

S2 Description of the analytical model

Here we provide a more complete overview of the analytical model, where Table S1 provides an overview of the notation used.

S2.1 Fitness expressions

To recap from the main text, the expected number w_{ij} of offspring who successfully establish themselves in a patch with i helpers and are born from a mutant adult breeder in a patch with a total number of j helpers is given by

$$w_{ij} = f_j \left(1 - h_j^\bullet\right) \left[\frac{n_b (1 - d) s_{j \rightarrow i} (h_j^\circ)}{C(h_j^\circ; \mathbf{h}, j)} + d \sum_{k=0}^{n_{h,\max}} u_k \frac{n_b s_{k \rightarrow i} (h_k)}{C(h_k; \mathbf{h}, k)} \right], \quad (\text{S1})$$

where f_j reflects the total number of surviving newborns produced by the focal adult breeder, a proportion $1 - h_j^\bullet$ of which develop as juvenile reproductives (rather than helpers). These juvenile reproductives then go on to compete for any of the n_b available breeding positions in the natal patch with probability $1 - d$ (first part in straight brackets), or in a random, remote patch with probability d (second part in straight brackets), where u_k reflects the population-wide frequency of patches currently containing $0 \leq k \leq n_{h,\max}$ helpers. Philopatric reproductives compete with a total number of $C(h_j^\circ; \mathbf{h}, j)$ philopatric and immigrant offspring (see eq. [S2] in the Online Supplement), which is a function of (i) the average tendency h_j° expressed by any locally born newborn to develop as a helper, (ii) the population wide tendencies $\mathbf{h} = [h_0, h_1, \dots, h_{n_{h,\max}}]$ to become helpers in any remote patch and (iii) the current number of helpers j in the local patch. Finally, after successful establishment, the probability that the newly established breeder is accompanied by i helpers in the next generation is then given by $s_{j \rightarrow i}(h_j^\circ)$ (see eq. 1). The expected number of offspring who successfully compete in remote patch can then be derived in a similar fashion.

The total number of competing reproductive juveniles in a local patch $C(h_j^\circ; \mathbf{h}, j)$ currently containing j helpers (where h_j° reflects the average tendency to develop as a helper) is given by

$$C(h_j^\circ; \mathbf{h}, j) = n_b f_j \left(1 - h_j^\circ\right) (1 - d) + d \sum_{k=0}^{n_{h,\max}} u_k n_b f_k (1 - h_k), \quad (\text{S2})$$

where the first part of the expression above describes the total number of philopatric, non-helping offspring produced by all members of the local group. The second part is the total number of juveniles who migrate to the current patch (with probability d), originating from a patch that currently contains k helpers (with probability p_k) resulting in a patch-level production of $n_b f_k (1 - h_k)$ reproductives.

| Symbol | Explanation |
|--|--|
| A | Resident transition matrix |
| a_{ij} | Entry in row i and column j of the resident transition matrix |
| B | Mutant transition matrix |
| b_{ij} | Entry in row i and column j of the mutant transition matrix |
| $C(h_j; \mathbf{h}, j)$ | Total number of competing reproductive juveniles in a patch which currently contains j helpers |
| d | Juvenile dispersal probability |
| f_i | Fecundity of a mother breeding in a patch which currently contains i helpers |
| ϕ_0, ϕ_1, ϕ_2 | Parameters of the fecundity function f_i |
| h_i | Probability that a juvenile born from a resident mother develops as a helper in a patch currently containing i helpers |
| h_i^\bullet | Probability that a juvenile born from a mutant mother develops as a helper in a patch currently containing i helpers |
| h_i° | Probability that any juvenile born on a mutant mother's patch develops as a helper in a patch currently containing i helpers |
| \mathbf{h} | Strategy vectors $\mathbf{h} = [h_0, h_1, \dots, h_{n_{h,\max}}]$ for the helping tendency expressed by offspring born from resident mothers |
| $\mathbf{h}^\bullet, \mathbf{h}^\circ$ | Strategy vectors for the average helping tendency expressed by offspring born from focal mutant mothers and offspring born from any mother in the mutant's patch |
| m_i | Mortality probability of an adult breeder (see section S2.6) |
| n_b | Number of breeders in each local patch |
| n_h | Number of helpers in a local patch (varies among patches) |
| $n_{h,\max}$ | Maximum possible number of helpers in each local patch |
| $Q_{i,t}$ | Coefficient of consanguinity between two breeding mothers on a patch containing i helpers at time t |
| \hat{Q}_i | Equilibrium coefficient of consanguinity |
| $r_{\text{local},i}$ | Relatedness between focal mutant breeder in a patch containing i helpers and any of the breeders in the focal's patch (including herself) |
| $s_{i \rightarrow j}$ | Probability that a patch which contains i helpers at time t will contain j helpers at time $t + 1$ (see eq. 1) |
| u_i | Stable class frequency of patches currently containing i helpers |
| v_i | Reproductive value of an adult breeder in a patch currently containing i helpers |
| w_{ij} | Expected number of offspring born who successfully establish themselves in a patch with j helpers born from a mutant focal mother in patch with i helpers |

Table S1: Notation of the model on the evolution of helping.

The mutant transition matrix \mathbf{B} , with elements $b_{ij} = w_{ij}$, then comprises the transition probabilities between all the different classes of mutants

$$\mathbf{B} = \begin{bmatrix} w_{00} & w_{01} & \cdots & w_{0n_{h,\max}} \\ w_{10} & w_{11} & \cdots & w_{1n_{h,\max}} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_{h,\max}0} & w_{n_{h,\max}1} & \cdots & w_{n_{h,\max}n_{h,\max}} \end{bmatrix}, \quad (\text{S3})$$

while

$$\mathbf{A} = \mathbf{B}|_{h^\bullet=h^\circ=h} \quad (\text{S4})$$

denotes the resident transition matrix which comprises the transition probabilities evaluated at the population average.

S2.2 Selection gradients

Recapping from the main text, we obtain the following term for the selection gradient \mathcal{H}_k

$$\mathcal{H}_k = V_k \sum_{i=0}^{n_{h,\max}} \sum_{j=0}^{n_{h,\max}} v_i u_j \left[\frac{\partial b_{ij}}{\partial h_k^\bullet} + \frac{\partial b_{ij}}{\partial h_k^\circ} r_{\text{local},j} \right] \Bigg|_{h^\bullet=h^\circ=h}, \quad (\text{S5})$$

where V_k is a term that is proportional to the amount of additive genetic variance in the helping tendency h_k . Next, v_i and u_j are the individual reproductive values and stable class frequencies of adult breeders which are in a patch with i helpers, which are obtained from the dominant left and right eigenvectors of the resident transition matrix (see eq. [S4]). Finally, the relatedness coefficient $r_{\text{local},j}$ reflects the relatedness between a focal adult breeder and all breeders in the local patch including herself (see eq. [S7]).

S2.2.1 Selection gradient for unconditional help

The model for unconditional helping behaviour is identical to the model with developmental plasticity, with the exception that we replace h_i with h for all $i \in 0, 1, \dots, n_{h,\max}$. Consequently, the selection gradient for the unconditional strategy \mathcal{H} is given by

$$\mathcal{H} = V \sum_{i=0}^{n_{h,\max}} \sum_{j=0}^{n_{h,\max}} v_i u_j \left[\frac{\partial b_{ij}}{\partial h^\bullet} + \frac{\partial b_{ij}}{\partial h^\circ} r_{\text{local},j} \right] \Bigg|_{h^\bullet=h^\circ=h}. \quad (\text{S6})$$

S2.3 Relatedness

The relatedness coefficient $r_{\text{local},j}$ reflects relatedness between a focal breeder and all breeders in the local patch including herself, which is given by

$$r_{\text{local},j} = \frac{1}{n_b} + \frac{n_b - 1}{n_b} \hat{Q}_j, \quad (\text{S7})$$

With probability $1/n_b$ the same breeder is sampled, so that identity is equal to 1. Alternatively, with probability $(n_b - 1)/n_b$, two distinct breeders are sampled, and the probability that both carry identical alleles is given by the equilibrium coefficient of consanguinity \hat{Q}_j in demes containing j breeders. The latter is calculated by solving the following recursion equation for all $j \in 0, 1, \dots, n_{\text{h,max}}$:

$$Q_{j,t+1} = \sum_{k=0}^{n_{\text{h,max}}} \frac{u_k s_{k \rightarrow j}(h_k)}{\sum_{\ell=0}^{n_{\text{h,max}}} u_\ell s_{\ell \rightarrow j}(h_\ell)} g_k(h_k)^2 \left(\frac{1}{n_b} + \frac{n_b - 1}{n_b} Q_{k,t} \right), \quad (\text{S8})$$

where $u_k s_{k \rightarrow j}(h_k) / \sum_{\ell=0}^{n_{\text{h,max}}} u_\ell s_{\ell \rightarrow j}(h_\ell)$ reflects the probability that a patch containing j helpers at time $t + 1$ was a patch containing k helpers at time t , in which case two natally born individuals are sampled with probability $g_k(h_k)^2$. Here $g_k(h_k)$ is the probability that a philopatric juvenile successfully establishes itself as an adult breeder. With probability $1/n_b$ one of the sampled breeders is born from the other sampled breeder, in which case the coefficient of consanguinity is 1. With probability $(n_b - 1)/n_b$ however, the sampled breeder has a different parent, in which case the coefficient of consanguinity is equal to $Q_{k,t}$. The probability $g_k(h_k)$ that a philopatric juvenile successfully establishes itself as an adult breeder is then given by

$$g_k(h_k) = \frac{n_b f_k (1 - d) (1 - h_k)}{C(h_k; \mathbf{h}, k)}. \quad (\text{S9})$$

S2.4 Numerical solutions

According to a standard result (e.g., [3–5]), the trait values in the next generation are then given by

$$\begin{bmatrix} h_{0,t+1} \\ h_{1,t+1} \\ \vdots \\ h_{n_{\text{h,max}},t+1} \end{bmatrix} = \begin{bmatrix} h_{0,t} \\ h_{1,t} \\ \vdots \\ h_{n_{\text{h,max}},t} \end{bmatrix} + \begin{bmatrix} \mathcal{H}_0 \\ \mathcal{H}_1 \\ \vdots \\ \mathcal{H}_{n_{\text{h,max}}} \end{bmatrix}, \quad (\text{S10})$$

where C reflects genetic variation due to mutation. To find the convergence stable evolutionary strategy [6, 7], we then iterate the above dynamic from the point $\mathbf{h}_{t=0} = [0.01]$ until convergence, where $|h_{i,t+1} - h_{i,t}| < 1 \times 10^{-8}$, using an algorithm written in C++ (see main text). During each timestep t , we solve for the equilibrium values of the left and right eigenvectors and the coefficients of consanguinity, given the updated values of \mathbf{h}_t . For the evolution of unconditional helping, eq. (S10) is given by $h_{t+1} = h_t + \mathcal{H}$ (see eq. [S6]).

S2.5 Individual-based simulations

We also ran some individual-based simulations to corroborate our results. To this end, we simulated a population of $n_p = 2500$ patches, each containing $n_b = 2$ hermaphroditic breeders. At the start of each generation, each breeder chooses a random sperm donor among the n individuals in the local patch (including itself). Upon fertilization, each breeder produces a number of offspring according to the same equation as in the analytical model, $f_j = (1/n_b)(\phi_0 + \phi_1 j^{\phi_2})$ (see the section “Evolutionary dynamics” in the main text), where we choose to multiply f_j by a fecundity parameter $K = 60$ to prevent local extinctions. Each offspring has $n_{h,\max} + 1$ unlinked, haploid and autosomal gene loci, which correspond to the helping tendencies $h_0, h_1, \dots, h_{n_{h,\max}}$. Upon inheritance, each locus independently mutates with probability $\mu = 0.01$, which involves adding a random number from a normal distribution with mean 0 and variance 0.0004 to its current allelic value (i.e., a continuum-of-alleles-model, [8]).

The number of helping offspring $n_{\text{juv,help},j}$ produced by a parent with helping trait h_j is then drawn from a Poisson distribution, with mean $f_j h_j$ while the number of reproductive offspring is given by $f_j - n_{\text{juv,help},j}$. A proportion d of all reproductive offspring is added to a pool of dispersers, which are randomized and then evenly distributed over all the $n_p = 2500$ patches, while the remainder of reproductive offspring competes for breeding positions in the natal patch.

As fecundity f_j is very large in the individual-based simulations (to prevent local extinction), the number of helpers is always much larger than the number of helping positions available. To approximate the analytical model therefore, we assume that all juvenile helpers experience a mortality stage (reflecting, for example, mortality during queing), during which individuals die with a probability of $m_{\text{juv}} = 0.95$ before becoming an adult helper. (An alternative approach in which fecundity f_j is small and extinct patches are recolonized by offspring born from randomly selected remote parents resulted in similar outcomes.) After juvenile survival, the number of helpers in the local patch in the next timestep is then the total number of surviving helping juveniles produced by all breeders. Subsequent to helper recruitment, we fill the vacant breeding positions by randomly sampling from the philopatric and immigrant juvenile reproductives, after which the cycle is repeated. Simulations ran for 40 000 generations, which was a sufficient amount of time for values of h_i to reach their equilibria. Simulations are coded in C++ and the source code is available on zenodo: <https://doi.org/10.5281/zenodo.1421729>

S2.6 Overlapping generations

Here we relax the assumption of nonoverlapping generations by assuming that adult breeders die with probability $0 < m \leq 1$. For the sake of tractability, we assume that helpers are recruited anew during each timestep. The modified version of eq. (2) is

then

$$w_{ij} = (1 - m)s_{j \rightarrow i}(h_j^\circ) + f_j(1 - h_j^\bullet) \left[\frac{n_b m (1 - d) s_{j \rightarrow i}(h_j^\circ)}{C(h_j^\circ; \mathbf{h}, j)} + d \sum_{k=0}^{n_{h, \max}} u_k \frac{n_b m s_{k \rightarrow i}(h_k)}{C(h_k; \mathbf{h}, k)} \right]. \quad (\text{S11})$$

The other equations are identical, except for the recursions of the coefficients of consanguinity (eq. [S12]), which are now given by (e.g., see [9])

$$Q_{j,t+1} = \sum_{k=0}^{n_{h, \max}} \frac{u_k s_{k \rightarrow j}(h_k)}{\sum_{\ell=0}^{n_{h, \max}} u_\ell s_{\ell \rightarrow j}(h_\ell)} \left[(1 - m)^2 Q_{k,t} + 2(1 - m) m g_k(h_k) \left(\frac{1}{n_b} + \frac{n_b - 1}{n_b} Q_{k,t} \right) + m^2 g_k(h_k)^2 \left(\frac{1}{n_b} + \frac{n_b - 1}{n_b} Q_{k,t} \right) \right]. \quad (\text{S12})$$

The first part in straight brackets reflects a scenario where two breeders are sampled which have both survived (each with probability $1 - m$), since the previous timestep and the probability that both have identical alleles is given by $Q_{k,t}$. The second part reflects the probability that one breeder is newborn, who has replaced a breeder that died (with probability m), while the other breeder has survived. In this case, the newborn breeder is born from the other sampled breeder with probability $1/n_b$ (hence both sampled breeders carry identical alleles with probability 1), while with probability $(n_b - 1)/n_b$ it was born from one of the other breeders (hence both sampled breeders carry identical alleles with probability $Q_{k,t}$). Finally, both sampled breeders are newborn with probability m^2 , which recapitulates the expression in eq. [S12].

References

- [1] Hamilton, W. D., 1964 The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–16. doi:10.1016/0022-5193(64)90038-4.
- [2] Rousset, F., 2004 *Genetic structure in subdivided populations*. Princeton: Princeton University Press.
- [3] Taylor, P. D., 1996 The selection differential in quantitative genetics and ESS models. *Evolution* **50**, 2106–2110. doi:10.2307/2410769.
- [4] Dieckmann, U. & Law, R., 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579–612. doi:10.1007/BF02409751.
- [5] Geritz, S. A. H., Kisdi, É., Meszéna, G. & Metz, J. A. J., 1998 Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* **12**, 35–57. doi:10.1023/a:1006554906681.

- [6] Christiansen, F. B., 1991 On conditions for evolutionary stability for a continuously varying character. *Am. Nat.* **138**, 37–50. doi:[10.1086/285203](https://doi.org/10.1086/285203).
- [7] Taylor, P. D., 1990 Allele-frequency change in a class-structured population. *Am. Nat.* **135**, 95–106. doi:[10.1086/285034](https://doi.org/10.1086/285034).
- [8] Kimura, M. & Crow, J. F., 1964 The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.
- [9] Pen, I., 2000 Reproductive effort in viscous populations. *Evolution* **54**, 293–297. doi:[10.1111/j.0014-3820.2000.tb00030.x](https://doi.org/10.1111/j.0014-3820.2000.tb00030.x).