



HAL
open science

Imperfect Strategy Transmission Can Reverse the Role of Population Viscosity on the Evolution of Altruism

Florence Débarre

► **To cite this version:**

Florence Débarre. Imperfect Strategy Transmission Can Reverse the Role of Population Viscosity on the Evolution of Altruism. *Dynamic Games and Applications*, 2019, 10, pp.732-763. 10.1007/s13235-019-00326-y . hal-02406568

HAL Id: hal-02406568

<https://hal.science/hal-02406568>

Submitted on 9 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

F. Débarre*

*Sorbonne Université, CNRS, Université Paris Est Créteil, INRA, IRD, Institute of Ecology and Environmental sciences - Paris, IEES-Paris, UMR 7618, 4 place Jussieu, 75005 Paris, France

florence.debarre@normalesup.org

+33 1 44 27 56 71

ORCID: 0000-0003-2497-833X

Abstract

Population viscosity, *i.e.*, low emigration out of the natal deme, leads to high within-deme relatedness, which is beneficial to the evolution of altruistic behavior when social interactions take place among deme-mates. However, a detrimental side-effect of low emigration is the increase in competition among related individuals. The evolution of altruism depends on the balance between these opposite effects. This balance is already known to be affected by details of the life cycle; we show here that it further depends on the fidelity of strategy transmission from parents to their offspring. We consider different life cycles and identify thresholds of parent-offspring strategy transmission inaccuracy, above which higher emigration can increase the frequency of altruists maintained in the population. Predictions were first obtained analytically assuming weak selection and equal deme sizes, then confirmed with stochastic simulations relaxing these assumptions. Contrary to what happens with perfect strategy transmission from parent to offspring, our results show that higher emigration can be favorable to the evolution of altruism.

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

Acknowledgements

Thanks to Charles Mullon and Jorge Peña for detailed comments on a previous version of the manuscript (CM suggested the $\mathcal{B}R - \mathcal{C}$ decomposition; JP suggested the shortcut calculation of eq. (A17c)). Sébastien Lion suggested using the R vs. \mathcal{C}/\mathcal{B} comparison to interpret the result.

This work was funded by a ANR-14-ACHN-0003-01 grant.

1 Introduction

2 In his pioneering work on the evolution of social behavior, Hamilton suggested that al-
3 truisitic behavior would be associated to limited dispersal (Hamilton, 1964, p. 10). This
4 notion, that tighter links between individuals are beneficial to the evolution of altruism,
5 has been shown to hold in a number of population structures (see *e.g.* Ohtsuki et al.,
6 2006; Taylor et al., 2007a; Lehmann et al., 2007; Allen et al., 2017). The rationale is that al-
7 truism is favored when altruists interact more with altruists than defectors do (Hamilton,
8 1975, p. 141; Fletcher & Doebeli, 2009), a condition that is met in viscous populations,
9 *i.e.*, populations with limited dispersal.

10 Yet, living next to your kin also implies competing against them (West et al., 2002;
11 Platt & Bever, 2009), which is detrimental to the evolution of altruism. The evolution of
12 social traits hence depends on the balance between the positive effects of interactions
13 with related individuals and the detrimental consequences of kin competition. Under
14 specific conditions, the two effects can even compensate each other, thereby annihi-
15 lating the impact of population viscosity on the evolution of altruism. First identified
16 with computer simulations (Wilson et al., 1992), this cancellation result was analyzed
17 by Taylor (1992a) in a model with synchronous generations (*i.e.*, Wright-Fisher model)
18 and a subdivided population of constant, infinite size. The cancellation result was later
19 extended to heterogeneous populations (Rodrigues & Gardner, 2012, with synchronous
20 generations and infinite population size), and other life cycles, with generic regular pop-
21 ulation structures (Taylor et al., 2011, with synchronous generations but also with con-
22 tinuous generations and Birth-Death updating). However, small changes in the model's
23 assumptions, such as overlapping generations (Taylor & Irwin, 2000) or the presence of
24 empty sites (Alizon & Taylor, 2008) can tip the balance in the favor of altruism. This high
25 dependence on life cycle specificities highlights the difficulty of making general state-
26 ments about the role of spatial structure on the evolution of altruism.

27 Three different life cycles are classically used in studies on altruism in structured
28 populations: Wright-Fisher, where the whole population is renewed at each time step,
29 and two Moran life cycles (Birth-Death and Death-Birth), where a single individual dies
30 and is replaced at each time step. We will consider the three of them in this study, be-
31 cause even though they differ by seemingly minor details, they are known to have very
32 different outcomes in models with perfect parent-offspring transmission (*e.g.*, Taylor,
33 1992a; Rousset, 2004; Ohtsuki et al., 2006; Lehmann et al., 2007; Taylor, 2010).

34 A large number of studies on the evolution of social behavior consider simple pop-
35 ulation structures (typically, homogeneous populations *sensu* Taylor et al. (2007a)) and
36 often also infinite population sizes (but see Allen et al., 2017, for results on any struc-
37 ture). These studies also make use of weak selection approximations, and commonly
38 assume rare (*e.g.*, Leturque & Rousset, 2002; Taylor et al., 2007b; Tarnita & Taylor, 2014;
39 Chen et al., 2019) or absent mutation (for models assuming infinite population sizes, or
40 models concentrating on fixation probabilities; see Lehmann & Rousset, 2014; Van Cleve,
41 2015, for recent reviews). These simplifying assumptions are often a necessary step to-
42 wards obtaining explicit analytical results. Simple population structures (*e.g.*, regular
43 graphs, or subdivided populations with demes of equal sizes) help reduce the dimen-
44 sionality of the system under study, in particular when the structure of the population
45 displays symmetries such that all sites behave the same way in expectation. Weak se-

46 lection approximations are crucial for disentangling spatial moments (Lion, 2016), that
47 is, changes in global *vs.* local frequencies (though they can in some cases be relaxed, as
48 in Mullon & Lehmann, 2014). Mutation, however, is usually ignored by classical models
49 of inclusive fitness because these models assume infinite population sizes, so that there
50 is no need to add mechanisms that restore genetic diversity (Tarnita & Taylor, 2014). In
51 populations of finite size, this diversifying effect can be obtained thanks to mutation.

52 When strategy transmission is purely genetic, it makes sense to assume that muta-
53 tion is relatively infrequent. Even in this case, though, mutations from “social” to “non-
54 social” types cannot always be neglected. For instance, experiments with the bacteria
55 *Pseudomonas fluorescens* have identified transitions between populations dominated by
56 the ancestral “solitary” Smooth Morph type and mat-forming “social” Wrinkly Spread-
57 ers, that can be re-invaded by Smooth Morphs not contributing to the formation of the
58 mat (hence described as “cheaters”). The transitions between the different types are due
59 to spontaneous mutations occurring over the timescale of the experiment (Hammer-
60 schmidt et al., 2014). In addition to genetic transmission, a social strategy can also be
61 culturally transmitted from parent to offspring. In this case, “rebellion” (as in Frank’s Re-
62 bellious Child Model (Frank, 1997)), *i.e.*, adopting a social strategy different from one’s
63 parents, does not have to be infrequent. Since it is known that imperfect strategy trans-
64 mission can alter the evolutionary dynamics of social traits, in particular in spatially
65 structured populations (see *e.g.*, Allen et al., 2012; Débarre, 2017, for graph-structured
66 populations), it is therefore important to understand the impact of imperfect strategy
67 transmission on the evolution of social behavior.

68 Here, we want to explore the consequences of imperfect strategy transmission from
69 parents to their offspring on the evolution of altruistic behavior in subdivided popula-
70 tions¹. The question was tackled by Frank (1997), but with a non “fully dynamic model”
71 (Frank, 1997, legend of Fig.7). Relatedness was treated like a parameter, which precluded
72 the exploration of the effects of population viscosity on the evolution altruism.

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

78 **Model and methods**

79 **Assumptions**

80 We consider a population of total size N , subdivided into N_D demes connected by dis-
81 persal, each deme hosting exactly n individuals (*i.e.*, each deme contains n sites, each
82 of which is occupied by exactly one individual; $nN_D = N$). Each site has a unique label
83 i , $1 \leq i \leq N$. There are two types of individuals in the population, altruists and defec-
84 tors. The type of the individual living at site i ($1 \leq i \leq N$) is given by an indicator variable
85 X_i , equal to 1 if the individual is an altruist, and to 0 if it is a defector. The state of the

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.

86 entire population is given by a vector $\mathbf{X} = \{X_i\}_{1 \leq i \leq N}$. For a given population state \mathbf{X} , the
 87 proportion of altruists is $\bar{X} = \sum_{i=1}^N X_i / N$. All symbols are summarized in table A1.

88 Reproduction is asexual. The offspring of altruists are altruists themselves with prob-
 89 ability $1 - \mu_{1 \rightarrow 0}$, and are defectors otherwise ($0 < \mu_{1 \rightarrow 0} \leq 1/2$). Similarly, the offspring of
 90 defectors are defectors with probability $1 - \mu_{0 \rightarrow 1}$, and are altruists otherwise ($0 < \mu_{0 \rightarrow 1} \leq$
 91 $1/2$). Our calculations will be simpler if we introduce the following change of parameters:
 92

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

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

93 The composite parameter v corresponds to the expected frequency of altruists in the
 94 population at the mutation-drift balance (*i.e.*, in the absence of selection; see Appendix A
 95 for details). We call v the “mutation bias” parameter. Parameter μ is the sum of the two
 96 mutation probabilities. In the absence of selection, at the mutation-drift equilibrium,
 97 the correlation between offspring type and their parent’s type is $1 - \mu$ (see Appendix A for
 98 details for the calculation). We call μ the mutation intensity.

99 An individual of type X_k expresses a social phenotype $\phi_k = \delta X_k$, where δ is assumed
 100 to be small ($\delta \ll 1$). This assumption of small phenotypic differences leads to weak selec-
 101 tion. This type of weak selection is called “ δ -weak selection” in Wild & Traulsen (2007).
 102 Social interactions take place within each deme; a focal individual interacts with its $n - 1$
 103 other deme-mates. We assume that social interactions affect individual fecundity; f_k de-
 104 notes the fecundity of the individual at site k ($1 \leq k \leq N$), which depends on deme com-
 105 position. We denote by b the sum of the marginal effects of deme-mates’ phenotypes on
 106 the fecundity of a focal individual, and by $-c$ the marginal effect of a focal individual’s
 107 phenotype on its own fecundity ($c \leq b$; see system (A22) for formal definitions).

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

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

113 with $0 < m < 1 - \frac{1}{N_D}$. This upper bound is here to ensure that within-deme relatedness R ,
 114 which will be defined later in the article, remains positive. When the emigration prob-
 115 ability m is equal to the upper bound $1 - \frac{1}{N_D}$, the population is effectively well-mixed
 116 ($d_{\text{in}} = d_{\text{out}}$).

117 We denote by $B_i = B_i(\mathbf{X}, \delta)$ the expected number of successful offspring of the indi-
 118 vidual living at site i (“successful” means alive at the next time step), and by $D_i = D_i(\mathbf{X}, \delta)$
 119 the probability that the individual living at site i dies. Both depend on the state of the
 120 population \mathbf{X} , but also on the way the population is updated from one time step to the
 121 next, *i.e.*, on the chosen life cycle (also called updating rule). Because this term appears
 122 in our calculations, we also define

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

123 This is a particular definition of fitness, where the number of offspring produced (B_i) is
124 scaled by the parent-offspring type correlation ($1 - \mu$).

125 We will specifically explore three different life cycles. At the beginning of each step of
126 each life cycle, all individuals produce a large (effectively infinite) number of offspring,
127 in proportion to their fecundity; some of these offspring can be mutated. Then these
128 juveniles move, within the parental deme or outside of it, and land on a site. The next
129 events occurring during the time step depend on the life cycle:

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

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

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

137 Previous studies have shown that, when social interactions affect fecundity, altruism is
138 disfavored under the Moran Birth-Death and Wright-Fisher life cycles, because the ex-
139 pected frequency of altruists under these life cycles is lower than what it would be in the
140 absence of selection (*e.g.*, Taylor, 1992a, 2010; Taylor et al., 2011; Débarre, 2017). How-
141 ever, we are interested in the actual value of the expected proportion of altruists in the
142 population, not just whether it is higher or lower than the neutral expectation. This is
143 why we are still considering the Moran Birth-Death and Wright-Fisher life cycles in this
144 study.

145 **Methods**

146 **Analytical part**

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

153 The formula involves quantities that can be identified as neutral probabilities of
154 identity by descent Q_{ij} . These quantities correspond to the probability that individu-
155 als living at site i and j share a common ancestor and that no mutation occurred on
156 either lineage since that ancestor, in a model with no selection ($\delta = 0$) and with mutation
157 intensity μ ; this is the “mutation definition” of identity by descent (Rousset & Billiard,
158 2000). In a subdivided population like the one we consider, there are only three possible
159 values of Q_{ij} :

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

160 These neutral probabilities of identity by descent depend on the chosen life cycle, and
 161 are also computed by taking the long-term expectation of conditional expectations after
 162 one time step (see Appendix C.1 and C.2 and supplementary Mathematica file (Wolfram
 163 Research, Inc., 2017).)

164 Stochastic simulations

165 To check our results and also relax some key assumptions, we ran stochastic simula-
 166 tions. The simulations were run for 10^8 generations (one generation is one time step for
 167 the Wright-Fisher life cycle, and N time steps for the Moran life cycles). For each set of
 168 parameters and life cycle, we estimated the long-term frequency of altruists by sampling
 169 the population every 10^3 generations and computing the average frequency of altruists.
 170 All scripts are available at
 171 <https://flodebarre.github.io/SocEvolSubdivPop/>

172 Results

173 Expected frequencies of altruists for each life cycle

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

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

176 with W as defined in eq. (3). Calculations leading to eq. (5) are presented in Appendix B;
 177 notations are recapitulated in table A1. In particular, B^* is the expected number of off-
 178 spring produced by an adult, in the absence of selection (when $\delta = 0$; $B^* = 1$ for the
 179 Wright-Fisher life cycle and $B^* = 1/N$ for the Moran life cycles). Subscript “•” denotes a
 180 focal individual itself, and “in” a deme-mate. Partial derivatives are evaluated for $\delta = 0$.

181 The expected frequency of altruists in the population is approximated, under weak
 182 selection ($\delta \ll 1$), by the sum of what it would be in the absence of selection ($\mathbb{E}_0[\bar{X}] = \nu$,
 183 first term in eq. (5)), plus a deviation from this value, scaled by δ . The $-C$ term cor-
 184 responds to the effects of a change of a focal individual’s phenotype on its own fitness
 185 (with the fitness definition given in eq. (3)). The B term corresponds to the sum of the ef-
 186 fects of the change of deme-mates’ phenotypes on an individual’s fitness. It is multiplied
 187 by R , which is relatedness.

188 The parametrization proposed in eq. (1) allows us to decouple the effects of the two
 189 new mutation parameters, ν and μ . The mutation bias ν , which was defined in eq. (1a),
 190 does not affect the sign of the second (“deviation”) term in eq. (5); it only appears in the
 191 $\nu(1-\nu)$ product. The mutation intensity μ , however, affects the values of W , Q_{in} and
 192 Q_{out} . The presence of μ at the denominator in eq. (5) may look ominous; however, both
 193 R and $(1 - Q_{\text{out}})/\mu$ have a finite limit when $\mu \rightarrow 0$.

194 The different terms depend on the chosen life cycle. We first focus on relatedness R .

195 Relatedness R

196 Within-deme relatedness R depends on the number of individuals that are born at each
 197 time step, and hence on the chosen life cycle. In a Moran life cycle (denoted by M), one
 198 individual is updated at each time step, while under a Wright-Fisher life cycle (denoted
 199 by WF), N individuals – the whole population – are updated at each time step. The for-
 200 mulas for relatedness, R^M and R^{WF} , calculated for any number of demes N_D and muta-
 201 tion intensity μ , are presented in Appendix C.2 (eq. (A44) and eq. (A50)). When we let the
 202 number of demes go to infinity ($N_D \rightarrow \infty$) and the intensity of mutation be vanishingly
 203 small ($\mu \rightarrow 0$), we recover the classical formulas for relatedness as limit cases (eq. (A45)
 204 and eq. (A51)).

205 The effects of emigration m and mutation intensity μ on relatedness are represented
 206 in figure 1. For $0 < m < 1 - 1/N_D$, within-deme relatedness is positive, and it decreases
 207 with m and with μ (the mutation bias ν has no effect). The effect of the mutation inten-
 208 sity μ on relatedness is strongest at low emigration probabilities m . As m increases, the
 209 relatedness values for different mutation intensities get closer, until they all hit zero for
 210 $m = 1 - 1/N_D$ (which is the upper bound for the emigration values that we consider, a
 211 value such that there is no proper population subdivision anymore).

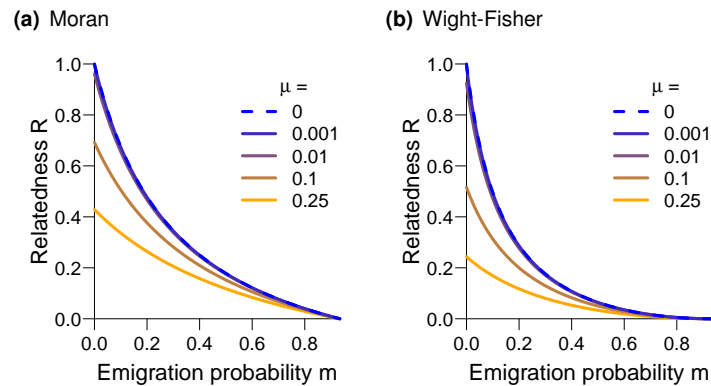


Figure 1: Within-deme relatedness of pairs of individuals R , as a function of the emigration probability m , for different values of the mutation probability μ (from 0 [blue] to 0.25 [orange]), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals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

212 Primary and secondary effects

213 We now turn to the \mathcal{B} and $-\mathcal{C}$ terms of eq. (5), which also depend on the chosen life cycle.
 214 We further decompose these terms into primary (subscript P) and secondary (subscript
 215 S) effects (West & Gardner, 2010):

$$\begin{aligned}
 \mathcal{B} &= \mathcal{B}_P & + & \mathcal{B}_S, \\
 -\mathcal{C} &= \underbrace{-\mathcal{C}_P}_{\text{Primary effect}} & + & \underbrace{-\mathcal{C}_S}_{\text{Secondary effect}}.
 \end{aligned}
 \tag{6}$$

216 Primary effects correspond to unmediated consequences of interactions (they are in-
 217 cluded in $\frac{\partial W}{\partial f}$). Secondary effects correspond to consequences of interactions mediated

218 by other individuals, including competition.

219 Primary effects

220 Primary effects are the same for all the life cycles that we consider:

$$\mathcal{B}_p^{\text{BD}} = \mathcal{B}_p^{\text{DB}} = \mathcal{B}_p^{\text{WF}} = (1 - \mu)b, \quad (7a)$$

$$-\mathcal{C}_p^{\text{BD}} = -\mathcal{C}_p^{\text{DB}} = -\mathcal{C}_p^{\text{WF}} = (1 - \mu)(-c), \quad (7b)$$

221 and they do not depend on the emigration probability m (see Appendix B.2 for details of
222 the calculations).

223 As we have seen above, the relatedness terms R^{M} and R^{WF} decrease with m (keeping
224 $m < 1 - 1/N_D$; see figure 1). Consequently, if we ignored secondary effects, we would
225 conclude that the expected frequency of altruists in the population $\mathbb{E}[\bar{X}]$ decreases as
226 the emigration probability m increases. However, secondary effects play a role as well.

227 Secondary effects

228 Secondary effects take competition into account, that is, how the change in the fecun-
229 dity of an individual affects the fitness of another one. As shown already in models with
230 nearly perfect strategy transmission (Grafen & Archetti, 2008), competition terms de-
231 pend on the chosen life cycle, because life cycle details affect the distance at which com-
232 petitive effects are felt. Given the way the model is formulated, $-\mathcal{C}_S = \mathcal{B}_S/(n - 1)$ holds
233 for all the life cycles that we consider (see Appendix B.2 for details of the calculations).

234 Under the Moran Birth-Death life cycle, both the probability of reproducing and the
235 probability of dying depend on the composition of the population. We obtain the fol-
236 lowing secondary effects:

$$-\mathcal{C}_S^{\text{BD}} = \frac{\mathcal{B}_S^{\text{BD}}}{n - 1} = -(b - c) \left(-\frac{\mu}{N} + \frac{1 - m}{n} \right). \quad (8a)$$

237 The competitive effects are the same for the Moran Death-Birth and Wright-Fisher
238 life cycles. In both cases, the probabilities of dying are constant, so we can factor $(1 - \mu)$
239 in the equations:

$$-\mathcal{C}_S^{\text{DB}} = \frac{\mathcal{B}_S^{\text{DB}}}{n - 1} = -\mathcal{C}_S^{\text{WF}} = \frac{\mathcal{B}_S^{\text{WF}}}{n - 1} = -(b - c)(1 - \mu) \left(\frac{(1 - m)^2}{n} + \frac{m^2}{N - n} \right). \quad (8b)$$

240 These secondary effects (eq. (8a) and eq. (8b)) remain negative for the range of emi-
241 gration values that we consider ($0 < m < 1 - 1/N_D$), and increase with m . In other words,
242 the intensity of competition decreases as emigration m increases.

243 While the value of these secondary effects increases with emigration m , relatedness
244 R , by which they are eventually multiplied in eq. (5), decreases with m . We therefore can-
245 not determine the overall effect of emigration m on the expected frequency of altruists
246 in the population by inspecting the different terms of eq. (5) in isolation. For each life
247 cycle, we need to consider the entire equations to know the overall effect of the emigra-
248 tion probability m on the expected frequency of altruists $\mathbb{E}[\bar{X}]$ and on how it is affected
249 by the (in)fideliy of parent-offspring transmission μ .

250 **Changes of the expected frequency of altruists with the emigration probability**

251 m

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

254 **Moran Birth-Death**

255 For the Moran Birth-Death life cycle, we find that the expected frequency of altruists
256 $\mathbb{E}[\bar{X}]$ is a monotonic function of the emigration probability m . The direction of the
257 change depends on the value of the mutation probability μ compared to a threshold
258 value 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
259 with m . The critical value μ_c^{BD} is given by

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

260 (recall that N is the total size of the population, $N = nN_D$.) This result is illustrated in
261 figure 2(b); with the parameters of the figure, $\mu_c^{\text{BD}} \approx 0.026$. The threshold value increases
262 with both deme size n and number of demes N_D , up to a maximum value $1 - \sqrt{1 - c/b}$
263 (equal to 0.034 with the parameters of figure 2(b).)

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

266 **Moran Death-Birth**

267 The relationship between $\mathbb{E}[\bar{X}]$ and m is a bit more complicated for the Moran Death-
268 Birth life cycle. For simplicity, we concentrate on what happens starting from low emi-
269 gration probabilities (*i.e.*, on the sign of the slope of $\mathbb{E}[\bar{X}]$ as a function of m when $m \rightarrow 0$).
270 If the benefits b provided by altruists are relatively low ($b < c(n + 1)$), $\mathbb{E}[\bar{X}]$ initially in-
271 creases with m provided the mutation probability μ is greater than a threshold value μ_c^{DB}
272 given in eq. (10) below; otherwise, when the benefits are high enough, $\mathbb{E}[\bar{X}]$ initially in-
273 creases with m for any value of μ . Combining these results, we write

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

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

276 When $\mu > \mu_c^{\text{DB}}$, the expected frequency of altruists $\mathbb{E}[\bar{X}]$ reaches a maximum at an
277 emigration probability m_c^{DB} (whose complicated equation is given in the supplementary
278 Mathematica file), as can be seen in figure 2(a). When the mutation probability gets close
279 to 0 ($\mu \rightarrow 0$), m_c^{DB} also gets close to 0.

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

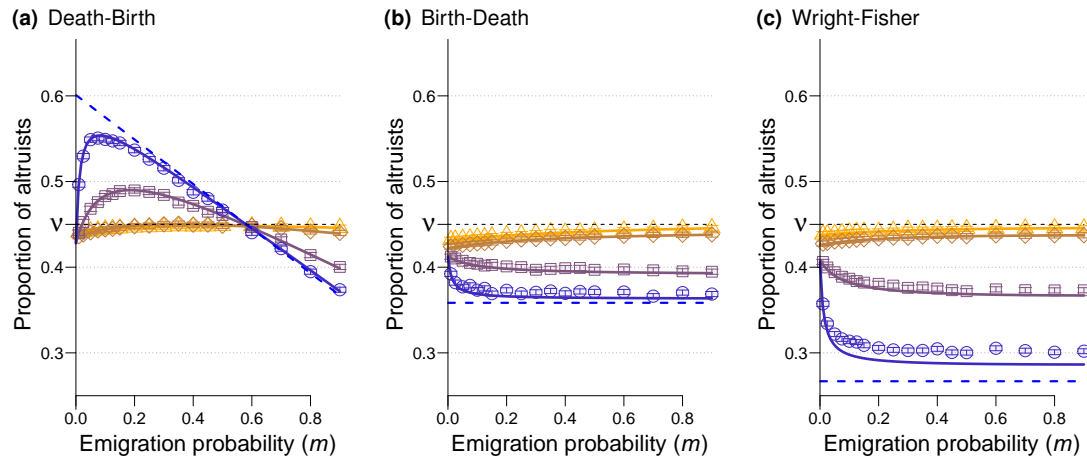


Figure 2: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different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

283 Wright-Fisher

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

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

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

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

291 Relaxing key assumptions

292 To derive our analytical results, we had to make a number of simplifying assumptions,
 293 such as the fact that selection is weak ($\delta \ll 1$), and the fact that the structure of the pop-
 294 ulation is regular (all demes have the same size n). We checked with numerical simula-
 295 tions the robustness of our results when these key assumptions are relaxed.

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

298 **Heterogeneity in deme sizes** To relax the assumption of equal deme sizes, we ran-
 299 domly drew deme sizes at the beginning of simulations, with sizes ranging from 2 to 6

300 individuals and on average $\bar{n} = 4$ individuals per deme as previously. As shown in fig-
301 ure A2, the patterns initially obtained with a homogeneous population structure are ro-
302 bust when the structure is heterogeneous.

303 **No self-replacement** For the Moran model, it may seem odd that an offspring can re-
304 place its own parent (which can occur since $d_{ii} \neq 0$). Figure A3, plotted with disper-
305 sal probabilities preventing immediate replacement of one's own parent (for all sites i ,
306 $d_{ii} = d_{\text{self}} = 0$; $d_{\text{in}} = (1 - m)/(n - 1)$ for two different sites in the same deme, d_{out} remain-
307 ing unchanged), confirms that this does affect our conclusions.

308 **Infinite number of demes** Our results are obtained in a population of finite size (the
309 figures are drawn with $N_D = 15$ demes), but still hold when the size of the population is
310 larger. Figure 3(b) shows the range of emigration and mutation values such that altruism
311 is favored, plotted also for $N_D \rightarrow \infty$.

312 **Same graphs for dispersal and social interactions** Compared to graphs classically used
313 in evolutionary graph theory (*e.g.*, regular random graphs, grids), the island model is par-
314 ticular because the interaction graph and the dispersal graph are different: interactions
315 take place only within demes ($e_{\text{out}} = 0$), while offspring can disperse out of their natal
316 deme ($d_{\text{out}} > 0$). One may wonder whether our result depends on this difference be-
317 tween the two graphs. Figure A4 shows that the result still holds when the dispersal and
318 interaction graphs are the same. In this figure indeed, we let a proportion m (equal to
319 the dispersal probability) of interactions occur outside of the deme where the individu-
320 als live, and set d_{self} , the probability of self replacement, equal to 0, so that the dispersal
321 and interactions graphs are the same. Our conclusions remain unchanged.

322 Discussion

323 The expected frequency of altruists in a subdivided population can increase 324 with the probability of emigration

325 Assuming that the transmission of a social strategy (being an altruist or a defector) from a
326 parent to its offspring could be imperfect, we found that the expected frequency of altru-
327 ists maintained in a population could increase with the probability m of emigration out
328 of the parental deme, a parameter tuning population viscosity. This result can seem sur-
329 prising, because it contradicts the conclusions obtained under the assumption of nearly
330 perfect strategy transmission (*i.e.*, in the case of genetic transmission, when mutation is
331 very weak or absent). Under nearly perfect strategy transmission indeed, increased pop-
332 ulation viscosity (*i.e.*, decreased emigration probability) is either neutral (Taylor, 1992a,
333 and dashed lines in figures 2(b)–(c)) or favorable (Taylor et al., 2007a, and dashed lines
334 in figure 2(a)) to the evolution of altruistic behavior.

335 Quantitative vs. qualitative measures

336 Often, evolutionary success is measured qualitatively, by comparing a quantity (an ex-
337 pected frequency, or, in models with no mutation, a probability of fixation) to the value

338 it would have in the absence of selection. In our model, this amounts to saying that
339 altruism is favored whenever $\mathbb{E}[\bar{X}] > \nu$ (ν is plotted as a horizontal dashed line in fig-
340 ure 2). Some of our conclusions change if we use this qualitative measure of evolution-
341 ary success: Under the Moran Birth-Death and Wright-Fisher life cycles, population vis-
342 cosity does not promote the evolution of altruism – actually, these two life cycles can-
343 not ever promote altruistic behavior for any regular population structure (Taylor et al.,
344 2011), whichever the probability of mutation (Débarre, 2017). However, under a Moran
345 Death-Birth life cycle (figure 2(a)), altruism can be favored only at intermediate emi-
346 gration probabilities. Starting for initially low values of m , increasing the emigration
347 probability can still favor the evolution of altruism under this qualitative criterion (see
348 figure 3(b).)

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

350 To better understand the role played by the mutation intensity μ , we focus on the qual-
351 itative condition for the evolution of altruism ($\mathbb{E}[\bar{X}] > \nu$); and on the Death-Birth life
352 cycle, since this qualitative condition is not satisfied in the two other life cycles. Having
353 made sure that $\mathcal{B}^{\text{DB}} > 0$ (as shown in the supplementary Mathematical file), the qualita-
354 tive condition for altruism to be favored is given by

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

355 With the Death-Birth life cycle, the $\mathcal{C}^{\text{DB}}/\mathcal{B}^{\text{DB}}$ ratio does not change with the mutation
356 probability μ (the $(1 - \mu)$ factors simplified out), but the ratio decreases with the emi-
357 gration probability m (with $0 < m < 1 - 1/N_D$; see the thick black curve in figure 3(a)).
358 This decrease of the $\mathcal{C}^{\text{DB}}/\mathcal{B}^{\text{DB}}$ ratio is due to secondary effects (competition) diminish-
359 ing as emigration increases. Relatedness, on the other hand, decreases with both μ and
360 m (see figure 3(a)). We need to explain the effect of the emigration probability m on
361 condition (12) for different values of mutation intensity μ .

362 When the emigration probability m is high, relatedness gets closer to zero for all val-
363 ues of mutation intensity μ , while the $\mathcal{C}^{\text{DB}}/\mathcal{B}^{\text{DB}}$ remains positive; condition (12) is not
364 satisfied. On the other hand, when the emigration probability m is vanishingly small,
365 $\lim_{m \rightarrow 0} R^{\text{M}} \leq \lim_{m \rightarrow 0} \frac{\mathcal{C}^{\text{DB}}}{\mathcal{B}^{\text{DB}}}$, the two only being equal when $\mu = 0$. Hence, condition (12)
366 is satisfied for vanishingly low m only when strategy transmission is perfect. Finally, as
367 m increases to intermediate values, the $\frac{\mathcal{C}^{\text{DB}}}{\mathcal{B}^{\text{DB}}}$ ratio decreases with a steeper slope than
368 relatedness R , so that the curves can cross provided the mutation probability μ is not
369 too high, *i.e.*, that R was not initially too low already. Hence, for no too high mutation
370 intensity, there is a range of emigration values m such that condition (12) is satisfied.

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

372 The result, that frequency of altruists can increase with the emigration probability m ,
373 may seem counterintuitive. It is the case because verbal explanations for the evolution
374 of altruism often rely on primary effects only. Relatedness R decreases with m , so it may
375 be tempting to conclude that increases in the emigration probability m are necessarily
376 detrimental to the evolution of altruism. However, secondary effects play an opposite

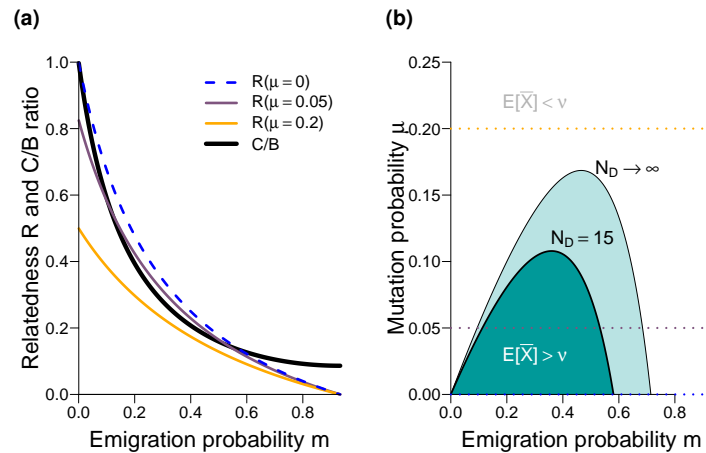


Figure 3: Understanding the effect of emigration m on whether altruism is favored in the Death-Birth life-cycle. (a) Comparison of the C/B ratio (thick black curve) and relatedness R (thin curves) for different values of the mutation probability μ (same color code as previously). (b) (m, μ) combinations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

377 role, as competition decreases with m , and the effect is strongest at low values of m (see
 378 the black curve on figure 3(a); in the absence of secondary effects, it would just be a
 379 horizontal line).

380 Secondary effects are less straightforward to understand than primary effects, and
 381 yet they play a crucial role for social evolution in spatially structured populations. Com-
 382 petition among relatives is for instance the reason for Taylor (1992b)'s cancellation result.
 383 Similarly, the qualitative differences between the Moran Birth-Death and Moran Death-
 384 Birth life cycles is explained by the different scales of competition that the two life cycle
 385 produce (Grafen & Archetti, 2008; Débarre et al., 2014). Secondary effects are also behind
 386 the evolution of social behaviors such as spite (West & Gardner, 2010).

387 How small is small and how large is large?

388 Our results were derived under the assumption of weak selection, assuming that the phe-
 389 notypic difference between altruists and defectors is small ($\delta \ll 1$). We considered any
 390 fidelity of transmission (any μ between 0 and 1) and population size. However, most
 391 models considering subdivided populations assume nearly perfect strategy transmis-
 392 sion ($\mu \rightarrow 0$) and infinite population sizes (number of demes $N_D \rightarrow \infty$). The point is
 393 technical, but it is important to know that the order in which these limits are taken mat-
 394 ters, *i.e.*, one needs to specify how small μ and δ are compared to the inverse size of
 395 the population $1/N$. This is in particular the case for the probability of identity by de-
 396 scendent of two individuals in different demes, Q_{out} : if we first take the small mutation limit,
 397 $\lim_{\mu \rightarrow 0} Q_{\text{out}} = 0$, while if we first take the large population limit, $\lim_{N \rightarrow \infty} Q_{\text{out}} = 1$ (see
 398 Appendix C.2 for details). This remark complements findings by Sample & Allen (2017),
 399 who highlighted the quantitative differences between different orders of weak selection

400 and large population limits.

401 **Imperfect transmission and Rebellious Children**

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

407 Mutation was also previously included in models investigating the maintenance of
408 cooperative microorganisms in the presence of cheaters (Brockhurst et al., 2007; Frank,
409 2010). In both of these models however, only loss-of-function mutation was considered,
410 which corresponds to setting the mutation bias at $\nu = 0$ in our model. This means that
411 the all-cheaters state is absorbing; no matter how favored cooperators may otherwise be,
412 in the long run, a finite population will only consist of cheaters.

413 **Cultural transmission**

414 Strategy transmission does not have to be genetic: it can be cultural. In our model, strat-
415 egy transmission occurs upon reproduction, so this is a case of vertical cultural trans-
416 mission.

417 The model could nevertheless be interpreted as a representation of horizontal trans-
418 mission, if we described reproduction as an instance of an individual convincing another
419 one to update its strategy. The Moran Death-Birth model can be interpreted as a modi-
420 fied imitation scheme (Boyd & Richerson, 2002; Ohtsuki et al., 2006; Traulsen et al., 2009)
421 – with a specific function specifying who is imitated –, with mutation (Kandori et al.,
422 1993), or as a voter model (Schneider et al., 2016). First, we choose uniformly at random
423 an individual who may change its strategy; with probability μ the individual chooses a
424 random strategy (altruistic with probability ν), and with probability $1 - \mu$ it imitates an-
425 other individual. Who is imitated depends on the distance to the focal individual (with
426 probability m it is a random individual in another deme) and on the “fecundities” of
427 those individuals (as shown in table A2). With this interpretation of the updating rule
428 however, there is not reproduction nor death anymore.

429 It remains to be investigated how imperfect strategy transmission would affect the
430 effect of population viscosity on the evolution of altruism in a model implementing both
431 reproduction and horizontal cultural transmission (as in Lehmann et al., 2008). Such a
432 model could then contrast the effects of imperfect genetic transmission and imperfect
433 horizontal cultural transmission.

434 **Coevolution of dispersal and social behavior**

435 This work also raises the question of what would happen if dispersal (e.g., the emigra-
436 tion probability m) could evolve as well. Recent work on the topic has shown that un-
437 der some conditions disruptive selection could take place, leading to a polymorphism
438 between sessile altruists and mobile defectors (Parvinen, 2013; Mullon et al., 2017)—
439 though more complex coevolutionary patterns can be obtained when considering the

440 coevolution of altruism and mobility instead of natal dispersal, and unsaturated popu-
441 lations (Le Galliard et al., 2005). The assumptions of these studies however differ from
442 ours in important ways, in that they consider continuous traits and use an adaptive dy-
443 namics framework, where, notably, mutations are assumed to be very rare. It remains to
444 be investigated how non-rare and potentially large mutations would affect their result.

445 Literature Cited

- 446 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolution*
447 62(6):1335–1344.
- 448 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak, M. A.
449 2017: Evolutionary dynamics on any population structure. *Nature* 544(7649):227–230.
- 450 Allen, B. & Nowak, M. A. 2014: Games on graphs. *EMS surveys in mathematical sciences*
451 1(1):113–151.
- 452 Allen, B. & Tarnita, C. E. 2014: Measures of success in a class of evolutionary models with
453 fixed population size and structure. *Journal of Mathematical Biology* 68(1):109–143.
- 454 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects evolu-
455 tionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105. *Evolution of*
456 *Cooperation*.
- 457 Boyd, R. & Richerson, P. J. 2002: Group beneficial norms can spread rapidly in a struc-
458 tured population. *Journal of theoretical biology* 215(3):287–296.
- 459 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at intermediate
460 disturbance. *Current Biology* 17(9):761–765.
- 461 Chen, X.; Åke Brännström & Dieckmann, U. 2019: Parent-preferred dispersal promotes
462 cooperation in structured populations. *Proceedings of the Royal Society B: Biological*
463 *Sciences* 286(1895):20181949.
- 464 Cockerham, C. C. & Weir, B. 1987: Correlations, descent measures: drift with migration
465 and mutation. *Proceedings of the National Academy of Sciences* 84(23):8512–8514.
- 466 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution of social
467 behavior in structured populations. *Journal of Theoretical Biology* 420:26 – 35.
- 468 Débarre, F.; Hauert, C. & Doebeli, M. 2014: Social evolution in structured populations.
469 *Nature Communications* 5.
- 470 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evolution of
471 altruism. *Proceedings of the Royal Society B: Biological Sciences* 276(1654):13–19.
- 472 Frank, S. A. 1997: The Price equation, Fisher’s fundamental theorem, kin selection, and
473 causal analysis. *Evolution* 51(6):1712–1729.
- 474 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transactions of*
475 *the Royal Society of London B: Biological Sciences* 365(1552):2515–2522.
- 476 Grafen, A. & Archetti, M. 2008: Natural selection of altruism in inelastic viscous homoge-
477 neous populations. *Journal of Theoretical Biology* 252(4):694 – 710.
- 478 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of Theoretical*
479 *Biology* 7(1):1 – 16.

- 480 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolutionary
481 genetics. *Biosocial anthropology* 53:133–55.
- 482 Hammerschmidt, K.; Rose, C. J.; Kerr, B. & Rainey, P. B. 2014: Life cycles, fitness decou-
483 pling and the evolution of multicellularity. *Nature* 515(7525):75–79.
- 484 Kandori, M.; Mailath, G. J. & Rob, R. 1993: Learning, mutation, and long run equilibria in
485 games. *Econometrica* 61(1):29–56.
- 486 Le Galliard, J.; Ferrière, R. & Dieckmann, U. 2005: Adaptive evolution of social traits: Ori-
487 gin, trajectories, and correlations of altruism and mobility. *The American Naturalist*
488 165(2):206–224. PMID: 15729651.
- 489 Lehmann, L.; Feldman, M. W. & Foster, K. R. 2008: Cultural transmission can inhibit
490 the evolution of altruistic helping. *The American Naturalist* 172(1):12–24. PMID:
491 18500938.
- 492 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and harming
493 on graphs: the return of the inclusive fitness effect. *Journal of Evolutionary Biology*
494 20(6):2284–2295.
- 495 Lehmann, L. & Rousset, F. 2014: The genetical theory of social behaviour. *Philosophical*
496 *Transactions of the Royal Society of London B: Biological Sciences* 369(1642).
- 497 Leturque, H. & Rousset, F. 2002: Dispersal, kin competition, and the ideal free distribu-
498 tion in a spatially heterogeneous population. *Theoretical Population Biology* 62(2):169
499 – 180.
- 500 Lion, S. 2016: Moment equations in spatial evolutionary ecology. *Journal of theoretical*
501 *biology* 405:46–57.
- 502 Mullon, C.; Keller, L. & Lehmann, L. 2017: Co-evolution of dispersal with behaviour
503 favours social polymorphism. *bioRxiv* .
- 504 Mullon, C. & Lehmann, L. 2014: The robustness of the weak selection approximation
505 for the evolution of altruism against strong selection. *Journal of evolutionary biology*
506 27(10):2272–2282.
- 507 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule for the evolu-
508 tion of cooperation on graphs and social networks. *Nature* 441(7092):502–505.
- 509 Parvinen, K. 2013: Joint evolution of altruistic cooperation and dispersal in a metapopu-
510 lation of small local populations. *Theoretical population biology* 85:12–19.
- 511 Platt, T. G. & Bever, J. D. 2009: Kin competition and the evolution of cooperation. *Trends*
512 *in Ecology & Evolution* 24(7):370 – 377.
- 513 Rodrigues, A. M. M. & Gardner, A. 2012: Evolution of helping and harming in heteroge-
514 neous populations. *Evolution* 66(7):2065–2079.
- 515 Rousset, F. 2004: *Genetic Structure and Selection in Subdivided Populations*. Princeton
516 University Press, Princeton, NJ.

- 517 Rousset, F. & Billiard, S. 2000: A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *Journal of Evolutionary Biology* 13(5):814–825.
- 520 Sample, C. & Allen, B. 2017: The limits of weak selection and large population size in evolutionary game theory. *Journal of mathematical biology* pages 1–33.
- 522 Schneider, D. M.; Martins, A. B. & de Aguiar, M. A. 2016: The mutation–drift balance in spatially structured populations. *Journal of theoretical biology* 402:9–17.
- 524 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors in finite structured population models. *The American Naturalist* 184(4):477–488.
- 526 Taylor, P. 1992a: Altruism in viscous populations—an inclusive fitness model. *Evolutionary ecology* 6(4):352–356.
- 528 Taylor, P. 2010: Birth–death symmetry in the evolution of a social trait. *Journal of Evolutionary Biology* 23(12):2569–2578.
- 530 Taylor, P.; Lillicrap, T. & Cownden, D. 2011: Inclusive fitness analysis on mathematical groups. *Evolution* 65(3):849–859.
- 532 Taylor, P. D. 1992b: Inclusive fitness in a homogeneous environment. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 249(1326):299–302.
- 534 Taylor, P. D.; Day, T. & Wild, G. 2007a: Evolution of cooperation in a finite homogeneous graph. *Nature* 447(7143):469–472.
- 536 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation probability in homogeneous structured populations. *Journal of Theoretical Biology* 249(1):101 – 110.
- 538 Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic behavior. *Evolution* 54(4):1135–1141.
- 540 Traulsen, A.; Hauert, C.; De Silva, H.; Nowak, M. A. & Sigmund, K. 2009: Exploration dynamics in evolutionary games. *Proceedings of the National Academy of Sciences* 106(3):709–712.
- 543 Van Cleve, J. 2015: Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology* 103:2 – 26.
- 545 West, S. A. & Gardner, A. 2010: Altruism, spite, and greenbeards. *Science* 327(5971):1341–1344.
- 547 West, S. A.; Pen, I. & Griffin, A. S. 2002: Cooperation and competition between relatives. *Science* 296(5565):72–75.
- 549 Wild, G. & Traulsen, A. 2007: The different limits of weak selection and the evolutionary dynamics of finite populations. *Journal of Theoretical Biology* 247(2):382 – 390.
- 551 Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely viscous populations? *Evolutionary Ecology* 6(4):331–341.
- 553 Wolfram Research, Inc. 2017: *Mathematica*, Version 11.1. Champaign, IL, 2017.

554 **Figures**

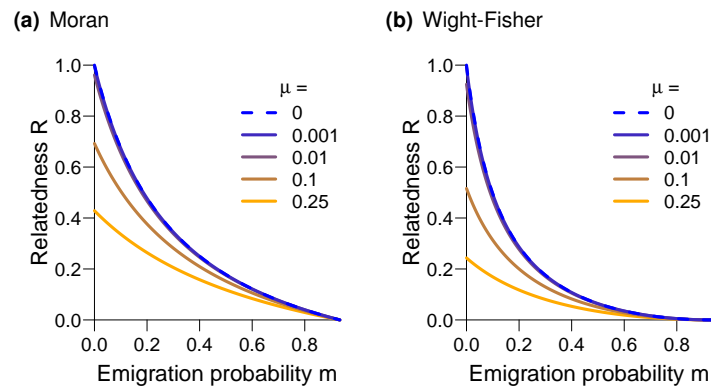


Figure 1: Within-deme relatedness of pairs of individuals R , as a function of the emigration probability m , for different values of the mutation probability μ (from 0 [blue] to 0.25 [orange]), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals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

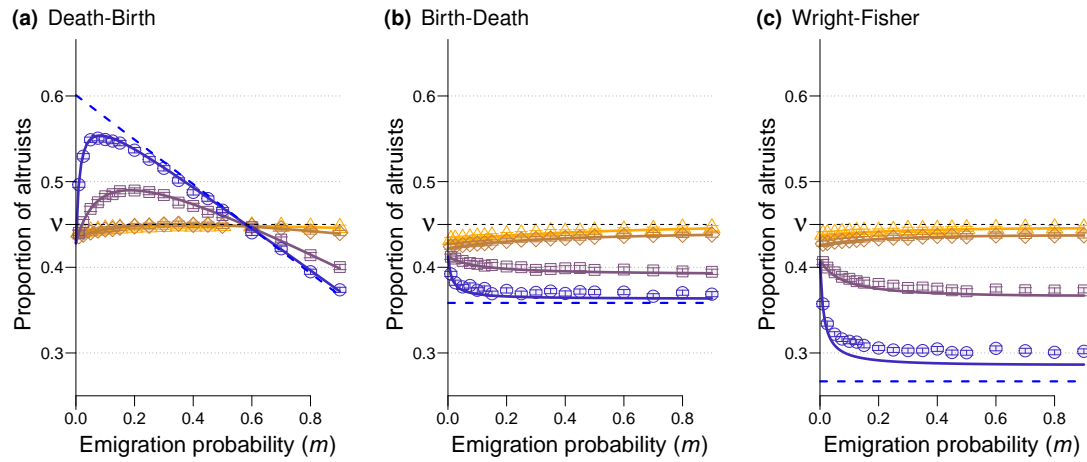


Figure 2: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different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

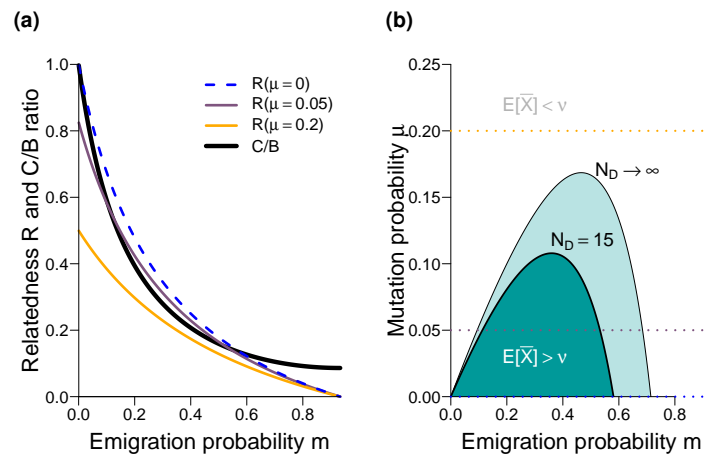


Figure 3: Understanding the effect of emigration m on whether altruism is favored in the Death-Birth life-cycle. (a) Comparison of the C/B ratio (thick black curve) and relatedness R (thin curves) for different values of the mutation probability μ (same color code as previously). (b) (m, μ) combinations for which $\mathbb{E}[\bar{X}] > v$. 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.

555 **Supplementary figures**

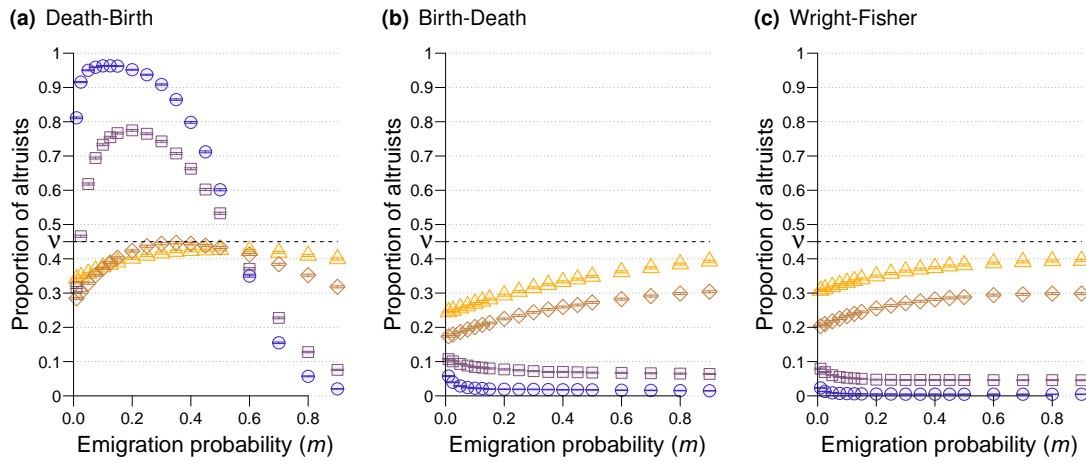


Figure A1: Equivalent of figure 2 (simulations only) but with strong selection ($\delta = 0.1$); please note the change of scale on the vertical axis. All other parameters and legends are identical to those of figure 2 (increasing mutation probabilities from blue dots to orange triangles).

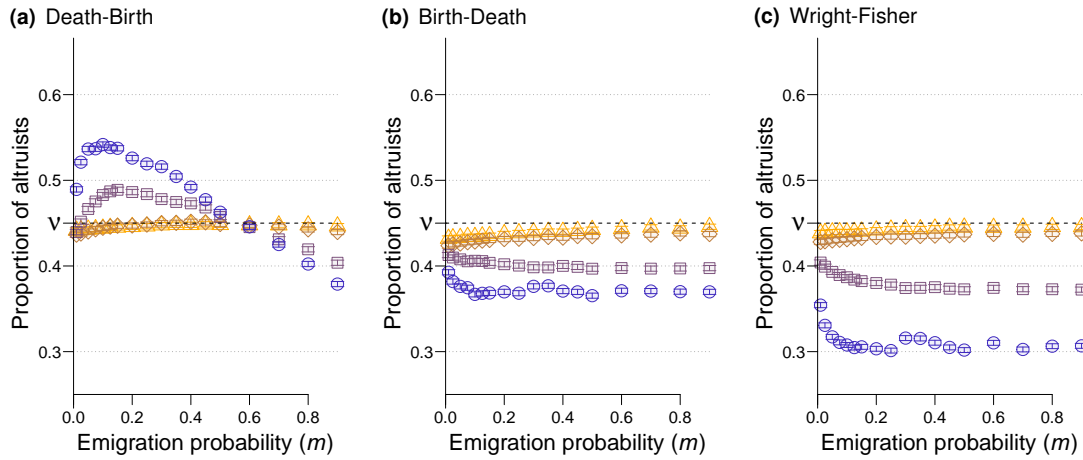


Figure A2: Equivalent of figure 2 (simulations only) but with a heterogeneous population structure: deme sizes range from 1 to 5 individuals per deme, the average deme size is 4 as in figure 2; all other parameters and legend are identical to those of figure 2.

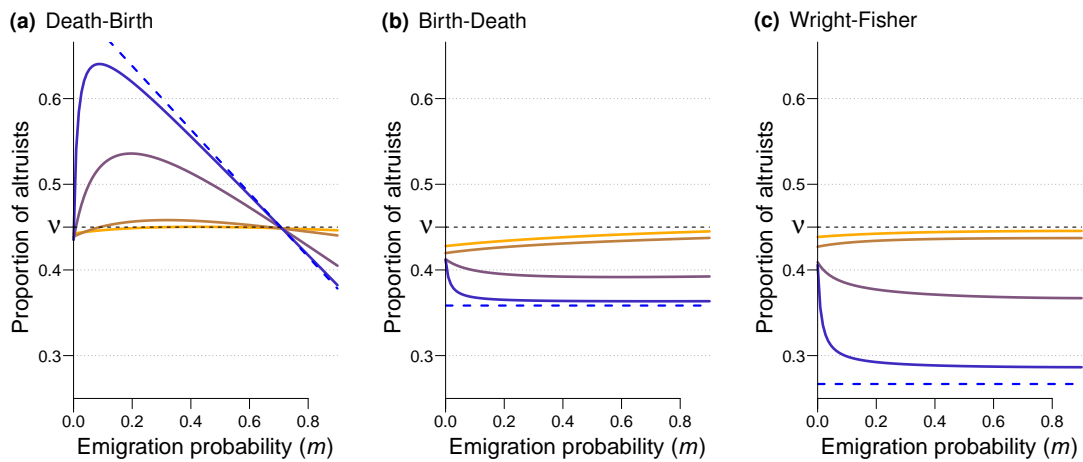


Figure A3: Equivalent of figure 2 (analysis only), with no self-replacement ($d_{ii} = d_{\text{self}} = 0$ for all sites).

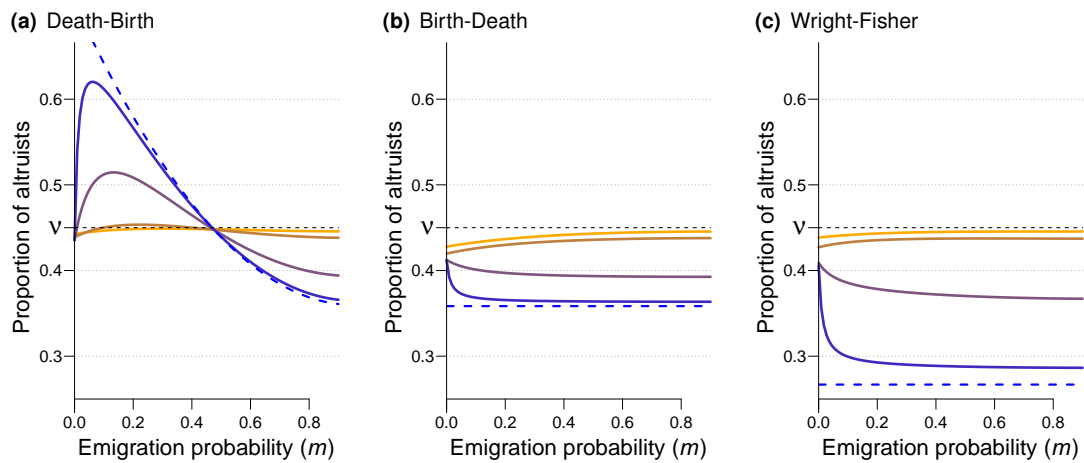


Figure A4: Equivalent of figure 2 (analysis only), with equal dispersal and interaction graphs (*i.e.*, no self-replacement [$d_{ii} = d_{\text{self}} = 0$ for all sites], and a proportion m of the interactions occurring outside of the home deme).

556 **Supplementary Table**

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

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

557

Appendix

558 A Mutation parameters

559 In the main text, we first introduce effective mutation parameters: $\mu_{1 \rightarrow 0}$, the probabil-
560 ity that an altruist has defector offspring, and $\mu_{0 \rightarrow 1}$, the probability that a defector has
561 altruist offspring.

562 A.1 Expected frequency of altruists at the mutation-drift balance

563 We assume that there is no selection acting ($\delta = 0$), but that there still are two types of
564 individuals in the population.

565 Let Y be the type of a randomly chosen individual ($Y = 1$ if the individual is an altruist
566 and $Y = 0$ if it is a defector) in the population, given a proportion y of altruists in the
567 population. In expectation, we have

$$\mathbb{E}[Y] = y. \quad (\text{A1a})$$

568 Let Y' be the type of a randomly chosen individual at the next time step, given the fre-
569 quency y at the previous time step. This randomly chosen individual is altruist if its par-
570 ent was (which happens with probability y) and it did not mutate (probability $1 - \mu_{1 \rightarrow 0}$),
571 or if its parent was not altruist (probability $1 - y$), but the offspring mutated into one
572 (probability $\mu_{0 \rightarrow 1}$). We obtain

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

573 The expected frequency of altruists at the mutation-drift balance, denoted by v , is found
574 by solving $\mathbb{E}[Y] = \mathbb{E}[Y']$. We obtain

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

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

576 We can then compute the parent-offspring type correlation at the mutation-drift bal-
577 ance. First, let us compute the parent-offspring covariance:

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

578 Remember that Y and Y' are indicator variables and therefore take value in $\{0, 1\}$, so
579 that $Y^2 = Y$ (likewise for Y'). Then, the standard deviations are given by

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

580 and

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

581 Finally, the parent-offspring correlation is given by

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

582 using the formulas eq. (A3)–(A5), and replacing ν by its value (mutation-drift equilib-
583 rium, eq. (A2)), we obtain

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

584 A.3 Redefining the mutation scheme

585 With the new mutation parameters μ and ν , we can describe the mutation scheme dif-
586 ferently.

587 If we denote by X_i the type of a given parent, then the expected type of one of its
588 offspring is

$$\begin{aligned}\mathbb{E}[X'_i|X_i] &= X_i(1 - \mu_{1\rightarrow 0}) + (1 - X_i)\mu_{0\rightarrow 1} \\ &= X_i(1 - (\mu_{1\rightarrow 0} + \mu_{0\rightarrow 1})) + \mu_{0\rightarrow 1}.\end{aligned}\tag{A7a}$$

589 Replacing $\mu_{1\rightarrow 0}$ and $\mu_{0\rightarrow 1}$ by equivalent combinations of μ and ν as defined in eq. (A6)
590 and eq. (A2), *i.e.*,

$$\mu_{1\rightarrow 0} = \mu(1 - \nu) \text{ and } \mu_{0\rightarrow 1} = \mu\nu,\tag{A7b}$$

591 then eq. (A7a) becomes

$$\mathbb{E}[X'_i|X_i] = X_i(1 - \mu) + \mu\nu.\tag{A7c}$$

592 We can redefine the mutation scheme and interpret eq. (A7c) as follows. Parents transmit
593 their strategy to their offspring with probability $1 - \mu$; with probability μ , offspring do not
594 inherit their strategy from their parent but instead get one randomly: with probability ν ,
595 they become altruists, with probability $1 - \nu$ they become defectors. With this alternative
596 description, we can call “mutants” individuals who have the same type as their parent.

597 B Expected frequency of altruists

598 B.1 For a generic life cycle

599 We want to compute the expected proportion of altruists in the population. We represent
 600 the state of the population at a given time t using indicator variables $X_i(t)$, $1 \leq i \leq N$,
 601 equal to 1 if the individual living at site i at time t is an altruist, and equal to 0 if it is
 602 a defector; these indicator variables are gathered in a N -long vector $\mathbf{X}(t)$. The set of all
 603 possible population states is $\Omega = \{0, 1\}^N$. The proportion of altruists in the population is
 604 written $\bar{X}(t) = 1/N \sum_{i=1}^N X_i(t)$. We denote by $B_{ji}(\mathbf{X}(t), \delta)$, written B_{ji} for simplicity, the
 605 probability that the individual at site j at time $t + 1$ is the newly established offspring
 606 of the individual living at site i at time t . The expected number of successful offspring
 607 produced by the individual living at site i at time t is given by $B_i = \sum_{j=1}^N B_{ji}$. We denote
 608 by $D_i(\mathbf{X}(t), \delta)$ (D_i for simplicity) the probability that the individual living at site i at time
 609 t has been replaced (*i.e.*, died) at time $t + 1$. These quantities depend on the chosen life
 610 cycle and on the state of the population; they are given in table A2 for each of the life
 611 cycles that we consider.

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

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

612 Since a dead individual is immediately replaced by one new individual (*i.e.*, popula-
 613 tion size remains constant and equal to N),

$$D_i = \sum_{j=1}^N B_{ij} \tag{A8a}$$

614 holds for all sites i and all life cycles.

615 The structure of the population is also such that in the absence of selection ($\delta = 0$, so
 616 that $f_i = 1$ for all sites $1 \leq i \leq N$), all individuals have the same probability of dying and
 617 the same probability of having successful offspring (*i.e.*, of having offspring that become
 618 adults at the next time step), so that

$$D_i^0 = \sum_{j=1}^N B_{ji}^0 = B_i^0 =: B^*, \tag{A8b}$$

619 where the 0 subscript means that the quantities are evaluated for $\delta = 0$. This also implies
 620 that B_i^0 and D_i^0 do not depend on the state \mathbf{X} of the population. For the Moran life cy-
 621 cles, $B_i^* = 1/N$, while for the Wright-Fisher life cycle, $B_i^* = 1$. (The difference between
 622 eq. (A8b) and eq. (A8a) is that we are now considering offspring produced by i landing
 623 on j).

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

$$\mathbb{E}[\bar{X}(t+1)|\mathbf{X}(t)] = \frac{1}{N} \sum_{i=1}^N [B_i(1-\mu)X_i + (1-D_i)X_i + B_i\mu\nu]. \quad (\text{A9a})$$

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

630 A lost strategy can always be created again by mutation, so there is no absorbing
 631 population state. There exists a stationary distribution of population states (Theorem 1
 632 in Allen & Tarnita (2014)). In other words, for large times t , the expected frequency of
 633 altruists does not change anymore (of course, realized frequencies keep changing over
 634 time). We denote by $\xi(\mathbf{X}, \delta, \mu)$ the probability that the population is in state \mathbf{X} , given
 635 the strength of selection δ and the mutation probability μ . Taking the expectation of
 636 eq. (A9a) ($\mathbb{E}[\bar{X}] = \sum_{\mathbf{X} \in \Omega} \bar{X} \xi(\mathbf{X}, \delta, \mu)$), we obtain, after reorganizing:

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

637 Now, we use the assumption of weak selection ($\delta \ll 1$) and consider the first-order
 638 expansion of eq. (A10) for δ close to 0.

$$\begin{aligned} 0 &= \frac{1}{N} \sum_{\mathbf{X} \in \Omega} \left[\sum_{i=1}^N (B_i^0(1-\mu)X_i - D_i^0 X_i) + \sum_{i=1}^N B_i^0 \mu \nu \right] \xi(\mathbf{X}, 0, \mu) \\ &+ \frac{\delta}{N} \sum_{\mathbf{X} \in \Omega} \left[\sum_{i=1}^N \left(\frac{\partial B_i(1-\mu) - D_i}{\partial \delta} X_i \right) + \sum_{i=1}^N \frac{\partial B_i}{\partial \delta} \mu \nu \right] \xi(\mathbf{X}, 0, \mu) \\ &+ \frac{\delta}{N} \sum_{\mathbf{X} \in \Omega} \left[\sum_{i=1}^N (B_i^0(1-\mu)X_i - D_i^0 X_i) + \sum_{i=1}^N B_i^0 \mu \nu \right] \frac{\partial \xi(\mathbf{X}, \delta, \mu)}{\partial \delta} + O(\delta^2), \end{aligned} \quad (\text{A11})$$

639 where all the derivatives are evaluated for $\delta = 0$. The first line of eq. (A11) is equal to
 640 zero, because $B_i^0 - D_i^0 = 0$ (eq. (A8b)), and because in the absence of selection ($\delta = 0$),
 641 the expected state of every site i is $\mathbb{E}_0[X_i] = \sum_{\mathbf{X} \in \Omega} X_i \xi(\mathbf{X}, 0, \mu) = \nu$ (by definition of ν ,
 642 see Appendix A.1). The second term of the second line is zero, because for all the life
 643 cycles that we consider, the total number of births in the population during one time
 644 step ($\sum_{i=1}^N B_i$) does not depend on population phenotypic composition (it is exactly 1
 645 death for the Moran life cycles, and exactly N for the Wright-Fisher life cycle); since it is
 646 a constant, its derivative is 0. The third line simplifies by noting again that $B_i^0 = D_i^0$ (first

647 term), and that $\sum_{X \in \Omega} \frac{\partial \xi(\mathbf{X}, \delta, \mu)}{\partial \delta} = 0$ since ξ is a probability distribution (so the second term
648 is zero). Eq. (A11) then becomes

$$0 = \frac{\delta}{N} \sum_{i=1}^N \left[\sum_{X \in \Omega} \left(\frac{\partial B_i}{\partial \delta} (1 - \mu) - \frac{\partial D_i}{\partial \delta} \right) X_i \xi(\mathbf{X}, 0, \mu) - \sum_{X \in \Omega} \mu B^* X_i \frac{\partial \xi}{\partial \delta} \right] + O(\delta^2), \quad (\text{A12})$$

649 where the derivatives are evaluated at $\delta = 0$. For conciseness, we define

$$W_i = (1 - \mu) B_i + (1 - D_i), \quad (\text{A13})$$

650 a measure of fitness counting offspring only when they are unmutated (in the sense of
651 the alternate mutation scheme described in Appendix A.3). With this, using the expect-
652 ation notation, and denoting by $\mathbb{E}_0[\cdot]$ expectations under $\delta = 0$, we can rewrite and reor-
653 ganize eq. (A12) as

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \mathbb{E}_0 \left[\frac{\partial W_i}{\partial \delta} X_i \right] + O(\delta^2). \quad (\text{A14})$$

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

$$\begin{aligned} \frac{\partial W_i}{\partial \delta} &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} \frac{\partial \phi_k}{\partial \delta} \\ &= \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} X_k, \end{aligned} \quad (\text{A15})$$

656 by definition of ϕ_k ($\phi_k = \delta X_k$), and where the derivatives are evaluated for all $\phi_i = 0$,
657 $1 \leq i \leq N$. Introducing the notation $P_{ij} = \mathbb{E}_0[X_i X_j]$ (expected state of a pair of sites),
658 eq. (A14) becomes

$$\delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} P_{ik} + O(\delta^2). \quad (\text{A16})$$

659 We note that $P_{ii} = \mathbb{E}_0[X_i X_i] = \mathbb{E}_0[X_i] = \nu$ (X_i being an indicator variable, it is either equal
660 to 0 or 1, so $X_i^2 = X_i$). Given that the size of the population is fixed ($\sum_{i=1}^N (B_i - D_i) = 0$),
661 and given that the total number of births does not depend on population composition
662 in the life cycles that we consider, we have

$$\sum_{i=1}^N \frac{\partial W_i}{\partial \delta} \Big|_{\delta=0} = 0. \quad (\text{A17a})$$

663 Using the decomposition in eq. (A15), which is valid for any population composition,
664 and so in particular for $\mathbf{X} = \mathbf{1}$, eq. (A17a) becomes

$$\sum_{i=1}^N \sum_{k=1}^N \frac{\partial W_i}{\partial \phi_k} \Big|_{\phi=0} = 0. \quad (\text{A17b})$$

665 So far, we have not used the specificities of the population structure that we consider.
666 First, the population is homogeneous (*sensu* Taylor et al., 2007a). Because this popula-
667 tion homogeneity, eq. (A17b) is valid for all i (not just their sum). Secondly, we are con-

668 sidering an island model. Once we have fixed a focal individual i , in expectation there
 669 are only three types of individuals: the focal itself (denoted by “•”), $n - 1$ other individu-
 670 als in the focal’s deme (denoted by “in”), and $N - n$ individuals in other demes (denoted
 671 by “out”). With these considerations, eq. (A17b) becomes

$$\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} + (N - n) \frac{\partial W_i}{\partial \phi_{\text{out}}} = 0. \quad (\text{A17c})$$

672 (as previously shown by (Rousset & Billiard, 2000, p.817–818)). Using this island model-
 673 specific notation, eq. (A16) becomes

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

674 Injecting eq. (A17c) into eq. (A16), we obtain

$$\delta \mu_B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\frac{\partial W_i}{\partial \phi_i} + (n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}} \frac{P_{\text{in}} - P_{\text{out}}}{P_{ii} - P_{\text{out}}} \right) (P_{ii} - P_{\text{out}}) + O(\delta^2). \quad (\text{A18})$$

675 We can also replace the P terms as follows:

$$\begin{aligned} P_{ij} &= Q_{ij} \nu + (1 - Q_{ij}) \nu^2 \\ &= \nu^2 + \nu(1 - \nu) Q_{ij}. \end{aligned} \quad (\text{A19})$$

676 In Appendix C.1, using recursions on P_{ij} , we will see that Q_{ij} can be interpreted as a
 677 probability of identity by descent, *i.e.*, the probability that the individuals at sites i and j
 678 have a common ancestor and that no mutation (using the alternative mutation scheme
 679 described in Appendix A.3) has occurred on either lineage since the ancestor. Replacing
 680 the P terms with eq. (A19), and noting that $Q_{ii} = 1$, eq. (A18) becomes

$$\delta \mu_B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} = \frac{\delta}{N} \sum_{i=1}^N \left(\underbrace{\frac{\partial W_i}{\partial \phi_i}}_{-C} + \underbrace{(n - 1) \frac{\partial W_i}{\partial \phi_{\text{in}}}}_B \underbrace{\frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}}_R \right) (1 - Q_{\text{out}}) \nu(1 - \nu) + O(\delta^2). \quad (\text{A20})$$

681 We can further decompose the derivatives, now using the fecundities f_ℓ as interme-
 682 diate variables, *i.e.*,

$$\frac{\partial W_i}{\partial \phi_k} = \sum_{\ell=1}^N \frac{\partial W_i}{\partial f_\ell} \frac{\partial f_\ell}{\partial \phi_k}. \quad (\text{A21})$$

683 The term $\frac{\partial f_\ell}{\partial \phi_k}$ is the marginal effect of a change in the phenotype of the individual
 684 living at site k on the fecundity of the individual living at site ℓ . By assumption, social
 685 interactions take place within demes only, so whenever sites ℓ and k are in different
 686 demes, we have $\frac{\partial f_\ell}{\partial \phi_k} = \frac{\partial f_\ell}{\partial \phi_{\text{out}}} = 0$. We then need to characterize the effect of one’s own
 687 phenotype (*i.e.*, $k = \ell$) and of another deme-mate’s phenotype (k and ℓ being different

688 sites in the same deme) on fecundity. For this, we define b and c so that:

$$\left. \frac{\partial f_\ell}{\partial \phi_\ell} \right|_{\delta=0} = -c, \quad (\text{A22a})$$

$$\left. \frac{\partial f_\ell}{\partial \phi_{in}} \right|_{\delta=0} = \frac{b}{n-1}. \quad (\text{A22b})$$

689 Eq. (A20) then becomes (using notation \bullet to refer to the focal individual itself, and where
690 $W = W_i$, since the derivatives are the same for all i):

$$\begin{aligned} \delta \mu B^* \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} &= \delta \nu (1 - \nu) (1 - Q_{out}) \times \\ &\left(\underbrace{\frac{\partial W}{\partial f_\bullet} (-c) + \frac{\partial W}{\partial f_{in}} b}_{-c} + \underbrace{\left(\frac{\partial W}{\partial f_\bullet} b + (n-1) \frac{\partial W}{\partial f_{in}} (-c) + (n-2) \frac{\partial W}{\partial f_{in}} b \right)}_{\mathcal{B}} \underbrace{\frac{Q_{in} - Q_{out}}{1 - Q_{out}}}_{\mathcal{R}} \right) + O(\delta^2). \end{aligned} \quad (\text{A23})$$

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

692 Finally, we write a first-order approximation of the expected frequency of altruists in
693 the population:

$$\mathbb{E}[\bar{X}] = \mathbb{E}_0[\bar{X}] + \delta \left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0} + O(\delta^2). \quad (\text{A24})$$

694 The first term, $\mathbb{E}_0[\bar{X}]$, is the expected frequency in the absence of selection; it is equal
695 to ν (as introduced in eq. (A2)). The derivative $\left. \frac{\partial \mathbb{E}[\bar{X}]}{\partial \delta} \right|_{\delta=0}$ is obtained from eq. (A23). We
696 then need to replace the B_i and D_i terms by their formulas for each life cycle; they are
697 given in table A2. This is how the expected frequency of altruists in the population is
698 approximated.

699 B.2 Derivatives for the specific life cycles

700 We use the formulas presented in table A2 and the definition of $W = W_i$ given in eq. (A13)
701 for each life cycle. In eq. (A26), eq. (A28) and eq. (A30), the first lines within parentheses
702 correspond to primary effects, and the second line to secondary effects.

703 **Moran Birth-Death** Under this life cycle, we obtain

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

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

704 With these derivatives, eq. (5) becomes

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{+(b-c)\left(\frac{\mu}{N} - \frac{1-m}{n}\right)} \right)}_{-C^{\text{BD}}} + \underbrace{\left(\frac{(1-\mu)b}{+(b-c)(n-1)\left(\frac{\mu}{N} - \frac{1-m}{n}\right)} \right)}_{B^{\text{BD}}} \underbrace{\left(\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M} \right)}_{R^M} \right], \quad (\text{A26})$$

705 In addition, for both Moran life cycles, we have $B_M^* = 1/N$. The secondary effects (sec-
706 ond line in the parentheses in eq. (A26)) include competitive effects on the probability
707 of reproducing, and consequences of social interactions on the probability that a given
708 individual dies. Note that the secondary effects remain negative for the realistic range of
709 emigration values that we consider (*i.e.*, $m < 1 - 1/N_D$).

710 **Moran Death-Birth** Under this life cycle, we obtain

$$\left. \frac{\partial W^{\text{DB}}}{\partial f} \right|_{\delta=0} = \frac{1-\mu}{N} \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A27a})$$

$$\left. \frac{\partial W^{\text{DB}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -\frac{1-\mu}{N} \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A27b})$$

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

$$\mathbb{E}[\bar{X}] \approx v + \frac{\delta}{\mu} v(1-v)(1-Q_{\text{out}}^M) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{-(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{-C^{\text{DB}}} + \underbrace{\left(\frac{(1-\mu)b}{-(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{B^{\text{DB}}} \underbrace{\left(\frac{Q_{\text{in}}^M - Q_{\text{out}}^M}{1 - Q_{\text{out}}^M} \right)}_{R^M} \right], \quad (\text{A28})$$

712 With this life cycle, Death occurs first, and the probability of dying is independent from
713 the state of the population (since we assume that social interactions affect fecundity. We
714 can therefore factor $(1-\mu)$ in all terms. The primary effects (first lines in the parentheses)
715 remain the same as with the Birth-Death life cycle. However, the Death-Birth life cycle
716 leads to different secondary effects compared to the Birth-Death life cycle: competition
717 occurs at a different scale (Grafen & Archetti, 2008). Finally, with this life cycle as we
718 defined it, the probabilities of identity by descent Q are the same as with the Birth-Death
719 model.

720 **Wright-Fisher** Under this life cycle, we obtain

$$\left. \frac{\partial W^{\text{WF}}}{\partial f} \right|_{\delta=0} = (1-\mu) \left[1 - \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right) \right], \quad (\text{A29a})$$

$$\left. \frac{\partial W^{\text{WF}}}{\partial f_{\text{in}}} \right|_{\delta=0} = -(1-\mu) \left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n} \right). \quad (\text{A29b})$$

721 For the Wright-Fisher life cycle, we have $B_{WF}^* = 1$. Replacing the derivatives presented in
 722 eq. (A29) into eq. (5), we obtain

$$\mathbb{E}[\bar{X}] \approx \nu + \frac{\delta}{\mu} \nu(1-\nu)(1-Q_{out}^{WF}) \times \left[\underbrace{\left(\frac{(1-\mu)(-c)}{-(b-c)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{-C^{WF}} + \underbrace{\left(\frac{(1-\mu)b}{-(b-c)(n-1)(1-\mu)\left(\frac{(1-m)^2}{n} + \frac{m^2}{N-n}\right)} \right)}_{B^{WF}} \underbrace{\left(\frac{Q_{in}^{WF} - Q_{out}^{WF}}{1 - Q_{out}^{WF}} \right)}_{R^{WF}} \right], \quad (A30)$$

723 The only – but important – difference between eq. (A30) and eq. (A28) is the value of
 724 the probabilities of identity by descent Q , because the number of individuals that are
 725 updated at each time step differs.

726 C Probabilities of identity by descent

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

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

731 C.1.1 Moran model

732 These calculations apply to both the Death-Birth and Birth-Death updating rules.
733 In a Moran model, exactly one individual dies and one individual reproduces during one
734 time step. Given a state \mathbf{X} at time t , at time $t + 1$ both sites i and $j \neq i$ are occupied by
735 altruists, if i it was the case at time t and neither site was replaced by a non-altruist (first
736 term in eq. (A31)), or ij if exactly one of the two sites was occupied by a non-altruist at
737 time t , but the site was replaced by an altruist (second and third terms of eq. (A31)):

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1)|X(t) = \mathbf{X}] &= X_i X_j \left(1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - \nu)) \right) \\ &+ X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu\nu) \\ &+ X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu\nu). \end{aligned} \quad (\text{A31})$$

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

$$P_{ij} = \frac{1}{2} \left(\sum_{k=1}^N (1 - \mu) (d_{kj} P_{ki} + d_{ki} P_{kj}) \right) + \mu\nu^2 \quad (i \neq j), \quad (\text{A32})$$

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

742 Now we substitute $P_{ij} = \nu^2 + \nu(1 - \nu)Q_{ij}$ in eq. (A32), we obtain

$$Q_{ij} = \frac{1}{2} \sum_{k=1}^N (1 - \mu) (d_{ki} Q_{kj} + d_{kj} Q_{ki}), \quad (\text{A33})$$

743 and we realize that Q_{ij} is the probability that the individuals at sites i and $j \neq i$ are iden-
744 tical by descent (e.g., Taylor et al. (2011), equation above (S1.11); Allen & Nowak (2014)
745 eq. (4)). To compute it indeed, we need to pick which site was last updated (i or j with
746 equal probabilities: $1/2$), then sum over the possible parent (k); the other individual
747 needs to be identical by descent to the parent (Q_{kj} , Q_{ki}), disperse to the considered site
748 (d_{ki} , d_{kj}), and no mutation should have occurred ($1 - \mu$).

749 **C.1.2 Wright-Fisher model**

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & \sum_{k,\ell=1}^N d_{ki} d_{\ell j} \left(X_k X_\ell (1 - \mu + \mu\nu)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu\nu)(\mu\nu) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu\nu)^2 \right) \end{aligned} \quad (\text{A34})$$

752 The first term of eq. (A34) corresponds to both parents being altruists, and having altruist
753 offspring; the second line corresponds to exactly one parent being altruist, and the third
754 line to both parents being non-altruists (in this latter case, the two offspring have to be
755 both mutants to be altruists).

756 Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k,\ell=1}^N (P_{kl}(1 - \mu)^2) + (2 - \mu)\mu\nu^2. \quad (\text{A35})$$

757 Replacing P_{ij} by $\nu^2 + \nu(1 - \nu)Q_{ij}$, eq. (A35) becomes

$$Q_{ij} = \sum_{k,\ell=1}^N d_{ki} d_{\ell j} Q_{k\ell} (1 - \mu)^2. \quad (\text{A36})$$

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

760 C.2 Probabilities of identity by descent in a subdivided population

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

$$Q_{ij} = \begin{cases} 1 & \text{when } i = j; \\ Q_{\text{in}} & \text{when } i \neq j \text{ and both sites are in the same deme;} \\ Q_{\text{out}} & \text{when sites } i \text{ and } j \text{ are in different demes.} \end{cases} \quad (\text{A37})$$

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

766 When the number of demes is infinite, Q_{in} is relatively easily obtained using recur-
 767 rence equations and noting that $Q_{\text{out}} = 0$. However, writing the recurrence equations for
 768 Q_{in} and Q_{out} is much more tedious for finite populations. Hence, for finite populations,
 769 we will use formulas already derived in Débarre (2017) for “two-dimensional population
 770 structures”. The name comes from the fact that we only need two types of transforma-
 771 tions to go from any site to any other site in the population: permutations on the deme
 772 index, and permutations on the within-deme index.

773 We rewrite site labels ($1 \leq i \leq N$) as (ℓ_1, ℓ_2) , where ℓ_1 is the index of the deme ($1 \leq \ell_1 \leq$
 774 N_D) and ℓ_2 the position of the site within the deme ($1 \leq \ell_2 \leq n$). Then, we introduce
 775 notations \tilde{d}_{i_1, i_2} and \tilde{Q}_{i_1, i_2} , that correspond to the dispersal probability and probability of
 776 identity by descent to a site at distances i_1 and i_2 in the among-demes and within-deme
 777 dimensions (e.g., $\tilde{d}_{i_1, i_2} = d_{j_1, j_2}$ with $j_1 = i_1$ and $j_2 = i_2$).

778 Also, in this section, we distinguish between $d_{\text{self}} = d_{ii}$ and d_{in} (in the main text,
 779 $d_{\text{self}} = d_{\text{in}}$).

780 C.2.1 Moran model

781 In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1, r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1, q_2}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{A38a})$$

782 with

$$\tilde{D}_{q_1, q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1, \ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{A38b})$$

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

$$\begin{aligned}\tilde{D}_{q_2} &= d_{\text{self}} + \sum_{\ell_2=1}^{N_2-1} d_{\text{in}} \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) + \sum_{\ell_1=1}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} d_{\text{out}} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2}(N_2 - 1) + (1 - \delta_{q_2})(-1)) d_{\text{in}} + (\delta_{q_1}(N_1 - 1) + (1 - \delta_{q_1})(-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}.\end{aligned}\tag{A39a}$$

785 (δ_q is equal to 1 when q is equal to 0 modulo the relevant dimension, and to 0 other-
 786 wise). So for the three types of distances that we need to consider (distance 0, distance
 787 to another deme-mate, distance to individual in another deme), and with $N_1 = N_D$ and
 788 $N_2 = n$, we obtain

$$\tilde{D}_0 = 1, \tag{A40a}$$

$$\tilde{D}_0^{q_1} = 1 - m - \frac{m}{N_D - 1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \tag{A40b}$$

$$\tilde{D}_{q_2}^{q_1} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \tag{A40c}$$

789 So for \tilde{Q} , using system (A40) in eq. (A38a),

$$\begin{aligned}\tilde{Q}_{r_2} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu) \tilde{D}_0} + \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_0^{q_2}} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right. \\ &\quad + \sum_{q_1=1}^{N_1-1} \frac{1}{1 - (1 - \mu) \tilde{D}_0^{q_1}} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_2}^{q_1}} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu)} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_2} N_2 - 1) \right. \\ &\quad + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (\delta_{r_1} N_1 - 1) \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_1} N_1 - 1)(\delta_{r_2} N_2 - 1) \right].\end{aligned}\tag{A41}$$

790 In particular,

$$\begin{aligned}\tilde{Q}_0 &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (n - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})} (N_D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (N_D - 1)(n - 1) \right] \\ &= 1.\end{aligned}\tag{A42a}$$

791 We find λ'_M using eq. (A42a). Let's now go back to eq. (A41): when $r_1 = 0$, the two indi-
 792 viduals are in the same deme. The two individuals are different when $r_2 \neq 0$, and so:

$$Q_{\text{in}} = \frac{\mu\lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})}(D - 1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(D - 1)(-1) \right]. \quad (\text{A42b})$$

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

$$Q_{\text{out}} = \frac{\mu\lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{N_D - 1})}(-1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} \right]. \quad (\text{A42c})$$

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

$$Q_{\text{in}}^M = \frac{(1 - \mu)(m + \mu(N_D(1 - m) - 1))}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}, \quad (\text{A43a})$$

$$Q_{\text{out}}^M = \frac{(1 - \mu)m}{(1 - \mu)m(N_D\mu(n - 1) + 1) + (N_D - 1)\mu(\mu(n - 1) + 1)}. \quad (\text{A43b})$$

795 The probability that two different deme-mates are identical by descent, Q_{in}^M , decreases
 796 monotonically with the emigration probability m , while Q_{out}^M monotonically increases
 797 with m (see figure A5(a)).

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

803 Relatedness R was defined in eq. (A20) as

$$R = \frac{Q_{\text{in}} - Q_{\text{out}}}{1 - Q_{\text{out}}}.$$

804 Using eq. (A43), relatedness under the Moran model is given by

$$R^M = \frac{(1 - \mu)(N_D(1 - m) - 1)}{N_D(1 - \mu)m(n - 1) + (N_D - 1)(1 + \mu(n - 1))}. \quad (\text{A44})$$

805 When there is an infinite number of demes ($N_D \rightarrow \infty$) and mutation is vanishingly small
 806 ($\mu \rightarrow 0$), we recover

$$\lim_{\mu \rightarrow 0} \lim_{N_D \rightarrow \infty} R^M = \lim_{N_D \rightarrow \infty} \lim_{\mu \rightarrow 0} R^M = \frac{1 - m}{1 + m(n - 1)}. \quad (\text{A45})$$

807 **C.2.2 Wright-Fisher**

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

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1-\mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right), \quad (\text{A46})$$

809 with \tilde{D} given in eq. (A38b). In a subdivided population, with $N_1 = N_D$ and $N_2 = n$, this
810 becomes

$$\begin{aligned} \tilde{Q}_{r_1 r_2} &= \frac{1}{N} \left[\frac{\mu \lambda'_{WF}}{1 - (1-\mu)^2 (\tilde{D}_0)^2} + \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1-\mu)^2 (\tilde{D}_0)^2} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right. \\ &\quad + \sum_{q_1=1}^{N_1-1} \frac{\mu \lambda'_{WF}}{1 - (1-\mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1-\mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1-\mu)^2} + \frac{1}{1 - (1-\mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \right. \\ &\quad + \frac{1}{1 - (1-\mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \\ &\quad \left. + \frac{1}{1 - (1-\mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_1} N_1 - 1) (\delta_{q_2} N_2 - 1) \right] \\ &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1-\mu)^2} + \frac{1}{1 - (1-\mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \delta_{q_1} N_1 \right. \\ &\quad \left. + \frac{1}{1 - (1-\mu)^2 (1 - m - \frac{m}{N_D-1})^2} (\delta_{q_1} N_1 - 1) \right]. \quad (\text{A47}) \end{aligned}$$

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

$$1 = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1-\mu)^2} + \frac{1}{1 - (1-\mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (N_2 - 1) N_1 + \frac{1}{1 - (1-\mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \quad (\text{A48a})$$

812 Then from eq. (A47) we deduce

$$Q_{\text{in}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1-\mu)^2} - \frac{1}{1 - (1-\mu)^2 (d_{\text{self}} - d_{\text{in}})^2} N_1 + \frac{1}{1 - (1-\mu)^2 (1 - m - \frac{m}{N_D-1})^2} (N_1 - 1) \right]. \quad (\text{A48b})$$

813 and

$$Q_{\text{out}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1-\mu)^2} - \frac{1}{1 - (1-\mu)^2 (1 - m - \frac{m}{d-1})^2} \right]. \quad (\text{A48c})$$

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

$$Q_{\text{in}}^{\text{WF}} = \frac{-N_D + M_1 + M_2}{(n-1)N_D + M_1 + M_2}, \quad (\text{A49a})$$

$$Q_{\text{out}}^{\text{WF}} = \frac{-\frac{1}{N_D-1}M_1 + M_2}{(n-1)N_D + M_1 + M_2}, \quad (\text{A49b})$$

815 with

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

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

818 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
819 the opposite pattern. The threshold value m_c^{WF} corresponds to an emigration probability
820 so high that $d_{\text{in}} = d_{\text{out}}$.

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

824 Also, because more sites (all of them, actually) are updated at each time step, Q_{in} is
825 lower for the Wright-Fisher updating than for a Moran updating, under which only one
826 site is updated at each time step (compare figure A5(a) and A5(b)).

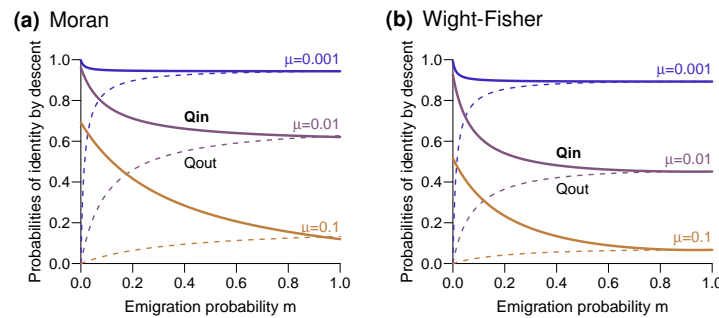


Figure A5: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in} , full curves) and two individuals in different demes (Q_{out} , dashed curves), as a function of the emigration probability m , for different values of the mutation probability μ (0.001, 0.01,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

827 Combining the formulas presented in eq. (A49), we obtain

$$R^{\text{WF}} = \frac{(1 - N_D(1 - m))^2(1 - \mu)^2}{D^{\text{WF}}}, \quad (\text{A50})$$

828 with

$$D^{\text{WF}} = 1 - N_D(2(1 + m(n-1)) - N_D(1 + (2-m)m(n-1))) - 2\mu \\ + 2(N_D(N_D(1-m) - 2)(1-m)(n-1) + n)\mu - (1 - N_D(1-m))^2(n-1)\mu^2.$$

829 When the number of demes is very large and mutation is vanishingly small, eq. (A50)
830 reduces to

$$\lim_{\mu \rightarrow 0} \lim_{N_D \rightarrow \infty} R^{\text{WF}} = \lim_{N_D \rightarrow \infty} \lim_{\mu \rightarrow 0} R^{\text{WF}} = \frac{(1-m)^2}{1 + (2-m)m(n-1)}. \quad (\text{A51})$$