

Article Type: Letter

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

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

Short Running Title: Mutation and altruism in subdivided populations.

Keywords: Altruism, Subdivided population, Mutation, Migration, Cooperation, Island model.

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

Check ω removes entirely + explain delta in table

1 Abstract

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

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

18 **Impact Summary**

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

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 (1992a) 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

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

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

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

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

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

110 2 Model and methods

111 2.1 Assumptions

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

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

parent but instead get one randomly: with probability v , they become altruists,
with probability $1 - v$ they become defectors. We call the parameter v the muta-
tion bias.

An individual of type X_k expresses a social phenotype $\phi_k = \delta X_k$, where δ is
assumed to be small ($\delta \ll 1$). Social interactions take place within each deme;
each individual interacts with the $n - 1$ other deme members. Denoting by e_{ij}
the interaction probability between individuals living at sites i and j , we have
 $e_{ij} = 1/(n - 1)$ if i and j are two different sites in the same deme, and $e_{ij} = 0$
otherwise. We assume that social interactions affect individual fecundity; f_k de-
notes the fecundity of the individual at site k . The baseline fecundity, *i.e.* indi-
vidual fecundity when no altruists are present in the population, is set equal to 1.
We denote by b the marginal effect of a deme-mate's phenotype on the fecundity
of a focal individual, and by $-c$ the marginal effect of a focal individual's pheno-
type on its own fecundity ($c \leq b$). With these assumptions and notation, at the
first order in δ , the fecundity of the individual living at site k is given by

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

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

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

We denote by $B_i = B_i(\mathbf{X}, \delta)$ the expected number of successful offspring of the
individual living at site i (successful means alive at the next time step), and by

147 $D_i = D_i(\mathbf{X}, \delta)$ the probability that the individual living at site i dies. Both depend
 148 on the state of the population \mathbf{X} , but also on the way the population is updated
 149 from one time step to the next, *i.e.*, on the chosen life-cycle (also called updating
 150 rule). We will specifically explore three different life-cycles. At the beginning of
 151 each step of each life-cycle, all individuals produce offspring, that can be mu-
 152 tated; then these juveniles move, within the parental deme or outside of it, and
 153 land on a site. The next events occurring during the time step depend on the
 154 life-cycle:

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

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

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

163 2.2 Methods

164 2.2.1 Analytical part

165 To derive the expected (*i.e.*, long-term) proportion of altruists in the population,
 166 we use the toolbox presented in Débarre (2017), which is valid for any regular
 167 population and any life-cycle. Calculation details are given in Appendix A; they
 168 go as follows. First, we write an equation for the expected frequency of altruists
 169 in the population at time $t + 1$, conditional on the composition of the population
 170 at time t :

$$\mathbb{E}[\bar{X}(t+1)|\mathbf{X}(t)] = \frac{1}{N} \sum_{i=1}^N \underbrace{(B_i(1-\mu) + (1-D_i))}_{W_i} X_i + \frac{1}{N} \sum_{i=1}^N B_i \mu \nu. \quad (3)$$

171 The first term in eq. (3) correspond to unmutated offspring of altruists and sur-
 172 viving altruists, while the second term corresponds to new altruists created by
 173 mutation. We denote by W_i the expected number of individuals at the next time
 174 step who are either unmutated offspring of the individual living at site i , or this
 175 individual itself. We also note that the total number of births ($\sum_{i=1}^N B_i$) does
 176 not depend on population composition for the life-cycles that we consider (it
 177 is equal to 1 for the Moran life-cycles, and to N for Wright-Fisher.)

178 We then take the expectation of eq. (3) over all possible compositions of the pop-
 179 ulations, and consider large times t . After this, we write a first order expansion of
 180 the expression that we have obtained, using the assumption that the phenotypic
 181 difference δ is small (which also implies that selection is weak). By doing so, we
 182 let appear quantities that can be identified as neutral probabilities of identity
 183 by descent Q_{ij} , *i.e.*, the probability that individuals living at site i and j share
 184 a common ancestor and that no mutation occurred on either lineage since that
 185 ancestor, in a model with no selection ($\omega = 0$) – this is the “mutation definition”
 186 of identity by descent (Rousset & Billiard, 2000).

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

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

193 2.2.2 Stochastic simulations

194 We also ran stochastic simulations (coded in C). The simulations were run for 10^8
 195 generations (one generation is one time step for the Wright-Fisher life-cycle, and
 196 N time steps for the Moran life-cycles). For each set of parameters and life-cycle,
 197 using R (R Core Team, 2015), we estimated the long-term frequency of altruists

198 by sampling the population every 10^3 generations and computing the average
 199 frequency of altruists.

200 All scripts are available at

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

202 **3 Results**

203 **3.1 Probabilities of identity by descent**

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

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

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

211 **3.1.1 Moran updating**

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

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

214 Given the law of total probabilities, we first consider the site that was last up-
 215 dated (1/2 chance that it was j rather than i); then we consider each potential

parent k , weighted by the dispersal probabilities d_{kj} . Then the individuals at sites i and j are identical by descent (IBD) if i and j 's parent were IBD (Q_{ki}^M) and if no mutation occurred ($1 - \mu$). Replacing the dispersal probabilities d_{ij} by their values (given in eq. (2)), we eventually obtain (see Appendix A.2 for calculation steps):

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

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

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

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

3.1.2 Wright-Fisher updating

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

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

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

237 Replacing the dispersal probabilities d_{ij} by their values (given in eq. (2)) and
 238 skipping calculation steps (but see Appendix A.2 for details), we obtain:

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

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

239 with

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

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

242 In the Wright-Fisher life-cycle, $Q_{\text{in}}^{\text{WF}}$ decreases until $m = m_c^{\text{WF}} = \frac{d-1}{d}$, then in-
 243 creases again, while $Q_{\text{out}}^{\text{WF}}$ follows the opposite pattern. The threshold value m_c^{WF}
 244 corresponds to an emigration probability so high that an individual's offspring is
 245 as likely to land in its parent's deme as in any other deme (*i.e.*, $d_{\text{in}} = d_{\text{out}}$).

246 The two probabilities of identity by descent go to 1 when the mutation prob-
 247 ability μ is very small ($\mu \rightarrow 0$), except if we first assume that the number of demes
 248 is very large ($N_D \rightarrow \infty$); for instance, with this life-cycle as well, $\lim_{N_D \rightarrow \infty} Q_{\text{out}}^{\text{WF}} =$
 249 0.

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

253 3.2 Expected frequencies of altruists for each life-cycle

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

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

256 (Calculations leading to eq. (9) are presented in Appendix A.)

257 The mutation bias v corresponds to the expected proportion of altruists in the
258 population in the absence of selection (*i.e.*, when $\omega = 0$); ω is the parameter
259 that scales the effects of interactions between individuals, which is assumed to
260 be small. The subscript _D refers to “direct” effects, and the subscript _I to “in-
261 direct” effects. “Direct” effects involve effects on primary beneficiaries of the
262 benefits (b) and costs (c) of social interactions (West & Gardner, 2010), *i.e.*, so-
263 cial interactants (for the benefits b) and the focal individuals themselves (for the
264 costs c). “Indirect” effects corresponds to effects on secondary interactants, *i.e.*,
265 to (kin) competition. By providing a benefit to a deme-mate and thereby in-
266 creasing its fecundity, a focal altruist indirectly harms others by reducing their
267 relative fecundity (β_I term in eq. (9)); by having a reduced fecundity due to the
268 cost of altruism, a focal altruist indirectly favors others by increasing their rela-
269 tive fecundity (γ_I term).

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

272 3.2.1 Direct effects

273 Direct (/primary) effects are similar for the three life-cycles; the only difference
274 is the value of probabilities of identity by descent Q (as seen in the previous sec-

tion, they differ between Moran and Wright-Fisher life-cycles):

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

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

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

For both benefits and costs, direct effects only count when there is no mutation (hence the $(1 - \mu)$ factors). Direct effects of benefits b (eq. (10a) and eq. (10b)) only count if the interaction takes place with an individual who is identical by descent. With the population structure that we consider, social interactions only occur within demes, so only Q_{in} is present in eq. (10a) and eq. (10b). On the other hand, the direct effect of the fecundity cost c (eq. (10c)) does not depend on the type of interactant, since the same cost c is paid by altruists irrespective of the interactant's identity.

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

3.2.2 Indirect effects

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

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

$$\begin{aligned}\beta_1^{\text{BD}} &= (1 - m) \left(\frac{n-1}{n} Q_{\text{in}}^{\text{M}} + \frac{1}{n} \right) + m Q_{\text{out}}^{\text{M}} - \mu \frac{1 + (n-1)Q_{\text{in}}^{\text{M}} + n(d-1)Q_{\text{out}}^{\text{M}}}{nd} \\ &= \gamma_1^{\text{BD}}.\end{aligned}\quad (11a)$$

303 (Calculation details are presented in Appendix A.)

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

311 Replacing Q_{in} and Q_{out} by their formula for the Moran life-cycle (eq. (6)), we
 312 conclude that $\beta_1^{\text{BD}} = \gamma_1^{\text{BD}}$ are decreasing functions of the emigration probabil-
 313 ity m (calculations in the supplementary Mathematica file).

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

318 descent to it. We obtain

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

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

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

326 Replacing Q_{in} and Q_{out} by their formulas given in eq. (6), we can conclude
 327 that $\beta_I^{\text{DB}} = \gamma_I^{\text{DB}}$ first decreases with the emigration probability m , and increases
 328 again after a threshold value m'_c , which is smaller than $m_c^{\text{WF}} = (d-1)/d$ (calcu-
 329 lation details are presented in the supplementary Mathematica file).

330 **Wright-Fisher** With this life-cycle, generations are synchronous and all indi-
 331 viduals again all have the same survival probability (now equal to 0 at all sites).
 332 As a result, the formulas for β_I^{WF} and γ_I^{WF} are the same as β_I^{DB} and γ_I^{DB} , except
 333 that instead of Q_{in}^{M} and $Q_{\text{out}}^{\text{M}}$, we need to use $Q_{\text{in}}^{\text{WF}}$ and $Q_{\text{out}}^{\text{WF}}$ (given in eq. (8)). Once
 334 this is done, we see that $\beta_I^{\text{WF}} = \gamma_I^{\text{WF}}$ first decreases with the emigration probab-
 335 ity m , and increases again after the threshold value $m_c^{\text{WF}} = (d-1)/d$. This em-
 336 igration threshold was identified above as the emigration probability such that
 337 offspring have an equal chance of landing in their natal deme or in any other
 338 deme, *i.e.*, $d_{\text{in}} = d_{\text{out}}$ (calculation details are presented in the supplementary
 339 Mathematica file.)

340 3.3 Identifying threshold values of the mutation probability μ

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

347 3.3.1 Moran Birth-Death

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

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

353 This result is illustrated in figure 2(b); with the parameters of the figure, $\mu_c^{\text{BD}} \approx$
 354 0.026.

355 3.3.2 Moran Death-Birth

356 The relationship between $\mathbb{E}[\bar{X}]$ and m is a bit more complicated for this life-
 357 cycle. For simplicity, we concentrate on what happens starting from low emi-
 358 gration probabilities (*i.e.*, the sign of the slope of $\mathbb{E}[\bar{X}]$ as a function of m when
 359 $m \rightarrow 0$). If the benefits b provided by altruists are relatively low ($b < c(n + 1)$),
 360 $\mathbb{E}[\bar{X}]$ initially increases with m provided the mutation probability μ is greater
 361 than a threshold value μ_c^{DB} given in eq. (13) below; otherwise, when the benefits
 362 are high enough, $\mathbb{E}[\bar{X}]$ initially increases with m for any value of μ . Combining

363 these results, we write

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

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

365 The expected frequency of altruists $\mathbb{E}[\bar{X}]$ then reaches a maximum at an em-
 366 igration probability m_c^{DB} (whose complicated equation is given in the supple-
 367 imentary Mathematica file), as can be seen in figure 2(a). When the mutation
 368 probability gets close to 0 ($\mu \rightarrow 0$), m_c^{DB} also gets close to 0,

369 3.3.3 Wright-Fisher

370 The expected frequency of altruists in the population reaches an extremum when
 371 $m = m_c^{\text{WF}} = \frac{d-1}{d}$. This extremum is a maximum when the mutation probability
 372 is higher than a threshold value μ_c^{WF} given by

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

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

374 3.4 Relaxing key assumptions

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

381 To relax the assumption of equal deme sizes, we randomly drew deme sizes
 382 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_{out} remaining unchanged), confirms that this does affect our conclusions.

4 Discussion

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

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

Quantitative vs. qualitative measures

We used a quantitative measure, the expected frequency of altruists in the population ($\mathbb{E}[\bar{X}]$), to explore how non-zero mutation probabilities altered the impact

408 of population viscosity. Often however, evolutionary success is measured quali-
 409 tatively, by comparing a quantity (an expected frequency, or, in models with no
 410 mutation, a probability of fixation) to the value it would have in the absence of
 411 selection. In our model, this amounts to saying that altruism is favored whenever
 412 $\mathbb{E}[\bar{X}] > \nu$ (ν is plotted as a horizontal dashed line in figure 2). Some of our con-
 413 clusions change if we switch to this qualitative measure of evolutionary success:
 414 Under the Moran Birth-Death and Wright-Fisher life-cycles, population viscosity
 415 does not promote the evolution of altruism – actually, these two life-cycles can-
 416 not ever promote altruistic behavior for any regular population structure (Taylor
 417 et al., 2011), whichever the probability of mutation (Débarre, 2017). However,
 418 under a Moran Death-Birth life-cycle, altruism can be favored only at interme-
 419 diate emigration probabilities (figure 2(a)): increased emigration can still favor
 420 the evolution of altruism under this qualitative criterion.

421 **The result is due to indirect (/secondary) effects**

422 To explain how the frequency of altruists can increase with the emigration prob-
 423 ability m , let us go back to the decomposition of the expected frequency of al-
 424 truists in the population $\mathbb{E}[\bar{X}]$ into different terms (eq. (9)). For all the life-cycles
 425 that we consider, the direct effect of helping others (β_D) decreases with emigra-
 426 tion m , while the direct effect of the cost of helping (γ_D) does not change with
 427 m . If we (erroneously) considered only direct effects, we would conclude that
 428 the expected proportion of altruists decreases with the emigration probability
 429 m , because an increase in m reduces the probability that two interactants (two
 430 deme-mates in this model) are identical by descent. But this explanation ignores
 431 indirect, competitive, effects. In the three life-cycles that we considered, $\beta_I = \gamma_I$,
 432 so the overall indirect effects are given by $-(b - c)\beta_I$. Hence, any increase of
 433 $\mathbb{E}[\bar{X}]$ with m is driven by β_I . Indirect effects correspond to competition: helping
 434 another individual indirectly harms others – even the individual who is provid-

ing 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 (μ), in particular because probabilities of identity by descent depend on μ .

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

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 μ 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 μ , but also ω , 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.

459 **Imperfect transmission and Rebellious Children**

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

466 Mutation was also previously included in models investigating the mainte-
467 nance of cooperative microorganisms in the presence of cheaters (Brockhurst
468 et al., 2007; Frank, 2010). In both of these models however, only loss-of-function
469 mutation was considered, which corresponds to setting the mutation bias at
470 $\nu = 0$ in our model. This means that the all-cheaters state is absorbing; no matter
471 how favored cooperators may otherwise be, in the long run, a finite population
472 will only consist of cheaters.

473 **Interactions and dispersal graphs**

474 Compared to graphs classically used in evolutionary graph theory (*e.g.*, regular
475 random graphs, grids), the island model is particular because the interaction
476 graph and the dispersal graph are different: interactions take place only within
477 demes ($e_{\text{out}} = 0$), while offspring can disperse out of their natal deme ($d_{\text{out}} > 0$).
478 One may wonder whether our result depends on this difference between the two
479 graphs. Figure S4 shows that the result still holds when the dispersal and inter-
480 action graphs are the same. In this figure indeed, we let a proportion m (equal
481 to the dispersal probability) of interactions occur outside of the deme where the
482 individuals live, and set d_{self} , the probability of self replacement, equal to 0, so
483 that the dispersal and interactions graphs are the same.

484 **Coevolution of dispersal and social behavior**

485 This work also raises the question of what would happen if dispersal (*e.g.*, the
486 emigration probability m) could evolve as well. Recent work on the topic has
487 shown that under some conditions disruptive selection could take place, lead-
488 ing to a polymorphism between sessile altruists and mobile defectors (Parvinen,
489 2013; Mullon et al., 2017). The assumptions of these studies however differ from
490 ours in important ways, in that they consider continuous traits and use an adap-
491 tive dynamics framework, where, notably, mutations are assumed to be very
492 rare. It remains to be investigated how non-rare and potentially large mutations
493 would affect their result.

494 References

- 495 Alizon, S. & Taylor, P. 2008: Empty sites can promote altruistic behavior. *Evolu-*
 496 *tion* 62(6):1335–1344.
- 497 Allen, B.; Lippner, G.; Chen, Y.-T.; Fotouhi, B.; Momeni, N.; Yau, S.-T. & Nowak,
 498 M. A. 2017: Evolutionary dynamics on any population structure. *Nature*
 499 544(7649):227–230.
- 500 Allen, B.; Traulsen, A.; Tarnita, C. E. & Nowak, M. A. 2012: How mutation affects
 501 evolutionary games on graphs. *Journal of Theoretical Biology* 299:97 – 105.
 502 *Evolution of Cooperation.*
- 503 Brockhurst, M. A.; Buckling, A. & Gardner, A. 2007: Cooperation peaks at inter-
 504 mediate disturbance. *Current Biology* 17(9):761–765.
- 505 Cockerham, C. C. & Weir, B. 1987: Correlations, descent measures: drift with
 506 migration and mutation. *Proceedings of the National Academy of Sciences*
 507 84(23):8512–8514.
- 508 Débarre, F. 2017: Fidelity of parent-offspring transmission and the evolution
 509 of social behavior in structured populations. *Journal of Theoretical Biology*
 510 420:26 – 35.
- 511 Débarre, F.; Hauert, C. & Doebeli, M. 2014: Social evolution in structured popu-
 512 lations. *Nature Communications* 5.
- 513 Fletcher, J. A. & Doebeli, M. 2009: A simple and general explanation for the evo-
 514 lution of altruism. *Proceedings of the Royal Society B: Biological Sciences*
 515 276(1654):13–19.
- 516 Frank, S. A. 1997: The Price equation, Fisher’s fundamental theorem, kin selec-
 517 tion, and causal analysis. *Evolution* 51(6):1712–1729.
- 518 Frank, S. A. 2010: Microbial secretor–cheater dynamics. *Philosophical Transac-*
 519 *tions of the Royal Society of London B: Biological Sciences* 365(1552):2515–
 520 2522.
- 521 Grafen, A. & Archetti, M. 2008: Natural selection of altruism in inelastic viscous
 522 homogeneous populations. *Journal of Theoretical Biology* 252(4):694 – 710.
- 523 Hamilton, W. 1964: The genetical evolution of social behaviour. i. *Journal of*
 524 *Theoretical Biology* 7(1):1 – 16.
- 525 Hamilton, W. D. 1975: Innate social aptitudes of man: an approach from evolu-
 526 tionary genetics. *Biosocial anthropology* 53:133–55.
- 527 Lehmann, L.; Keller, L. & Sumpter, D. J. T. 2007: The evolution of helping and
 528 harming on graphs: the return of the inclusive fitness effect. *Journal of Evolu-*
 529 *tionary Biology* 20(6):2284–2295.

530 Lehmann, L. & Rousset, F. 2014: The genetical theory of social behaviour. Philo-
531 sophical Transactions of the Royal Society of London B: Biological Sciences
532 369(1642).

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

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

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

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

543 Ohtsuki, H.; Hauert, C.; Lieberman, E. & Nowak, M. A. 2006: A simple rule
544 for the evolution of cooperation on graphs and social networks. *Nature*
545 441(7092):502–505.

546 Parvinen, K. 2013: Joint evolution of altruistic cooperation and dispersal in a
547 metapopulation of small local populations. *Theoretical population biology*
548 85:12–19.

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

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

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

556 Sample, C. & Allen, B. 2017: The limits of weak selection and large population
557 size in evolutionary game theory. *Journal of mathematical biology* pages 1–
558 33.

559 Tarnita, C. E. & Taylor, P. D. 2014: Measures of relative fitness of social behaviors
560 in finite structured population models. *The American Naturalist* 184(4):477–
561 488.

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

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

- 566 Taylor, P. D. 1992b: Inclusive fitness in a homogeneous environment. Pro-
567 ceedings of the Royal Society of London. Series B: Biological Sciences
568 249(1326):299–302.
- 569 Taylor, P. D.; Day, T. & Wild, G. 2007a: Evolution of cooperation in a finite homo-
570 geneous graph. *Nature* 447(7143):469–472.
- 571 Taylor, P. D.; Day, T. & Wild, G. 2007b: From inclusive fitness to fixation proba-
572 bility in homogeneous structured populations. *Journal of Theoretical Biology*
573 249(1):101 – 110.
- 574 Taylor, P. D. & Irwin, A. J. 2000: Overlapping generations can promote altruistic
575 behavior. *Evolution* 54(4):1135–1141.
- 576 Van Cleve, J. 2015: Social evolution and genetic interactions in the short and long
577 term. *Theoretical Population Biology* 103:2 – 26.
- 578 West, S. A. & Gardner, A. 2010: Altruism, spite, and greenbeards. *Science*
579 327(5971):1341–1344.
- 580 West, S. A.; Pen, I. & Griffin, A. S. 2002: Cooperation and competition between
581 relatives. *Science* 296(5565):72–75.
- 582 Wilson, D. S.; Pollock, G. B. & Dugatkin, L. A. 1992: Can altruism evolve in purely
583 viscous populations? *Evolutionary Ecology* 6(4):331–341.
- 584 Wolfram Research, Inc. 2017: *Mathematica*, Version 11.1. Champaign, IL, 2017.

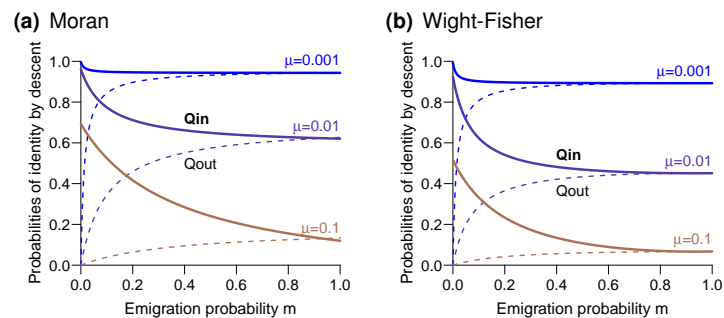


Figure 1: Probabilities of identity by descent, for two different individuals within the same deme (Q_{in} , full curves) and two individuals in different demes (Q_{out} , dashed curves), as a function of the emigration probability m , for different values of the mutation probability μ (0.001, 0.01, 0.1), and for the two types of life-cycles ((a): Moran, (b): Wright-Fisher). Other parameters: $n = 4$ individuals per deme, $N_D = 15$ demes.

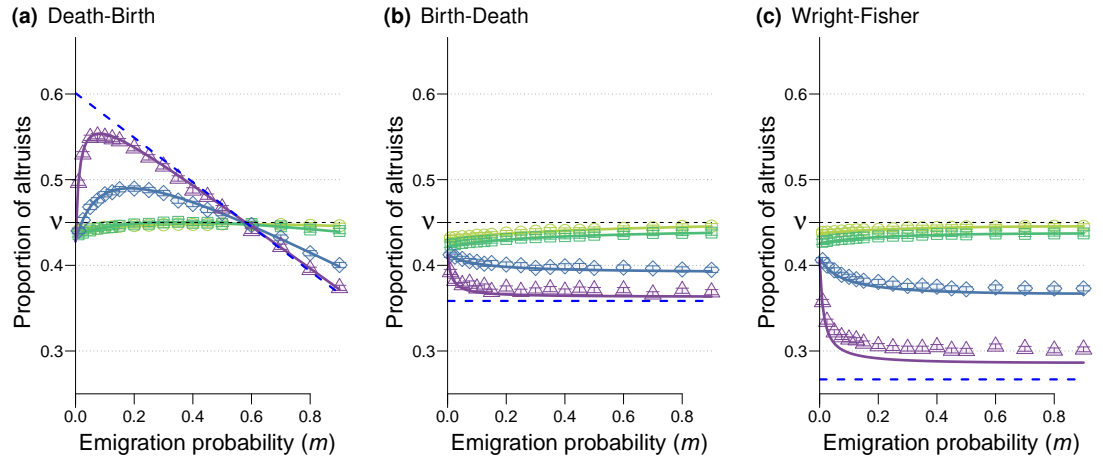


Figure 2: Expected proportion of altruists under weak selection, as a function of the emigration probability m , for different mutation values ($\mu = 0.001$ (blue, dots), 0.01 (purple, squares), 0.1 (brown, diamonds), 0.25 (orange, triangles); the dashed blue lines correspond to $\mu = 0$) and different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

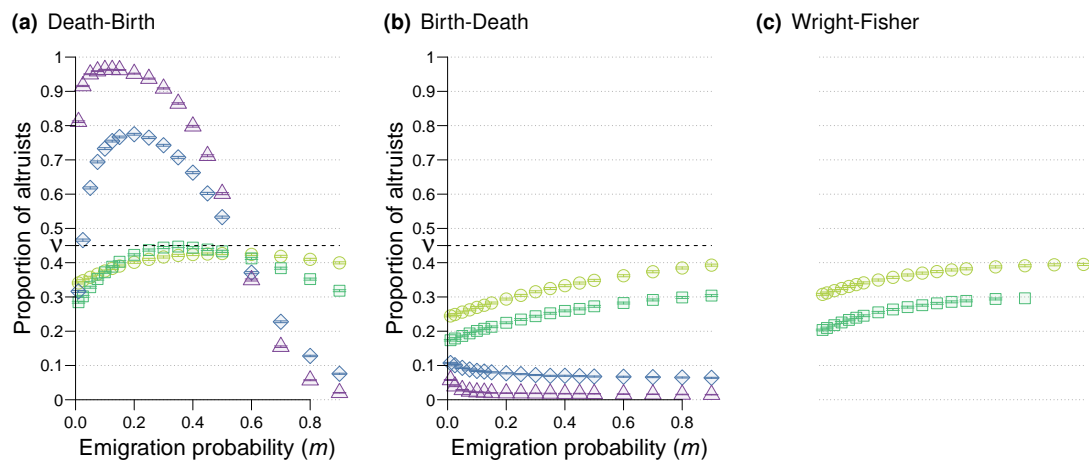


Figure S1: Equivalent of figure 2 (simulations only) but with strong selection ($\omega = 0.1$); please note the change of scale on the vertical axis. All other parameters and legends are identical to those of figure 2 (increasing mutation probabilities from blue dots to orange triangles).

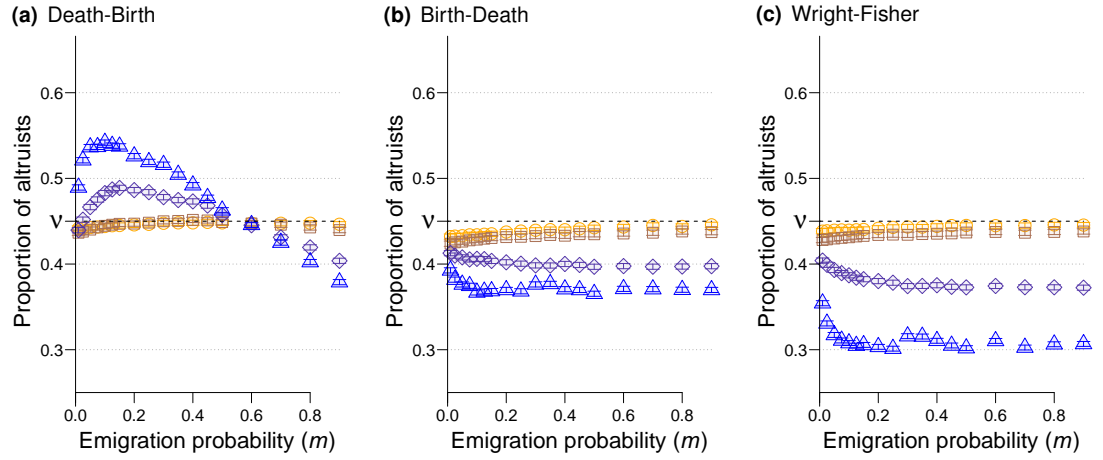


Figure S2: Equivalent of figure 2 (simulations only) but with a heterogeneous population structure: deme sizes range from 1 to 5 individuals per deme, the average deme size is 4 as in figure 2; all other parameters and legend are identical to those of figure 2.

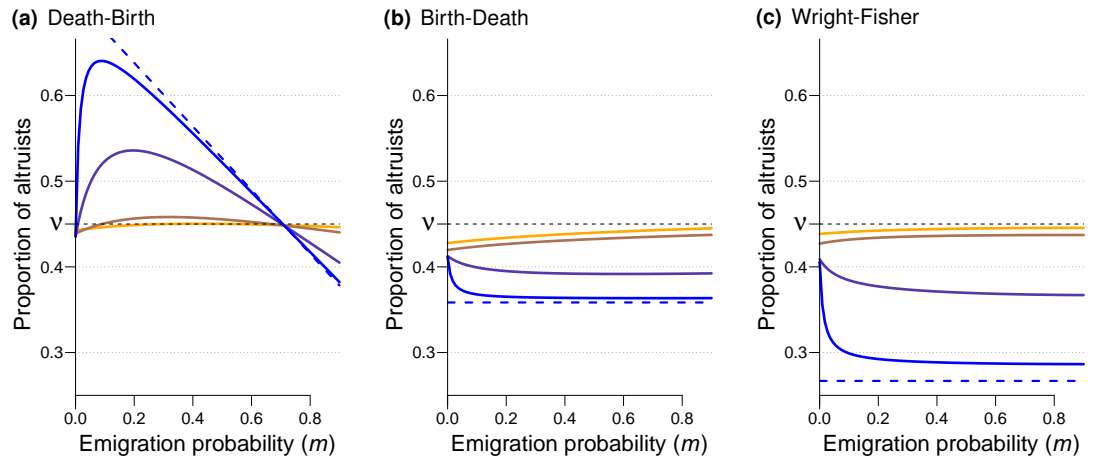


Figure S3: Equivalent of figure 2 (analysis only), with no self-replacement ($d_{ii} = d_{\text{self}} = 0$ for all sites).

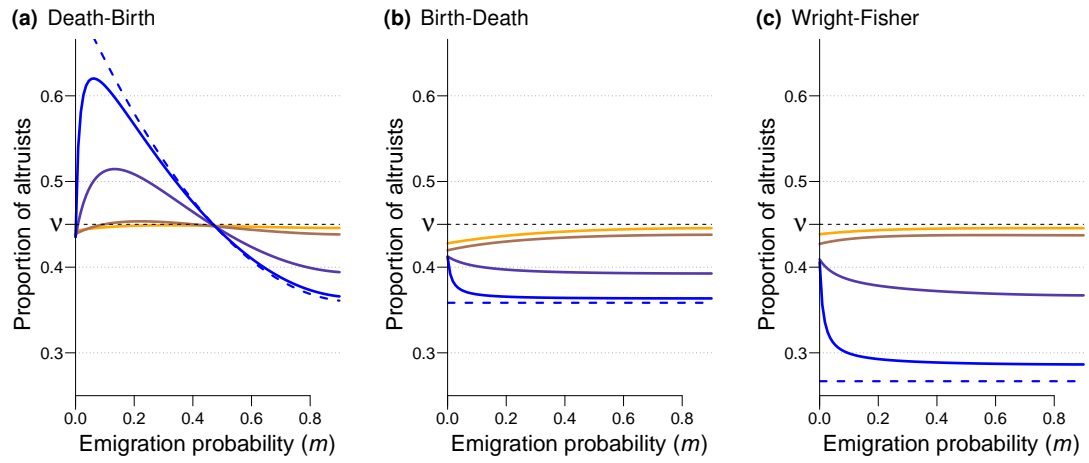


Figure S4: Equivalent of figure 2 (analysis only), with equal dispersal and interaction graphs (*i.e.*, no self-replacement [$d_{ii} = d_{\text{self}} = 0$ for all sites], and a proportion m of the interactions occurring outside of the home deme).

b	Fecundity benefit given by altruists to social interactants
c	Fecundity cost paid by altruists
d_{ij}	Dispersal probability from site i to site j
e_{ij}	Interaction probability from site i to site j
n	Deme size
N_D	Number of demes
N	Total population size ($N = N_D n$)
m	Emigration probability
Q_{ij}	(Long-term) Probability of identity by descent of individuals at sites i and j
X_i	Indicator variable, equal to 1 if site i is occupied by an altruist, to 0 otherwise (r.v.)
\bar{X}	Frequency of altruists in the population (r.v.)
β	Term associated to the benefits b
γ	Term associated to the costs c
μ	Mutation probability
ν	Mutation bias: probability that mutant is altruist
ω	Parameter scaling the relative effect of social interactions on fecundity
D	Subscript corresponding to direct/primary effects
I	Subscript corresponding to indirect/secondary effects
in	Subscript used when $i \neq j$ and the two sites are in the same deme
out	Subscript used when the two sites i and j are in different demes
self	Subscript used when $i = j$
BD	Superscript corresponding to the Moran Birth-Death model
DB	Superscript corresponding to the Moran Death-Birth model
M	Superscript corresponding to a Moran model
WF	Superscript corresponding to the Wright-Fisher model

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

Appendix

A Expected frequency of altruists

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 A.

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.

A.1 For a generic life-cycle

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 S2 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

Table S2: 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. (1).

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

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

610 holds for all sites i . The structure of the population is also such that in the ab-
 611 sence of selection ($\omega = 0$, so that $f_i = 1$ for all sites $1 \leq i \leq N$), all individuals have
 612 the same probability of dying and the same probability of having successful off-
 613 spring (*i.e.*, of having offspring that become adults at the next time step), so that
 614

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

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

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

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

622 The first term within the brackets corresponds to births: the type of the individ-
 623 ual living at i at time $t + 1$ depends on the type of its parent (living at site j), and
 624 on whether mutation occurred. The second term in the brackets of eq. (A.2a)
 625 corresponds to the survival of the individual living at site i .

626 Given that there is no absorbing population state (a lost strategy can always
 627 be recreated by mutation), there is a stationary distribution of population states;
 628 the expected frequency of altruists does not change anymore for large times t
 629 (realized frequencies of course keep changing). We denote by $\xi(\mathbf{X}, \omega, \mu)$ the prob-
 630 ability that the population is in state \mathbf{X} , given the strength of selection ω and the
 631 mutation probability μ . Taking the expectation of eq. (A.2a) ($\mathbb{E}[\bar{X}] = \sum_{\mathbf{X} \in \Omega} \bar{X} \xi(\mathbf{X}, \omega, \mu)$),
 632 we obtain, after reorganizing:

$$0 = \frac{1}{N} \sum_{\mathbf{X} \in \Omega} \sum_{i=1}^N \left[\sum_{j=1}^N B_{ij} (X_j(1-\mu) + \mu\nu) - D_i X_i \right] \xi(\mathbf{X}, \omega, \mu). \quad (\text{A.3})$$

633 Now, we use the assumption of weak selection ($\omega \ll 1$) and consider the first-
 634 order expansion of eq. (A.3) for ω close to 0. First, we note that in the absence
 635 of selection ($\omega = 0$), the population is at a mutation-drift balance; the expected
 636 state of every site i is then $\mathbb{E}_0[X_i] = \sum_{\mathbf{X} \in \Omega} X_i \xi(\mathbf{X}, 0, \mu) = \nu$ (recall that ν is the mu-
 637 tation bias parameter). Secondly, we further expand derivatives of B_{ji} and D_i
 638 thanks to the chain rule, using the variables f_k ($1 \leq k \leq N$), corresponding to in-
 639 dividual fecundities (also, recall that $f_k = 1$ when $\omega = 0$). Thirdly, we note that for

all the life-cycles that we consider, the total number of deaths in the population during one time step does not depend on population composition (it is exactly 1 death for the Moran life-cycles, and exactly N for the Wright-Fisher life-cycle), so that $\sum_{i,j=1}^N B_{ij}$ does not depend on ω . After simplification and reorganization, the first order expansion of eq. (A.3) yields

$$0 = \frac{1}{N} \sum_{i,k=1}^N \left[\frac{\partial \left(\sum_{j=1}^N (1-\mu) B_{ji} - D_i \right)}{\partial f_k} \right]_{f_k=1} \times \left(\sum_{\ell=1}^N e_{\ell k} b \sum_{X \in \Omega} X_{\ell} X_i \xi(\mathbf{X}, 0, \mu) - c \sum_{X \in \Omega} X_k X_i \xi(\mathbf{X}, 0, \mu) \right) - B^* \mu \frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} + O(\omega^2). \quad (\text{A.4})$$

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 expected state of the pair of sites (i, j) , evaluated in the absence of selection ($\omega = 0$). We can also replace these terms by

$$P_{ij} = v^2 + v(1-v)Q_{ij}. \quad (\text{A.5})$$

In Appendix A.2, we will see that recursions on P_{ij} reveal that Q_{ij} can be interpreted as a probability of identity by descent, *i.e.*, the probability that the individuals at sites i and j have a common ancestor and that no mutation has occurred on either lineage since the ancestor.

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

$$\mathbb{E}[\bar{X}] = v + \omega \frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} + O(\omega^2), \quad (\text{A.6})$$

where $\frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0}$ is obtained from eq. (A.4). We then need to replace the B_{ij} and D_j terms by their formulas for each life-cycle (given in table S2), and the d_{ij} and e_{ij} terms by their formulas (given in eq. (2)) and eq. (??), respectively). For each life-cycle we can group terms as

$$\frac{\partial \mathbb{E}[\bar{X}]}{\partial \omega} \Big|_{\omega=0} \approx \frac{v(1-v)}{\mu} [b(\beta_D - \beta_I) - c(\gamma_D - \gamma_I)], \quad (\text{A.7})$$

where D terms come from the numerators of B_{ij} and D_i , and I terms come from the denominator of B_{ij} and D_i ; replacing B_{ij} and D_i by their formulas given in table S2, we obtain the following sets of equations for each life-cycle:

Moran Birth-Death

$$\beta_D^{BD} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^M, \quad (\text{A.8a})$$

$$\beta_I^{BD} = \sum_{j,k,\ell=1}^N \left(\frac{d_{\ell j}}{N} - \frac{\mu}{N^2} \right) e_{k\ell} Q_{jk}^M, \quad (\text{A.8b})$$

$$\gamma_D^{BD} = 1 - \mu, \quad (\text{A.8c})$$

$$\gamma_I^{BD} = \sum_{j,k=1}^N \left(\frac{d_{kj}}{N} - \frac{\mu}{N^2} \right) Q_{jk}^M. \quad (\text{A.8d})$$

Moran Death-Birth

$$\beta_D^{DB} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^M, \quad (\text{A.9a})$$

$$\beta_I^{DB} = (1 - \mu) \sum_{i,j,k,\ell=1}^N \frac{d_{ji} d_{\ell i}}{N} e_{k\ell} Q_{jk}^M, \quad (\text{A.9b})$$

$$\gamma_D^{DB} = 1 - \mu, \quad (\text{A.9c})$$

$$\gamma_I^{DB} = (1 - \mu) \sum_{i,j,k=1}^N \frac{d_{ji} d_{ki}}{N} Q_{jk}^M. \quad (\text{A.9d})$$

Wright-Fisher

$$\beta_D^{WF} = \sum_{k,\ell=1}^N \frac{1-\mu}{N} e_{k\ell} Q_{\ell k}^{WF}, \quad (\text{A.10a})$$

$$\beta_I^{WF} = (1 - \mu) \sum_{i,j,k,\ell=1}^N \frac{d_{ji} d_{\ell i}}{N} e_{k\ell} Q_{jk}^{WF}, \quad (\text{A.10b})$$

$$\gamma_D^{WF} = 1 - \mu, \quad (\text{A.10c})$$

$$\gamma_I^{WF} = (1 - \mu) \sum_{i,j,k=1}^N \frac{d_{ji} d_{ki}}{N} Q_{jk}^{WF}. \quad (\text{A.10d})$$

System (A.10) is the same set of equations as for the Moran Death-Birth model (system (A.1)), except for the values of probabilities of identity by descent... that we now need to compute.

A.2 Probabilities of identity by descent

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

669 A.2.1 Moran model

670 In a Moran model, exactly one individual dies and one individual reproduces
 671 during one time step. Given a state \mathbf{X} at time t , at time $t + 1$ both sites i and
 672 $j \neq i$ are occupied by altruists, if i it was the case at time t and neither site was
 673 replaced by a non-altruist (first term in eq. (A.11)), or ij if exactly one of the two
 674 sites was occupied by a non-altruist at time t , but the site was replaced by an
 675 altruist (second and third terms of eq. (A.11)):

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & X_i X_j \left(1 - \sum_{k=1}^N \frac{1}{N} (d_{ki} + d_{kj}) ((1 - X_k)(1 - \mu) + \mu(1 - \nu)) \right) \\ & + X_i(1 - X_j) \sum_{k=1}^N \frac{1}{N} d_{kj} (X_k(1 - \mu) + \mu\nu) \\ & + X_j(1 - X_i) \sum_{k=1}^N \frac{1}{N} d_{ki} (X_k(1 - \mu) + \mu\nu). \end{aligned} \quad (\text{A.11})$$

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

$$P_{ij} = \frac{1}{2} \left(\sum_{k=1}^N (1 - \mu) (d_{kj} P_{ki} + d_{ki} P_{kj}) \right) + \mu\nu^2 \quad (i \neq j), \quad (\text{A.12})$$

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

679 Now we substitute $P_{ij} = \nu^2 + \nu(1 - \nu)Q_{ij}$ in eq. (A.12), we obtain

$$Q_{ij} = \frac{1}{2} \sum_{k=1}^N (1 - \mu) (d_{ki} Q_{kj} + d_{kj} Q_{ki}), \quad (\text{A.13})$$

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

685 A.2.2 Wright-Fisher model

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

$$\begin{aligned} \mathbb{E}[X_i X_j(t+1) | X(t) = \mathbf{X}] = & \sum_{k, \ell=1}^N d_{ki} d_{\ell j} \left(X_k X_\ell (1 - \mu + \mu\nu)^2 \right. \\ & + (X_k(1 - X_\ell) + (1 - X_k)X_\ell) (1 - \mu + \mu\nu)(\mu\nu) \\ & \left. + (1 - X_k)(1 - X_\ell)(\mu\nu)^2 \right) \end{aligned} \quad (\text{A.14})$$

688 The first term of eq. (A.14) corresponds to both parents being altruists, and hav-
 689 ing altruist offspring; the second line corresponds to exactly one parent being
 690 altruist, and the third line to both parents being non-altruists (in this latter case,
 691 the two offspring have to be both mutants to be altruists).
 692 Taking the expectation and simplifying, we obtain

$$P_{ij} = \sum_{k,\ell=1}^N (P_{kl}(1-\mu)^2) + (2-\mu)\mu v^2. \quad (\text{A.15})$$

693 Replacing P_{ij} by $v^2 + v(1-v)Q_{ij}$, eq. (A.15) becomes

$$Q_{ij} = \sum_{k,\ell=1}^N d_{ki}d_{\ell j}Q_{k\ell}(1-\mu)^2. \quad (\text{A.16})$$

694 Again, Q_{ij} corresponds to a probability of identity by descent: the individuals at
 695 sites i and j are identical by descent if their parents were and if neither mutated
 696 $((1-\mu)^2)$.

697 B In a subdivided population

698 B.1 β and γ

699 Now, we need to adapt the results presented in Appendix A to our structure of
 700 interest, a subdivided population, with dispersal and interaction probabilities
 701 given by eq. (2) and eq. (??). For the β and γ terms, we use a brute-force ap-
 702 proach, replacing d_{ij} and e_{ij} by their values in a subdivided population, and
 703 simplifying the equations (for instance, there are 60 different cases to consider
 704 for the four sums that appear in β_1^{DB} , shown in the table in section B.4 below).
 705 The calculations and subsequent simplifications are detailed in the supplemen-
 706 tary Mathematica file, and the results are presented in the main text (system (10),
 707 eq. (11a), and eq. (11b)).

708 B.2 Probabilities of identity by descent

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

715 We rewrite site labels ($1 \leq i \leq N$) as (ℓ_1, ℓ_2) , where ℓ_1 is the index of the deme
 716 ($1 \leq \ell_1 \leq N_D$) and ℓ_2 the position of the site within the deme ($1 \leq \ell_2 \leq n$). Then,
 717 we introduce notations $\tilde{d}_{i_1}^{i_2}$ and $\tilde{Q}_{i_1}^{i_2}$, that correspond to the dispersal probab-
 718 ity and probability of identity by descent to a site at distances i_1 and i_2 in the
 719 among-demes and within-deme dimensions (e.g., $\tilde{d}_{i_1}^{i_2} = d_{j_1 j_1 + i_1}^{j_2' j_2 + i_2}$).

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

722 B.2.1 Moran model

723 In Débarre (2017), it was shown that

$$\tilde{Q}_{r_1}^{r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_M}{1 - (1 - \mu) \tilde{D}_{q_1}^{q_2}} \exp\left(i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(i \frac{2\pi q_2 r_2}{N_2}\right) \quad (\text{B.17a})$$

724 with

$$\tilde{D}_{q_1}^{q_2} = \sum_{\ell_1=0}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} \tilde{d}_{\ell_1}^{\ell_2} \exp\left(-i \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 \ell_2}{N_2}\right), \quad (\text{B.17b})$$

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

$$\begin{aligned}\tilde{D}_{q_1} &= d_{\text{self}} + \sum_{\ell_2=1}^{N_2-1} d_{\text{in}} \exp\left(-\iota \frac{2\pi q_2 \ell_2}{N_2}\right) + \sum_{\ell_1=1}^{N_1-1} \sum_{\ell_2=0}^{N_2-1} d_{\text{out}} \exp\left(-\iota \frac{2\pi q_1 \ell_1}{N_1}\right) \exp\left(-\iota \frac{2\pi q_2 \ell_2}{N_2}\right) \\ &= d_{\text{self}} + (\delta_{q_2}(N_2 - 1) + (1 - \delta_{q_2})(-1)) d_{\text{in}} + (\delta_{q_1}(N_1 - 1) + (1 - \delta_{q_1})(-1)) (\delta_{q_2} N_2) d_{\text{out}} \\ &= d_{\text{self}} + (\delta_{q_2} N_2 - 1) d_{\text{in}} + (\delta_{q_1} N_1 - 1) \delta_{q_2} N_2 d_{\text{out}}.\end{aligned}\quad (\text{B.18a})$$

727 (δ_q is equal to 1 when q is equal to 0 modulo the relevant dimension, and to 0
728 otherwise). So for the three types of distances that we need to consider (distance
729 0, distance to another deme-mate, distance to individual in another deme), and
730 with $N_1 = N_D$ and $N_2 = n$, we obtain

$$\tilde{D}_0 = 1, \quad (\text{B.19a})$$

$$\tilde{D}_0^{q_1} = 1 - m - \frac{m}{d-1} \quad (q_1 \not\equiv 0 \pmod{N_1}), \quad (\text{B.19b})$$

$$\tilde{D}_{q_1}^{q_2} = d_{\text{self}} - d_{\text{in}} \quad (q_2 \not\equiv 0 \pmod{N_2}). \quad (\text{B.19c})$$

731 So for \tilde{Q} , using system (B.19) in eq. (B.17a),

$$\begin{aligned}\tilde{Q}_{r_1} &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu) \tilde{D}_0} + \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_0^{q_2}} \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right. \\ &\quad + \sum_{q_1=1}^{N_1-1} \frac{1}{1 - (1 - \mu) \tilde{D}_0^{q_1}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \\ &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{1}{1 - (1 - \mu) \tilde{D}_{q_1}^{q_2}} \exp\left(-\iota \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-\iota \frac{2\pi q_2 r_2}{N_2}\right) \right] \\ &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{1 - (1 - \mu)} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_2} N_2 - 1) \right. \\ &\quad + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})} (\delta_{r_1} N_1 - 1) \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (\delta_{r_1} N_1 - 1)(\delta_{r_2} N_2 - 1) \right].\end{aligned}\quad (\text{B.20})$$

732 In particular,

$$\begin{aligned}\tilde{Q}_0 &= \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (n - 1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})} (D - 1) \right. \\ &\quad \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} (D - 1)(n - 1) \right] \\ &= 1.\end{aligned}\quad (\text{B.21a})$$

733 We find λ'_M using the eq. (B.21a). Going back to eq. (B.20), when $r_1 = 0$, the two
 734 individuals are in the same deme. They are different when $r_2 \neq 0$, and so:

$$Q_{\text{in}} = \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})}(D-1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(D-1)(-1) \right]. \quad (\text{B.21b})$$

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

$$Q_{\text{out}} = \frac{\mu \lambda'_M}{N} \left[\frac{1}{\mu} + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})}(-1) + \frac{1}{1 - (1 - \mu)(1 - m - \frac{m}{d-1})}(-1) \right. \\ \left. + \frac{1}{1 - (1 - \mu)(d_{\text{self}} - d_{\text{in}})} \right]. \quad (\text{B.21c})$$

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

738 B.3 Wright-Fisher

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

$$\tilde{Q}_{r_1 r_2} = \frac{1}{N} \sum_{q_1=0}^{N_1-1} \sum_{q_2=0}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{D}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right), \quad (\text{B.22})$$

with $\tilde{\mathcal{D}}$ given in eq. (B.17b). In a subdivided population, with $N_1 = N_D$ and $N_2 = n$, this becomes

$$\begin{aligned}
 \tilde{Q}_{r_1} &= \frac{1}{N} \left[\frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_0)^2} + \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_2})^2} \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right. \\
 &\quad + \sum_{q_1=1}^{N_1-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \\
 &\quad \left. + \sum_{q_1=1}^{N_1-1} \sum_{q_2=1}^{N_2-1} \frac{\mu \lambda'_{WF}}{1 - (1 - \mu)^2 (\tilde{\mathcal{D}}_{q_1})^2} \exp\left(-i \frac{2\pi q_1 r_1}{N_1}\right) \exp\left(-i \frac{2\pi q_2 r_2}{N_2}\right) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \right. \\
 &\quad + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (\delta_{q_1} N_1 - 1) \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_1} N_1 - 1) (\delta_{q_2} N_2 - 1) \right] \\
 &= \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (\delta_{q_2} N_2 - 1) \delta_{q_1} N_1 \right. \\
 &\quad \left. + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (\delta_{q_1} N_1 - 1) \right]. \tag{B.23}
 \end{aligned}$$

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

$$1 = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} + \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} (N_2 - 1) N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (N_1 - 1) \right]. \tag{B.24a}$$

Then from eq. (B.23) we deduce

$$Q_{\text{in}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (d_{\text{self}} - d_{\text{in}})^2} N_1 + \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} (N_1 - 1) \right]. \tag{B.24b}$$

and

$$Q_{\text{out}} = \frac{\mu \lambda'_{WF}}{N} \left[\frac{1}{1 - (1 - \mu)^2} - \frac{1}{1 - (1 - \mu)^2 (1 - m - \frac{m}{d-1})^2} \right]. \tag{B.24c}$$

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

747 **B.4 Unpacking β_i^{DB}**

748 The table below contains all combinations for i, j, k, l involved in the four sums. (i, j) : means
 749 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_{self}	d_{self}	e_{self}	1
2	$j = i$	$k = i$	$l \neq i; l \in G_i$	$(i = j = k, l)$	$n - 1$	d_{self}	d_{in}	e_{in}	1
3	$j = i$	$k = i$	$l \notin G_i$	$(i = j = k), (l)$	$N - n$	d_{self}	d_{out}	e_{out}	1
4	$j = i$	$k \neq i; k \in G_i$	$l = i$	$(i = j = l, k)$	$n - 1$	d_{self}	d_{self}	e_{in}	Q_{in}
5	$j = i$	$k \neq i; k \in G_i$	$l = k$	$(i = j, k = l)$	$n - 1$	d_{self}	d_{in}	e_{self}	Q_{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_{self}	d_{in}	e_{in}	Q_{in}
7	$j = i$	$k \neq i; k \in G_i$	$l \notin G_i$	$(i = j, k), (l)$	$(n - 1)(N - n)$	d_{self}	d_{out}	e_{out}	Q_{in}
8	$j = i$	$k \notin G_i$	$l = i = j$	$(i = j = l), (k)$	$(N - n)$	d_{self}	d_{self}	e_{out}	Q_{out}
9	$j = i$	$k \notin G_i$	$l \neq i, l \in G_i$	$(i = j, l), (k)$	$(N - n)(n - 1)$	d_{self}	d_{in}	e_{out}	Q_{out}
10	$j = i$	$k \notin G_i$	$l = k$	$(i = j), (k = l)$	$(N - n)$	d_{self}	d_{out}	e_{self}	Q_{out}
11	$j = i$	$k \notin G_i$	$l \neq k; l \in G_k$	$(i = j), (k, l)$	$(N - n)(n - 1)$	d_{self}	d_{out}	e_{in}	Q_{out}
12	$j = i$	$k \notin G_i$	$l \notin G_i, G_k$	$(i = j), (k), (l)$	$(N - n)(N - 2n)$	d_{self}	d_{out}	e_{out}	Q_{out}
13	$j \neq i, j \in G_i$	$k = i$	$l = i$	$(i = k = l, j)$	$(n - 1)$	d_{in}	d_{self}	e_{self}	Q_{in}
14	$j \neq i, j \in G_i$	$k = i$	$l = j$	$(i = k, j = l)$	$(n - 1)$	d_{in}	d_{in}	e_{in}	Q_{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_{in}	d_{in}	e_{in}	Q_{in}
16	$j \neq i, j \in G_i$	$k = i$	$l \notin G_i$	$(i = k, j), (l)$	$(n - 1)(N - n)$	d_{in}	d_{out}	e_{out}	Q_{in}
17	$j \neq i, j \in G_i$	$k = j$	$l = i$	$(i = l, j = k)$	$(n - 1)$	d_{in}	d_{self}	e_{in}	1
18	$j \neq i, j \in G_i$	$k = j$	$l = j$	$(i, j = k = l)$	$(n - 1)$	d_{in}	d_{in}	e_{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_{in}	d_{in}	e_{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_{in}	d_{out}	e_{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_{in}	d_{self}	e_{in}	Q_{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_{in}	d_{in}	e_{in}	Q_{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_{in}	d_{in}	e_{self}	Q_{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_{in}	d_{in}	e_{in}	Q_{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_{in}	d_{out}	e_{out}	Q_{in}
26	$j \neq i, j \in G_i$	$k \notin G_i$	$l = i$	$(i = l, j), (k)$	$(n - 1)(N - n)$	d_{in}	d_{self}	e_{out}	Q_{out}
27	$j \neq i, j \in G_i$	$k \notin G_i$	$l = j$	$(i, j = l), (k)$	$(n - 1)(N - n)$	d_{in}	d_{in}	e_{out}	Q_{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_{in}	d_{in}	e_{out}	Q_{out}
29	$j \neq i, j \in G_i$	$k \notin G_i$	$l = k$	$(i, j), (k = l)$	$(n - 1)(N - n)$	d_{in}	d_{out}	e_{self}	Q_{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_{in}	d_{out}	e_{in}	Q_{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_{in}	d_{out}	e_{out}	Q_{out}
32	$j \notin G_i$	$k = i$	$l = i$	$(i = k = l), (j)$	$(N - n)$	d_{out}	d_{self}	e_{self}	Q_{out}
33	$j \notin G_i$	$k = i$	$l \neq i; l \in G_i$	$(i = k, l), (j)$	$(N - n)(n - 1)$	d_{out}	d_{in}	e_{in}	Q_{out}
34	$j \notin G_i$	$k = i$	$l = j$	$(i = k), (j = l)$	$(N - n)$	d_{out}	d_{out}	e_{out}	Q_{out}
35	$j \notin G_i$	$k = i$	$l \neq j; l \in G_j$	$(i = k), (j, l)$	$(N - n)(n - 1)$	d_{out}	d_{out}	e_{out}	Q_{out}
36	$j \notin G_i$	$k = i$	$l \notin G_i, G_j$	$(i = k), (j), (l)$	$(N - n)(N - 2n)$	d_{out}	d_{out}	e_{out}	Q_{out}
37	$j \notin G_i$	$k \neq i; k \in G_i$	$l = i$	$(i = l, k), (j)$	$(N - n)(n - 1)$	d_{out}	d_{self}	e_{in}	Q_{out}
38	$j \notin G_i$	$k \neq i; k \in G_i$	$l = k$	$(i, k = l), (j)$	$(N - n)(n - 1)$	d_{out}	d_{in}	e_{self}	Q_{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_{out}	d_{in}	e_{in}	Q_{out}
40	$j \notin G_i$	$k \neq i; k \in G_i$	$l = j$	$(i, k), (j = l)$	$(N - n)(n - 1)$	d_{out}	d_{out}	e_{out}	Q_{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_{out}	d_{out}	e_{out}	Q_{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_{out}	d_{out}	e_{out}	Q_{out}
43	$j \notin G_i$	$k = j$	$l = i$	$(i = l), (j = k)$	$(N - n)$	d_{out}	d_{self}	e_{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_{out}	d_{in}	e_{out}	1
45	$j \notin G_i$	$k = j$	$l = j$	$(i), (j = k = l)$	$(N - n)$	d_{out}	d_{out}	e_{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_{out}	d_{out}	e_{in}	1
47	$j \notin G_i$	$k = j$	$l \notin G_i, G_j$	$(i), (j = k), (l)$	$(N - n)(N - 2n)$	d_{out}	d_{out}	e_{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_{out}	d_{self}	e_{out}	Q_{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_{out}	d_{in}	e_{out}	Q_{in}
50	$j \notin G_i$	$k \neq j; k \in G_j$	$l = j$	$(i), (j = l, k)$	$(N - n)(n - 1)$	d_{out}	d_{out}	e_{in}	Q_{in}
51	$j \notin G_i$	$k \neq j; k \in G_j$	$l = k$	$(i), (j, k = l)$	$(N - n)(n - 1)$	d_{out}	d_{out}	e_{self}	Q_{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_{out}	d_{out}	e_{in}	Q_{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_{out}	d_{out}	e_{out}	Q_{in}
54	$j \notin G_i$	$k \notin G_i, G_j$	$l = i$	$(i = l), (j), (k)$	$(N - n)(N - 2n)$	d_{out}	d_{self}	e_{out}	Q_{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_{out}	d_{in}	e_{out}	Q_{out}
56	$j \notin G_i$	$k \notin G_i, G_j$	$l = j$	$(i), (j = l), (k)$	$(N - n)(N - 2n)$	d_{out}	d_{out}	e_{out}	Q_{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_{out}	d_{out}	e_{out}	Q_{out}
58	$j \notin G_i$	$k \notin G_i, G_j$	$l = k$	$(i), (j), (k = l)$	$(N - n)(N - 2n)$	d_{out}	d_{out}	e_{self}	Q_{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_{out}	d_{out}	e_{in}	Q_{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_{out}	d_{out}	e_{out}	Q_{out}