fitness of individual  $ij$  given that a group division occurs just after time  $t$ , while  $w_{ij}^A$  is the fitness of  $ij$  given that a migration occurs just after time  $t$ . Because we only need the first derivative of  $w_{ij}$  with respect to  $z_{ij}$  (taken in  $z_{ij} = z = 0$ ), it is sufficient to express  $w_{ij}^Q$  and  $w_{ij}^A$  to the first order in  $z_{ij}$  and  $z$ .

From the assumptions on the costs and benefits of helping of the model of T&N, the probability that individual  $ij$  is chosen for reproduction is given by (to the first order in  $z_{ij}$  and  $z$ ):

$$\frac{1}{n_g N} [1 + (b-c)(z_{ij} - z)]. \quad (32)$$

After reproduction, group  $i$  is made of  $N+1$  individuals. If the group splits, the two offspring groups survive regulation with probability  $1 - 1/n_g$ , in which case the two offspring



of individual  $ij$  will be the ancestors of all the individuals in the two offspring groups of group  $i$  (at the next time unit) with probability  $2/[(N+1)N]$  (in which case the fitness of  $ij$  is  $2N$ ), while only one of his offspring will be the ancestor of all the individuals in one of the two offspring groups of group  $i$  with probability  $4(N-1)/[(N+1)N]$  (in which case the fitness of  $ij$  is  $N$ ). With probability  $1/n_g$ , only one offspring group of group  $i$  survives regulation, in which case one of the two offspring of individual  $ij$  will be the ancestor of all the individuals in the group with probability  $2/(N+1)$  (in which case the fitness of  $ij$  is  $N$ ). Therefore, if individual  $ij$  is chosen for reproduction and if group  $i$  splits, the fitness of  $ij$  is  $2N(2n_g - 1)/[n_g(N+1)]$ .

The probability that another individual from group  $i$  is chosen for reproduction is (given that  $z_{ij} = z_i$ , and to the first order in  $z_{ij}$  and  $z$ ):

$$\frac{N-1}{n_g N} [1 + (b-c)(z_{ij} - z)]. \quad (33)$$

In that case, and if group  $i$  splits, individual  $ij$  will be the ancestor of all the individuals in one of the two offspring groups of group  $i$  with probability  $2/(N+1)$  if both offspring groups survive regulation, while  $ij$  will be the ancestor of all individuals in the offspring group of group  $i$  with probability  $1/(N+1)$  if only one of the offspring groups survives regulation (in both cases the fitness of  $ij$  is  $N$ ). Therefore, if another individual than  $ij$  is chosen for reproduction in group  $i$  and if the group splits, the fitness of  $ij$  is  $N(2n_g - 1)/[n_g(N+1)]$ .

Finally, the probability that an individual from another group is chosen for reproduction is (to the first order in  $z_{ij}$  and  $z$ ):

$$1 - \frac{1}{n_g} [1 + (b-c)(z_{ij} - z)]. \quad (34)$$

In that case (and if this other group splits), group  $i$  persists with probability  $1 - 1/n_g$ , in which case individual  $ij$  will be the ancestor of all the  $N$  individuals in group  $i$  at the next time unit with probability  $1/N$ . Hence, if an individual from another group than  $i$  is chosen for reproduction, the fitness of  $ij$  is  $1 - 1/n_g$ .

Combining all three cases just described gives:

$$w_{ij}^Q = 1 + \frac{1}{n_g} (b-c)(z_{ij} - z). \quad (35)$$

We now express  $w_{ij}^A$ , which is the fitness of individual  $ij$ , given that a migration occurs just after time  $t$ . With a probability given by equation 32, individual  $ij$  is chosen for reproduction.

One of his offspring remains in group  $i$ , while the other migrates. With probability  $1/n_g$  the migrant returns in its natal group, in which case with probability  $1 - 1/N$  the offspring that remained in group  $i$  will survive regulation. Each of the surviving offspring of individual  $ij$  will be the ancestor of all the  $N$  individuals in group  $i$  with probability  $1/N$  at the next time unit. Hence, the fitness of individual  $ij$  is given by  $2 - 1/N$  when the migrant offspring returns to the natal deme. With probability  $1 - 1/n_g$  the migrant enters a different group than group  $i$ . Then, from eq. 27 and eq. 2, the immigrant will be the ancestor of all the  $N$  individuals in group  $k$  at the next time unit with probability (to the first order in  $z_{ij}$  and  $z_k$ ):

$$\pi_g(z_{ij}, z_k) = \frac{1}{N} - \frac{1}{2} \left( c + \frac{(b-c)}{N} \right) (z_{ij} - z_k). \quad (36)$$

Thus, the fitness of individual  $ij$ , given that he is chosen for reproduction and that one of his offspring migrates to another deme is:

$$1 + \frac{1}{n_g - 1} \sum_{k \neq i} \left[ 1 - \frac{1}{2} (b + c(N-1)) (z_{ij} - z_k) \right] \quad (37)$$

which, given that

$$\frac{1}{n_g - 1} \sum_{k \neq i} z_k = \frac{n_g z - z_i}{n_g - 1} \quad (38)$$

and that  $z_i = z_{ij}$ , simplifies to:

$$2 - \frac{n_g}{2(n_g - 1)} (b + c(N-1)) (z_{ij} - z). \quad (39)$$

With a probability given by equation 33 it is another individual than  $ij$  that is chosen from group  $i$  for reproduction. If this reproductive events is followed by a migration, one of the offspring of the reproducing individual remains in group  $i$ , while the other migrates. The migrant returns to its natal group with probability  $1/n_g$ , in which case  $ij$  will be the ancestor of all the individuals of group  $i$  with probability  $1/N$  if he survives regulation, which occurs with probability  $1 - 1/N$  (in which case his fitness is  $N$ ). If the migrant reaches another group (probability  $1 - 1/n_g$ ), individual  $ij$  will be the ancestor of all the individuals of group  $i$  at the next time unit with probability  $1/N$  (in which case his fitness is also  $N$ ).

Finally, with a probability given by equation 34, an individual from another group (say group  $k$ ) is chosen for reproduction. If one of his offspring migrates, he will arrive in group  $i$  with probability  $1/n_g$ . Then, with probability  $1 - 1/N$ , individual  $ij$  survives regulation and

with probability  $1 - \pi_g(z_k, z_{ij})$  the immigrant lineage eventually disappears from group  $i$ . In that case, individual  $ij$  will be the ancestor of all the  $N$  individuals of group  $i$  at the next time unit with probability  $1/(N - 1)$ . From all this, the fitness of individual  $ij$ , given that an individual from another group is chosen for reproduction, and that this reproduction is followed by a migration is given by:

$$1 - \frac{1}{n_g} \pi_g(z_{ij}, z_k) \quad (40)$$

Taking all cases into account and simplifying, one obtains that  $w_{ij}^\Lambda$  is given by (to the first order in  $z_{ij}$  and  $z$ ):

$$w_{ij}^\Lambda = 1 - \frac{c}{n_g} (z_{ij} - z). \quad (41)$$

Since all groups are genetically homogeneous at time units  $t > 0$  ( $R = 1$ ), we can write  $z_{ij} = z_i$ , the term  $\partial w_{ij}^p / \partial z_{ij} + R \partial w_{ij}^p / \partial z_i$  that appears in equation 21 is then given directly by  $\partial w_{ij}^p / \partial z_{ij}$ . Equations 31, 35 and 41 give:

$$\frac{\partial w_{ij}^p}{\partial z_{ij}} = \frac{q(b - c) - \lambda c}{(\lambda + q)n_g}. \quad (42)$$

Finally, one obtains from equations 20, 24, 27, 30 and 42 that the inclusive fitness effect of allele  $A$  is positive when:

$$\frac{b}{c} > 1 + \frac{N + \frac{n_g \lambda}{q}}{n_g - 2 - \frac{\lambda}{Nq}}, \quad (43)$$

which is precisely eq. 34 of T&N and represents their exact result, holding for any group size and number. If  $n_g \gg 1$  and  $N \gg 1$ , the inclusive fitness effect boils down to

$$\frac{b}{c} > 1 + \frac{N}{n_g} + \frac{\lambda}{q}. \quad (44)$$

## Typical kin selection model

### Relatedness

We calculate here the terms  $R$  and  $w_{ij}$  appearing in the inclusive fitness effect for the life-cycle without overlapping generations described in our main text (eq. 6). The coefficient of relatedness is obtained from eq. 11 and involves expressions for the probabilities of identity between two homologous genes sampled with replacement from the same group  $Q_0^R(t)$  and

sampled with replacement from the total population  $Q_1^R(t)$ , where  $t$  corresponds to the time between two reproductive events during which group division, regulation and migration occurred. With probability  $1/N$  the same individual is sampled twice from the same group and with probability  $1/n_g$  the same group is sampled twice from the population. Whereby:

$$\begin{aligned} Q_0^R(t) &= \frac{1}{N} + \left( \frac{N-1}{N} \right) Q_0^D(t), \\ Q_1^R(t) &= \frac{1}{n_g} Q_0^R(t) + \left( \frac{n_g-1}{n_g} \right) Q_1^D(t), \end{aligned} \quad (45)$$

where is  $Q_0^D(t)$  the probability of identity between two homologous genes sampled from different individuals in the same group and  $Q_1^D(t)$  is the probability of identity between two homologous genes sampled from two individuals residing in different groups.

Note that  $Q_0^R(t)$  and  $Q_1^R(t)$  in eq. 45 can also be interpreted as the probabilities of identity between different offspring sampled in the same and different groups right after reproduction and regulation but before migration, and expressed in terms of the identities between individuals of the parental generation. After reproduction and regulation, each individual of the offspring generation migrates independently from each other with probability  $\lambda$  to another random group. The probability  $q_s$  that two individuals sampled in the same group after dispersal descend from the same group before dispersal is given by:

$$q_s = (1 - \lambda)^2 + \frac{\lambda^2}{n_g - 1}, \quad (46)$$

while the probability  $q_d$  that two individuals sampled in different groups after dispersal descend from the same group before dispersal is:

$$q_d = \frac{1 - q_s}{n_g - 1}, \quad (47)$$

We then have:

$$\begin{aligned} Q_0^D(t) &= q_s Q_0^R(t-1) + (1 - q_s) Q_1^R(t-1) \\ Q_1^D(t) &= q_d Q_0^R(t-1) + (1 - q_d) Q_1^R(t-1). \end{aligned} \quad (48)$$

Using the fact that  $Q_0^R(0) = 0$  and  $Q_1^D(0) = 0$ , one obtains after simplification:

$$R = \frac{n_g - 1}{N n_g - (N - 1)(n_g - 1)(q_s - q_d) - 1}. \quad (49)$$

Substituting the explicit values of  $q_s$  and  $q_d$  into the last equation, keeping only terms of leading order when the number of groups becomes large and the migration rate becomes

small ( $n_g \gg 1$  and  $\lambda \ll 1$ ), relatedness simplifies to:

$$R = 1 - (N - 1) \left( \frac{1}{n_g} + 2\lambda \right). \quad (50)$$

### Fitness function

The fitness  $w_{ij}$  of individual  $ij$  is given by the ratio of the individual's fecundity relative to average fecundity in the population. To the first order in  $z_i$ ,  $z_{ij}$  and  $z$ , this is:

$$w_{ij} = 1 - \left( c + \frac{b}{N-1} \right) z_{ij} + \frac{bN}{N-1} z_i - (b-c) z \quad (51)$$

Substituting this fitness function and eq. 49 into eq. 13, we find that allele  $A$  is selected for when:

$$\frac{b}{c} > 1 + \frac{N(1 + n_g(1 - q_s))}{n_g q_s - 2}. \quad (52)$$

In a panmictic population (i.e.,  $\lambda = 1 - 1/n_g$ ), this condition of invasion becomes:

$$\frac{b}{c} > 1 - n_g N. \quad (53)$$

If  $b > c$ , the equality can be satisfied only when  $b < 0$ . Helping group neighbors can no longer evolve, but harming them at a direct fecundity cost can.

## References

- [1] Rousset, F. (2004) *Genetic structure and selection in subdivided populations*. (Princeton University Press, Princeton, NJ).
- [2] Rousset, F. (2003) *Journal of Theoretical Biology* **221**, 665–668.
- [3] Rousset, F & Billiard, S. (2000) *Journal of Evolutionary Biology* **13**, 814–825.
- [4] Hamilton, W. D. (1964) *Journal of Theoretical Biology* **7**, 1–16.