

Mon titre

# 1 Introduction

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

Yet, living next to your kin also implies competing against them (West et al., 2002). The evolution of social traits hence depends on the balance between the positive effects of interactions with related individuals and the detrimental consequences of kin competition. Under specific conditions, the two effects can even compensate each other, thereby annihilating the impact of population viscosity on the evolution of altruism. First identified with computer simulations (Wilson et al., 1992), this cancellation result was analyzed by Taylor (1992) in a model with synchronous generations (Wright-Fisher model) and a subdivided population of constant, infinite size, and was later extended to heterogeneous populations (Rodrigues & Gardner, 2012, with synchronous generations and infinite population size), and other life-cycles and regular population structures (Taylor et al., 2011, with synchronous generations but also with continuous generations and Birth-Death updating). However, small changes in the model's assumptions, such as overlapping generations (Taylor & Irwin, 2000) or the presence of empty sites (Alizon & Taylor, 2008) can tip the balance back in the favor of altruism. This high dependence on life-cycle specificities highlights the difficulty of making general statements about the role of spatial structure on the evolution of altruism.

Another limitation of mechanistic models is the necessity of simplifying assumptions to obtain analytical results. A large number of studies on the evolution of social behavior consider simple population structures (typically, homogeneous populations *sensu* Taylor et al. (2007a)) and often also infinite population sizes (but see Allen et al., 2017, for results on any structure); they make use of weak selection approximations, and commonly assume rare (*e.g.*, Leturque & Rousset, 2002; Taylor et al., 2007b; Tarnita & Taylor, 2014) or absent mutation (models assuming infinite population sizes, or models concentrating on fixation probabilities). Simple population structures (*e.g.*, regular graphs, or subdivided populations with demes of equal sizes) help reduce the dimensionality of the system under study: this is for instance the case when the structure of the population displays symmetries such that all sites behave the same way in expec-

tation. Weak selection approximations are also crucial for disentangling spatial moments (Lion, 2016), that is, changes in global *vs.* local frequencies (though they can in some cases be relaxed Mullon & Lehmann, 2014). Finally, as highlighted by Tarnita & Taylor (2014), classical models of inclusive fitness assume infinite population sizes, which maintains diversity; in finite population this effect is obtained thanks to (rare) mutation. The aim of this study is to explore whether and how imperfect strategy transmission from parents to their offspring affects the impact of population viscosity on the evolution of altruistic behavior in subdivided populations.

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

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

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

## 2 Model and methods

### 2.1 Assumptions

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

Reproduction is asexual. Parents transmit their strategy to their offspring with probability  $1 - \mu$ ; this transmission can be genetic or cultural (vertical cultural transmission), but for simplicity, we refer to the parameter  $\mu$  as a mutation probability. With probability  $\mu$ , offspring do not inherit their strategy from their parent but instead get one randomly: with probability  $p$ , they become altruists,

76 with probability  $1 - p$  they become defectors. We call the parameter  $p$  the mu-  
77 tation bias.

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

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

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

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

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

88 Although our assumptions may seem restrictive (fecundity benefits are uncon-  
89 ditional, *i.e.*, the same which ever the type of the recipient; the fecundity effects  
90 are additive, *i.e.*, the effect of interacting with  $k$  altruists is  $k$  times the effect of  
91 interacting with one altruist), the same fecundities are obtained with a generic  
92 fecundity function, after linearization, under the assumption that altruists and  
93 defectors are phenotypically close (see Appendix A for details).

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

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

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

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

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

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

## 113 2.2 Methods

### 114 2.2.1 Analytical part

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

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

130 Check results with regular results

### 131 2.2.2 Stochastic simulations

132 We also run stochastic simulations (coded in C). The simulations are run for  $10^8$   
133 generations (one generation is one time step for the Wright-Fisher life-cycle, and  
134  $N$  time steps for the Moran life-cycles). For each set of parameters and life-cycle,  
135 we estimate the long-term frequency of altruists by sampling the population every  
136  $10^3$  generations and computing the average frequency of altruists.

add Code  
availability

### 137 3 Results

#### 138 3.1 Probabilities of identity by descent

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

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

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

##### 142 3.1.1 Moran updating

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

$$Q_{ij}^M = \frac{1-\mu}{2} \sum_{k=1}^N (d_{kj} Q_{ki}^M + d_{ki} Q_{kj}^M). \quad (5)$$

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

{eq:QM}

$$Q_{\text{in}}^M = \frac{(1-\mu)(m + \mu(N_D(1-m) - 1))}{(1-\mu)m(N_D\mu(n-1) + 1) + (N_D - 1)\mu(\mu(n-1) + 1)}, \quad (6a)$$

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

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

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

### 160 3.1.2 Wright-Fisher updating

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

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

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

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

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

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

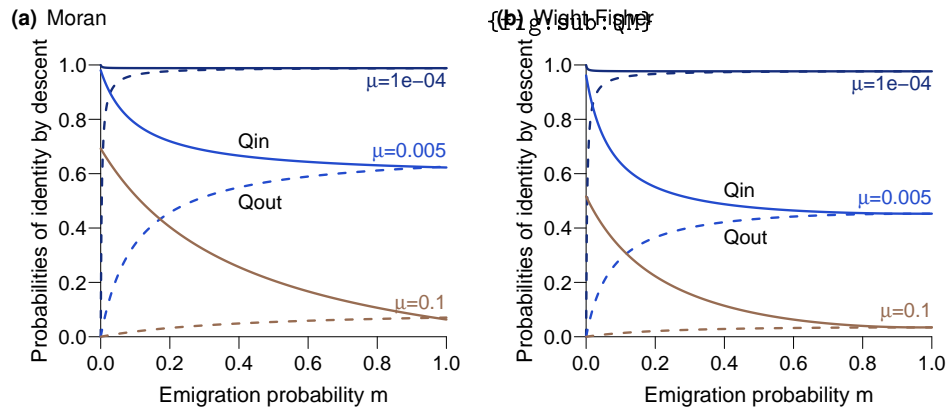
169 with

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

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

174 The two probabilities of identity by descent go to 1 when the mutation prob-  
 175 ability  $\mu$  is very small ( $\mu \rightarrow 0$ ), except if we first assume that the number of demes  
 176 is very large ( $N_D \rightarrow \infty$ ); for instance, with this life-cycle as well,  $\lim_{N_D \rightarrow \infty} Q_{\text{out}}^{\text{WF}} =$   
 177 0.

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



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



## 181 3.2 Expected frequencies of altruists for each life-cycle

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

$$\mathbb{E}[\bar{X}] \approx p + \omega \frac{p(1-p)}{\mu} [b(\beta_D - \beta_I) - c(\gamma_D - \gamma_I)]. \quad (9) \quad \{\text{eq:EXapprox}\}$$

184 (Calculations leading to eq. (9) are presented in the Appendix). The mutation  
185 bias  $p$  corresponds to the expected proportion of altruists in the population  
186 in the absence of selection (*i.e.*, when  $\omega = 0$ );  $\omega$  is the parameter that scales  
187 the effects of interactions between individuals and is assumed to be small; the  
188 subscript  $_D$  refers to “direct” effects, and the subscript  $_I$  to “indirect” effects.  
189 These indirect effects correspond to (kin) competition: by providing a benefit  
190 to a deme-mate and thereby increasing its fecundity, a focal altruist indirectly  
191 harms others by reducing their relative fecundity ( $\beta_I$  term in eq. (9)); by having  
192 a reduced fecundity due to the cost of altruism, a focal altruist indirectly favors  
193 others by increasing their relative fecundity ( $\gamma_I$  term).

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

### 196 3.2.1 Direct effects

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

{eq:directeffects}

$$\beta_D^{BD} = \beta_D^{DB} = (1 - \mu) Q_{in}^M, \quad (10a) \quad \{\text{eq:bBDD}\}$$

$$\beta_D^{WF} = (1 - \mu) Q_{in}^{WF}; \quad (10b) \quad \{\text{eq:bWFD}\}$$

$$\gamma_D^{BD} = \gamma_D^{DB} = \gamma_D^{WF} = 1 - \mu. \quad (10c) \quad \{\text{eq:cBDD}\}$$

200 For both benefits and costs, direct effects only count when there is no mutation  
201 ( $1 - \mu$ ). Direct effects of benefits (b) only count if the interaction takes place with  
202 an individual who is identical by descent; social interactions occur only within  
203 demes, hence the presence of  $Q_{in}$  in eq. (10a) and eq. (10b). The direct effect of  
204 the fecundity cost  $c$  however does not depend on the type of interactant.

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

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

### 3.2.2 Indirect effects

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

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

$$\begin{aligned}\beta_1^{\text{BD}} &= (1 - m) \left( \frac{n-1}{n} Q_{\text{in}}^{\text{M}} + \frac{1}{n} \right) + m Q_{\text{out}}^{\text{M}} - \mu \frac{1 + (n-1)Q_{\text{in}}^{\text{M}} + n(d-1)Q_{\text{out}}^{\text{M}}}{nd} \\ &= \gamma_{\text{D}}^{\text{BD}}.\end{aligned}\tag{11a} \quad \{\text{eq:bBDI}\}$$

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

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

### 3.2.3 Moran Death-Birth

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

$$\begin{aligned}
\beta_1^{\text{DB}} &= (1 - \mu) \left[ \left( \frac{1}{n} + \frac{(n-1)Q_{\text{in}}^{\text{M}}}{n} \right) \left( (1-m)^2 + \frac{m^2}{(d-1)} \right) \right. \\
&\quad \left. + m \left( 2(1-m) + (d-2) \frac{m}{(d-1)} \right) Q_{\text{out}}^{\text{M}} \right] \\
&= \gamma_1^{\text{DB}}
\end{aligned} \tag{11b} \quad \{\text{eq:bDBI}\}$$

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

242 Here again,  $\beta_1 = \gamma_1$ , so the balance between indirect benefits and indirect  
 243 costs does not change when the emigration probability  $m$  increases.

244 Replacing  $Q_{\text{in}}$  and  $Q_{\text{out}}$  by their formulas given in eq. (6), we can conclude  
 245 that  $\beta_1 = \gamma_1$  first decreases with the emigration probability  $m$ , and increases  
 246 again after a threshold value  $m'_c$  (given in the appendix;  $m'_c < (d-1)/d$ ).

Mathematica

name

### 247 3.2.4 Wright-Fisher

248 With this life-cycle, generations are synchronous and all individuals again all  
 249 have the same survival probability (now equal to 0). As a result, the formulas  
 250 for  $\beta_1^{\text{WF}}$  and  $\gamma_1^{\text{WF}}$  are the same as  $\beta_1^{\text{DB}}$  and  $\gamma_1^{\text{DB}}$ , except that instead of  $Q_{\text{in}}^{\text{M}}$  and  
 251  $Q_{\text{out}}^{\text{M}}$ , we need to use  $Q_{\text{in}}^{\text{WF}}$  and  $Q_{\text{out}}^{\text{WF}}$  (given in eq. (8)). Once this is done, we see  
 252 that  $\beta_1^{\text{WF}} = \gamma_1^{\text{WF}}$  first decreases with the emigration probability  $m$ , and increases  
 253 again after the threshold value  $m'_c = (d-1)/d$  (which was identified previously  
 254 as the emigration probability such that offspring have an equal chance of land-  
 255 ing in their natal deme or in any other deme, *i.e.*,  $d_{\text{in}} = d_{\text{out}}$ ).

## 256 3.3 Identifying threshold values of the mutation probability $\mu$

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

### 262 3.3.1 Moran Birth-Death

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

265 pends on the value of the mutation probability  $\mu$  compared to a threshold value  
 266  $\mu_c^{\text{BD}}$ . 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  
 267 with  $m$ ;  $\mu_c^{\text{BD}}$  is given by

$$\mu_c^{\text{BD}} = 1 - \frac{b - c + \sqrt{(b - c)(4b(nd)^2 + b - c)}}{2bnd} \quad (12) \quad \{\text{eq:mucBD}\}$$

268 This result is illustrated in figure 2(b).

donner la  
valeur

### 269 3.3.2 Moran Death-Birth

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

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

278 The expected frequency of altruists  $\mathbb{E}[\bar{X}]$  reaches a maximum for an emigration  
 279 probability  $m_c^{\text{DB}}$  (whose complicated equation is in the **appendix**), as can be seen  
 280 in figure 2(a). The limit of this critical emigration probability  $m_c^{\text{DB}}$  when  $\mu \rightarrow 0$  is  
 281 0: we recover the result that more emigration is detrimental to the evolution of  
 282 altruism when the mutation probability is either null or vanishingly small.

do

attention  
order of  
limits

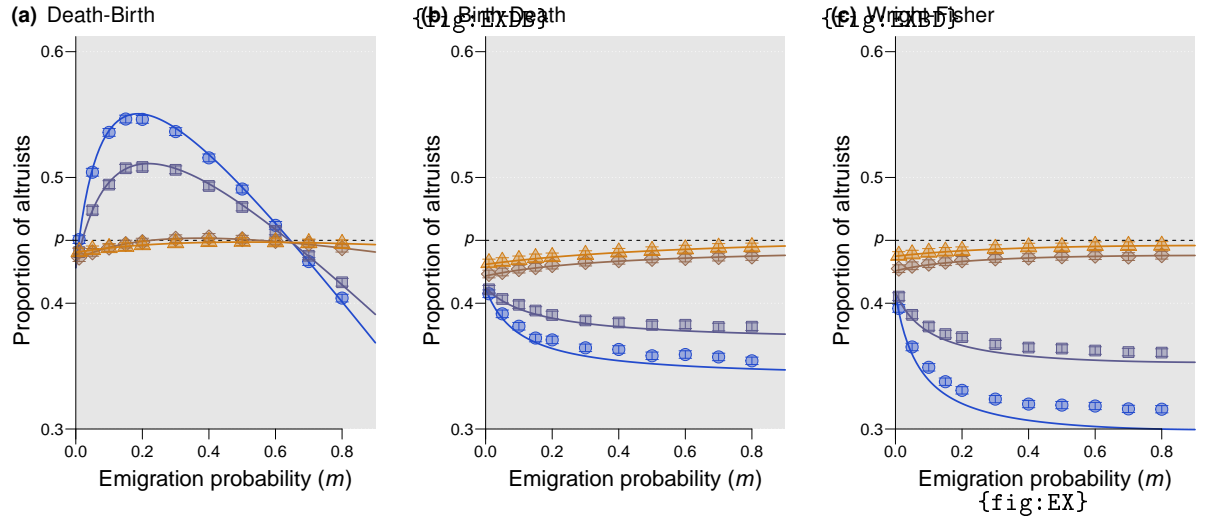
### 283 3.3.3 Wright-Fisher

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

$$\mu_c^{\text{WF}} = 1 - \sqrt{1 - \frac{c}{b}}, \quad (14)$$

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

appendix



**Figure 2:** Expected proportion of altruists under weak selection, as a function of the emigration probability  $m$ , for different mutation values  $\mu = 0.005$  (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), and 0.25 (orange, triangles) and life-cycles ((a) Moran Death-Birth, (b) Moran Birth Death, (c) Wright-Fisher). The curves are the analytical results, the points are the output of numerical simulations. Parameters:  $\omega = 0.005$ ,  $p = 0.45$ ,  $b = 15$ ,  $c = 1$ ,  $n = 4$  individuals per deme,  $N_D = 30$  demes.

### 288 3.4 Relaxing key assumptions

289 To derive our analytical results, we had to make a number of simplifying as-  
290 sumptions, such as the fact that selection is weak ( $\omega \ll 1$ ), and the fact that the  
291 structure of the population is regular (all demes have the same size  $n$ ). We ex-  
292 plored with numerical simulations the effect of relaxing these key assumptions.  
293 The patterns that we identified hold when selection is strong (see figure S1, done  
294 with  $\omega = 0.1$ ), but also when the demes have different sizes. Deme sizes are  
295 drawn randomly at the beginning of a simulation; the range from 1 to 5 individ-  
296 uals per deme and the average size is 4 individuals as in the other figures.. Here  
297 as well, the same patterns hold as those obtained with a homogeneous structure  
298 (figure S2). Add effect of  $d_{\text{self}}$ .

## 299 4 Discussion

300 Assuming that the transmission of a social strategy (being an altruist or a de-  
301 fector) from a parent to its offspring could be imperfect, we found that the ex-  
302 pected frequency of altruists maintained in the population could increase with  
303 the probability of emigration out of the parental deme, a parameter tuning pop-  
304 ulation viscosity. This result can seem surprising, because it contradicts the con-  
305 clusions obtained under the assumption of nearly perfect strategy transmission  
306 (*i.e.*, in the case of genetic transmission, that mutation is very weak or absent),  
307 where increased population viscosity (decreased emigration probability) is ei-  
308 ther neutral (Taylor, 1992) or favorable (Taylor et al., 2007a) to the evolution of  
309 altruistic behavior.

310 Adding non zero mutation probability altruism increases with emigration.

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

322 Go back to the decomposition of the different terms, we see that increase  
323 of  $\mathbb{E}[\bar{X}]$  with  $m$  is driven by the  $\beta_1$  term. To simplify the explanations, let us  
324 consider that the number of demes is large: in this case,  $Q_{\text{out}}$  is vanishingly small

le pb c'est  
ptet sim-  
plement  
que 1 in-  
dividu  
ca pose  
probleme!

todo

325 and so terms involving it can be omitted. Let us also assume that there is no  
326 direct cost to being an altruist ( $c = 0$ ).

327 Problems of orders of limits, especially when  $d \rightarrow \infty$  and  $\mu \rightarrow 0$ . Need to  
328 specify how small the mutation probability is compared to the size of the popu-  
329 lation.

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

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

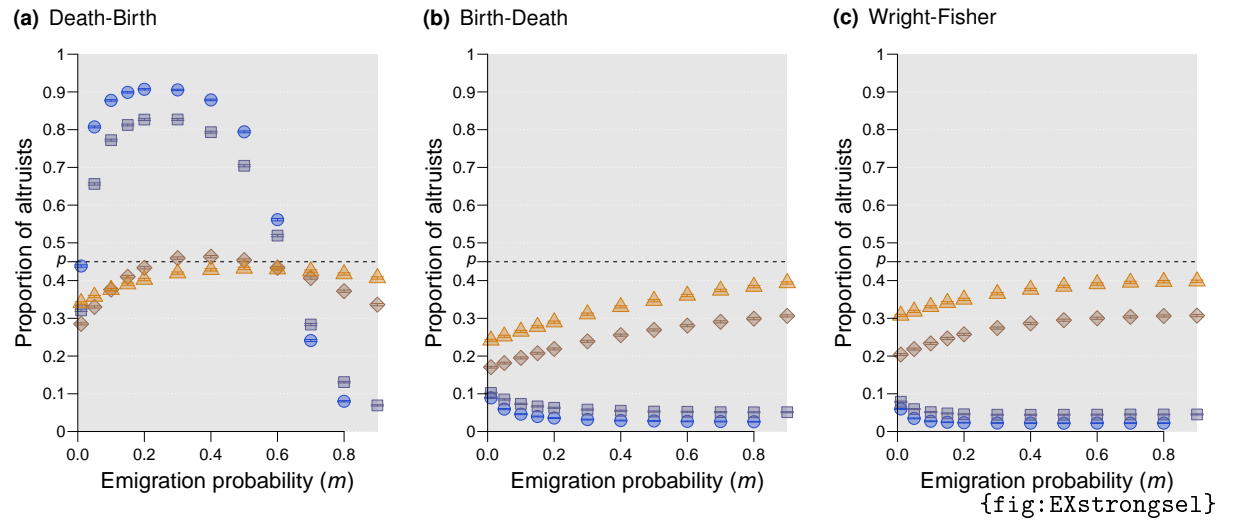
## 344 **References**

- 345 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*  
346 *tion* 62(6):1335–1344.
- 347 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,  
348 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*  
349 544(7649):227–230.
- 350 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects  
351 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.  
352 *Evolution of Cooperation.*
- 353 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-  
354 mediate disturbance. *Current Biology* 17(9):761–765.
- 355 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution  
356 of social behavior in structured populations. *Journal of Theoretical Biology*  
357 420:26 – 35.
- 358 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-  
359 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*  
360 276(1654):13–19.
- 361 Frank, S. A. 1997: The price equation, fisher's fundamental theorem, kin selec-  
362 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 363 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*  
364 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–  
365 2522.
- 366 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*  
367 *Theoretical Biology* 7(1):1 – 16.
- 368 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-  
369 tionary genetics. *Biosocial anthropology* 53:133–55.
- 370 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and  
371 harming on graphs: the return of the inclusive fitness effect. *Journal of Evolu-*  
372 *tionary Biology* 20(6):2284–2295.
- 373 Leturque, H. & Rousset, F. 2002: Dispersal, kin competition, and the ideal free  
374 distribution in a spatially heterogeneous population. *Theoretical Population*  
375 *Biology* 62(2):169 – 180.

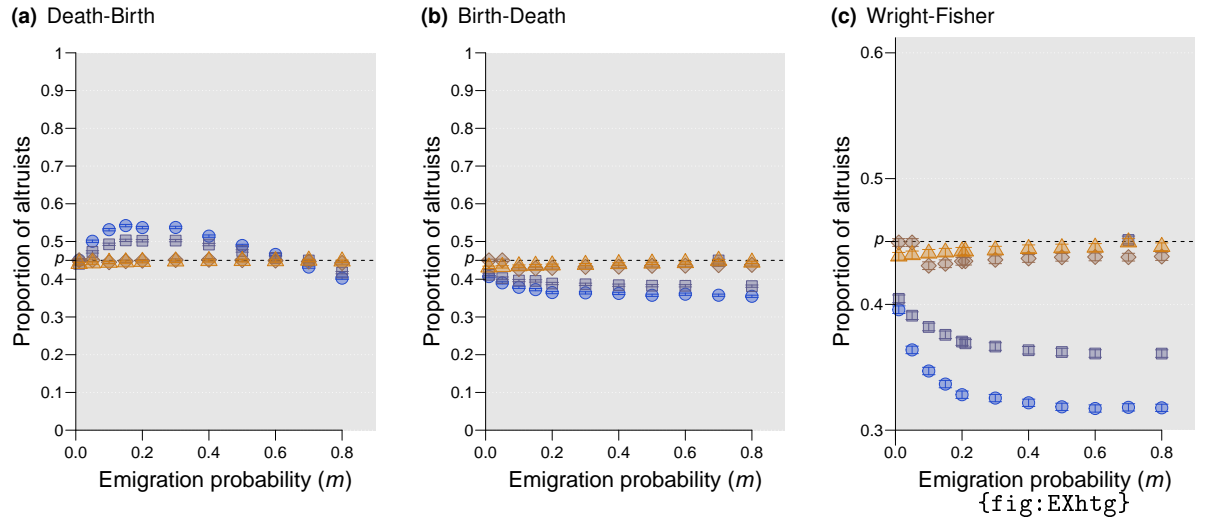


- 376 Lion, S. 2016: Moment equations in spatial evolutionary ecology. *Journal of the-*  
377 *oretical biology* 405:46–57.
- 378 Mullon, C. & Lehmann, L. 2014: The robustness of the weak selection approxi-  
379 *mation for the evolution of altruism against strong selection. Journal of evo-*  
380 *lutionary biology* 27(10):2272–2282.
- 381 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule  
382 *for the evolution of cooperation on graphs and social networks. Nature*  
383 441(7092):502–505.
- 384 Rodrigues, A. M. M. & Gardner, A. 2012: Evolution of helping and harming in  
385 *heterogeneous populations. Evolution* 66(7):2065–2079.
- 386 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors  
387 *in finite structured population models. The American Naturalist* 184(4):477–  
388 488.
- 389 Taylor, P. 1992: Altruism in viscous populations—an inclusive fitness model.  
390 *Evolutionary ecology* 6(4):352–356.
- 391 Taylor, P.; Lillicrap, T. & Cownden, D. 2011: Inclusive fitness analysis on mathe-  
392 *matical groups. Evolution* 65(3):849–859.
- 393 Taylor, P. D.; Day, T. & Wild, G. 2007a: Evolution of cooperation in a finite homo-  
394 *geneous graph. Nature* 447(7143):469–472.
- 395 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-  
396 *bility in homogeneous structured populations. Journal of Theoretical Biology*  
397 249(1):101 – 110.
- 398 Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic  
399 *behavior. Evolution* 54(4):1135–1141.
- 400 West, S. A.; Pen, I. & Griffin, A. S. 2002: Cooperation and competition between  
401 *relatives. Science* 296(5565):72–75.
- 402 Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely  
403 *viscous populations? Evolutionary Ecology* 6(4):331–341.

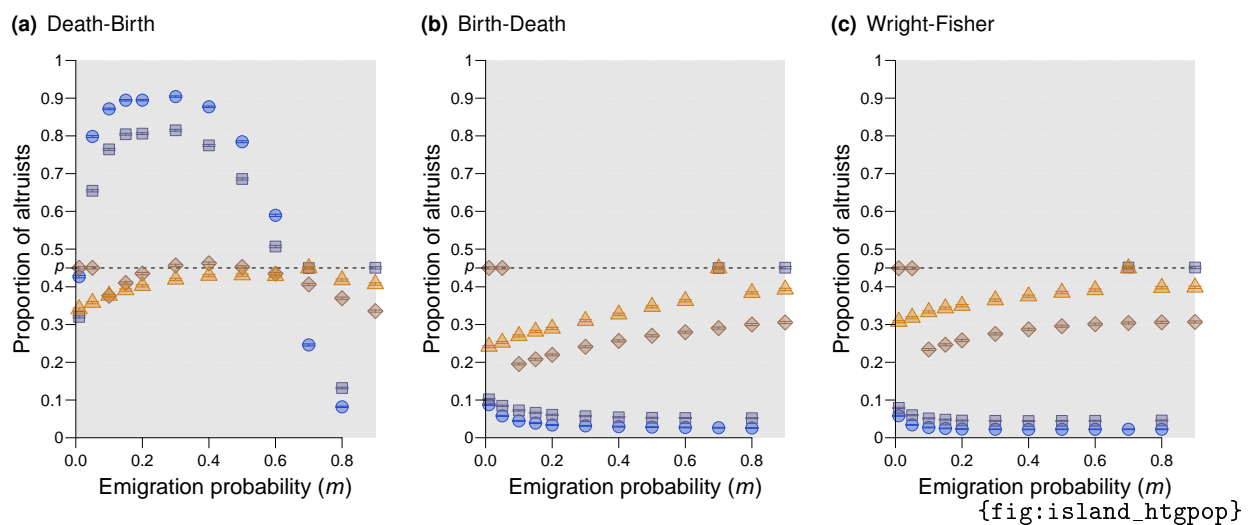
## Supplementary figures



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



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



**Figure S3:** Strong selection, heterogeneous population

## A Fecundity and weak selection approximation

{sec:app:F}

Here we show that the fecundity function presented in eq. (2) can be obtained from a generic fecundity function, under the assumption of small phenotypic differences between altruists and defectors. Let us denote by  $\phi_i$  the phenotype of the individual living at site  $i$ , and assume that the phenotypic value of altruists and defectors differs by  $\omega \ll 1$ , so that

$$\phi_i = \phi^{(0)} + \omega X_i, \quad (\text{A.1}) \quad \{\text{eq:app:phidef}\}$$

where  $\phi^{(0)}$  is the phenotype of defectors (and  $\phi_0 + \omega$  the phenotype of altruists). Then we consider a generic fecundity function, that depends on the focal individual's phenotype, and on the phenotype of all the individuals that it interacts with, weighted by the probability of interaction ( $e_{ki}$  for an individual at site  $k$ ,  $1 \leq k \leq N$ ). For instance, the fecundity of an individual at site  $i$  is given by a function with  $N + 1$  arguments

$$F_i = \mathcal{F}(e_{1,i}\phi_1, \dots, e_{N,i}\phi_N; \phi_i),$$

which, using eq. (A.1), becomes

$$F_i = \mathcal{F}(\phi^{(0)} + \omega X_1, \dots, \phi^{(0)} + \omega X_N; \phi^{(0)} + \omega X_i). \quad (\text{A.2}) \quad \{\text{eq:app:Fi}\}$$

We then write a first-order expansion of eq. (A.2) for  $\omega \ll 1$ :

$$F_i = \mathcal{F}(\phi^{(0)}, \dots, \phi^{(0)}; \phi^{(0)}) + \omega \left[ \sum_{k=1}^N e_{ki} X_k \partial_{(k)} \mathcal{F}|_{\omega=0} + X_i \partial_{(N+1)} \mathcal{F}|_{\omega=0} \right] + O(\omega^2), \quad (\text{A.3}) \quad \{\text{eq:app:DLF}\}$$

where  $\partial_{(k)} \mathcal{F}|_{\omega=0}$  is the derivative of  $\mathcal{F}$  with respect to its  $k^{\text{th}}$  argument, evaluated at  $\omega = 0$ . The first term in eq. (A.3) is the fecundity of individual  $i$  when there is not a single altruist in the population. Given the chosen structure of the population, all individuals have the same number of social interactions, so they end up having the same fecundity when the population is fixed for the defector type. Without loss of generality, we set this baseline fecundity equal to 1. If we now consider that it does not matter where the benefits of social interactions come from, only that they are received by the focal individual, then  $\partial_{(k)} \mathcal{F}$  is the same for all  $k$ ,  $1 \leq k \leq N$ ; let us denote it by  $b$ . If we denote by  $-c$  the marginal effect of the focal individual's phenotype on its own fecundity (*i.e.*,  $-c = \partial_{N+1} \mathcal{F}|_{\omega=0}$ ), then we obtain

$$F_i = 1 + \omega \left( b \sum_{k=1}^N e_{ki} X_k - c X_i \right) + O(\omega^2), \quad (\text{A.4})$$

which is equal to  $f_i$  as defined in eq. (2), neglecting terms in  $\omega^2$  and higher.

## B Expected frequency of altruists

{sec:app:EX}

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

### B.1 For a generic life-cycle

{sec:app:generic}

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

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

{tab:BD}

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

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

$$D_i = \sum_{j=1}^N B_{ij} \quad (\text{B.5a}) \quad \{\text{eq:DBequiv}\}$$

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

$$D_i^0 = \sum_{j=1}^N B_{ji}^0 = B^*, \quad (\text{B.5b}) \quad \{\text{eq:DBRV}\}$$

456 where the  $^0$  subscript means that the quantities are evaluated for  $\omega = 0$ ; this also  
 457 implies that  $B_{ij}^0$  and  $D_i^0$  do not depend on the state  $\mathbf{X}$  of the population. For the  
 458 Moran life-cycles,  $B^* = 1/N$ , while for the Wright-Fisher life-cycle,  $B^* = 1$ . (The  
 459 difference with eq. (B.5a) is that we are now considering offspring produced by  
 460  $i$  landing on  $j$ ).

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

$$\mathbb{E}[\bar{X}(t+1)|\mathbf{X}(t)] = \frac{1}{N} \sum_{i=1}^N \left[ \sum_{j=1}^N B_{ij} (X_j(1-\mu) + \mu p) + (1 - D_i) X_i \right]. \quad (\text{B.6a}) \quad \{\text{eq:conditionalchange}\}$$

463 The first term within the brackets corresponds to births: the type of the individ-  
 464 ual living at  $i$  at time  $t + 1$  depends on the type of its parent (living at site  $j$ ), and  
 465 on whether mutation occurred. The second term in the brackets of eq. (B.6a)  
 466 corresponds to the survival of the individual living at site  $i$ .

467 Given that there is no absorbing population state (a lost strategy can always  
 468 be recreated by mutation), there is a stationary distribution of population states,  
 469 so that the expected frequency of altruists does not change anymore for large  
 470 times  $t$  (realized frequencies of course keep changing). We denote by  $\xi(\mathbf{X}, \omega, \mu)$   
 471 the probability that the population is in state  $\mathbf{X}$ , given the strength of selection  
 472  $\omega$  and the mutation probability  $\mu$ . Taking the expectation of eq. (B.6a) ( $\mathbb{E}[\bar{X}] =$   
 473  $\sum_{\mathbf{X} \in \Omega} \bar{X} \xi(\mathbf{X}, \omega, \mu)$ ), we obtain, after reorganizing:

$$0 = \frac{1}{N} \sum_{\mathbf{X} \in \Omega} \sum_{i=1}^N \left[ \sum_{j=1}^N B_{ij} (X_j(1-\mu) + \mu p) - D_i X_i \right] \xi(\mathbf{X}, \omega, \mu). \quad (\text{B.7}) \quad \{\text{eq:statdist}\}$$

474 Now, we use the assumption of weak selection ( $\omega \ll 1$ ) and consider the first-  
 475 order expansion of eq. (B.7) for  $\omega$  close to 0. First, we note that in the absence of

476 selection ( $\omega = 0$ ), the population is at a mutation-drift balance, and the expected  
 477 state of every site  $i$  is then  $\mathbb{E}_0[X_i] = \sum_{X \in \Omega} X_i \xi(X, 0, \mu) = p$ , the mutation bias.  
 478 Secondly, we further expand derivatives of  $B_{ji}$  and  $D_i$  thanks to the chain rule,  
 479 using the variables  $f_k$  ( $1 \leq k \leq N$ ), corresponding to individual fecundities (also,  
 480 recall that  $f_k = 1$  when  $\omega = 0$ ). Thirdly, we note that for all the life-cycles that we  
 481 consider, the number of deaths in the population during one time step does not  
 482 depend on population composition (exactly 1 death for the Moran life-cycles,  
 483 and exactly  $N$  for the Wright-Fisher life-cycle), so that  $\partial \sum_{i,j=1}^N B_{ij} / \partial \omega$  does not  
 484 depend on  $\omega$ . After simplification and reorganization, the first order expansion  
 485 of eq. (B.7) yields

$$\begin{aligned}
 0 = & \frac{1}{N} \sum_{i,k=1}^N \left[ \frac{\partial \left( \sum_{j=1}^N (1-\mu) B_{ji} - D_i \right)}{\partial f_k} \right]_{f_k=1} \\
 & \times \left( \sum_{\ell=1}^N e_{\ell k} b \sum_{X \in \Omega} X_{\ell} X_i \xi(\mathbf{X}, 0, \mu) - c \sum_{X \in \Omega} X_k X_i \xi(\mathbf{X}, 0, \mu) \right) \\
 & - B^* \mu \frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} + O(\omega^2). \tag{B.8} \quad \{\text{eq:weaksel1}\}
 \end{aligned}$$

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

$$P_{ij} = p^2 + p(1-p)Q_{ij}. \tag{B.9} \quad \{\text{eq:QP}\}$$

489 In [appendix B.2](#), we will see that recursions on  $P_{ij}$  will reveal that  $Q_{ij}$  can be  
 490 interpreted as a probability of identity by descent, *i.e.*, the probability that the  
 491 individuals at sites  $i$  and  $j$  have a common ancestor and that no mutation has  
 492 occurred on either lineage since the ancestor.

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

$$\mathbb{E}[\bar{X}] = p + \omega \frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} + O(\omega^2), \tag{B.10} \quad \{\text{eq:EXgeneric}\}$$

495 where  $\frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0}$  is obtained from eq. (B.8). We then need to replace the  $B_{ij}$  and  
 496  $D_j$  terms by their formulas for each life-cycle (given in table S1), and the  $d_{ij}$  and  
 497  $e_{ij}$  terms by their formulas (given in eq. (3)) and eq. (1), respectively). For each  
 498 life-cycle we can group terms as

$$\frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} \approx \frac{p(1-p)}{\mu} [b(\beta_D - \beta_I) - c(\gamma_D - \gamma_I)], \tag{B.11}$$



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

### Moran Birth-Death

$$\beta_D^{\text{BD}} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^{\text{M}}, \quad (\text{B.12a}) \quad \{\text{eq:EXBDsums}\}$$

$$\beta_I^{\text{BD}} = \sum_{j,k,\ell=1}^N \left( \frac{d_{\ell j}}{N} - \frac{\mu}{N^2} \right) e_{k\ell} Q_{jk}^{\text{M}}, \quad (\text{B.12b})$$

$$\gamma_D^{\text{BD}} = 1 - \mu, \quad (\text{B.12c})$$

$$\gamma_I^{\text{BD}} = \sum_{j,k=1}^N \left( \frac{d_{kj}}{N} - \frac{\mu}{N^2} \right) Q_{jk}^{\text{M}}. \quad (\text{B.12d})$$

### Moran Death-Birth

$$\beta_D^{\text{DB}} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^{\text{M}}, \quad (\text{B.13a})$$

$$\beta_I^{\text{DB}} = (1-\mu) \sum_{i,j,k,\ell=1}^N \frac{d_{ji} d_{\ell i}}{N} e_{k\ell} Q_{jk}^{\text{M}}, \quad (\text{B.13b})$$

$$\gamma_D^{\text{DB}} = 1 - \mu, \quad (\text{B.13c})$$

$$\gamma_I^{\text{DB}} = (1-\mu) \sum_{i,j,k=1}^N \frac{d_{ji} d_{ki}}{N} Q_{jk}^{\text{M}}. \quad (\text{B.13d})$$

### Wright-Fisher

$$\beta_D^{\text{WF}} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^{\text{WF}}, \quad (\text{B.14a})$$

$$\beta_I^{\text{WF}} = (1-\mu) \sum_{i,j,k,\ell=1}^N \frac{d_{ji} d_{\ell i}}{N} e_{k\ell} Q_{jk}^{\text{WF}}, \quad (\text{B.14b})$$

$$\gamma_D^{\text{WF}} = 1 - \mu, \quad (\text{B.14c})$$

$$\gamma_I^{\text{WF}} = (1-\mu) \sum_{i,j,k=1}^N \frac{d_{ji} d_{ki}}{N} Q_{jk}^{\text{WF}}, \quad (\text{B.14d})$$

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

## B.2 Probabilities of identity by descent

{sec:app:IBD}

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

### B.2.1 Moran model

In a Moran model, exactly one individual died and one individual reproduces during one time step. Given a state  $\mathbf{X}$  at time  $t$ , at time  $t + 1$  both sites  $i$  and  $j \neq i$  are occupied by altruists (or say mutants, since there is no selection and hence no benefits or costs provided by altruists), if *i*) it was the case at time  $t$  and neither site was replaced by a non-mutant (first term in eq. (B.15)), or *ii*) if exactly one of the two sites was occupied by a non-mutant at time  $t$ , but the site was replaced by a mutant (second and third terms of eq. (B.15)):

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & X_i X_j \left( 1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - p)) \right) \\ & + X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu p) \\ & + X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu p). \end{aligned} \quad (\text{B.15}) \quad \{\text{eq:app:PijM1}\}$$

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

$$P_{ij} = \frac{1}{2} \left( \sum_{k=1}^N (1 - \mu) (d_{kj} P_{ki} + d_{ki} P_{kj}) \right) + \mu p^2 \quad (i \neq j), \quad (\text{B.16}) \quad \{\text{eq:app:PijM}\}$$

while  $P_{ii} = p$ .

Now we substitute  $P_{ij} = p^2 + p(1 - p)Q_{ij}$  in eq. (B.16), obtain

$$Q_{ij} = \frac{1}{2} \sum_{k=1}^N (1 - \mu) (d_{ki} Q_{kj} + d_{kj} Q_{ki}), \quad (\text{B.17}) \quad \{\text{eq:app:QijM}\}$$

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

## 527 **B.2.2 Wright-Fisher model**

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & \sum_{k, \ell=1}^N d_{ki} d_{\ell j} \left( X_k X_\ell (1 - \mu + \mu p)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu p)(\mu p) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu p)^2 \right) \end{aligned} \quad (\text{B.18}) \quad \{\text{eq:app:PijWF1}\}$$

530 Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k, \ell=1}^N (P_{kl}(1 - \mu)^2) + (2 - \mu)\mu p^2. \quad (\text{B.19}) \quad \{\text{eq:app:PijWF}\}$$

531 Replacing  $P_{ij}$  by  $p^2 + p(1 - p)Q_{ij}$ , eq. (B.19) becomes

$$Q_{ij} = \sum_{k, \ell=1}^N d_{ki} d_{\ell j} Q_{k\ell} (1 - \mu)^2. \quad (\text{B.20}) \quad \{\text{eq:app:QijWF}\}$$

532 Again,  $Q_{ij}$  corresponds to a probability of identity by descent: the individuals at  
 533 sites  $i$  and  $j$  are identical by descent if their parents were and if neither mutated  
 534  $((1 - \mu)^2)$ .

## 535 C In a subdivided population

### 536 C.1 $\beta$ and $\gamma$

537 Now, we need to adapt the results presented in [appendix B](#) to our structure of  
 538 interest, a subdivided population, with dispersal and interaction probabilities  
 539 given by eq. (3) and eq. (1). For the  $\beta$  and  $\gamma$  terms, we use a brute-force approach,  
 540 replacing  $d_{ij}$  and  $e_{ij}$  by their values in a subdivided population, and simplifying  
 541 the equations (for instance, there are 60 different cases to consider for the four  
 542 sums that appear in  $\beta_1^{\text{DB}}$ ). The calculations are detailed in an accompanying  
 543 Mathematica file, [and the results are presented in the main text.](#)

todo

### 544 C.2 Probabilities of identity by descent

545 For the probabilities of identity by descent, we could also use a brute-force ap-  
 546 proach, but calculations are faster if we use formulas derived in Débarre (2017)  
 547 for “two-dimensional population structures”. The name comes from the fact  
 548 that we only need two types of transformations to go from any site to any other  
 549 site in the population: permutations on the deme index, and permutations on  
 550 the within-deme index. We introduce notations  $\tilde{d}_i$  and  $\tilde{Q}_i$ , that correspond to  
 551 the dispersal probability to a site at distance  $i$  (e.g., for all  $j, 1 \leq j \leq N$ ,  $\tilde{d}_1 =$   
 552  $d_{j,j+1}$ ) and the probability of identity by descent with a site at distance  $i$  (e.g.,  
 553 for all  $j, 1 \leq j \leq N$ ,  $\tilde{Q}_1 = Q_{j,j+1}$ ), respectively. Finally, we can rewrite site labels  
 554 ( $1 \leq i \leq N$ ) as  $(l_1, l_2)$ , where  $l_1$  is the number of the deme ( $1 \leq l_1 \leq N_D$ ) and  $l_2$  the  
 555 position of the site within the deme ( $1 \leq l_2 \leq n$ ).

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

#### 558 C.2.1 Moran model

559 In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{C.21a}) \quad \{\text{eq:app:Q2DM}\}$$

560 with

$$\tilde{D}_{q_1} = \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} \tilde{d}_{l_1} \exp\left(-i \frac{2\pi q_1 l_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 l_2}{N_2}\right), \quad (\text{C.21b}) \quad \{\text{eq:app:D2D}\}$$

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

$$\begin{aligned}\tilde{D}_{q_1} &= d_{\text{self}} + \sum_{l_2=1}^{N_2-1} d_{\text{in}} \exp\left(-i \frac{2\pi q_2 l_2}{N_2}\right) + \sum_{l_1=1}^{N_1-1} \sum_{l_2=0}^{N_2-1} d_{\text{out}} \exp\left(-i \frac{2\pi q_1 l_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 l_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2}(N_2 - 1) + (1 - \delta_{q_2})(-1)) d_{\text{in}} + (\delta_{q_1}(N_1 - 1) + (1 - \delta_{q_1})(-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}.\end{aligned}\quad (\text{C.22a})$$

563 ( $\delta_q$  is equal to 1 when  $q$  is equal to 0 modulo the relevant dimension, and to 0  
 564 otherwise). So for the three types of distances that we need to consider (distance  
 565 0, distance to another deme-mate, distance to individual in another deme), and  
 566 with  $N_1 = N_D$  and  $N_2 = n$ , we obtain {eq:app:Dsystem}

$$\tilde{D}_0 = 1, \quad (\text{C.23a})$$

$$\tilde{D}_{q_1} = 1 - m - \frac{m}{d-1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \quad (\text{C.23b})$$

$$\tilde{D}_{q_1} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \quad (\text{C.23c})$$

567 So for  $\tilde{Q}$ , using system (C.23) in eq. (C.21a),

$$\begin{aligned}\tilde{Q}_{r_1} &= \frac{\mu \lambda'_M}{N} \left[ \frac{1}{1 - (1 - \mu) \tilde{D}_0} + \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_2}} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) + \sum_{q_1=1}^{N_1-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \right. \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_M}{N} \left[ \frac{1}{1 - (1 - \mu)} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_2} N_2 - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})} (\delta_{r_1} N_1 - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_1} N_1 - 1)(\delta_{r_2} N_2 - 1) \right].\end{aligned}\quad (\text{C.24}) \quad \{\text{eq:app:Q2DMsol}\}$$

568 In particular,

$$\begin{aligned}\tilde{Q}_0 &= \frac{\mu \lambda'_M}{N} \left[ \frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (n - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})} (D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (D - 1)(n - 1) \right] \\ &= 1.\end{aligned}\quad (\text{C.25a}) \quad \{\text{eq:app:Q2D1}\}$$

569 We find  $\lambda'_M$  using the eq. (C.25a). Going back to eq. (C.24), when  $r_1 = 0$ , the two  
 570 individuals are in the same deme. They are different when  $r_2 \neq 0$ , and so:

$$Q_{\text{in}} = \frac{\mu\lambda'_M}{N} \left[ \frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})}(D-1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(D-1)(-1) \right]. \quad (\text{C.25b})$$

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

$$Q_{\text{out}} = \frac{\mu\lambda'_M}{N} \left[ \frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})}(-1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} \right]. \quad (\text{C.25c})$$

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

### 574 **C.3 Wright-Fisher**

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

$$\tilde{Q}_{r_1} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu\lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right), \quad (\text{C.26})$$

576 with  $\tilde{D}$  given in eq. (C.21b). In a subdivided population, with  $N_1 = N_D$  and  $N_2 =$   
 577  $n$ , this becomes

$$\begin{aligned}
 \tilde{Q}_{r_1 r_2} &= \frac{1}{N} \left[ \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_0)^2} + \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_2})^2} \exp \left( -i \frac{2\pi q_2 r_2}{N_2} \right) \right. \\
 &\quad + \sum_{q_1=1}^{N_1-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_1})^2} \exp \left( -i \frac{2\pi q_1 r_1}{N_1} \right) \\
 &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_1})^2} \exp \left( -i \frac{2\pi q_1 r_1}{N_1} \right) \exp \left( -i \frac{2\pi q_2 r_2}{N_2} \right) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[ \frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \right. \\
 &\quad + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (\delta_{q_1} N_1 - 1) \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_1} N_1 - 1) (\delta_{q_2} N_2 - 1) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[ \frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \delta_{q_1} N_1 \right. \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (\delta_{q_1} N_1 - 1) \right]. \tag{C.27} \quad \{\text{eq:app:Q2DWFsol}\}
 \end{aligned}$$

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

$$1 = \frac{\mu \lambda'_{WF}}{N} \left[ \frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (N_2 - 1) N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (N_1 - 1) \right]. \tag{C.28a}$$

579 Then from eq. (C.27) we deduce

$$Q_{\text{in}} = \frac{\mu \lambda'_{WF}}{N} \left[ \frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (N_1 - 1) \right]. \tag{C.28b}$$

580 and

$$Q_{\text{out}} = \frac{\mu \lambda'_{WF}}{N} \left[ \frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} \right]. \tag{C.28c}$$

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