The coevolution of cooperation and socially-mediated dispersal: a model

Iris Prigent[1] and Charles Mullon[1,*]

[1]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[*]Correspondence: charles.mullon@unil.ch

Abstract. Limited dispersal can promote the evolution of cooperation by increasing relatedness between social partners. However it also intensifies kin competition, potentially cancelling the benefits of helping. Here, we analyse a model in which individuals evolve both (i) the probability of cooperating within social groups as adults, and (ii) the dispersal probability of juveniles conditional on the number of adults that have cooperated in the group, leading to a reaction norm for dispersal. We show that cooperation and socially-mediated dispersal coevolve such that individuals disperse more from cooperative groups, reducing kin competition and thereby favouring further cooperation. This evolutionary feedback allows cooperation to evolve even when it would be entirely disfavoured under unconditional dispersal. In some cases, selection leads to the long-term coexistence of full cooperators and full defectors, each expressing a distinct dispersal reaction norm: cooperators disperse less on average but are more responsive to the social environment. These outcomes are driven by selection on indirect fitness effects and do not require kin recognition or spatial memory. Our results show how social behaviour and dispersal plasticity can coevolve through a feedback that enhances cooperation among relatives and mitigates kin competition.

1 Introduction

From microbes that secrete beneficial compounds (West et al., 2006) to alloparental care in arthropods (Yip and Rayor, 2014; Viera and Agnarsson, 2017), fish (Wisenden, 1999), birds (Skutch, 1961; Ekman, 2006), and mammals (Riedman, 1982; Clutton-Brock, 2002, 2021), cooperative behaviours are found across the tree of life. One of the key mechanisms that explain how such helping traits evolve is kin selection, which occurs when individuals are more likely to interact with others that are genetically related to them at the loci underlying the behaviour (Hamilton, 1963, 1964; Maynard Smith, 1964; Lehmann and Keller, 2006; West et al., 2007). This is the case under limited dispersal, which generates population structure and ensures that the recipients of help are more likely to be relatives (Hamilton, 1964). But limited dispersal also increases local competition between relatives, which can reduce or even cancel out the benefits of helping (Queller, 1992; West et al., 2002; Taylor, 1992; Lehmann and Rousset, 2010).

Because dispersal alleviates local competition, it can itself be favoured by kin selection. When dispersing reduces competition among relatives, selection may favour greater dispersal even when it is individually costly (Hamilton and May, 1977; Comins et al., 1980; Frank, 1986; Taylor, 1988; Gandon and Rousset, 1999; Gandon and Michalakis, 1999). But dispersal also reduces relatedness, weakening the kin-selected benefits of cooperation (Lehmann and Rousset, 2010). This trade-off between kin competition and relatedness can be mitigated when individuals express social behaviour conditionally on dispersal status. In such cases, selection tends to favour helping in philopatric individuals and defection in dispersers (El Mouden and Gardner, 2008). Several models have shown that associations between social behaviour and dispersal can support the stable coexistence of sessile cooperators and dispersive defectors (Hochberg et al., 2008; Wakano et al., 2009), and that such associations can emerge through gradual evolution (Purcell et al., 2012; Parvinen, 2013; Mullon et al., 2016, 2018; Prigent and Mullon, 2023). However, these models typically assume that dispersal does not depend on an individual's social environment.

In nature, dispersal is often conditionally expressed. Individuals adjust their dispersal behaviour in response to a range of internal or external cues, including body condition (Dufty and Belthoff, 2001; Clobert et al., 2009), local density (Lambin et al., 2001; Matthysen, 2005), and social context such as aggression from dominant individuals (Lawson Handley and Perrin, 2007; for reviews: Clobert et al., 2004; Matthysen, 2012). Theoretical models have investigated how such dispersal plasticity can evolve in response to patch quality (Crespi and Taylor, 1990), parental age (Ronce et al., 1998, 2000), or local crowding (Poethke and Hovestadt, 2002; Kisdi, 2004). In particular, Kanwal and Gardner (2022) showed that when individuals have evolved adaptive density-dependent dispersal, kin competition effects are abolished, which in turn facilitates the evolution of helping. Yet dispersal may depend on environmental cues other than density (Clobert et al., 2012), and the consequences of such plasticity for the evolution of cooperation remain understudied.

Here, we investigate the joint evolution of cooperation, defined as a contribution to a local common good, and socially-mediated dispersal, whereby individuals adjust dispersal based on the social composition of their group (i.e. the number of groupmates who have cooperated versus defected). The interaction between these traits introduces feedbacks that make the direction of selection difficult to anticipate. On the one hand, dispersing more in cooperative groups can help export the additional offspring produced by cooperation, thereby reducing kin competition among relatives (Kanwal and Gardner, 2022). On the other hand, dispersing away from defectors may protect cooperators from exploitation and increase assortment between like types (Aktipis, 2004). Additionally, conditional dispersal may also benefit defectors, for instance by allowing them to avoid other defectors and disperse into groups richer in cooperators. Whether such dynamics ultimately favour cooperation, defection, or their coexistence is the question we address here.

2 Model

We consider a population subdivided into a large number of groups, each of size N, and with the following life-cycle (Fig. 1). First, individuals interact socially within their group. Each individual independently contributes to a local common good with probability z at a personal cost C. The resulting common good is shared equally among all group members, each receiving a benefit B_k/N, where k denotes the number of cooperators in the group. Individuals then reproduce according to their fecundity: $1+\delta(B_k/N-C)$ for cooperators, and $1+\delta(B_k/N)$ for defectors, where δ scales the effect of cooperation on fecundity. Offspring disperse with probability d_k, which depends on the number k of cooperators in their parental group (i.e. among the adults). Dispersal follows the island model: when an offspring disperses, it survives with probability $1-c_{\mathrm{d}}$ and settles uniformly at random into one of the other groups. Finally, offspring compete locally to replace the adults (Wright-Fisher process).

To investigate the joint evolution of cooperation and conditional dispersal, we use an approach inspired by quantitative genetics for group-structured populations (Mullon and Lehmann, 2019; see Appendix A for details). Individuals are characterised by $N+2$ evolving traits: the probability z of cooperating, and $N+1$ probabilities d_k of dispersal conditional on the number of cooperators in the group. Assuming that each trait evolves through small-effect mutations at a pleiotropic locus (continuum-of-alleles model), the change in the vector of population mean trait values, $\overline{\boldsymbol{x}}=$

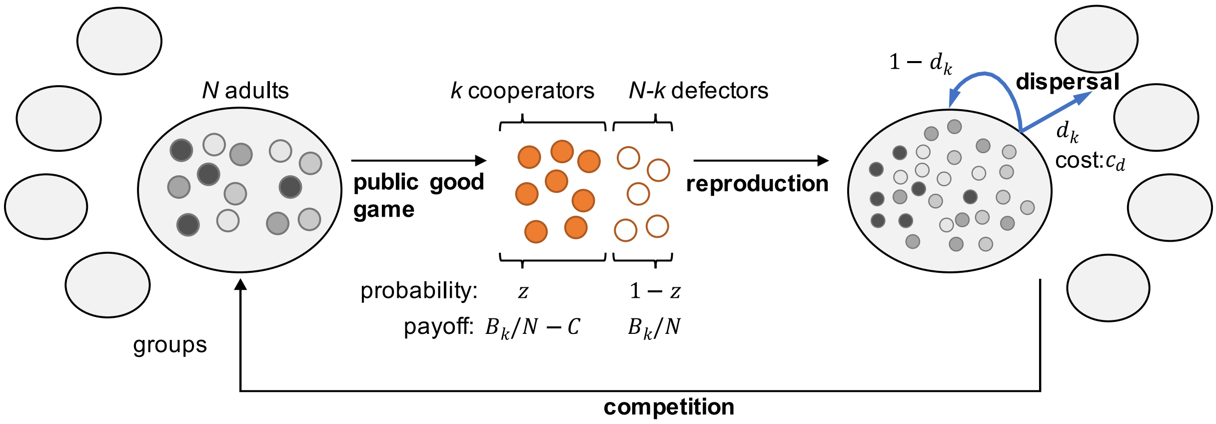

Figure 1: A model for the co-evolution of common good contribution and socially-mediated dispersal. Each adult either contributes to the common good, i.e. cooperates (with probability z, at cost C) or defects (with probability $1-z$). The group-level benefit B_k is shared equally among N group members, where k is the number of cooperators. Individuals reproduce based on their fecundity, and each offspring either stays or disperses with probability d_k, depending on k. Dispersal is costly and surviving migrants settle uniformly at random into other groups (island model). Adults die, and offspring compete locally for the N breeding spots in each group. See main text for details.

$(\overline{z}, \overline{d}_0, \ldots, \overline{d}_N)$, over one generation is given by

$$\Delta\overline{\boldsymbol{x}} = \mathbf{G}\,\boldsymbol{s}(\overline{\boldsymbol{x}}), \tag{1}$$

where $\mathbf{G}$ is the additive genetic variance-covariance matrix, and $\boldsymbol{s}(\overline{\boldsymbol{x}})$ is the vector of selection gradients. The gradient $\boldsymbol{s}(\overline{\boldsymbol{x}})$ points in the direction favoured by selection in multitrait space, while $\mathbf{G}$ determines how this selection translates into evolutionary change (Lande, 1979).

The matrix $\mathbf{G}$ itself evolves, but this can be neglected as long as $\overline{\boldsymbol{x}}$ is still under directional selection owing to a separation of timescales between changes in $\mathbf{G}$, which are slow, and in $\overline{\boldsymbol{x}}$, which are fast, when mutations are rare with weak effects (Iwasa et al., 1991; Leimar, 2009; Débarre et al., 2014; Mullon and Lehmann, 2019). That is, if one is interested in long-term evolution, the change in $\mathbf{G}$ can be neglected unless the mean in the population is such that directional selection has vanished, i.e. when $\overline{\boldsymbol{x}} = \boldsymbol{x}^*$ such that $\boldsymbol{s}(\boldsymbol{x}^*) = \mathbf{0}$. We refer to such a trait vector $\boldsymbol{x}^*$ as a singular strategy.

If additionally, the population converges according to eq. (1) toward $\boldsymbol{x}^*$ for any $\mathbf{G}$-matrix, $\boldsymbol{x}^*$ is said to be "strongly convergence stable" (Leimar, 2009). Here, we will assume that mutations affect each trait independently of one another such that $\mathbf{G}$ can be considered as a diagonal matrix during this convergence. If the population evolves to a singular strategy $\overline{\boldsymbol{x}} = \boldsymbol{x}^*$ in this case, then we will refer to $\boldsymbol{x}^*$ simply as "convergence stable". If such convergence occurs, subsequent selection on $\mathbf{G}$ when $\overline{\boldsymbol{x}} = \boldsymbol{x}^*$ determines whether the trait distribution in the population remains unimodally distributed around

$\overline{\boldsymbol{x}}$ (selection is stabilising) or becomes multi-modal such that the population becomes polymorphic (selection is disruptive; Débarre et al., 2014; Mullon and Lehmann, 2019).

We derive the relevant expressions for the selection gradients $\boldsymbol{s}(\overline{\boldsymbol{x}})$ and for changes in $\mathbf{G}$ in Appendix C, based on individual fitness and relatedness coefficients defined in Appendix B. A summary of the main results is given below.

3 Results

3.1 Socially-mediated dispersal is conducive to the emergence of cooperation

We begin by analysing the emergence of cooperation in a population where dispersal is socially-mediated but does not evolve. That is, we assume the dispersal probabilities $\overline{d}_0, \ldots, \overline{d}_N$ are fixed, although they may differ from one another. We ask when selection favours an increase in the probability of cooperation, starting from a resident population of pure defectors. This corresponds to a positive selection gradient on z when the population expresses $\overline{\boldsymbol{x}}^{(0)} = (0, \overline{d}_0, \ldots, \overline{d}_N)$, i.e. when $s_1(\overline{\boldsymbol{x}}^{(0)}) > 0$. Assuming weak effects on fecundity (i.e. small δ), this condition becomes:

$$\left[\frac{B_1}{N} - C\right] + \kappa\,(N-1)\,\frac{B_1}{N} + A > 0, \tag{2}$$

where κ and A depend on group size N, dispersal cost c_d, and the dispersal strategies in groups with zero or one cooperator (i.e. $\overline{d}_0$ and $\overline{d}_1$). Full derivations are given in Appendix C.1 (see in particular eqs. C-16 and C-18–C-20).

Condition (2) decomposes selection on cooperation into three components: (i) $B_1/N - C$ is the direct effect on the actor's own fecundity, which increases by B_1/N from the common good and decreases by C from individual cost. (ii) $\kappa(N-1)B_1/N$ is the effect on the fecundity of neighbours, weighted by a scaled relatedness coefficient κ. This term reflects the indirect effect of cooperation, accounting for the fact that local competition reduces the evolutionary benefit of helping kin. κ summarises how relatedness, which favours the evolution of cooperation, balances against kin competition, which hinders it (e.g. Lehmann and Rousset, 2010; Van Cleve, 2015). When $\kappa > 0$, the benefits of helping outweigh the costs of competing with kin; when $\kappa < 0$, the reverse holds; and when $\kappa = 0$, benefits and costs cancel each other (as in Taylor, 1992). (iii) A captures selection on cooperation via its effect on dispersal. When $\overline{d}_0 \neq \overline{d}_1$, switching from defection to cooperation affects the dispersal behaviour of both the actor's and its neighbours' offspring. This alters the intensity of kin competition and may therefore be favoured or disfavoured by selection even when $C = 0$ and $B_k = 0$. A positive value of A means that cooperation reduces local competition through its effect on dispersal and thus tends to be favoured; a negative value indicates the opposite.

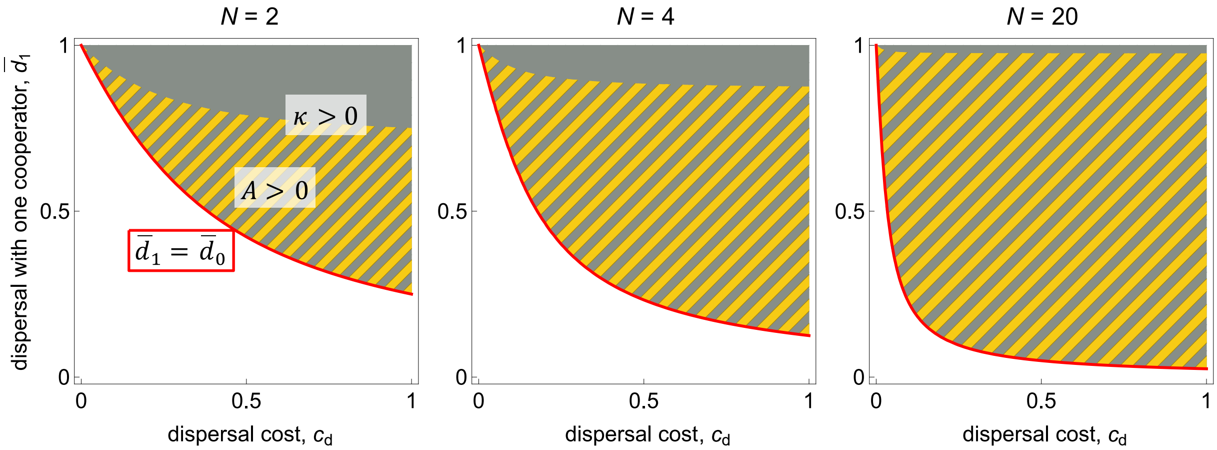

Figure 2: Socially-mediated dispersal strategies that favour the emergence of cooperation. Parameter regions of $\bar{d}_1$ and dispersal cost c_d for which $A > 0$ (yellow-striped) and $\kappa > 0$ (grey region) in eq. (2) (drawn from eqs. C-18–C-20). Dispersal in the absence of cooperators is assumed to be at its evolutionarily stable value, $\bar{d}_0 = d^*$ (eq. 3). The red curve indicates $\bar{d}_1 = \bar{d}_0 = d^*$. Panels show different group sizes ($N = 2$, 4, and 20). Cooperation is favoured when $A > 0$ or $\kappa > 0$, which requires $\bar{d}_1 > \bar{d}_0$; that is, individuals must disperse more readily in the presence of one cooperator than in its absence.

In the absence of socially-mediated dispersal ($\bar{d}_0 = \bar{d}_1$), the effect of cooperation on dispersal vanishes entirely: $A = \kappa = 0$ (eq. C-21, Appendix C.1). In this case, condition (2) reduces to the classical requirement $B_1/N - C > 0$, meaning that cooperation evolves only if the direct fecundity benefit outweighs the cost (for homogeneous groups and non-overlapping generations, e.g. Taylor, 1992).

By contrast, when dispersal is socially-mediated ($\bar{d}_0 \neq \bar{d}_1$), cooperation can be favoured even when this condition fails (i.e. when $B_1/N - C < 0$), provided that $\kappa > 0$ and/or $A > 0$. To explore when this occurs, let us assume that in a population of pure defectors ($\bar{z} = 0$), dispersal in the absence of cooperation ($\bar{d}_0$) is at the evolutionarily stable value,

$$d^* = \frac{2}{1 + 2c_d N + \sqrt{1 + 4c_d N(N-1)}}, \tag{3}$$

which corresponds to the stable strategy for unconditional dispersal in a homogeneous population (Ajar, 2003). We then find that both A and κ are positive only if $\bar{d}_1 > \bar{d}_0$ (Fig. 2, yellow-striped and grey regions above the red curve). In other words, cooperation is favoured when individuals disperse more readily in the presence of a cooperator. This is because such a reaction norm for dispersal reduces kin competition. When a focal individual cooperates, it increases the number of offspring produced by relatives. If these offspring, and the cooperator's own, disperse more than average, they compete less with one another, and the benefits of helping are less eroded by local competition. This alleviation of kin competition promotes the evolution of cooperation. Socially-mediated dispersal can also reduce kin competition among defectors and therefore favour the invasion of defectors in a population of full cooperators, although the conditions for this are more restrictive (requiring $\overline{d}_{N-1} > \overline{d}_N$, Appendix C.2 for details).

The above suggests that depending on its nature, socially-mediated dispersal can favour or hinder cooperation evolution. This raises the question of which dispersal strategy is actually favoured by selection. To address this, we investigate next the joint evolution of all traits, $\overline{\boldsymbol{x}} = (\overline{z}, \overline{d}_0, \ldots, \overline{d}_N)$, under three typical common goods games (i.e. different payoff functions B_k).

3.2 Linear returns: transitions between no and full cooperation

When the benefits of the common good increase linearly with the number of cooperators, the payoff function is $B_k = b \times k$ for some fixed parameter $b > 0$ (Fig. 3a). When dispersal among groups is unconditional, $\overline{z}$ evolves to one only if $b/N - C > 0$, and to zero otherwise (Fig. 3b; e.g. Peña et al., 2015). Meanwhile, unconditional dispersal evolves toward the evolutionarily stable equilibrium d^* (eq. 3).

To gain insights into the evolution of socially-mediated dispersal and its effect on cooperation, we first perform a numerical stability analysis (using Mathematica function FindRoot[]) to look for singular strategies $\boldsymbol{x}^* = (z^*, d_0^*, \ldots, d_N^*)$ with all values between zero and one, such that directional selection vanishes on all traits (i.e., such that $\boldsymbol{s}(\boldsymbol{x}^*) = \boldsymbol{0}$). For all model parameters tested, we could not find such a strategy (Appendix D.1.1 for details). The absence of singular strategies suggests that evolutionary dynamics also lead to either full or no cooperation when conditional dispersal coevolves with cooperation.

To determine whether socially-mediated dispersal allows cooperation to fix when initially absent (i.e. $\overline{z}$ going from 0 to 1) or to be lost when initially fixed ($\overline{z}$ going from 1 to 0), we tracked the joint evolutionary dynamics of all traits using eq. (1) (Appendix D.2.1 for procedure). Fig. 3c shows one such trajectory, in which cooperation evolves to fixation even though direct fecundity effects favour defection ($b/N - C < 0$). We repeated this analysis across a wide range of initial dispersal strategies and

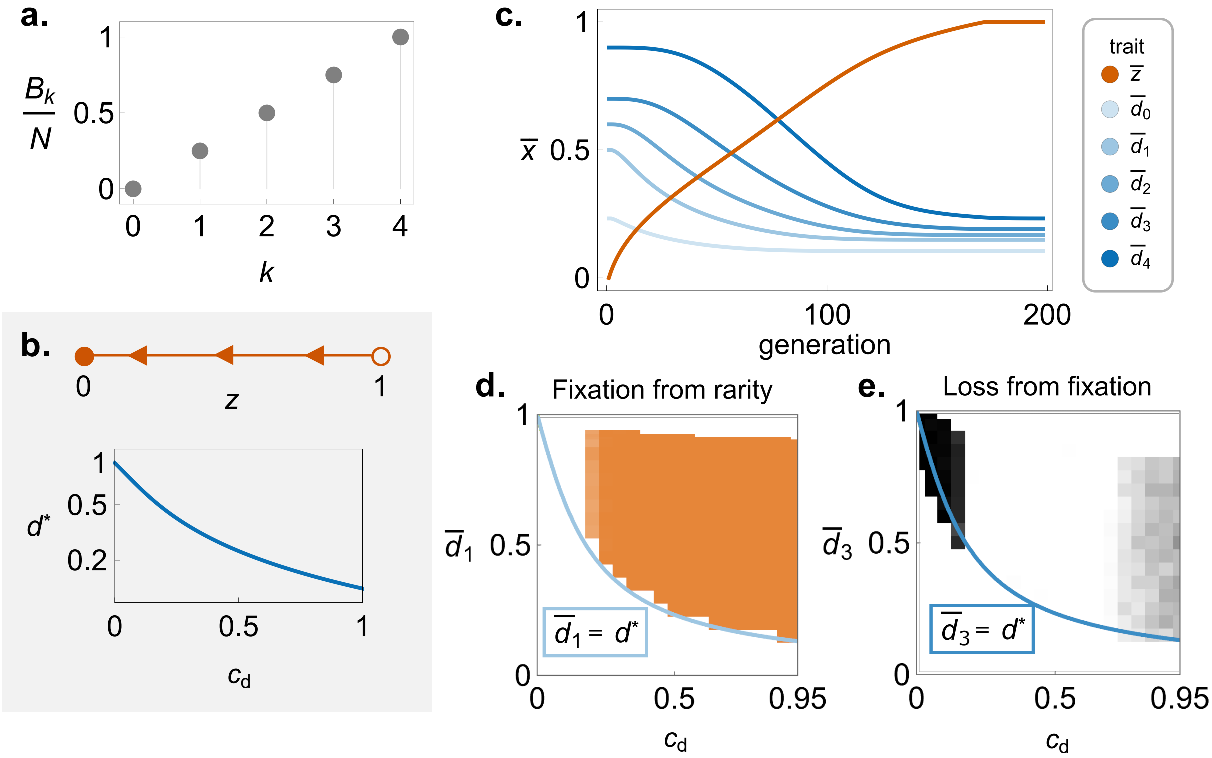

Figure 3: Evolutionary dynamics of cooperation and socially-mediated dispersal under linear benefits. a. Individual payoff function B_k/N increasing linearly with the number of cooperators k ($B_k = b \times k$). **b.** Top: evolution of cooperation when dispersal is fixed and unconditional (with $C > b/N$, leading to full defection). Bottom: evolutionarily stable unconditional dispersal strategy, d^* (from eq. 3). **c.** Joint evolutionary dynamics of all traits according to eq. (1) with $\mathbf{G}$ proportional to the identity matrix. Initial dispersal strategy: $\bar{\boldsymbol{x}} = (0, d^*, 0.5, 0.6, 0.7, 0.9)$. Other parameters: $N = 4$, $\delta = 1$, $b = 1$, $C = 0.3$ such that $b/N - C = -0.05 < 0$. This shows that cooperation can emerge and fix even if $b/N - C < 0$ when socially-mediated dispersal coevolves. **d.** Proportion of initial dispersal strategies that lead to invasion and fixation of cooperation (colour scale), for each combination of dispersal cost c_d and dispersal probability $\bar{d}_1$. Each value is computed from 100 deterministic evolutionary trajectories (as in panel c) starting from $\bar{z} = 0$ with $\bar{d}_0 = d^*$ and the indicated value of $\bar{d}_1$; the remaining dispersal traits are drawn randomly between 0 and 1 (Appendix D.2.2). Cooperation often fixes when $d_1 > d^*$, except when c_d is small. Other parameters as in panel c. **e.** Proportion of initial dispersal strategies that lead to invasion and fixation of defection (grey scale), for each combination of dispersal cost c_d and dispersal probability $\bar{d}_{N-1} = \bar{d}_3$. Each value is computed from 100 deterministic evolutionary trajectories (as in panel c) starting from $\bar{z} = 1$ with $\bar{d}_4 = d^*$ and the indicated value of $\bar{d}_3$, while the remaining dispersal traits are drawn randomly between 0 and 1 (Appendix D.2.2). Despite direct fecundity effects favouring defection ($b/N - C < 0$), the evolution of socially mediated dispersal often maintains cooperation. Other parameters as in panel c.

dispersal costs. Starting from a population of defectors ($\bar{z} = 0$), we fixed dispersal in the absence of cooperators at its evolutionarily stable value ($\bar{d}_0 = d^*$) and varied dispersal in groups with one cooperator ($\bar{d}_1$), while drawing the remaining dispersal traits randomly between 0 and 1. Fig. 3d shows, for each combination of c_d and $\bar{d}_1$, the proportion of initial dispersal strategies that lead to invasion and fixation of cooperation. This reveals that cooperation often invades and fixes even when $b/N - C < 0$, provided dispersal cost c_d is not too small and initial dispersal is such that $\overline{d}_1 > \overline{d}_0$. By contrast, the other initial dispersal probabilities ($\overline{d}_2$ and so on) have little effect.

We then repeated the same analysis starting from a population of full cooperators ($\overline{z} = 1$), fixing dispersal in the absence of defectors at its evolutionarily stable value $\overline{d}_N = d^*$ and varying dispersal in groups with one defector $\overline{d}_{N-1}$, while drawing the remaining initial values for the dispersal traits randomly between 0 and 1. This allowed us to ask whether socially mediated dispersal maintains cooperation once it has fixed, or instead allows it to be lost. Fig. 3e shows that cooperation is maintained across most of the parameter space (white space). However, when dispersal costs c_d are either very low or very high and $\overline{d}_{N-1} > \overline{d}_N$, defection can invade and lead to the collapse of cooperation.

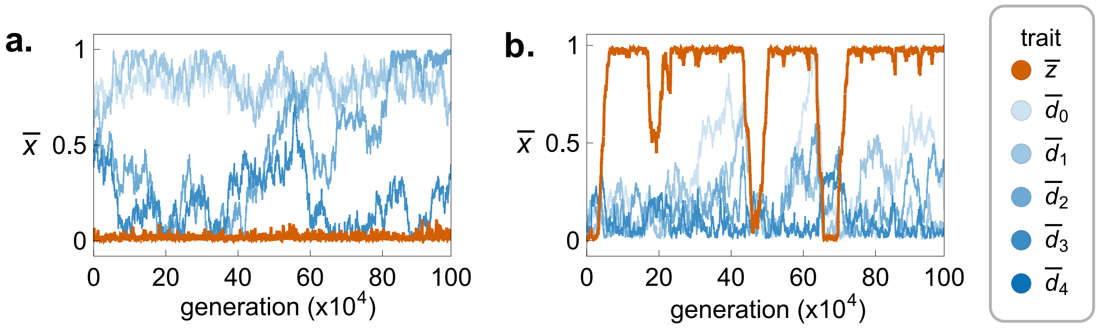

Figure 4: Stochastic transitions between cooperation and defection. Individual-based simulations showing transitions between full defection and full cooperation driven by genetic drift in socially-mediated dispersal (Appendix E.1 for procedure). Orange: average probability of cooperation $\overline{z}$; blue: average dispersal probabilities $\overline{d}_k$. **a.** Low dispersal cost ($c_\mathrm{d} = 0.05$): the population remains trapped in defection. **b.** High dispersal cost ($c_\mathrm{d} = 0.85$): the population switches between defection and cooperation, but spends longer periods in the cooperative state. Other parameters: $N = 4$, $\delta = 1$, $b = 1$, $C = 0.3$.

The existence of initial dispersal strategies that lead to the fixation of cooperation when absent (Fig. 3d) and others that lead to its loss when present (Fig. 3e) suggests that the population may transition between full cooperation and full defection. This is especially likely when dispersal cost is high, because both outcomes are then accessible under different initial dispersal strategies (Fig. 3de). Such switches require conditional dispersal to keep evolving even when cooperation is either absent or fixed. In these states, all dispersal traits except $\overline{d}_0$ (when $\overline{z} = 0$) or $\overline{d}_N$ (when $\overline{z} = 1$) are selectively neutral. They can therefore change via genetic drift. To test whether such drift in socially-mediated dispersal can drive switches between the two social states, we performed individual-based simulations in a population with few groups (125 groups; Appendix E.1 for procedure). As expected, full cooperation does not evolve when c_d is low (Fig. 4a). By contrast, when c_d is high, the population switches between cooperation and defection, while spending longer in the cooperative state (Fig. 4b).

3.3 Accelerating returns: facilitating the transition to full cooperation

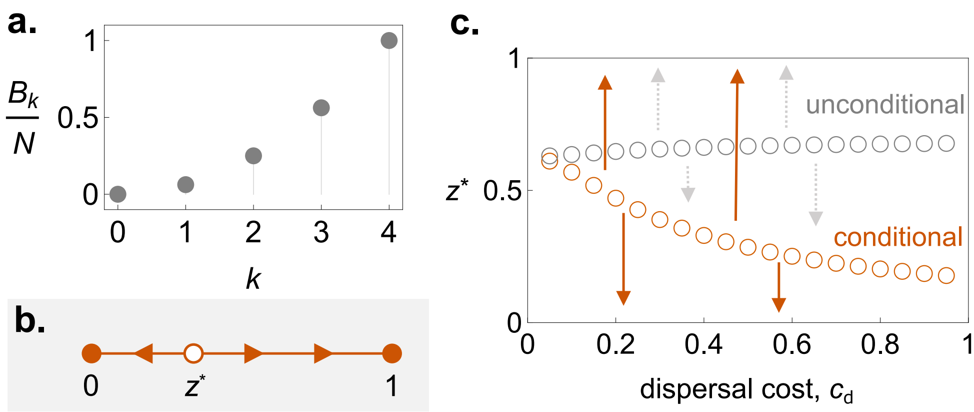

Figure 5: Evolutionary dynamics under accelerating benefits. a. Individual payoff function $B_k/N = b \times (k/N)^u$, which accelerates with the number of cooperators when $u > 1$ (here, $u = 2$). **b.** Evolutionary dynamics of cooperation when dispersal is fixed and unconditional. Cooperation either evolves to fixation or is lost, depending on whether the initial condition is above or below a critical threshold z^*. **c.** Singular values of z^* as a function of c_d. In grey: z^* under unconditional dispersal. In orange: z^* when socially mediated dispersal jointly evolves with cooperation. The evolution of socially mediated dispersal lowers the threshold for cooperation, particularly when the cost of dispersal is high (see Appendix D.1.1 for details). Other parameters: $N = 4$, $b = 1$, $C = 0.3$, $u = 2$.

We now consider cases in which cooperation yields accelerating benefits for fecundity (Fig. 5a). When dispersal is fixed and unconditional, such returns generate evolutionary bistability: there exists a singular strategy z^* that is not convergence stable, such that the evolutionary outcome depends on the initial level of cooperation. If the population starts with $\bar{z} < z^*$, $\bar{z}$ evolves to zero; if $\bar{z} > z^*$, $\bar{z}$ evolves to one (Peña et al., 2015) (Fig. 5b).

To examine whether evolutionary dynamics are similar when socially mediated dispersal also evolves, we first performed a stability analysis to identify singular strategies $\boldsymbol{x}^* = (z^*, d_0^*, \ldots, d_N^*)$ and to assess their convergence stability across parameter values (see Appendix D.1.1 for details). We find that singular strategies $\boldsymbol{x}^*$ exist but none are convergence stable. For all parameter sets tested, the singular value of cooperation, z^*, is lower when socially mediated dispersal can evolve than when dispersal is unconditional, particularly when the cost of dispersal is high (Fig. 5c; Supp. Fig. S3a-c). This suggests that under socially mediated dispersal, cooperation can evolve from a lower threshold, thereby increasing the basin of attraction of the cooperative outcome.

In addition to lowering this threshold, we next tested whether cooperation could also evolve from $\bar{z} = 0$ when it jointly evolves with socially mediated dispersal as observed under linear payoffs (Section 3.2). To do so, we tracked the joint evolutionary dynamics of all traits using eq. (1) (Appendix D.2.2

for details). We started from a range of randomly chosen initial dispersal strategies and considered two initial conditions: $\overline{z} = 0$, to test whether cooperation can invade, and $\overline{z} = 1$, to test whether cooperation can be lost. For each initial dispersal strategy, we recorded whether the population evolved toward fixation of cooperation or defection.

When the cost of dispersal c_d is high, the evolutionary outcome depends on the initial dispersal strategy: some strategies allow cooperation to invade and reach fixation when initially absent, whereas others allow defection to invade and drive cooperation to collapse when it is initially fixed (Fig. 6a). Strategies in which individuals initially disperse more in the presence of cooperators promote the evolution of cooperation (Fig. 6b), whereas strategies in which individuals initially disperse more in the presence of defectors promote defection (Fig. 6c). Moreover, a larger fraction of initial dispersal strategies leads to the fixation of cooperation than to its loss (Fig. 6a).

These results are consistent with our analyses in Sections 3.1-3.2. They confirm that selection favours social behaviours that modulate dispersal in response to cooperation – particularly when the cost of dispersal is high – because such behaviours reduce kin competition and thereby increase the likelihood of evolutionary transitions toward the cooperative state.

3.4 Diminishing returns: dispersal reaction norms and their variation

Finally, we investigate the case in which the benefits of cooperation saturate with the number of cooperators (Fig. 7a). When dispersal is unconditional, the population gradually evolves towards an intermediate probability of cooperation, $0 < z^* < 1$, which is maintained by stabilising selection (Peña et al., 2015) (Fig. 7b).

We performed numerical analyses to identify singular strategies $\boldsymbol{x}^* = (z^*, d_0^*, \ldots, d_N^*)$ and assess their convergence stability across parameter values. We find that such strategies can exist and, when they do, they are convergence stable. In these cases, the population therefore converges to an intermediate probability of cooperation z^*, as under unconditional dispersal. However, this value is always higher than under unconditional dispersal, especially when the cost of dispersal is high and the payoff saturates slowly (Fig. 7c; Supp. Fig. S3).

We also tracked the evolutionary dynamics of all traits using eq. (1) to determine the evolutionary outcome across parameter values, including cases in which no interior singular strategy exists. Fig. 7c-d shows that the population can converge towards three types of equilibrium. First, when the cost of

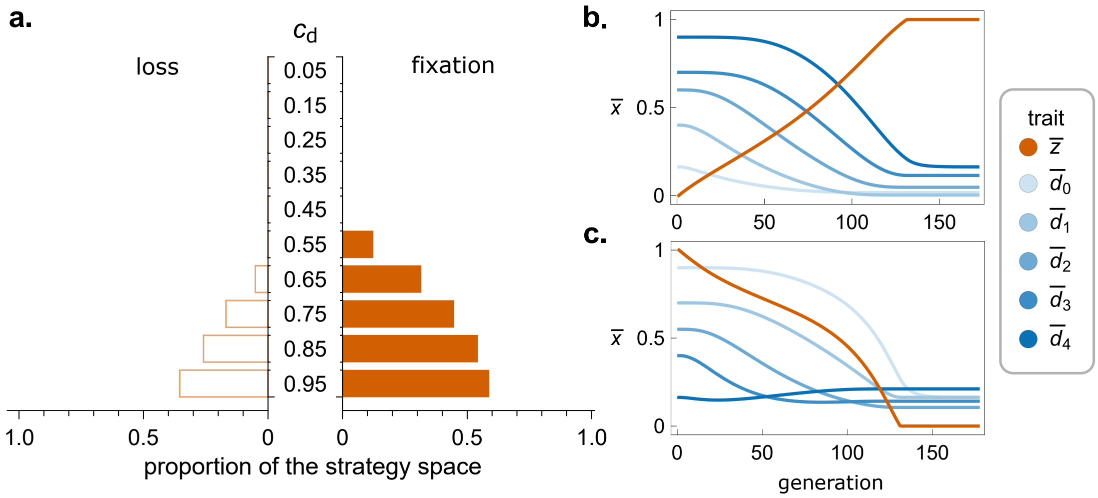

Figure 6: Socially mediated dispersal allows evolution to bypass the cooperation threshold under accelerating benefits. a. Proportion of the space of initial dispersal strategies leading to the invasion and fixation of cooperation (right-hand side; see example in panel b) and defection (left-hand side; see example in panel c; Appendix D.2.2 for details). **b.** Example of evolutionary dynamics leading to invasion and fixation of cooperation. Under unconditional dispersal, cooperation cannot invade because the population remains below the cooperation threshold. When dispersal is socially mediated, however, cooperation can invade and evolve to fixation (here starting from $\bar{\boldsymbol{x}} = (0, d^*, 0.4, 0.6, 0.7, 0.9)$, with $c_\mathrm{d} = 0.75$ and d^* given by eq. 3). **c.** Example of evolutionary dynamics leading to invasion and fixation of defection. Under unconditional dispersal, defection cannot invade but when dispersal is socially mediated defection can invade and evolve to fixation (here starting from $\bar{\boldsymbol{x}} = (1, 0.9, 0.7, 0.55, 0.4, d^*)$, with $c_\mathrm{d} = 0.75$). Other parameters: $N = 4$, $\delta = 1$, $b = 1$, $C = 0.3$, $u = 2$.

dispersal c_d is small, the equilibrium corresponds to the interior singular strategy identified above, in which all traits take intermediate values between 0 and 1 and directional selection has subsided on all traits ($\boldsymbol{s}(\boldsymbol{x}^*) = \boldsymbol{0}$). Second, when the cost of dispersal is intermediate, the equilibrium probability of cooperation remains intermediate, but some dispersal traits are being continuously pushed against the boundary of trait space by directional selection (shown here at $c_\mathrm{d} = 0.35$, where $\bar{d}_0$ and $\bar{d}_1$ are pushed to 0 while $0 < z^* < 1$). Finally, when c_d is large, directional selection drives the population to full cooperation, such that $\bar{z}$ is pushed to 1.

When selection favours an intermediate probability of cooperation $0 < z^* < 1$, the number of cooperators in a group fluctuates owing to sampling effects. Selection can then shape a dispersal reaction norm with respect to the number of cooperators by modifying the dispersal probabilities $\bar{d}_k$. We find that, under this reaction norm, individuals express higher dispersal probabilities in groups with more cooperators (Fig. 7d; Supp. Fig. S4). As expected, a higher cost of dispersal leads to lower overall dispersal, whereas intermediate costs produce steeper reaction norms (Fig. 7d). Further analyses of the dispersal strategy favoured by directional selection across parameter values (Appendix D.3.2;

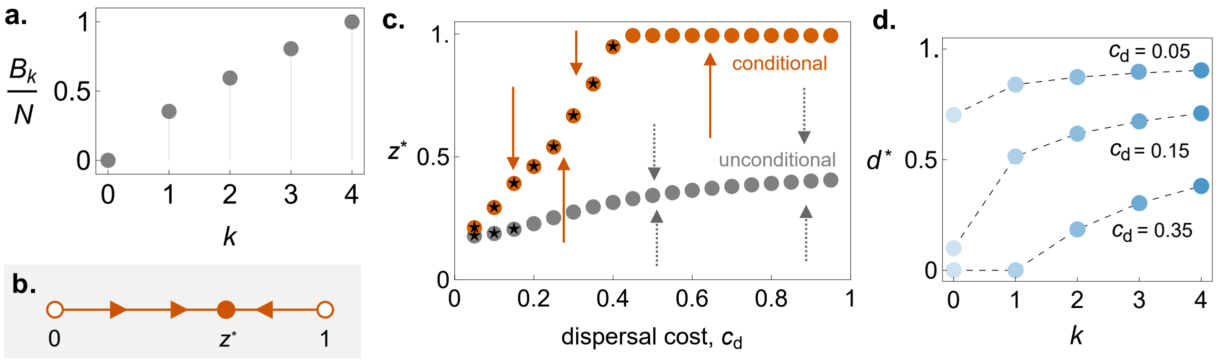

Figure 7: Evolutionary dynamics under saturating benefits. **a.** Individual payoff function $B_k/N = b \times (k/N)^u$, which saturates with the number of cooperators when $u < 1$ (here $u = 0.75$). **b.** Evolutionary dynamics of cooperation when dispersal is fixed and unconditional. Cooperation converges to an intermediate singular trait value z^*. **c.** Equilibrium probability of cooperation z toward which the population converges under directional selection. In orange: with socially mediated dispersal. In grey: with unconditional dispersal. Black stars denote singular strategies at which selection is disruptive (see Appendix D.1.1 for details). **d.** Equilibrium dispersal probabilities toward which the population converges under directional selection. Socially mediated dispersal evolves such that dispersal increases with the number k of cooperators in the group. Parameters used: $N = 4$, $\delta = 3$, $b = 1$, $C = 0.3$, $u = 0.75$.

Supp. Fig. S4) show that the dispersal response to the number of cooperators scales with the effect of cooperation on fecundity, such that $\overline{d}_k - \overline{d}_{k-1}$ scales with $B_k - B_{k-1}$. In fact, the number of offspring that do not disperse remains constant across groups with different numbers of cooperators (Fig. S5). This pattern is consistent with the "constant non-disperser principle" (Crespi and Taylor, 1990) (or the "constant philopater hypothesis", Rodrigues and Gardner, 2016), which states that the evolution of conditional dispersal leads parents to produce the same number of non-dispersing offspring across conditions.

Up to this point, we have focused on directional selection, which acts on trait means $\overline{\boldsymbol{x}}$. If these means converge to singular values, genetic (co)variances captured by **G** then start changing (Débarre et al., 2014; Mullon and Lehmann, 2019). In particular, analysing the change in **G** allows us to determine whether the population will either remain unimodally distributed around the singular strategy owing to stabilising selection, or instead become polymorphic owing to disruptive selection (Débarre et al., 2014; Mullon and Lehmann, 2019; Appendix D.1.1 here for method). We find that when the benefits of the common good saturate rapidly, selection is stabilising (Supp. Fig. S3d). However, when the benefits of cooperation saturate slowly (denoted by stars in Fig. 7d and Supp. Fig. S3d), selection is disruptive. Selection may also be disruptive when dispersal is unconditional, but the range of parameter values for which selection is disruptive is larger under socially mediated dispersal (compare grey and orange starred values in Fig. 7c and Fig. S3d).

To determine the evolutionary outcome under disruptive selection, we performed individual-based simulations (Appendix E.1 for the procedure). These simulations show the emergence of two morphs: full cooperators (expressing $z = 1$) and full defectors (expressing $z = 0$), each with distinct dispersal strategies (Fig. 8). Cooperators are overall less dispersive than defectors, such that they preferentially cooperate with relatives, whereas defectors preferentially defect against non-relatives, consistent with previous studies under unconditional dispersal (Mullon et al., 2018). However, we find here that the dispersal reaction norms of the two morphs also differ: the response to the number of cooperators is steeper for cooperators than for defectors (Fig. 8b). When cooperators are rare in a group, they produce few related competitors locally, so their offspring benefit from remaining in their natal group. As the number of cooperators increases, both relatedness and local offspring production increase, intensifying kin competition and favouring greater dispersal. Defectors respond less strongly because their relatedness to other group members decreases as the number of cooperators increases. Although higher cooperation increases local productivity, which favours exporting more offspring, this effect quickly saturates, resulting in a weaker dispersal response (see Fig. S7 for further comparisons).

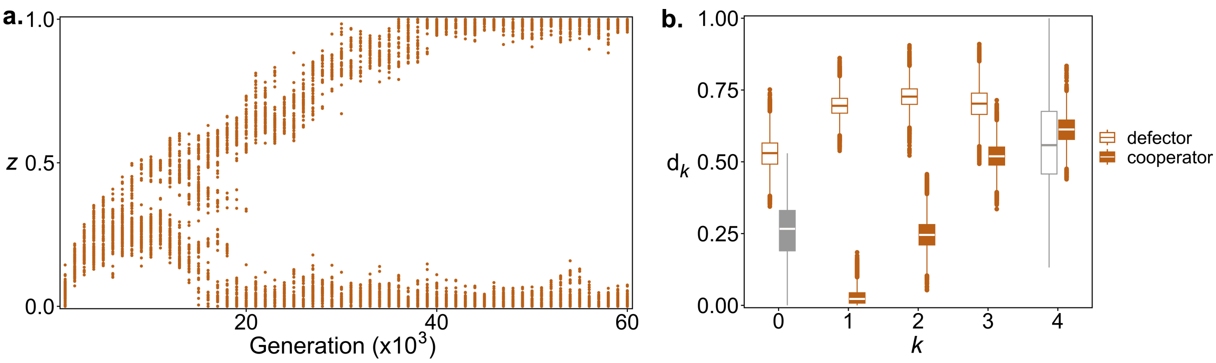

Figure 8: Emergence and coexistence of pure cooperators and pure defectors in individual-based simulations (Appendix E.1 for the procedure). **a.** Evolutionary dynamics of the probability of cooperating z. Each dot represents the trait value of a randomly sampled individual every 1,000 generations. The trait distribution splits into two branches, corresponding to near full cooperators ($z \approx 1$) and near full defectors ($z \approx 0$). **b.** Conditional dispersal strategies d_k expressed by full cooperators (filled boxes) and full defectors (empty boxes). Each boxplot shows the distribution of trait values in the population at 50 time points separated by 1,000 generations after the emergence of the two types. Cooperators disperse less than defectors and respond more strongly to their social environment. Note that cooperators are rarely found in groups with no cooperators (because their parent is very likely to have cooperated), so selection on trait d_0 is very weak in cooperators, which results in large variance in this trait (the boxplot is therefore greyed out). Conversely, defectors are rarely found in groups composed only of cooperators, so selection on their trait d_N is very weak (the boxplot is also greyed out). Parameters used: $N = 4$, $\delta = 3$, $b = 1$, $C = 0.3$, $u = 0.75$, $c_d = 0.15$.

4 Discussion

It may seem intuitive that individuals should leave groups with few cooperators, since social behaviour is heritable and poor social environments are likely to persist (Aktipis, 2004). However, this intuition neglects the effect of kin competition. When cooperation increases local offspring production, remaining in a cooperative group intensifies competition among relatives (Taylor, 1992; Frank, 1998; Rousset, 2004). In this case, dispersing more in the presence of cooperators reduces kin competition and increases inclusive fitness (Kanwal and Gardner, 2022). In our analysis, this effect selects for a counterintuitive reaction norm: individuals evolve to leave when others are cooperative. This socially-mediated dispersal reduces the inclusive fitness cost of helping by limiting local competition among kin, thereby allowing cooperation to spread even when its direct benefits are insufficient to favour it on their own. As cooperation increases, it further strengthens selection for conditional dispersal, establishing a positive feedback between social behaviour and dispersal strategy.

The evolution of socially-mediated dispersal favours cooperation across a range of social scenarios. When the benefits of cooperation increase linearly, cooperation can invade and fix even when it is entirely disfavoured under fixed dispersal (Fig. 3c). When benefits accelerate, as when individuals coordinate or complement each other (Sumpter, 2010), cooperation is typically disfavoured when rare (Peña et al., 2015). In this case, the evolution of dispersal plasticity lowers the threshold level of cooperation above which invasion is likely (Fig. 5c) and can even reduce this threshold to zero (Fig. 6b). When benefits saturate, so that additional cooperators yield diminishing returns (Foster, 2004; Doebeli and Hauert, 2005; Sumpter, 2010), socially-mediated dispersal leads to higher equilibrium levels of cooperation, and in some cases favours full cooperation (Fig. 7c).

The boost for cooperation provided by the evolution of socially-mediated dispersal depends on the cost of dispersal, with stronger effects on cooperation when the cost is high. This is because high dispersal cost increases kin competition, so that dispersing more in cooperative groups becomes especially effective at alleviating competition among kin and increasing the inclusive fitness benefits of helping. When dispersal is inexpensive, by contrast, kin competition is weaker, and socially-mediated dispersal may even facilitate the spread of defection. These results are consistent with the idea that cooperation is more likely to evolve in harsh or unpredictable environments, where successful dispersal is rare (Emlen, 1982). However, the mechanism here is different from commonly invoked explanations: cooperation is favoured not because dispersal is blocked and helping is the best available option, but because cooperation increases kin competition, which can then be relieved through socially-mediated dispersal.

The joint evolution of cooperation and socially-mediated dispersal can lead to a stable polymorphism, where individuals that always cooperate coexist with individuals that always defect (Fig. 8a). Full cooperators and full defectors evolve distinct dispersal strategies: cooperators increase dispersal sharply in cooperative groups, whereas defectors disperse more on average and respond more weakly to the social environment (Fig. 8b). Previous models have shown that social polymorphisms can be stabilised by differences in unconditional dispersal such that cooperators interact more often with relatives and defectors with non-relatives (Wakano et al., 2009; Purcell et al., 2012; Parvinen, 2013; Mullon et al., 2018). Our results show that conditional dispersal can play the same role, but here the resulting variation in dispersal arises not solely due to genetic differences but from gene-by-environment interactions. Such interactions have been documented in several systems, including the lizard *Uta stansburiana* (Sinervo et al., 2006), the plant *Arabidopsis thaliana* (Donohue et al., 2005), and the ciliate *Tetrahymena thermophila* (Pennekamp et al., 2014; Matthysen, 2012), though none involve helping by others as an environmental factor.

Our result that socially mediated dispersal can allow cooperation to evolve under conditions where it would otherwise be absent raises the question of how such dispersal strategies arise in the first place. In our model, we showed that this could occur through genetic drift. However, it seems unlikely that the sensory or regulatory machinery underlying such conditional behaviours would evolve entirely through neutral processes. A more plausible scenario is that initial plasticity in dispersal evolved in response to other environmental or ecological cues, such as local density, and was subsequently co-opted for other cues. Previous models have shown that density-dependent dispersal can evolve to reduce kin competition under demographic stochasticity (Crespi and Taylor, 1990; Kisdi, 2004), and such plasticity has been documented in many systems (Matthysen, 2005). If cooperation increases local density–as it does in our model by increasing offspring production–then dispersal cues based on density may indirectly correlate with the number of cooperators.

Various other factors may also correlate with the number of cooperators or the level of cooperation in a group. Examples include the body mass of individuals in the group (because helping increases the mass of recipients, Hatchwell et al., 2004; Russell et al., 2007; Lloyd et al., 2009; or because larger individuals help more, Rotics et al., 2025; Siegmann et al., 2021), the amount of public good available in the group (West and Buckling, 2003; Grinsted and Field, 2018), or the state of shared resources.

Such external cues can be detected by organisms ranging from bacteria to large vertebrates and can influence their movement decisions (Daniels et al., 2004; Debeffe et al., 2012; Clobert et al., 2012). In these cases, conditional dispersal may promote or hinder the evolution of cooperation, depending on whether dispersal increases or decreases with cues correlated with cooperative behaviour.

More generally, our model assumes that individuals can adjust their dispersal behaviour in response to social traits expressed by others, in particular to the level of helping in the group. To our knowledge, direct evidence that the amount of helping in a group affects dispersal decisions remains limited. However, helping can influence dispersal indirectly, through correlated cues such as group size or body condition. In meerkats, for example, helping increases group size by boosting reproductive success and survival, and subordinates are more likely to disperse from larger groups (Ozgul et al., 2014; Maag et al., 2018). Helpers also increase pup mass at independence, and heavier pups disperse earlier (Russell et al., 2007). Similar patterns are found in banded mongooses, where helping behaviour increases group size and is associated with higher rates of eviction and dispersal, especially under resource limitation (Cant et al., 2001; Nichols et al., 2012; Cant, 2003; Cant et al., 2013). Although it remains unclear whether dispersal is an adaptive response to the number of helpers *per se*, these examples suggest that helping behaviours can influence dispersal indirectly, which could facilitate their coevolution.

5 Conclusion

In conclusion, kin selection provides a general explanation for the evolution of helping within social groups across the tree of life, and limited dispersal is often invoked as the ecological condition that generates the necessary kin structure. But limited dispersal is a double-edged sword: while it increases relatedness, it can also intensify kin competition. Here, we have shown that when dispersal evolves as a socially-mediated trait, this trade-off can be offset–and even reversed–via a feedback between social and dispersal behaviours, such that more cooperative individuals increase the production of related offspring while avoiding competition among them. Counterintuitively, the strategy that promotes cooperation is one in which individuals disperse more when their social partners are more cooperative.

References

Ajar, É. 2003. Analysis of disruptive selection in subdivided populations. *BMC Evolutionary Biology* 3:22.

Aktipis, C. A. 2004. Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology* 231:249–260.

Avila, P. and C. Mullon. 2023. Evolutionary game theory and the adaptive dynamics approach: adaptation where individuals interact. *Philosophical Transactions of the Royal Society B: Biological Sciences* 378:20210502.

Bhatia, R. 2007. Positive Definite Matrices. Princeton University Press, Princeton, NJ.

Cant, M. A. 2003. Patterns of helping effort in co-operatively breeding banded mongooses (Mungos mungo). *Journal of Zoology* 259:115–121.

Cant, M. A., E. Otali, and F. Mwanguhya. 2001. Eviction and dispersal in co-operatively breeding banded mongooses (Mungos mungo). *Journal of Zoology* 254:155–162.

Cant, M. A., E. Vitikainen, and H. J. Nichols. 2013. Demography and social evolution of banded mongooses. In *Advances in the Study of Behavior*, vol. 45, pp. 407–445. Elsevier.

Clobert, J., M. Baguette, T. G. Benton, and J. M. Bullock (eds.). 2012. Dispersal Ecology and Evolution. Oxford University Press, Oxford.

Clobert, J., R. A. Ims, and F. Rousset. 2004. Causes, mechanisms and consequences of dispersal. In Hanski, I. and O. E. Gaggiotti (eds.), *Ecology, Genetics and Evolution of Metapopulations*, pp. 307–335. Elsevier Academic Press, Amsterdam.

Clobert, J., J.-F. Le Galliard, J. Cote, S. Meylan, and M. Massot. 2009. Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters* 12:197–209.

Clutton-Brock, T. 2002. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* 296:69–72.

Clutton-Brock, T. 2021. Social evolution in mammals. *Science* 373:eabc9699.

Comins, H. N., W. D. Hamilton, and R. M. May. 1980. Evolutionarily stable dispersal strategies. *Journal of Theoretical Biology* 82:205–230.

Crespi, B. J. and P. D. Taylor. 1990. Dispersal rates under variable patch density. *The American Naturalist* 135:48–62.

Daniels, R., J. Vanderleyden, and J. Michiels. 2004. Quorum sensing and swarming migration in bacteria. *FEMS Microbiology Reviews* 28:261–289.

Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. Multidimensional (co)evolutionary stability. *The American Naturalist* 184:158–171.

Debeffe, L., N. Morellet, B. Cargnelutti, B. Lourtet, R. Bon, J.-M. Gaillard, and A. J. M. Hewison. 2012. Condition-dependent natal dispersal in a large herbivore: heavier animals show a greater propensity to disperse and travel further. *Journal of Animal Ecology* 81:1327–1337.

Doebeli, M. and C. Hauert. 2005. Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology Letters* 8:748–766.

Donohue, K., C. R. Polisetty, and N. J. Wender. 2005. Genetic basis and consequences of niche construction: plasticity-induced genetic constraints on the evolution of seed dispersal in Arabidopsis thaliana. *The American Naturalist* 165:537–550.

Dufty, Jr., A. M. and J. R. Belthoff. 2001. Proximate mechanisms of natal dispersal: the role of body condition and hormones. In Clobert, J., E. Danchin, A. A. Dhondt, and J. D. Nichols (eds.), *Dispersal*, pp. 217–229. Oxford University Press, Oxford.

Ekman, J. 2006. Family living among birds. *Journal of Avian Biology* 37:289–298.

El Mouden, C. and A. Gardner. 2008. Nice natives and mean migrants: the evolution of dispersal-dependent social behaviour in viscous populations. *Journal of Evolutionary Biology* 21:1480–1491.

Emlen, S. T. 1982. The evolution of helping. i. an ecological constraints model. *The American Naturalist* 119:29–39.

Foster, K. R. 2004. Diminishing returns in social evolution: the not-so-tragic commons. *Journal of Evolutionary Biology* 17:1058–1072.

Frank, S. A. 1986. Dispersal polymorphisms in subdivided populations. *Journal of Theoretical Biology* 122:303–309.

Frank, S. A. 1998. Foundations of Social Evolution. Princeton University Press, Princeton, NJ.

Gandon, S. 1999. Kin competition, the cost of inbreeding and the evolution of dispersal. *Journal of Theoretical Biology* 200:345–364.

Gandon, S. and Y. Michalakis. 1999. Evolutionarily stable dispersal rate in a metapopulation with extinctions and kin competition. *Journal of Theoretical Biology* 199:275–290.

Gandon, S. and F. Rousset. 1999. Evolution of stepping-stone dispersal rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 266:2507–2513.

Grinsted, L. and J. Field. 2018. Predictors of nest growth: diminishing returns for subordinates in the paper wasp Polistes dominula. *Behavioral Ecology and Sociobiology* 72:88.

Hamilton, W. D. 1963. The evolution of altruistic behavior. *The American Naturalist* 97:354–356.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. i. *Journal of Theoretical Biology* 7:1–16.

Hamilton, W. D. and R. M. May. 1977. Dispersal in stable habitats. *Nature* 269:578–581.

Hatchwell, B. J., A. F. Russell, A. D. MacColl, D. J. Ross, M. K. Fowlie, and A. McGowan. 2004. Helpers increase long-term but not short-term productivity in cooperatively breeding long-tailed tits. *Behavioral Ecology* 15:1–10.

Hochberg, M. E., D. J. Rankin, and M. Taborsky. 2008. The coevolution of cooperation and dispersal in social groups and its implications for the emergence of multicellularity. *BMC Evolutionary Biology* 8:238.

Iwasa, Y., A. Pomiankowski, and S. Nee. 1991. The evolution of costly mate preferences ii. the "handicap" principle. *Evolution* 45:1431–1442.

Kanwal, J. and A. Gardner. 2022. Population viscosity promotes altruism under density-dependent dispersal. *Proceedings of the Royal Society B* 289:20212668.

Karlin, S. 1968. Equilibrium Behaviour of Population Genetic Models with Non-Random Mating. *J. Appl. Probab.* 5:231–313.

Kisdi, É. 2004. Conditional dispersal under kin competition: extension of the hamilton–may model to brood size-dependent dispersal. *Theoretical Population Biology* 66:369–380.

Lambin, X., J. Aars, and S. B. Piertney. 2001. Dispersal, intraspecific competition, kin competition and kin facilitation: a review of the empirical evidence. In Clobert, J., E. Danchin, A. A. Dhondt, and J. D. Nichols (eds.), *Dispersal*, pp. 110–122. Oxford University Press, Oxford.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33:402–416.

Lawson Handley, L. and N. Perrin. 2007. Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology* 16:1559–1578.

Lehmann, L. and F. Balloux. 2007. Natural selection on fecundity variance in subdivided populations: kin selection meets bet hedging. *Genetics* 176:361–377.

Lehmann, L. and L. Keller. 2006. The evolution of cooperation and altruism–a general framework and a classification of models. *Journal of Evolutionary Biology* 19:1365–1376.

Lehmann, L. and F. Rousset. 2010. How life history and demography promote or inhibit the evolution of helping behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:2599–2617.

Leimar, O. 2005. The evolution of phenotypic polymorphism: randomized strategies versus evolutionary branching. *The American Naturalist* 165:669–681.

Leimar, O. 2009. Multidimensional convergence stability. *Evolutionary Ecology Research* 11:191–208.

Lloyd, P., W. A. Taylor, M. A. Du Plessis, and T. E. Martin. 2009. Females increase reproductive investment in response to helper-mediated improvements in allo-feeding, nest survival, nestling provisioning and post-fledging survival in the Karoo scrub-robin Cercotrichas coryphaeus. *Journal of Avian Biology* 40:400–411.

Maag, N., G. Cozzi, T. Clutton-Brock, and A. Ozgul. 2018. Density-dependent dispersal strategies in a cooperative breeder. *Ecology* 99:1932–1941.

Matthysen, E. 2005. Density-dependent dispersal in birds and mammals. *Ecography* 28:403–416.

Matthysen, E. 2012. Multicausality of dispersal: a review. In Clobert, J., M. Baguette, T. G. Benton, and J. M. Bullock (eds.), *Dispersal Ecology and Evolution*, pp. 3–18. Oxford University Press, Oxford.

Maynard Smith, J. 1964. Group selection and kin selection. *Nature* 201:1145–1147.

Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary Stability of Jointly Evolving Traits in Subdivided Populations. *The American Naturalist* 188:175–195.

Mullon, C., L. Keller, and L. Lehmann. 2018. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology & Evolution* 2:132–140.

Mullon, C. and L. Lehmann. 2019. An evolutionary quantitative genetics model for phenotypic (co)variance under limited dispersal, with an application to socially synergistic traits. *Evolution* 73:1695–1728.

Mullon, C., J. Y. Wakano, and H. Ohtsuki. 2021. Coevolutionary dynamics of genetic traits and their long-term extended effects under non-random interactions. *Journal of Theoretical Biology* 525:110750.

Nichols, H., M. Bell, S. Hodge, and M. Cant. 2012. Resource limitation moderates the adaptive suppression of subordinate breeding in a cooperatively breeding mongoose. *Behavioral Ecology* 23:635–642.

Ohtsuki, H. 2010. Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. *Evolution* 64:3344–3353.

Ozgul, A., A. W. Bateman, S. English, T. Coulson, and T. H. Clutton-Brock. 2014. Linking body mass and group dynamics in an obligate cooperative breeder. *Journal of Animal Ecology* 83:1357–1366.

Parvinen, K. 2013. Joint evolution of altruistic cooperation and dispersal in a metapopulation of small local populations. *Theoretical Population Biology* 85:12–19.

Peña, J., G. Nöldeke, and L. Lehmann. 2015. Evolutionary dynamics of collective action in spatially structured populations. *Journal of Theoretical Biology* 382:122–136.

Pennekamp, F., K. A. Mitchell, A. Chaine, and N. Schtickzelle. 2014. Dispersal propensity in Tetrahymena thermophila ciliates—a reaction norm perspective. *Evolution* 68:2319–2330.

Poethke, H. J. and T. Hovestadt. 2002. Evolution of density–and patch–size–dependent dispersal rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269:637–645.

Prigent, I. and C. Mullon. 2023. The molding of intraspecific trait variation by selection under ecological inheritance. *Evolution* 77:2144–2161.

Purcell, J., A. Brelsford, and L. Avilés. 2012. Co-evolution between sociality and dispersal: the role of synergistic cooperative benefits. *Journal of Theoretical Biology* 312:44–54.

Queller, D. C. 1992. Does population viscosity promote kin selection? *Trends in Ecology & Evolution* 7:322–324.

Riedman, M. L. 1982. The evolution of alloparental care and adoption in mammals and birds. *The Quarterly Review of Biology* 57:405–435.

Rodrigues, A. M. and A. Gardner. 2016. The constant philopater hypothesis: a new life history invariant for dispersal evolution. *Journal of Evolutionary Biology* 29:153–166.

Ronce, O., J. Clobert, and M. Massot. 1998. Natal dispersal and senescence. *Proceedings of the National Academy of Sciences* 95:600–605.

Ronce, O., S. Gandon, and F. Rousset. 2000. Kin selection and natal dispersal in an age-structured population. *Theoretical Population Biology* 58:143–159.

Rotics, S., H. M. Bensch, Y. S. Resheff, T. Clutton-Brock, and M. Zöttl. 2025. Workload distribution in wild Damaraland mole-rat groups. *Philosophical Transactions of the Royal Society B: Biological Sciences* 380:20230276.

Rousset, F. 2004. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, NJ.

Russell, A., A. Young, G. Spong, N. Jordan, and T. Clutton-Brock. 2007. Helpers increase the reproductive potential of offspring in cooperative meerkats. *Proceedings of the Royal Society B: Biological Sciences* 274:513–520.

Schreiber, S. J. 2015. Unifying within-and between-generation bet-hedging theories: an ode to jh gillespie. *The American Naturalist* 186:792–796.

Siegmann, S., R. Feitsch, D. W. Hart, N. C. Bennett, D. J. Penn, and M. Zöttl. 2021. Naked mole-rats (Heterocephalus glaber) do not specialise in cooperative tasks. *Ethology* 127:850–864.

Sinervo, B., R. Calsbeek, T. Comendant, C. Both, C. Adamopoulou, and J. Clobert. 2006. Genetic and maternal determinants of effective dispersal: the effect of sire genotype and size at birth in side-blotched lizards. *The American Naturalist* 168:88–99.

Skutch, A. F. 1961. Helpers among birds. *The Condor* 63:198–226.

Starrfelt, J. and H. Kokko. 2012. Bet-hedging – a triple trade-off between means, variances and correlations. *Biological Reviews* 87:742–755.

Sumpter, D. J. 2010. Collective animal behavior. Princeton University Press.

Taylor, P. D. 1988. An inclusive fitness model for dispersal of offspring. *Journal of Theoretical Biology* 130:363–378.

Taylor, P. D. 1992. Altruism in viscous populations - an inclusive fitness model. *Evolutionary Ecology* 6:352–356.

Van Cleve, J. 2015. Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology* 103:2–26.

Viera, C. and I. Agnarsson. 2017. Parental care and sociality. In Viera, C. and M. O. Gonzaga (eds.), *Behaviour and Ecology of Spiders: Contributions from the Neotropical Region*, pp. 351–381. Springer International Publishing, Cham.

Wakano, J. Y. and L. Lehmann. 2014. Evolutionary branching in deme-structured populations. *Journal of Theoretical Biology* 351:83–95.

Wakano, J. Y., M. A. Nowak, and C. Hauert. 2009. Spatial dynamics of ecological public goods. *Proceedings of the National Academy of Sciences* 106:7910–7914.

West, S. A. and A. Buckling. 2003. Cooperation, virulence and siderophore production in bacterial parasites. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270:37–44.

West, S. A., A. S. Griffin, and A. Gardner. 2007. Evolutionary explanations for cooperation. *Current Biology* 17:R661–R672.

West, S. A., A. S. Griffin, A. Gardner, and S. P. Diggle. 2006. Social evolution theory for microorganisms. *Nature Reviews Microbiology* 4:597–607.

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

Wisenden, B. D. 1999. Alloparental care in fishes. *Reviews in Fish Biology and Fisheries* 9:45–70.

Wolfram Research, Inc. 2016. Mathematica. Champaign, Illinois.

Yip, E. C. and L. S. Rayor. 2014. Maternal care and subsocial behaviour in spiders. *Biological Reviews* 89:427–449.

A Method

We are interested in the joint evolution of $N+2$ traits (where recall N is the number of individuals in the group): the probability z of cooperating and contributing to the common good and $N+1$ dispersal probabilities conditional on the number of cooperators, $d_0, d_1, \dots, d_N$. Since all these traits are probabilities, each must be between zero and one. The phenotypic space we are considering is thus the unit hypercube in $(N+2)$-dimensions, which we denote by $\mathscr{X}$. We arbitrarily order traits such that z is on the first axis of $\mathscr{X}$, d_0 on the second, d_1 on the third and so forth. Starting with a population with average phenotype $\overline{\boldsymbol{x}} = (\overline{z}, \overline{d}_0, \dots, \overline{d}_N) \in \mathscr{X}$ and small phenotypic variance, the coevolution of multiple traits that experience rare mutations with weak effects can be decomposed in two steps in the island model of dispersal (Mullon and Lehmann, 2019), which we briefly describe in this section.

A.1 Directional selection

A.1.1 Dynamics of the vector of average trait values

First, the population evolves under directional selection. The change over one generation in the population average phenotype due to directional selection is given by

$$\Delta\overline{\boldsymbol{x}} = \mathbf{G}\,\boldsymbol{s}(\overline{\boldsymbol{x}}), \tag{A-4}$$

where $\mathbf{G}$ is the variance-covariance matrix (i.e. the (a, b) entry of this matrix, G_{ab}, is the (co)variance between traits indexed a and b); and $\boldsymbol{s}(\overline{\boldsymbol{x}}) = (s_1(\overline{\boldsymbol{x}}), s_2(\overline{\boldsymbol{x}}), \dots, s_{N+2}(\overline{\boldsymbol{x}}))$ is a vector that collects the selection gradient on each trait (i.e., the ath entry of this vector, $\boldsymbol{s}(\overline{\boldsymbol{x}})_a = s_a(\overline{\boldsymbol{x}})$ is the selection gradient on trait indexed a). The selection gradient vector points in the direction favoured by selection in phenotypic space when the population mean is $\overline{\boldsymbol{x}}$, and the effect of this directional selection on the mean of each trait is constrained by the available genetic variation, which is captured by $\mathbf{G}$ in eq. (1) (Lande, 1979). The variance-covariance matrix $\mathbf{G}$ also evolves under directional selection, but when mutations have weak effects, this evolution is slow compared to the evolution of mean phenotype $\overline{\boldsymbol{x}}$ (Lande, 1979; Débarre et al., 2014; Mullon and Lehmann, 2019). The evolution of $\mathbf{G}$ can therefore be ignored until directional selection has vanished.

A.1.2 Convergence stability

As the population evolves under directional selection, it may eventually converge to a singular phenotype $\boldsymbol{x}^*$ that sits within the interior of the phenotypic space. Such an interior singular point satisfies

$$\boldsymbol{s}(\boldsymbol{x}^*) = \mathbf{0}. \tag{A-5}$$

For the population mean to converge to $\boldsymbol{x}^*$ whatever the variance-covariance matrix $\mathbf{G}$ (under eq. A-4), it is sufficient that the $(N+2) \times (N+2)$ Jacobian matrix $\mathbf{J}(\boldsymbol{x}^*)$ with (a, b)-entry

$$J_{ab}(\boldsymbol{x}^*) = \left.\frac{\partial s_a(\boldsymbol{x})}{\partial x_b}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}, \tag{A-6}$$

is negative-definite, i.e., such that $(\mathbf{J}(\boldsymbol{x}^*) + \mathbf{J}(\boldsymbol{x}^*)^{\mathrm{T}})/2$ has eigenvalues that are all negative (Leimar, 2005). A singular point of this type is commonly referred to as a (strongly) convergence stable strategy (Leimar, 2009).

A.1.3 Selection gradient under limited dispersal

In the island model of dispersal, the selection gradient on a trait ordered as $a \in \{1, \dots, N+2\}$ (or trait a for short) is given by

$$s_a(\overline{\boldsymbol{x}}) = \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})}{\partial x_{i,a}}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}} \\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} + (N-1)\, r_2^{\circ}(\overline{\boldsymbol{x}}) \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})}{\partial x_{j,a}}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}} \\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}}, \tag{A-7}$$

where $w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})$ is the fitness (i.e., the expected number of successful offspring) produced by a focal individual arbitrarily indexed as individual $i \in \{1, \dots N\}$ with phenotype $\boldsymbol{x}_i = (z_i, d_{0,i}, d_{1,i}, \dots, d_{N,i})$, when the other individuals in the focal group have phenotypes $\boldsymbol{x}_{-i} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \dots, \boldsymbol{x}_N)$ and the rest of the population is considered to be monomorphic for $\overline{\boldsymbol{x}}$ (Rousset, 2004). The first fitness derivative in eq. (A-7) is the effect of an infinitesimal change in trait a in the focal individual on its own fitness, and the second fitness derivative, the effect of a trait change in a neighbour indexed as $j \in \{1, \dots N\}$ (with $j \neq i$) on focal fitness. This latter effect is weighted by the neutral coefficient of pairwise relatedness, $r_2^{\circ}(\overline{\boldsymbol{x}})$, which in the infinite island model of dispersal is equivalent to the probability that in a population monomorphic for $\overline{\boldsymbol{x}}$ and where individuals have the same fecundity, two individuals sampled without replacement from a group are identical-by-descent (IBD). We detail how to compute fitness and relatedness for our model in sections B.1 and B.2 below.

A.2 Stabilising and disruptive selection

A.2.1 Dynamics of the variance-covariance matrix

Once the population average has converged to a singular strategy, $\bar{\boldsymbol{x}} = \boldsymbol{x}^*$, the phenotypic distribution may either remain unimodally distributed around this average under stabilising selection, or become polymorphic due to disruptive selection (Leimar, 2009; Débarre et al., 2014; Mullon and Lehmann, 2019). These outcomes depend on the evolution of the variance-covariance matrix whose change over one generation at a convergence stable strategy is

$$\Delta\mathbf{G} = \mathbf{M} + \mathbf{G}\,\mathbf{H}(\boldsymbol{x}^*)\,\mathbf{G}, \tag{A-8}$$

where $\mathbf{M}$ is a variance-covariance matrix that captures the effects of mutation (its (a,b)-entry M_{ab} is the product between the probability that an offspring mutates and the (co)variance between the effect of a mutation on traits a and b), and $\mathbf{H}(\boldsymbol{x}^*)$ is a $(N+2)\times(N+2)$ Hessian matrix that captures whether selection is stabilising or disruptive when the population expresses the average phenotype $\boldsymbol{x}^*$. Specifically, when $\mathbf{H}(\boldsymbol{x}^*)$ is negative-definite, selection is stabilising. In this case, the dynamics of $\mathbf{G}$ (eq. A-8) reach an equilibrium $\mathbf{G}^*$ where the forces of mutation and stabilising selection are balanced:

$$\mathbf{G}^* = \mathbf{M}\left[\mathbf{M}^{-1}\left(-\mathbf{H}(\boldsymbol{x}^*)\right)^{-1}\right]^{-1/2}, \tag{A-9}$$

with the operation $\boldsymbol{X}^{-1/2}$ denoting the square root of $\boldsymbol{X}$ such that all the eigenvalues of $\boldsymbol{X}^{-1/2}$ are positive (Bhatia, 2007). The distribution of phenotypes in the population at equilibrium then has mean $\boldsymbol{x}^*$ and variance-covariance $\mathbf{G}^*$. By contrast, when $\mathbf{H}(\boldsymbol{x}^*)$ has at least one positive eigenvalue, selection is disruptive: $\mathbf{G}$ diverges under eq. (A-8), the phenotypic variance between individuals increases and the population becomes bimodally distributed (Leimar, 2009). In this case, we use individual-based simulations (described in Appendix E.1) to investigate the nature of the different morphs that evolve.

A.2.2 Hessian matrix under limited dispersal

In the infinite island model of dispersal, the Hessian matrix $\mathbf{H}(\boldsymbol{x}^*)$ has (a,b)-entry given by

$$\begin{aligned}
h_{ab}(\boldsymbol{x}^*) = {} & \left.\frac{\partial^2 w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{i,a} x_{i,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} + (N-1) r_2^\circ(\boldsymbol{x}^*) \left.\frac{\partial^2 w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{j,a} x_{j,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} \\
& + (N-1) r_2^\circ(\boldsymbol{x}^*) \left(\left.\frac{\partial^2 w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{i,a} x_{j,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} + \left.\frac{\partial^2 w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{j,a} x_{i,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} \right) \\
& + (N-1)(N-2) r_3^\circ(\boldsymbol{x}^*) \left.\frac{\partial^2 w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{j,a} x_{k,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} \\
& + (N-1) \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{j,a}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} \left.\frac{\partial r_2(\boldsymbol{y}, \boldsymbol{x}^*)}{\partial y_b}\right|_{\boldsymbol{y}=\boldsymbol{x}^*} + (N-1) \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}^*)}{\partial x_{j,b}}\right|_{\substack{\boldsymbol{x}_i=\boldsymbol{x}^* \\ \boldsymbol{x}_{-i}=\boldsymbol{x}^*}} \left.\frac{\partial r_2(\boldsymbol{y}, \boldsymbol{x}^*)}{\partial y_a}\right|_{\boldsymbol{y}=\boldsymbol{x}^*},
\end{aligned} \tag{A-10}$$

where $r_3^\circ(\boldsymbol{x}^*)$ is the neutral coefficient of three-way relatedness, i.e. the probability that in a population monomorphic for $\boldsymbol{x}^*$ and where individuals have the same fecundity, three individuals sampled without replacement from a group have a common ancestor that resided in that same group; and $\partial r_2(\boldsymbol{y}, \boldsymbol{x}^*)/\partial y_a$ is the effect of a change in trait a on pairwise relatedness (Mullon et al., 2016; Mullon and Lehmann, 2019). The different second-order fitness derivatives in eq. (A-10) measure the effects on the fitness of a focal individual of joint changes in traits a and b within individuals (first line of eq. A-10) and between individuals (second and third line of eq. A-10) that belong to the same group. These effects are weighted by relevant coefficients of relatedness between the focal individual and the individuals in which trait changes occur. The last line of eq. (A-10) consists of the product between the indirect effect of one trait with the effect of the other trait on relatedness (Mullon et al., 2016; Mullon and Lehmann, 2019; Avila and Mullon, 2023 for more details and interpretation).

B Fitness and relatedness

The approach described in Appendix A requires: (i) the individual fitness function, $w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})$; (ii) neutral pairwise and three-way relatedness coefficients, $r_2^\circ(\overline{\boldsymbol{x}})$ and $r_3^\circ(\overline{\boldsymbol{x}})$; and (iii) the first order effect of traits on relatedness, $\partial r_2(\boldsymbol{y}, \boldsymbol{x}^*)/\partial y_a$. We describe these quantities in this appendix.

B.1 Fitness

B.1.1 Conditional fitness

Suppose that at some generation in a given group, we know how each individual has behaved towards the common good, i.e., we know whether individual $i \in \{1,\dots,N\}$ has cooperated and contributed to the common good or not. To describe this outcome, we let $c_i = 1$ if individual i has cooperated, and 0 if it has not; and collect these indicator variables for all individuals in the focal group into the vector $\boldsymbol{c} = (c_1, c_2, \dots, c_N)$. The number of cooperators in the group is $k = \sum_{j=1}^{N} c_j$. The fecundity of an individual in a group with k cooperators, given its own contribution c_i, can be written as

$$f(k, c_i) = 1 + \delta(B_k/N - Cc_i), \tag{B-1}$$

where δ scales the effect of cooperation on fecundity. With this notation, the fitness of a focal individual i conditional on $\boldsymbol{c}$ can then be expressed as the sum,

$$w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c}) = \underbrace{N\frac{(1-d_{k,i})f(k,c_i)}{\sum_{j=1}^{N}(1-d_{k,j})f(k,c_j) + F_{\text{d}}(\overline{\boldsymbol{x}})}}_{w_{\text{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})} + \underbrace{\sum_{k'=0}^{N} N\frac{d_{k,i}(1-c_{\text{d}})f(k,c_i)}{F_{\text{p}}(\overline{\boldsymbol{x}}, k') + F_{\text{d}}(\overline{\boldsymbol{x}})}\Pr(k'|N, \overline{\boldsymbol{x}})}_{w_{\text{d}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}, \tag{B-2}$$

where the first term, $w_{\text{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})$, is the expected number of offspring that settle locally in the focal group (i.e., the philopatric component of fitness), and the second, $w_{\text{d}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})$, the expected number of offspring that settle in other groups through dispersal (i.e., the dispersal component of fitness), both of these quantities being conditional on $\boldsymbol{c}$. We detail both of these quantities in the next two sections.

B.1.2 Conditional philopatric fitness

The philopatric component of fitness consists of the ratio of the number of offspring of the focal individual that remain in the focal group, $(1-d_{k,i})f(k,c_i)$, to the total number of offspring that enter competition in the group. This latter term is composed of two terms, $\sum_{j=1}^{N}(1-d_{k,j})f(k,c_j)$, which is the total number of offspring that are born and remain in the focal group, and $F_{\text{d}}(\overline{\boldsymbol{x}})$, which is the number of offspring that immigrated into the focal group from other groups (and that are considered to be monomorphic for $\overline{\boldsymbol{x}}$). The number of immigrant offspring can be computed as

$$F_{\text{d}}(\overline{\boldsymbol{x}}) = \sum_{k'=0}^{N} \underbrace{\overline{d}_{k'}(1-c_{\text{d}})\left\{k' f(k',1) + (N-k') f(k',0)\right\}}_{\substack{\text{offspring coming from} \\ \text{a group with } k' \text{ cooperators}}} \Pr(k'|N,\overline{\boldsymbol{x}}), \tag{B-3}$$

where

$$\Pr(k'|N,\overline{\boldsymbol{x}}) = \binom{N}{k'} \overline{z}^{k'} (1-\overline{z})^{N-k'}, \tag{B-4}$$

is the frequency of groups with k' cooperators in a population monomorphic for $\overline{\boldsymbol{x}}$ (or equivalently, the probability that there are k' contributors in a group monomorphic for $\overline{\boldsymbol{x}}$).

B.1.3 Conditional dispersal fitness

The dispersal component of fitness, $w_{\text{d}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})$ (eq. B-2), consists of the expected number of offspring of the focal individual that successfully settle in a group with k' cooperators:

$$N \frac{d_{k,i}(1-c_{\text{d}}) f(k, c_i)}{F_{\text{p}}(\overline{\boldsymbol{x}}, k') + F_{\text{d}}(\overline{\boldsymbol{x}})}, \tag{B-5}$$

averaged over the distribution $\Pr(k'|N,\overline{\boldsymbol{x}})$ of groups in a population monomorphic for $\overline{\boldsymbol{x}}$ (eq. B-4). The ratio in eq. (B-5) consists of the number of offspring produced by the focal individual that disperse and survive dispersal, $d_{k,i}(1-c_{\text{d}}) f(k, c_i)$, to the total number of offspring that enter competition in a group with k' contributors, which is composed of

$$F_{\text{p}}(\overline{\boldsymbol{x}}, k') = (1-\overline{d}_{k'})\left\{k' f(k',1) + (N-k') f(k',0)\right\} \tag{B-6}$$

offspring that have remained in their natal group (with k' cooperators), and $F_{\text{d}}(\overline{\boldsymbol{x}})$ that have emigrated from other groups (see eq. B-3).

B.1.4 Unconditional fitness

Unconditional fitness is calculated by marginalising conditional fitness, $w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{x}|\boldsymbol{c})$ (eq. B-2), over the probability mass function of contributions to the common good in the focal group,

$$w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}) = \sum_{\boldsymbol{c}\in\mathscr{C}} w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c}) \Pr(\boldsymbol{c}|\boldsymbol{x}), \tag{B-7}$$

where $\mathscr{C} = \{(c_1, c_2, \dots, c_N) : c_i \in \{0,1\}, i = 1, \dots, N\}$ is the space of all possible outcomes for contribution to the common good. Given all the traits expressed in the group, the probability that the realised contribution to the common good is $\boldsymbol{c}$ is

$$\Pr(\boldsymbol{c}|\boldsymbol{x}) = \prod_{j=1}^{N} z_j^{c_j} (1 - z_j)^{1-c_j}. \tag{B-8}$$

B.2 Relatedness

The analysis of selection described in Appendix A also requires the pairwise and three-way neutral relatedness coefficients, $r_2^\circ(\overline{\boldsymbol{x}})$ and $r_3^\circ(\overline{\boldsymbol{x}})$, as well as the effect of traits on pairwise relatedness, $\partial r_2(\boldsymbol{y}, \overline{\boldsymbol{x}})/\partial y_a$. We characterise these for our model using standard coalescent arguments (e.g., Karlin, 1968; Rousset, 2004).

B.2.1 Neutral relatedness coefficients

Let $r_{2,t+1}^\circ(\overline{\boldsymbol{x}})$ be the probability that two individuals sampled from a group at generation $t+1$ are identical-by-descent (IBD) in a population monomorphic for $\overline{\boldsymbol{x}}$. This $r_{2,t+1}^\circ(\overline{\boldsymbol{x}})$ can be conditioned on the number k of individuals that cooperated at generation t as,

$$r_{2,t+1}^\circ(\overline{\boldsymbol{x}}) = \sum_{k=0}^{N} r_{2,k,t+1}^\circ(\overline{\boldsymbol{x}}) \Pr(k|N, \overline{\boldsymbol{x}}), \tag{B-9}$$

where $r_{2,k,t+1}^\circ(\overline{\boldsymbol{x}})$ is the probability that two individuals sampled from the group at generation $t+1$ are IBD given k. Assuming that the effect of cooperation on fecundity is weak (i.e. that B_k and C are of the order of a small parameter δ), this $r_{2,k,t+1}^\circ(\overline{\boldsymbol{x}})$ can be expressed as

$$r_{2,k,t+1}^\circ(\overline{\boldsymbol{x}}) = (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \left(\frac{1}{N} + \frac{N-1}{N} r_{2,t}^\circ(\overline{\boldsymbol{x}}) \right) + \mathcal{O}(\delta) \tag{B-10}$$

where

$$m_k^\circ(\overline{\boldsymbol{x}}) = \left. \frac{F_\text{d}(\overline{\boldsymbol{x}})}{F_\text{p}(\overline{\boldsymbol{x}}, k) + F_\text{d}(\overline{\boldsymbol{x}})} \right|_{\delta=0} \tag{B-11}$$

is the probability that a randomly sampled individual from the group is an immigrant in a population monomorphic for $\overline{\boldsymbol{x}}$ and where $\delta = 0$, given that k individuals cooperated in this group at the previous generation, i.e., $m_k^\circ(\overline{\boldsymbol{x}})$ is the (conditional) backward probability of dispersal (Gandon, 1999). The parameter δ tunes the effect of cooperation on fecundity (i.e. is of the order of B_k and C). The right-hand side of eq. (B-10) can be read as follows. For two individuals to share a common ancestor, they must both be philopatric, which occurs with probability $(1-m_k^\circ(\overline{\boldsymbol{x}}))^2$. Then, with probability $1/N$ these two individuals have the same parent, in which case they are IBD, or with probability $(N-1)/N$ they have different parents who themselves share a common ancestor with probability $r_{2,t}^\circ(\overline{\boldsymbol{x}})$.

Substituting eq. (B-10) into eq. (B-9), we then solve for

$$r_{2,t+1}^\circ(\overline{\boldsymbol{x}}) = r_{2,t}^\circ(\overline{\boldsymbol{x}}) = r_2^\circ(\overline{\boldsymbol{x}}), \tag{B-12}$$

obtaining

$$r_2^\circ(\overline{\boldsymbol{x}}) = \frac{\langle(1-m_k^\circ(\overline{\boldsymbol{x}}))^2\rangle}{N-(N-1)\langle(1-m_k^\circ(\overline{\boldsymbol{x}}))^2\rangle} + \mathcal{O}(\delta), \tag{B-13}$$

where

$$\langle(1-m_k^\circ(\overline{\boldsymbol{x}}))^2\rangle = \sum_{k=0}^{N}(1-m_k^\circ(\overline{\boldsymbol{x}}))^2 \Pr(k|N,\overline{\boldsymbol{x}}) = (1-m^\circ(\overline{\boldsymbol{x}}))^2 + \sigma_{\mathrm{m}}^2(\overline{\boldsymbol{x}}) \tag{B-14}$$

is the unconditional probability of sampling two philopatric individuals within the same group under neutrality, with

$$\begin{aligned} m^\circ(\overline{\boldsymbol{x}}) &= \sum_{k=0}^{N} m_k^\circ(\overline{\boldsymbol{x}}) \Pr(k|N,\overline{\boldsymbol{x}}) \\ \sigma_{\mathrm{m}}^2(\overline{\boldsymbol{x}}) &= \sum_{k=0}^{N} (m_k^\circ(\overline{\boldsymbol{x}}) - m^\circ(\overline{\boldsymbol{x}}))^2 \Pr(k|N,\overline{\boldsymbol{x}}) \end{aligned} \tag{B-15}$$

as the mean and variance in the proportion of immigrants in a group. Note that when $\sigma_{\mathrm{m}}^2(\overline{\boldsymbol{x}}) = 0$, eq. (B-13) reduces to the classical expression for pairwise relatedness under the Wright-Fisher model of reproduction and the island model of dispersal (e.g. eq. 12a in Ohtsuki, 2010). Eq.(B-13) shows that variance in the number of immigrants (here arising from socially-mediated dispersal and variation in the number of cooperators) increases relatedness. This is because some groups receive (proportionally) very few immigrants, and these groups contribute disproportionately to relatedness due to terms of the form $(1-m_k^\circ(\overline{\boldsymbol{x}}))^2$.

Following a similar argument, the probability $r_{3,t+1}^\circ(\overline{\boldsymbol{x}})$ that three individuals sampled from a group at generation $t+1$ are IBD in a population monomorphic for $\overline{\boldsymbol{x}}$ can be expressed as

$$r_{3,t+1}^\circ(\overline{\boldsymbol{x}}) = \sum_{k=0}^{N}(1-m_k^\circ(\overline{\boldsymbol{x}}))^3\left(\frac{1}{N^2} + 3\frac{N-1}{N^2}r_{2,k,t}^\circ(\overline{\boldsymbol{x}}) + \frac{(N-1)(N-2)}{N^2}r_{3,k,t}^\circ(\overline{\boldsymbol{x}})\right)\Pr(k|N,\overline{\boldsymbol{x}}) + \mathcal{O}(\delta), \tag{B-16}$$

where $r_{3,k,t}^\circ(\overline{\boldsymbol{x}})$ is the probability that three individuals sampled from the group at generation $t+1$ are

IBD given k (such that $r^{\circ}_{3,t}(\overline{\boldsymbol{x}}) = \sum_{k=0}^{N} r^{\circ}_{3,k,t}(\overline{\boldsymbol{x}})\,\Pr(k|N,\overline{\boldsymbol{x}})$). Solving the above for

$$r^{\circ}_{3,t+1}(\overline{\boldsymbol{x}}) = r^{\circ}_{3,t}(\overline{\boldsymbol{x}}) = r^{\circ}_{3}(\overline{\boldsymbol{x}}), \tag{B-17}$$

yields

$$r^{\circ}_{3}(\overline{\boldsymbol{x}}) = \frac{\langle(1-m^{\circ}_{k}(\overline{\boldsymbol{x}}))^{3}\rangle(1+(N-1)r^{\circ}_{2}(\overline{\boldsymbol{x}}))}{N^{2}-(N-1)(N-2)\langle(1-m^{\circ}_{k}(\overline{\boldsymbol{x}}))^{3}\rangle} + \mathcal{O}(\delta), \tag{B-18}$$

where

$$\langle(1-m^{\circ}_{k}(\overline{\boldsymbol{x}}))^{3}\rangle = \sum_{k=0}^{N}(1-m^{\circ}_{k}(\overline{\boldsymbol{x}}))^{3}\,\Pr(k|N,\overline{\boldsymbol{x}}) = (1-m^{\circ}(\overline{\boldsymbol{x}}))^{3} + 3(1-m^{\circ}(\overline{\boldsymbol{x}}))\sigma^{2}_{\mathrm{m}}(\overline{\boldsymbol{x}}) - \gamma_{\mathrm{m}}(\overline{\boldsymbol{x}}) \tag{B-19}$$

is the unconditional probability of sampling three philopatric individuals within the same group under neutrality, where

$$\gamma_{\mathrm{m}}(\overline{\boldsymbol{x}}) = \sum_{k=0}^{N}(m^{\circ}_{k}(\overline{\boldsymbol{x}}) - m^{\circ}(\overline{\boldsymbol{x}}))^{3}\,\Pr(k|N,\overline{\boldsymbol{x}}) \tag{B-20}$$

is the third-order centred moment of the proportion of immigrants in a group. When $\sigma^{2}_{\mathrm{m}}(\overline{\boldsymbol{x}}) = \gamma_{\mathrm{m}}(\overline{\boldsymbol{x}}) = 0$, eq. (B-18) reduces to previous expressions for three-way relatedness under the Wright-Fisher model of reproduction and the island model of dispersal (e.g. eq. 12b in Ohtsuki, 2010). Eq. (B-18) shows that negative $\gamma_{\mathrm{m}}(\overline{\boldsymbol{x}})$ increases three-way relatedness. This is because negative skew in the proportion of immigrants means that extremely high-immigration groups are rare, while low-immigration groups are relatively common. Groups with very high immigration strongly reduce the probability that three sampled individuals are all philopatric, because this probability scales with $(1-m^{\circ}_{k}(\overline{\boldsymbol{x}}))^{3}$. When the right tail of the immigration distribution is reduced (negative skew), these patches occur less often, increasing the average probability that three sampled individuals are all philopatric and therefore increasing $r^{\circ}_{3}(\overline{\boldsymbol{x}})$.

B.2.2 Effect of traits on pairwise relatedness

We derive $\partial r_{2}(\boldsymbol{y},\overline{\boldsymbol{x}})/\partial y_{a}$ in Appendix F. We show there that it can be computed as

$$\frac{\partial r_{2}(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_{a}} = N\tilde{r}^{\circ}_{2}(\overline{\boldsymbol{x}})\left[\tilde{r}^{\circ}_{2}(\overline{\boldsymbol{x}})\frac{\partial w^{(2)}_{\mathrm{p}}(\boldsymbol{x}_{i},\boldsymbol{x}_{-i},\overline{\boldsymbol{x}})}{\partial x_{i,a}} + (N-1)\tilde{r}^{\circ}_{3}(\overline{\boldsymbol{x}})\frac{\partial w^{(2)}_{\mathrm{p}}(\boldsymbol{x}_{i},\boldsymbol{x}_{-i},\overline{\boldsymbol{x}})}{\partial x_{j,a}}\right] + \mathcal{O}(\delta), \tag{B-21}$$

where:

$$w^{(2)}_{\mathrm{p}}(\boldsymbol{x}_{i},\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}) = \sum_{\boldsymbol{c}\in\mathscr{C}} w_{\mathrm{p}}(\boldsymbol{x}_{i},\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})^{2}\Pr(\boldsymbol{c}|\boldsymbol{x}), \tag{B-22}$$

with $\Pr(\boldsymbol{c}|\boldsymbol{x})$ given by eq. (B-8) and $w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})$ by eq. (B-2);

$$\tilde{r}_2^{\circ}(\overline{\boldsymbol{x}}) = \frac{1}{N} + \frac{N-1}{N} r_2^{\circ}(\overline{\boldsymbol{x}}), \tag{B-23}$$

is the probability that two individuals sampled with replacement from a group are IBD under neutrality;

$$\tilde{r}_3^{\circ}(\overline{\boldsymbol{x}}) = \frac{2}{N} r_2^{\circ}(\overline{\boldsymbol{x}}) + \frac{N-2}{N} r_3^{\circ}(\overline{\boldsymbol{x}}), \tag{B-24}$$

is the probability under neutrality that three individuals, the second sampled with replacement of the first individual and the third without, are all IBD (Ohtsuki, 2010).

C Directional selection

In this section, we derive the results presented in section 2 of the main text, in particular we derive eq. (2) of the main text.

C.1 Invasion of cooperation

C.1.1 Direct effect of cooperation

The direct fitness effect of trait z is

$$\begin{aligned} \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}} \\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} &= \left.\frac{\partial}{\partial z_i}\left(\sum_{\boldsymbol{c}\in\mathscr{C}} w(\boldsymbol{x}_i, \boldsymbol{x}, \overline{\boldsymbol{x}}|\boldsymbol{c}) \Pr(\boldsymbol{c}|\boldsymbol{x})\right)\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}} \\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} \\ &= \sum_{\boldsymbol{c}\in\mathscr{C}} w(\boldsymbol{x}_i, \boldsymbol{x}, \overline{\boldsymbol{x}}|\boldsymbol{c}) \left.\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}} \\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}}, \end{aligned} \tag{C-1}$$

where conditional fitness, $w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})$ (eq. B-2), does not depend on trait z; and $\partial \Pr(\boldsymbol{c}|\boldsymbol{x})/\partial z_i$ is the effect of a change in the trait z_i of the focal on the probability that contribution $\boldsymbol{c}$ is realised. This probability (given by eq. B-8) can be decomposed as

$$\Pr(\boldsymbol{c}|\boldsymbol{x}) = z_i^{c_i}(1-z_i)^{1-c_i} \underbrace{\prod_{\substack{j=1 \\ j\neq i}}^{N} z_j^{c_j}(1-z_j)^{1-c_j}}_{\Pr(\boldsymbol{c}_{-i}|\boldsymbol{x}_{-i})}, \tag{C-2}$$

such that

$$\left.\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} = \begin{cases} \Pr^{\circ}(\boldsymbol{c}_{-i}|\overline{\boldsymbol{x}}) & \text{when } c_i = 1 \\ -\Pr^{\circ}(\boldsymbol{c}_{-i}|\overline{\boldsymbol{x}}) & \text{when } c_i = 0. \end{cases} \tag{C-3}$$

where $\Pr^{\circ}(\boldsymbol{c}_{-i}|\overline{\boldsymbol{x}}) = \prod_{j\neq i} \overline{z}^{c_j}(1-\overline{z})^{1-c_j} = \Pr(k|N-1, \overline{\boldsymbol{x}})$ with $k = \sum_{i\neq j} c_i$.

Substituting eq. (C-3) into eq. (C-1) we obtain

$$\left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} = \sum_{k=0}^{N-1} \left(w^{\circ}(c_i = 1, \overline{\boldsymbol{x}}|k+1) - w^{\circ}(c_i = 0, \overline{\boldsymbol{x}}|k) \right) \Pr(k|N-1, \overline{\boldsymbol{x}}), \tag{C-4}$$

where $w^{\circ}(c_i = 1, \overline{\boldsymbol{x}}|k+1)$ is the fitness of a cooperator in a group with $k+1$ cooperators (in a population of average trait $\overline{\boldsymbol{x}}$), and $w^{\circ}(c_i = 0, \overline{\boldsymbol{x}}|k)$ the fitness of a defector in a group with k cooperators (in a population of average trait $\overline{\boldsymbol{x}}$). Hence, the direct effect of cooperation (eq. C-4) is the expected fitness gained by a focal individual from cooperating instead of defecting, marginalised over the probability of finding k cooperators among its neighbours.

As the direct effect of cooperation depends on the distribution of cooperators and defectors across groups, it is complicated to compute and to interpret. But in a population entirely made of defectors, all individuals apart from the focal do not contribute to the common good (i.e. $\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}^{(0)} = (0, \overline{d}_0, \overline{d}_1, \ldots, \overline{d}_N)$), meaning that the probability of there being any cooperator in any group (apart from the focal individual) is 0. The direct effect of cooperation (eq. C-4) simplifies to

$$\begin{aligned} \left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}^{(0)})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(0)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(0)}}} &= w^{\circ}(c_i = 1, \overline{\boldsymbol{x}}^{(0)}|1) - 1 \\ &= w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}^{(0)}|\boldsymbol{e}_i) - 1 \end{aligned} \tag{C-5}$$

where

$$\boldsymbol{e}_i = (0, \ldots, 0, \underbrace{1}_{i-\text{entry}}, 0, \ldots, 0) \tag{C-6}$$

is a vector of length N whose entries are all zeros, except the i^{th} entry, which is one. Hence, the direct effect of the probability of cooperation z is simply the difference between the fitness of an individual who is the sole cooperator in the population, and the fitness of the rest of the individuals in the population (all defectors), which is one.

Substituting conditional fitness (eq. B-2) evaluated at $\overline{\boldsymbol{x}}^{(0)}$ into eq. (C-5), the direct fitness effect of cooperation reads as

$$\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(0)})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(0)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(0)}}} = \left[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-1\right] - C\left[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})/N\right] + B_1/N\left[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})\right] + \mathcal{O}(\delta^2). \tag{C-7}$$

The quantity $w_k^{\circ}(\boldsymbol{x})$ is the (conditional) fitness of an individual in the absence of selection (i.e., when $C = B_k = 0$ for all k), given there are k cooperators in its group in a population monomorphic for $\boldsymbol{x}$; and $\lambda_k^{\circ}(\boldsymbol{x}) = (1-m_k^{\circ}(\overline{\boldsymbol{x}}))^2$ is the probability that two offspring born in a group with k cooperators will compete in that same group after dispersal in a population monomorphic for $\boldsymbol{x}$.

The first term within square brackets in eq. (C-7) is the marginal fitness effect of socially-mediated dispersal. It captures the fitness effect of dispersing owing to there being one cooperator in the group. This first term reflects that in the presence of socially-mediated dispersal, selection on cooperation depends on how the effect of cooperation on dispersal influences fitness. This knock-on effect is positive when

$$w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-1 = \frac{1-\overline{d}_1}{1-\overline{d}_1+(1-c_{\mathrm{d}})\overline{d}_0} + \frac{\overline{d}_1(1-c_{\mathrm{d}})}{1-c_{\mathrm{d}}\overline{d}_0} - 1 > 0; \tag{C-8}$$

or equivalently when

$$\min\{\overline{d}_0, 1-c_{\mathrm{d}}\overline{d}_0\} < \overline{d}_1 < \max\{\overline{d}_0, 1-c_{\mathrm{d}}\overline{d}_0\}. \tag{C-9}$$

This condition shows that for most of the parameter space (when $\overline{d}_0 < 1-c_{\mathrm{d}}\overline{d}_0$), greater dispersal propensity when there is one cooperator than when there is none ($\overline{d}_0 < \overline{d}_1$) increases fitness relative to one. This is because dispersal typically reduces local competition for breeding spots and thus increases fitness (Hamilton and May, 1977). If, however, the cost to dispersal, c_{d}, and baseline dispersal, $\overline{d}_0$, are high (so that $1-c_{\mathrm{d}}\overline{d}_0 < d_0$), then greater dispersal propensity when there is one cooperator than when there is none ($\overline{d}_0 < \overline{d}_1$) decreases fitness due to high mortality during dispersal.

The second term of eq. (C-7), $-C[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})/N]$, captures the marginal effect for the focal individual of paying a cost to contribute to the common good on its own fitness. By paying this cost and producing fewer offspring, the focal decreases its fitness by an amount $-Cw_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})$, but also increases it by a smaller amount $C\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})/N$. This increase is due to the fact that its offspring reduced local competition, simply by being fewer (see eq. 70 of Van Cleve, 2015, for example). Conversely, the last term of eq. (C-7) reflects that by contributing to the common good, the focal individual increases its fecundity by B_1/N, which increases its fitness by an amount $(B_1/N)w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})$, and decreases it by an amount $(B_1/N)\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})$ due to increased local competition (eq. 70 of Van Cleve, 2015).

C.1.2 Indirect effect of cooperation

The indirect fitness effect of trait z is given by

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} &= \frac{\partial}{\partial z_j}\left(\sum_{\boldsymbol{c}\in\mathscr{C}} w(\boldsymbol{x}_i,\boldsymbol{x},\overline{\boldsymbol{x}}|\boldsymbol{c})\Pr(\boldsymbol{c}|\boldsymbol{x})\right)\Bigg|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} \\ &= \sum_{\boldsymbol{c}\in\mathscr{C}} w(\boldsymbol{x}_i,\boldsymbol{x},\overline{\boldsymbol{x}}|\boldsymbol{c})\left.\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}}\end{aligned} \tag{C-10}$$

where

$$\left.\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} = \begin{cases}\Pr^{\circ}(\boldsymbol{c}_{-j};\overline{\boldsymbol{x}}) & \text{when } c_j=1\\ -\Pr^{\circ}(\boldsymbol{c}_{-j};\overline{\boldsymbol{x}}) & \text{when } c_j=0\end{cases} \tag{C-11}$$

with $\Pr^{\circ}(\boldsymbol{c}_{-j};\overline{\boldsymbol{x}}) = \prod_{k\neq j}\overline{z}^{c_k}(1-\overline{z})^{1-c_k}$ with $k=\sum_{i\neq j} c_i$. Using eq. (C-11) into eq. (C-10) we obtain

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}}} &= (1-\overline{z})\sum_{k=0}^{N-2}\Big(w^{\circ}(c_i=0,\overline{\boldsymbol{x}}^{(0)}|k+1) - w^{\circ}(c_i=0,\overline{\boldsymbol{x}}^{(0)}|k)\Big)\Pr(k|N-2,\overline{\boldsymbol{x}}) \\ &+\overline{z}\sum_{k=0}^{N-2}\Big(w^{\circ}(c_i=1,\overline{\boldsymbol{x}}^{(0)}|k+2) - w^{\circ}(c_i=1,\overline{\boldsymbol{x}}^{(0)}|k+1)\Big)\Pr(k|N-2,\overline{\boldsymbol{x}})\end{aligned} \tag{C-12}$$

such that the fitness effect of a neighbour increasing its probability of cooperation can be written as the average fitness gained by a focal (which itself cooperates with probability $\overline{z}$) from having one more cooperator in its group.

Like the direct fitness effect, the indirect fitness effect also depends on the distribution of cooperators and defectors across groups, but simplifies in a population entirely made of defectors. In such a population, the indirect fitness effect is given by

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(0)})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(0)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(0)}}} &= w^{\circ}(c_i=0,\overline{\boldsymbol{x}}^{(0)}|1)-1 \\ &= w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(0)}|\boldsymbol{e}_j)-1,\end{aligned} \tag{C-13}$$

i.e., by the difference between the fitness of a non-contributing individual in a group with a single contributor to the common good and the fitness of the rest of the individuals in the population (all defectors with fitness one). Substituting eq. (B-2) with $\boldsymbol{c}=\boldsymbol{e}_j$ and $\boldsymbol{x}=\overline{\boldsymbol{x}}^{(0)}$ into eq. (C-13) gives

$$\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(0)})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(0)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(0)}}} = \Big[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-1\Big] + C\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})/N + B_1/N\Big[w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-\lambda_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})\Big]+\mathscr{O}(\delta^2) \tag{C-14}$$

when cooperation has small effects on fecundity. This equation decomposes the indirect fitness effect in three terms. The first squared brackets captures the fitness effect of dispersing owing to there being one cooperator in a group, when cooperation gives no fecundity benefits. The second term of eq. (C-14) shows that by paying a fecundity cost C, a cooperator increases the fitness of a group-neighbour by $C\lambda_1^\circ(\overline{\boldsymbol{x}}^{(0)})/N$, as a smaller offspring number decreases local competition. Meanwhile, the last term of eq. (C-14) captures the indirect fitness effect of the benefit generated by the common good, which depends on the balance between an increase in fecundity and in local competition (according to $B_1/N[w_1^\circ(\overline{\boldsymbol{x}}^{(0)}) - \lambda_1^\circ(\overline{\boldsymbol{x}}^{(0)})]$).

C.1.3 The selection gradient

Substituting eqs. (C-7) and (C-14) into eq. (A-7), we obtain the selection gradient on the probability of contributing to the common good, z, when the population is monomorphic for $\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}^{(0)} = (0, d_0, d_1, \dots, d_N)$. This is

$$s_1(\overline{\boldsymbol{x}}^{(0)}) = N\bar{r}_2^\circ(\overline{\boldsymbol{x}}^{(0)})\left[w_1^\circ(\overline{\boldsymbol{x}}^{(0)}) - 1\right] - C\left[w_1^\circ(\overline{\boldsymbol{x}}^{(0)}) - r_{2;1}(\overline{\boldsymbol{x}}^{(0)})\right] + B_1\left[w_1^\circ(\overline{\boldsymbol{x}}^{(0)})\bar{r}_2^\circ(\overline{\boldsymbol{x}}^{(0)}) - r_{2;1}(\overline{\boldsymbol{x}}^{(0)})\right] + \mathcal{O}(\delta^2), \tag{C-15}$$

where recall : $\bar{r}_2^\circ(\overline{\boldsymbol{x}}) = 1/N + (N-1)r_2^\circ(\overline{\boldsymbol{x}})/N$ is the probability that in the absence of selection and in a population monomorphic for $\overline{\boldsymbol{x}}$, two offspring sampled in a group before dispersal are related (i.e., have a common ancestor that resided in that same group); and $r_{2;k}(\overline{\boldsymbol{x}}) = \lambda_k^\circ(\overline{\boldsymbol{x}})\bar{r}_2^\circ(\overline{\boldsymbol{x}})$ is the (conditional) probability that in the absence of selection and in a population monomorphic for $\boldsymbol{x}$, two related offspring born in a group with k cooperators will compete in that same group after dispersal. This quantity $r_{2;k}(\overline{\boldsymbol{x}})$ can thus be thought of as a measure of the intensity of kin competition, which, as shown in eq. (C-15), discounts the effects of cooperation.

We can express how the balance between relatedness and kin competition influences the evolution of cooperation more succinctly by re-arranging eq. (C-15) as

$$s_1(\overline{\boldsymbol{x}}^{(0)}) = \rho_1(\overline{\boldsymbol{x}}^{(0)})\Big(A_1(\overline{\boldsymbol{x}}^{(0)}) + [B_1/N - C] + (N-1)\kappa_1(\overline{\boldsymbol{x}}^{(0)})B_1/N\Big) + \mathcal{O}(\delta^2), \tag{C-16}$$

where

$$\rho_1(\overline{\boldsymbol{x}}^{(0)}) = w_1^\circ(\overline{\boldsymbol{x}}^{(0)}) - r_{2;1}(\overline{\boldsymbol{x}}^{(0)}) > 0 \tag{C-17}$$

is a positive coefficient that scales the magnitude of the selection gradient;

$$A_1(\overline{\boldsymbol{x}}^{(0)}) = N\bar{r}_2^{\circ}(\overline{\boldsymbol{x}}^{(0)})\frac{w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-1}{w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-r_{2;1}(\overline{\boldsymbol{x}}^{(0)})}, \tag{C-18}$$

is the relative fitness effect of socially-mediated dispersal when $C = B_k = 0$ for all k; and

$$\kappa_1(\overline{\boldsymbol{x}}^{(0)}) = \frac{w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})r_2(\overline{\boldsymbol{x}}^{(0)})-r_{2;1}(\overline{\boldsymbol{x}}^{(0)})}{w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)})-r_{2;1}(\overline{\boldsymbol{x}}^{(0)})}, \tag{C-19}$$

is the "scaled relatedness coefficient" (Lehmann and Rousset, 2010; eq. 74 of Van Cleve, 2015). The first term in the parentheses of eq. (C-16), $A_1(\overline{\boldsymbol{x}}^{(0)})$, captures selection on cooperation solely due to its effect on dispersal. The second term, $[B_1/N-C]$, is the effect of cooperation by the focal individual on its own fecundity. The last term within brackets is the effect of cooperation by group members other than the focal on the fecundity of the focal, weighted by the scaled relatedness coefficient $\kappa_1(\overline{\boldsymbol{x}}^{(0)})$. This coefficient (eq. C-19) is pairwise relatedness ($r_2(\overline{\boldsymbol{x}}^{(0)})$) reduced by the effect of kin competition, $r_{2;1}(\overline{\boldsymbol{x}}^{(0)})$. Scaled relatedness thus summarises how relatedness, which favours the evolution of cooperation, balances against kin competition, which hinders the evolution of cooperation. The selection gradient (eq. C-16) is positive when the sum between brackets is positive, giving us eq. (2) of the main text as required (where $A_1(\overline{\boldsymbol{x}}^{(0)}) = A$ and $\kappa_1(\overline{\boldsymbol{x}}^{(0)}) = \kappa$). The quantities $\rho_1(\overline{\boldsymbol{x}}^{(0)})$, $A_1(\overline{\boldsymbol{x}}^{(0)})$ and $\kappa_1(\overline{\boldsymbol{x}}^{(0)})$ can be expressed in terms of group size (N), dispersal cost (c_d), and dispersal strategies $\overline{d}_0$ and $\overline{d}_1$, by substituting

$$w_1^{\circ}(\overline{\boldsymbol{x}}^{(0)}) = \frac{1-\overline{d}_1}{1-\overline{d}_1+(1-c_\text{d})\overline{d}_0}+\frac{\overline{d}_1(1-c_\text{d})}{1-c_\text{d}\overline{d}_0} \tag{C-20a}$$

$$r_2(\overline{\boldsymbol{x}}^{(0)}) = \frac{(1-m_0)^2}{N-(N-1)(1-m_0)^2}, \quad \text{with} \quad m_0 = \frac{(1-c_\text{d})\overline{d}_0}{1-\overline{d}_0+(1-c_\text{d})\overline{d}_0} \tag{C-20b}$$

$$r_{2;1}(\overline{\boldsymbol{x}}^{(0)}) = (1-m_1)^2\left(\frac{1}{N}+\frac{N-1}{N}r_2(\overline{\boldsymbol{x}}^{(0)})\right), \quad \text{with} \quad m_1 = \frac{(1-c_\text{d})\overline{d}_0}{1-\overline{d}_1+(1-c_\text{d})\overline{d}_0} \tag{C-20c}$$

into eqs. (C-17) – (C-19).

Eq. (C-16) is consistent with previously derived expressions of the selection gradient on cooperation in the absence of socially-mediated dispersal (for homogeneous groups and non-overlapping generations). Indeed, substituting $\overline{d}_1 = \overline{d}_0$ into eq. (C-20) which is in turn substituted into eqs. (C-17) –

(C-19), we obtain

$$\rho_1(\overline{\boldsymbol{x}}^{(0)}) = 1 - r_2(\overline{\boldsymbol{x}}^{(0)}) \tag{C-21}$$

$$A_1(\overline{\boldsymbol{x}}^{(0)}) = 0 \tag{C-22}$$

$$\kappa_1(\overline{\boldsymbol{x}}^{(0)}) = 0 \tag{C-23}$$

in line with eq. (5) of Taylor (1992), for example.

C.2 Invasion of defection

In this section, we determine the conditions under which defection can invade once cooperation has fixed, following the same steps as in the previous section. To do so we compute the selection gradient on z when the population is monomorphic for $\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}^{(1)} = (1, \overline{d}_0, \overline{d}_1, \ldots, \overline{d}_N)$. If this gradient is negative, then defection can invade.

C.2.1 Direct effect of cooperation

The direct effect (eq. C-4) at $\overline{\boldsymbol{x}}^{(1)}$ simplifies to

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}^{(1)})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i = \overline{\boldsymbol{x}}^{(1)} \\ \boldsymbol{x}_{-i} = \overline{\boldsymbol{x}}^{(1)}}} &= 1 - w^{\circ}(c_i = 0, \overline{\boldsymbol{x}}^{(1)} | N-1) \\ &= 1 - w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}^{(1)} | \boldsymbol{e}_i')\end{aligned} \tag{C-24}$$

where

$$\boldsymbol{e}_i' = (1, \ldots, 1, \underbrace{0}_{i-\text{entry}}, 1, \ldots, 1) \tag{C-25}$$

is a vector of length N whose entries are all ones, except the i^{th} entry, which is zero. Substituting conditional fitness (eq. B-2) evaluated at $\overline{\boldsymbol{x}}^{(1)}$ into eq. (C-24), the direct fitness effect becomes

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}^{(1)})}{\partial z_i}\right|_{\substack{\boldsymbol{x}_i = \overline{\boldsymbol{x}}^{(1)} \\ \boldsymbol{x}_{-i} = \overline{\boldsymbol{x}}^{(1)}}} &= \left[1 - w_{N-1}^{\circ}(\overline{\boldsymbol{x}}^{(1)})\right] - C\left[w_{N-1}^{\circ}(\overline{\boldsymbol{x}}^{(1)}) - \lambda_{N-1}^{\circ}(\overline{\boldsymbol{x}}^{(1)})/N\right] \\ &\quad + (B_N - B_{N-1})/N\left[w_{N-1}^{\circ}(\overline{\boldsymbol{x}}^{(1)}) - \lambda_{N-1}^{\circ}(\overline{\boldsymbol{x}}^{(1)})\right] + \mathcal{O}(\delta^2).\end{aligned} \tag{C-26}$$

The first term within square brackets of eq. (C-26) is

$$1 - w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) = 1 - \frac{1-\overline{d}_{N-1}}{1-\overline{d}_{N-1}-(1-c_{\text{d}})\overline{d}_N} + \frac{\overline{d}_{N-1}(1-c_{\text{d}})}{1-c_{\text{d}}\overline{d}_N}, \tag{C-27}$$

which captures the direct fitness effect of z via its effect on dispersal (ignoring the effects of defection on fecundity). The rest of eq. (C-26) captures the direct fitness effect of z via its effect on fecundity.

C.2.2 Indirect effect of cooperation

The indirect fitness effect (eq. C-12) at $\overline{\boldsymbol{x}}^{(1)}$ meanwhile reads as

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(1)})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(1)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(1)}}} &= 1 - w^{\circ}(c_i=1,\overline{\boldsymbol{x}}^{(1)}|N-1)\\ &= 1 - w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(1)}|\boldsymbol{e}'_j).\end{aligned} \tag{C-28}$$

Substituting eq. (B-2) with $\boldsymbol{c}=\boldsymbol{e}'_j$ (eq. C-25) and $\boldsymbol{x}=\overline{\boldsymbol{x}}^{(1)}$ into eq. (C-28) gives

$$\begin{aligned}\left.\frac{\partial w(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}^{(1)})}{\partial z_j}\right|_{\substack{\boldsymbol{x}_i=\overline{\boldsymbol{x}}^{(1)}\\ \boldsymbol{x}_{-i}=\overline{\boldsymbol{x}}^{(1)}}} &= \left[1 - w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)})\right] + C\lambda^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)})\\ &\quad + (B_N - B_{N-1})/N\left[w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) - \lambda^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)})\right] + \mathcal{O}(\delta^2),\end{aligned} \tag{C-29}$$

where the first term captures the indirect fitness effect of z through its effect on dispersal, while the remaining terms capture its effects through fecundity.

C.2.3 The selection gradient

Substituting eqs. (C-26) and (C-29) into eq. (A-7), we obtain the selection gradient on z in a population of full cooperators,

$$s_1(\overline{\boldsymbol{x}}^{(1)}) = \rho_{N-1}(\overline{\boldsymbol{x}}^{(1)})\Big(A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) + [(B_N - B_{N-1})/N - C] + (N-1)\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})\,(B_N - B_{N-1})/N\Big) + \mathcal{O}(\delta^2), \tag{C-30}$$

where

$$\rho_{N-1}(\overline{\boldsymbol{x}}^{(1)}) = w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) - r_{2;N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0 \tag{C-31}$$

$$A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) = N\bar{r}^{\circ}_2(\overline{\boldsymbol{x}}^{(1)})\frac{1-w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)})}{w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) - r_{2;N-1}(\overline{\boldsymbol{x}}^{(1)})}, \tag{C-32}$$

is the relative fitness effect of socially-mediated dispersal when $C = B_k = 0$ for all k; and

$$\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)}) = \frac{w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)})r_2(\overline{\boldsymbol{x}}^{(1)}) - r_{2;N-1}(\overline{\boldsymbol{x}}^{(1)})}{w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) - r_{2;N-1}(\overline{\boldsymbol{x}}^{(1)})}, \tag{C-33}$$

is the scaled relatedness coefficient. The quantities $\rho_{N-1}(\overline{\boldsymbol{x}}^{(1)})$, $A_{N-1}(\overline{\boldsymbol{x}}^{(1)})$ and $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})$ can be expressed as

$$w^{\circ}_{N-1}(\overline{\boldsymbol{x}}^{(1)}) = \frac{1-\overline{d}_{N-1}}{1-\overline{d}_{N-1}+(1-c_{\mathrm{d}})\overline{d}_N} + \frac{\overline{d}_{N-1}(1-c_{\mathrm{d}})}{1-c_{\mathrm{d}}\overline{d}_N} \tag{C-34a}$$

$$r_2(\overline{\boldsymbol{x}}^{(1)}) = \frac{(1-m_N)^2}{N-(N-1)(1-m_N)^2}, \quad \text{with} \quad m_N = \frac{(1-c_{\mathrm{d}})\overline{d}_N}{1-\overline{d}_N+(1-c_{\mathrm{d}})\overline{d}_N} \tag{C-34b}$$

$$r_{2;N-1}(\overline{\boldsymbol{x}}^{(1)}) = (1-m_{N-1})^2\left(\frac{1}{N}+\frac{N-1}{N}r_2(\overline{\boldsymbol{x}}^{(1)})\right), \quad \text{with} \quad m_{N-1} = \frac{(1-c_{\mathrm{d}})\overline{d}_N}{1-\overline{d}_{N-1}+(1-c_{\mathrm{d}})\overline{d}_N} \tag{C-34c}$$

in terms of group size (N), dispersal cost (c_{d}), and dispersal strategies $\overline{d}_{N-1}$ and $\overline{d}_N$.

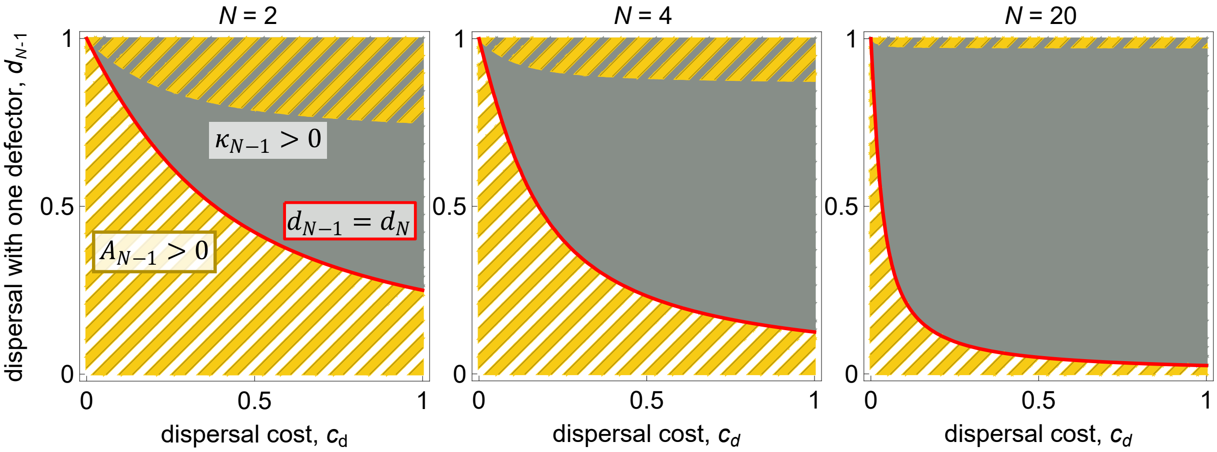

Figure S1: Socially-mediated dispersal strategies for which cooperation has positive indirect effects. Parameter regions of $\overline{d}_{N-1}$ and dispersal cost c_{d} for which $A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$ (yellow-striped) and $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$ (grey region). The regions are drawn from substituting eq. (C-34) into eqs. (C-32) and (C-33), assuming that dispersal in the absence of defectors is at its evolutionarily stable value $\overline{d}_N = d^*$ (eq. 3). The red curve shows $\overline{d}_{N-1} = \overline{d}_N$. The different graphs correspond to different group sizes, $N = 2, 4$ and 20. This shows that for all values of $\overline{d}_{N-1}$, at least one indirect effect of cooperation favours its maintenance (i.e. $A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$ or $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$ always).

Substituting eq. (C-34) into eqs. (C-31)–(C-33) and assuming no socially-mediated dispersal (i.e. $\overline{d}_{N-1} = \overline{d}_N$), the selection gradient in a population of cooperators (eq. C-30) simplifies to being proportional to $(B_N - B_{N-1})/N - C$, as expected. Fig. S1 then shows how socially-mediated dispersal affects selection on z through $A_{N-1}(\overline{\boldsymbol{x}}^{(1)})$ and $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})$ when the population is fixed for full cooperation. The yellow-striped region shows where $A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$, while the grey region shows where $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)}) > 0$.

Fig. S1 shows that socially-mediated dispersal tends to have antagonistic effects on the invasion of defection. When $\overline{d}_{N-1} > \overline{d}_N$, a rare defector induces greater dispersal from its group than would occur in a fully cooperative group. This can favour defection through $A_{N-1}(\overline{\boldsymbol{x}}^{(1)})$, because fewer offspring from that group remain to compete locally. However, the same increase in dispersal makes $\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})$ favour the maintenance of cooperation, because more offspring pay the mortality cost of dispersal. When $\overline{d}_{N-1} < \overline{d}_N$, the effects are reversed. The overall balance is captured by the net effect of socially-mediated dispersal on selection for cooperation: $A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) + (N-1)\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})\,(B_N - B_{N-1})\,/\,N$. When this quantity is negative, socially-mediated dispersal favours the invasion of defection. Fig. S2 shows that this occurs mainly when $\overline{d}_{N-1} > \overline{d}_N$, indicating that the reduction in local competition caused by increased dispersal often outweighs the other effect.

These antagonistic effects on selection for defection contrast with the invasion of cooperation. In that case, $A_1(\overline{\boldsymbol{x}}^{(0)})$ and $\kappa_1(\overline{\boldsymbol{x}}^{(0)})$ tend to have the same sign, so the two effects of socially-mediated dispersal usually reinforce one another (eq. 2, Fig. 2). This is why socially-mediated dispersal favours the invasion of defection over a narrower range of conditions than it favours the invasion of cooperation.

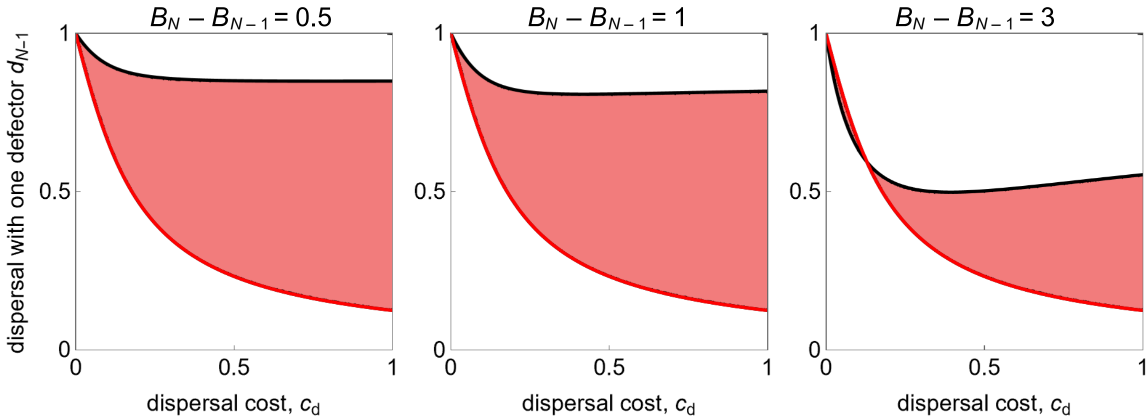

Figure S2: Socially-mediated dispersal strategies that favour the invasion of defection. Parameter regions of dispersal probability $\overline{d}_{N-1}$ and dispersal cost, c_d, for which the indirect effects of cooperation favour defection, i.e. $A_{N-1}(\overline{\boldsymbol{x}}^{(1)}) + (N-1)\kappa_{N-1}(\overline{\boldsymbol{x}}^{(1)})\,(B_N - B_{N-1})\,/\,N < 0$. The regions are drawn assuming that dispersal in the absence of defectors is at its evolutionarily stable value, $\overline{d}_N = d^*$ (eq. 3). The red curve shows $\overline{d}_{N-1} = \overline{d}_N$. The different graphs correspond to different effect sizes of cooperation on fecundity $(B_N - B_{N-1})$. Other parameter: $N = 4$.

D Numerical analyses and results

Here we detail the numerical analyses used to investigate the joint evolution of all traits and derive the results presented in sections 3.2–3.4 of the main text, in particular Figs. 3c-e, 5c, 6b-c, and 7c-d. All analyses were performed in Mathematica (Wolfram Research, Inc., 2016), and we provide the Mathematica notebooks used for these analyses. Our numerical analyses consisted of two main approaches that we describe in the next two sections (D.1 and D.2). We also extend the analyses presented in the main text by exploring additional parameter combinations (section D.3).

D.1 Singular strategies and their stability

We first searched for singular strategies and, when they existed, assessed their convergence and evolutionary stability using the methods described in Appendix A.

D.1.1 Computation of singular strategies

We used different numerical procedures depending on the shape of the returns from cooperation.

For linear returns, with payoff function $B_k = bk$, we searched for interior singular strategies by solving $\boldsymbol{s}(\boldsymbol{x}^*) = 0$ for $\boldsymbol{x}^* = (z^*, d_0^*, \ldots, d_N^*) \in \mathscr{X}$ using `FindRoot[]` in Mathematica. We fixed group size at $N = 4$ and considered $b = 0.75$, 1, and 1.25. For each value of b, we tested 5 000 randomly chosen parameter combinations, with c_d and C independently drawn from a uniform distribution on $[0,1]$. We found no interior singular strategy for any of the 15 000 parameter combinations tested.

For non-linear returns, we used the payoff function $B_k = Nb(k/N)^u$, where $0 < u < 1$ gives saturating returns and $u > 1$ gives accelerating returns. We again searched for interior singular strategies by solving $\boldsymbol{s}(\boldsymbol{x}^*) = 0$ numerically using `FindRoot[]` in Mathematica (see figure legends for parameter values).

Under accelerating returns, this procedure identified interior singular strategies with all trait values strictly between zero and one for all parameter combinations tested.

Under saturating returns, interior singular strategies with all trait values strictly between zero and one were not always found. In these cases, we instead followed the deterministic evolutionary dynamics of the mean trait vector until the population reached an equilibrium at which directional selection had vanished for all traits not at the boundary of trait space, i.e. $s_a(\boldsymbol{x}^*) = 0$ for all traits satisfying $0 < x_a^* < 1$ (see section D.2.1). We initiated these iterations either from $\boldsymbol{x} = (0.1, d^*, d^*, d^*, d^*, d^*)$, where d^* is the evolutionarily stable unconditional dispersal probability given by eq. 3, or from an equilibrium found for a nearby parameter combination. The latter procedure was used only to improve numerical efficiency by starting the search close to the expected equilibrium.

D.1.2 Convergence stability

We assessed the convergence stability of each interior singular strategy found using `FindRoot[]` by computing the eigenvalues of the Jacobian matrix $\mathbf{J}(\boldsymbol{x}^*)$, whose elements are given by eq. (A-6). A singular strategy was classified as convergence unstable when at least one eigenvalue had a positive real part, and convergence stable when all eigenvalues had negative real parts. Non-interior equilibria obtained by iterating the selection gradient were instead classified directly from the evolutionary trajectories, as these equilibria acted as attractors of the deterministic dynamics.

Under accelerating returns, the singular strategies we found were convergence unstable (left-hand side of Fig. S3). Directional selection therefore drives the population away from the singular strategy towards either full cooperation ($\overline{z} = 1$) or no cooperation ($\overline{z} = 0$), depending on the initial trait values.

Under saturating returns, the singular strategies were convergence stable (right-hand side of Fig. S3), so directional selection drives the population towards them.

D.1.3 Evolutionary stability

For convergence stable singular strategies (with all trait values strictly between zero and one), we assessed evolutionary stability by computing the eigenvalues of the Hessian matrix $\mathbf{H}(\boldsymbol{x}^*)$, whose elements are given by eq. (A-10). For each parameter combination, we evaluated this matrix by substituting the expressions for unconditional fitness (eq. (B-7)), pairwise relatedness (eq. (B-13)), three-way relatedness (eq. (B-18)), and the effect of traits on relatedness (eq. (B-21)) into eq. (A-10). Selection was classified as disruptive when at least one eigenvalue of $\mathbf{H}(\boldsymbol{x}^*)$ was positive. In this case, the singular strategy is not evolutionarily stable and phenotypic variation can increase around it.

As mentioned in section D.1.1, some evolutionary attractors of directional selection under saturating returns involved an intermediate value of cooperation but one or more dispersal probabilities at the boundary of trait space, most often $d_k^* = 0$ for some k. In these cases, we assessed disruptive selection only along the unconstrained trait dimensions. Specifically, for any boundary trait indexed by a such that $x_a^* = d_{a-1}^* = 0$ or 1, we removed the corresponding row and column from $\mathbf{H}(\boldsymbol{x}^*)$ and computed the eigenvalues of the reduced matrix. A positive leading eigenvalue of this reduced matrix indicates that selection is disruptive along at least one trait dimension not constrained by the boundary.

D.2 Evolutionary dynamics under directional selection

The second numerical approach was to iterate the recursion for the change in the multidimensional vector of mean trait values, given by eq. (1) of the main text.

D.2.1 Following the selection gradient

To obtain deterministic evolutionary trajectories under directional selection (Figs. 3c and 6b-c of the main text), we iterated

$$\overline{\boldsymbol{x}}_{t+1} = \overline{\boldsymbol{x}}_t + \Delta\overline{\boldsymbol{x}}_t, \tag{D-35}$$

thereby tracking the vector of mean trait values across generations.

For each parameter set and initial phenotype $\overline{\boldsymbol{x}}_0$ specified in the figure legends, we computed $\Delta\overline{\boldsymbol{x}}_t$ from eq. (1). We assumed that traits varied independently, so that the additive genetic variance-covariance matrix $\boldsymbol{G}$ was diagonal, with entries 0.1 on the diagonal and 0 elsewhere. The selection gradient $\boldsymbol{s}(\overline{\boldsymbol{x}}_t)$ was computed by substituting the expressions for unconditional fitness (eq. B-7) and pairwise relatedness (eq. B-13) into eq. A-7. Under these assumptions, each trait x_a changed by $\Delta x_a = 0.1\, s_a(\overline{\boldsymbol{x}}_t)$. After each iteration, trait values were truncated to remain within the admissible trait space: $0 \leq x_a \leq 1$.

We iterated eq. (D-35) until the Euclidean distance between successive mean trait vectors, $\|\overline{\boldsymbol{x}}_{t+1} - \overline{\boldsymbol{x}}_t\|$, was smaller than 10^{-8}. The resulting endpoint was either an interior singular strategy, where directional selection had vanished on all traits, or an evolutionary attractor of directional selection lying on the boundary of trait space, where one or more traits could no longer change in the direction favoured by selection.

D.2.2 Basins of attraction for full cooperation and full defection

To obtain Figs. 3d-e and 6a, we repeated the deterministic procedure described in section D.2.1 across many initial dispersal strategies. For each initial strategy, we recorded whether the trajectory led to full cooperation ($z = 1$) or full defection ($z = 0$). This allowed us to estimate the proportion of sampled initial dispersal strategies that led to the invasion and fixation of cooperation from a population of full defectors, or to the invasion and fixation of defection from a population of full cooperators.

The initial conditions were chosen in the same way across these analyses. When starting from full defection, we set dispersal in the absence of cooperators to the evolutionarily stable unconditional dispersal probability, $\overline{d}_0 = d^*$ (eq. 3). When starting from full cooperation, we instead set dispersal in the absence of defectors to $\overline{d}_N = d^*$. For Fig. 6a, all remaining dispersal probabilities were sampled independently from a uniform distribution on $[0,1]$, and 5000 initial dispersal strategies were tested for each value of c_d.

For Figs. 3d-e, we fixed one additional dispersal probability to examine its effect on the outcome. In Fig. 3d, trajectories started from full defection with $\overline{d}_0 = d^*$ and a specified value of $\overline{d}_1$, while the remaining dispersal probabilities were sampled uniformly on $[0,1]$. In Fig. 3e, trajectories started from full cooperation with $\overline{d}_N = d^*$ and a specified value of $\overline{d}_{N-1}$, while the remaining dispersal probabilities were sampled uniformly on $[0,1]$. For each combination of c_d and the specified dispersal probability, we tested 100 randomly chosen initial dispersal strategies.

D.3 Supplementary results

D.3.1 Singular probability of cooperation

Using the method described in section D.1.1, we examined how the singular probability of cooperation changes with payoff and with group size.

The degree of non-linearity u in common-good payoff affects the singular value of cooperation both under unconditional dispersal and under socially-mediated dispersal (Fig. S3a,d). Under unconditional dispersal, stronger non-linearity favours cooperation in the expected direction: stronger accelerating returns $u > 1$ increase the basin of attraction of full cooperation (compare grey circles and grey triangles in Fig. S3a), whereas stronger saturating returns $u < 1$ lead to higher equilibrium levels of cooperation (compare grey circles and grey triangles in Fig. S3d). Under socially-mediated dispersal, this pattern holds mainly when dispersal costs are low. When dispersal costs are high, cooperation is favoured more strongly when the payoff is closer to linearity (compare orange circles and orange triangles in Fig. S3a,d). This is consistent with the linear-payoff results in section 3.2, where socially-mediated dispersal can strongly promote cooperation.

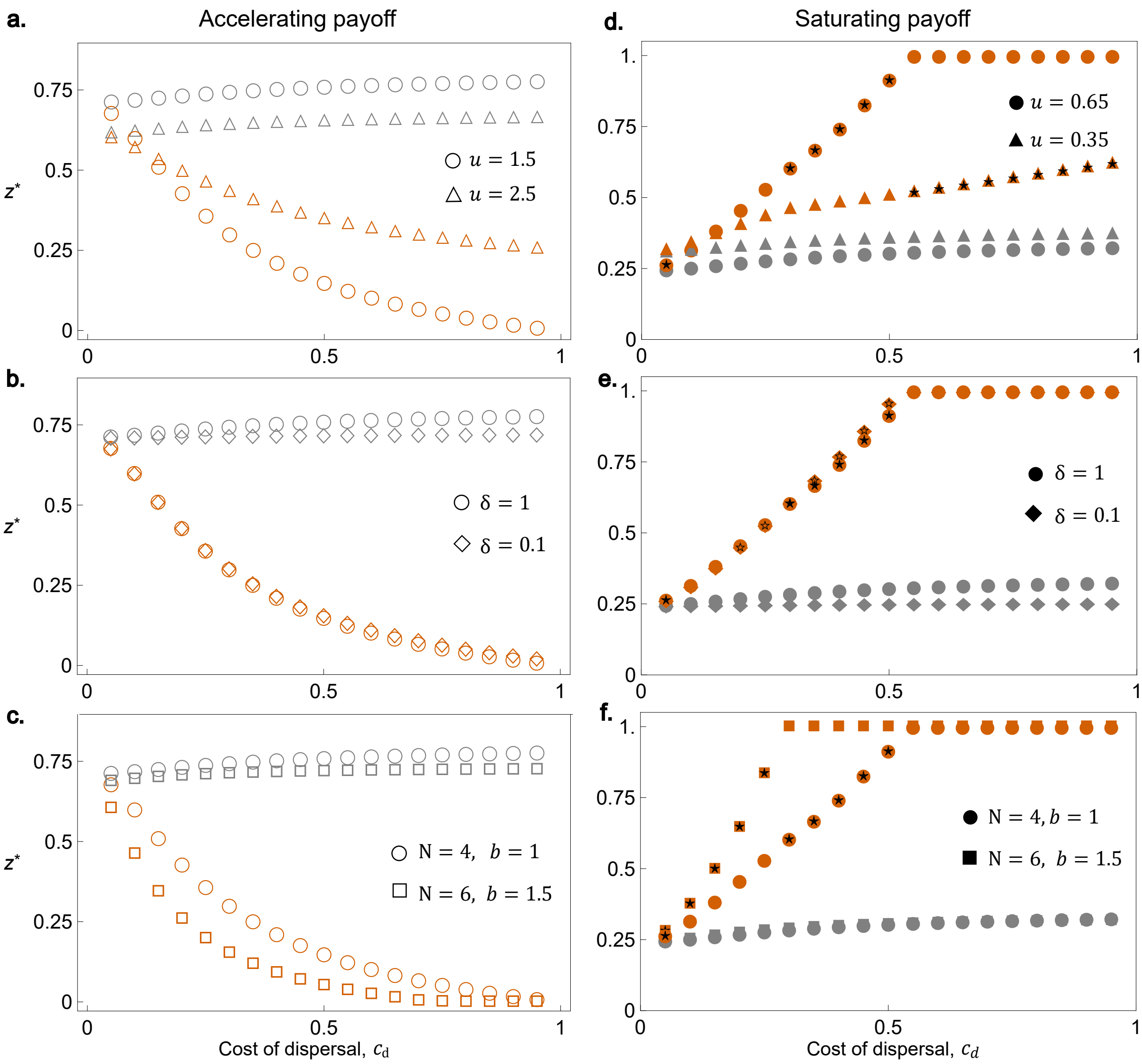

Figure S3: Singular probability of cooperation in non-linear common-good games. Grey markers show singular strategy z^* when cooperation coevolves with unconditional dispersal; orange markers show results when cooperation coevolves with socially-mediated dispersal. Panels a–c correspond to accelerating returns ($u > 1$), whereas panels d–f correspond to saturating returns ($u < 1$). Under accelerating returns, the singular probability of cooperation is convergence unstable and corresponds to the threshold separating trajectories that evolve towards full cooperation or full defection. Under saturating returns, it is convergence stable and corresponds to the equilibrium probability of cooperation. Open markers denote convergence-unstable singular strategies and filled markers convergence-stable singular strategies. Black stars in panels d–f indicate convergence-stable singular strategies at which selection is disruptive. **a,d.** Effect of the non-linearity of the common-good payoff, controlled by u. **b,e.** Effect of the strength of selection on cooperation, controlled by δ. **c,f.** Effect of group size N, holding b/N constant. Unless otherwise indicated, parameter values are $b = 1$, $C = 0.3$, $\delta = 1$, $N = 4$, and $u = 1.5$ for panels a–c and $u = 0.65$ for panels d–f.

We then varied the strength of selection on cooperation by changing δ, which scales the effect of cooperation on fecundity. This had little effect on the singular probability of cooperation when dispersal was socially-mediated (orange markers in Fig. S3b,e). By contrast, under unconditional dispersal, increasing δ affected the singular probability of cooperation in opposite directions for accelerating $u > 1$ and saturating $u < 1$ returns. Under accelerating returns, increasing δ reduced the basin of attraction of full cooperation (compare grey circles and grey diamonds in Fig. S3b). Under saturating returns, it increased the equilibrium level of cooperation (compare grey circles and grey diamonds in Fig. S3e). This pattern is consistent with bet-hedging: individuals may trade lower mean number of offspring for lower variance in offspring number (Starrfelt and Kokko, 2012; Schreiber, 2015). Selection for reduced variance becomes stronger when δ is larger, because larger effects of cooperation on fecundity generate greater differences in reproductive output among groups, and when dispersal is limited, because offspring produced within the same group are more likely to compete locally rather than having these differences averaged across the population (Lehmann and Balloux, 2007). In our model, this effect favours defection under accelerating returns but favours cooperation under saturating returns.

Finally, we examined the effect of group size N on the singular probability of cooperation while holding b/N constant. Larger groups slightly favoured cooperation under both accelerating and saturating returns, but the effect was small (Fig. S3c,f). This is because increasing N while holding b/N constant increases the total payoff available in fully cooperative groups, B_N/N, but leaves the average marginal effect of an individual cooperator on fecundity broadly comparable across group sizes.

D.3.2 Equilibrium probabilities for socially-mediated dispersal

Under diminishing returns, cooperation can converge to an intermediate value ($0 < z^* < 1$). In this case, groups vary in the number of cooperators they contain, so selection can act on the different dispersal probabilities $\overline{d}_0, \ldots, \overline{d}_N$. Here, we examine how several parameters affect the socially-mediated dispersal strategy favoured by directional selection.

We find that the shape of the dispersal reaction norm follows the shape of the common-good payoff. When benefits saturate rapidly ($u \ll 1$), the dispersal response to the number of cooperators also saturates rapidly (Fig. S4a). This is because the marginal gain in fecundity from adding one cooperator determines the marginal increase in kin competition, and therefore the benefit of increasing dispersal. At the evolutionary attractor, the relative number of offspring remaining in their natal group is approximately constant across social environments, while the number of dispersing offspring increases with the number of cooperators (Fig. S5).

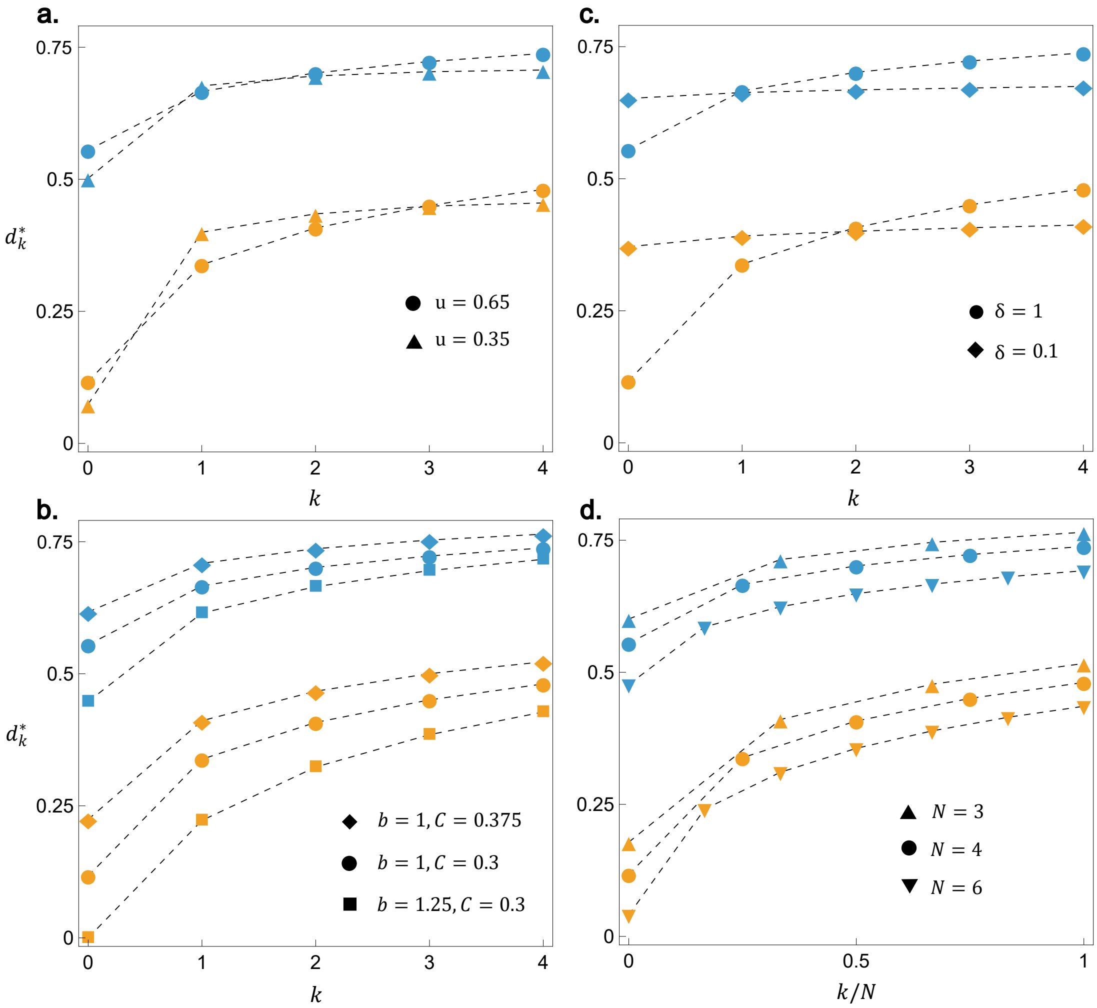

Figure S4: Equilibrium dispersal probabilities under diminishing returns. The dispersal strategy d_k for different k favoured by directional selection is shown as a function of the number of cooperators in the group. In all cases, dispersal increases with the number of cooperators. The steepness of this response tracks the effect of cooperation on fecundity: dispersal increases more strongly when the marginal benefit of an additional cooperator is larger. **a.** Effect of the non-linearity of the common-good payoff, controlled by u. **b.** Effect of the cost C and benefit b of cooperation, varied independently. **c.** Effect of the strength of selection on cooperation, controlled by δ. **d.** Effect of group size, N. Unless otherwise indicated, parameter values are $b = 1$, $N = 4$, $C = 0.3$, and $u = 0.65$. Blue markers: $c_d = 0.1$; orange markers: $c_d = 0.25$.

Increasing the benefit parameter b makes dispersal more sensitive to the number of cooperators (Fig. S4b), because cooperative groups produce more offspring and therefore generate stronger kin competition. Increasing the cost C has the opposite effect: by reducing the fecundity of cooperative groups, it weakens the increase in kin competition associated with cooperation and flattens the dispersal reaction norm. When both costs and benefits are increased together, thereby increasing the overall strength of selection on cooperation, the dispersal response to the number of cooperators becomes steeper (Fig. S4c).

Finally, group size has a small negative effect on dispersal probability (Fig. S4d). Larger groups reduce kin competition, and therefore weaken selection for dispersal. This negative effect is consistent with results for unconditional dispersal (Ajar, 2003).

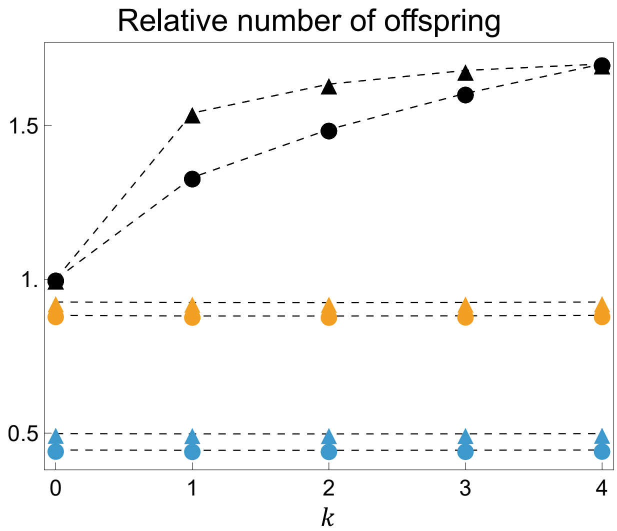

Figure S5: Relative number of offspring and philopatric offspring at the equilibrium dispersal strategy. Black markers show the relative number of offspring produced in groups with different numbers of cooperators. Coloured markers show the corresponding relative number of offspring that remain in their natal group (blue: $c_d = 0.1$; orange: $c_d = 0.25$). Although total offspring production increases with the number of cooperators, the number of philopatric offspring remains constant across social environments. This is consistent with the constant non-disperser principle, whereby conditional dispersal evolves so that increases in local productivity are compensated by greater dispersal. Parameter values correspond to those used in Fig. S4a ($b = 1$, $N = 4$, $C = 0.3$; triangular markers: $u = 0.35$; circular markers: $u = 0.65$).

E Individual-based simulations

E.1 Simulation procedure

Here we describe the individual-based simulations used to obtain Figs. 4b-c and 8 of the main text. Simulations were implemented in Julia. We simulated a finite population subdivided into N_p groups, with $N_p = 125$ for Fig. 4 and $N_p = 4000$ for Fig. 8. Each group $l \in \{1, \dots, N_p\}$ contained $N = 4$ individuals. Each individual i in group l was characterised by a phenotype $x_{i,l} = (z_{i,l}, d_{0,i,l}, \dots, d_{N,i,l})$, where $z_{i,l}$ is its probability of cooperating and $d_{k,i,l}$ is its dispersal probability when k individuals cooperate in its group. All trait values were constrained to lie between 0 and 1.

Each simulation started from a monomorphic population of defectors, with $z = 0$ for all individuals.

All dispersal traits were initially set to the evolutionarily stable unconditional dispersal probability, $d_k = d^* = 2/[1 + 2c_\mathrm{d}N + \sqrt{1 + 4c_\mathrm{d}N(N-1)}]$ for all $k \in \{0, \dots, N\}$ (Ajar, 2003).

Each generation followed the life cycle described in Fig. 1. First, each individual cooperated with probability $z_{i,l}$. We denote the realised action of individual i in group l by $c_{i,l}$, with $c_{i,l} = 1$ for cooperation and $c_{i,l} = 0$ for defection, and write $k = \sum_{j=1}^{N} c_{j,l}$ for the number of cooperators in the group. The fecundity of individual i in group l was then

$$f_l(k, c_i) = 1 + \delta(B_{k,l}/N - Cc_{i,l}), \tag{E-36}$$

where

$$B_{k,l} = bN\left(\frac{1}{N}\sum_{j=1}^{N} c_{j,l}\right)^u \tag{E-37}$$

is the group-level benefit from the common good.

We then formed the next generation by sampling parents for each breeding position in each group (weighted sampling). When filling group l', each individual i from group l was assigned weight

$$\omega_{i,l} = \begin{cases} f_l(k, c_i)(1 - d_{k,i,l}), & \text{if } l = l', \\[2ex] \dfrac{f_l(k, c_i)d_{k,i,l}(1 - c_\mathrm{d})}{N_\mathrm{p} - 1}, & \text{if } l \neq l'. \end{cases} \tag{E-38}$$

These weights account for fecundity, dispersal, dispersal mortality, and the probability that a dispersing offspring reaches the focal group. Parents were sampled in proportion to these weights until all breeding positions in all groups were filled.

Each offspring mutated with probability $\mu = 0.01$. When mutation occurred, each trait was perturbed independently by adding a normally distributed deviate with mean 0 and standard deviation 0.01. Trait values after mutation were truncated to remain within $[0, 1]$.

We ran each simulation for 150 000 generations. We recorded the population mean phenotype every generation and the phenotype of every individual every 1 000 generations.

E.2 Supplementary results

Here, we present supplementary results from the individual-based simulations.

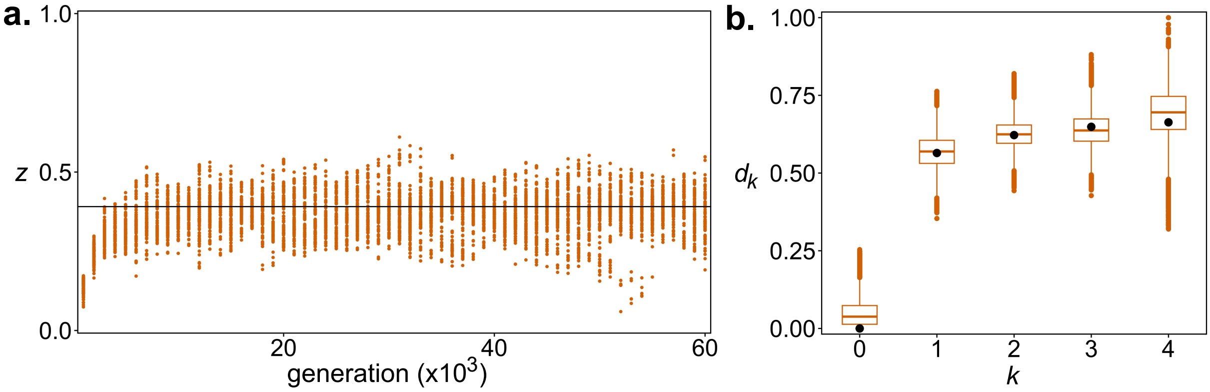

Figure S6: Results from stochastic individual-based simulations under stabilising selection. a. Evolutionary dynamics of the probability of cooperation z. Each dot corresponds to a randomly sampled individual, recorded every 1 000 generations. The population first evolves towards the singular strategy under directional selection. Once the singular probability of cooperation z^* is reached (horizontal line), the population remains centred around this value and no social polymorphism emerges. **b.** Distribution of dispersal probabilities after convergence to the singular strategy. Boxplots show the distribution of trait values in the final generation. Trait values remain centred around the corresponding singular dispersal probabilities (black dots), consistent with stabilising selection maintaining a unimodal phenotypic distribution. Parameters: $N = 4$, $\delta = 3$, $b = 1$, $C = 0.3$, $u = 0.45$.

Fig. S6 illustrates the outcome when selection is stabilising at the singular strategy. The population first evolves towards the singular strategy under directional selection and then remains centred around it. Accordingly, the distribution of trait values remains unimodal, with individuals fluctuating around the singular values calculated analytically for cooperation and dispersal (horizontal line in Fig. S6a; black dots in Fig. S6b).

Fig. S7 provides additional details on the dispersal strategies expressed by cooperators and defectors when disruptive selection leads to the emergence of a social polymorphism. Across all parameter combinations considered, cooperators disperse less on average than defectors. Cooperators also show a stronger dispersal response to the social environment, with dispersal increasing more steeply with the number of cooperators in the group. This difference is particularly pronounced when dispersal isn't too costly. As expected, both morphs evolve higher dispersal probabilities when the cost of dispersal is lower.

F Trait effect on relatedness

Here we derive $\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})/\partial y_a$, the effect of a change in trait a on relatedness for our model. To do so, we use a similar construction as Mullon et al. (2021) (see their Appendix A for details).

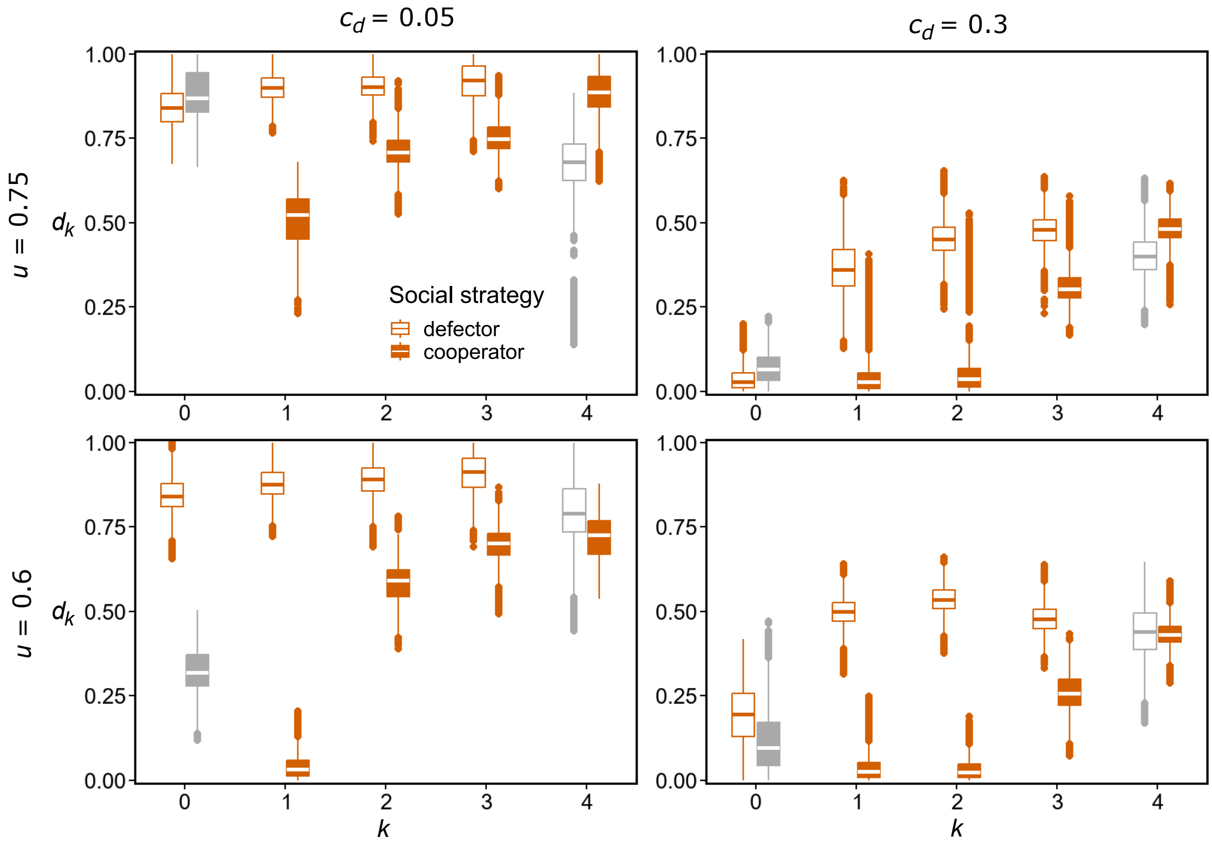

Figure S7: Dispersal strategies expressed by cooperators and defectors under disruptive selection. Boxplots show the distribution of dispersal probabilities in the final generation for cooperators (filled) and defectors (empty), for different values of u and c_{d} (see figure). Across all parameter combinations shown, cooperators disperse less on average than defectors but exhibit a steeper dispersal response to the number of cooperators in the group. Greyed distributions correspond to dispersal traits that are rarely expressed by a given morph and therefore experience very weak selection. Specifically, cooperators are rarely found in groups containing only defectors, whereas defectors are rarely found in groups containing only cooperators, causing the corresponding dispersal probabilities to evolve approximately neutrally.

F.1 Preliminaries

Consider a mutant lineage, initiated by a single individual expressing a phenotype $\boldsymbol{y} = (z_{\mathrm{y}}, d_{0,\mathrm{y}}, \ldots, d_{N,\mathrm{y}})$, while the rest of the population is monomorphic for $\overline{\boldsymbol{x}}$. Consider now a focal mutant group, that is a representative group with at least one mutant. Let us write $M_t \in \{1, ..., N\}$ for the random variable for the number of mutants at generation t in such a mutant group, and $\Pr(M_t)$ be the probability that M_t is realised.

Pairwise relatedness between mutants is the expected probability that one individual randomly sampled among the neighbours of a mutant is also a mutant. This can be expressed as,

$$r_2(\boldsymbol{y},\overline{\boldsymbol{x}}) = \sum_{t=0}^{\infty}\sum_{M_t=1}^{N}\frac{M_t-1}{N-1}\psi(M_t) \tag{F-39}$$

where

$$\psi(M_t) = \frac{[M_t/N]\,\mathrm{Pr}(M_t)}{\underbrace{\sum_{t=0}^{\infty}\sum_{M_t=1}^{N}[M_t/N]\,\mathrm{Pr}(M_t)}_{=n_{\mathrm{M}}}} \tag{F-40}$$

is the probability that an individual sampled randomly from the mutant local lineage lives at generation t (eq. A4 in Mullon et al., 2021). We can condition this probability as

$$\psi(M_t) = \sum_{M_{t-1}=1}^{N}\frac{M_t}{N}\mathrm{Pr}(M_t|M_{t-1})\frac{\mathrm{Pr}(M_{t-1})}{n_{\mathrm{M}}}, \tag{F-41}$$

where $\mathrm{Pr}(M_t|M_{t-1})$ is the probability that there are M_t mutants in the group at generation t given there were M_{t-1} at generation $t-1$. Substituting eq. (F-41) into eq. (F-39), we obtain

$$r_2(\boldsymbol{y},\overline{\boldsymbol{x}}) = \sum_{t=0}^{\infty}\sum_{M_{t-1}=1}^{N}\frac{\mathrm{Pr}(M_{t-1})}{n_{\mathrm{M}}}\sum_{M_t=1}^{N}\frac{M_t-1}{N-1}\frac{M_t}{N}\mathrm{Pr}(M_t|M_{t-1}). \tag{F-42}$$

The innermost sum is the conditional probability that two individuals sampled at generation t are mutants, given M_{t-1}.

To go further and use notation previously introduced, let us write

$$\boldsymbol{x}_{t-1} = \Big(\underbrace{\boldsymbol{y},\ldots,\boldsymbol{y}}_{M_{t-1}},\underbrace{\overline{\boldsymbol{x}},\ldots,\overline{\boldsymbol{x}}}_{N-M_{t-1}}\Big) \tag{F-43}$$

for the vector collecting the phenotypes of individuals in the mutant group at generation $t-1$, and recall that $\boldsymbol{c} = (c_1, c_2, \ldots, c_N)$ (with $c_i \in \{0,1\}$) collects the actions of these individuals (in the same order as $\boldsymbol{x}_{t-1}$). With this notation, we can condition the probability that two individuals sampled at generation t are mutants given M_{t-1} as

$$\sum_{M_t=1}^{N}\frac{M_t-1}{N-1}\frac{M_t}{N}\mathrm{Pr}(M_t|M_{t-1}) = \sum_{\boldsymbol{c}\in\mathscr{C}}\left[\sum_{M_t=1}^{N}\frac{M_t-1}{N-1}\frac{M_t}{N}\mathrm{Pr}(M_t|\boldsymbol{c},M_{t-1})\right]\mathrm{Pr}(\boldsymbol{c}|\boldsymbol{x}_{t-1}), \tag{F-44}$$

where

$$\mathrm{Pr}(\boldsymbol{c}|\boldsymbol{x}_t) = \prod_{j=1}^{N} z_j^{c_j}(1-z_j)^{1-c_j} = \prod_{j=1}^{M_t} z_{\mathrm{y}}^{c_j}(1-z_{\mathrm{y}})^{1-c_j}\prod_{l=M_t+1}^{N}\bar{z}^{c_l}(1-\bar{z})^{1-c_l} \tag{F-45}$$

is the probability of observing $\boldsymbol{c}$ given $\boldsymbol{x}_t$ (eq. B-8).

Now, the innermost sum in eq. (F-44) can be expanded into

$$\sum_{M_t=1}^{N} \frac{M_t-1}{N-1}\frac{M_t}{N}\Pr(M_t|\boldsymbol{c},M_{t-1}) = \frac{1}{N(N-1)}\left(\sum_{M_t=1}^{N} M_t^2 \Pr(M_t|\boldsymbol{c},M_{t-1}) - \sum_{M_t=1}^{N} M_t \Pr(M_t|\boldsymbol{c},M_{t-1})\right), \tag{F-46}$$

revealing that it depends on the first two moments of M_t, given the number of mutants at generation $t-1$ (M_{t-1}) and their realised contribution to the common good ($\boldsymbol{c}$). As the life-cycle considered is Wright-Fisher, the number of mutants at generation t follows a binomial distribution with N trials and a probability of success of $M_{t-1}\, w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})/N$, where

$$w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1}) = \frac{1}{M_{t-1}}\sum_{i=1}^{M_{t-1}} w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c}) \tag{F-47}$$

is the expected number of mutant offspring that successfully settle in their natal group given $\boldsymbol{c}$ and M_{t-1}. This is the average philopatric fitness of a mutant in a group with M_{t-1} mutants and $\boldsymbol{c}$ the contributions to the common good, as defined in eq. (B-2) (first under-braced term) with $\boldsymbol{x}_i = \boldsymbol{y}$ and $\boldsymbol{x}_{-i}$ as a vector made of (M_t-1) entries $\boldsymbol{y}$ and $(N-M_t)$ entries $\overline{\boldsymbol{x}}$.

Using the properties of the binomial distribution, we have

$$\begin{aligned}\sum_{M_t=1}^{N} M_t \Pr(M_t|\boldsymbol{c},M_{t-1}) &= M_{t-1}\, w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})\\ \sum_{M_t=1}^{N} M_t^2 \Pr(M_t|\boldsymbol{c},M_{t-1}) &= M_{t-1}\, w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})\left(1-w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})\frac{M_{t-1}}{N}\right)\\ &\quad + \left(M_{t-1}\, w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})\right)^2.\end{aligned} \tag{F-48}$$

Substituting these into eq. (F-46), we obtain

$$\sum_{M_t=1}^{N} \frac{M_t-1}{N-1}\frac{M_t}{N}\Pr(M_t|\boldsymbol{c},M_{t-1}) = \frac{M_{t-1}^2}{N^2}\, w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})^2. \tag{F-49}$$

In turn, we plug the above into eq. (F-44) to get

$$\begin{aligned}\sum_{M_t=1}^{N} \frac{M_t-1}{N-1}\frac{M_t}{N}\Pr(M_t|M_{t-1}) &= \frac{M_{t-1}^2}{N^2}\underbrace{\sum_{\boldsymbol{c}\in\mathscr{C}} w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_{t-1})^2 \Pr(\boldsymbol{c}|\boldsymbol{x}_{t-1})}_{=w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_{t-1})}\\ &= \left(\frac{1}{N}+\frac{N-1}{N}\frac{M_{t-1}-1}{N-1}\right)\frac{M_{t-1}}{N}\, w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_{t-1}),\end{aligned} \tag{F-50}$$

which is then substituted into eq. (F-42) to obtain

$$\begin{aligned} r_2(\boldsymbol{y},\overline{\boldsymbol{x}}) &= \sum_{t=0}^{\infty}\sum_{M_{t-1}=1}^{N} w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_{t-1})\left(\frac{1}{N}+\frac{N-1}{N}\frac{M_{t-1}-1}{N-1}\right)\psi(M_{t-1}) \\ &= \sum_{t=0}^{\infty}\sum_{M_{t}=1}^{N} w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_{t})\left(\frac{1}{N}+\frac{N-1}{N}\frac{M_{t}-1}{N-1}\right)\psi(M_{t}). \end{aligned} \tag{F-51}$$

F.2 Derivative

F.2.1 Total derivative

To obtain the effect of trait y_a on relatedness, we take the derivative of eq. (F-51) and evaluate it at $\overline{\boldsymbol{x}}$, which using the product rule gives

$$\begin{aligned} \left.\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a}\right|_{\boldsymbol{y}=\overline{\boldsymbol{x}}} &= \langle(1-m_k^{\circ}(\overline{\boldsymbol{x}}))^2\rangle\left[\frac{\partial}{\partial y_a}\left[\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{\psi(M_t)}{N}\right]+\frac{N-1}{N}\underbrace{\frac{\partial}{\partial y_a}\left[\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{M_t-1}{N-1}\psi(M_t)\right]}_{=\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})/\partial y_a}\right] \\ &+\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\left(\frac{\partial}{\partial y_a}w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_t)\right)\frac{M_t}{N}\psi(M_t)^{\circ}+\mathcal{O}(\delta), \end{aligned} \tag{F-52}$$

where $k=\sum_i c_i$ is the number of cooperators (see eq. B-11 for definition of $m_k^{\circ}(\overline{\boldsymbol{x}})$) and all derivatives are evaluated at $\boldsymbol{y}=\overline{\boldsymbol{x}}$ (here and thereafter we drop this notation for readability); and where $\langle(1-m_k^{\circ}(\overline{\boldsymbol{x}}))^2\rangle+\mathcal{O}(\delta)=w_{\mathrm{p}}^{(2)}(\overline{\boldsymbol{x}},\overline{\boldsymbol{x}}|M_t)$ is given by eq. (B-14) and $\psi(M_t)^{\circ}$ denotes $\psi(M_t)$ under neutrality (i.e. in the absence of trait variation and when $\delta=0$).

Now since $\psi(M_t)$ is a probability mass function over t and M_t (eq. F-40), $\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\psi(M_t)=1$ holds, such that

$$\frac{\partial}{\partial y_a}\left[\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{\psi(M_t)}{N}\right]=0, \tag{F-53}$$

which substituted into eq. (F-52) gives us

$$\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a}=\langle(1-m_k^{\circ}(\overline{\boldsymbol{x}}))^2\rangle\frac{N-1}{N}\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a}+\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\left(\frac{\partial}{\partial y_a}w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_t)\right)\frac{M_t}{N}\psi(M_t)^{\circ}+\mathcal{O}(\delta). \tag{F-54}$$

Solving the above for $\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})/y_a$, we obtain

$$\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a}=N\bar{r}_2^{\circ}(\overline{\boldsymbol{x}})\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\left(\frac{\partial}{\partial y_a}w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_t)\right)\frac{M_t}{N}\psi(M_t)^{\circ}, \tag{F-55}$$

where

$$\bar{r}_2^{\circ}(\overline{\boldsymbol{x}}) = \frac{1}{N} + \frac{N-1}{N} r_2^{\circ}(\overline{\boldsymbol{x}}) = \frac{1}{N-(N-1)\langle(1-m_k^{\circ}(\overline{\boldsymbol{x}}))^2\rangle} + \mathcal{O}(\delta) \tag{F-56}$$

is the probability that two individuals sampled with replacement from a group are IBD under neutrality.

F.2.2 Derivative of philopatric fitness effects

Let us now derive $\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_t)/\partial y_a$. We first use the product rule to obtain

$$\begin{aligned}\frac{\partial}{\partial y_a} w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\overline{\boldsymbol{x}}|M_t) &= \frac{\partial}{\partial y_a} \sum_{\boldsymbol{c}\in\mathscr{C}} w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_t)^2 \Pr(\boldsymbol{c}|\boldsymbol{x}) \\ &= 2\sum_{\boldsymbol{c}\in\mathscr{C}} \frac{\partial w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_t)}{\partial y_a}(1-m_k^{\circ}(\overline{\boldsymbol{x}}))\Pr{}^{\circ}(\boldsymbol{c}|\overline{\boldsymbol{x}}) + \sum_{\boldsymbol{c}\in\mathscr{C}}(1-m_k^{\circ}(\overline{\boldsymbol{x}}))\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial y_a} + \mathcal{O}(\delta),\end{aligned} \tag{F-57}$$

where $(1-m_k^{\circ}(\overline{\boldsymbol{x}})) + \mathcal{O}(\delta) = w_{\mathrm{p}}(\overline{\boldsymbol{x}},\overline{\boldsymbol{x}}|\boldsymbol{c},M_t)$. Using the chain rule, we then differentiate $w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_t)$ (given by eq. F-47) to obtain

$$\begin{aligned}\frac{\partial w_{\mathrm{p}}(\boldsymbol{y},\overline{\boldsymbol{x}}|\boldsymbol{c},M_t)}{\partial y_a} &= \frac{1}{M_t}\frac{\partial}{\partial y_a}\left[\sum_{i=1}^{M_t} w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})\right] \\ &= \frac{1}{M_t}\left[\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}}\frac{\partial x_{i,a}}{\partial y_a} + \sum_{\substack{j=1\\j\neq i}}^{N}\left(\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}\frac{\partial x_{j,a}}{\partial y_a}\right) + \sum_{\substack{j=1\\j\neq i}}^{M_t}\left(\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j,\boldsymbol{x}_{-j},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}}\frac{\partial x_{i,a}}{\partial y_a}\right)\right. \\ &\quad\left. + \sum_{\substack{j=1\\j\neq i}}^{M_t}\left(\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j,\boldsymbol{x}_{-j},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}\frac{\partial x_{j,a}}{\partial y_a}\right) + \sum_{\substack{j=1\\j\neq i}}^{M_t}\sum_{\substack{h=1\\h\neq i\\h\neq j}}^{N}\left(\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j,\boldsymbol{x}_{-j},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{h,a}}\frac{\partial x_{h,a}}{\partial y_a}\right)\right].\end{aligned} \tag{F-58}$$

Recall that $\boldsymbol{x}_i = \boldsymbol{y}$, and $\boldsymbol{x}_{-i}$ is a vector made of $(M_t - 1)$ entries $\boldsymbol{y}$ and $(N - M_t)$ entries $\overline{\boldsymbol{x}}$. Thus, $\partial x_{i,a}/\partial y_a = \partial x_{j,a}/\partial y_a = 1$ for any $1 < j \leq M_t$ and $\partial x_{j,a}/\partial y_a = 0$ otherwise. Furthermore, all mutants are equivalent (there is neither age nor class structure in the population), such that the marginal effect of a trait change in a focal individual on its own fitness is

$$\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j,\boldsymbol{x}_{-j},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}} = \frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} \tag{F-59}$$

and the marginal effect of a trait change in a neighbour of a focal individual on the fitness of that focal is

$$\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-j}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}} = \frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} = \frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_j, \boldsymbol{x}_{-j}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{h,a}}. \tag{F-60}$$

Using eq. (F-59) and eq. (F-60), eq. (F-58) readily simplifies to

$$\begin{aligned}\frac{\partial w_{\mathrm{p}}(\boldsymbol{y}, \overline{\boldsymbol{x}}|\boldsymbol{c}, M_t)}{\partial y_a} &= \frac{1}{M_t}\left[\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} + (M_t - 1)\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}} + (M_t - 1)\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}\right.\\ &\quad\left. + (M_t - 1)\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} + (M_t - 1)(M_t - 2)\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}\right]\\ &= \frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} + (M_t - 1)\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}.\end{aligned} \tag{F-61}$$

F.2.3 Derivative of distribution of cooperators

Let us now differentiate $\Pr(\boldsymbol{c}|\boldsymbol{x}_t)$ using similar reasoning. First, we use the chain rule to write

$$\begin{aligned}\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial y_a} &= \sum_{i=1}^{N} \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}} \frac{\partial x_{i,a}}{\partial y_a}\\ &= \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}} + \sum_{\substack{j=1\\ j\neq i}}^{M_t} \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}}.\end{aligned} \tag{F-62}$$

Here, all mutants are not equivalent, as some may have cooperated and not others ($\partial \Pr(\boldsymbol{c}|\boldsymbol{x}_t)/\partial x_{i,a} = \partial \Pr(\boldsymbol{c}|\boldsymbol{x}_t)/\partial x_{j,a}$ if and only if $c_i = c_j$). Substituting eq. (F-62) into the second term of eq. (F-57), we obtain

$$\sum_{\boldsymbol{c}\in\mathscr{C}} (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial y_a} = \sum_{\boldsymbol{c}\in\mathscr{C}} (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}} + \sum_{\substack{j=1\\ j\neq i}}^{M_t} \sum_{\boldsymbol{c}\in\mathscr{C}} (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}}. \tag{F-63}$$

As all mutants cooperate with the same probability z_{y}, the effect of a change in trait a in one mutant summed over all possible contributions to the common good is the same for all mutants, i.e., for $j \neq j'$

$$\sum_{\boldsymbol{c}\in\mathscr{C}} (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}} = \sum_{\boldsymbol{c}\in\mathscr{C}} (1 - m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{a,j'}}, \tag{F-64}$$

meaning that eq. (F-63) simplifies to

$$(1-m_k^\circ(\bar{\boldsymbol{x}}))^2\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial y_a}=\sum_{\boldsymbol{c}\in\mathscr{C}}(1-m_k^\circ(\bar{\boldsymbol{x}}))^2\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}}+(M_t-1)\sum_{\boldsymbol{c}\in\mathscr{C}}(1-m_k^\circ(\bar{\boldsymbol{x}}))^2\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}}. \tag{F-65}$$

F.2.4 Putting it all together

Substituting eq. (F-61) and eq. (F-65) into eq. (F-57), we obtain

$$\begin{aligned}\frac{\partial}{\partial y_a}w_{\mathrm{p}}^{(2)}(\boldsymbol{y},\bar{\boldsymbol{x}}|M_t)=&\,2\sum_{\boldsymbol{c}\in\mathscr{C}}\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}}(1-m_k^\circ(\bar{\boldsymbol{x}}))\Pr{}^\circ(\boldsymbol{c}|\bar{\boldsymbol{x}})+\sum_{\boldsymbol{c}\in\mathscr{C}}(1-m_k^\circ(\bar{\boldsymbol{x}}))^2\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}}\\&+(M_t-1)\left[2\sum_{\boldsymbol{c}\in\mathscr{C}}\frac{\partial w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}}(1-m_k^\circ(\bar{\boldsymbol{x}}))\Pr{}^\circ(\boldsymbol{c}|\bar{\boldsymbol{x}})+\sum_{\boldsymbol{c}\in\mathscr{C}}(1-m_k^\circ(\bar{\boldsymbol{x}}))^2\frac{\partial \Pr(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}}\right]\\&+\mathcal{O}(\delta)\\=&\,\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{i,a}}+(M_t-1)\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{j,a}}+\mathcal{O}(\delta),\end{aligned} \tag{F-66}$$

where

$$w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})=\sum_{\boldsymbol{c}\in\mathscr{C}}\left(w_{\mathrm{p}}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}}|\boldsymbol{c})\right)^2\Pr(\boldsymbol{c}|\boldsymbol{x}). \tag{F-67}$$

We then substitute eq. (F-66) into eq. (F-55) and obtain

$$\begin{aligned}\frac{\partial r_2(\boldsymbol{y},\bar{\boldsymbol{x}})}{\partial y_a}=&\,N\bar{r}_2^\circ(\bar{\boldsymbol{x}})\left[\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{i,a}}\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{M_t}{N}\psi(M_t)^\circ+\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{j,a}}\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}(M_t-1)\frac{M_t}{N}\psi(M_t)^\circ\right]\\&+\mathcal{O}(\delta)\\=&\,N\bar{r}_2^\circ(\bar{\boldsymbol{x}})\left[\bar{r}_2^\circ(\bar{\boldsymbol{x}})\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{i,a}}+(N-1)\bar{r}_3^\circ(\bar{\boldsymbol{x}})\frac{\partial w_{\mathrm{p}}^{(2)}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\bar{\boldsymbol{x}})}{\partial x_{j,a}}\right]+\mathcal{O}(\delta).\end{aligned} \tag{F-68}$$

where

$$\bar{r}_2^\circ(\bar{\boldsymbol{x}})=\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{M_t}{N}\psi(M_t)^\circ \tag{F-69}$$

is the probability under neutrality that two individuals sampled with replacement from the same group are IBD (given by eq. B-23); and

$$\bar{r}_3^\circ(\bar{\boldsymbol{x}})=\sum_{t=0}^{\infty}\sum_{M_t=0}^{N}\frac{M_t}{N}\frac{M_t-1}{N-1}\psi(M_t)^\circ \tag{F-70}$$

is the probability under neutrality that three individuals, the second sampled with replacement of the first individual and the third without, are all IBD. In the Wright-Fisher island model, this is

$$\tilde{r}_3^\circ(\overline{\boldsymbol{x}}) = \frac{2}{N} r_2^\circ(\overline{\boldsymbol{x}}) + \frac{N-2}{N} r_3^\circ(\overline{\boldsymbol{x}}) \tag{F-71}$$

(Ohtsuki, 2010). Eq. (F-68) is eq. (B-21), as required.

F.3 Connection to previous results

We can connect eq. (F-68) to previous results concerning the effects of trait on relatedness when dispersal is unconditional (e.g. Wakano and Lehmann, 2014). Using the product rule, eq. (F-68) gives us

$$\begin{aligned}
\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a} = N\bar{r}_2^\circ(\overline{\boldsymbol{x}}) \Bigg[& \bar{r}_2^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} (1-m_k^\circ(\overline{\boldsymbol{x}})) \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} \mathrm{Pr}^\circ(\boldsymbol{c}|\overline{\boldsymbol{x}}) \\
& + (N-1)\tilde{r}_3^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} (1-m_k^\circ(\overline{\boldsymbol{x}})) \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}} \mathrm{Pr}^\circ(\boldsymbol{c}|\overline{\boldsymbol{x}}) \Bigg] \\
& + N\bar{r}_2^\circ(\overline{\boldsymbol{x}}) \Bigg[\bar{r}_2^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} (1-m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \mathrm{Pr}(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{i,a}} + (N-1)\tilde{r}_3^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} (1-m_k^\circ(\overline{\boldsymbol{x}}))^2 \frac{\partial \mathrm{Pr}(\boldsymbol{c}|\boldsymbol{x})}{\partial x_{j,a}} \Bigg] \\
& + \mathcal{O}(\delta).
\end{aligned} \tag{F-72}$$

Assuming that dispersal is unconditional ($d_0 = \ldots = d_k = \ldots = d_N$), we have $1 - m_k^\circ(\overline{\boldsymbol{x}}) = 1 - m$ for all k, such that eq. (F-72) simplifies to

$$\begin{aligned}
\frac{\partial r_2(\boldsymbol{y},\overline{\boldsymbol{x}})}{\partial y_a} &= N\bar{r}_2^\circ(\overline{\boldsymbol{x}})(1-m) \Bigg[\bar{r}_2^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{i,a}} \mathrm{Pr}^\circ(\boldsymbol{c}|\overline{\boldsymbol{x}}) \\
&\quad + (N-1)\tilde{r}_3^\circ(\overline{\boldsymbol{x}}) \sum_{\boldsymbol{c}\in\mathscr{C}} \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})}{\partial x_{j,a}} \mathrm{Pr}^\circ(\boldsymbol{c}|\overline{\boldsymbol{x}}) \Bigg] + \mathcal{O}(\delta). \\
&= N\bar{r}_2^\circ(\overline{\boldsymbol{x}})(1-m) \Bigg[\bar{r}_2^\circ(\overline{\boldsymbol{x}}) \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}})}{\partial x_{i,a}} + (N-1)\tilde{r}_3^\circ(\overline{\boldsymbol{x}}) \frac{\partial w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}}{\partial x_{j,a}} \Bigg] + \mathcal{O}(\delta).
\end{aligned} \tag{F-73}$$

where $w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}) = \sum_{\boldsymbol{c}\in\mathscr{C}} w_\mathrm{p}(\boldsymbol{x}_i,\boldsymbol{x}_{-i},\overline{\boldsymbol{x}}|\boldsymbol{c})\mathrm{Pr}^\circ(\boldsymbol{c}|\overline{\boldsymbol{x}})$ here. It is straightforward then to see that eq. (F-73) using the definition of $\bar{r}_2^\circ(\overline{\boldsymbol{x}})$ (eq. B-23) and $\tilde{r}_3^\circ(\overline{\boldsymbol{x}})$ (eq. F-70) under unconditional dispersal is equal to eq. (28) in Wakano and Lehmann (2014) or eq. (15) in Mullon et al. (2018).