Mon titre

1 Introduction

In his pioneering work on the evolution of social behavior, Hamilton suggested that altruistic behavior would be associated to limited dispersal (Hamilton, 1964, p. 10). This notion, that tighter links between individuals favor the evolution of altruism, has been shown to hold in a number of population structures (see *e.g.* Ohtsuki et al., 2006; Taylor et al., 2007; Lehmann et al., 2007). The rationale that altruism is favored when altruists interact more with altruists than defectors do (Hamilton, 1975, p. 141; Fletcher & Doebeli, 2009), a condition that is met in viscous populations, *i.e.*, populations with limited dispersal.

Yet, living next to your kin also implies competing against them. The evolution of social traits hence depends on the balance between the positive effects of interactions with related individuals and the detrimental consequences of kin competition. Under specific conditions, the two effects can even compensate each other, thereby annihilating the impact of population viscosity on the evolution of altruism. First identified with computer simulations (Wilson et al., 1992), this cancellation result was analyzed by Taylor (1992) in a model with synchronous generations (Wright-Fisher model) and a subdivided population of constant, infinite size, and was later extended to heterogeneous populations (Rodrigues & Gardner, 2012, with synchronous generations and infinite population size), and other life-cycles and regular population structures (Taylor et al., 2011, with synchronous generations but also with continuous generations and Birth-Death updating). However, small changes in the model's assumptions, such as overlapping generations (Taylor & Irwin, 2000) or the presence of empty sites (Alizon & Taylor, 2008) can tip the balance back in the favor of altruism. This high dependence on life-cycle specificities highlights the difficulty of making general statements about the role of spatial structure on the evolution of altruism.

Another limitation of mechanistic models is the necessity of simplifying assumptions to obtain analytical results. A large number of studies on the evolution of social behavior consider simple population structures (typically, homogeneous populations *sensu* Taylor et al. (2007)) and often also infinite population sizes (but see Allen et al., 2017, for results on any structure); they make use of weak selection approximations, and commonly assume rare or absent mutation. Simple population structures (*e.g.*, regular graphs, or subdivided populations with demes of equal sizes) help reduce the dimensionality of the system under study: this is for instance the case when the structure of the population displays symmetries such that all sites behave the same way in expectation. Weak selection approximations are also crucial for disentangling spatial moments (Lion, 2016), that is, changes in global *vs.* local frequencies. Finally, as

highlighted by Tarnita & Taylor (2014), classical models of inclusive fitness assume infinite population sizes, which maintains diversity; this effect is obtained thanks to (rare) mutation in finite populations. The aim of this study is to explore whether and how imperfect strategy transmission from parents to their offspring affects the impact of population viscosity on the evolution of altruistic behavior in subdivided populations.

When strategy transmission is purely genetic, it makes sense to assume that mutation is relatively weak. A social strategy can however also be culturally transmitted from parent to offspring, in which case "rebellion" (as in Frank's Rebellious Child Model (Frank, 1997)) can be frequent. For simplicity though, we will keep using the word "mutation", keeping in mind that strategy transmission does not have to be genetic.

cite somewhere (Allen et al., 2012; Débarre, 2017)

In this study, we consider three different life-cycles (Wright-Fisher, Moran Birth-Death and Moran Death-Birth), compute for each of them the expected (*i.e.*, long-term) frequency of altruists in the population, and check our findings with numerical simulations. Our results reveal that imperfect strategy transmission from parent to offspring can qualitatively alter the way population viscosity affects the expected frequency of altruists in the population.

2 Model and methods

2.1 Assumptions

We consider a population of size N, subdivided into N_D demes, each hosting exactly n individuals (*i.e.*, containing n sites, each of which is occupied by exactly one individual; we have $nN_D = N$). Each site has a unique label i, $1 \leq i \leq N$. There are two types of individuals in the population, altruists and defectors. Reproduction is asexual. Parents transmit their strategy to their offspring with probability $1-\mu$; this transmission can be genetic or cultural (vertical cultural transmission), but for simplicity, we refer to the parameter μ as a mutation probability. With probability μ, offspring do not inherit their strategy from their parent but instead get one randomly: with probability p, they become altruists, with probability $1-p$ they become defectors. We call the parameter p the mutation bias.

Social interactions take place within each deme; each individual interacts with the $n-1$ other deme members. We assume that social interactions affect individual fecundity, whose baseline is set equal to 1. Each interaction with an altruist increases an individual's fecundity by $\omega \mathsf{b}$; altruists pay a fecundity cost $\omega \mathsf{c}$ ($\mathsf{c} \leq \mathsf{b}$). The parameter ω scales the relative effect of social interactions on fecundity, and is assumed to be small ($\omega \ll 1$).

Denoting by e_{ij} the interaction probability between individuals living at sites i and j, we have

$$e_{ij} = \begin{cases} 0 & \text{if } i = j; \\ \frac{1}{n-1} & \text{if } i \neq j \text{ and both sites are in the same deme;} \\ 0 & \text{if the two sites are in different demes.} \end{cases} \quad (1) \quad \{\texttt{eq:defE}\}$$

attention, maybe rather $1/(n-1)$

Given our assumptions and with this notation, the fecundity of the individual living at site k is given by

$$f_k(\mathbf{X}, \omega) = 1 + \omega \left(\sum_{\ell=1}^{N} e_{\ell k} \mathsf{b} X_\ell - \mathsf{c} X_k \right). \quad (2) \quad \{\texttt{eq:deff}\}$$

Although our assumptions may seem restrictive (fecundity benefits are unconditional, *i.e.*, the same which ever the type of the recipient; the fecundity effects are additive, *i.e.*, the effect of interacting with k altruists is k times the effect of interacting with one altruist), the same fecundities are obtained with a generic fecundity function, after linearization, under the assumption that altruists and defectors are phenotypically close (see APPENDIX for details).

Offspring remain in the parental deme with probability $1-m$; when they do, they land on any site of the deme with equal probability (including the very site of their parent). With probability m, offspring emigrate to a different deme, chosen uniformly at random among the other demes. Denoting by d_{ij} the probability of moving from site i to site j, we have

$$d_{ij} = \begin{cases} d_{\text{in}} = \frac{1-m}{n} & \text{if both sites are in the same deme;} \\ d_{\text{out}} = \frac{m}{(N_D-1)n} & \text{if the two sites are in different demes.} \end{cases} \quad (3) \quad \{\texttt{eq:defD}\}$$

The way the population is updated from one time step to the next depends on the chosen life-cycle (updating rule). We will specifically explore three different life-cycles. At the beginning of each step of each life-cycle, all individuals produce offspring, that can be mutated; then these juveniles move, within the parental deme or outside of it, and land on a site. The next events occurring during the time step depend on the life-cycle:

Moran Birth-Death : One of the newly created juveniles is chosen at random; it kills the adult who was living at the site, and replaces it; all other juveniles die.

Moran Death-Birth : One of the adults is chosen to die (uniformly at random among all adults). It is replaced by one of the juveniles who had landed in its site. All other juveniles die.

Wright-Fisher : All the adults die. At each site of the entire population, one of the juveniles that landed there is chosen and establishes at the site.

2.2 Methods

2.2.1 Analytical part

To derive the expected (*i.e.*, long-term) proportion of altruists in the population, we use the toolbox presented in Débarre (2017), which is valid for any regular population and any life-cycle. Calculation details are given in Appendix A; they go as follows. First, we write an equation for the expected frequency of altruists in the population at time $t+1$, conditional on the composition of the population at time t; we then take the expectation of this quantity, for large times t. After this, we use the assumption that selection is weak ($\omega \ll 1$) and write a first order expansion of the expression that we have obtained. By doing so, we let appear quantities that can be identified as neutral probabilities of identity by descent Q_{ij}, *i.e.*, the probability that individuals living at site i and j share a common ancestor and that no mutation occurred on either lineage since that ancestor, in a model with no selection ($\omega = 0$).

These neutral probabilities of identity by descent depend on the chosen life-cycle, and are also computed by taking the long-term expectation of conditional expectations after one time step (see Appendix A.2).

Check results with regular results

2.2.2 Stochastic simulations

We also run stochastic simulations (coded in `C`). The simulations are run for 10^8 generations (one generation is one time step for the Wright-Fisher life-cycle, and N time steps for the Moran life-cycles). For each set of parameters and life-cycle, we estimate the long-term frequency of altruists by sampling the population every 10^3 generations and computing the average frequency of altruists.

add Code availability

3 Results

3.1 Probabilities of identity by descent

Because of the structure of the population, there are only three types of pairs of individuals, and hence three different values of Q_{ij}:

$$Q_{ij} = \begin{cases} 1 & \text{when } i = j; \\ Q_{\text{in}} & \text{when } i \neq j \text{ and both sites are in the same deme;} \\ Q_{\text{out}} & \text{when sites } i \text{ and } j \text{ are in different demes.} \end{cases} \tag{4}$$

Their values depend on the type of life-cycle that we consider.

3.1.1 Moran updating

Under the Moran life-cycles, probabilities of identity by descent satisfy, for any pair of sites i and $j \neq i$,

$$Q^{\text{M}}_{ij} = \frac{1-\mu}{2} \sum_{k=1}^{N} \left(d_{kj} Q^{\text{M}}_{ki} + d_{ki} Q^{\text{M}}_{kj} \right). \tag{5}$$

Each site is equally likely to have been the latest one that was updated (say it is j); the sum is over the potential parents k, weighted by the dispersal probabilities d_{kj}; the individuals at sites i and j are identical by descent if i and j's parent were (Q^{M}_{ki}) and if no mutation occurred ($1-\mu$). Replacing the dispersal probabilities d_{ij} by their values (eq. (3)), we eventually obtain (see Appendix A.2 for calculation steps): {eq:QM}

$$Q^{\text{M}}_{\text{in}} = \frac{(1-\mu)\left(m + \mu(N_D(1-m)-1)\right)}{(1-\mu)m(N_D\mu(n-1)+1) + (N_D-1)\mu(\mu(n-1)+1)}, \tag{6a}$$

$$Q^{\text{M}}_{\text{out}} = \frac{(1-\mu)m}{(1-\mu)m(N_D\mu(n-1)+1) + (N_D-1)\mu(\mu(n-1)+1)}. \tag{6b}$$

The probability that two different deme-mates are identical by descent, Q^{M}_{in}, monotonically decreases with the emigration probability m, while $Q^{\text{M}}_{\text{out}}$ monotonically increases with m (see figure 1(a)).

When the mutation probability μ is vanishingly small ($\mu \to 0$), both Q^{M}_{in} and $Q^{\text{M}}_{\text{out}}$ are equal to 1: in the absence of mutation indeed, the population ends up fixed for one of the two types, and all individuals are identical by descent. We however obtain a different result if we first assume that the size of the population is infinite ($N_D \to \infty$), because the order of limits matters. For instance, $\lim_{d\to\infty} Q^{M}_{\text{out}} = 0$.

3.1.2 Wright-Fisher updating

Under a Wright-Fisher life-cycle, generations are synchronous, all individuals are replaced at each time step. Probabilities of identity by descent satisfy, for any pair of sites i and $j \neq i$

$$Q_{ij}^{\mathrm{WF}} = (1-\mu)^2 \sum_{k,\ell=1}^{N} d_{ki} d_{\ell j} Q_{kl}^{\mathrm{WF}}. \tag{7}$$

The sum is over all possible parents of i and j, weighted by the dispersal probabilities to sites i and j; the individuals at sites i and j are identical by descent if their parents were ($Q_{k\ell}$) and if neither mutated ($(1-\mu)^2$).

Replacing the dispersal probabilities d_{ij} by their values (eq. (3)) and skipping calculation steps (but see Appendix A.2 for details), we obtain: {eq:QWF}

$$Q_{\mathrm{in}}^{\mathrm{WF}} = \frac{-N_D + M_1 + M_2}{(n-1)N_D + M_1 + M_2}, \tag{8a}$$

$$Q_{\mathrm{out}}^{\mathrm{WF}} = \frac{-\frac{1}{N_D-1} M_1 + M2}{(n-1)N_D + M_1 + M_2}, \tag{8b}$$

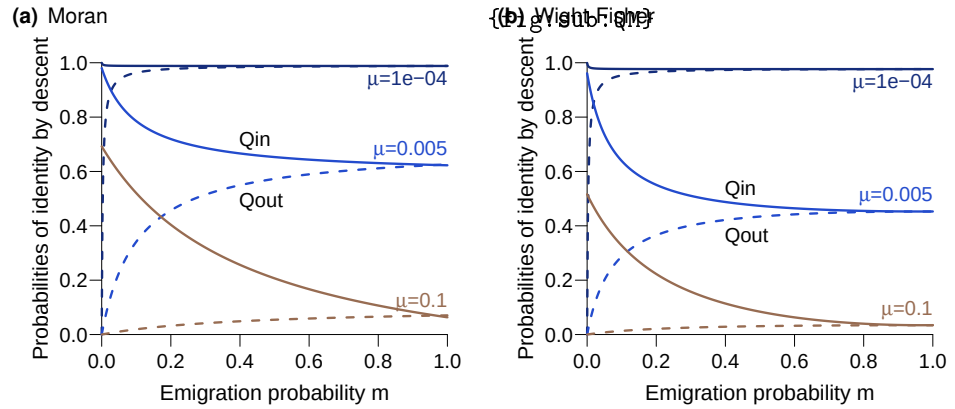
with

$$M_1 = \frac{N_D - 1}{1 - \frac{(1-\mu)^2 (N_D(1-m)-1)^2}{(N_D-1)^2}} \text{ and } M_2 = \frac{1}{1-(1-\mu)^2}.$$

Here, $Q_{\mathrm{in}}^{\mathrm{WF}}$ decreases until $m = m_c^{\mathrm{WF}} = \frac{d-1}{d}$, then increases again, while Q_{out}^{WF} follows the opposite pattern. The threshold value m_c^{WF} corresponds to an emigration probability so high that an individual's offspring is as likely to land in its parent's deme as in any other deme (*i.e.*, $d_{\mathrm{in}} = d_{\mathrm{out}}$).

The two probabilities of identity by descent go to 1 when the mutation probability μ is very small ($\mu \to 0$), except if we first assume that the number of demes is very large ($N_D \to \infty$); for instance, with this life-cycle as well, $\lim_{N_D \to \infty} Q_{\mathrm{out}}^{\mathrm{WF}} = 0$.

Also, because more sites (all of them, actually) are updated at each time step, Q_{in} is lower for the Wright-Fisher updating than for a Moran updating, under which only one site is updated at each time step (compare figure 1(a) and 1(b)).



{fig:sub:M} {fig:sub:WF} {fig:sub:QWF} {fig:Q}

Figure 1: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in}, full curves) and two individuals in different demes (Q_{out}, dashed curves), for different values of the mutation probability μ (10^{-4}, 0.005, 0.1), and for the two types of life-cycles: Moran (a) and Wright-Fisher (b). Other parameters: $n = 4$ individuals per deme, $N_D = 30$ demes.

3.2 Expected frequencies of altruists for each life-cycle

For each of the life-cycles that we consider, the expected frequency of altruists in the population, $\mathbb{E}[\overline{X}]$, can be approximated as

$$\mathbb{E}[\overline{X}] \approx p + \omega \frac{p(1-p)}{\mu} \left[\mathsf{b} \left(\beta_{\mathrm{D}} - \beta_{\mathrm{I}} \right) - \mathsf{c} \left(\gamma_{\mathrm{D}} - \gamma_{\mathrm{I}} \right) \right]. \tag{9}$$ {eq:EXapprox}

(Calculations leading to eq. (9) are presented in the Appendix). The mutation bias p corresponds to the expected proportion of altruists in the population in the absence of selection (*i.e.*, when $\omega = 0$); ω is the parameter that scales the effects of interactions between individuals and is assumed to be small; the subscript $_{\mathrm{D}}$ refers to "direct" effects, and the subscript $_{\mathrm{I}}$ to "indirect" effects. These indirect effects correspond to (kin) competition: by providing a benefit to a deme-mate and thereby increasing its fecundity, a focal altruist indirectly harms others by reducing their relative fecundity (β_{I} term in eq. (9)); by having a reduced fecundity due to the cost of altruism, a focal altruist indirectly favors others by increasing their relative fecundity (γ_{I} term).

We now present the values of these different terms for the three life-cycles under study.

3.2.1 Direct effects

Direct effects are similar for the three life-cycles; the only difference is the value of probabilities of identity by descent Q, that differ between Moran and Wright-Fisher life-cycles, as seen in the previous section: {eq:directeffects}

$$\beta_{\mathrm{D}}^{\mathrm{BD}} = \beta_{\mathrm{D}}^{\mathrm{DB}} = \left(1-\mu\right) Q_{\mathrm{in}}^{\mathrm{M}}, \tag{10a}$$ {eq:bBDD}

$$\beta_{\mathrm{D}}^{\mathrm{WF}} = \left(1-\mu\right) Q_{\mathrm{in}}^{\mathrm{WF}}; \tag{10b}$$ {eq:bWFD}

$$\gamma_{\mathrm{D}}^{\mathrm{BD}} = \gamma_{\mathrm{D}}^{\mathrm{BD}} = \gamma_{\mathrm{D}}^{\mathrm{WF}} = 1-\mu. \tag{10c}$$ {eq:cBDD}

For both benefits and costs, direct effects only count when there is no mutation $(1-\mu)$. Direct effects of benefits (b) only count if the interaction takes place with an individual who is identical by descent; social interactions occur only within demes, hence the presence of Q_{in} in eq. (10a) and eq. (10b). The direct effect of the fecundity cost c however does not depend on the type of interactant.

As seen in the previous section, $Q_{\mathrm{in}}^{\mathrm{M}}$ and $Q_{\mathrm{in}}^{\mathrm{WF}}$ decrease with the emigration probability m (actually only until $m = \frac{d-1}{d}$ for the latter). Consequently, the magnitude of the direct (beneficial) effects of benefits b provided by altruists (β_{D}) decreases, while the direct (costly) effects (γ_{D}) due to the direct cost of altruism c are constant. As a result, if we only consider direct effects, more emi-

gration m is detrimental to the evolution of altruistic behaviour. But there are also indirect effects at play.

3.2.2 Indirect effects

Indirect effects are collateral effects on other individuals; they depend on the type of life-cycle, but always involve individuals who are identical by descent.

Moran Birth-Death Changing the fecundity of a focal individual has two kinds of indirect effects on others: *i*) it changes their probability of being the one chosen to reproduce – this affects all individuals in the population who are identical by descent to the focal, and *ii*) it changes their probability of dying because the number of offspring landing in their site changes – this affects individuals in the population who can send offspring at the same locations as the focal and are identical-by-descent to it. For this life-cycle, the indirect effects are (calculation details are presented in the Appendix)

$$\begin{aligned}\beta_{\mathrm{I}}^{\mathrm{BD}} &= (1-m)\left(\frac{n-1}{n}Q_{\mathrm{in}}^{\mathrm{M}}+\frac{1}{n}\right)+m\,Q_{\mathrm{out}}^{\mathrm{M}}-\mu\frac{1+(n-1)Q_{\mathrm{in}}^{\mathrm{M}}+n(d-1)Q_{\mathrm{out}}^{\mathrm{M}}}{nd} \\ &= \gamma_{\mathrm{D}}^{\mathrm{BD}}.\end{aligned} \tag{11a}$$ {eq:bBDI}

The formulas are the same for the indirect effects associated to b and to c; in other words, the balance between the two indirect effects remains the same when the emigration probability changes. The term $\left(\frac{n-1}{n}Q_{\mathrm{in}}^{\mathrm{M}}+\frac{1}{n}\right)$, which will appear again later, corresponds to the probability that two individuals sampled with replacement from the same deme are identical by descent. Indirect effects are indeed also felt by the focal individual itself (*e.g.*, increasing the fecundity of another individual implies decreasing one's own relative fecundity).

Replacing Q_{in} and Q_{out} by their formula for the Moran life-cycle (eq. (6)), we conclude that both are decreasing functions of the emigration probability m.

3.2.3 Moran Death-Birth

With this life-cycle, death comes first and every individual in the population has the same survival probability $(1/N)$. The indirect consequences of changing a focal individual's fecundity affect all individuals who can send their offspring to the same locations as the focal, and who are identical by descent to it. We obtain

$$\begin{aligned}\beta_{\mathrm{I}}^{\mathrm{DB}} &= (1-\mu)\left[\left(\frac{1}{n}+\frac{(n-1)Q_{\mathrm{in}}^{\mathrm{M}}}{n}\right)\left((1-m)^2+\frac{m^2}{(d-1)}\right)\right.\\ &\qquad\left.+\,m\left(2(1-m)+(d-2)\frac{m}{(d-1)}\right)Q_{\mathrm{out}}^{\mathrm{M}}\right] \qquad (11b)\\ &= \gamma_{\mathrm{I}}^{\mathrm{DB}}\end{aligned}$$

{eq:bDBI}

The first term within the brackets in eq. (11b) corresponds individuals from the same deme whose offspring either does not emigrate, or emigrate to the same deme. The second term corresponds to individuals from different demes who end up in the same deme (either one of their demes, or a third deme).

Here again, $\beta_{\mathrm{I}} = \gamma_{\mathrm{I}}$, so the balance between indirect benefits and indirect costs does not change when the emigration probability m increases.

Replacing Q_{in} and Q_{out} by their formulas given in eq. (6), we can conclude that $\beta_{\mathrm{I}} = \gamma_{\mathrm{I}}$ first decreases with the emigration probability m, and increases again after a threshold value m_c' (given in the appendix; $m_c' < (d-1)/d$).

Mathematica name

3.2.4 Wright-Fisher

With this life-cycle, generations are synchronous and all individuals again all have the same survival probability (now equal to 0). As a result, the formulas for $\beta_{\mathrm{I}}^{\mathrm{WF}}$ and $\gamma_{\mathrm{I}}^{\mathrm{WF}}$ are the same as $\beta_{\mathrm{I}}^{\mathrm{DB}}$ and $\gamma_{\mathrm{I}}^{\mathrm{WF}}$, except that instead of $Q_{\mathrm{in}}^{\mathrm{M}}$ and $Q_{\mathrm{out}}^{\mathrm{M}}$, we need to use $Q_{\mathrm{in}}^{\mathrm{WF}}$ and $Q_{\mathrm{out}}^{\mathrm{WF}}$ (given in eq. (8)). Once this is done, we see that $\beta_{\mathrm{I}}^{\mathrm{WF}} = \gamma_{\mathrm{I}}^{\mathrm{WF}}$ first decreases with the emigration probability m, and increases again after the threshold value $m_c^{\mathrm{WF}} = (d-1)/d$ (which was identified previously as the emigration probability such that offspring have an equal chance of landing in their natal deme or in any other deme, *i.e.*, $d_{\mathrm{in}} = d_{\mathrm{out}}$).

3.3 Identifying threshold values of the mutation probability μ

In the previous section, we investigated the impact of changes in the emigration probability m on each of the terms that make up the expected frequency of altruists $\mathbb{E}[\overline{X}]$. Now we need to combine these different terms to focus on the quantity we are eventually interested in, $\mathbb{E}[\overline{X}]$. The rather lengthy formulas that we obtain are relegated to the appendix, and we concentrate here on the results.

3.3.1 Moran Birth-Death

For this life-cycle, we find that the expected frequency of altruists $\mathbb{E}[\overline{X}]$ is a monotonic function of the emigration probability m; the direction of the change de-

pends on the value of the mutation probability μ compared to a threshold value μ_c^{BD}. When $\mu < \mu_c^{\mathrm{BD}}$, $\mathbb{E}[\overline{X}]$ decreases with m, while when $\mu > \mu_c^{\mathrm{BD}}$, $\mathbb{E}[\overline{X}]$ increases with m; μ_c^{BD} is given by

$$\mu_c^{\mathrm{BD}} = 1 - \frac{\mathsf{b}-\mathsf{c}+\sqrt{(\mathsf{b}-\mathsf{c})\left(4\mathsf{b}(nd)^2+\mathsf{b}-\mathsf{c}\right)}}{2\mathsf{b}nd} \tag{12}$$ {eq:mucBD}

This result is illustrated in figure 2(b).

donner la valeur

3.3.2 Moran Death-Birth

The relationship between $\mathbb{E}[\overline{X}]$ and m is a bit more complicated for this life-cycle. For simplicity, we concentrate on what happens starting from low emigration probabilities (*i.e.*, the sign of the slope of $\mathbb{E}[\overline{X}]$ as a function of m when $m \to 0$). If the benefits b provided by altruists are relatively low ($\mathsf{b} < \mathsf{c}(n+1)$), $\mathbb{E}[\overline{X}]$ initially increases with m provided the mutation probability μ is greater than a threshold value μ_c^{DB} given in eq. (13) below; otherwise, when the benefits are high enough, $\mathbb{E}[\overline{X}]$ initially increases with m for any value of μ. Combining these results, we write

$$\mu_c^{\mathrm{DB}} = \begin{cases} \dfrac{\mathsf{b}-(n+1)\mathsf{c}}{(n-1)\mathsf{c}-(2n-1)\mathsf{b}} & \text{if } \mathsf{b} < \mathsf{c}(n+1), \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$ {eq:mucDB}

The expected frequency of altruists $\mathbb{E}[\overline{X}]$ reaches a maximum for an emigration probability m_c^{DB} (whose complicatedequation is in the appendix), as can be seen in figure 2(a). The limit of this critical emigration probability m_c^{DB} when $\mu \to 0$ is 0: we recover the result that more emigration is detrimental to the evolution of altruism when the mutation probability is either null or vanishingly small.

do

attention order of limits

appendix

3.3.3 Wright-Fisher

The expected frequency of altruists in the population reaches an extremum when $m = m_c^{\mathrm{WF}} = \frac{d-1}{d}$. This extremum is a maximum when the mutation probability is higher than a threshold value μ_c^{WF} given by

$$\mu_c^{\mathrm{WF}} = 1 - \sqrt{1 - \frac{\mathsf{c}}{\mathsf{b}}}, \tag{14}$$

and it is a minimum otherwise (see figure 2(c)).

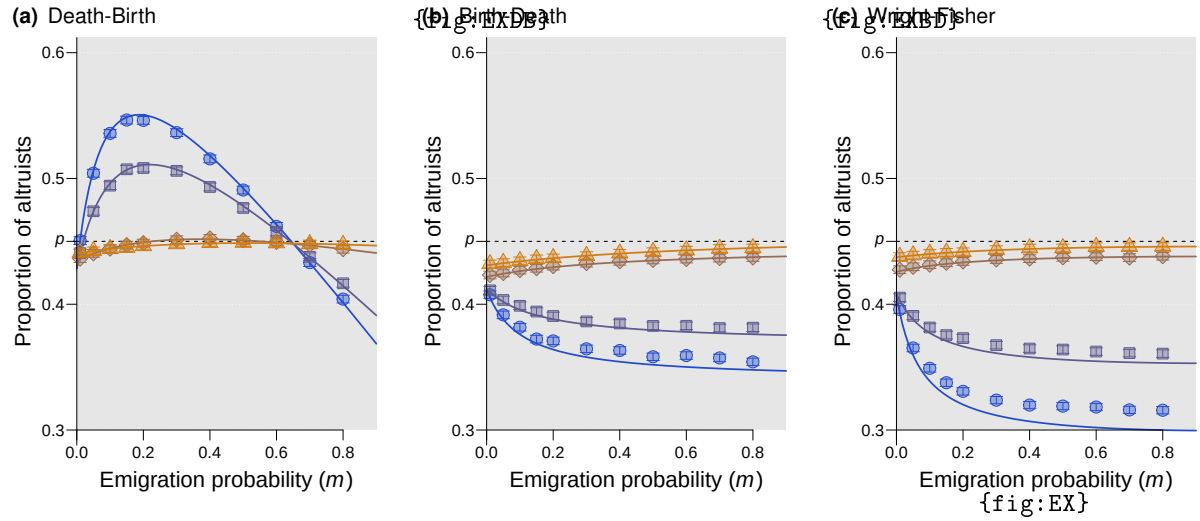


Figure 2: Weak selection. Parameters: $\omega = 0.005$, $b = 15$, $c = 1$, ndemes, size, nreps. NOTE simulations running with 0.005 for mu and with 0.8 for mig.

3.4 Relaxing key assumptions

To derive our analytical results, we had to make a number of simplifying assumptions, such as the fact that selection is weak ($\omega \ll 1$), and the fact that the structure of the population is regular (all demes have the same size n). We explored with numerical simulations the effect of relaxing these key assumptions. The patterns that we identified hold when selection is strong (see figure **??**, done with $\omega = 0.1$), but also when the demes have different sizes. Deme sizes are drawn randomly at the beginning of a simulation; the range from 1 to 5 individuals per deme and the average size is 4 individuals as in the other figures.. Here as well, the same patterns hold as those obtained with a homogeneous structure (figure S2). Addeffect of d_{self}.

le pb c'est ptet simplement que 1 individu ca pose probleme!

todo

4 Discussion

Adding non zero mutation probability altruism increases with emigration.

We used a quantitative measure, $\mathbb{E}[\overline{X}]$, to explore how non-zero mutation probabilities altered the impact of population viscosity. Often, evolutionary success is measured qualitatively, by comparing a quantity (an expected frequency, or, in models with no mutation, a probability of fixation) to the value it would have in the absence of selection; in our case, this amount to saying that altruism is favored whenever $\mathbb{E}[\overline{X}] > p$. Under this condition, population viscosity does not promote the evolution of altruism under the Moran Birth-Death and Wright-Fisher (actually, these two life-cycles cannot ever promote altruistic behavior for any regular population structure (Taylor et al., 2011), whichever the probability of mutation (Débarre, 2017)). However, under a Moran Death-Birth life-cycle, altruism can be only favored at intermediate emigration probabilities.

Go back to the decomposition of the different terms, we see that increase of $\mathbb{E}[\overline{X}]$ with m is driven by the β_{I} term. To simplify the explanations, let us consider that the number of demes is large: in this case, Q_{out} is vanishingly small and so terms involving it can be omitted. Let us also assume that there is no direct cost to being an altruist ($\mathsf{c} = 0$).

Problems of orders of limits, especially when $d \to \infty$ and $\mu \to 0$. Need to specify how small the mutation probability is compared to the size of the population.

Our model bears resemblance to the Rebellious Child Model by Frank (1997), who studied the evolution of a vertically transmitted cultural trait in an asexually reproducing population. In his analysis however, and as acknowledged in the legend of his Figure 7, the model is not fully dynamic because relatedness r is treated as a fixed parameter, which does not depend on mutation. In our mechanistic treatment, r does depend on the mutation probability μ because probabilities of identity by descent do. Mutation was also previously included in models investigating the maintenance of cooperative microorganisms in the presence of cheaters (Brockhurst et al., 2007; Frank, 2010). In both of these models however, only loss-of-function mutation was considered (in our model, this is obtained by setting the mutation biais at $p = 0$). This means that the all-cheaters state is absorbing, and that no matter how favored cooperators may otherwise be, in the long run a finite population will only consist of cheaters.

Voter model Ayana Graphs et dire que on peut avoir e = d. Faire figure.

References

Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. Evolution 62(6):1335–1344.

Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak, M. A. 2017: Evolutionary dynamics on any population structure. Nature 544(7649):227–230.

Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects evolutionary games on graphs. Journal of Theoretical Biology 299:97 – 105. Evolution of Cooperation.

Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at intermediate disturbance. Current Biology 17(9):761–765.

Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution of social behavior in structured populations. Journal of Theoretical Biology 420:26 – 35.

Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evolution of altruism. Proceedings of the Royal Society B: Biological Sciences 276(1654):13–19.

Frank, S. A. 1997: The price equation, fisher's fundamental theorem, kin selection, and causal analysis. Evolution 51(6):1712–1729.

Frank, S. A. 2010: Microbial secretor–cheater dynamics. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365(1552):2515–2522.

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

Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolutionary genetics. Biosocial anthropology 53:133–55.

Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and harming on graphs: the return of the inclusive fitness effect. Journal of Evolutionary Biology 20(6):2284–2295.

Lion, S. 2016: Moment equations in spatial evolutionary ecology. Journal of theoretical biology 405:46–57.

Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule for the evolution of cooperation on graphs and social networks. Nature 441(7092):502–505.

Rodrigues, A. M. M. & Gardner, A. 2012: Evolution of helping and harming in heterogeneous populations. Evolution 66(7):2065–2079.

Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors in finite structured population models. The American Naturalist 184(4):477–488.

Taylor, P. 1992: Altruism in viscous populations—an inclusive fitness model. Evolutionary ecology 6(4):352–356.

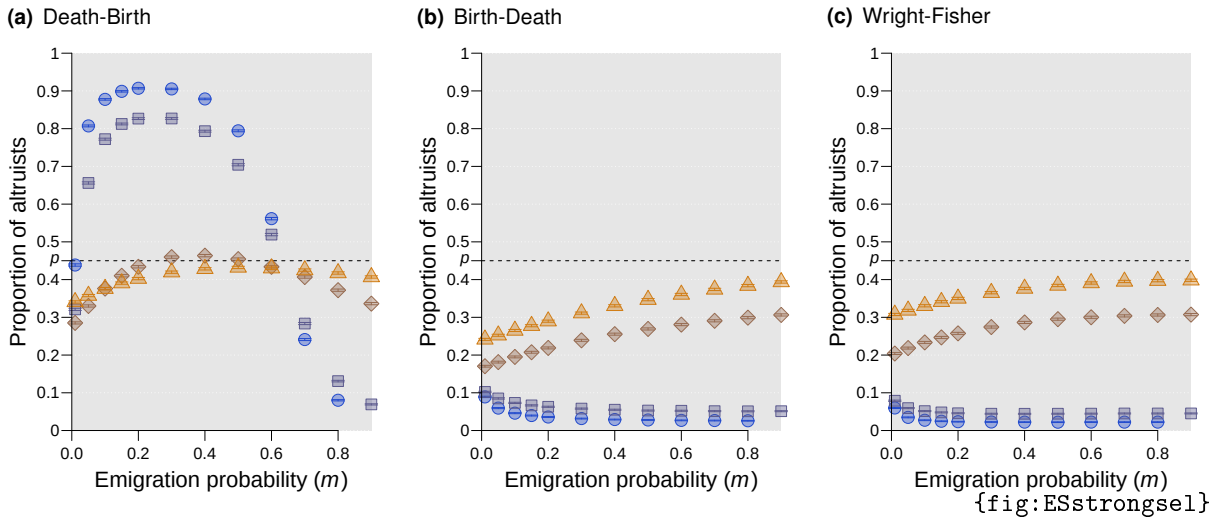
Taylor, P.; Lillicrap, T. & Cownden, D. 2011: Inclusive fitness analysis on mathematical groups. Evolution 65(3):849–859.

Taylor, P. D.; Day, T. & Wild, G. 2007: Evolution of cooperation in a finite homogeneous graph. Nature 447(7143):469–472.

Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic behavior. Evolution 54(4):1135–1141.

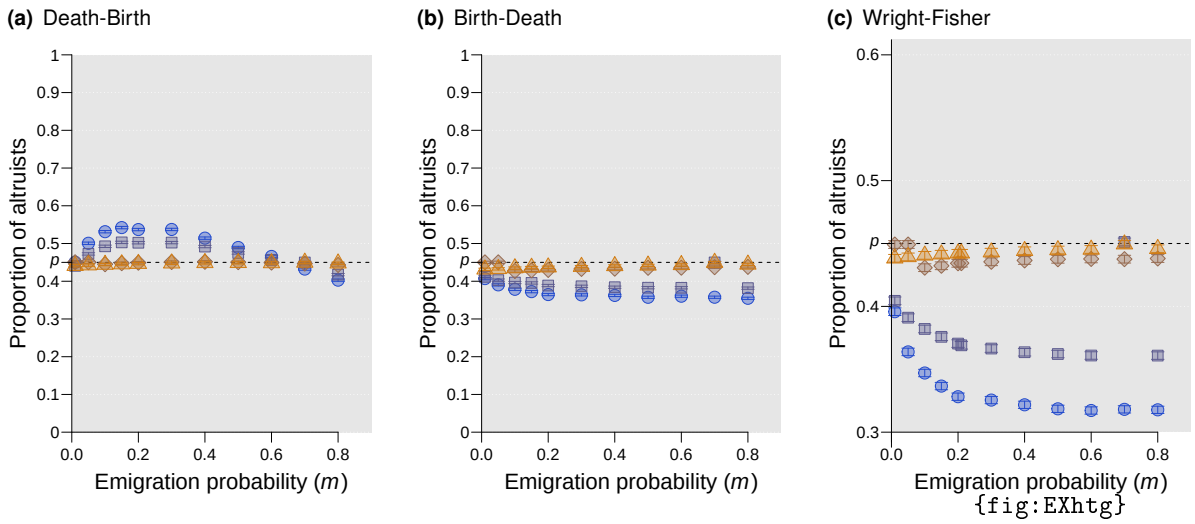
Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely viscous populations? Evolutionary Ecology 6(4):331–341.

Supplementary figures



{fig:ESstrongsel}

Figure S1: Equivalent of figure 2 but with strong selection ($\omega = 0.1$); all other parameters and legend are identical to those of figure 2.



{fig:EXhtg}

Figure S2: Equivalent of figure 2 but with a heterogeneous population structure: deme sizes range form 1 to 5 individuals per deme, the average deme size is 4 as in figure 2; all other parameters and legend are identical to those of figure 2.

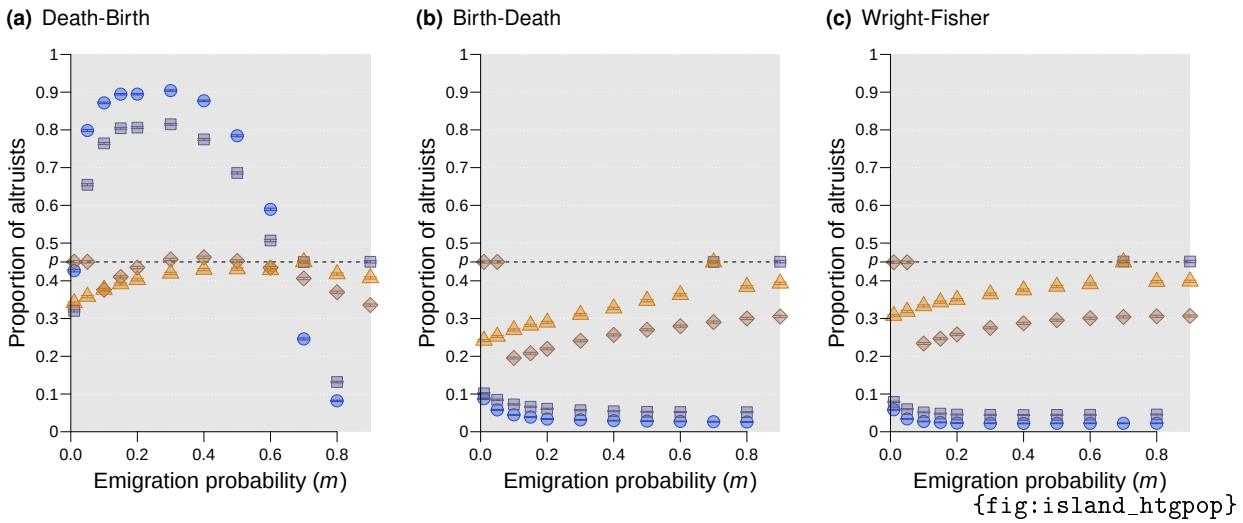


Figure S3: Strong selection, heterogeneous population

A Expected frequency of altruists {sec:app:EX}

Note: The calculation steps are the same as the ones presented in Débarre (2017); they are presented here so that the article is self-contained, but there are no new results in section A. In this section, we work with a generic regular population structure (with symmetries such that all individuals behave the same way in expectation), of which the island model is a particular case.

A.1 For a generic life-cycle {sec:app:generic}

We want to compute the expected proportion of altruists in the population. Some steps can be done without specifying the life-cycle. We represent the state of the population at a given time t using indicator variables $X_i(t)$, $1 \leq i \leq N$, equal to 1 if the individual living at site i at time t is an altruist, and equal to 0 if it is a defector; these indicator variables are gathered in a N-long vector $\mathbf{X}(t)$. The set of all possible population states is $\Omega = \{0,1\}^N$. The proportion of altruists in the population is written $\overline{X}(t) = \sum_{i=1}^{N} X_i(t)$. We denote by $B_{ji}(X(t),\omega)$, written B_{ji} for simplicity, the probability that the individual at site j at time $t+1$ is the newly established offspring of the individual living at site i at time t. We denote by $D_i(X(t),\omega)$ (D_i for simplicity) the probability that the individual living at site i at time t has been replaced (*i.e.*, died) at time $t+1$. Both quantities depend on the chosen life-cycle and on the state of the population; they are given in table S1 for each of the life-cycles that we consider.

Life-cycle	B_{ij}	D_i
Moran Birth-Death	$d_{ji}\dfrac{f_j}{\sum_{k=1}^{N} f_k}$	$\dfrac{\sum_{j=1}^{N} d_{ji} f_j}{\sum_{k=1}^{N} f_k}$
Moran Death-Birth	$\dfrac{1}{N}\dfrac{d_{ji} f_j}{\sum_{k=1}^{N} d_{ki} f_k}$	$\dfrac{1}{N}$
Wright-Fisher	$\dfrac{d_{ji} f_j}{\sum_{k=1}^{N} d_{ki} f_k}$	1

{tab:BD}

Table S1: Formulas of B_{ij} and D_i for each of the life-cycle that we consider; f_i (shorthand notation for $f_i(X,\omega)$) is the fecundity of the individual living at site i, as defined in eq. (2).

Since a dead individual is immediately replaced by one new individual,

$$D_i = \sum_{j=1}^{N} B_{ij} \tag{A.1a}$$ {eq:DBequiv}

holds for all sites i. The structure of the population is also such that in the absence of selection ($\omega = 0$, so that $f_i = 1$ for all sites $1 \leq i \leq N$), all individuals have the same probability of dying and the same probability of having successful offspring (*i.e.*, offspring that become adults), so that

$$D_i^0 = \sum_{j=1}^{N} B_{ji}^0 = B^*, \tag{A.1b}$$ {eq:DBRV}

where the ^{0} subscript means that the quantities are evaluated for $\omega = 0$; this also implies that B_{ij}^0 and D_i^0 do not depend on the state $\mathbf{X}$ of the population. For the Moran life-cycles, $B^* = 1/N$, while for the Wright-Fisher life-cycle, $B^* = 1$. (The difference with eq. (A.1a) is that we are now considering offspring produced by i landing on j).

Given that the population is in state $\mathbf{X}(t)$ at time t, the expected frequency of altruists at time $t+1$ is given by

$$\mathbb{E}\left[\overline{X}(t+1)|\mathbf{X}(t)\right] = \frac{1}{N}\sum_{i=1}^{N}\left[\sum_{j=1}^{N} B_{ij}\left(X_j(1-\mu)+\mu p\right) + (1-D_i)X_i\right]. \tag{A.2a}$$ {eq:conditionalchange}

The first term within the brackets corresponds to births: the type of the individual living at i at time $t+1$ depends on the type of its parent (living at site j), and on whether mutation occurred. The second term in the brackets of eq. (A.2a) corresponds to the survival of the individual living at site i.

Given that there is no absorbing population state (a lost strategy can always be recreated by mutation), there is a stationary distribution of population states, so that the expected frequency of altruists does not change anymore for large times t (realized frequencies of course keep changing). We denote by $\xi(\mathbf{X},\omega,\mu)$ the probability that the population is in state $\mathbf{X}$, given the strength of selection ω and the mutation probability μ. Taking the expectation of eq. (A.2a) ($\mathbb{E}\left[\overline{X}\right] = \sum_{X\in\Omega}\overline{X}\xi(\mathbf{X},\omega,\mu)$), we obtain, after reorganizing:

$$0 = \frac{1}{N}\sum_{X\in\Omega}\sum_{i=1}^{N}\left[\sum_{j=1}^{N} B_{ij}\left(X_j(1-\mu)+\mu p\right) - D_i X_i\right]\xi(\mathbf{X},\omega,\mu). \tag{A.3}$$ {eq:statdist}

Now, we use the assumption of weak selection ($\omega \ll 1$) and consider the first-order expansion of eq. (A.3) for ω close to 0. First, we note that in the absence of

selection ($\omega = 0$), the population is at a mutation-drift balance, and the expected state of every site i is then $\mathbb{E}_0[X_i] = \sum_{X\in\Omega} X_i \xi(X,0,\mu) = p$, the mutation bias. Secondly, we further expand derivatives of B_{ji} and D_i thanks to the chain rule, using the variables f_k ($1 \le k \le N$), corresponding to individual fecundities (also, recall that $f_k = 1$ when $\omega = 0$). Thirdly, we note that for all the life-cycles that we consider, the number of deaths in the population during one time step does not depend on population composition (exactly 1 death for the Moran life-cycles, and exactly N for the Wright-Fisher life-cycle), so that $\partial \sum_{i,j=1}^{N} B_{ij}/\partial\omega$ does not depend on ω. After simplification and reorganization, the first order expansion of eq. (A.3) yields

$$0 = \frac{1}{N}\sum_{i,k=1}^{N}\left[\left.\frac{\partial\left(\sum_{j=1}^{N}(1-\mu)B_{ji} - D_i\right)}{\partial f_k}\right|_{f_k=1} \times\left(\sum_{\ell=1}^{N} e_{\ell k}\mathsf{b}\sum_{X\in\Omega} X_\ell X_i \xi(\mathbf{X},0,\mu) - \mathsf{c}\sum_{X\in\Omega} X_k X_i \xi(\mathbf{X},0,\mu)\right)\right] \quad \text{(A.4)} \quad \texttt{\{eq:weaksel1\}}$$
$$- B^*\mu \left.\frac{\partial\mathbb{E}[\overline{X}]}{\partial\omega}\right|_{\omega=0} + O(\omega^2).$$

The terms $\sum_{X\in\Omega} X_i X_j \xi(\mathbf{X},0,\mu)$, that we will also denote by P_{ij}, correspond to the expected state of the pair of sites $(i,\ j)$, evaluated in the absence of selection ($\omega = 0$). We can also replace these terms by

$$P_{ij} = p^2 + p(1-p)Q_{ij}. \quad \text{(A.5)} \quad \texttt{\{eq:QP\}}$$

In appendix A.2, we will see that recursions on P_{ij} will reveal that Q_{ij} can be interpreted as a probability of identity by descent, *i.e.*, the probability that the individuals at sites i and j have a common ancestor and that no mutation has occurred on either lineage since the ancestor.

Finally, we obtain a first-order approximation of the expected frequency of altruists in the population with

$$\mathbb{E}[\overline{X}] = p + \omega \left.\frac{\partial\mathbb{E}[\overline{X}]}{\partial\omega}\right|_{\omega=0} + O(\omega^2), \quad \text{(A.6)} \quad \texttt{\{eq:EXgeneric\}}$$

where $\left.\frac{\partial\mathbb{E}[\overline{X}]}{\partial\omega}\right|_{\omega=0}$ is obtained from eq. (A.4). We then need to replace the B_{ij} and D_j terms by their formulas for each life-cycle (given in table S1), and the d_{ij} and e_{ij} terms by their formulas (given in eq. (3)) and eq. (1), respectively). For each life-cycle we can group terms as

$$\left.\frac{\partial\mathbb{E}[\overline{X}]}{\partial\omega}\right|_{\omega=0} \approx \frac{p(1-p)}{\mu}\left[\mathsf{b}(\beta_D - \beta_I) - \mathsf{c}(\gamma_D - \gamma_I)\right], \quad \text{(A.7)}$$

where D terms come from the numerators of B_{ij} and D_i, and I terms come from the denominator of B_{ij} and D_i; replacing B_{ij} and D_i by their formulas given in table S1, we obtain

Moran Birth-Death

$$\beta_{\mathrm{D}}^{\mathrm{BD}} = \sum_{k,\ell=1}^{N} \frac{1-\mu}{N} e_{k\ell}\, Q_{\ell k}^{\mathrm{M}}, \tag{A.8a}$$

{eq:EXBDsums}

$$\beta_{\mathrm{I}}^{\mathrm{BD}} = \sum_{j,k,\ell=1}^{N} \left(\frac{d_{\ell j}}{N} - \frac{\mu}{N^2} \right) e_{k\ell}\, Q_{jk}^{\mathrm{M}}, \tag{A.8b}$$

$$\gamma_{\mathrm{D}}^{\mathrm{BD}} = 1-\mu, \tag{A.8c}$$

$$\gamma_{\mathrm{I}}^{\mathrm{BD}} = \sum_{j,k=1}^{N} \left(\frac{d_{kj}}{N} - \frac{\mu}{N^2} \right) Q_{jk}^{\mathrm{M}}. \tag{A.8d}$$

Moran Death-Birth

$$\beta_{\mathrm{D}}^{\mathrm{DB}} = \sum_{k,\ell=1}^{N} \frac{1-\mu}{N} e_{k\ell}\, Q_{\ell k}^{\mathrm{M}}, \tag{A.9a}$$

$$\beta_{\mathrm{I}}^{\mathrm{DB}} = (1-\mu) \sum_{i,j,k,\ell=1}^{N} \frac{d_{ji} d_{\ell i}}{N} e_{k\ell}\, Q_{jk}^{\mathrm{M}}, \tag{A.9b}$$

$$\gamma_{\mathrm{D}}^{\mathrm{DB}} = 1-\mu, \tag{A.9c}$$

$$\gamma_{\mathrm{I}}^{\mathrm{DB}} = (1-\mu) \sum_{i,j,k=1}^{N} \frac{d_{ji} d_{ki}}{N} Q_{jk}^{\mathrm{M}}. \tag{A.9d}$$

Wright-Fisher

$$\beta_{\mathrm{D}}^{\mathrm{WF}} = \sum_{k,\ell=1}^{N} \frac{1-\mu}{N} e_{k\ell}\, Q_{\ell k}^{\mathrm{WF}}, \tag{A.10a}$$

$$\beta_{\mathrm{I}}^{\mathrm{WF}} = (1-\mu) \sum_{i,j,k,\ell=1}^{N} \frac{d_{ji} d_{\ell i}}{N} e_{k\ell}\, Q_{jk}^{\mathrm{WF}}, \tag{A.10b}$$

$$\gamma_{\mathrm{D}}^{\mathrm{WF}} = 1-\mu, \tag{A.10c}$$

$$\gamma_{\mathrm{I}}^{\mathrm{WF}} = (1-\mu) \sum_{i,j,k=1}^{N} \frac{d_{ji} d_{ki}}{N} Q_{jk}^{\mathrm{WF}}, \tag{A.10d}$$

which is the same set of equations as for the Moran Death-Birth model, except for the values of probabilities of identity by descent... that we now need to compute.

A.2 Probabilities of identity by descent {sec:app:IBD}

Here we show the link between the expected state of a pair of sites P_{ij} and probabilities of identity by descent Q_{ij}. In our derivation of $\mathbb{E}[\overline{X}]$, P_{ij} is the quantity that appears, but most studies use Q_{ij}.

A.2.1 Moran model

In a Moran model, exactly one individual died and one individual reproduces during one time step. Given a state $\mathbf{X}$ at time t, at time $t+1$ both sites i and $j \neq i$ are occupied by altruists (or say mutants, since there is no selection and hence no benefits or costs provided by altruists), if *i*) it was the case at time t and neither site was replaced by a non-mutant (first term in eq. (A.11)), or *ii*) if exactly one of the two sites was occupied by a non-mutant at time t, but the site was replaced by a mutant (second and third terms of eq. (A.11)):

$$\begin{aligned}\mathbb{E}\big[X_iX_j(t+1)|X(t)=\mathbf{X}\big] =& X_iX_j\left(1-\sum_{k=1}^{N}\frac{1}{N}\big(d_{ki}+d_{kj}\big)\big((1-X_k)(1-\mu)+\mu(1-p)\big)\right)\\ &+X_i(1-X_j)\sum_{k=1}^{N}\frac{1}{N}d_{kj}\big(X_k(1-\mu)+\mu p\big)\\ &+X_j(1-X_i)\sum_{k=1}^{N}\frac{1}{N}d_{ki}\big(X_k(1-\mu)+\mu p\big).\end{aligned} \tag{A.11}$$ {eq:app:PijM1}

We take the expectation of this quantity, and consider that the stationary distribution is reached ($t\to\infty$); then $\mathbb{E}\big[X_iX_j(t+1)\big]=\mathbb{E}\big[X_iX_j(t)\big]$, and we obtain

$$P_{ij}=\frac{1}{2}\left(\sum_{k=1}^{N}(1-\mu)\big(d_{kj}P_{ki}+d_{ki}P_{kj}\big)\right)+\mu p^2 \qquad (i\neq j), \tag{A.12}$$ {eq:app:PijM}

while $P_{ii}=p$.

Now we substitute $P_{ij}=p^2+p(1-p)Q_{ij}$ in eq. (A.12), obtain

$$Q_{ij}=\frac{1}{2}\sum_{k=1}^{N}(1-\mu)\big(d_{ki}Q_{kj}+d_{kj}Q_{ki}\big), \tag{A.13}$$ {eq:app:QijM}

and realize that Q_{ij} is the probability that the individuals at sites i and $j\neq i$ are identical by descent. To compute it indeed, we need to pick which site was last updated (equal probabilities), then who was the parent (k); the other individual needs to be identical by descent to the parent, and no mutation should have occurred $(1-\mu)$.

A.2.2 Wright-Fisher model

In a Wright-Fisher model, all individuals are replaced at each time step, so we directly consider the state of the parents, so we have:

$$\begin{aligned}\mathbb{E}\big[X_i X_j(t+1)|X(t)=\mathbf{X}\big] = \sum_{k,\ell=1}^{N} d_{ki}d_{\ell j}\Big(& X_k X_\ell (1-\mu+\mu p)^2 \\ & + (X_k(1-X_\ell)+(1-X_k)X_\ell)\,(1-\mu+\mu p)(\mu p) \\ & + (1-X_k)(1-X_\ell)(\mu p)^2\Big) \end{aligned} \tag{A.14}$$ {eq:app:PijWF1}

Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k,\ell=1}^{N} \big(P_{kl}(1-\mu)^2\big) + (2-\mu)\mu p^2. \tag{A.15}$$ {eq:app:PijWF}

Replacing P_{ij} by $p^2 + p(1-p)Q_{ij}$, eq. (A.15) becomes

$$Q_{ij} = \sum_{k,\ell=1}^{N} d_{ki}d_{\ell j}Q_{k\ell}(1-\mu)^2. \tag{A.16}$$ {eq:app:QijWF}

Again, Q_{ij} corresponds to a probability of identity by descent: the individuals at sites i and j are identical by descent if their parents were and if neither mutated ($(1-\mu)^2$).

B In a subdivided population

B.1 β and γ

Now, we need to adapt the results presented in appendix A to our structure of interest, a subdivided population, with dispersal and interaction probabilities given by eq. (3) and eq. (1). For the β and γ terms, we use a brute-force approach, replacing d_{ij} and e_{ij} by their values in a subdivided population, and simplifying the equations (for instance, there are 60 different cases to consider for the four sums that appear in $\beta_{\mathrm{I}}^{\mathrm{DB}}$). The calculations are detailed in an accompanying Mathematica file, and the results are presented in the main text. todo

B.2 Probabilities of identity by descent

For the probabilities of identity by descent, we could also use a brute-force approach, but calculations are faster if we use formulas derived in Débarre (2017) for "two-dimensional population structures". The name comes from the fact that we only need two types of transformations to go from any site to any other site in the population: permutations on the deme index, and permutations on the within-deme index. We introduce notations $\tilde{d}_i$ and $\tilde{Q}_i$, that correspond to the dispersal probability to a site at distance i (*e.g.*, for all $j, 1 \leq j \leq N$, $\tilde{d}_1 = d_{j,j+1}$) and the probability of identity by descent with a site at distance i (*e.g.*, for all $j, 1 \leq j \leq N$, $\tilde{Q}_1 = Q_{j,j+1}$), respectively. Finally, we can rewrite site labels ($1 \leq i \leq N$) as (l_1, l_2), where l_1 is the number of the deme ($1 \leq l_1 \leq N_D$) and l_2 the position of the site within the deme ($1 \leq l_2 \leq n$).

Also, in this section, we distinguish between $d_{\mathrm{self}} = d_{ii}$ and d_{in} (in the main text, $d_{\mathrm{self}} = d_{\mathrm{in}}$).

B.2.1 Moran model

In Débarre (2017), it was shown that

$$\tilde{\mathcal{Q}}_{\substack{r_1 \\ r_2}} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1-(1-\mu)\tilde{\mathcal{D}}_{\substack{q_1 \\ q_2}}} \exp\left(\imath \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(\imath \frac{2\pi q_2 r_2}{N_2}\right) \tag{B.17a}$$ {eq:app:Q2DM}

with

$$\tilde{\mathcal{D}}_{\substack{q_1 \\ q_2}} = \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} \tilde{d}_{\substack{l_1 \\ l_2}} \exp\left(-\imath \frac{2\pi q_1 l_1}{N_1}\right) \exp\left(-\imath \frac{2\pi q_2 l_2}{N_2}\right), \tag{B.17b}$$ {eq:app:D2D}

and λ'_M such that $\tilde{\mathcal{Q}}_{\substack{0\\0}} = 1$. Let us first compute $\tilde{\mathcal{D}}_{\substack{q_1\\q_2}}$ in the case of a subdivided population, with $N_1 = N_D$ and $N_2 = n$:

$$\begin{aligned}\tilde{\mathcal{D}}_{\substack{q_1\\q_2}} &= d_{\text{self}} + \sum_{l_2=1}^{N_2-1} d_{\text{in}} \exp\left(-\imath\frac{2\pi q_2 l_2}{N_2}\right) + \sum_{l_1=1}^{N_1-1}\sum_{l_2=0}^{N_2-1} d_{\text{out}} \exp\left(-\imath\frac{2\pi q_1 l_1}{N_1}\right)\exp\left(-\imath\frac{2\pi q_2 l_2}{N_2}\right)\\ &= d_{\text{self}} + \left(\delta_{q_2}(N_2-1) + (1-\delta_{q_2})(-1)\right) d_{\text{in}} + \left(\delta_{q_1}(N_1-1) + (1-\delta_{q_1})(-1)\right)\left(\delta_{q_2} N_2\right) d_{\text{out}}\\ &= d_{\text{self}} + \left(\delta_{q_2} N_2 - 1\right) d_{\text{in}} + \left(\delta_{q_1} N_1 - 1\right)\delta_{q_2} N_2 d_{\text{out}}. \qquad \text{(B.18a)}\end{aligned}$$

(δ_q is equal to 1 when q is equal to 0 modulo the relevant dimension, and to 0 otherwise). So for the three types of distances that we need to consider (distance 0, distance to another deme-mate, distance to individual in another deme), and with $N_1 = N_D$ and $N_2 = n$, we obtain {eq:app:Dsystem}

$$\tilde{\mathcal{D}}_{\substack{0\\0}} = 1, \qquad \text{(B.19a)}$$

$$\tilde{\mathcal{D}}_{\substack{q_1\\0}} = 1 - m - \frac{m}{d-1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \qquad \text{(B.19b)}$$

$$\tilde{\mathcal{D}}_{\substack{q_1\\q_2}} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \qquad \text{(B.19c)}$$

So for $\tilde{\mathcal{Q}}$, using system (B.19) in eq. (B.17a),

$$\begin{aligned}\tilde{\mathcal{Q}}_{\substack{r_1\\r_2}} &= \frac{\mu\lambda'_M}{N}\left[\frac{1}{1-(1-\mu)\tilde{\mathcal{D}}_{\substack{0\\0}}} + \sum_{q_2=1}^{N_2-1}\frac{1}{1-(1-\mu)\tilde{\mathcal{D}}_{\substack{0\\q_2}}}\exp\left(-\imath\frac{2\pi q_2 r_2}{N_2}\right) + \sum_{q_1=1}^{N_1-1}\frac{1}{1-(1-\mu)\tilde{\mathcal{D}}_{\substack{q_1\\0}}}\exp\left(-\imath\frac{2\pi q_1 r_1}{N_1}\right)\right.\\ &\qquad \left. + \sum_{q_1=1}^{N_1-1}\sum_{q_2=1}^{N_2-1}\frac{1}{1-(1-\mu)\tilde{\mathcal{D}}_{\substack{q_1\\q_2}}}\exp\left(-\imath\frac{2\pi q_1 r_1}{N_1}\right)\exp\left(-\imath\frac{2\pi q_2 r_2}{N_2}\right)\right]\\ &= \frac{\mu\lambda'_M}{N}\left[\frac{1}{1-(1-\mu)} + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(\delta_{r_2}N_2 - 1) + \frac{1}{1-(1-\mu)(1-m-\frac{m}{d-1})}(\delta_{r_1}N_1 - 1)\right.\\ &\qquad \left. + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(\delta_{r_1}N_1 - 1)(\delta_{r_2}N_2 - 1)\right]. \qquad \text{(B.20)}\end{aligned}$$ {eq:app:Q2DMsol}

In particular,

$$\begin{aligned}\tilde{\mathcal{Q}}_{\substack{0\\0}} &= \frac{\mu\lambda'_M}{N}\left[\frac{1}{\mu} + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(n-1) + \frac{1}{1-(1-\mu)(1-m-\frac{m}{d-1})}(D-1)\right.\\ &\qquad \left. + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(D-1)(n-1)\right]\\ &= 1. \qquad \text{(B.21a)}\end{aligned}$$ {eq:app:Q2D1}

We find λ'_M using the eq. (B.21a). Going back to eq. (B.20), when $r_1 = 0$, the two individuals are in the same deme. They are different when $r_2 \not\equiv 0$, and so:

$$Q_{\text{in}} = \frac{\mu\lambda'_M}{N}\left[\frac{1}{\mu} + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(-1) + \frac{1}{1-(1-\mu)(1-m-\frac{m}{d-1})}(D-1) + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(D-1)(-1)\right]. \tag{B.21b}$$

And when $r_1 \not\equiv 0$, the two individuals are in different demes:

$$Q_{\text{out}} = \frac{\mu\lambda'_M}{N}\left[\frac{1}{\mu} + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}(-1) + \frac{1}{1-(1-\mu)(1-m-\frac{m}{d-1})}(-1) + \frac{1}{1-(1-\mu)(d_{\text{self}}-d_{\text{in}})}\right]. \tag{B.21c}$$

With $d_{\text{self}} = d_{\text{in}} = (1-m)/n$, we recover the equations given in the main text (system (6)).

B.3 Wright-Fisher

For the Wright-Fisher updating, the equation for $\tilde{Q}$ is different:

$$\tilde{\mathcal{Q}}_{\substack{r_1\\r_2}} = \frac{1}{N}\sum_{q_1=0}^{N_1-1}\sum_{q_2=0}^{N_2-1}\frac{\mu\lambda'_{WF}}{1-(1-\mu)^2(\tilde{\mathcal{D}}_{\substack{q_1\\q_2}})^2}\exp\left(-\imath\frac{2\pi q_1 r_1}{N_1}\right)\exp\left(-\imath\frac{2\pi q_2 r_2}{N_2}\right), \tag{B.22}$$

with $\tilde{D}$ given in eq. (B.17b). In a subdivided population, with $N_1 = N_D$ and $N_2 = n$, this becomes

$$\begin{aligned}
\tilde{\mathcal{Q}}_{\substack{r_1\\r_2}} &= \frac{1}{N}\Bigg[\frac{\mu\lambda'_{WF}}{1-(1-\mu)^2(\tilde{\mathcal{D}}_{\substack{0\\0}})^2} + \sum_{q_2=1}^{N_2-1}\frac{\mu\lambda'_{WF}}{1-(1-\mu)^2(\tilde{\mathcal{D}}_{\substack{0\\q_2}})^2}\exp\left(-\imath\frac{2\pi q_2 r_2}{N_2}\right)\\
&\quad + \sum_{q_1=1}^{N_1-1}\frac{\mu\lambda'_{WF}}{1-(1-\mu)^2(\tilde{\mathcal{D}}_{\substack{q_1\\0}})^2}\exp\left(-\imath\frac{2\pi q_1 r_1}{N_1}\right)\\
&\quad + \sum_{q_1=1}^{N_1-1}\sum_{q_2=1}^{N_2-1}\frac{\mu\lambda'_{WF}}{1-(1-\mu)^2(\tilde{\mathcal{D}}_{\substack{q_1\\q_2}})^2}\exp\left(-\imath\frac{2\pi q_1 r_1}{N_1}\right)\exp\left(-\imath\frac{2\pi q_2 r_2}{N_2}\right)\Bigg]\\
&= \frac{\mu\lambda'_{WF}}{N}\Bigg[\frac{1}{1-(1-\mu)^2} + \frac{1}{1-(1-\mu)^2(d_{\text{self}}-d_{\text{in}})^2}(\delta_{q_2 N_2}-1)\\
&\quad + \frac{1}{1-(1-\mu)^2(1-m-\frac{m}{d-1})^2}(\delta_{q_1 N_1}-1)\\
&\quad + \frac{1}{1-(1-\mu)^2(d_{\text{self}}-d_{\text{in}})^2}(\delta_{q_1 N_1}-1)(\delta_{q_2 N_2}-1)\Bigg]\\
&= \frac{\mu\lambda'_{WF}}{N}\Bigg[\frac{1}{1-(1-\mu)^2} + \frac{1}{1-(1-\mu)^2(d_{\text{self}}-d_{\text{in}})^2}(\delta_{q_2 N_2}-1)\delta_{q_1 N_1}\\
&\quad + \frac{1}{1-(1-\mu)^2(1-m-\frac{m}{d-1})^2}(\delta_{q_1 N_1}-1)\Bigg].
\end{aligned} \tag{B.23}$$

{eq:app:Q2DWFsol}

To find λ'_{WF}, we solve $\tilde{\mathcal{Q}}_{\substack{0\\0}} = 1$, *i.e.*,

$$1 = \frac{\mu\lambda'_{WF}}{N}\left[\frac{1}{1-(1-\mu)^2} + \frac{1}{1-(1-\mu)^2(d_{\text{self}}-d_{\text{in}})^2}(N_2-1)N_1 + \frac{1}{1-(1-\mu)^2(1-m-\frac{m}{d-1})^2}(N_1-1)\right]. \tag{B.24a}$$

Then from eq. (B.23) we deduce

$$Q_{\text{in}} = \frac{\mu\lambda'_{WF}}{N}\left[\frac{1}{1-(1-\mu)^2} - \frac{1}{1-(1-\mu)^2(d_{\text{self}}-d_{\text{in}})^2}N_1 + \frac{1}{1-(1-\mu)^2(1-m-\frac{m}{d-1})^2}(N_1-1)\right]. \tag{B.24b}$$

and

$$Q_{\text{out}} = \frac{\mu\lambda'_{WF}}{N}\left[\frac{1}{1-(1-\mu)^2} - \frac{1}{1-(1-\mu)^2(1-m-\frac{m}{d-1})^2}\right]. \tag{B.24c}$$

With $d_{\text{self}} = d_{\text{in}} = (1-m)/n$, we recover the equations given in the main text (system (8)).