

Imperfect strategy transmission can reverse the role of population viscosity on the evolution of altruism

Manuscript elements: Figure 1, figure 2, table 1, online appendices A and B (including figure A1 and figure A2). Figure 2 is to print in color.

Keywords: Altruism, Subdivided population, Mutation, Migration, Cooperation, Island model, Wright-Fisher, Moran.

Manuscript type: Article.

Short Running Title: Mutation and altruism in subdivided populations.

Keywords: Altruism, Subdivided population, Mutation, Migration, Cooperation, Island model, Wright-Fisher, Moran.

| | | |
|--------------------|----------------|------------|
| Word Count: | Abstract | 148 |
| | Impact summary | 285 words |
| | Total | 4213 words |

Prepared using the suggested L^AT_EX template for *Am. Nat.*

1 **Abstract**

2 Population viscosity, *i.e.*, low emigration out of the natal deme, leads to high
3 within-deme relatedness, which is beneficial to the evolution of altruistic behav-
4 ior when social interactions take place among deme-mates. However, a detri-
5 mental side-effect of low emigration is the increase in competition among re-
6 lated individuals. The evolution of altruism depends on the balance between
7 these opposite effects. This balance is already known to be affected by details
8 of the life-cycle; we show here that it further depends on the fidelity of strategy
9 transmission from parents to their offspring. We consider different life-cycles
10 and identify thresholds of parent-offspring strategy transmission inaccuracy, above
11 which higher emigration can increase the frequency of altruists maintained in
12 the population. **EXPLAIN RESULT** Predictions were first obtained analytically
13 assuming weak selection and equal deme sizes, then confirmed with stochastic
14 simulations relaxing these assumptions. This result challenges the notion that
15 the evolution of altruism **REMOVE REQUIRE** requires limited dispersal.

16 **Introduction**

17 In his pioneering work on the evolution of social behavior, Hamilton suggested
18 that altruistic behavior would be associated to limited dispersal (Hamilton, 1964,
19 p. 10). This notion, that tighter links between individuals are beneficial to the
20 evolution of altruism, has been shown to hold in a number of population struc-
21 tures (see *e.g.* Allen et al., 2017; Lehmann et al., 2007; Ohtsuki et al., 2006; Taylor
22 et al., 2007a). The rationale is that altruism is favored when altruists interact
23 more with altruists than defectors do (Hamilton, 1975, p. 141; Fletcher & Doe-
24 beli, 2009), a condition that is met in viscous populations, *i.e.*, populations with
25 limited dispersal.

26 Yet, living next to your kin also implies competing against them (West et al.,
27 2002), which is detrimental to the evolution of altruism. The evolution of so-
28 cial traits hence depends on the balance between the positive effects of inter-
29 actions with related individuals and the detrimental consequences of kin com-
30 petition. Under specific conditions, the two effects can even compensate each
31 other, thereby annihilating the impact of population viscosity on the evolution
32 of altruism. First identified with computer simulations (Wilson et al., 1992), this
33 cancellation result was analyzed by Taylor (1992a) in a model with synchronous
34 generations (*i.e.*, Wright-Fisher model) and a subdivided population of constant,
35 infinite size. The cancellation result was later extended to heterogeneous pop-
36 ulations (Rodrigues & Gardner, 2012, with synchronous generations and infinite
37 population size), and other life-cycles, with generic regular population struc-
38 tures (Taylor et al., 2011, with synchronous generations but also with continuous
39 generations and Birth-Death updating). However, small changes in the model's
40 assumptions, such as overlapping generations (Taylor & Irwin, 2000) or the pres-
41 ence of empty sites (Alizon & Taylor, 2008) can tip the balance back in the favor
42 of altruism. This high dependence on life-cycle specificities highlights the dif-
43 ficulty of making general statements about the role of spatial structure on the

44 evolution of altruism. In this study, we will consider three different life-cycles:
45 Wright-Fisher, where the whole population is renewed at each time step, and
46 two Moran life-cycles (Birth-Death and Death-Birth), where a single individual
47 dies and is replaced at each time step. These life-cycles are classically used in
48 studies on altruism in structured populations, and are already known to have
49 different outcomes in models with perfect parent-offspring transmission (*e.g.*,
50 Lehmann et al., 2007; Ohtsuki et al., 2006; Rousset, 2004; Taylor, 1992a, 2010).

51 A large number of studies on the evolution of social behavior consider simple
52 population structures (typically, homogeneous populations *sensu* Taylor et al.
53 (2007a)) and often also infinite population sizes (but see Allen et al., 2017, for
54 results on any structure). These studies also make use of weak selection approx-
55 imations, and commonly assume rare (*e.g.*, Leturque & Rousset, 2002; Tarnita
56 & Taylor, 2014; Taylor et al., 2007b) or absent mutation (for models assuming
57 infinite population sizes, or models concentrating on fixation probabilities; see
58 Lehmann & Rousset, 2014; Van Cleve, 2015, for recent reviews). Often, these sim-
59 plifying assumptions are a necessary step towards obtaining explicit analytical
60 results. Although artificial, simple population structures (*e.g.*, regular graphs, or
61 subdivided populations with demes of equal sizes) help reduce the dimension-
62 ality of the system under study, in particular when the structure of the popula-
63 tion displays symmetries such that all sites behave the same way in expectation.
64 Weak selection approximations are crucial for disentangling spatial moments
65 (Lion, 2016), that is, changes in global *vs.* local frequencies (though they can in
66 some cases be relaxed, as in Mullon & Lehmann, 2014). Mutation, however, is
67 usually ignored by classical models of inclusive fitness because these models as-
68 sume infinite population sizes, so that there is no need to add mechanisms that
69 restore genetic diversity (Tarnita & Taylor, 2014). In populations of finite size,
70 this diversifying effect can be obtained thanks to mutation.

71 When strategy transmission is purely genetic, it makes sense to assume that

72 mutation is relatively weak. A social strategy can however also be culturally
73 transmitted from parent to offspring, in which case “rebellion” (as in Frank’s Re-
74 bellious Child Model (Frank, 1997)) does not have to be rare. Imperfect strategy
75 transmission can alter evolutionary dynamics, in particular in spatially struc-
76 tured populations (see *e.g.*, Allen et al., 2012; Débarre, 2017, for graph-structured
77 populations). Here, we want to explore the consequences of imperfect strategy
78 transmission from parents to their offspring on the evolution of altruistic behav-
79 ior in subdivided populations¹. The question was tackled by Frank (1997), but
80 with a “non fully” dynamic model. His method, done “in the spirit of compar-
81 ative statics” (Frank, 1997, p.1721)², precluded the exploration of the effects of
82 population viscosity on the evolution altruism.

83 For each of the three life-cycles that we consider, we compute the expected
84 (*i.e.*, long-term) frequency of altruists maintained in a subdivided population,
85 and investigate how this frequency is affected by mutation and emigration. We
86 find that, contrary to what happens with perfect strategy transmission, higher
87 emigration can increase the expected frequency of altruists in the population.

88 **Model and methods**

89 **Assumptions**

90 We consider a population of size N , subdivided into N_D demes connected by
91 dispersal, each deme hosting exactly n individuals (*i.e.*, each deme contains n
92 sites, each of which is occupied by exactly one individual; we have $nN_D = N$).
93 Each site has a unique label i , $1 \leq i \leq N$. There are two types of individuals in
94 the population, altruists and defectors. The type of the individual living at site i
95 ($1 \leq i \leq N$) is given by an indicator variable X_i , equal to 1 if the individual is an
96 altruist, and to 0 if it is a defector. The state of the entire population is given by

¹Note that for the sake of concision, we use the word “mutation” throughout the paper, keeping in mind that strategy transmission does not have to be genetic.

²Terms with quotes are citations from Frank’s article.

97 a N -long vector \mathbf{X} . For a given population state \mathbf{X} , the proportion of altruists is
 98 $\bar{X} = \sum_{i=1}^N X_i$. All symbols are summarized in table A1.

99 Reproduction is asexual. The offspring of altruists are altruists themselves
 100 with probability $1 - \mu_{1 \rightarrow 0}$, and are defectors otherwise ($0 < \mu_{1 \rightarrow 0} \leq 1/2$). Similarly,
 101 the offspring of defectors are defectors with probability $1 - \mu_{0 \rightarrow 1}$, and are altruists
 102 otherwise ($0 < \mu_{0 \rightarrow 1} \leq 1/2$). Our calculations will be simpler if we introduce the
 103 following change of parameters:

$$v = \frac{\mu_{0 \rightarrow 1}}{\mu_{1 \rightarrow 0} + \mu_{0 \rightarrow 1}} \quad (0 < v < 1), \text{ and} \quad (1a)$$

$$\mu = \mu_{1 \rightarrow 0} + \mu_{0 \rightarrow 1} \quad (0 < \mu \leq 1). \quad (1b)$$

104 Parents transmit their strategy to their offspring with probability $1 - \mu$; this trans-
 105 mission can be genetic or cultural (vertical cultural transmission), but for sim-
 106 plicity, we refer to the parameter μ as a mutation probability. With probability
 107 μ , offspring do not inherit their strategy from their parent but instead get one
 108 randomly: with probability v , they become altruists, with probability $1 - v$ they
 109 become defectors. We call the parameter v the mutation bias.

110 An individual of type X_k expresses a social phenotype $\phi_k = \delta X_k$, where δ is
 111 assumed to be small ($\delta \ll 1$). Social interactions take place within each deme,
 112 benefits are shared with the $n - 1$ other deme-mates. We assume that social in-
 113 teractions affect individual fecundity; f_k denotes the fecundity of the individual
 114 at site k . We denote by b the sum of the marginal effects of deme-mates' phe-
 115 notypes on the fecundity of a focal individual, and by $-c$ the marginal effect of
 116 a focal individual's phenotype on its own fecundity ($c \leq b$; see system (A20) for
 117 formal definitions).

118 Offspring remain in the parental deme with probability $1 - m$; when they
 119 do, they land on any site of the deme with equal probability (including the very
 120 site of their parent). With probability m , offspring emigrate to a different deme,

121 chosen uniformly at random among the other demes. Denoting by d_{ij} the prob-
 122 ability of moving from site i to site j , we have

$$d_{ij} = \begin{cases} d_{\text{in}} = \frac{1-m}{n} & \text{if both sites are in the same deme;} \\ d_{\text{out}} = \frac{m}{(N_D-1)n} & \text{if the two sites are in different demes,} \end{cases} \quad (2)$$

123 with $0 < m < 1 - \frac{1}{N_D}$. (This upper bound is here to ensure that within-deme
 124 relatedness R , which will be defined later in the article, remains positive.)

125 We denote by $B_i = B_i(\mathbf{X}, \delta)$ the expected number of successful offspring of the
 126 individual living at site i (successful means alive at the next time step), and by
 127 $D_i = D_i(\mathbf{X}, \delta)$ the probability that the individual living at site i dies. Both depend
 128 on the state of the population \mathbf{X} , but also on the way the population is updated
 129 from one time step to the next, *i.e.*, on the chosen life-cycle (also called updating
 130 rule). We also define

$$W_i := (1 - \mu) B_i + 1 - D_i, \quad (3)$$

131 which is a definition of fitness were only unmutated offspring are counted.

132 We will specifically explore three different life-cycles. At the beginning of
 133 each step of each life-cycle, all individuals produce offspring, that can be mu-
 134 tated; then these juveniles move, within the parental deme or outside of it, and
 135 land on a site. The next events occurring during the time step depend on the
 136 life-cycle:

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

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

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

145 **Methods**

146 **Analytical part**

147 The calculation steps to obtain the expected (*i.e.*, long-term) proportion of al-
 148 truists are given in Appendix B. They go as follows: first, we write an equation for
 149 the expected frequency of altruists in the population at time $t + 1$, conditional
 150 on the composition of the population at time t ; we then take the expectation of
 151 this quantity and consider large times t . After this, we write a first order expan-
 152 sion for phenotypic differences δ close to 0 (this corresponds to weak selection
 153 approximation).

154 The formula involves quantities that can be identified as neutral probabili-
 155 ties of identity by descent Q_{ij} , *i.e.*, the probability that individuals living at site
 156 i and j share a common ancestor and that no mutation occurred on either lin-
 157 eage since that ancestor, in a model with no selection ($\delta = 0$; this is the “muta-
 158 tion definition” of identity by descent (Rousset & Billiard, 2000).) In a subdivided
 159 population like ours, there are three possible values of Q_{ij} :

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

160 These neutral probabilities of identity by descent depend on the chosen life-
 161 cycle, and are also computed by taking the long-term expectation of conditional
 162 expectations after one time step (see Appendix C and C and supplementary Math-
 163 ematica file (Wolfram Research, Inc., 2017).)

164 Stochastic simulations

165 We also ran stochastic simulations (coded in C). The simulations were run for 10^8
 166 generations (one generation is one time step for the Wright-Fisher life-cycle, and
 167 N time steps for the Moran life-cycles). For each set of parameters and life-cycle,
 168 using R (R Core Team, 2015), we estimated the long-term frequency of altruists
 169 by sampling the population every 10^3 generations and computing the average
 170 frequency of altruists. All scripts are available at
 171 <https://flodebarre.github.io/SocEvolSubdivPop/>

172 Results

173 Expected frequencies of altruists for each life-cycle

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

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu B^*} v(1-v)(1-Q_{\text{out}}) \times \left[\underbrace{\frac{\partial W}{\partial f_{\bullet}}(-c) + \frac{\partial W}{\partial f_{\text{in}}}b}_{-\mathcal{C}} + \underbrace{\left(\frac{\partial W}{\partial f_{\bullet}}b + (n-1)\frac{\partial W}{\partial f_{\text{in}}}(-c) + (n-2)\frac{\partial W}{\partial f_{\text{in}}}b \right)}_{\mathcal{B}} \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right], \quad (5)$$

176 with W as defined in eq. (3). (Calculations leading to eq. (5) are presented in Ap-
 177 pendix B; notations are recapitulated in table A1)

178 The mutation bias v corresponds to the expected proportion of altruists in the
 179 population in the absence of selection (*i.e.*, when $\delta = 0$); δ is the parameter that
 180 scales the strength of selection ($\delta \ll 1$). The $-\mathcal{C}$ term groups the effects corre-
 181 sponding to the effects of a change of a focal individual's phenotype on its own
 182 fitness (with the fitness definition given in eq. (3).) The \mathcal{B} term corresponds to
 183 the sum of the effects on an individual's fitness of the change of deme-mates'
 184 phenotypes. It is multiplied by R , which is relatedness. Also note the overall
 185 $(1 - Q_{\text{out}})$ factor. All of these terms depend on the chosen life-cycle, and on pa-

rameters such as the mutation probability μ and the emigration probability m .

We now present the values of these different terms for the three life-cycles under study. The equations are obtained by replacing the derivatives of W by their specific formulas for each life-cycle, which are presented in Appendix B. The formulas of probabilities of identity by descent are more complicated, so not substituted, but they can be found in eq. (A38) and eq. (A42). The additional superscripts correspond to the life-cycles (WF: Wright-Fisher, M: Moran (either), DB: Moran Death-Birth, BD: Moran Birth-Death.) Finally, for each life-cycle, we further decompose the $-C$ and B terms into primary (subscript P) and secondary (subscript S) effects (West & Gardner, 2010). Primary effects correspond to unmediated consequences of interactions (they are included in $\frac{\partial W}{\partial f_i}$), while secondary effects correspond to consequences of interactions mediated by other individuals, including competition. Primary and secondary effects correspond to what was called “direct” and “indirect” effects, respectively, in Débarre (2017); Débarre et al. (2014), but the terminology was updated for more consistency with other authors. All equations are of the form

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}) \left[\underbrace{\begin{pmatrix} -C_P \\ -C_S \end{pmatrix}}_{-C} + \underbrace{\begin{pmatrix} B_P \\ +B_S \end{pmatrix}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right]. \quad (6)$$

Moran Birth-Death With the Birth-Death life-cycle, eq. (5) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\begin{pmatrix} (1-\mu)(-c) \\ +(b-c)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{-C^{\text{BD}}} + \underbrace{\begin{pmatrix} (1-\mu)b \\ +(b-c)(n-1)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{B^{\text{BD}}} \underbrace{\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M}}_{R^M} \right], \quad (7)$$

The secondary effects (second line in the parentheses) include competitive effects on the probability of reproducing, and consequences of social interactions

205 on the probability that a given individual dies. Note that the secondary effects
 206 remain negative for the realistic range of emigration values that we consider (*i.e.*,
 207 $m < 1 - 1/N_D$).

208 **Moran Death-Birth** With the Death-Birth life-cycle, eq. (5) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{-(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{-\mathcal{C}^{\text{DB}}} + \underbrace{\left(\frac{(1-\mu)b}{-(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{\mathcal{B}^{\text{DB}}} \underbrace{\left(\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M} \right)}_{R^M} \right], \quad (8)$$

209 With this life-cycle, Death occurs first, and the probability of dying is indepen-
 210 dent from the state of the population (since we assume that social interactions
 211 affect fecundity. We can therefore factor $(1-\mu)$ in all terms. The primary ef-
 212 fects (first lines in the parentheses) remain the same as with the Birth-Death
 213 life-cycle. However, the Death-Birth life-cycle leads to different secondary ef-
 214 fects compared to the Birth-Death life-cycle: competition occurs at a different
 215 scale (Grafen & Archetti, 2008). Finally, with this life-cycle as we defined it, the
 216 probabilities of identity by descent are the same as with the Birth-Death model.

Wright-Fisher

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^{\text{WF}}) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{-(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{-\mathcal{C}^{\text{WF}}} + \underbrace{\left(\frac{(1-\mu)b}{-(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{\mathcal{B}^{\text{WF}}} \underbrace{\left(\frac{Q_{\text{in}}^{\text{WF}} - Q_{\text{out}}^{\text{WF}}}{1 - Q_{\text{out}}^{\text{WF}}} \right)}_{R^{\text{WF}}} \right], \quad (9)$$

217 The only – but important – different between eq. (9) and eq. (8) is the value of the
 218 probabilities of identity by descent, because the number of individuals that are

219 updated at each time step differs.

220 **Primary effects**

221 Let us first focus on primary effects, shown on the first lines in the parentheses
222 in eq. (7)–eq. (9). They are the same for the three life-cycles that we consider,
223 and they do not depend on the emigration probability m . At the same time, as
224 shown on figure 1, the relatedness terms R^M and R^{WF} decrease with m (keeping
225 $m < 1 - \frac{1}{N_D}$; see figure 1) Consequently, if we ignored secondary effects, we would
226 conclude that even with imperfect strategy transmission ($\mu > 0$), increasing the
227 emigration probability m decreases the expected frequency of altruists in the
228 population (as our intuition may suggest already.) However, secondary effects
229 play a role as well.

230 **Changes with the emigration probability m**

231 For the three life-cycles, the secondary effects are negative (with $0 < m < 1 - \frac{1}{N_D}$)
232 and they increase with the emigration probability m . In other words, these detri-
233 mental secondary effects weaken when the emigration probability increases. This
234 effect goes against the reduction of relatedness R as m increases. Hence, we
235 need to consider the entire equations to know the overall effect of the emigra-
236 tion probability m on the expected frequency of altruists $\mathbb{E}[\bar{X}]$ and on how it is
237 affected by the (in)fidelity of parent-offspring transmission μ (see Figure 2).

238 The rather lengthy formulas that we obtain are relegated to the Appendix and
239 supplementary Mathematica file, and we concentrate here on the results.

240 **Moran Birth-Death**

241 For the Moran Birth-Death life-cycle, we find that the expected frequency of al-
242 truists $\mathbb{E}[\bar{X}]$ is a monotonic function of the emigration probability m ; the direc-
243 tion of the change depends on the value of the mutation probability μ compared

244 to a threshold value μ_c^{BD} . When $\mu < \mu_c^{\text{BD}}$, $\mathbb{E}[\bar{X}]$ decreases with m , while when
 245 $\mu > \mu_c^{\text{BD}}$, $\mathbb{E}[\bar{X}]$ increases with m . The critical value μ_c^{BD} is given by

$$\mu_c^{\text{BD}} = 1 - \frac{b - c + \sqrt{(b - c)(4bN^2 + b - c)}}{2bN} \quad (10)$$

246 (recall that N is the total size of the population, $N = nN_D$.) This result is illus-
 247 trated in figure 2(b); with the parameters of the figure, $\mu_c^{\text{BD}} \approx 0.026$. The thresh-
 248 old value increases with both deme size n and number of demes N_D , up to a
 249 maximum value $1 - \sqrt{1 - c/b}$ (equal to 0.034 with our parameters.)

250 With this life-cycle however, the expected frequency of altruists $\mathbb{E}[\bar{X}]$ remains
 251 lower than v , its value in the absence of selection (*i.e.*, when $\delta = 0$).

252 **Moran Death-Birth**

253 The relationship between $\mathbb{E}[\bar{X}]$ and m is a bit more complicated for the Moran
 254 Death-Birth life-cycle. For simplicity, we concentrate on what happens starting
 255 from low emigration probabilities (*i.e.*, on the sign of the slope of $\mathbb{E}[\bar{X}]$ as a func-
 256 tion of m when $m \rightarrow 0$). If the benefits b provided by altruists are relatively low
 257 ($b < c(n + 1)$), $\mathbb{E}[\bar{X}]$ initially increases with m provided the mutation probability
 258 μ is greater than a threshold value μ_c^{DB} given in eq. (11) below; otherwise, when
 259 the benefits are high enough, $\mathbb{E}[\bar{X}]$ initially increases with m for any value of μ .
 260 Combining these results, we write

$$\mu_c^{\text{DB}} = \begin{cases} \frac{(n+1)c - b}{(2n-1)b - (n-1)c} & \text{if } b < c(n+1), \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

261 When $b < c(n + 1)$, the mutation threshold does not depend on the number of
 262 demes N_D , but increases when the size of the demes n increases. In figure 2(a),
 263 the parameters are such that $\mu_c^{\text{DB}} = 0$.

264 When $\mu > \mu_c^{\text{DB}}$, the expected frequency of altruists $\mathbb{E}[\bar{X}]$ reaches a maximum

at an emigration probability m_c^{DB} (whose complicated equation is given in the supplementary Mathematica file), as can be seen in figure 2(a). When the mutation probability gets close to 0 ($\mu \rightarrow 0$), m_c^{DB} also gets close to 0.

With the Death-Birth life-cycle, the expected frequency of altruists is higher than its neutral value v for intermediate values of the emigration probability m (unless $\mu \rightarrow 0$, in which case the lower bound tends to 0).

Wright-Fisher

Under a Wright-Fisher updating, the expected frequency of altruists in the population reaches an extremum at the highest admissible emigration value $m = 1 - \frac{1}{N_D}$. This extremum is a maximum when the mutation probability is higher than a threshold value μ_c^{WF} given by

$$\mu_c^{\text{WF}} = 1 - \sqrt{1 - \frac{c}{b}}, \quad (12)$$

and it is a minimum otherwise. With the parameters of figure 2(c), $\mu_c^{\text{WF}} = 0.034$.

With the Wright-Fisher life-cycle however, the expected frequency of altruists remains below its value in the absence of selection, v .

Interpreting the effect of m on $\mathbb{E}[\bar{X}]$

The result, that higher emigration probabilities can actually favor altruistic behavior, may seem surprising, even though it is confirmed by simulations. The result may appear counterintuitive because explanations for the effect of population viscosity on the evolution of altruism often focus on primary effects. The role played by secondary effects is harder to grasp. To better understand the role played by the mutation probability μ , we now focus on a qualitative condition for the evolution of altruism and on the Death-Birth life-cycle:

$$\mathbb{E}[\bar{X}] > v \Leftrightarrow R^{\text{M}} > \frac{C^{\text{DB}}}{\mathcal{B}^{\text{DB}}}. \quad (13)$$

(having made sure that $\mathcal{B}^{\text{DB}} > 0$, as shown in the supplementary Mathematical file). With this life-cycle, the $\mathcal{C}^{\text{DB}}/\mathcal{B}^{\text{DB}}$ ratio does not change with the mutation probability μ , but it decreases with the emigration probability m ($0 < m < 1 - 1/N_D$). This decrease of the $\mathcal{C}^{\text{DB}}/\mathcal{B}^{\text{DB}}$ ratio is due to secondary effects (competition) diminishing as emigration increases. Relatedness, on the other hand, decreases with both μ and m (see figure 3(a)).

When the emigration probability m is vanishingly small, $\lim_{m \rightarrow 0} R^{\text{M}} \leq \lim_{m \rightarrow 0} \frac{\mathcal{C}^{\text{DB}}}{\mathcal{B}^{\text{DB}}}$, the two only being equal when $\mu = 0$. Hence, it is only when strategy transmission is perfect that condition (13) is satisfied for vanishingly low m . Then, as m increases, the $\frac{\mathcal{C}^{\text{DB}}}{\mathcal{B}^{\text{DB}}}$ ratio and relatedness R do not decrease with the same slope. Provided the mutation probability μ is not too high, *i.e.*, that R initially is not too low already, there can be a range of emigration values m such that condition (13) is satisfied (*i.e.*, the R curve is higher than the \mathcal{C}/\mathcal{B} curve in figure 3(a)).

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 ($\delta \ll 1$), 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.

When selection is strong, the patterns that we identified not only still hold but are even more marked, as shown on figure A1.

To relax the assumption of equal deme sizes, we randomly drew deme sizes at the beginning of simulations, with sizes ranging from 2 to 6 individuals and on average $\bar{n} = 4$ individuals per deme as previously. As shown in figure A2, the patterns initially obtained with a homogeneous population structure are robust when the structure is heterogeneous.

For the Moran model, it may seem odd that an offspring can replace its own parent (which can occur since $d_{ii} \neq 0$). Figure A3, plotted with dispersal prob-

abilities preventing immediate replacement of one's own parent (for all sites i ,
 $d_{ii} = d_{\text{self}} = 0$; $d_{\text{in}} = (1 - m)/(n - 1)$ for two different sites in the same deme, d_{out}
remaining unchanged), confirms that this does affect our conclusions.

The results are obtained in a population of finite size (the figures are done
with $N_D = 15$ demes), but still hold when the size of the population is larger (see
e.g., figure 3(b), showing the range of emigration and mutation values such that
altruism is favored, plotted also for $N_D \rightarrow \infty$).

Compared to graphs classically used in evolutionary graph theory (*e.g.*, regu-
lar random graphs, grids), the island model is particular because the interaction
graph and the dispersal graph are different: interactions take place only within
demes ($e_{\text{out}} = 0$), while offspring can disperse out of their natal deme ($d_{\text{out}} > 0$).
One may wonder whether our result depends on this difference between the two
graphs. Figure A4 shows that the result still holds when the dispersal and inter-
action graphs are the same. In this figure indeed, we let a proportion m (equal
to the dispersal probability) of interactions occur outside of the deme where the
individuals live, and set d_{self} , the probability of self replacement, equal to 0, so
that the dispersal and interactions graphs are the same. Our conclusions remain
unchanged.

Discussion

The expected frequency of altruists in a subdivided population can in- crease with the probability of emigration

Assuming that the transmission of a social strategy (being an altruist or a defec-
tor) from a parent to its offspring could be imperfect, we found that the expected
frequency of altruists maintained in a population could increase with the prob-
ability m of emigration out of the parental deme, a parameter tuning population
viscosity. This result can seem surprising, because it contradicts the conclusions

340 obtained under the assumption of nearly perfect strategy transmission (*i.e.*, in
 341 the case of genetic transmission, when mutation is very weak or absent). Under
 342 nearly perfect strategy transmission indeed, increased population viscosity (*i.e.*,
 343 decreased emigration probability) is either neutral (Taylor, 1992a, and dashed
 344 lines in figures 2(b)–(c)) or favorable (Taylor et al., 2007a, and dashed lines in
 345 figure 2(a)) to the evolution of altruistic behavior.

346 **Quantitative vs. qualitative measures**

347 Often, evolutionary success is measured qualitatively, by comparing a quantity
 348 (an expected frequency, or, in models with no mutation, a probability of fixation)
 349 to the value it would have in the absence of selection. In our model, this amounts
 350 to saying that altruism is favored whenever $\mathbb{E}[\bar{X}] > v$ (v is plotted as a horizon-
 351 tal dashed line in figure 2). Some of our conclusions change if we switch to this
 352 qualitative measure of evolutionary success: Under the Moran Birth-Death and
 353 Wright-Fisher life-cycles, population viscosity does not promote the evolution of
 354 altruism – actually, these two life-cycles cannot ever promote altruistic behavior
 355 for any regular population structure (Taylor et al., 2011), whichever the probabil-
 356 ity of mutation (Débarre, 2017). However, under a Moran Death-Birth life-cycle
 357 (figure 2(a)), altruism can be favored only at intermediate emigration probabil-
 358 ities. Starting for initially low values of m , increasing the emigration probability
 359 can still favor the evolution of altruism under this qualitative criterion (see fig-
 360 ure 3(b).)

361 **The result is due to secondary effects**

362 The result, that frequency of altruists can increase with the emigration proba-
 363 bility m , may seem counterintuitive. It is the case because verbal explanations
 364 for the evolution of altruism often rely on primary effects only. Relatedness R
 365 decreases with m , so it may be tempting to conclude that increases in the em-

366 igration probability m are necessarily detrimental to the evolution of altruism.
367 However, secondary effects play an opposite role, as competition decreases with
368 m . To further explain the relative weight of the detrimental and beneficial conse-
369 quences of increases in the emigration probability m , let us focus on the Death-
370 Birth life-cycle and consider the qualitative criterion for evolutionary success
371 ($\mathbb{E}[\bar{X}] > v$, *i.e.* $R > C/B$; figure 3.)

372 When parent-offspring strategy transmission is nearly perfect ($\mu \rightarrow 0$), for
373 vanishingly small emigration probabilities ($m \rightarrow 0$), both R and the C/B ratio
374 tend to 1. An increase in the mutation probability μ reduces R while leaving
375 C/B unchanged. In other words, for vanishingly small emigration probabilities,
376 altruism is favored by selection only when transmission fidelity is nearly perfect.
377 Let us now consider that benefits b of social interactions are high enough for
378 altruism to be favored at low m when $\mu \rightarrow 0$ (as in figure 3(a)). Starting from
379 low values of m , small increases in m have a stronger effect on the C/B ratio
380 than on relatedness R : local competition is initially so strong that the beneficial
381 reduction in competition caused by an increase in m initially predominates over
382 the detrimental reduction in relatedness R . The opposite holds for much higher
383 values of m : competition is already small enough that reducing it further does
384 not outweigh the reduction in relatedness R .

385 Secondary effects are less straightforward to understand than primary ef-
386 fects, and yet they play a crucial role for social evolution in spatially structured
387 populations. Competition among relatives is for instance the reason for Taylor
388 (1992b)’s cancellation result. Similarly, the qualitative differences between the
389 Moran Birth-Death and Moran Death-Birth life-cycles is explained by the differ-
390 ent scales of competition that the two life-cycle produce (Débarre et al., 2014;
391 Grafen & Archetti, 2008). Secondary effects are also behind the evolution of so-
392 cial behaviors such as spite (West & Gardner, 2010).

393 **How small is small and how large is large?**

394 Our results were derived under the assumption of weak selection, assuming that
395 the phenotypic difference between altruists and defectors is small ($\delta \ll 1$). We
396 considered any fidelity of transmission (any μ between 0 and 1) and population
397 size. However, most models considering subdivided populations assume nearly
398 perfect strategy transmission ($\mu \rightarrow 0$) and infinite population sizes (number of
399 demes $N_D \rightarrow \infty$). The point is technical, but it is important to know that the or-
400 der in which these limits are taken matters, *i.e.*, one needs to specify how small
401 μ and δ are compared to the inverse size of the population. This remark com-
402 plements findings by Sample & Allen (2017), who highlighted the quantitative
403 differences between different orders of weak selection and large population lim-
404 its.

405 **Imperfect transmission and Rebellious Children**

406 Our model bears resemblance to the Rebellious Child Model by Frank (1997),
407 who studied the evolution of a vertically transmitted cultural trait in an asexually
408 reproducing population. In Frank's model, however, relatedness r is treated as
409 a fixed parameter (as acknowledged in the legend of Figure 7 in Frank (1997)).
410 Our model is mechanistic; relatedness r necessarily depends on the mutation
411 probability μ , because probabilities of identity by descent do.

412 Mutation was also previously included in models investigating the mainte-
413 nance of cooperative microorganisms in the presence of cheaters (Brockhurst
414 et al., 2007; Frank, 2010). In both of these models however, only loss-of-function
415 mutation was considered, which corresponds to setting the mutation bias at
416 $\nu = 0$ in our model. This means that the all-cheaters state is absorbing; no matter
417 how favored cooperators may otherwise be, in the long run, a finite population
418 will only consist of cheaters.

419 **Cultural transmission**

420 Strategy transmission does not have to be genetic: it can be cultural. In our
421 model, strategy transmission occurs upon reproduction, so this is a case of ver-
422 tical cultural transmission.

423 The model could nevertheless be interpreted as a representation of horizon-
424 tal transmission, if we described reproduction as an instance of an individual
425 convincing another one to update its strategy. The Moran Death-Birth model
426 can be interpreted as a modified imitation scheme (Boyd & Richerson, 2002; Oht-
427 suki et al., 2006) – with a specific function specifying who is imitated –, with mu-
428 tation (Kandori et al., 1993). First, we choose uniformly at random an individual
429 who may change its strategy; with probability μ the individual chooses a random
430 strategy (altruistic with probability ν), and with probability $1 - \mu$ it imitates an-
431 other individual. Who is imitated depends on the distance to the focal individual
432 (with probability m it is a random individual in another deme) and on the “fe-
433 cundities” of those individuals (as shown in table A2). With this interpretation of
434 the updating rule however, there is not reproduction nor death anymore.

435 It remains to be investigated how imperfect strategy transmission would af-
436 fect the effect of population viscosity on the evolution of altruism in a model im-
437 plementing both reproduction and horizontal cultural transmission (as in Lehmann
438 et al., 2008). Such a model could then contrast the effects of imperfect genetic
439 transmission and imperfect horizontal cultural transmission.

440 **Coevolution of dispersal and social behavior**

441 This work also raises the question of what would happen if dispersal (*e.g.*, the
442 emigration probability m) could evolve as well. Recent work on the topic has
443 shown that under some conditions disruptive selection could take place, lead-
444 ing to a polymorphism between sessile altruists and mobile defectors (Mullon
445 et al., 2017; Parvinen, 2013). The assumptions of these studies however differ

446 from ours in important ways, in that they consider continuous traits and use
447 an adaptive dynamics framework, where, notably, mutations are assumed to be
448 very rare. It remains to be investigated how non-rare and potentially large mu-
449 tations would affect their result.

450 **Acknowledgements**

451 Thanks to Charles Mullan for detailed comments on a previous version of the
452 manuscript, and for suggesting the $\mathcal{BR}-\mathcal{C}$ decomposition. At ESEB 2017, Sébastien
453 Lion suggested using the R vs. \mathcal{C}/\mathcal{B} comparison to interpret the result. This work
454 was funded by a ANR-14-ACHN-0003-01 grant.

455 Literature Cited

- 456 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*
457 *tion* 62(6):1335–1344.
- 458 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,
459 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*
460 544(7649):227–230.
- 461 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects
462 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.
463 *Evolution of Cooperation.*
- 464 Boyd, R. & Richerson, P. J. 2002: Group beneficial norms can spread rapidly in a
465 structured population. *Journal of theoretical biology* 215(3):287–296.
- 466 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-
467 mediate disturbance. *Current Biology* 17(9):761–765.
- 468 Cockerham, C. C. & Weir, B. 1987: Correlations, descent measures: drift with
469 migration and mutation. *Proceedings of the National Academy of Sciences*
470 84(23):8512–8514.
- 471 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution
472 of social behavior in structured populations. *Journal of Theoretical Biology*
473 420:26 – 35.
- 474 Débarre, F.; Hauert, C. & Doebeli, M. 2014: Social evolution in structured popu-
475 lations. *Nature Communications* 5.
- 476 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-
477 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*
478 276(1654):13–19.
- 479 Frank, S. A. 1997: The Price equation, Fisher’s fundamental theorem, kin selec-
480 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 481 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*
482 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–
483 2522.
- 484 Grafen, A. & Archetti, M. 2008: Natural selection of altruism in inelastic viscous
485 homogeneous populations. *Journal of Theoretical Biology* 252(4):694 – 710.
- 486 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*
487 *Theoretical Biology* 7(1):1 – 16.
- 488 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-
489 tionary genetics. *Biosocial anthropology* 53:133–55.
- 490 Kandori, M.; Mailath, G. J. & Rob, R. 1993: Learning, mutation, and long run
491 equilibria in games. *Econometrica* 61(1):29–56.

492 Lehmann, L.; Feldman, M. & Foster, K. 2008: Cultural transmission can inhibit
 493 the evolution of altruistic helping. *The American Naturalist* 172(1):12–24.
 494 PMID: 18500938.

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

498 Lehmann, L. & Rousset, F. 2014: The genetical theory of social behaviour. *Philo-*
 499 *sophical Transactions of the Royal Society of London B: Biological Sciences*
 500 369(1642).

501 Leturque, H. & Rousset, F. 2002: Dispersal, kin competition, and the ideal free
 502 distribution in a spatially heterogeneous population. *Theoretical Population*
 503 *Biology* 62(2):169 – 180.

504 Lion, S. 2016: Moment equations in spatial evolutionary ecology. *Journal of the-*
 505 *oretical biology* 405:46–57.

506 Mullon, C.; Keller, L. & Lehmann, L. 2017: Co-evolution of dispersal with be-
 507 haviour favours social polymorphism. *bioRxiv* .

508 Mullon, C. & Lehmann, L. 2014: The robustness of the weak selection approxi-
 509 mation for the evolution of altruism against strong selection. *Journal of evo-*
 510 *lutionary biology* 27(10):2272–2282.

511 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule
 512 for the evolution of cooperation on graphs and social networks. *Nature*
 513 441(7092):502–505.

514 Parvinen, K. 2013: Joint evolution of altruistic cooperation and dispersal in a
 515 metapopulation of small local populations. *Theoretical population biology*
 516 85:12–19.

517 R Core Team. 2015: R: A Language and Environment for Statistical Computing.
 518 R Foundation for Statistical Computing, Vienna, Austria.

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

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

523 Rousset, F. & Billiard, S. 2000: A theoretical basis for measures of kin selection in
 524 subdivided populations: finite populations and localized dispersal. *Journal of*
 525 *Evolutionary Biology* 13(5):814–825.

526 Sample, C. & Allen, B. 2017: The limits of weak selection and large population
 527 size in evolutionary game theory. *Journal of mathematical biology* pages 1–
 528 33.

529 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors
530 in finite structured population models. *The American Naturalist* 184(4):477–
531 488.

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

534 Taylor, P. 2010: Birth–death symmetry in the evolution of a social trait. *Journal of*
535 *Evolutionary Biology* 23(12):2569–2578.

536 Taylor, P.; Lillicrap, T. & Cownden, D. 2011: Inclusive fitness analysis on mathe-
537 matical groups. *Evolution* 65(3):849–859.

538 Taylor, P. D. 1992b: Inclusive fitness in a homogeneous environment. Pro-
539 ceedings of the Royal Society of London. Series B: Biological Sciences
540 249(1326):299–302.

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

543 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-
544 bility in homogeneous structured populations. *Journal of Theoretical Biology*
545 249(1):101 – 110.

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

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

550 West, S. A. & Gardner, A. 2010: Altruism, spite, and greenbeards. *Science*
551 327(5971):1341–1344.

552 West, S. A.; Pen, I. & Griffin, A. S. 2002: Cooperation and competition between
553 relatives. *Science* 296(5565):72–75.

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

556 Wolfram Research, Inc. 2017: Mathematica, Version 11.1. Champaign, IL, 2017.

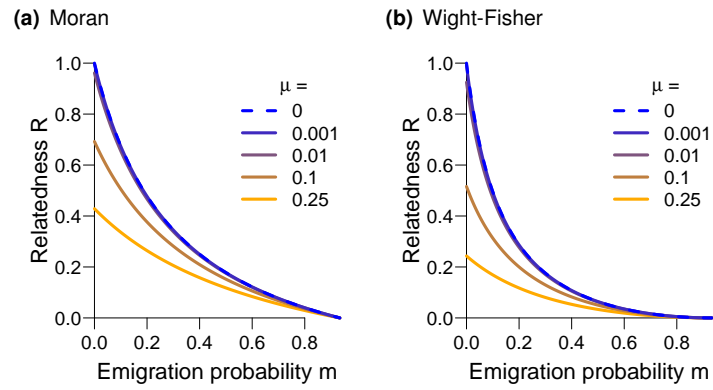


Figure 1: Within-deme relatedness of pairs of individuals, as a function of the emigration probability m , for different values of the mutation probability μ (from 0 [blue] to 0.25 [orange]), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals per deme, $N_D = 15$ demes.

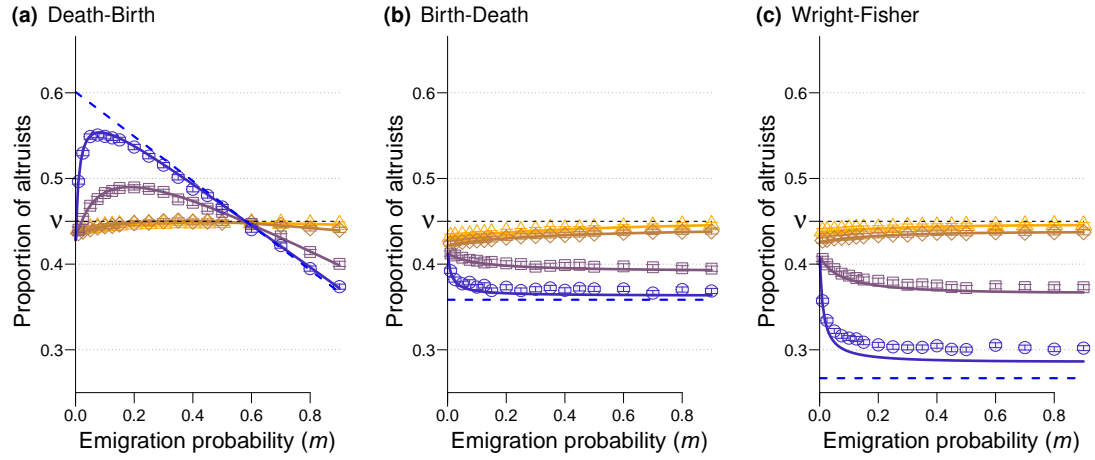


Figure 2: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different life-cycles ((a) Moran Death-Birth, (b) Moran Birth Death, (c) Wright-Fisher). The curves are the analytical results, the points are the output of numerical simulations. Parameters: $\delta = 0.005$, $v = 0.45$, $b = 15$, $c = 1$, $n = 4$ individuals per deme, $N_D = 15$ demes.

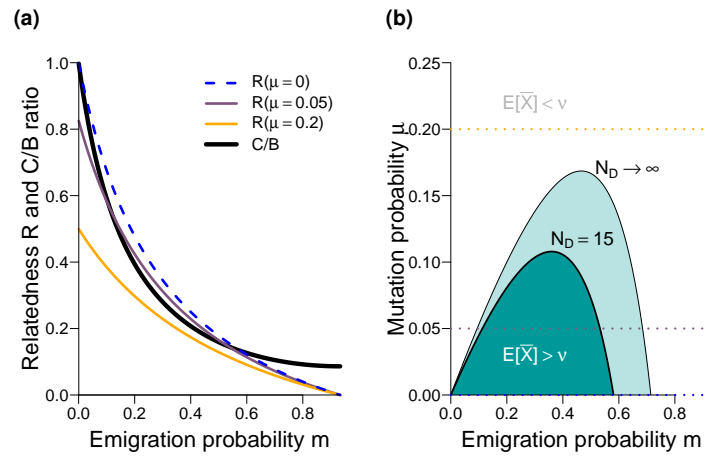


Figure 3: Understanding the effect of emigration m on whether altruism is favored in the Death-Birth life-cycle. (a) Comparison of the C/B ratio (thick black curve) and relatedness R (thin curves) for different values of the mutation probability μ (same color code as previously). (b) (m, μ) combinations for which $E[\bar{X}] > v$. The dotted horizontal lines correspond to the mutation values used in panel (a). Unless specified, all other parameters are the same as in figure 2.

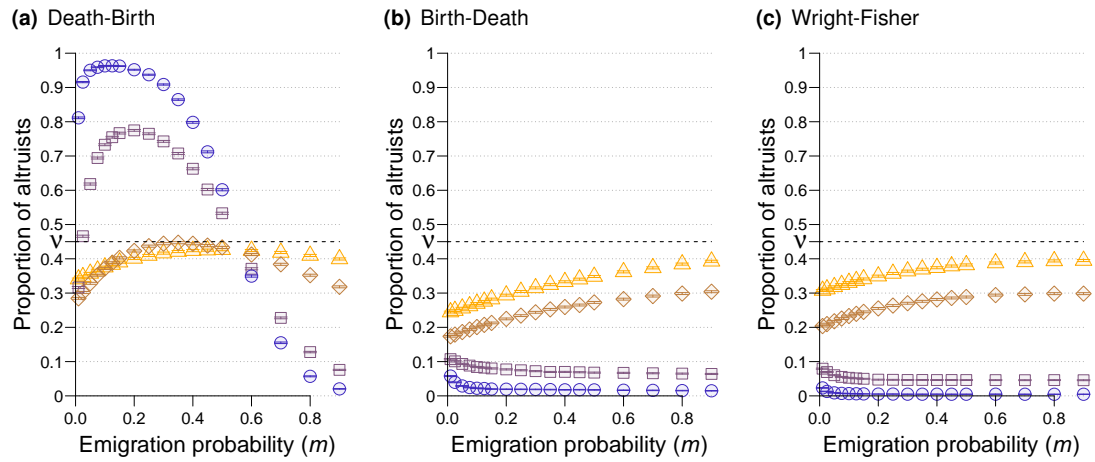


Figure A1: Equivalent of figure 2 (simulations only) but with strong selection ($\delta = 0.1$); please note the change of scale on the vertical axis. All other parameters and legends are identical to those of figure 2 (increasing mutation probabilities from blue dots to orange triangles).

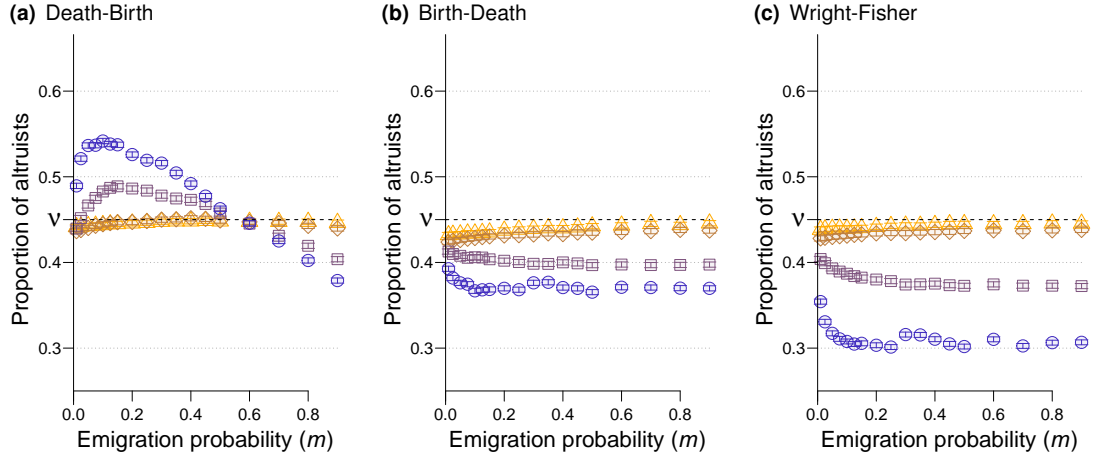


Figure A2: Equivalent of figure 2 (simulations only) but with a heterogeneous population structure: deme sizes range from 1 to 5 individuals per deme, the average deme size is 4 as in figure 2; all other parameters and legend are identical to those of figure 2.

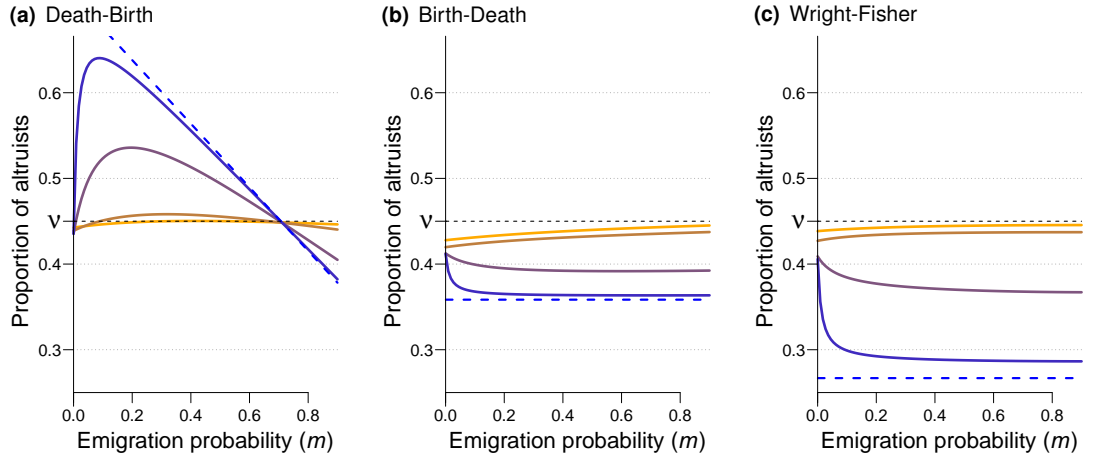


Figure A3: Equivalent of figure 2 (analysis only), with no self-replacement ($d_{ii} = d_{\text{self}} = 0$ for all sites).

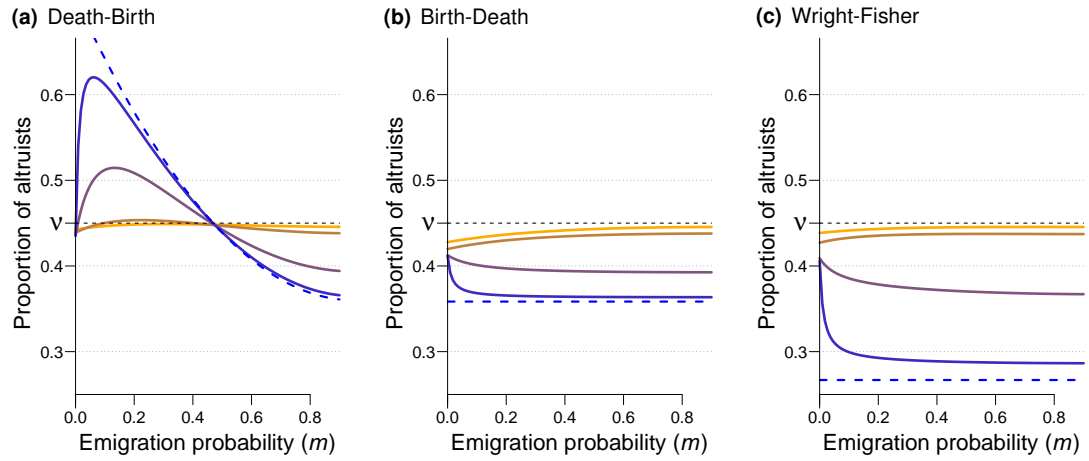


Figure A4: Equivalent of figure 2 (analysis only), with equal dispersal and interaction graphs (*i.e.*, no self-replacement [$d_{ii} = d_{\text{self}} = 0$ for all sites], and a proportion m of the interactions occurring outside of the home deme).

Supplementary Table

| | |
|---------------|---|
| b | Sum of the marginal effects of deme-mates' phenotypes on focal individual's fecundity (benefit) |
| \mathcal{B} | Sum of the marginal effects of deme-mates' phenotypes on the fitness W of a focal individual |
| B_i | Expected number of successful offspring of the individual living at site i (r.v.) |
| B^* | Value of B_i for all sites, in the absence of selection ($\delta = 0$) |
| c | Marginal effect of a focal individual's phenotype on its own fecundity (cost) |
| \mathcal{C} | Marginal effect of an individual's phenotype on its own fitness W |
| d_{ij} | Dispersal probability from site i to site j |
| D_i | Probability that the individual currently living at site i is dead at the end of the time step (r.v.) |
| e_{ij} | Interaction probability from site i to site j |
| f_i | Fecundity of the individual currently living at site i (r.v.) |
| n | Deme size |
| N_D | Number of demes |
| N | Total population size ($N = N_D n$) |
| m | Emigration probability |
| P_{ij} | (Long-term) Expected state of the pair of sites (i, j) |
| Q_{ij} | (Long-term) Probability of identity by descent of individuals at sites i and j |
| R | Pairwise within-deme relatedness (see eq. (5)) |
| W_i | Measure of fitness, counting offspring only when unmutated (see eq. (3)) |
| X_i | Indicator variable, equal to 1 if site i is occupied by an altruist, to 0 otherwise (r.v.) |
| \bar{X} | Frequency of altruists in the population (r.v.) |
| δ | Phenotypic distance between altruists and defectors; strength of selection |
| ϕ_i | Phenotype of the individual living at site i ; $\phi_i = \delta X_i$ (r.v.) |
| μ | Mutation probability |
| ν | Mutation bias: probability that mutant is altruist |
| P | Subscript corresponding to primary effects |
| S | Subscript corresponding to secondary effects |
| • | Subscript used to denote a focal individual |
| in | Subscript used when $i \neq j$ and the two sites are in the same deme |
| out | Subscript used when the two sites i and j are in different demes |
| self | Subscript used when $i = j$ |
| 0 | Sub- or superscript meaning that a quantity is evaluated at $\delta = 0$ |
| BD | Superscript corresponding to the Moran Birth-Death model |
| DB | Superscript corresponding to the Moran Death-Birth model |
| M | Superscript corresponding to a Moran model |
| WF | Superscript corresponding to the Wright-Fisher model |

Table A1: List of symbols. "r.v." means *random variable*.

Appendix

561 **A Mutation parameters**

562 In the main text, we first introduce effective mutation parameters: $\mu_{1 \rightarrow 0}$, the
563 probability that an altruist has defector offspring, and $\mu_{0 \rightarrow 1}$, the probability that
564 a defector has altruist offspring.

565 **A.1 Expected frequency of altruists at the mutation drift balance**

566 Let Y be the type of a randomly chosen individual in the population, and let
567 Y' be the type of a randomly chosen individual at the next time step. Given a
568 frequency v of altruists in the population, we have

$$\mathbb{E}[Y] = v, \quad (\text{A1a})$$

$$\mathbb{E}[Y'] = v(1 - \mu_{1 \rightarrow 0}) + (1 - v)\mu_{0 \rightarrow 1}. \quad (\text{A1b})$$

569 The expected frequency of altruists is found by solving $\mathbb{E}[Y] = \mathbb{E}[Y']$, and we
570 obtain

$$v = \frac{\mu_{0 \rightarrow 1}}{\mu_{1 \rightarrow 0} + \mu_{0 \rightarrow 1}}. \quad (\text{A2})$$

571 **A.2 Parent-offspring correlation at the mutation drift balance**

572 We can then compute the parent-offspring type correlation at the mutation-drift
573 balance. First, let us compute the parent-offspring covariance:

$$\begin{aligned} \text{Cov}[Y Y'] &= \mathbb{E}[Y Y'] - \mathbb{E}[Y']\mathbb{E}[Y] \\ &= v(1 - \mu_{1 \rightarrow 0}) - (v(1 - \mu_{1 \rightarrow 0}) + (1 - v)\mu_{0 \rightarrow 1})v \\ &= v(1 - v)(1 - \mu_{1 \rightarrow 0} - \mu_{0 \rightarrow 1}). \end{aligned} \quad (\text{A3})$$

574 Then, the standard deviations are given by

$$\begin{aligned} \sigma_Y &= \sqrt{\mathbb{E}[Y^2] - \mathbb{E}[Y]^2} = \sqrt{\mathbb{E}[Y] - \mathbb{E}[Y]^2} \\ &= \sqrt{v(1 - v)}, \end{aligned} \quad (\text{A4})$$

575 and

$$\begin{aligned} \sigma_{Y'} &= \sqrt{\mathbb{E}[Y'^2] - \mathbb{E}[Y']^2} = \sqrt{\mathbb{E}[Y'] - \mathbb{E}[Y']^2} \\ &= \sqrt{v(1 - v)(1 - \mu_{1 \rightarrow 0} - \mu_{0 \rightarrow 1}) - (v(1 - v)(1 - \mu_{1 \rightarrow 0} - \mu_{0 \rightarrow 1}))^2}. \end{aligned} \quad (\text{A5})$$

576 The parent-offspring correlation is given by

$$\text{Corr}[Y Y'] = \frac{\text{Cov}[Y Y']}{\sigma_Y \sigma_{Y'}};$$

577 using the formulas eq. (A3)–(A5), and replacing v by its value (mutation-drift
578 equilibrium, eq. (A2)), we obtain

$$\text{Corr}[Y Y'] = 1 - (\mu_{1 \rightarrow 0} + \mu_{0 \rightarrow 1}). \quad (\text{A6})$$

579 B Expected frequency of altruists

580 For a generic life-cycle

581 We want to compute the expected proportion of altruists in the population. We
 582 represent the state of the population at a given time t using indicator variables
 583 $X_i(t)$, $1 \leq i \leq N$, equal to 1 if the individual living at site i at time t is an altru-
 584 ist, and equal to 0 if it is a defector; these indicator variables are gathered in a
 585 N -long vector $\mathbf{X}(t)$. The set of all possible population states is $\Omega = \{0, 1\}^N$. The
 586 proportion of altruists in the population is written $\bar{X}(t) = \sum_{i=1}^N X_i(t)$. We denote
 587 by $B_{ji}(X(t), \delta)$, written B_{ji} for simplicity, the probability that the individual at
 588 site j at time $t+1$ is the newly established offspring of the individual living at site
 589 i at time t . The expected number of successful offspring produced by the indi-
 590 vidual living at site i at time t is given by $B_i = \sum_{j=1}^N B_{ji}$. We denote by $D_i(X(t), \delta)$
 591 (D_i for simplicity) the probability that the individual living at site i at time t has
 592 been replaced (*i.e.*, died) at time $t+1$. These quantities depend on the chosen
 593 life-cycle and on the state of the population; they are given in table A2 for each
 594 of the life-cycles that we consider.

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

Table A2: Formulas of B_{ji} and D_i for each of the life-cycles that we consider; f_i (shorthand notation for $f_i(X, \delta)$) is the fecundity of the individual living at site i , and d_{ji} is a dispersal probability, given in eq. (2).

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

$$D_i = \sum_{j=1}^N B_{ij} \quad (\text{A7a})$$

596 holds for all sites i . The structure of the population is also such that in the ab-
 597 sence of selection ($\delta = 0$, so that $f_i = 1$ for all sites $1 \leq i \leq N$), all individuals have
 598 the same probability of dying and the same probability of having successful off-
 599 spring (*i.e.*, of having offspring that become adults at the next time step), so that

$$D_i^0 = \sum_{j=1}^N B_{ji}^0 = B_i^0 =: B^*, \quad (\text{A7b})$$

where the 0 subscript means that the quantities are evaluated for $\delta = 0$. 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 between eq. (A7b) and eq. (A7a) 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

$$\mathbb{E}[\bar{X}(t+1)|\mathbf{X}(t)] = \frac{1}{N} \sum_{i=1}^N [B_i(1-\mu)X_i + (1-D_i)X_i + B_i\mu\nu]. \quad (\text{A8a})$$

The first term within the brackets corresponds to births of unmutated offspring from parents who are altruists (X_i). The second term corresponds to the survival of altruists. The third term corresponds to the births of mutants who became altruists (which occurs with probability ν), whichever the type of the parent.

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; the expected frequency of altruists does not change anymore for large times t (realized frequencies of course keep changing). We denote by $\xi(\mathbf{X}, \delta, \mu)$ the probability that the population is in state \mathbf{X} , given the strength of selection δ and the mutation probability μ . Taking the expectation of eq. (A8a) ($\mathbb{E}[\bar{X}] = \sum_{\mathbf{X} \in \Omega} \bar{X} \xi(\mathbf{X}, \delta, \mu)$), we obtain, after reorganizing:

$$0 = \frac{1}{N} \sum_{\mathbf{X} \in \Omega} \left[\sum_{i=1}^N (B_i(1-\mu)X_i - D_iX_i) + \sum_{i=1}^N B_i\mu\nu \right] \xi(\mathbf{X}, \delta, \mu). \quad (\text{A9})$$

Now, we use the assumption of weak selection ($\delta \ll 1$) and consider the first-order expansion of eq. (A9) for δ close to 0. First, we note that in the absence of selection ($\delta = 0$), the population is at a mutation-drift balance; the expected state of every site i is then $\mathbb{E}_0[X_i] = \sum_{\mathbf{X} \in \Omega} X_i \xi(\mathbf{X}, 0, \mu) = \nu$ (recall that ν is the mutation bias parameter). Secondly, we note that for all the life-cycles that we consider, the total number of deaths in the population during one time step does not depend on population composition (it is exactly 1 death for the Moran life-cycles, and exactly N for the Wright-Fisher life-cycle), so that $\sum_{i,j=1}^N B_{ij}$ does not depend on δ . Consequently, the last term of eq. (A9) vanishes when we take its derivative with respect to δ , for the life-cycles that we consider.

The first order expansion of eq. (A9) yields

$$0 = \frac{\delta}{N} \sum_{i=1}^N \left[\sum_{\mathbf{X} \in \Omega} \left(\frac{\partial B_i}{\partial \delta} (1-\mu) - \frac{\partial D_i}{\partial \delta} \right) X_i \xi(\mathbf{X}, 0, \mu) - \sum_{\mathbf{X} \in \Omega} \mu B^* X_i \frac{\partial \xi}{\partial \delta} \right] + O(\delta^2), \quad (\text{A10})$$

where the derivatives are evaluated at $\delta = 0$, and where we used eq. (A7b) to simplify the second term. For conciseness, we define

$$W_i = (1-\mu)B_i + (1-D_i), \quad (\text{A11})$$

a measure of fitness counting offspring only when they are unmutated. With

632 this, using the expectation notation, and denoting by $\mathbb{E}_0[\cdot]$ expectations under
633 $\delta = 0$, we can rewrite and reorganize eq. (A10) as

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \mathbb{E}_0 \left[\frac{\partial W_i}{\partial \delta} X_i \right] + O(\delta^2). \quad (\text{A12})$$

634 Now, we use a first time the law of total probabilities, taking individual pheno-
635 types ϕ_k are intermediate variables:

$$\begin{aligned} \frac{\partial W_i}{\partial \delta} &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} \frac{\partial \phi_k}{\partial \delta} \\ &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} X_k, \end{aligned} \quad (\text{A13})$$

636 by definition of ϕ_k , and where the derivatives are evaluated for all $\phi_i = 0$. With
637 the notation $P_{ij} = \mathbb{E}_0[X_i X_j]$, eq. (A12) becomes

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} P_{ik} + O(\delta^2). \quad (\text{A14})$$

638 So far, we have not used the specificities of the population structure that we con-
639 sider. Once we have fixed a focal individual i , in expectation there are only three
640 types of individuals: the focal itself, $n - 1$ other individuals in the focal's deme
641 (denoted by in), and $N - n$ individuals in other demes (denoted by out). We note
642 that given that the size of the population is fixed ($\sum_{i=1}^N (B_i - D_i) = 0$), and given
643 that the total number of births does not depend on population composition in
644 the life-cycles that we consider,

$$\sum_{i=1}^N \frac{\partial W_i}{\partial \delta} = 0,$$

645 which we can rewrite as (Rousset & Billiard, 2000)

$$\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} + (N - n) \frac{\partial W_i}{\partial \phi_{\text{out}}} = 0. \quad (\text{A15})$$

646 With this, eq. (A14) becomes

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} \frac{P_{\text{in}} - P_{\text{out}}}{P_{ii} - P_{\text{out}}} \right) (P_{ii} - P_{\text{out}}) + O(\delta^2). \quad (\text{A16})$$

647 We can also replace the P terms by

$$\begin{aligned} P_{ij} &= Q_{ij} \nu + (1 - Q_{ij}) \nu^2 \\ &= \nu^2 + \nu(1 - \nu) Q_{ij}. \end{aligned} \quad (\text{A17})$$

648 In Appendix C, we will see using recursions on P_{ij} that Q_{ij} can be interpreted
649 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. Eq. (A16) becomes

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\underbrace{\frac{\partial W_i}{\partial \phi_i}}_{-C} + \underbrace{(n-1) \frac{\partial W_i}{\partial \phi_{\text{in}}}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) (1 - Q_{\text{out}}) \nu (1 - \nu) + O(\delta^2). \quad (\text{A18})$$

We can further decompose the derivatives, now using the fecundities f_ℓ as intermediate variables, *i.e.*,

$$\frac{\partial W_i}{\partial \phi_k} = \sum_{\ell=1}^N \frac{\partial W_i}{\partial f_\ell} \frac{\partial f_\ell}{\partial \phi_k}. \quad (\text{A19})$$

With our notation, and given that social interactions take place within demes and affect fecundity, we have

$$\left. \frac{\partial f_\ell}{\partial \phi_\ell} \right|_{\delta=0} = -c, \quad (\text{A20a})$$

$$\left. \frac{\partial f_\ell}{\partial \phi_{\text{in}}} \right|_{\delta=0} = \frac{b}{n-1}, \quad (\text{A20b})$$

$$\left. \frac{\partial f_\ell}{\partial \phi_{\text{out}}} \right|_{\delta=0} = 0. \quad (\text{A20c})$$

Eq. (A18) then becomes (using notation \bullet to refer to the focal individual itself, and where $W = W_i$, since the derivatives are the same for all i):

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \delta \nu (1 - \nu) (1 - Q_{\text{out}}) \times \left(\underbrace{\left(\frac{\partial W}{\partial f_\bullet} (-c) + \frac{\partial W}{\partial f_{\text{in}}} b \right)}_{-C} + \underbrace{\left(\frac{\partial W}{\partial f_\bullet} b + (n-1) \frac{\partial W}{\partial f_{\text{in}}} (-c) + (n-2) \frac{\partial W}{\partial f_{\text{in}}} b \right)}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) + O(\delta^2). \quad (\text{A21})$$

(As previously, all derivatives are evaluated at $\delta = 0$.)

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

$$\mathbb{E}[\bar{X}] = \nu + \delta \left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0} + O(\delta^2), \quad (\text{A22})$$

where $\left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0}$ is obtained from eq. (A21). We then need to replace the B_i and D_i terms by their formulas for each life-cycle (given in table A2).

Derivatives for the specific life-cycles

Using the formulas presented in table A2, and the definition of $W = W_i$ given in eq. (A11), we obtain the following equations:

Moran Birth-Death

$$\left. \frac{\partial W^{\text{BD}}}{\partial f_{\bullet}} \right|_{\delta=0} = (1-\mu) \left(\frac{1}{N} - \frac{1}{N^2} \right) - \left(\frac{1-m}{nN} - \frac{1}{N^2} \right) = \frac{1-\mu}{N} + \frac{\mu}{N^2} - \frac{1-m}{nN}, \quad (\text{A23a})$$

$$\left. \frac{\partial W^{\text{BD}}}{\partial f_{\text{in}}} \right|_{\delta=0} = (1-\mu) \left(-\frac{1}{N^2} \right) - \left(\frac{1-m}{nN} - \frac{1}{N^2} \right) = \frac{\mu}{N^2} - \frac{1-m}{nN}. \quad (\text{A23b})$$

Moran Death-Birth

$$\left. \frac{\partial W^{\text{DB}}}{\partial f_{\bullet}} \right|_{\delta=0} = \frac{1-\mu}{N} \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A24a})$$

$$\left. \frac{\partial W^{\text{DB}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -\frac{1-\mu}{N} \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A24b})$$

666 In addition, for both Moran life-cycles, we have $B_{\text{M}}^* = 1/N$.

Wright-Fisher

$$\left. \frac{\partial W^{\text{WF}}}{\partial f_{\bullet}} \right|_{\delta=0} = (1-\mu) \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A25a})$$

$$\left. \frac{\partial W^{\text{WF}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -(1-\mu) \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A25b})$$

667 For the Wright-Fisher life-cycle, we have $B_{\text{WF}}^* = 1$.

668 Combining these equations with eq. (A22) and eq. (A21) given eqs. (7)–(9) in
669 the main text.

670 C Probabilities of identity by descent

671 Expected state of pairs of sites and probabilities of identity by descent

672 Here we show the link between the expected state of a pair of sites P_{ij} and prob-
673 abilities of identity by descent Q_{ij} . In our derivation of $\mathbb{E}[\bar{X}]$, P_{ij} is the quantity
674 that appears, but most studies use Q_{ij} . Both are evaluated in the absence of
675 selection ($\delta = 0$).

676 Moran model

677 In a Moran model, exactly one individual dies and one individual reproduces
678 during one time step. Given a state \mathbf{X} at time t , at time $t + 1$ both sites i and
679 $j \neq i$ are occupied by altruists, if i it was the case at time t and neither site was
680 replaced by a non-altruist (first term in eq. (A26)), or ij if exactly one of the two
681 sites was occupied by a non-altruist at time t , but the site was replaced by an
682 altruist (second and third terms of eq. (A26)):

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & X_i X_j \left(1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - \nu)) \right) \\ & + X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu\nu) \\ & + X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu\nu). \end{aligned} \quad (\text{A26})$$

683 We take the expectation of this quantity, and consider that the stationary dis-
684 tribution is reached ($t \rightarrow \infty$); then $\mathbb{E}[X_i X_j(t+1)] = \mathbb{E}[X_i X_j(t)]$, and we obtain

$$P_{ij} = \frac{1}{2} \left(\sum_{k=1}^N (1 - \mu) (d_{kj} P_{ki} + d_{ki} P_{kj}) \right) + \mu\nu^2 \quad (i \neq j), \quad (\text{A27})$$

685 while $P_{ii} = \nu$.

686 Now we substitute $P_{ij} = \nu^2 + \nu(1 - \nu)Q_{ij}$ in eq. (A27), we obtain

$$Q_{ij} = \frac{1}{2} \sum_{k=1}^N (1 - \mu) (d_{ki} Q_{kj} + d_{kj} Q_{ki}), \quad (\text{A28})$$

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

692 **Wright-Fisher model**

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & \sum_{k, \ell=1}^N d_{ki} d_{\ell j} \left(X_k X_\ell (1 - \mu + \mu v)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu v)(\mu v) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu v)^2 \right) \end{aligned} \quad (\text{A29})$$

695 The first term of eq. (A29) corresponds to both parents being altruists, and hav-
 696 ing altruist offspring; the second line corresponds to exactly one parent being
 697 altruist, and the third line to both parents being non-altruists (in this latter case,
 698 the two offspring have to be both mutants to be altruists).

699 Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k, \ell=1}^N (P_{kl}(1 - \mu)^2) + (2 - \mu)\mu v^2. \quad (\text{A30})$$

700 Replacing P_{ij} by $v^2 + v(1 - v)Q_{ij}$, eq. (A30) becomes

$$Q_{ij} = \sum_{k, \ell=1}^N d_{ki} d_{\ell j} Q_{k\ell} (1 - \mu)^2. \quad (\text{A31})$$

701 Again, Q_{ij} corresponds to a probability of identity by descent: the individuals at
 702 sites i and j are identical by descent if their parents were and if neither mutated
 703 $((1 - \mu)^2)$.

704 **Probabilities of identity by descent in a subdivided population**

705 Two individuals are said to be identical by descent if there has not been any mu-
 706 tation on either lineage since their common ancestor. Because of the structure
 707 of the population, there are only three types of pairs of individuals, and hence
 708 three different values of the probabilities of identity by descent of pairs of sites
 709 Q_{ij} :

$$Q_{ij} = \begin{cases} 1 & \text{when } i = j; \\ Q_{\text{in}} & \text{when } i \neq j \text{ and both sites are in the same deme;} \\ Q_{\text{out}} & \text{when sites } i \text{ and } j \text{ are in different demes.} \end{cases} \quad (\text{A32})$$

710 The values of Q_{in} and Q_{out} depend on the type of life-cycle that we consider.

711 Here, we will use formulas derived in Débarre (2017) for “two-dimensional
 712 population structures”. The name comes from the fact that we only need two
 713 types of transformations to go from any site to any other site in the population:
 714 permutations on the deme index, and permutations on the within-deme index.
 715 We rewrite site labels ($1 \leq i \leq N$) as (ℓ_1, ℓ_2) , where ℓ_1 is the index of the deme ($1 \leq$
 716 $\ell_1 \leq N_D$) and ℓ_2 the position of the site within the deme ($1 \leq \ell_2 \leq n$). Then, we
 717 introduce notations \tilde{d}_{i_1, i_2} and \tilde{Q}_{i_1, i_2} , that correspond to the dispersal probability and
 718 probability of identity by descent to a site at distances i_1 and i_2 in the among-
 719 demes and within-deme dimensions (e.g., $\tilde{d}_{i_1, i_2} = d_{j_1, j_1 + i_1, j_2' j_2 + i_2}$).

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

722 **Moran model**

723 In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1, r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1, q_2}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{A33a})$$

724 with

$$\tilde{D}_{q_1, q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1, \ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{A33b})$$

725 and λ'_M such that $\tilde{Q}_0 = 1$. Let us first compute \tilde{D}_{q_1, q_2} in the case of a subdivided
 726 population, with $N_1 = N_D$ and $N_2 = n$:

$$\begin{aligned} \tilde{D}_{q_1, q_2} &= d_{\text{self}} + \sum_{\ell_2=1}^{N_2-1} d_{\text{in}} \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) + \sum_{\ell_1=1}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} d_{\text{out}} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2} (N_2 - 1) + (1 - \delta_{q_2}) (-1)) d_{\text{in}} + (\delta_{q_1} (N_1 - 1) + (1 - \delta_{q_1}) (-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}. \end{aligned} \quad (\text{A34a})$$

727 (δ_q is equal to 1 when q is equal to 0 modulo the relevant dimension, and to 0
 728 otherwise). So for the three types of distances that we need to consider (distance
 729 0, distance to another deme-mate, distance to individual in another deme), and
 730 with $N_1 = N_D$ and $N_2 = n$, we obtain

$$\tilde{D}_0 = 1, \quad (\text{A35a})$$

$$\tilde{D}_{q_1} = 1 - m - \frac{m}{N_D - 1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \quad (\text{A35b})$$

$$\tilde{D}_{q_1} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \quad (\text{A35c})$$

731 So for \tilde{Q} , using system (A35) in eq. (A33a),

$$\begin{aligned} \tilde{Q}_{r_1} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu) \tilde{D}_0} + \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_2}} \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right. \\ &\quad + \sum_{q_1=1}^{N_1-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu)} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_2} N_2 - 1) \right. \\ &\quad + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (\delta_{r_1} N_1 - 1) \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_1} N_1 - 1)(\delta_{r_2} N_2 - 1) \right]. \quad (\text{A36}) \end{aligned}$$

732 In particular,

$$\begin{aligned} \tilde{Q}_0 &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (n - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (N_D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (N_D - 1)(n - 1) \right] \\ &= 1. \quad (\text{A37a}) \end{aligned}$$

733 We find λ'_M using eq. (A37a). Let's now go back to eq. (A36): when $r_1 = 0$, the two
 734 individuals are in the same deme. They are different when $r_2 \neq 0$, and so:

$$\begin{aligned} Q_{\text{in}} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (D - 1)(-1) \right]. \quad (\text{A37b}) \end{aligned}$$

735 And when $r_1 \neq 0$, the two individuals are in different demes:

$$Q_{\text{out}} = \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})}(-1) + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} \right]. \quad (\text{A37c})$$

736 With $d_{\text{self}} = d_{\text{in}} = (1 - m)/n$, we eventually obtain:

$$Q_{\text{in}}^M = \frac{(1 - \mu)(m + \mu(N_D(1 - m) - 1))}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}, \quad (\text{A38a})$$

$$Q_{\text{out}}^M = \frac{(1 - \mu)m}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}. \quad (\text{A38b})$$

737 The probability that two different deme-mates are identical by descent, Q_{in}^M , de-
 738 creases monotonically with the emigration probability m , while Q_{out}^M monoton-
 739 ically increases with m (see figure A5(a)).

740 When the mutation probability μ is vanishingly small ($\mu \rightarrow 0$), both Q_{in}^M and
 741 Q_{out}^M are equal to 1: in the absence of mutation indeed, the population ends up
 742 fixed for one of the two types, and all individuals are identical by descent. Note
 743 that we obtain a different result if we first assumed that the size of the popu-
 744 lation is infinite ($N_D \rightarrow \infty$), because the order of limits matters; for instance,
 745 $\lim_{d \rightarrow \infty} Q_{\text{out}}^M = 0$.

746 **Wright-Fisher**

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

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right), \quad (\text{A39})$$

748 with $\tilde{\mathcal{D}}$ given in eq. (A33b). In a subdivided population, with $N_1 = N_D$ and $N_2 = n$,
 749 this becomes

$$\begin{aligned}
 \tilde{Q}_{r_1 r_2} &= \frac{1}{N} \left[\frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_0)^2} + \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_2})^2} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right. \\
 &\quad + \sum_{q_1=1}^{N_1-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \\
 &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \right. \\
 &\quad + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_1} N_1 - 1) (\delta_{q_2} N_2 - 1) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \delta_{q_1} N_1 \right. \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \right]. \tag{A40}
 \end{aligned}$$

750 To find λ'_{WF} , we solve $\tilde{Q}_0 = 1$, i.e.,

$$1 = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (N_2 - 1) N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \tag{A41a}$$

751 Then from eq. (A40) we deduce

$$Q_{\text{in}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \tag{A41b}$$

752 and

$$Q_{\text{out}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} \right]. \tag{A41c}$$

753 With $d_{\text{self}} = d_{\text{in}} = (1 - m)/n$, we obtain:

$$Q_{\text{in}}^{\text{WF}} = \frac{-N_D + M_1 + M_2}{(n - 1)N_D + M_1 + M_2}, \tag{A42a}$$

$$Q_{\text{out}}^{\text{WF}} = \frac{-\frac{1}{N_D-1} M_1 + M_2}{(n - 1)N_D + M_1 + M_2}, \tag{A42b}$$

754 with

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

(These formulas are compatible with, *e.g.*, results presented by Cockerham & Weir (1987), adapted for haploid individuals).

In the Wright-Fisher life-cycle, $Q_{\text{in}}^{\text{WF}}$ decreases until $m = m_c^{\text{WF}} = \frac{N_D - 1}{N_D}$, while $Q_{\text{out}}^{\text{WF}}$ follows the opposite pattern. The threshold value m_c^{WF} corresponds to an emigration probability so high that $d_{\text{in}} = d_{\text{out}}$.

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

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 A5(a) and A5(b)).

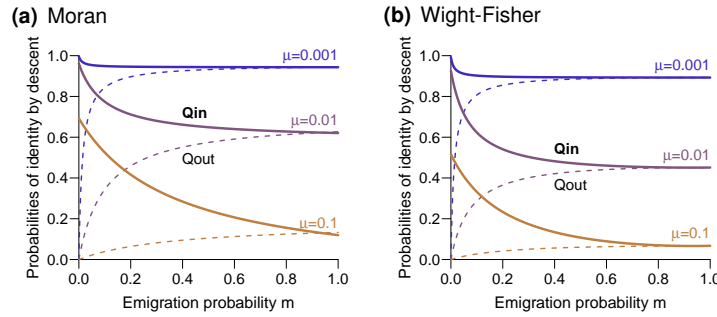


Figure A5: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in} , full curves) and two individuals in different demes (Q_{out} , dashed curves), as a function of the emigration probability m , for different values of the mutation probability μ (0.001, 0.01, 0.1), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals per deme, $N_D = 15$ demes.