

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.

Word Count:	Abstract	172
	Impact summary	282 words
	Total	4234 words

Check ω removes entirely + explain delta in table

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. At the same time
5 however, it increases competition among related individuals. The evolution of
6 altruism depends on the balance between these opposite effects. This balance
7 is already known to be affected by details of the life-cycle; we show here that
8 it further depends on the fidelity of strategy transmission from parents to their
9 offspring. We consider different life-cycles (Wright-Fisher, with synchronous
10 non-overlapping generations, Moran Death-Birth and Moran Birth-Death, both
11 with exactly one individual dying and reproducing at each time step) and we
12 identify thresholds of parent-offspring strategy transmission inaccuracy, above
13 which the effect of population viscosity on the frequency of altruists maintained
14 in the population qualitatively changes. Analytical predictions are first obtained
15 analytically under weak selection and equal deme sizes, then confirmed with

16 stochastic simulations relaxing these assumptions. This result challenges the
17 notion that the evolution of altruism requires limited dispersal.

18 **Impact Summary**

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

44 1 Introduction

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

53 Yet, living next to your kin also implies competing against them (West et al.,
54 2002). The evolution of social traits hence depends on the balance between the
55 positive effects of interactions with related individuals and the detrimental con-
56 sequences of kin competition. Under specific conditions, the two effects can
57 even compensate each other, thereby annihilating the impact of population vis-
58 cosity on the evolution of altruism. First identified with computer simulations
59 (Wilson et al., 1992), this cancellation result was analyzed by Taylor (1992a) in
60 a model with synchronous generations (*i.e.*, Wright-Fisher model) and a sub-
61 divided population of constant, infinite size. The cancellation result was later
62 extended to heterogeneous populations (Rodrigues & Gardner, 2012, with syn-
63 chronous generations and infinite population size), and other life-cycles, with
64 generic regular population structures (Taylor et al., 2011, with synchronous gen-
65 erations but also with continuous generations and Birth-Death updating). How-
66 ever, small changes in the model's assumptions, such as overlapping generations
67 (Taylor & Irwin, 2000) or the presence of empty sites (Alizon & Taylor, 2008) can
68 tip the balance back in the favor of altruism. This high dependence on life-
69 cycle specificities highlights the difficulty of making general statements about
70 the role of spatial structure on the evolution of altruism. In this study, we will

71 consider three different life-cycles: Wright-Fisher, where the whole population
72 is renewed at each time step, and two Moran life-cycles (Birth-Death and Death-
73 Birth), where a single individual dies and is replaced at each time step.

74 A large number of studies on the evolution of social behavior consider simple
75 population structures (typically, homogeneous populations *sensu* Taylor et al.
76 (2007a)) and often also infinite population sizes (but see Allen et al., 2017, for
77 results on any structure). These studies also make use of weak selection ap-
78 proximations, and commonly assume rare (*e.g.*, Leturque & Rousset, 2002; Tay-
79 lor et al., 2007b; Tarnita & Taylor, 2014) or absent mutation (for models assum-
80 ing infinite population sizes, or models concentrating on fixation probabilities
81 Lehmann & Rousset, 2014; Van Cleve, 2015, for recent reviews). Often, these sim-
82 plifying assumptions are a necessary step towards obtaining explicit analytical
83 results. Although artificial, simple population structures (*e.g.*, regular graphs, or
84 subdivided populations with demes of equal sizes) help reduce the dimension-
85 ality of the system under study, in particular when the structure of the popula-
86 tion displays symmetries such that all sites behave the same way in expectation.
87 Weak selection approximations are crucial for disentangling spatial moments
88 (Lion, 2016), that is, changes in global *vs.* local frequencies (though they can in
89 some cases be relaxed, as in Mullon & Lehmann, 2014). Mutation, however, is
90 usually ignored by classical models of inclusive fitness because these models as-
91 sume infinite population sizes, so that there is no need to add mechanisms that
92 restore genetic diversity (Tarnita & Taylor, 2014). In populations of finite size,
93 this diversifying effect can be obtained thanks to mutation.

94 When strategy transmission is purely genetic, it makes sense to assume that
95 mutation is relatively weak. A social strategy can however also be culturally
96 transmitted from parent to offspring, in which case “rebellion” (as in Frank’s Re-
97 bellious Child Model (Frank, 1997)) does not have to be rare. Imperfect strategy
98 transmission can alter evolutionary dynamics, in particular in spatially struc-

99 tured populations (see *e.g.*, Allen et al., 2012; Débarre, 2017, for graph-structured
 100 populations). Here, we want to explore the consequences of imperfect strategy
 101 transmission from parents to their offspring on the evolution of altruistic behav-
 102 ior in subdivided populations. For the sake of concision, we use the word “mu-
 103 tation” throughout the paper, keeping in mind that strategy transmission does
 104 not have to be genetic.

105 For each of the three life-cycles that we consider, we compute the expected
 106 (*i.e.*, long-term) frequency of altruists maintained in a subdivided population,
 107 and investigate how it is affected by mutation and emigration. We find that, con-
 108 trary to what happens with perfect strategy transmission, higher emigration can
 109 increase the expected frequency of altruists in the population.

110 2 Model and methods

111 2.1 Assumptions

112 We consider a population of size N , subdivided into N_D demes, each hosting
 113 exactly n individuals (*i.e.*, each deme contains n sites, each of which is occupied
 114 by exactly one individual; we have $nN_D = N$). Each site has a unique label i ,
 115 $1 \leq i \leq N$. There are two types of individuals in the population, altruists and
 116 defectors. The type of the individual living at site i ($1 \leq i \leq N$) is given by an
 117 indicator variable X_i , equal to 1 if the individual is an altruist, and to 0 if it is a
 118 defector. The state of the entire population is given by a N -long vector \mathbf{X} . For a
 119 given population state \mathbf{X} , the proportion of altruists is $\bar{X} = \sum_{i=1}^N X_i$. All symbols
 120 are summarized in table S1.

121 Reproduction is asexual. Parents transmit their strategy to their offspring
 122 with probability $1 - \mu$; this transmission can be genetic or cultural (vertical cul-
 123 tural transmission), but for simplicity, we refer to the parameter μ as a mutation
 124 probability. With probability μ , offspring do not inherit their strategy from their

parent but instead get one randomly: with probability ν , they become altruists, with probability $1 - \nu$ they become defectors. We call the parameter ν the mutation bias.

An individual of type X_k expresses a social phenotype $\phi_k = \delta X_k$, where δ is assumed to be small ($\delta \ll 1$). Social interactions take place within each deme, benefits are shared with the $n - 1$ other deme-mates. We assume that social interactions affect individual fecundity; f_k denotes the fecundity of the individual at site k . We denote by b the marginal effect of a deme-mate's phenotype on the fecundity of a focal individual, and by $-c$ the marginal effect of a focal individual's phenotype on its own fecundity ($c \leq b$).

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

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

with $0 < m < 1 - \frac{1}{N_D}$ (the upper bound implies $d_{\text{in}} > d_{\text{out}}$).

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

$$W_i := (1 - \mu) B_i + 1 - D_i, \quad (2) \quad \{\text{eq: def W}\}$$

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

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

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

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

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

161 2.2 Methods

162 2.2.1 Analytical part

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

170 The formula involves quantities that can be identified as neutral probabili-
171 ties of identity by descent Q_{ij} , *i.e.*, the probability that individuals living at site i
172 and j share a common ancestor and that no mutation occurred on either lineage

173 since that ancestor, in a model with no selection ($\omega = 0$; this is the “mutation def-
174 inition” of identity by descent (Rousset & Billiard, 2000).)

175 These neutral probabilities of identity by descent depend on the chosen life-
176 cycle, and are also computed by taking the long-term expectation of conditional
177 expectations after one time step (see Appendix B.1 and B.2).

178 All the results obtained analytically were checked numerically using specific
179 population structures (see supplementary Mathematica file (Wolfram Research,
180 Inc., 2017).)

181 2.2.2 Stochastic simulations

182 We also ran stochastic simulations (coded in C). The simulations were run for 10^8
183 generations (one generation is one time step for the Wright-Fisher life-cycle, and
184 N time steps for the Moran life-cycles). For each set of parameters and life-cycle,
185 using R (R Core Team, 2015), we estimated the long-term frequency of altruists
186 by sampling the population every 10^3 generations and computing the average
187 frequency of altruists.

188 All scripts are available at

189 <https://github.com/flodebarre/SocEvolSubdivPop/tree/master/Programs>

190 3 Results

191 3.1 Expected frequencies of altruists for each life-cycle

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

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu B^*} v(1-v)(1-Q_{\text{out}}) \times \left[\underbrace{\frac{\partial W}{\partial f_{\bullet}}(-c) + \frac{\partial W}{\partial f_{\text{in}}}b}_{-C} + \underbrace{\left(\frac{\partial W}{\partial f_{\bullet}}b + (n-1)\frac{\partial W}{\partial f_{\text{in}}}(-c) + (n-2)\frac{\partial W}{\partial f_{\text{in}}}b \right)}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right], \quad (3) \quad \{\text{eq:EXapprox}\}$$

194 with W as defined in eq. (2). (Calculations leading to eq. (3) are presented in Ap-
195 pendix A.)

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

206 We now present the values of these different terms for the three life-cycles
207 under study. The equations are obtained by replacing the derivatives of W by
208 their specific formulas for each life-cycle, which are presented in Appendix A.2.
209 The formulas of probabilities of identity by descent are more complicated, so
210 not substituted, but they can be found in eq. (B.32) and eq. (B.36). The addi-
211 tional superscripts correspond to the life-cycles (WF: Wright-Fisher, M: Moran
212 (either), DB: Moran Death-Birth, BD: Moran Birth-Death.) Finally, for each life-
213 cycle, we further decompose the $-\mathcal{C}$ and \mathcal{B} terms into primary (subscript P) and
214 secondary (subscript S) effects TODO CITE Gardner. Primary effects correspond
215 to unmediated, direct consequences of interactions (they are included in $\frac{\partial W}{\partial f_i}$),
216 while secondary effects correspond to consequences of interactions mediated
217 by other individuals, competition. They correspond to what was called “direct”
218 and “indirect” effects in (Débarre et al., 2014; Débarre, 2017), but the terminol-
219 ogy was updated for more consistency with other authors. All equations are of

220 the form

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}) \left[\underbrace{\begin{pmatrix} -C_P \\ -C_S \end{pmatrix}}_{-C} + \underbrace{\begin{pmatrix} B_P \\ +B_S \end{pmatrix}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right]. \quad (4) \quad \{\text{eq:EXBD}\}$$

221 **Moran Birth-Death** With the Birth-Death life-cycle, eq. (3) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\begin{pmatrix} (1-\mu)(-c) \\ +(b-c)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{-C^{\text{BD}}} + \underbrace{\begin{pmatrix} (1-\mu)b \\ +(b-c)(n-1)\left(\frac{\mu}{N} - \frac{1-m}{n}\right) \end{pmatrix}}_{B^{\text{BD}}} \underbrace{\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M}}_{R^M} \right], \quad (5) \quad \{\text{eq:EXBD}\}$$

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

227 **Moran Death-Birth** With the Death-Birth life-cycle, eq. (3) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\begin{pmatrix} (1-\mu)(-c) \\ -(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right) \end{pmatrix}}_{-C^{\text{DB}}} + \underbrace{\begin{pmatrix} (1-\mu)b \\ -(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right) \end{pmatrix}}_{B^{\text{DB}}} \underbrace{\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M}}_{R^M} \right], \quad (6) \quad \{\text{eq:EXDB}\}$$

228 With this life-cycle, Death occurs first, and the probability of dying is indepen-
 229 dent from the state of the population (since we assume that social interactions
 230 affect fecundity TODO cite D and Nakam). We can therefore factor a $1 - \mu$ term
 231 in all terms. The primary effects (first lines in the parentheses) remain the same

as with the Birth-Death life-cycle. However, the Death-Birth life-cycle leads to different secondary effects compared to the Birth-Death life-cycle: competition occurs at a different scale (Grafen & Archetti, 2008). Finally, with this life-cycle as we defined it, the probabilities of identity by descent are the same as with the Birth-Death model.

Wright-Fisher

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

The only – but important – different between eq. (7) and eq. (6) is the value of the probabilities of identity by descent, because the number of individuals that are updated at each time step differs.

3.2 Primary effects

Let us first focus on primary effects, shown on the first lines in the parentheses in eq. (5)–eq. (7). They are the same for the three life-cycles that we consider, and they do not depend on the emigration probability m . At the same time, the relatedness terms R^{M} and R^{WF} decrease with m (keeping $m < \frac{N_D-1}{N_D}$.) Consequently, ignoring secondary effects, we would conclude that even with imperfect strategy transmission ($\mu > 0$), increasing the emigration probability m decreases the expected frequency of altruists in the population (as our intuition may suggest already.) However, secondary effects play a role as well.

249 3.3 Changes with m

250 For the three life-cycles, the secondary effects are negative (with $0 < m < \frac{N_D-1}{N_D}$)
 251 and they increase with the emigration probability m . In other words, these detri-
 252 mental secondary effects weaken when the emigration probability increases. This
 253 goes against the reduction of relatedness R as m increases. Hence, we need to
 254 consider the entire equations to know the overall effect of the emigration proba-
 255 bility m on the expected frequency of altruists $\mathbb{E}[\bar{X}]$.

256 How the secondary effects change with m and combining both we cannot
 257 say anything.

258 Figures with both.

259 Identify critical values.

260 3.3.1 Direct effects

261 Direct (/primary) effects are similar for the three life-cycles; the only difference
 262 is the value of probabilities of identity by descent Q (as seen in the previous sec-
 263 tion, they differ between Moran and Wright-Fisher life-cycles):

{eq:directeffects}

$$\beta_D^{BD} = \beta_D^{DB} = (1 - \mu) Q_{in}^M, \quad (8a) \quad \text{{eq:bBDD}}$$

$$\beta_D^{WF} = (1 - \mu) Q_{in}^{WF}; \quad (8b) \quad \text{{eq:bWFD}}$$

$$\gamma_D^{BD} = \gamma_D^{DB} = \gamma_D^{WF} = 1 - \mu. \quad (8c) \quad \text{{eq:cBDD}}$$

264 For both benefits and costs, direct effects only count when there is no mutation
 265 (hence the $(1 - \mu)$ factors). Direct effects of benefits b (eq. (8a) and eq. (8b)) only
 266 count if the interaction takes place with an individual who is identical by de-
 267 scent. With the population structure that we consider, social interactions only
 268 occur within demes, so only Q_{in} is present in eq. (8a) and eq. (8b). On the other
 269 hand, the direct effect of the fecundity cost c (eq. (8c)) does not depend on the
 270 type of interactant, since the same cost c is paid by altruists irrespective of the

271 interactant's identity.

272 As seen in the previous section, Q_{in}^{M} and $Q_{\text{in}}^{\text{WF}}$ decrease with the emigration
273 probability m (actually only until $m = \frac{d-1}{d}$ for the latter). Consequently, the mag-
274 nitude of the direct (beneficial) effects of benefits b provided by altruists (β_{D})
275 decreases when the emigration probability m increases, while the direct (detrimental)
276 effects (γ_{D}) due to the direct cost of altruism c are constant. As a result,
277 if we only considered direct effects, we would conclude that more emigration m
278 is detrimental to the evolution of altruistic behaviour. However, there are also
279 indirect effects at play.

280 3.3.2 Indirect effects

281 Indirect (/secondary) effects are collateral effects on other individuals; they de-
282 pend on the type of life-cycle, and always involve individuals who are identical
283 by descent.

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

$$\begin{aligned} \beta_{\text{I}}^{\text{BD}} &= (1 - m) \left(\frac{n-1}{n} Q_{\text{in}}^{\text{M}} + \frac{1}{n} \right) + m Q_{\text{out}}^{\text{M}} - \mu \frac{1 + (n-1) Q_{\text{in}}^{\text{M}} + n(d-1) Q_{\text{out}}^{\text{M}}}{nd} \\ &= \gamma_{\text{I}}^{\text{BD}}. \end{aligned} \quad (9a) \quad \{\text{eq: bBDI}\}$$

291 (Calculation details are presented in Appendix A.)

292 The formulas are the same for the indirect effects associated to b and to c ; in
293 other words, the balance between the two indirect effects remains the same when

the emigration probability changes. The term $(\frac{n-1}{n}Q_{\text{in}}^M + \frac{1}{n})$, which will appear again later, corresponds to the probability that two individuals sampled with replacement from the same deme are identical by descent. Indirect effects are indeed also felt by the focal individual itself (*e.g.*, increasing the fecundity of another individual implies decreasing one's own relative fecundity).

Replacing Q_{in} and Q_{out} by their formula for the Moran life-cycle (eq. (B.32)), we conclude that $\beta_I^{\text{BD}} = \gamma_I^{\text{BD}}$ are decreasing functions of the emigration probability m (calculations in the supplementary Mathematica file).

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

$$\begin{aligned} \beta_I^{\text{DB}} &= (1 - \mu) \left[\left(\frac{1}{n} + \frac{(n-1)Q_{\text{in}}^M}{n} \right) \left((1-m)^2 + \frac{m^2}{(d-1)} \right) \right. \\ &\quad \left. + Q_{\text{out}}^M \left(2m(1-m) + (d-2)\frac{m^2}{(d-1)} \right) \right] \\ &= \gamma_I^{\text{DB}} \end{aligned} \tag{9b} \quad \{\text{eq: bDBI}\}$$

The brackets in eq. (9b) contain a sum of two terms. The first term corresponds to two individuals from the same deme (with replacement) whose offspring either do not emigrate, or emigrate together to the same deme. The second term corresponds to individuals initially from different demes who end up in the same deme (either one of their home demes, or a third deme).

Here again, $\beta_I = \gamma_I$, so the balance between indirect benefits and indirect costs does not change when the emigration probability m increases.

Replacing Q_{in} and Q_{out} by their formulas given in eq. (B.32), we can conclude that $\beta_I^{\text{DB}} = \gamma_I^{\text{DB}}$ first decreases with the emigration probability m , and increases again after a threshold value m'_c , which is smaller than $m_c^{\text{WF}} = (d-1)/d$ (calcu-

317 lation details are presented in the supplementary Mathematica file).

318 **Wright-Fisher** With this life-cycle, generations are synchronous and all indi-
 319 viduals again all have the same survival probability (now equal to 0 at all sites).
 320 As a result, the formulas for β_1^{WF} and γ_1^{WF} are the same as β_1^{DB} and γ_1^{WF} , except
 321 that instead of Q_{in}^{M} and $Q_{\text{out}}^{\text{M}}$, we need to use $Q_{\text{in}}^{\text{WF}}$ and $Q_{\text{out}}^{\text{WF}}$ (given in eq. (B.36)).
 322 Once this is done, we see that $\beta_1^{\text{WF}} = \gamma_1^{\text{WF}}$ first decreases with the emigration
 323 probability m , and increases again after the threshold value $m_c^{\text{WF}} = (d-1)/d$.
 324 This emigration threshold was identified above as the emigration probability
 325 such that offspring have an equal chance of landing in their natal deme or in
 326 any other deme, *i.e.*, $d_{\text{in}} = d_{\text{out}}$ (calculation details are presented in the supple-
 327 mentary Mathematica file.)

328 3.4 Identifying threshold values of the mutation probability μ

329 In the previous section, we investigated the impact of changes in the emigration
 330 probability m on each of the terms that make up the expected frequency of altru-
 331 ists $\mathbb{E}[\bar{X}]$. Now we need to combine these different terms to focus on the quantity
 332 we are eventually interested in, $\mathbb{E}[\bar{X}]$. The rather lengthy formulas that we ob-
 333 tain are relegated to the Appendix and supplementary Mathematica file, and we
 334 concentrate here on the results.

335 3.4.1 Moran Birth-Death

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

$$\mu_c^{\text{BD}} = 1 - \frac{b - c + \sqrt{(b - c)(4b(nd)^2 + b - c)}}{2bnd} \quad (10) \quad \{\text{eq:mucBD}\}$$

341 This result is illustrated in figure 1(b); with the parameters of the figure, $\mu_c^{\text{BD}} \approx$
 342 0.026.

343 3.4.2 Moran Death-Birth

344 The relationship between $\mathbb{E}[\bar{X}]$ and m is a bit more complicated for this life-
 345 cycle. For simplicity, we concentrate on what happens starting from low emi-
 346 gration probabilities (*i.e.*, the sign of the slope of $\mathbb{E}[\bar{X}]$ as a function of m when
 347 $m \rightarrow 0$). If the benefits b provided by altruists are relatively low ($b < c(n+1)$),
 348 $\mathbb{E}[\bar{X}]$ initially increases with m provided the mutation probability μ is greater
 349 than a threshold value μ_c^{DB} given in eq. (11) below; otherwise, when the benefits
 350 are high enough, $\mathbb{E}[\bar{X}]$ initially increases with m for any value of μ . Combining
 351 these results, we write

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

352 In figure 1(a), the parameters are such that $\mu_c^{\text{DB}} = 0$.

353 The expected frequency of altruists $\mathbb{E}[\bar{X}]$ then reaches a maximum at an em-
 354 igration probability m_c^{DB} (whose complicated equation is given in the supple-
 355 mentary Mathematica file), as can be seen in figure 1(a). When the mutation
 356 probability gets close to 0 ($\mu \rightarrow 0$), m_c^{DB} also gets close to 0,

357 3.4.3 Wright-Fisher

358 The expected frequency of altruists in the population reaches an extremum when
 359 $m = m_c^{\text{WF}} = \frac{d-1}{d}$. This extremum is a maximum when the mutation probability
 360 is higher than a threshold value μ_c^{WF} given by

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

361 and it is a minimum otherwise. With the parameters of figure 1(c), $\mu_c^{\text{WF}} = 0.034$.

362 3.5 Relaxing key assumptions

363 To derive our analytical results, we had to make a number of simplifying assump-
364 tions, such as the fact that selection is weak ($\omega \ll 1$), and the fact that the struc-
365 ture of the population is regular (all demes have the same size n). We explored
366 with numerical simulations the effect of relaxing these key assumptions. When
367 selection is strong, the patterns that we identified not only still hold but are even
368 more marked, as shown on figure S1.

369 To relax the assumption of equal deme sizes, we randomly drew deme sizes
370 at the beginning of simulations, with sizes ranging from 2 to 6 individuals and
371 on average $\bar{n} = 4$ individuals per deme as previously. As shown in figure S2, the
372 patterns initially obtained with a homogeneous population structure are robust
373 when the structure is heterogeneous.

374 For the Moran model, it may seem odd that an offspring can replace its own
375 parent (which can occur since $d_{ii} \neq 0$). Figure S3, plotted with dispersal prob-
376 abilities preventing immediate replacement of one's own parent (for all sites i ,
377 $d_{ii} = d_{\text{self}} = 0$; $d_{\text{in}} = (1 - m)/(n - 1)$ for two different sites in the same deme, d_{out}
378 remaining unchanged), confirms that this does affect our conclusions.

379 4 Discussion

380 The expected frequency of altruists in a subdivided population can in- 381 crease with the probability of emigration

382 Assuming that the transmission of a social strategy (being an altruist or a defec-
383 tor) from a parent to its offspring could be imperfect, we found that the expected
384 frequency of altruists maintained in a population could increase with the prob-
385 ability 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 1(b)–(c)) or favorable (Taylor et al., 2007a, and dashed lines in figure 1(a)) to the evolution of altruistic behavior.

Quantitative vs. qualitative measures

We used a quantitative measure, the expected frequency of altruists in the population ($\mathbb{E}[\bar{X}]$), to explore how non-zero mutation probabilities altered the impact of population viscosity. Often however, 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 to saying that altruism is favored whenever $\mathbb{E}[\bar{X}] > v$ (v is plotted as a horizontal dashed line in figure 1). Some of our conclusions change if we switch to this qualitative measure of evolutionary success: Under the Moran Birth-Death and Wright-Fisher life-cycles, population viscosity does not promote the evolution of altruism – actually, these two life-cycles can not ever promote altruistic behavior for any regular population structure (Taylor et al., 2011), whichever the probability of mutation (Débarre, 2017). However, under a Moran Death-Birth life-cycle, altruism can be favored only at intermediate emigration probabilities (figure 1(a)): increased emigration can still favor the evolution of altruism under this qualitative criterion.

The result is due to indirect (/secondary) effects

To explain how the frequency of altruists can increase with the emigration probability m , let us go back to the decomposition of the expected frequency of al-

412 altruists in the population $\mathbb{E}[\bar{X}]$ into different terms (eq. (3)). For all the life-cycles
 413 that we consider, the direct effect of helping others (β_D) decreases with emigra-
 414 tion m , while the direct effect of the cost of helping (γ_D) does not change with
 415 m . If we (erroneously) considered only direct effects, we would conclude that
 416 the expected proportion of altruists decreases with the emigration probability
 417 m , because an increase in m reduces the probability that two interactants (two
 418 deme-mates in this model) are identical by descent. But this explanation ignores
 419 indirect, competitive, effects. In the three life-cycles that we considered, $\beta_I = \gamma_I$,
 420 so the overall indirect effects are given by $-(b - c)\beta_I$. Hence, any increase of
 421 $\mathbb{E}[\bar{X}]$ with m is driven by β_I . Indirect effects correspond to competition: helping
 422 another individual indirectly harms others – even the individual who is provid-
 423 ing help is indirectly harmed. This competition can be diluted by increasing the
 424 emigration probability m . The overall effect of m on the expected frequency of
 425 altruists depends on the balance between direct and indirect effects. This bal-
 426 ance depends on the fidelity of parent-offspring transmission (μ), in particular
 427 because probabilities of identity by descent depend on μ .

428 Indirect (secondary) effects are less straightforward to understand than di-
 429 rect (/primary) effects, yet they play a crucial role for social evolution in spatially
 430 structured populations. Competition among relatives is for instance behind Tay-
 431 lor (1992b)'s cancellation result; similarly, the qualitative differences between
 432 the Moran Birth-Death and Moran Death-Birth life-cycles is explained by the dif-
 433 ferent scales of competition that the two life-cycle produce (Grafen & Archetti,
 434 2008; Débarre et al., 2014). Secondary effects are also behind the evolution of
 435 social behaviors such as spite (West & Gardner, 2010).

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

437 Our results were derived under the assumption of weak selection, assuming that
 438 the phenotypic difference between altruists and defectors is small ($\delta \ll 0$). We

439 considered any fidelity of transmission (any μ between 0 and 1) and population
440 size. However, most models considering subdivided populations assume nearly
441 perfect strategy transmission ($\mu \rightarrow 0$) and infinite population sizes (number of
442 demes $N_D \rightarrow \infty$). The order in which these limits are taken matters, *i.e.*, one
443 needs to specify how small μ , but also ω , are compared to the inverse size of
444 the population. This remark complements findings by Sample & Allen (2017),
445 who highlighted the quantitative differences between different orders of weak
446 selection and large population limits.

447 **Imperfect transmission and Rebellious Children**

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

454 Mutation was also previously included in models investigating the mainte-
455 nance of cooperative microorganisms in the presence of cheaters (Brockhurst
456 et al., 2007; Frank, 2010). In both of these models however, only loss-of-function
457 mutation was considered, which corresponds to setting the mutation bias at
458 $\nu = 0$ in our model. This means that the all-cheaters state is absorbing; no matter
459 how favored cooperators may otherwise be, in the long run, a finite population
460 will only consist of cheaters.

461 **Interactions and dispersal graphs**

462 Compared to graphs classically used in evolutionary graph theory (*e.g.*, regular
463 random graphs, grids), the island model is particular because the interaction
464 graph and the dispersal graph are different: interactions take place only within

465 demes ($e_{\text{out}} = 0$), while offspring can disperse out of their natal deme ($d_{\text{out}} > 0$).
466 One may wonder whether our result depends on this difference between the two
467 graphs. Figure S4 shows that the result still holds when the dispersal and inter-
468 action graphs are the same. In this figure indeed, we let a proportion m (equal
469 to the dispersal probability) of interactions occur outside of the deme where the
470 individuals live, and set d_{self} , the probability of self replacement, equal to 0, so
471 that the dispersal and interactions graphs are the same.

472 **Coevolution of dispersal and social behavior**

473 This work also raises the question of what would happen if dispersal (*e.g.*, the
474 emigration probability m) could evolve as well. Recent work on the topic has
475 shown that under some conditions disruptive selection could take place, lead-
476 ing to a polymorphism between sessile altruists and mobile defectors (Parvinen,
477 2013; Mullan et al., 2017). The assumptions of these studies however differ from
478 ours in important ways, in that they consider continuous traits and use an adap-
479 tive dynamics framework, where, notably, mutations are assumed to be very
480 rare. It remains to be investigated how non-rare and potentially large mutations
481 would affect their result.

482 References

- 483 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*
484 *tion* 62(6):1335–1344.
- 485 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,
486 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*
487 544(7649):227–230.
- 488 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects
489 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.
490 *Evolution of Cooperation.*
- 491 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-
492 mediate disturbance. *Current Biology* 17(9):761–765.
- 493 Cockerham, C. C. & Weir, B. 1987: Correlations, descent measures: drift with
494 migration and mutation. *Proceedings of the National Academy of Sciences*
495 84(23):8512–8514.
- 496 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution
497 of social behavior in structured populations. *Journal of Theoretical Biology*
498 420:26 – 35.
- 499 Débarre, F.; Hauert, C. & Doebeli, M. 2014: Social evolution in structured popu-
500 lations. *Nature Communications* 5.
- 501 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-
502 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*
503 276(1654):13–19.
- 504 Frank, S. A. 1997: The Price equation, Fisher’s fundamental theorem, kin selec-
505 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 506 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*
507 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–
508 2522.
- 509 Grafen, A. & Archetti, M. 2008: Natural selection of altruism in inelastic viscous
510 homogeneous populations. *Journal of Theoretical Biology* 252(4):694 – 710.
- 511 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*
512 *Theoretical Biology* 7(1):1 – 16.
- 513 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-
514 tionary genetics. *Biosocial anthropology* 53:133–55.
- 515 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and
516 harming on graphs: the return of the inclusive fitness effect. *Journal of Evolu-*
517 *tionary Biology* 20(6):2284–2295.

518 Lehmann, L. & Rousset, F. 2014: The genetical theory of social behaviour. Philo-
519 sophical Transactions of the Royal Society of London B: Biological Sciences
520 369(1642).

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

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

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

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

531 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule
532 for the evolution of cooperation on graphs and social networks. *Nature*
533 441(7092):502–505.

534 Parvinen, K. 2013: Joint evolution of altruistic cooperation and dispersal in a
535 metapopulation of small local populations. *Theoretical population biology*
536 85:12–19.

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

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

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

544 Sample, C. & Allen, B. 2017: The limits of weak selection and large population
545 size in evolutionary game theory. *Journal of mathematical biology* pages 1–
546 33.

547 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors
548 in finite structured population models. *The American Naturalist* 184(4):477–
549 488.

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

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

- 554 Taylor, P. D. 1992b: Inclusive fitness in a homogeneous environment. Pro-
555 ceedings of the Royal Society of London. Series B: Biological Sciences
556 249(1326):299–302.
- 557 Taylor, P. D.; Day, T. & Wild, G. 2007a: Evolution of cooperation in a finite homo-
558 geneous graph. *Nature* 447(7143):469–472.
- 559 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-
560 bility in homogeneous structured populations. *Journal of Theoretical Biology*
561 249(1):101 – 110.
- 562 Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic
563 behavior. *Evolution* 54(4):1135–1141.
- 564 Van Cleve, J. 2015: Social evolution and genetic interactions in the short and long
565 term. *Theoretical Population Biology* 103:2 – 26.
- 566 West, S. A. & Gardner, A. 2010: Altruism, spite, and greenbeards. *Science*
567 327(5971):1341–1344.
- 568 West, S. A.; Pen, I. & Griffin, A. S. 2002: Cooperation and competition between
569 relatives. *Science* 296(5565):72–75.
- 570 Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely
571 viscous populations? *Evolutionary Ecology* 6(4):331–341.
- 572 Wolfram Research, Inc. 2017: *Mathematica*, Version 11.1. Champaign, IL, 2017.

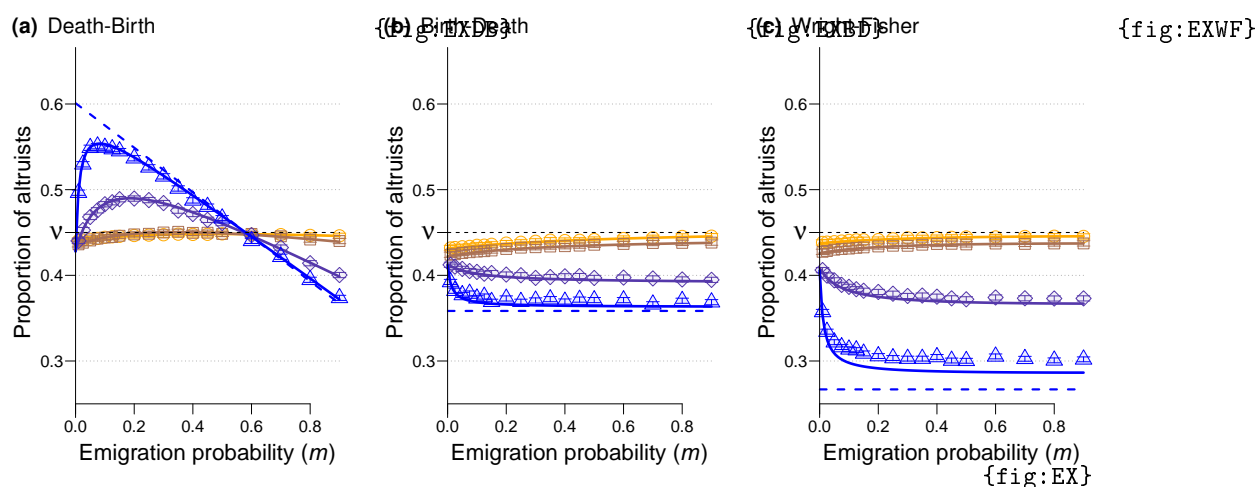


Figure 1: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different 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: $\omega = 0.005$, $\nu = 0.45$, $b = 15$, $c = 1$, $n = 4$ individuals per deme, $N_D = 15$ demes.

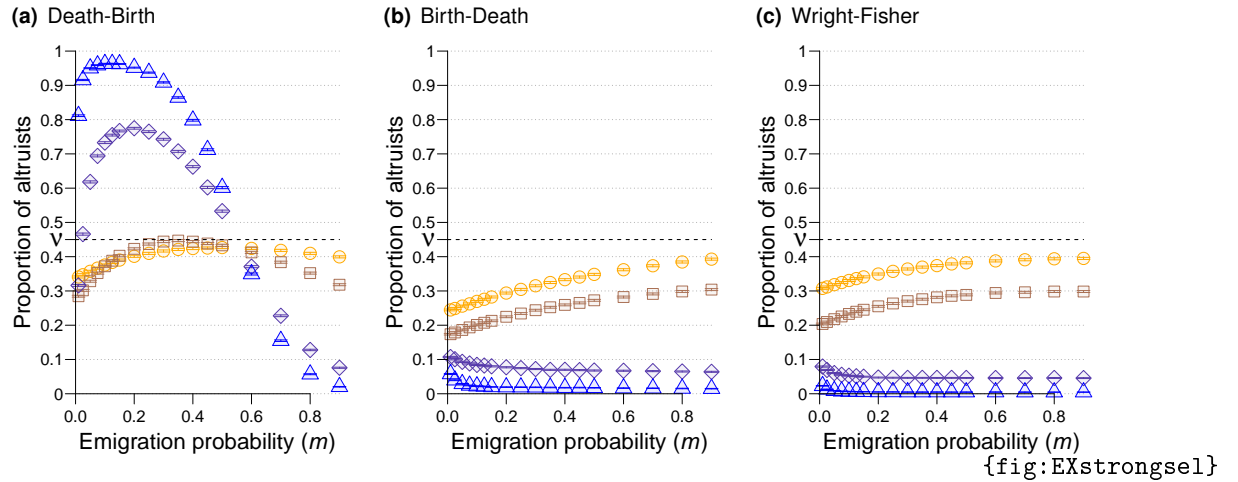


Figure S1: Equivalent of figure 1 (simulations only) but with strong selection ($\omega = 0.1$); please note the change of scale on the vertical axis. All other parameters and legends are identical to those of figure 1 (increasing mutation probabilities from blue dots to orange triangles).

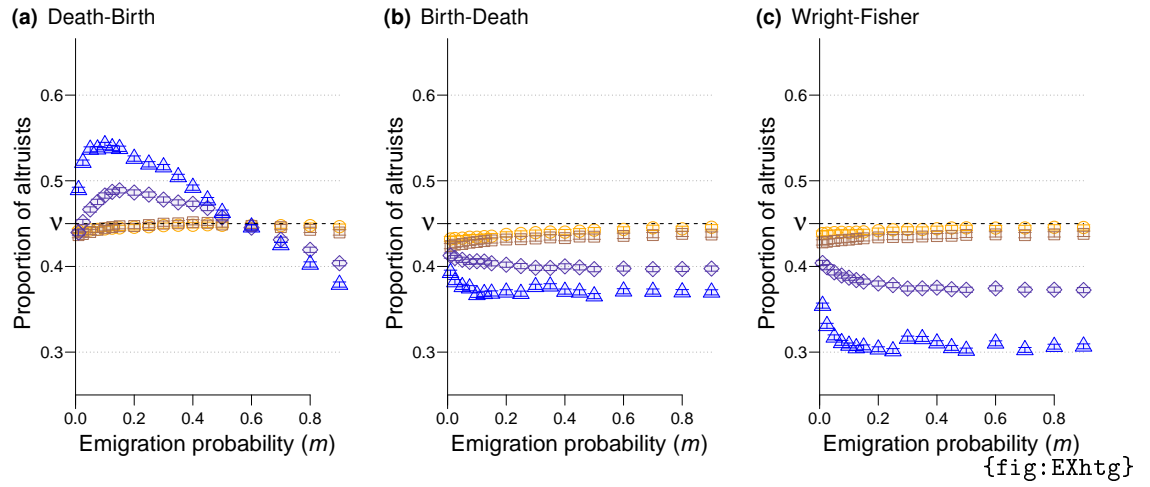


Figure S2: Equivalent of figure 1 (simulations only) but with a heterogeneous population structure: deme sizes range from 1 to 5 individuals per deme, the average deme size is 4 as in figure 1; all other parameters and legend are identical to those of figure 1.

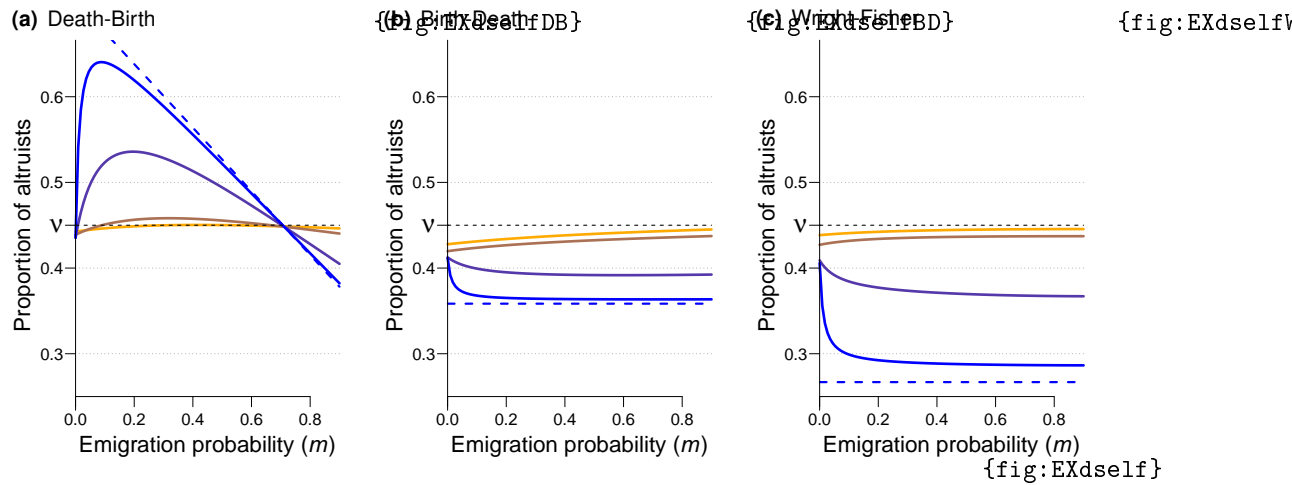


Figure S3: Equivalent of figure 1 (analysis only), with no self-replacement ($d_{ii} = d_{\text{self}} = 0$ for all sites).

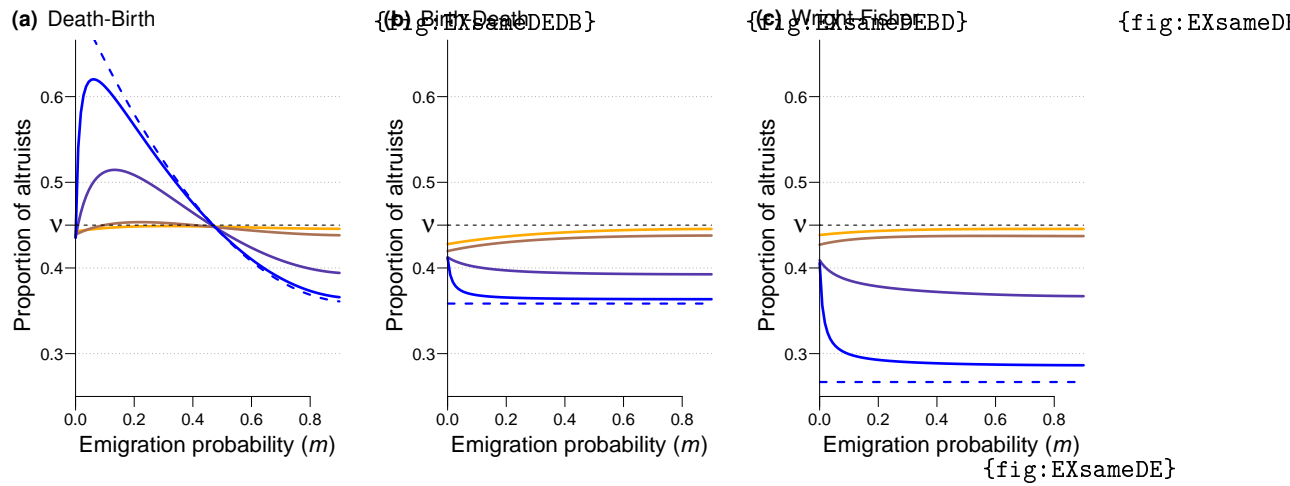


Figure S4: Equivalent of figure 1 (analysis only), with equal dispersal and interaction graphs (*i.e.*, no self-replacement [$d_{ii} = d_{\text{self}} = 0$ for all sites], and a proportion m of the interactions occurring outside of the home deme).

b	Fecundity benefit given by altruists to social interactants
c	Fecundity cost paid by altruists
d_{ij}	Dispersal probability from site i to site j
e_{ij}	Interaction probability from site i to site j
n	Deme size
N_D	Number of demes
N	Total population size ($N = N_D n$)
m	Emigration probability
Q_{ij}	(Long-term) Probability of identity by descent of individuals at sites i and j
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.)
β	Term associated to the benefits b
γ	Term associated to the costs c
μ	Mutation probability
ν	Mutation bias: probability that mutant is altruist
ω	Parameter scaling the relative effect of social interactions on fecundity
D	Subscript corresponding to direct/primary effects
I	Subscript corresponding to indirect/secondary effects
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 $i = j$
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

{tab:symbols}

Table S1: List of symbols. “r.v.” means *random variable*.

576

Appendix

A Expected frequency of altruists

{sec:app:EX}

A.1 For a generic life-cycle

{sec:app:generic}

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

{tab:BD}

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).

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

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

holds for all sites i . The structure of the population is also such that in the absence 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

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

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

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

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), which ever 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:

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

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

$$0 = \frac{\delta}{N} \sum_{i=1}^N \left[\sum_{\mathbf{X} \in \Omega} \left(\frac{\partial B_i}{\partial \delta} (1-\mu) - \frac{\partial D_i}{\partial \delta} \right) X_i \xi(\mathbf{X}, 0, \mu) - \sum_{\mathbf{X} \in \Omega} \mu B^* X_i \frac{\partial \xi}{\partial \delta} \right] + O(\delta^2), \quad (\text{A.4}) \quad \{\text{eq:weaksel0}\}$$

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

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

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

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \mathbb{E}_0 \left[\frac{\partial W_i}{\partial \delta} X_i \right] + O(\delta^2). \quad (\text{A.6}) \quad \{\text{eq:weaksel0reorg}\}$$

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

$$\begin{aligned} \frac{\partial W_i}{\partial \delta} &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} \frac{\partial \phi_k}{\partial \delta} \\ &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} X_k, \end{aligned} \quad (\text{A.7}) \quad \{\text{eq:totalproba1}\}$$

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

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} P_{ik} + O(\delta^2). \quad (\text{A.8}) \quad \{\text{eq:weaksel11}\}$$

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 since the size of the population is fixed ($\sum_{i=1}^N (B_i - D_i) = 0$), and since 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)

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

With this, eq. (A.8) becomes

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} \frac{P_{\text{in}} - P_{\text{out}}}{P_{ii} - P_{\text{out}}} \right) (P_{ii} - P_{\text{out}}) + O(\delta^2). \quad (\text{A.10}) \quad \{\text{eq:weaksel11CBRP}\}$$

646 We can also replace the P terms by

$$P_{ij} = \nu^2 + \nu(1 - \nu)Q_{ij}. \quad (\text{A.11}) \quad \{\text{eq:QP}\}$$

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

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\underbrace{\frac{\partial W_i}{\partial \phi_i}}_{-C} + \underbrace{(n-1) \frac{\partial W_i}{\partial \phi_{\text{in}}}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) (1 - Q_{\text{out}}) \nu (1 - \nu) + O(\delta^2). \quad (\text{A.12}) \quad \{\text{eq:weaksel1CBR}\}$$

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

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

653 With our notation, and given that social interactions take place within demes
 654 and affect fecundity, we have \{\text{eq:derivf}\}

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

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

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

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

$$\delta\mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \delta \nu (1 - \nu) (1 - Q_{\text{out}}) \times \left(\underbrace{\left(\frac{\partial W}{\partial f_\bullet} (-c) + \frac{\partial W}{\partial f_{\text{in}}} b \right)}_{-C} + \underbrace{\left(\frac{\partial W}{\partial f_\bullet} b + (n-1) \frac{\partial W}{\partial f_{\text{in}}} (-c) + (n-2) \frac{\partial W}{\partial f_{\text{in}}} b \right)}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) + O(\delta^2). \quad (\text{A.15}) \quad \{\text{eq:weaksel2}\}$$

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

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

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

660 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

661 D_i terms by their formulas for each life-cycle (given in table S2).

662 A.2 Derivatives for the specific life-cycles

{sec:app:dW}

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

{eq:dWBD}

Moran Birth-Death

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

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

{eq:dWDB}

Moran Death-Birth

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

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

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

{eq:dWWF}

Wright-Fisher

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

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

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

B Probabilities of identity by descent

B.1 Expected state of pairs of sites and probabilities of identity by descent

{sec:app:IBD}

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

B.1.1 Moran model

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & X_i X_j \left(1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - \nu)) \right) \\ & + X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu\nu) \\ & + X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu\nu). \end{aligned} \quad (\text{B.20}) \quad \{\text{eq:app:PijM1}\}$$

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

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

while $P_{ii} = \nu$.

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

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

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

690 B.1.2 Wright-Fisher model

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = \sum_{k, \ell=1}^N d_{ki} d_{\ell j} & \left(X_k X_\ell (1 - \mu + \mu v)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu v)(\mu v) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu v)^2 \right) \end{aligned} \quad (\text{B.23}) \quad \{\text{eq:app:PijWF1}\}$$

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

697 Taking the expectation and simplifying, we obtain

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

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

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

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

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

{sec:app:Qsubdiv}

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

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

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

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

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

B.2.1 Moran model

In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1 q_2}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{B.27a}) \quad \{\text{eq:app:Q2DM}\}$$

with

$$\tilde{D}_{q_1 q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1 \ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{B.27b}) \quad \{\text{eq:app:D2D}\}$$

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

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

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

{eq:app:Dsystem}

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

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

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

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

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

{eq:app:Q2DMsol}

730 In particular,

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

{eq:app:Q2D1}

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

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

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

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

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

{eq:QM}

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

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

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

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

744 B.2.2 Wright-Fisher

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

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

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

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

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

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

Then from eq. (B.34) we deduce

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

and

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

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

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

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

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)).

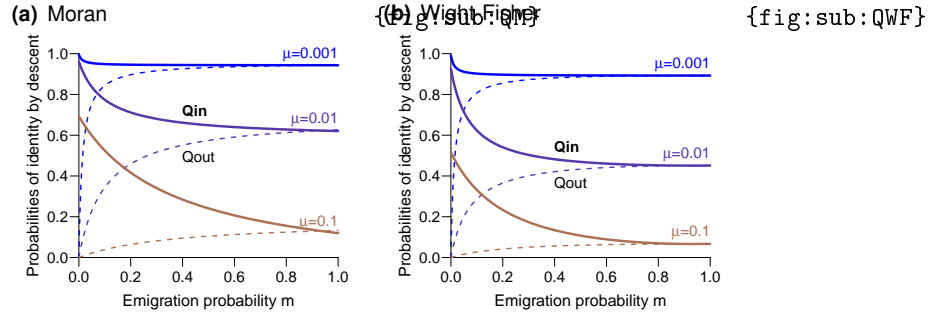


Figure S5: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in} , full curves) and two individuals in different demes (Q_{out} , dashed curves), as a function of the emigration probability m , for different values of the mutation probability μ (0.001, 0.01,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