

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 (*i.e.*, Wright-Fisher model) and a subdivided population of constant, infinite size. The cancellation result was later extended to heterogeneous populations (Rodrigues & Gardner, 2012, with synchronous generations and infinite population size), and other life-cycles, with generic 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. In this study, we will consider three different life-cycles: Wright-Fisher, where the whole population is renewed at each time step, and two Moran life-cycles (Birth-Death and Death-Birth), where a single individual dies and is replaced at each time step.

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). These studies also 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 (for models assuming infinite population sizes, or models concentrating on fixation probabilities). Often, these simplifying assumptions are a necessary step towards obtaining explicit analytical results. Although artificial, simple population structures (*e.g.*, regular graphs, or subdivided populations with demes of equal sizes) help reduce the dimensionality of the system under study, in particular when the structure of the population displays symmetries such that all sites behave the same way in expectation. Weak selection approximations are crucial for disentangling spatial moments (Lion, 2016), that is, changes in global *vs.* local frequen-

refs?

cies (though they can in some cases be relaxed, as in Mullan & Lehmann, 2014). Mutation, however, is usually ignored by classical models of inclusive fitness because these models assume infinite population sizes, so that there is no need to add mechanisms that restore genetic diversity (Tarnita & Taylor, 2014). In populations of finite size, this diversifying effect can be obtained thanks to mutation.

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)) does not have to be rare. Imperfect strategy transmission can alter evolutionary dynamics, in particular in spatially structured populations (see *e.g.*, Allen et al., 2012; Débarre, 2017, for graph-structured populations). Here, we want to explore the consequences of imperfect strategy transmission from parents to their offspring on the evolution of altruistic behavior in subdivided populations. 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.

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

## 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  $\nu$ , they become altruists, with probability  $1 - \nu$  they become defectors. We call the parameter  $\nu$  the mutation bias.

Social interactions take place within each deme; each individual interacts with the  $n - 1$  other deme members. We assume that social interactions affect individual fecundity, whose baseline is set equal to 1. Each interaction with an altruist increases an individual’s fecundity by  $\omega b$ ; altruists pay a fecundity cost

87  $\omega c$  ( $c \leq b$ ). The parameter  $\omega$  scales the relative effect of social interactions on  
 88 fecundity, and is assumed to be small ( $\omega \ll 1$ ).  
 89 Denoting by  $e_{ij}$  the interaction probability between individuals living at sites  $i$   
 90 and  $j$ , we have

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

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

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

93 Although our assumptions may seem restrictive: we assume that fecundity ben-  
 94 efits are unconditional, *i.e.*, the same which ever the type of the recipient, and  
 95 the fecundity effects are additive, *i.e.*, the effect of interacting with  $k$  altruists is  
 96  $k$  times the effect of interacting with one altruist. And yet, the same fecundities  
 97 are obtained with a generic fecundity function, after linearization, under the as-  
 98 sumption that altruists and defectors are phenotypically close (see Appendix A  
 99 for details).

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

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

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

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

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

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

## 119 2.2 Methods

### 120 2.2.1 Analytical part

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

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

136 Check results with regular results

### 137 2.2.2 Stochastic simulations

138 We also ran stochastic simulations (coded in C). The simulations were run for  $10^8$   
139 generations (one generation is one time step for the Wright-Fisher life-cycle, and  
140  $N$  time steps for the Moran life-cycles). For each set of parameters and life-cycle,  
141 we estimated the long-term frequency of altruists by sampling the population  
142 every  $10^3$  generations and computing the average frequency of altruists.

143 add Code availability, github? make repo public

## 144 3 Results

### 145 3.1 Probabilities of identity by descent

146 As we will see later, the expected frequencies of altruists in the population de-  
147 pend on probabilities of identity by descent of pairs of sites,  $Q_{ij}$ . Two individuals  
148 are said to be identical by descent if there has not been any mutation on either  
149 lineage since their common ancestor. Because of the structure of the popula-  
150 tion, there are only three types of pairs of individuals, and hence three different  
151 values of  $Q_{ij}$ :

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

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

### 153 3.1.1 Moran updating

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

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

156 Given the law of total probabilities, we first consider the site that was last up-  
157 dated (1/2 chance that it was  $j$  rather than  $i$ ); then we consider each potential  
158 parent  $k$ , weighted by the dispersal probabilities  $d_{kj}$ . Then the individuals at  
159 sites  $i$  and  $j$  are identical by descent (IBD) if  $i$  and  $j$ 's parent were IBD ( $Q_{ki}^M$ ) and  
160 if no mutation occurred ( $1-\mu$ ). Replacing the dispersal probabilities  $d_{ij}$  by their  
161 values (given in eq. (3)), we eventually obtain (see Appendix B.2 for calculation  
162 steps):

{eq:QM}

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

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

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

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

### 172 3.1.2 Wright-Fisher updating

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

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

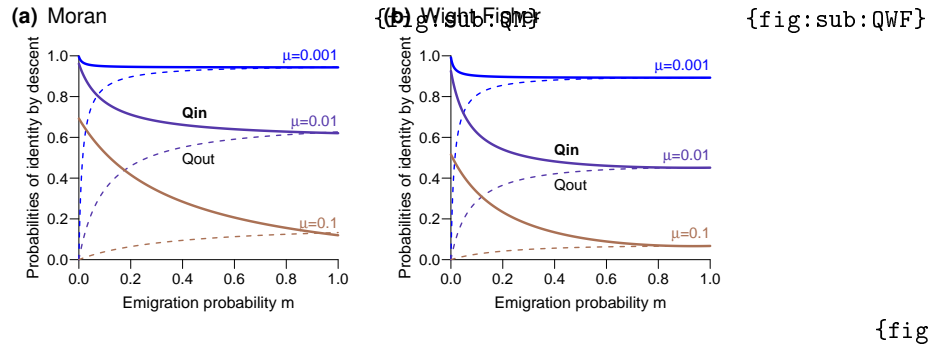
176 The sum is over all possible parents  $k$  and  $\ell$  of  $i$  and  $j$ , weighted by the disper-  
177 sal probabilities to sites  $i$  and  $j$ ; the individuals at sites  $i$  and  $j$  are identical by  
178 descent if their parents were ( $Q_{kl}$ ) and if neither mutated ( $(1-\mu)^2$ ).

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

{eq:QWF}

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

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



**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  $\mu$  (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.

181 with

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

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

186 The two probabilities of identity by descent go to 1 when the mutation prob-  
 187 ability  $\mu$  is very small ( $\mu \rightarrow 0$ ), except if we first assume that the number of demes  
 188 is very large ( $N_D \rightarrow \infty$ ); for instance, with this life-cycle as well,  $\lim_{N_D \rightarrow \infty} Q_{out}^{WF} =$   
 189 0.

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

## 193 3.2 Expected frequencies of altruists for each life-cycle

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

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

196 (Calculations leading to eq. (9) are presented in the Appendix B). The muta-  
197 tion bias  $v$  corresponds to the expected proportion of altruists in the popula-  
198 tion in the absence of selection (*i.e.*, when  $\omega = 0$ );  $\omega$  is the parameter that scales  
199 the effects of interactions between individuals and is assumed to be small; the  
200 subscript  $D$  refers to “direct” effects, and the subscript  $I$  to “indirect” effects.  
201 These indirect effects correspond to (kin) competition: by providing a benefit  
202 to a deme-mate and thereby increasing its fecundity, a focal altruist indirectly  
203 harms others by reducing their relative fecundity ( $\beta_I$  term in eq. (9)); by having  
204 a reduced fecundity due to the cost of altruism, a focal altruist indirectly favors  
205 others by increasing their relative fecundity ( $\gamma_I$  term).

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

### 208 3.2.1 Direct effects

209 Direct effects are similar for the three life-cycles; the only difference is the value  
210 of probabilities of identity by descent  $Q$  (as seen in the previous section, they  
211 differ between Moran and Wright-Fisher life-cycles):

$\{\text{eq:directeffects}\}$

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

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

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

212 For both benefits and costs, direct effects only count when there is no muta-  
213 tion ( $1 - \mu$ ). Direct effects of benefits  $b$  (eq. (10a) and eq. (10b)) only count if  
214 the interaction takes place with an individual who is identical by descent. With  
215 the population structure that we consider, social interactions only occur within  
216 demes, so only  $Q_{in}$  is present in eq. (10a) and eq. (10b). On the other hand, the  
217 direct effect of the fecundity cost  $c$  (eq. (10c)) does not depend on the type of  
218 interactant.

219 As seen in the previous section,  $Q_{in}^M$  and  $Q_{in}^{WF}$  decrease with the emigration  
220 probability  $m$  (actually only until  $m = \frac{d-1}{d}$  for the latter). Consequently, the mag-  
221 nitude of the direct (beneficial) effects of benefits  $b$  provided by altruists ( $\beta_D$ )  
222 decreases when the emigration probability  $m$  increases, while the direct (detrimental)  
223 effects ( $\gamma_D$ ) due to the direct cost of altruism  $c$  are constant. As a result,  
224 if we only consider direct effects, more emigration  $m$  is detrimental to the evo-  
225 lution of altruistic behaviour. But there are also indirect effects at play.



### 226 3.2.2 Indirect effects

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

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

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

236 (Calculation details are presented in Appendix B.)

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

244 Replacing  $Q_{\text{in}}$  and  $Q_{\text{out}}$  by their formula for the Moran life-cycle (eq. (6)), we  
245 conclude that  $\beta_I^{\text{BD}} = \gamma_I^{\text{BD}}$  are decreasing functions of the emigration probab-  
246 ility  $m$ .

Add Math-  
ematica

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

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

252 The brackets in eq. (11b) contain a sum of two terms. The first term corresponds  
253 two individuals from the same deme (with replacement) whose offspring either  
254 do not emigrate, or emigrate together to the same deme. The second term cor-  
255 responds to individuals initially from different demes who end up in the same  
256 deme (either one of their home demes, or a third deme).

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

259 Replacing  $Q_{in}$  and  $Q_{out}$  by their formulas given in eq. (6), we can conclude  
 260 that  $\beta_I^{DB} = \gamma_I^{DB}$  first decreases with the emigration probability  $m$ , and increases  
 261 again after a threshold value  $m'_c$  (given in the appendix;  $m'_c < (d-1)/d$ ).

Mathematica  
name

262 **Wright-Fisher** With this life-cycle, generations are synchronous and all indi-  
 263 viduals again all have the same survival probability (now equal to 0 at all sites).  
 264 As a result, the formulas for  $\beta_I^{WF}$  and  $\gamma_I^{WF}$  are the same as  $\beta_I^{DB}$  and  $\gamma_I^{DB}$ , except  
 265 that instead of  $Q_{in}^M$  and  $Q_{out}^M$ , we need to use  $Q_{in}^{WF}$  and  $Q_{out}^{WF}$  (given in eq. (8)). Once  
 266 this is done, we see that  $\beta_I^{WF} = \gamma_I^{WF}$  first decreases with the emigration probabil-  
 267 ity  $m$ , and increases again after the threshold value  $m_c^{WF} = (d-1)/d$  (which was  
 268 identified previously as the emigration probability such that offspring have an  
 269 equal chance of landing in their natal deme or in any other deme, *i.e.*,  $d_{in} = d_{out}$ ).

### 270 3.3 Identifying threshold values of the mutation probability $\mu$

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

#### 276 3.3.1 Moran Birth-Death

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

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

282 This result is illustrated in figure 2(b); with the parameters of the figure,  $\mu_c^{BD} \approx$   
 283 0.026.

#### 284 3.3.2 Moran Death-Birth

285 The relationship between  $\mathbb{E}[\bar{X}]$  and  $m$  is a bit more complicated for this life-  
 286 cycle. For simplicity, we concentrate on what happens starting from low emi-  
 287 gration probabilities (*i.e.*, the sign of the slope of  $\mathbb{E}[\bar{X}]$  as a function of  $m$  when  
 288  $m \rightarrow 0$ ). If the benefits  $b$  provided by altruists are relatively low ( $b < c(n+1)$ ),  
 289  $\mathbb{E}[\bar{X}]$  initially increases with  $m$  provided the mutation probability  $\mu$  is greater  
 290 than a threshold value  $\mu_c^{DB}$  given in eq. (13) below; otherwise, when the benefits

are high enough,  $\mathbb{E}[\bar{X}]$  initially increases with  $m$  for any value of  $\mu$ . Combining these results, we write

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

In figure 2(a), the parameters are such that  $\mu_c^{\text{DB}} = 0$ .

The expected frequency of altruists  $\mathbb{E}[\bar{X}]$  reaches a maximum at an emigration probability  $m_c^{\text{DB}}$  (whose complicated equation is in the [appendix](#)), as can be seen in figure 2(a). When the mutation probability gets close to 0 ( $\mu \rightarrow 0$ ),  $m_c^{\text{DB}}$  also gets close to 0,

what about when initially decreases with  $m$ ?

### 3.3.3 Wright-Fisher

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

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

do

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

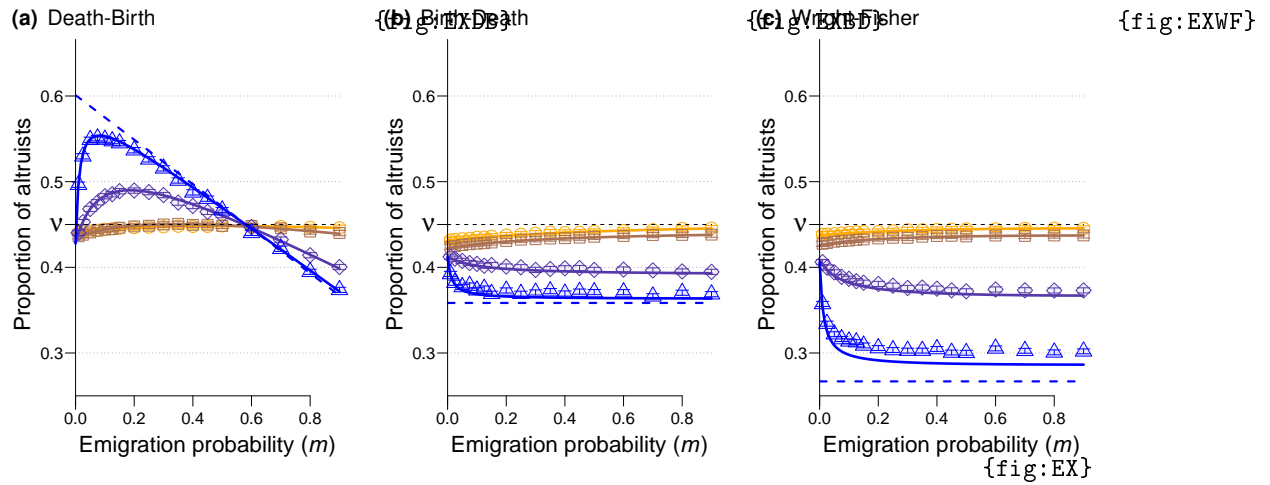
## 3.4 Relaxing key assumptions

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

Cite (Débarre, 2015)

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

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



**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:  $\omega = 0.005$ ,  $\nu = 0.45$ ,  $b = 15$ ,  $c = 1$ ,  $n = 4$  individuals per deme,  $N_D = 15$  demes.

## 320 4 Discussion

### 321 The expected frequency of altruists in a subdivided population 322 can increase with the probability of emigration

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

### 334 Quantitative vs. qualitative measures

335 We used a quantitative measure, the expected frequency of altruists in the popu-  
336 lation ( $\mathbb{E}[\bar{X}]$ ), to explore how non-zero mutation probabilities altered the impact  
337 of population viscosity. Often however, evolutionary success is measured quali-  
338 tatively, by comparing a quantity (an expected frequency, or, in models with no  
339 mutation, a probability of fixation) to the value it would have in the absence of  
340 selection. In our model, this amounts to saying that altruism is favored whenever  
341  $\mathbb{E}[\bar{X}] > v$  ( $v$  is plotted as a horizontal dashed line in figure 2). Some of our con-  
342 clusions change if we switch to this qualitative measure of evolutionary success:  
343 Under the Moran Birth-Death and Wright-Fisher life-cycles, population viscosity  
344 does not promote the evolution of altruism – actually, these two life-cycles can-  
345 not ever promote altruistic behavior for any regular population structure (Taylor  
346 et al., 2011), whichever the probability of mutation (Débarre, 2017). However,  
347 under a Moran Death-Birth life-cycle, altruism can be favored only at interme-  
348 diate emigration probabilities (figure 2(a)): increased emigration can still favor  
349 the evolution of altruism under this qualitative criterion.

### 350 The result is due to indirect, competitive, effects

351 To explain how the frequency of altruists can increase with the emigration prob-  
352 ability  $m$ , let us go back to the decomposition of the expected frequency of al-  
353 truists in the population  $\mathbb{E}[\bar{X}]$  into different terms (eq. (9)). For all the life-cycles  
354 that we consider, the direct effect of helping others ( $\beta_D$ ) decreases with emigra-  
355 tion  $m$ , while the direct effect of the cost of helping ( $\gamma_D$ ) does not change with  
356  $m$ . If we (erroneously) considered only direct effects, we would conclude that  
357 the expected proportion of altruists decreases with the emigration probability  
358  $m$ , because an increase in  $m$  reduces the probability that two interactants (two  
359 deme-mates in this model) are identical by descent. But this explanation ignores

indirect, competitive, effects. In the three life-cycles that we considered,  $\beta_I = \gamma_I$ , so the overall indirect effects are given by  $-(b - c)\beta_I$ . Hence, any increase of  $\mathbb{E}[\bar{X}]$  with  $m$  is driven by  $\beta_I$ . Indirect effects correspond to competition: helping another individual indirectly harms others – even the individual who is providing help is indirectly harmed. This competition can be diluted by increasing the emigration probability  $m$ . The overall effect of  $m$  on the expected frequency of altruists depends on the balance between direct and indirect effects. This balance depends on the fidelity of parent-offspring transmission ( $\mu$ ), in particular because probabilities of identity by descent depend on  $\mu$ .

Primary and secondary recipients West and Gardner 2010.

Competitive effects are less straightforward to explain than direct

## How small is small and how large is large?

Our results were derived under the assumption of weak selection, assuming that the phenotypic difference between altruists and defectors is small ( $\delta \ll 0$ ). We considered any fidelity of transmission (any  $\mu$  between 0 and 1) and population size. However, most models considering subdivided populations assume nearly perfect strategy transmission ( $\mu \rightarrow 0$ ) and infinite population sizes (number of demes  $N_D \rightarrow \infty$ ). The order in which these limits are taken matters, *i.e.*, one needs to specify how small  $\mu$ , but also  $\omega$ , are compared to the inverse size of the population. This remark complements findings by Sample & Allen (2017), who highlighted the quantitative differences between different orders of weak selection and large population limits.

## Imperfect transmission and Rebellious Children

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

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

Voter  
model  
Ayana  
Graphs  
et dire  
que on  
peut avoir  
 $e = d$ .  
Faire fig-  
ure avant  
de com-  
mencer a  
discuter

## 397 References

- 398 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*  
399 *tion* 62(6):1335–1344.
- 400 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,  
401 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*  
402 544(7649):227–230.
- 403 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects  
404 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.  
405 *Evolution of Cooperation.*
- 406 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-  
407 mediate disturbance. *Current Biology* 17(9):761–765.
- 408 Débarre, F. 2015: Fitness costs in spatially structured environments. *Evolution*  
409 69:1329–1335.
- 410 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution  
411 of social behavior in structured populations. *Journal of Theoretical Biology*  
412 420:26 – 35.
- 413 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-  
414 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*  
415 276(1654):13–19.
- 416 Frank, S. A. 1997: The price equation, fisher’s fundamental theorem, kin selec-  
417 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 418 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*  
419 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–  
420 2522.
- 421 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*  
422 *Theoretical Biology* 7(1):1 – 16.
- 423 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-  
424 tionary genetics. *Biosocial anthropology* 53:133–55.
- 425 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and  
426 harming on graphs: the return of the inclusive fitness effect. *Journal of Evolu-*  
427 *tionary Biology* 20(6):2284–2295.
- 428 Leturque, H. & Rousset, F. 2002: Dispersal, kin competition, and the ideal free  
429 distribution in a spatially heterogeneous population. *Theoretical Population*  
430 *Biology* 62(2):169 – 180.
- 431 Lion, S. 2016: Moment equations in spatial evolutionary ecology. *Journal of the-*  
432 *oretical biology* 405:46–57.

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

436 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule  
437 for the evolution of cooperation on graphs and social networks. *Nature*  
438 441(7092):502–505.

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

441 Sample, C. & Allen, B. 2017: The limits of weak selection and large population  
442 size in evolutionary game theory. *Journal of mathematical biology* pages 1–  
443 33.

444 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors  
445 in finite structured population models. *The American Naturalist* 184(4):477–  
446 488.

447 Taylor, P. 1992: Altruism in viscous populations—an inclusive fitness model. *Evo-*  
448 *lutionary ecology* 6(4):352–356.

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

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

453 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-  
454 bility in homogeneous structured populations. *Journal of Theoretical Biology*  
455 249(1):101 – 110.

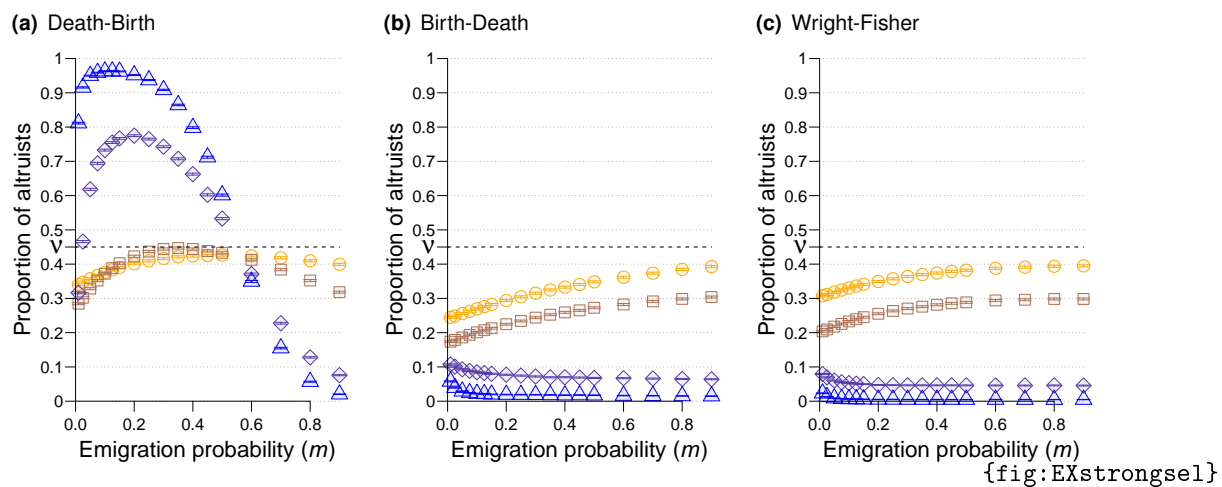
456 Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic  
457 behavior. *Evolution* 54(4):1135–1141.

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

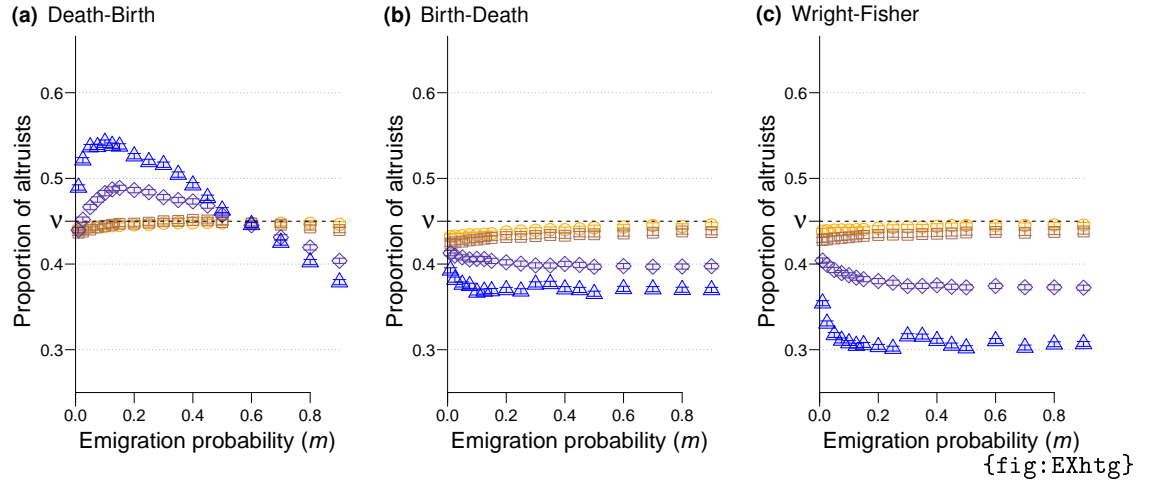
460 Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely  
461 viscous populations? *Evolutionary Ecology* 6(4):331–341.



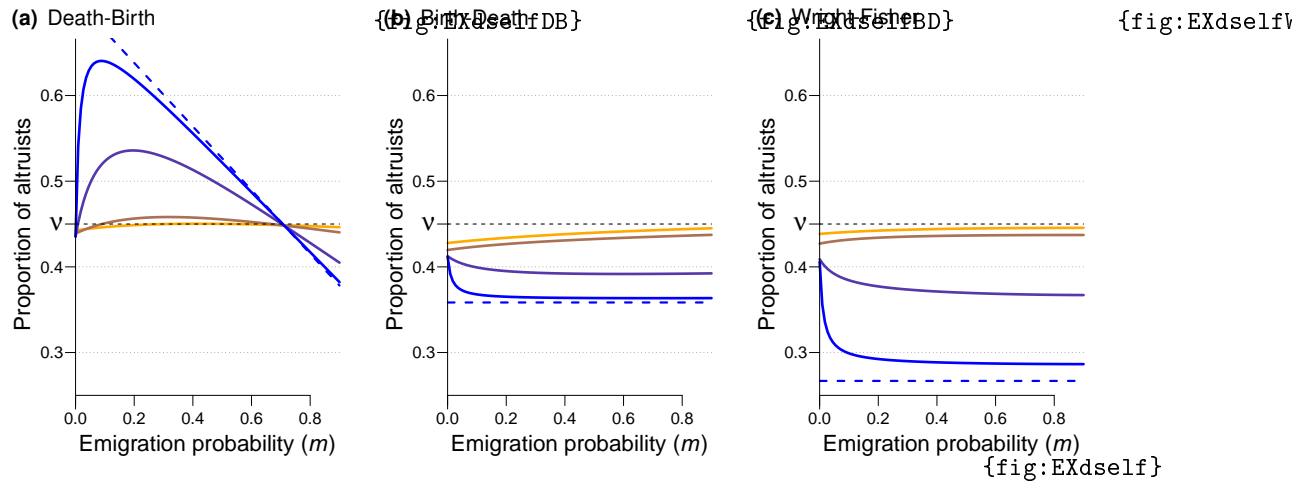
462 **Supplementary figures**



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



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



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

# Appendix

## 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).

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

$$F_i = \mathcal{F}(e_{1,i}(\phi^{(0)} + \omega X_1), \dots, e_{N,i}(\phi^{(0)} + \omega X_N); \phi^{(0)} + \omega X_i). \quad (\text{A.2}) \quad \{\text{eq:app:Fi}\}$$

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

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

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 Appendix 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. 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-cycles 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).

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

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

holds for all sites  $i$ . The structure of the population is also such that in the absence of selection ( $\omega = 0$ , so that  $f_i = 1$  for all sites  $1 \leq i \leq N$ ), all individuals have

514 the same probability of dying and the same probability of having successful off-  
 515 spring (*i.e.*, of having offspring that become adults at the next time step), so that  
 516

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

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

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

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

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

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

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

535 Now, we use the assumption of weak selection ( $\omega \ll 1$ ) and consider the first-  
 536 order expansion of eq. (B.7) for  $\omega$  close to 0. First, we note that in the absence  
 537 of selection ( $\omega = 0$ ), the population is at a mutation-drift balance; the expected  
 538 state of every site  $i$  is then  $\mathbb{E}_0[X_i] = \sum_{\mathbf{X} \in \Omega} X_i \xi(\mathbf{X}, 0, \mu) = \nu$  (recall that  $\nu$  is the mu-  
 539 tation bias parameter). Secondly, we further expand derivatives of  $B_{ji}$  and  $D_i$   
 540 thanks to the chain rule, using the variables  $f_k$  ( $1 \leq k \leq N$ ), corresponding to in-  
 541 dividual fecundities (also, recall that  $f_k = 1$  when  $\omega = 0$ ). Thirdly, we note that for  
 542 all the life-cycles that we consider, the total number of deaths in the population  
 543 during one time step does not depend on population composition (it is exactly  
 544 1 death for the Moran life-cycles, and exactly  $N$  for the Wright-Fisher life-cycle),  
 545 so that  $\sum_{i,j=1}^N B_{ij}$  does not depend on  $\omega$ . After simplification and reorganization,

546 the first order expansion of eq. (B.7) yields

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

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

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

550 In Appendix B.2, we will see that recursions on  $P_{ij}$  reveal that  $Q_{ij}$  can be inter-  
 551 preted as a probability of identity by descent, *i.e.*, the probability that the individ-  
 552 uals at sites  $i$  and  $j$  have a common ancestor and that no mutation has occurred  
 553 on either lineage since the ancestor.

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

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

556 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  
 557  $D_j$  terms by their formulas for each life-cycle (given in table S1), and the  $d_{ij}$  and  
 558  $e_{ij}$  terms by their formulas (given in eq. (3)) and eq. (1), respectively). For each  
 559 life-cycle we can group terms as

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

560 where D terms come from the numerators of  $B_{ij}$  and  $D_i$ , and I terms come from  
 561 the denominator of  $B_{ij}$  and  $D_i$ ; replacing  $B_{ij}$  and  $D_i$  by their formulas given in  
 562 table S1, we obtain the following sets of equations for each life-cycle:

{eq:EXBDsums}

### Moran Birth-Death

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

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

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

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

## Moran Death-Birth

{eq:EXDBsums}

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

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

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

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

## Wright-Fisher

{eq:EXWFsums}

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

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

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

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

System (B.14)s the same set of equations as for the Moran Death-Birth model (system (B.1)), except for the values of probabilities of identity by descent... that we now need to compute.

## 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  $E[\bar{X}]$ ,  $P_{ij}$  is the quantity that appears, but most studies use  $Q_{ij}$ . Both are evaluated in the absence of selection ( $\omega = 0$ ).

### B.2.1 Moran model

In a Moran model, exactly one individual dies and one individual reproduces during one time step. Given a state  $\mathbf{X}$  at time  $t$ , at time  $t + 1$  both sites  $i$  and  $j \neq i$  are occupied by altruists, if  $i$ ) it was the case at time  $t$  and neither site was replaced by a non-altruist (first term in eq. (B.15)), or  $ii$ ) if exactly one of the two sites was occupied by a non-altruist at time  $t$ , but the site was replaced by an

577 altruist (second and third terms of eq. (B.15)):

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

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

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

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

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

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

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

## 587 B.2.2 Wright-Fisher model

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

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

590 The first term of eq. (B.18) corresponds to both parents being altruists, and hav-  
591 ing altruist offspring; the second line corresponds to exactly one parent being  
592 altruist, and the third line to both parents being non-altruists (in this latter case,  
593 the two offspring have to be both mutants to be altruists).

594 Taking the expectation and simplifying, we obtain

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



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

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

596 Again,  $Q_{ij}$  corresponds to a probability of identity by descent: the individuals at  
597 sites  $i$  and  $j$  are identical by descent if their parents were and if neither mutated  
598  $((1 - \mu)^2)$ .

## 599 C In a subdivided population

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

601 Now, we need to adapt the results presented in Appendix B to our structure of  
 602 interest, a subdivided population, with dispersal and interaction probabilities  
 603 given by eq. (3) and eq. (1). For the  $\beta$  and  $\gamma$  terms, we use a brute-force ap-  
 604 proach, replacing  $d_{ij}$  and  $e_{ij}$  by their values in a subdivided population, and  
 605 simplifying the equations (for instance, there are 60 different cases to consider  
 606 for the four sums that appear in  $\beta_1^{\text{DB}}$ , shown in the table in section C.4 below).  
 607 The calculations and subsequent simplifications are detailed in the supplemen-  
 608 tary Mathematica file, and the results are presented in the main text (system (10),  
 609 eq. (11a), and eq. (11b)).

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

611 For the probabilities of identity by descent, we could also use a brute-force ap-  
 612 proach, but calculations are faster if we use formulas derived in Débarre (2017)  
 613 for “two-dimensional population structures”. The name comes from the fact that  
 614 we only need two types of transformations to go from any site to any other site  
 615 in the population: permutations on the deme index, and permutations on the  
 616 within-deme index.

617 We rewrite site labels ( $1 \leq i \leq N$ ) as  $(\ell_1, \ell_2)$ , where  $\ell_1$  is the index of the deme  
 618 ( $1 \leq \ell_1 \leq N_D$ ) and  $\ell_2$  the position of the site within the deme ( $1 \leq \ell_2 \leq n$ ). Then,  
 619 we introduce notations  $\tilde{d}_{i_1}^{i_2}$  and  $\tilde{Q}_{i_1}^{i_2}$ , that correspond to the dispersal probab-  
 620 ity and probability of identity by descent to a site at distances  $i_1$  and  $i_2$  in the  
 621 among-demes and within-deme dimensions (e.g.,  $\tilde{d}_{i_1}^{i_2} = d_{j_1 j_1 + i_1}^{j_2' j_2 + i_2}$ ).

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

#### 624 C.2.1 Moran model

625 In Débarre (2017), it was shown that

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

626 with

$$\tilde{D}_{q_1}^{q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1}^{\ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{C.21b}) \quad \{\text{eq:app:D2D}\}$$

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

$$\begin{aligned}\tilde{D}_{q_1} &= d_{\text{self}} + \sum_{\ell_2=1}^{N_2-1} d_{\text{in}} \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) + \sum_{\ell_1=1}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} d_{\text{out}} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2}(N_2 - 1) + (1 - \delta_{q_2})(-1)) d_{\text{in}} + (\delta_{q_1}(N_1 - 1) + (1 - \delta_{q_1})(-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}.\end{aligned}\quad (\text{C.22a})$$

629 ( $\delta_q$  is equal to 1 when  $q$  is equal to 0 modulo the relevant dimension, and to 0  
 630 otherwise). So for the three types of distances that we need to consider (distance  
 631 0, distance to another deme-mate, distance to individual in another deme), and  
 632 with  $N_1 = N_D$  and  $N_2 = n$ , we obtain {eq:app:Dsystem}

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

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

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

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

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

634 In particular,

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

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

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

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

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

638 With  $d_{\text{self}} = d_{\text{in}} = (1 - m)/n$ , we recover the equations given in the main text (sys-  
639 tem (6)).

### 640 C.3 Wright-Fisher

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

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

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

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

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

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

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

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

646 and

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

647 With  $d_{\text{self}} = d_{\text{in}} = (1 - m)/n$ , we recover the equations given in the main text (sys-  
648 tem (8)).

## C.4 Unpacking $\beta_I^{\text{DB}}$

{sec:app:betaI}

The table below contains all combinations for  $i, j, k, l$  involved in the four sums.  $(i, j)$ : means that  $i$  and  $j$  are different sites in the same deme;  $G_i$ : deme containing site  $i$ .

	$j$	$k$	$l$	Notation	Count	$d_{ji}$	$d_{li}$	$e_{kl}$	$Q_{jk}$
1	$j = i$	$k = i$	$l = i$	$(i = j = k = l)$	1	$d_{\text{self}}$	$d_{\text{self}}$	$e_{\text{self}}$	1
2	$j = i$	$k = i$	$l \neq i; l \in G_i$	$(i = j = k, l)$	$n - 1$	$d_{\text{self}}$	$d_{\text{in}}$	$e_{\text{in}}$	1
3	$j = i$	$k = i$	$l \notin G_i$	$(i = j = k), (l)$	$N - n$	$d_{\text{self}}$	$d_{\text{out}}$	$e_{\text{out}}$	1
4	$j = i$	$k \neq i; k \in G_i$	$l = i$	$(i = j = l, k)$	$n - 1$	$d_{\text{self}}$	$d_{\text{self}}$	$e_{\text{in}}$	$Q_{\text{in}}$
5	$j = i$	$k \neq i; k \in G_i$	$l = k$	$(i = j, k = l)$	$n - 1$	$d_{\text{self}}$	$d_{\text{in}}$	$e_{\text{self}}$	$Q_{\text{in}}$
6	$j = i$	$k \neq i; k \in G_i$	$l \neq i, k; l \in G_i$	$(i = j, k, l)$	$(n - 1)(n - 2)$	$d_{\text{self}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{in}}$
7	$j = i$	$k \neq i; k \in G_i$	$l \notin G_i$	$(i = j, k), (l)$	$(n - 1)(N - n)$	$d_{\text{self}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{in}}$
8	$j = i$	$k \notin G_i$	$l = i = j$	$(i = j = l), (k)$	$(N - n)$	$d_{\text{self}}$	$d_{\text{self}}$	$e_{\text{out}}$	$Q_{\text{out}}$
9	$j = i$	$k \notin G_i$	$l \neq i, l \in G_i$	$(i = j, l), (k)$	$(N - n)(n - 1)$	$d_{\text{self}}$	$d_{\text{in}}$	$e_{\text{out}}$	$Q_{\text{out}}$
10	$j = i$	$k \notin G_i$	$l = k$	$(i = j), (k = l)$	$(N - n)$	$d_{\text{self}}$	$d_{\text{out}}$	$e_{\text{self}}$	$Q_{\text{out}}$
11	$j = i$	$k \notin G_i$	$l \neq k; l \in G_k$	$(i = j), (k, l)$	$(N - n)(n - 1)$	$d_{\text{self}}$	$d_{\text{out}}$	$e_{\text{in}}$	$Q_{\text{out}}$
12	$j = i$	$k \notin G_i$	$l \notin G_i, G_k$	$(i = j), (k), (l)$	$(N - n)(N - 2n)$	$d_{\text{self}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
13	$j \neq i, j \in G_i$	$k = i$	$l = i$	$(i = k = l, j)$	$(n - 1)$	$d_{\text{in}}$	$d_{\text{self}}$	$e_{\text{self}}$	$Q_{\text{in}}$
14	$j \neq i, j \in G_i$	$k = i$	$l = j$	$(i = k, j = l)$	$(n - 1)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{in}}$
15	$j \neq i, j \in G_i$	$k = i$	$l \neq i, j; l \in G_i$	$(i = k, j, l)$	$(n - 1)(n - 2)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{in}}$
16	$j \neq i, j \in G_i$	$k = i$	$l \notin G_i$	$(i = k, j), (l)$	$(n - 1)(N - n)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{in}}$
17	$j \neq i, j \in G_i$	$k = j$	$l = i$	$(i = l, j = k)$	$(n - 1)$	$d_{\text{in}}$	$d_{\text{self}}$	$e_{\text{in}}$	1
18	$j \neq i, j \in G_i$	$k = j$	$l = j$	$(i, j = k = l)$	$(n - 1)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{self}}$	1
19	$j \neq i, j \in G_i$	$k = j$	$l \neq i, j; l \in G_i$	$(i, j = k, l)$	$(n - 1)(n - 2)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{in}}$	1
20	$j \neq i, j \in G_i$	$k = j$	$l \notin G_i$	$(i, j = k), (l)$	$(n - 1)(N - n)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{out}}$	1
21	$j \neq i, j \in G_i$	$k \neq i, j; k \in G_i$	$l = i$	$(i = l, j, k)$	$(n - 1)(n - 2)$	$d_{\text{in}}$	$d_{\text{self}}$	$e_{\text{in}}$	$Q_{\text{in}}$
22	$j \neq i, j \in G_i$	$k \neq i, j; k \in G_i$	$l = j$	$(i, j = l, k)$	$(n - 1)(n - 2)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{in}}$
23	$j \neq i, j \in G_i$	$k \neq i, j; k \in G_i$	$l = k$	$(i, j, k = l)$	$(n - 1)(n - 2)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{self}}$	$Q_{\text{in}}$
24	$j \neq i, j \in G_i$	$k \neq i, j; k \in G_i$	$l \neq i, j, k; l \in G_i$	$(i, j, k, l)$	$(n - 1)(n - 2)(n - 3)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{in}}$
25	$j \neq i, j \in G_i$	$k \neq i, j; k \in G_i$	$l \notin G_i$	$(i, j, k), (l)$	$(n - 1)(n - 2)(N - n)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{in}}$
26	$j \neq i, j \in G_i$	$k \notin G_i$	$l = i$	$(i = l, j), (k)$	$(n - 1)(N - n)$	$d_{\text{in}}$	$d_{\text{self}}$	$e_{\text{out}}$	$Q_{\text{out}}$
27	$j \neq i, j \in G_i$	$k \notin G_i$	$l = j$	$(i, j = l), (k)$	$(n - 1)(N - n)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{out}}$	$Q_{\text{out}}$
28	$j \neq i, j \in G_i$	$k \notin G_i$	$l \neq i, j; l \in G_i$	$(i, j, l), (k)$	$(n - 1)(N - n)(n - 2)$	$d_{\text{in}}$	$d_{\text{in}}$	$e_{\text{out}}$	$Q_{\text{out}}$
29	$j \neq i, j \in G_i$	$k \notin G_i$	$l = k$	$(i, j), (k = l)$	$(n - 1)(N - n)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{self}}$	$Q_{\text{out}}$
30	$j \neq i, j \in G_i$	$k \notin G_i$	$l \neq k; l \in G_k$	$(i, j), (k, l)$	$(n - 1)(N - n)(n - 1)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{in}}$	$Q_{\text{out}}$
31	$j \neq i, j \in G_i$	$k \notin G_i$	$l \notin G_i, G_k$	$(i, j), (k), (l)$	$(n - 1)(N - n)(N - 2n)$	$d_{\text{in}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
32	$j \notin G_i$	$k = i$	$l = i$	$(i = k = l), (j)$	$(N - n)$	$d_{\text{out}}$	$d_{\text{self}}$	$e_{\text{self}}$	$Q_{\text{out}}$
33	$j \notin G_i$	$k = i$	$l \neq i; l \in G_i$	$(i = k, l), (j)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{out}}$
34	$j \notin G_i$	$k = i$	$l = j$	$(i = k), (j = l)$	$(N - n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
35	$j \notin G_i$	$k = i$	$l \neq j; l \in G_j$	$(i = k), (j, l)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
36	$j \notin G_i$	$k = i$	$l \notin G_i, G_j$	$(i = k), (j), (l)$	$(N - n)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
37	$j \notin G_i$	$k \neq i; k \in G_i$	$l = i$	$(i = l, k), (j)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{self}}$	$e_{\text{in}}$	$Q_{\text{out}}$
38	$j \notin G_i$	$k \neq i; k \in G_i$	$l = k$	$(i, k = l), (j)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{self}}$	$Q_{\text{out}}$
39	$j \notin G_i$	$k \neq i; k \in G_i$	$l \neq i, k; l \in G_i$	$(i, k, l), (j)$	$(N - n)(n - 1)(n - 2)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{in}}$	$Q_{\text{out}}$
40	$j \notin G_i$	$k \neq i; k \in G_i$	$l = j$	$(i, k), (j = l)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
41	$j \notin G_i$	$k \neq i; k \in G_i$	$l \neq j; l \in G_j$	$(i, k), (j, l)$	$(N - n)(n - 1)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
42	$j \notin G_i$	$k \neq i; k \in G_i$	$l \notin G_i, G_j$	$(i, k), (j), (l)$	$(N - n)(n - 1)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
43	$j \notin G_i$	$k = j$	$l = i$	$(i = l), (j = k)$	$(N - n)$	$d_{\text{out}}$	$d_{\text{self}}$	$e_{\text{out}}$	1
44	$j \notin G_i$	$k = j$	$l \neq i; l \in G_i$	$(i, l), (j = k)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{out}}$	1
45	$j \notin G_i$	$k = j$	$l = j$	$(i), (j = k = l)$	$(N - n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{self}}$	1
46	$j \notin G_i$	$k = j$	$l \neq j; l \in G_j$	$(i), (j = k, l)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{in}}$	1
47	$j \notin G_i$	$k = j$	$l \notin G_i, G_j$	$(i), (j = k), (l)$	$(N - n)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	1
48	$j \notin G_i$	$k \neq j; k \in G_j$	$l = i$	$(i = l), (j, k)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{self}}$	$e_{\text{out}}$	$Q_{\text{in}}$
49	$j \notin G_i$	$k \neq j; k \in G_j$	$l \neq i; l \in G_i$	$(i, l), (j, k)$	$(N - n)(n - 1)(n - 1)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{out}}$	$Q_{\text{in}}$
50	$j \notin G_i$	$k \neq j; k \in G_j$	$l = j$	$(i), (j = l, k)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{in}}$	$Q_{\text{in}}$
51	$j \notin G_i$	$k \neq j; k \in G_j$	$l = k$	$(i), (j, k = l)$	$(N - n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{self}}$	$Q_{\text{in}}$
52	$j \notin G_i$	$k \neq j; k \in G_j$	$l \neq j, k; l \in G_j$	$(i), (j, k, l)$	$(N - n)(n - 1)(n - 2)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{in}}$	$Q_{\text{in}}$
53	$j \notin G_i$	$k \neq j; k \in G_j$	$l \notin G_i, G_j$	$(i), (j, k), (l)$	$(N - n)(n - 1)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{in}}$
54	$j \notin G_i$	$k \notin G_i, G_j$	$l = i$	$(i = l), (j), (k)$	$(N - n)(N - 2n)$	$d_{\text{out}}$	$d_{\text{self}}$	$e_{\text{out}}$	$Q_{\text{out}}$
55	$j \notin G_i$	$k \notin G_i, G_j$	$l \neq i; l \in G_i$	$(i, l), (j), (k)$	$(N - n)(N - 2n)(n - 1)$	$d_{\text{out}}$	$d_{\text{in}}$	$e_{\text{out}}$	$Q_{\text{out}}$
56	$j \notin G_i$	$k \notin G_i, G_j$	$l = j$	$(i), (j = l), (k)$	$(N - n)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
57	$j \notin G_i$	$k \notin G_i, G_j$	$l \neq j; l \in G_j$	$(i), (j, l), (k)$	$(N - n)(N - 2n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$
58	$j \notin G_i$	$k \notin G_i, G_j$	$l = k$	$(i), (j), (k = l)$	$(N - n)(N - 2n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{self}}$	$Q_{\text{out}}$
59	$j \notin G_i$	$k \notin G_i, G_j$	$l \neq k; l \in G_k$	$(i), (j), (k, l)$	$(N - n)(N - 2n)(n - 1)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{in}}$	$Q_{\text{out}}$
60	$j \notin G_i$	$k \notin G_i, G_j$	$l \notin G_i, G_j, G_k$	$(i), (j), (k), (l)$	$(N - n)(N - 2n)(N - 3n)$	$d_{\text{out}}$	$d_{\text{out}}$	$e_{\text{out}}$	$Q_{\text{out}}$