

Article Type: Letter

Article Title: Imperfect strategy transmission can reverse the role of population viscosity on the evolution of altruism .

Author: Florence Débarre – Centre Interdisciplinaire de Recherche en Biologie (CIRB), Collège de France, CNRS UMR 7241 - Inserm U1050, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France. florence.debarre@normalesup.org

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	283 words
	Total	3989 words

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. Predictions were first obtained analytically assuming weak se-
13 lection and equal deme sizes, then confirmed with stochastic simulations relax-
14 ing these assumptions. This result challenges the notion that the evolution of
15 altruism requires limited dispersal.

16 **Impact Summary**

17 The evolution of altruistic behavior has fascinated and puzzled evolutionary bi-
18 ologists for a long time: how can a strategy whereby individuals help others at
19 their own cost be maintained in a population? One answer is the fact that altru-
20 ists may interact with other altruists more often than non-altruists do, a situa-
21 tion made possible by spatial structure and low emigration. Low emigration in-
22 deed means that an individual is mostly surrounded by related individuals; when
23 social strategies are faithfully transmitted from parents to offspring, and social
24 interactions are local as well, then altruists interact mainly with other altruists.
25 However, this also means that related individuals have to compete against each
26 other. Whether altruism eventually evolves depends on the balance between
27 these beneficial and detrimental consequences of low emigration. Previous work
28 has shown that the balance depends on the life-cycle that the population under-
29 goes; under nearly perfect strategy transmission, low emigration goes from be-
30 ing neutral to the evolution of altruism (when generations are synchronous and
31 non-overlapping) to favorable. In this work, we show that this conclusion qual-
32 itatively changes when offspring do not necessarily adopt their parent's strat-
33 egy, that is, when strategy transmission is imperfect. This can be due to muta-
34 tion when transmission is genetic, but also to imperfect vertical cultural trans-
35 mission. We identify thresholds of strategy transmission infidelity, above which
36 higher emigration is more conducive to the evolution of altruism than low em-
37 igration. The predictions are first obtained mathematically under the restric-
38 tive assumptions that selection is weak and that all demes have the same size,
39 but are then confirmed with computer simulations relaxing these assumptions.
40 This work shows that the evolution of altruism does not require – and even can
41 be hampered by – low emigration.

42 1 Introduction

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

51 Yet, living next to your kin also implies competing against them (West et al.,
52 2002), which is detrimental to the evolution of altruism. The evolution of so-
53 cial traits hence depends on the balance between the positive effects of inter-
54 actions with related individuals and the detrimental consequences of kin com-
55 petition. Under specific conditions, the two effects can even compensate each
56 other, thereby annihilating the impact of population viscosity on the evolution
57 of altruism. First identified with computer simulations (Wilson et al., 1992), this
58 cancellation result was analyzed by Taylor (1992a) in a model with synchronous
59 generations (*i.e.*, Wright-Fisher model) and a subdivided population of constant,
60 infinite size. The cancellation result was later extended to heterogeneous pop-
61 ulations (Rodrigues & Gardner, 2012, with synchronous generations and infinite
62 population size), and other life-cycles, with generic regular population struc-
63 tures (Taylor et al., 2011, with synchronous generations but also with continuous
64 generations and Birth-Death updating). However, small changes in the model's
65 assumptions, such as overlapping generations (Taylor & Irwin, 2000) or the pres-
66 ence of empty sites (Alizon & Taylor, 2008) can tip the balance back in the favor
67 of altruism. This high dependence on life-cycle specificities highlights the dif-
68 ficulty of making general statements about the role of spatial structure on the

69 evolution of altruism. In this study, we will consider three different life-cycles:
70 Wright-Fisher, where the whole population is renewed at each time step, and
71 two Moran life-cycles (Birth-Death and Death-Birth), where a single individual
72 dies and is replaced at each time step. These life-cycles are classically used in
73 studies on altruism in structured populations, and are already known to have
74 different outcomes in models with perfect parent-offspring transmission (*e.g.*,
75 Taylor, 1992a; Rousset, 2004; Ohtsuki et al., 2006; Lehmann et al., 2007; Taylor,
76 2010).

77 A large number of studies on the evolution of social behavior consider simple
78 population structures (typically, homogeneous populations *sensu* Taylor et al.
79 (2007a)) and often also infinite population sizes (but see Allen et al., 2017, for
80 results on any structure). These studies also make use of weak selection approx-
81 imations, and commonly assume rare (*e.g.*, Leturque & Rousset, 2002; Taylor
82 et al., 2007b; Tarnita & Taylor, 2014) or absent mutation (for models assuming
83 infinite population sizes, or models concentrating on fixation probabilities; see
84 Lehmann & Rousset, 2014; Van Cleve, 2015, for recent reviews). Often, these sim-
85 plifying assumptions are a necessary step towards obtaining explicit analytical
86 results. Although artificial, simple population structures (*e.g.*, regular graphs, or
87 subdivided populations with demes of equal sizes) help reduce the dimension-
88 ality of the system under study, in particular when the structure of the popula-
89 tion displays symmetries such that all sites behave the same way in expectation.
90 Weak selection approximations are crucial for disentangling spatial moments
91 (Lion, 2016), that is, changes in global *vs.* local frequencies (though they can in
92 some cases be relaxed, as in Mullan & Lehmann, 2014). Mutation, however, is
93 usually ignored by classical models of inclusive fitness because these models as-
94 sume infinite population sizes, so that there is no need to add mechanisms that
95 restore genetic diversity (Tarnita & Taylor, 2014). In populations of finite size,
96 this diversifying effect can be obtained thanks to mutation.

97 When strategy transmission is purely genetic, it makes sense to assume that
 98 mutation is relatively weak. A social strategy can however also be culturally
 99 transmitted from parent to offspring, in which case “rebellion” (as in Frank’s Re-
 100 bellious Child Model (Frank, 1997)) does not have to be rare. Imperfect strategy
 101 transmission can alter evolutionary dynamics, in particular in spatially struc-
 102 tured populations (see *e.g.*, Allen et al., 2012; Débarre, 2017, for graph-structured
 103 populations). Here, we want to explore the consequences of imperfect strategy
 104 transmission from parents to their offspring on the evolution of altruistic behav-
 105 ior in subdivided populations¹. The question was tackled by Frank (1997), but
 106 with a non fully dynamic model. His method done “in the spirit of comparative
 107 statics” (p.1721) precluded the exploration of the effects of population viscosity
 108 on the evolution altruism.

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

114 2 Model and methods

115 2.1 Assumptions

116 We consider a population of size N , subdivided into N_D demes connected by
 117 dispersal, each deme hosting exactly n individuals (*i.e.*, each deme contains n
 118 sites, each of which is occupied by exactly one individual; we have $nN_D = N$).
 119 Each site has a unique label i , $1 \leq i \leq N$. There are two types of individuals in
 120 the population, altruists and defectors. The type of the individual living at site i
 121 ($1 \leq i \leq N$) is given by an indicator variable X_i , equal to 1 if the individual is an

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.

122 altruist, and to 0 if it is a defector. The state of the entire population is given by
 123 a N -long vector \mathbf{X} . For a given population state \mathbf{X} , the proportion of altruists is
 124 $\bar{X} = \sum_{i=1}^N X_i$. All symbols are summarized in table S1.

125 Reproduction is asexual. Parents transmit their strategy to their offspring
 126 with probability $1 - \mu$; this transmission can be genetic or cultural (vertical cul-
 127 tural transmission), but for simplicity, we refer to the parameter μ as a mutation
 128 probability. With probability μ , offspring do not inherit their strategy from their
 129 parent but instead get one randomly: with probability ν , they become altruists,
 130 with probability $1 - \nu$ they become defectors. We call the parameter ν the muta-
 131 tion bias.

132 An individual of type X_k expresses a social phenotype $\phi_k = \delta X_k$, where δ is
 133 assumed to be small ($\delta \ll 1$). Social interactions take place within each deme,
 134 benefits are shared with the $n - 1$ other deme-mates. We assume that social in-
 135 teractions affect individual fecundity; f_k denotes the fecundity of the individual
 136 at site k . We denote by b the sum of the marginal effects of deme-mates' phe-
 137 notypes on the fecundity of a focal individual, and by $-c$ the marginal effect of
 138 a focal individual's phenotype on its own fecundity ($c \leq b$) (see system (A.14) for
 139 formal definitions).

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

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

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

147 We denote by $B_i = B_i(\mathbf{X}, \delta)$ the expected number of successful offspring of the
 148 individual living at site i (successful means alive at the next time step), and by
 149 $D_i = D_i(\mathbf{X}, \delta)$ the probability that the individual living at site i dies. Both depend
 150 on the state of the population \mathbf{X} , but also on the way the population is updated
 151 from one time step to the next, *i.e.*, on the chosen life-cycle (also called updating
 152 rule). We also define

$$W_i := (1 - \mu)B_i + 1 - D_i, \quad (2)$$

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

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

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

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

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

167 2.2 Methods

168 2.2.1 Analytical part

169 The calculation steps to obtain the expected (*i.e.*, long-term) proportion of al-
 170 truists are given in Appendix A. They go as follows: first, we write an equation for
 171 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 and consider large times t . After this, we write a first order expansion for phenotypic differences δ close to 0 (this corresponds to weak selection approximation).

The formula involves 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 ($\delta = 0$; this is the “mutation definition” of identity by descent (Rousset & Billiard, 2000).) In a subdivided population like ours, there are three possible values of Q_{ij} :

$$Q_{ij} = \begin{cases} 1 & \text{when } i = j, \\ Q_{\text{in}} & \text{when } i \neq j \text{ and both sites are in the same deme,} \\ Q_{\text{out}} & \text{when both sites are in different demes.} \end{cases} \quad (3)$$

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 B.1 and B.2 and supplementary Mathematica file (Wolfram Research, Inc., 2017).)

2.2.2 Stochastic simulations

We also ran stochastic simulations (coded in C). The simulations were 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, using R (R Core Team, 2015), we estimated the long-term frequency of altruists by sampling the population every 10^3 generations and computing the average frequency of altruists. All scripts are available at <https://github.com/flodebarre/SocEvolSubdivPop/tree/master/Programs>

194 3 Results

195 3.1 Expected frequencies of altruists for each life-cycle

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

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu B^*} v(1-v)(1-Q_{\text{out}}) \times \left[\underbrace{\frac{\partial W}{\partial f_{\bullet}}(-c) + \frac{\partial W}{\partial f_{\text{in}}}b}_{-\mathcal{C}} + \underbrace{\left(\frac{\partial W}{\partial f_{\bullet}}b + (n-1)\frac{\partial W}{\partial f_{\text{in}}}(-c) + (n-2)\frac{\partial W}{\partial f_{\text{in}}}b \right)}_{\mathcal{B}} \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right], \quad (4)$$

198 with W as defined in eq. (2). (Calculations leading to eq. (4) are presented in Ap-
 199 pendix A; notations are recapitulated in table S1)

200 The mutation bias v corresponds to the expected proportion of altruists in the
 201 population in the absence of selection (*i.e.*, when $\delta = 0$); δ is the parameter that
 202 scales the strength of selection ($\delta \ll 1$). The $-\mathcal{C}$ term groups the effects corre-
 203 sponding to the effects of a change of a focal individual's phenotype on its own
 204 fitness (with the fitness definition given in eq. (2).) The \mathcal{B} term corresponds to
 205 effects on an individual's fitness of the change of a deme-mate's phenotype. It is
 206 multiplied by R , which is relatedness. Also note the overall $(1 - Q_{\text{out}})$ factor. All
 207 of these terms depend on the chosen life-cycle, and on parameters such as the
 208 mutation probability μ and the emigration probability m .

209 We now present the values of these different terms for the three life-cycles
 210 under study. The equations are obtained by replacing the derivatives of W by
 211 their specific formulas for each life-cycle, which are presented in Appendix A.2.
 212 The formulas of probabilities of identity by descent are more complicated, so
 213 not substituted, but they can be found in eq. (B.32) and eq. (B.36). The addi-
 214 tional superscripts correspond to the life-cycles (WF: Wright-Fisher, M: Moran
 215 (either), DB: Moran Death-Birth, BD: Moran Birth-Death.) Finally, for each life-
 216 cycle, we further decompose the $-\mathcal{C}$ and \mathcal{B} terms into primary (subscript P) and

217 secondary (subscript S) effects (West & Gardner, 2010). Primary effects corre-
 218 spond to unmediated consequences of interactions (they are included in $\frac{\partial W}{\partial f_i}$),
 219 while secondary effects correspond to consequences of interactions mediated
 220 by other individuals, competition. Primary and secondary effects correspond
 221 to what was called “direct” and “indirect” effects, respectively, in Débarre et al.
 222 (2014); Débarre (2017), but the terminology was updated for more consistency
 223 with other authors. All equations are of the form

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}) \left[\underbrace{\begin{pmatrix} -C_P \\ -C_S \end{pmatrix}}_{-C} + \underbrace{\begin{pmatrix} B_P \\ +B_S \end{pmatrix}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right]. \quad (5)$$

224 **Moran Birth-Death** With the Birth-Death life-cycle, eq. (4) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\begin{pmatrix} (1-\mu)(-c) \\ +(b-c)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{-C^{\text{BD}}} + \underbrace{\begin{pmatrix} (1-\mu)b \\ +(b-c)(n-1)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{B^{\text{BD}}} \underbrace{\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M}}_{R^M} \right], \quad (6)$$

225 The secondary effects (second line in the parentheses) include competitive ef-
 226 fects on the probability of reproducing, and consequences of social interactions
 227 on the probability that a given individual dies. Note that the secondary effects
 228 remain negative for the realistic range of emigration values that we consider (*i.e.*,
 229 $m < 1 - 1/N_D$).

230 **Moran Death-Birth** With the Death-Birth life-cycle, eq. (4) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{-(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{-C^{\text{DB}}} + \underbrace{\left(\frac{(1-\mu)b}{-(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{B^{\text{DB}}} \right] \underbrace{\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M}}_{R^M}, \quad (7)$$

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

Wright-Fisher

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

239 The only – but important – different between eq. (8) and eq. (7) is the value of the
 240 probabilities of identity by descent, because the number of individuals that are
 241 updated at each time step differs.

242 3.2 Primary effects

243 Let us first focus on primary effects, shown on the first lines in the parentheses
244 in eq. (6)–eq. (8). They are the same for the three life-cycles that we consider,
245 and they do not depend on the emigration probability m . At the same time, as
246 shown on figure 1, the relatedness terms R^M and R^{WF} decrease with m (keeping
247 $m < 1 - \frac{1}{N_D}$; see figure 1) Consequently, ignoring secondary effects, we would
248 conclude that even with imperfect strategy transmission ($\mu > 0$), increasing the
249 emigration probability m decreases the expected frequency of altruists in the
250 population (as our intuition may suggest already.) However, secondary effects
251 play a role as well.

252 3.3 Changes with the emigration probability m

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

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

262 3.3.1 Moran Birth-Death

263 For this life-cycle, we find that the expected frequency of altruists $\mathbb{E}[\bar{X}]$ is a mono-
264 tonic function of the emigration probability m ; the direction of the change de-
265 pends on the value of the mutation probability μ compared to a threshold value
266 μ_c^{BD} . When $\mu < \mu_c^{BD}$, $\mathbb{E}[\bar{X}]$ decreases with m , while when $\mu > \mu_c^{BD}$, $\mathbb{E}[\bar{X}]$ increases

267 with m . The critical value μ_c^{BD} is given by

$$\mu_c^{\text{BD}} = 1 - \frac{b - c + \sqrt{(b - c)(4bN^2 + b - c)}}{2bN} \quad (9)$$

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

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

274 3.3.2 Moran Death-Birth

275 The relationship between $\mathbb{E}[\bar{X}]$ and m is a bit more complicated for this life-
 276 cycle. For simplicity, we concentrate on what happens starting from low emigra-
 277 tion probabilities (*i.e.*, on the sign of the slope of $\mathbb{E}[\bar{X}]$ as a function of m when
 278 $m \rightarrow 0$). If the benefits b provided by altruists are relatively low ($b < c(n + 1)$),
 279 $\mathbb{E}[\bar{X}]$ initially increases with m provided the mutation probability μ is greater
 280 than a threshold value μ_c^{DB} given in eq. (10) below; otherwise, when the benefits
 281 are high enough, $\mathbb{E}[\bar{X}]$ initially increases with m for any value of μ . Combining
 282 these results, we write

$$\mu_c^{\text{DB}} = \begin{cases} \frac{(n+1)c - b}{(2n-1)b - (n-1)c} & \text{if } b < c(n+1), \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

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

286 When $\mu > \mu_c^{\text{DB}}$, the expected frequency of altruists $\mathbb{E}[\bar{X}]$ reaches a maximum
 287 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).

3.3.3 Wright-Fisher

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

$$\mu_c^{\text{WF}} = 1 - \sqrt{1 - \frac{c}{b}}, \quad (11)$$

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 .

3.4 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 is 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:

$$\mathbb{E}[\bar{X}] > v \Leftrightarrow R^{\text{M}} > \frac{C^{\text{DB}}}{B^{\text{DB}}}. \quad (12)$$

(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 eq. (12) 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 (12) is satisfied (*i.e.*, the R curve is higher than the \mathcal{C}/\mathcal{B} curve in figure 3(a)).

3.5 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 S1.

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 S2, 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 S3, 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.

Finally, 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 $N_D \rightarrow \infty$).

4 Discussion

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

Assuming that the transmission of a social strategy (being an altruist or a defector) 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 probability m of emigration out of the parental deme, a parameter tuning population viscosity. This result can seem surprising, because it contradicts the conclusions obtained under the assumption of nearly perfect strategy transmission (*i.e.*, in the case of genetic transmission, when mutation is very weak or absent). Under nearly perfect strategy transmission indeed, increased population viscosity (*i.e.*, decreased emigration probability) is either neutral (Taylor, 1992a, and dashed lines in figures 2(b)–(c)) or favorable (Taylor et al., 2007a, and dashed lines in figure 2(a)) to the evolution of altruistic behavior.

Quantitative vs. qualitative measures

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 model, this amounts

361 to saying that altruism is favored whenever $\mathbb{E}[\bar{X}] > v$ (v is plotted as a horizon-
 362 tal dashed line in figure 2). Some of our conclusions change if we switch to this
 363 qualitative measure of evolutionary success: Under the Moran Birth-Death and
 364 Wright-Fisher life-cycles, population viscosity does not promote the evolution of
 365 altruism – actually, these two life-cycles cannot ever promote altruistic behavior
 366 for any regular population structure (Taylor et al., 2011), whichever the probabil-
 367 ity of mutation (Débarre, 2017). However, under a Moran Death-Birth life-cycle
 368 (figure 2(a)), altruism can be favored only at intermediate emigration probabil-
 369 ities: increased emigration can still favor the evolution of altruism under this
 370 qualitative criterion (see figure 3(b).)

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

372 The result, that frequency of altruists can increase with the emigration proba-
 373 bility m , may seem counterintuitive. It is the case because verbal explanations
 374 for the evolution of altruism often rely on primary effects only. Relatedness R
 375 decreases with m , so it may be tempting to conclude that increases in the em-
 376 igration probability m are necessarily detrimental to the evolution of altruism.
 377 However, secondary effects play an opposite role, as competition decreases with
 378 m . To further explain the relative weight of the detrimental and beneficial conse-
 379 quences of increases in the emigration probability m , let us focus on the Death-
 380 Birth life-cycle and consider the qualitative criterion for evolutionary success
 381 ($\mathbb{E}[\bar{X}] > v$, *i.e.* $R > C/B$; figure 3.)

382 When parent-offspring strategy transmission is nearly perfect ($\mu \rightarrow 0$), for
 383 vanishingly small emigration probabilities ($m \rightarrow 0$), both R and the C/B ratio
 384 tend to 1. An increase in the mutation probability μ reduces R while leaving
 385 C/B unchanged. In other words, for vanishingly small emigration probabilities,
 386 altruism is favored by selection only when transmission fidelity is nearly perfect.
 387 Let us now consider that benefits b of social interactions are high enough for

388 altruism to be favored at low m when $\mu \rightarrow 0$ (as in figure 3(a)). Starting from
 389 low values of m , small increases in m have a stronger effect on the C/B ratio
 390 than on relatedness R : local competition is initially so strong that the beneficial
 391 reduction in competition caused by an increase in m initially predominates over
 392 the detrimental reduction in relatedness R . The opposite holds for much higher
 393 values of m : competition is already small enough that reducing it further does
 394 not outweigh the reduction in relatedness R .

395 Secondary effects are less straightforward to understand than primary ef-
 396 fects, and yet they play a crucial role for social evolution in spatially structured
 397 populations. Competition among relatives is for instance the reason for Taylor
 398 (1992b)’s cancellation result. Similarly, the qualitative differences between the
 399 Moran Birth-Death and Moran Death-Birth life-cycles is explained by the dif-
 400 ferent scales of competition that the two life-cycle produce (Grafen & Archetti,
 401 2008; Débarre et al., 2014). Secondary effects are also behind the evolution of
 402 social behaviors such as spite (West & Gardner, 2010).

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

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

415 **Imperfect transmission and Rebellious Children**

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

422 Mutation was also previously included in models investigating the mainte-
423 nance of cooperative microorganisms in the presence of cheaters (Brockhurst
424 et al., 2007; Frank, 2010). In both of these models however, only loss-of-function
425 mutation was considered, which corresponds to setting the mutation bias at
426 $\nu = 0$ in our model. This means that the all-cheaters state is absorbing; no matter
427 how favored cooperators may otherwise be, in the long run, a finite population
428 will only consist of cheaters.

429 **Interactions and dispersal graphs**

430 Compared to graphs classically used in evolutionary graph theory (*e.g.*, regular
431 random graphs, grids), the island model is particular because the interaction
432 graph and the dispersal graph are different: interactions take place only within
433 demes ($e_{\text{out}} = 0$), while offspring can disperse out of their natal deme ($d_{\text{out}} > 0$).
434 One may wonder whether our result depends on this difference between the two
435 graphs. Figure S4 shows that the result still holds when the dispersal and inter-
436 action graphs are the same. In this figure indeed, we let a proportion m (equal
437 to the dispersal probability) of interactions occur outside of the deme where the
438 individuals live, and set d_{self} , the probability of self replacement, equal to 0, so
439 that the dispersal and interactions graphs are the same. Our conclusions remain
440 unchanged.

441 **Coevolution of dispersal and social behavior**

442 This work also raises the question of what would happen if dispersal (*e.g.*, the
443 emigration probability m) could evolve as well. Recent work on the topic has
444 shown that under some conditions disruptive selection could take place, lead-
445 ing to a polymorphism between sessile altruists and mobile defectors (Parvinen,
446 2013; Mullan et al., 2017). The assumptions of these studies however differ from
447 ours in important ways, in that they consider continuous traits and use an adap-
448 tive dynamics framework, where, notably, mutations are assumed to be very
449 rare. It remains to be investigated how non-rare and potentially large mutations
450 would affect their result.

451 **Acknowledgements**

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

456 References

- 457 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*
458 *tion* 62(6):1335–1344.
- 459 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,
460 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*
461 544(7649):227–230.
- 462 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects
463 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.
464 *Evolution of Cooperation.*
- 465 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-
466 mediate disturbance. *Current Biology* 17(9):761–765.
- 467 Cockerham, C. C. & Weir, B. 1987: Correlations, descent measures: drift with
468 migration and mutation. *Proceedings of the National Academy of Sciences*
469 84(23):8512–8514.
- 470 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution
471 of social behavior in structured populations. *Journal of Theoretical Biology*
472 420:26 – 35.
- 473 Débarre, F.; Hauert, C. & Doebeli, M. 2014: Social evolution in structured popu-
474 lations. *Nature Communications* 5.
- 475 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-
476 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*
477 276(1654):13–19.
- 478 Frank, S. A. 1997: The Price equation, Fisher’s fundamental theorem, kin selec-
479 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 480 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*
481 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–
482 2522.
- 483 Grafen, A. & Archetti, M. 2008: Natural selection of altruism in inelastic viscous
484 homogeneous populations. *Journal of Theoretical Biology* 252(4):694 – 710.
- 485 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*
486 *Theoretical Biology* 7(1):1 – 16.
- 487 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-
488 tionary genetics. *Biosocial anthropology* 53:133–55.
- 489 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and
490 harming on graphs: the return of the inclusive fitness effect. *Journal of Evolu-*
491 *tionary Biology* 20(6):2284–2295.

492 Lehmann, L. & Rousset, F. 2014: The genetical theory of social behaviour. Philo-
493 sophical Transactions of the Royal Society of London B: Biological Sciences
494 369(1642).

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

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

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

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

505 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule
506 for the evolution of cooperation on graphs and social networks. *Nature*
507 441(7092):502–505.

508 Parvinen, K. 2013: Joint evolution of altruistic cooperation and dispersal in a
509 metapopulation of small local populations. *Theoretical population biology*
510 85:12–19.

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

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

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

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

520 Sample, C. & Allen, B. 2017: The limits of weak selection and large population
521 size in evolutionary game theory. *Journal of mathematical biology* pages 1–
522 33.

523 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors
524 in finite structured population models. *The American Naturalist* 184(4):477–
525 488.

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

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

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

532 Taylor, P. D. 1992b: Inclusive fitness in a homogeneous environment. *Pro-*
533 *ceedings of the Royal Society of London. Series B: Biological Sciences*
534 249(1326):299–302.

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

537 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-
538 *bility in homogeneous structured populations. Journal of Theoretical Biology*
539 249(1):101 – 110.

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

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

544 West, S. A. & Gardner, A. 2010: Altruism, spite, and greenbeards. *Science*
545 327(5971):1341–1344.

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

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

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

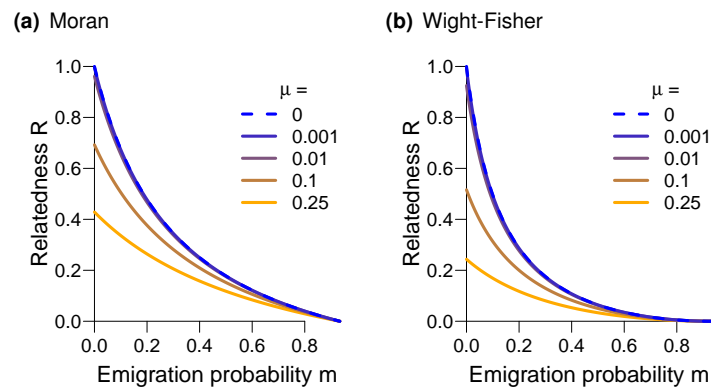


Figure 1: Within-deme relatedness of pairs of individuals, as a function of the emigration probability m , for different values of the mutation probability μ (from 0 [blue] to 0.25 [orange]), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals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

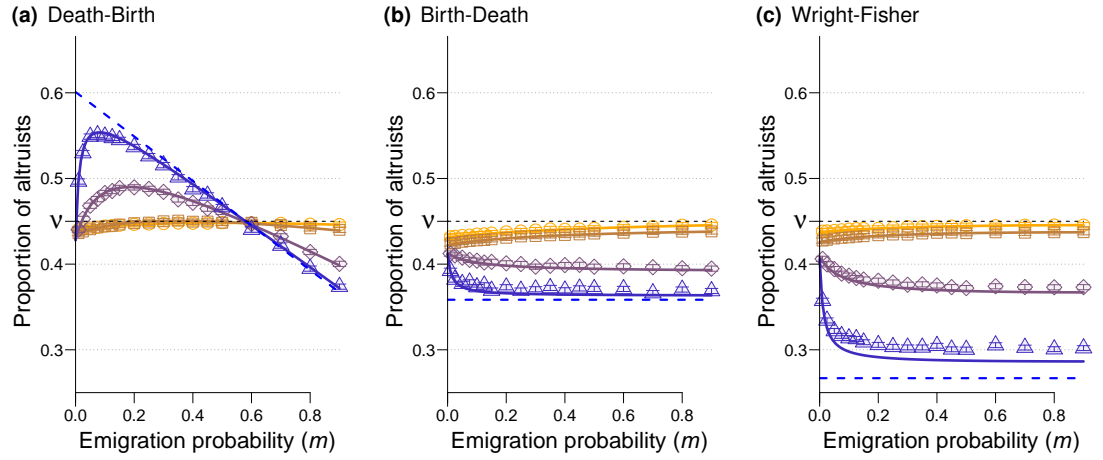


Figure 2: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different 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.

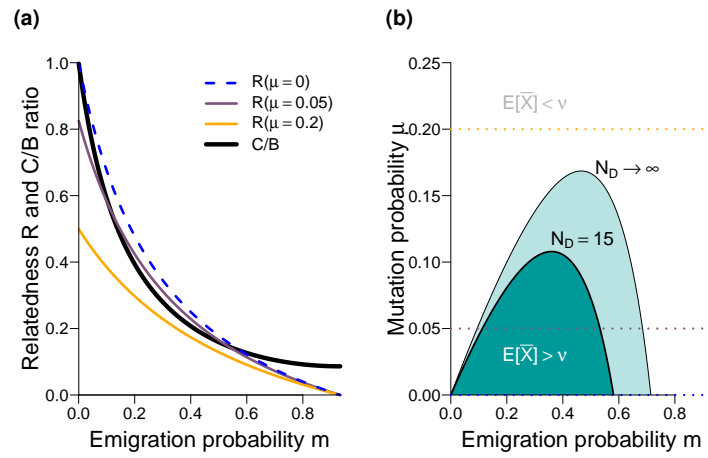


Figure 3: Understanding the effect of emigration m on whether altruism is favored in the Death-Birth life-cycle. (a) Comparison of relatedness R (thick red curve) and C/B ratio (thin curves) for different values of the mutation probability μ (same color code as previously). (b) (m, μ) combinations for 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.

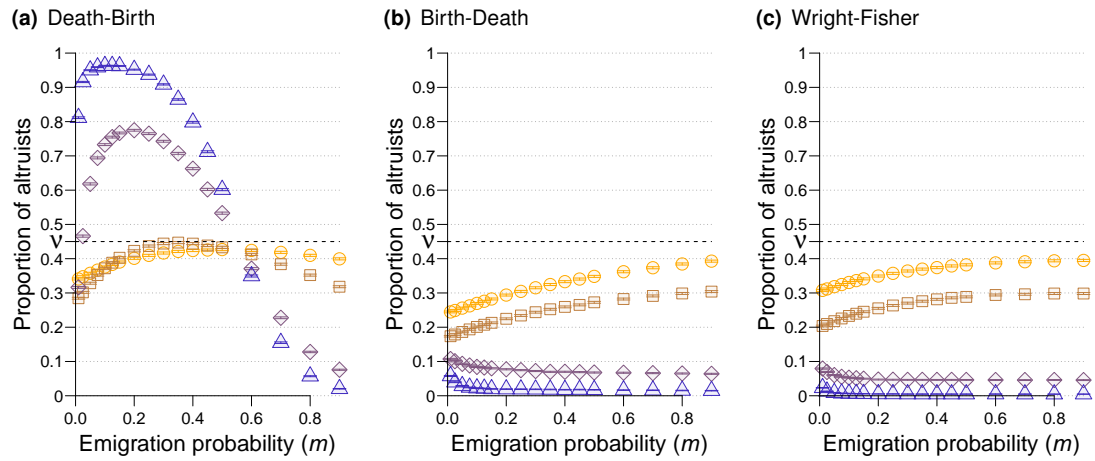


Figure S1: Equivalent of figure 2 (simulations only) but with strong selection ($\delta = 0.1$); please note the change of scale on the vertical axis. All other parameters and legends are identical to those of figure 2 (increasing mutation probabilities from blue dots to orange triangles).

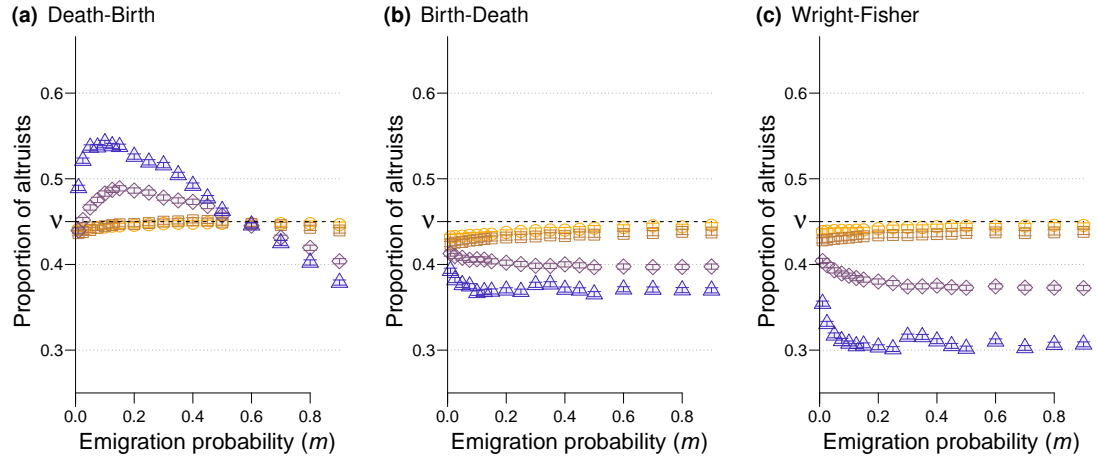


Figure S2: Equivalent of figure 2 (simulations only) but with a heterogeneous population structure: deme sizes range from 1 to 5 individuals per deme, the average deme size is 4 as in figure 2; all other parameters and legend are identical to those of figure 2.

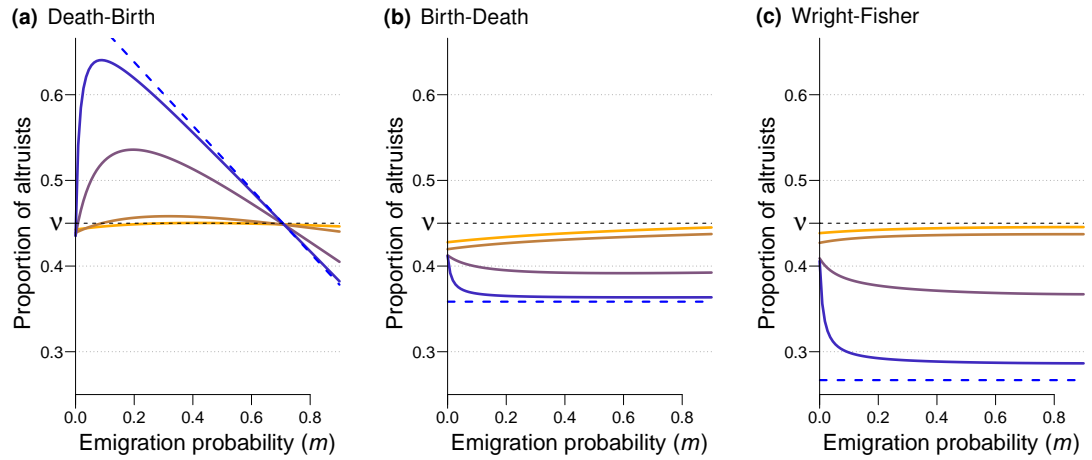


Figure S3: Equivalent of figure 2 (analysis only), with no self-replacement ($d_{ii} = d_{\text{self}} = 0$ for all sites).

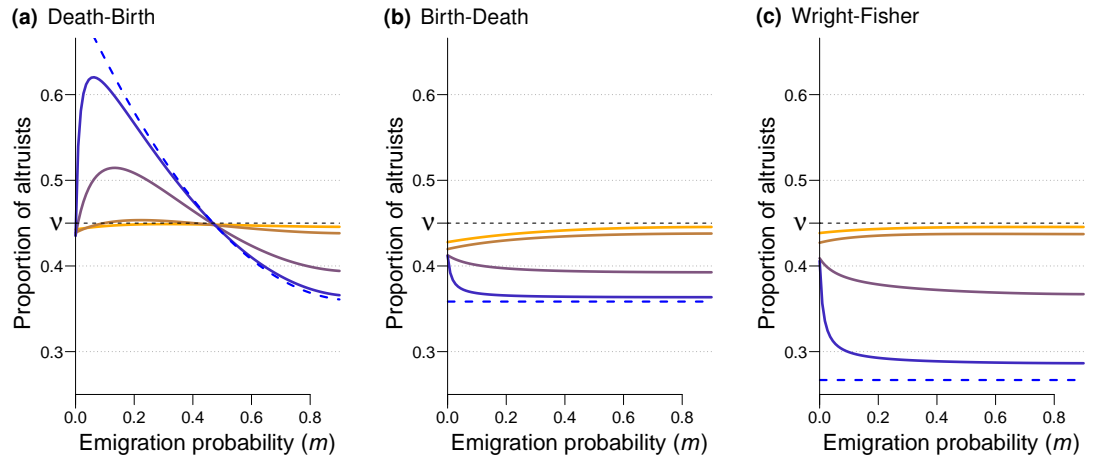


Figure S4: Equivalent of figure 2 (analysis only), with equal dispersal and interaction graphs (*i.e.*, no self-replacement [$d_{ii} = d_{\text{self}} = 0$ for all sites], and a proportion m of the interactions occurring outside of the home deme).

b	Marginal effect of a deme-mate's phenotype on the fecundity of a focal individual (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. (4))
W_i	Measure of fitness, counting offspring only when unmutated (see eq. (2))
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 S1: List of symbols. "r.v." means *random variable*.

554

Appendix

A Expected frequency of altruists

A.1 For a generic life-cycle

557 We want to compute the expected proportion of altruists in the population. We
 558 represent the state of the population at a given time t using indicator variables
 559 $X_i(t)$, $1 \leq i \leq N$, equal to 1 if the individual living at site i at time t is an altru-
 560 ist, and equal to 0 if it is a defector; these indicator variables are gathered in a
 561 N -long vector $\mathbf{X}(t)$. The set of all possible population states is $\Omega = \{0, 1\}^N$. The
 562 proportion of altruists in the population is written $\bar{X}(t) = \sum_{i=1}^N X_i(t)$. We denote
 563 by $B_{ji}(X(t), \delta)$, written B_{ji} for simplicity, the probability that the individual at
 564 site j at time $t+1$ is the newly established offspring of the individual living at site
 565 i at time t . The expected number of successful offspring produced by the indi-
 566 vidual living at site i at time t is given by $B_i = \sum_{j=1}^N B_{ji}$. We denote by $D_i(X(t), \delta)$
 567 (D_i for simplicity) the probability that the individual living at site i at time t has
 568 been replaced (*i.e.*, died) at time $t+1$. These quantities depend on the chosen
 569 life-cycle and on the state of the population; they are given in table S2 for each
 570 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 S2: 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. (1).

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

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

572 holds for all sites i . The structure of the population is also such that in the ab-
 573 sence of selection ($\delta = 0$, 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.*, of having offspring that become adults at the next time step), so that

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

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. (A.1b) and 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

$$\mathbb{E}[\bar{X}(t+1)|\mathbf{X}(t)] = \frac{1}{N} \sum_{i=1}^N [B_i(1-\mu)X_i + (1-D_i)X_i + B_i\mu\nu]. \quad (\text{A.2a})$$

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. (A.2a) ($\mathbb{E}[\bar{X}] = \sum_{\mathbf{X} \in \Omega} \bar{X} \xi(\mathbf{X}, \delta, \mu)$), we obtain, after reorganizing:

$$0 = \frac{1}{N} \sum_{\mathbf{X} \in \Omega} \left[\sum_{i=1}^N (B_i(1-\mu)X_i - D_iX_i) + \sum_{i=1}^N B_i\mu\nu \right] \xi(\mathbf{X}, \delta, \mu). \quad (\text{A.3})$$

Now, we use the assumption of weak selection ($\delta \ll 1$) and consider the first-order expansion of eq. (A.3) 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. (A.3) vanishes when we take its derivative with respect to δ , for the life-cycles that we consider.

The first order expansion of eq. (A.3) yields

$$0 = \frac{\delta}{N} \sum_{i=1}^N \left[\sum_{\mathbf{X} \in \Omega} \left(\frac{\partial B_i}{\partial \delta} (1-\mu) - \frac{\partial D_i}{\partial \delta} \right) X_i \xi(\mathbf{X}, 0, \mu) - \sum_{\mathbf{X} \in \Omega} \mu B^* X_i \frac{\partial \xi}{\partial \delta} \right] + O(\delta^2), \quad (\text{A.4})$$

where the derivatives are evaluated at $\delta = 0$, and where we used eq. (A.1b) to simplify the second term. For conciseness, we define

$$W_i = (1 - \mu)B_i + (1 - D_i), \quad (\text{A.5})$$

a measure of fitness counting offspring only when they are unmutated. With this, using the expectation notation, and denoting by $\mathbb{E}_0[\cdot]$ expectations under $\delta = 0$, we can rewrite and reorganize eq. (A.4) as

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \mathbb{E}_0 \left[\frac{\partial W_i}{\partial \delta} X_i \right] + O(\delta^2). \quad (\text{A.6})$$

Now, we use a first time the law of total probabilities, taking individual phenotypes ϕ_k are intermediate variables:

$$\begin{aligned} \frac{\partial W_i}{\partial \delta} &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} \frac{\partial \phi_k}{\partial \delta} \\ &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} X_k, \end{aligned} \quad (\text{A.7})$$

by definition of ϕ_k , and where the first derivative is evaluated for all $\phi_i = 0$. With the notation $P_{ij} = \mathbb{E}_0[X_i X_j]$, eq. (A.6) becomes

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} P_{ik} + O(\delta^2). \quad (\text{A.8})$$

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

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

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

$$\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} + (N - n) \frac{\partial W_i}{\partial \phi_{\text{out}}} = 0. \quad (\text{A.9})$$

With this, eq. (A.8) becomes

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} \frac{P_{\text{in}} - P_{\text{out}}}{P_{ii} - P_{\text{out}}} \right) (P_{ii} - P_{\text{out}}) + O(\delta^2). \quad (\text{A.10})$$

624 We can also replace the P terms by

$$\begin{aligned} P_{ij} &= Q_{ij}\nu + (1 - Q_{ij})\nu^2 \\ &= \nu^2 + \nu(1 - \nu)Q_{ij}. \end{aligned} \quad (\text{A.11})$$

625 In Appendix B.1, we will see that recursions on P_{ij} reveal that Q_{ij} can be inter-
626 preted as a probability of identity by descent, *i.e.*, the probability that the individ-
627 uals at sites i and j have a common ancestor and that no mutation has occurred
628 on either lineage since the ancestor. Eq. (A.10) becomes

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\underbrace{\frac{\partial W_i}{\partial \phi_i}}_{-C} + \underbrace{(n-1) \frac{\partial W_i}{\partial \phi_{\text{in}}}}_{B} \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) (1 - Q_{\text{out}})\nu(1 - \nu) + O(\delta^2). \quad (\text{A.12})$$

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

$$\frac{\partial W_i}{\partial \phi_k} = \sum_{\ell=1}^N \frac{\partial W_i}{\partial f_\ell} \frac{\partial f_\ell}{\partial \phi_k}. \quad (\text{A.13})$$

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

$$\left. \frac{\partial f_\ell}{\partial \phi_\ell} \right|_{\delta=0} = -c, \quad (\text{A.14a})$$

$$\left. \frac{\partial f_\ell}{\partial \phi_{\text{in}}} \right|_{\delta=0} = \frac{b}{n-1}, \quad (\text{A.14b})$$

$$\left. \frac{\partial f_\ell}{\partial \phi_{\text{out}}} \right|_{\delta=0} = 0. \quad (\text{A.14c})$$

633 Eq. (A.12) then becomes (using notation \bullet to refer to the focal individual itself,
634 and where $W = W_i$, since the derivatives are the same for all i):

$$\begin{aligned} \delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} &= \delta\nu(1 - \nu)(1 - Q_{\text{out}}) \times \\ &\quad \left(\underbrace{\left(\frac{\partial W}{\partial f_\bullet}(-c) + \frac{\partial W}{\partial f_{\text{in}}}b \right)}_{-C} + \underbrace{\left(\frac{\partial W}{\partial f_\bullet}b + (n-1) \frac{\partial W}{\partial f_{\text{in}}}(-c) + (n-2) \frac{\partial W}{\partial f_{\text{in}}}b \right)}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) + O(\delta^2). \end{aligned} \quad (\text{A.15})$$

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

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

$$\mathbb{E}[\bar{X}] = \nu + \delta \left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0} + O(\delta^2), \quad (\text{A.16})$$

638 where $\left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0}$ is obtained from eq. (A.15). We then need to replace the B_i and
 639 D_i terms by their formulas for each life-cycle (given in table S2).

640 A.2 Derivatives for the specific life-cycles

641 Using the formulas presented in table S2, and the definition of $W = W_i$ given in
 642 eq. (A.5), we obtain the following equations:

Moran Birth-Death

$$\left. \frac{\partial W^{\text{BD}}}{\partial f_{\bullet}} \right|_{\delta=0} = (1-\mu) \left(\frac{1}{N} - \frac{1}{N^2} \right) - \left(\frac{1-m}{nN} - \frac{1}{N^2} \right) = \frac{1-\mu}{N} + \frac{\mu}{N^2} - \frac{1-m}{nN}, \quad (\text{A.17a})$$

$$\left. \frac{\partial W^{\text{BD}}}{\partial f_{\text{in}}} \right|_{\delta=0} = (1-\mu) \left(-\frac{1}{N^2} \right) - \left(\frac{1-m}{nN} - \frac{1}{N^2} \right) = \frac{\mu}{N^2} - \frac{1-m}{nN}. \quad (\text{A.17b})$$

Moran Death-Birth

$$\left. \frac{\partial W^{\text{DB}}}{\partial f_{\bullet}} \right|_{\delta=0} = \frac{1-\mu}{N} \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A.18a})$$

$$\left. \frac{\partial W^{\text{DB}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -\frac{1-\mu}{N} \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A.18b})$$

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

Wright-Fisher

$$\left. \frac{\partial W^{\text{WF}}}{\partial f_{\bullet}} \right|_{\delta=0} = (1-\mu) \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A.19a})$$

$$\left. \frac{\partial W^{\text{WF}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -(1-\mu) \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A.19b})$$

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

645 Combining these equations with eq. (A.16) and eq. (A.15) given eqs. (6)–(8)
 646 in the main text.

647 B Probabilities of identity by descent

648 B.1 Expected state of pairs of sites and probabilities of identity by de- 649 scent

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

654 B.1.1 Moran model

655 In a Moran model, exactly one individual dies and one individual reproduces
656 during one time step. Given a state \mathbf{X} at time t , at time $t + 1$ both sites i and
657 $j \neq i$ are occupied by altruists, if i it was the case at time t and neither site was
658 replaced by a non-altruist (first term in eq. (B.20)), or ij if exactly one of the two
659 sites was occupied by a non-altruist at time t , but the site was replaced by an
660 altruist (second and third terms of eq. (B.20)):

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & X_i X_j \left(1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - \nu)) \right) \\ & + X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu\nu) \\ & + X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu\nu). \end{aligned} \quad (\text{B.20})$$

661 We take the expectation of this quantity, and consider that the stationary dis-
662 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

$$P_{ij} = \frac{1}{2} \left(\sum_{k=1}^N (1 - \mu) (d_{kj} P_{ki} + d_{ki} P_{kj}) \right) + \mu\nu^2 \quad (i \neq j), \quad (\text{B.21})$$

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

664 Now we substitute $P_{ij} = \nu^2 + \nu(1 - \nu)Q_{ij}$ in eq. (B.21), we obtain

$$Q_{ij} = \frac{1}{2} \sum_{k=1}^N (1 - \mu) (d_{ki} Q_{kj} + d_{kj} Q_{ki}), \quad (\text{B.22})$$

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

670 B.1.2 Wright-Fisher model

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & \sum_{k, \ell=1}^N d_{ki} d_{\ell j} \left(X_k X_\ell (1 - \mu + \mu v)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu v)(\mu v) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu v)^2 \right) \end{aligned} \quad (\text{B.23})$$

673 The first term of eq. (B.23) corresponds to both parents being altruists, and hav-
674 ing altruist offspring; the second line corresponds to exactly one parent being
675 altruist, and the third line to both parents being non-altruists (in this latter case,
676 the two offspring have to be both mutants to be altruists).

677 Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k, \ell=1}^N (P_{kl}(1 - \mu)^2) + (2 - \mu)\mu v^2. \quad (\text{B.24})$$

678 Replacing P_{ij} by $v^2 + v(1 - v)Q_{ij}$, eq. (B.24) becomes

$$Q_{ij} = \sum_{k, \ell=1}^N d_{ki} d_{\ell j} Q_{k\ell} (1 - \mu)^2. \quad (\text{B.25})$$

679 Again, Q_{ij} corresponds to a probability of identity by descent: the individuals at
680 sites i and j are identical by descent if their parents were and if neither mutated
681 $((1 - \mu)^2)$.

682 B.2 Probabilities of identity by descent in a subdivided population

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

$$Q_{ij} = \begin{cases} 1 & \text{when } i = j; \\ Q_{\text{in}} & \text{when } i \neq j \text{ and both sites are in the same deme;} \\ Q_{\text{out}} & \text{when sites } i \text{ and } j \text{ are in different demes.} \end{cases} \quad (\text{B.26})$$

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

689 Here, we will use formulas derived in Débarre (2017) for “two-dimensional
 690 population structures”. The name comes from the fact that we only need two
 691 types of transformations to go from any site to any other site in the population:
 692 permutations on the deme index, and permutations on the within-deme index.
 693 We rewrite site labels ($1 \leq i \leq N$) as (ℓ_1, ℓ_2) , where ℓ_1 is the index of the deme ($1 \leq$
 694 $\ell_1 \leq N_D$) and ℓ_2 the position of the site within the deme ($1 \leq \ell_2 \leq n$). Then, we
 695 introduce notations $\tilde{d}_{i_1 i_2}$ and $\tilde{Q}_{i_1 i_2}$, that correspond to the dispersal probability and
 696 probability of identity by descent to a site at distances i_1 and i_2 in the among-
 697 demes and within-deme dimensions (e.g., $\tilde{d}_{i_1 i_2} = d_{j_1 j_2' j_2 + i_2}$).

698 Also, in this section, we distinguish between $\tilde{d}_{\text{self}} = d_{ii}$ and d_{in} (in the main
 699 text, $d_{\text{self}} = d_{\text{in}}$).

700 B.2.1 Moran model

701 In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1 q_2}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{B.27a})$$

702 with

$$\tilde{D}_{q_1 q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1 \ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{B.27b})$$

703 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
 704 population, with $N_1 = N_D$ and $N_2 = n$:

$$\begin{aligned} \tilde{D}_{q_1 q_2} &= d_{\text{self}} + \sum_{\ell_2=1}^{N_2-1} d_{\text{in}} \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) + \sum_{\ell_1=1}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} d_{\text{out}} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2} (N_2 - 1) + (1 - \delta_{q_2}) (-1)) d_{\text{in}} + (\delta_{q_1} (N_1 - 1) + (1 - \delta_{q_1}) (-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}. \end{aligned} \quad (\text{B.28a})$$

(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

$$\tilde{D}_0 = 1, \quad (\text{B.29a})$$

$$\tilde{D}_{q_1} = 1 - m - \frac{m}{N_D - 1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \quad (\text{B.29b})$$

$$\tilde{D}_{q_1} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \quad (\text{B.29c})$$

So for \tilde{Q} , using system (B.29) in eq. (B.27a),

$$\begin{aligned} \tilde{Q}_{r_1} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu) \tilde{D}_0} + \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_2}} \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right. \\ &\quad + \sum_{q_1=1}^{N_1-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu)} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_2} N_2 - 1) \right. \\ &\quad + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (\delta_{r_1} N_1 - 1) \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_1} N_1 - 1)(\delta_{r_2} N_2 - 1) \right]. \quad (\text{B.30}) \end{aligned}$$

In particular,

$$\begin{aligned} \tilde{Q}_0 &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (n - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (N_D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (N_D - 1)(n - 1) \right] \\ &= 1. \quad (\text{B.31a}) \end{aligned}$$

We find λ'_M using the eq. (B.31a). Going back to eq. (B.30), when $r_1 = 0$, the two individuals are in the same deme. They are different when $r_2 \neq 0$, and so:

$$\begin{aligned} Q_{\text{in}} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (D - 1)(-1) \right]. \quad (\text{B.31b}) \end{aligned}$$

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

$$Q_{\text{out}} = \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (-1) + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} \right]. \quad (\text{B.31c})$$

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

$$Q_{\text{in}}^M = \frac{(1 - \mu)(m + \mu(N_D(1 - m) - 1))}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}, \quad (\text{B.32a})$$

$$Q_{\text{out}}^M = \frac{(1 - \mu)m}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}. \quad (\text{B.32b})$$

715 The probability that two different deme-mates are identical by descent, Q_{in}^M , de-
 716 creases monotonically with the emigration probability m , while Q_{out}^M monoton-
 717 ically increases with m (see figure S5(a)).

718 When the mutation probability μ is vanishingly small ($\mu \rightarrow 0$), both Q_{in}^M and
 719 Q_{out}^M are equal to 1: in the absence of mutation indeed, the population ends up
 720 fixed for one of the two types, and all individuals are identical by descent. Note
 721 that we obtain a different result if we first assumed that the size of the popu-
 722 lation is infinite ($N_D \rightarrow \infty$), because the order of limits matters; for instance,
 723 $\lim_{d \rightarrow \infty} Q_{\text{out}}^M = 0$.

724 B.2.2 Wright-Fisher

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

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right), \quad (\text{B.33})$$

726 with $\tilde{\mathcal{D}}$ given in eq. (B.27b). In a subdivided population, with $N_1 = N_D$ and $N_2 =$
 727 n , this becomes

$$\begin{aligned}
 \tilde{Q}_{r_1} &= \frac{1}{N} \left[\frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_0)^2} + \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_2})^2} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right. \\
 &\quad + \sum_{q_1=1}^{N_1-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \\
 &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \right. \\
 &\quad + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_1} N_1 - 1) (\delta_{q_2} N_2 - 1) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \delta_{q_1} N_1 \right. \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \right]. \tag{B.34}
 \end{aligned}$$

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

$$1 = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (N_2 - 1) N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \tag{B.35a}$$

729 Then from eq. (B.34) we deduce

$$Q_{\text{in}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \tag{B.35b}$$

730 and

$$Q_{\text{out}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} \right]. \tag{B.35c}$$

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

$$Q_{\text{in}}^{\text{WF}} = \frac{-N_D + M_1 + M_2}{(n - 1)N_D + M_1 + M_2}, \tag{B.36a}$$

$$Q_{\text{out}}^{\text{WF}} = \frac{-\frac{1}{N_D-1} M_1 + M_2}{(n - 1)N_D + M_1 + M_2}, \tag{B.36b}$$

732 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 S5(a) and S5(b)).

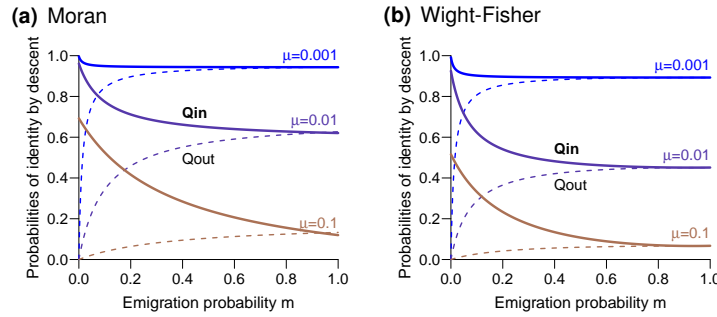


Figure S5: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in} , full curves) and two individuals in different demes (Q_{out} , dashed curves), as a function of the emigration probability m , for different values of the mutation probability μ (0.001, 0.01, 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.